4
2Ej

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

"ALGUNOS ASPECTOS SOBRE LA TEORIA DE APROXIMACION"

T E S I S

QUE PARA OBTENER EL TITULO DE:

M A T E M A T I C O

P R E S E N T A :

JULIO CESAR CEDILLO SANCHEZ

1996

TESIS CON
FALLA DE ORIGEN

VNIVERSIDAD NACIONAL
AVENMA DE
MEXICO

M. en C. Virginia Abrín Batule
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
P r e s e n t e

Comunicamos a usted que hemos revisado el trabajo de Tesis:

" Algunos aspectos sobre la Teoría de Aproximación ".

realizado por Cedillo Sánchez Julio César.

con número de cuenta 8838296-8 , pasante de la carrera de Matemáticas

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

México, D.F. a 30 de abril de 1996.

Director de Tesis Propietario	Dr. Guillermo Grabinsky Steider.
Propietario	Dra. María Emilia Caballero Acosta.
Propietario	Dr. Javier Paez Cárdenas.
Suplente	Act. Javier Fernández García.
Suplente	Dr. Jesús López Estrada

Consejo Departamental de Matemáticas

M. en C. Alejandro Bravo Mojica

INTRODUCCION

El presente trabajo está dedicado al estudio de la Teoría de Aproximación a través del cuál se pretende que el lector adquiera las nociones básicas para introducirse en esta rama de las matemáticas, para ello sólo se necesita que este familiarizado con el material de un primer curso de análisis real.

El trabajo se inicia con el estudio del Teorema de Aproximación de Weierstrass (TAW) (1885) del que se proporcionan varias demostraciones y se discuten algunas preguntas relacionadas con su contenido, como la existencia y caracterización de los aproximadores polinomiales uniformes (norma uniforme) para una función continua dada.

De lo anterior se van desarrollando algunos de los temas centrales en la Teoría de Aproximación.

El material se encuentra distribuido en 9 capítulos. En el primer capítulo se presenta el TAW, se hacen ver algunas diferencias entre dicho resultado y el Teorema de Taylor para funciones analíticas reales. Se proporcionan varias demostraciones de TAW entre otras la prueba original de Weierstrass (1885), y aquella debida a Bernstein (1912) y la de Lebesgue (1908). La demostración de Bernstein tiene gran interés teórico ya que introduce una clase de polinomios (polinomios de Bernstein) particularmente importante por el hecho que se conocen explícitamente a los aproximadores polinomiales uniformes de una función continua f, además de tener la virtud de que si f es de clase C^p los polinomios de Bernstein no sólo aproximan a f sino que también sus derivadas hacen lo propio con cada una de las p derivadas de f (ver (I.20)).

A pesar de esas cualidades, esta clase de polinomios tiene la desventaja de que el error uniforme de la aproximación a la función f converge muy lentamente a cero (Teorema de Voronosky (1932) (I.34)).

La demostración del Teorema de Weierstrass por Lebesgue se introduce para probar la equivalencia entre dicho resultado y una generalización del TAW, a saber, el Teorema de Stone-Weierstrass (1948) (II.11) al cuál se le dedica unas páginas y del que obtenemos algunos corolarios de interés a lo largo del trabajo (V.gr. II.15), así mismo el TAW nos permite obtener los Teoremas de extensión de Tietze (II.20) y un Teorema de interpolación de Walsh (II.22).

En la tercera parte del trabajo nos enfocamos en resolver preguntas sobre existencia de mejores aproximadores en general para un elemento de un espacio vectorial normado tomado de un subespacio vectorial de dimensión finita. La primera pregunta que se responde garantiza la existencia (III.4), a continuación examinamos su unicidad. Para tenerla se introduce una condición suficiente sobre la norma (norma estrictamente convexa, III.23), y que desafortunadamente no cumple la norma uniforme por lo que será necesario buscar una respuesta a través de la caracterización de los mejores aproximadores polinomiales (Teorema de Equioscilación de Chebyshev, 1859), con lo que se da una prueba de la existencia y unicidad de dichos aproximadores para una función continua f (III.34).

Una vez resueltas estas cuestiones debe observarse que en la práctica el cálculo del mejor aproximador suele ser muy complicado y sólo en algunos casos dicho cálculo puede darse explícitamente como lo es para la función $f(x) = x^{n+1}$ aproximandola con polinomios de grado $\leq n$ con lo que se introducen los polinomios de Chebyshev de primera especie, los que son pieza fundamental en la Teoría de aproximación e Interpolación por lo que se les dedica un capítulo completo en el presente trabajo (capítulo 4).

Se desarrollan los resultados obtenidos hasta el momento ahora para polinomios trigonométricos, los cuáles son introducidos en (IV.10) a partir de los cuáles se obtiene una demostración más del TAW (IV.13), se enuncia y demuestra el Teorema Trigonométrico de Weierstrass (V.3), así cómo en el capítulo V, se desarrolla toda la discusión encaminada a la formulación y demostración de los resultados trigonométricos correspondientes de Walsh (V.8) y de Equioscilación de Chebyshev (V.12).

En la última parte del trabajo nos dedicamos a estudiar la rapidez con la cuál los aproximadores considerados en capítulos anteriores tienden a las funciones que aproximan y esto se hace en términos del módulo de continuidad, lo que nos permitirá medir de alguna manera el orden correcto de la aproximación, se muestra un resultado relacionado con los polinomios de Bernstein (VI.5) con lo que se demuestra la lentitud de la aproximación. Con lo anterior nos dedicamos al estudio de los Teoremas polinomiales de Jackson los cuáles dan una respuesta sobre la conducta que tiene el error mínimo $E_n(f)$ cuando $n \to \infty$, para ello demostramos el Teorema de Korovkin (VI.13) a partir del cuál se desprende el Teorema polinomial de Jackson (VI.16) y un grupo de corolarios que nos dan una respuesta sobre $E_n(f)$ cuando la función pertenece a una clase de funciones.

Posteriormente se desarrolla la teoría análoga para polinomios trigonométricos haciendo uso de diversos núcleos de sumabilidad (Núcleo de Dirichlet (VII.4), Núcleo de Fejer (VII.5)), los que dan pie a la introducción del núcleo de Jackson (VII.9) parte fundamental en la demostración del Teorema Trigonométrico de Jackson (VII.11), del que se desprenden también un grupo de corolarios que nos dan una respuesta sobre el error mínimo $\widehat{E}_n(f)$ con polinomios trigonométricos.

El capítulo VIII estudia los resultados inversos a algunos corolarios del Teorema Trigonométrico de Jackson ((VIII.6) y (VIII.7)) para lo cuál se introducen dos desigualdades de interés en la Teoría de Aproximación , la desigualdad de Bernstein y Markoff ((VIII.3) y (VIII.4)).

En la sección 2 del capítulo VIII se introduce la clase de Zygmund que resuelve el problema de dar condiciones necesarias y suficientes para que $E_n(f) \leq \frac{c}{n}$ (VIII.11) y en la última parte del capítulo se prueba que $\frac{1}{n^\alpha}$ es el mejor orden de aproximación para $f \in Lip(\alpha)$ (VIII.2).

Para concluir con la exposición del material en el último capítulo se hace un breve estudio de la Teoría de Aproximación con otras familias , entre los temas que destacan estan los Teoremas de Müntz, que dan condiciones necesarias y suficientes para que una función $f \in \mathcal{L}^2$ se pueda aproximar en media cuadrática por combinaciones lineales finitas de funciones $\{x^{p_1}, x^{p_2}, \ldots\}$ con $p_n \in \mathbb{R}$ y $p_m \neq p_n$ si $m \neq n$ (IX.16), para ello se presentan una serie de interesantes resultados previos que son necesarios para la demostración del Teorema de Müntz (V.gr (IX.7), (IX.8)), también se expone una versión del Teorema en norma uniforme donde se dan sólo condiciones suficientes y se concluye el capítulo con una breve introducción a la aproximación uniforme de una función continua pero ahora con funciones racionales.

AGRADECIMIENTOS

Este pequeño logro esta dedicado a todas las personas que a lo largo de mi vida ayudaron a desarrollar mis estudios y mi formación como ser humano. Principalmente a mi madre, la cuál es todo para mi ya que con un sin número de acciones logró lo que ahora soy. ¡Mil gracias mamá!. Te amo, no puedo agradecerte todo, pero todo, sólo te digo que este triunfo es totalmente tuyo.

A mi padre, el cual siempre ha sido, la disciplina en mi vida, gracias por marcar el camino ya que sin tu energía no sería posible este momento. ¡Eres único!.

Por supuesto no pueden faltar de mencionar mis hermanas, Lizsully y Ana Karen, las cuales fueron, son y serán un motivo para superarme cada vez más para ser un buen hermano y amigo siempre.

También debo agardecer a los profesores que tuve el gusto de conocer, a todos les digo que pueden estar seguros de que aprovecharé todos sus conocimientos adquiridos para mi formación como profesionista hoy y siempre, muy especialmente dedico unas cuantas líneas a dos grandes profesores, pero sobre todo personas, que han dejado en mi una huella inborrable debido a sus extraordinarias cualidades, en primer término al Dr. Guillermo Grabinsky a quien le debo entre muchas cosas, la elaboración del trabajo, su paciencia así como todos aquellos aciertos que el lector puede encontrar en el mismo, gracias doctor, dentro de lo que puedo ofrecerle con sinceridad es una amistad por siempre. ¡Es super, doctor!.

A Javier Fernández, el profesor, es muy díficil expresar lo que siento por ti , pero sé que lo sabes, piensa que siempre te admiraré y apoyaré incondicionalmente. ¡Eres fantástico Javier!.

A Guillermo y Javier les deseo lo mejor. Asimismo debo agradecer los comentarios y sugerencias de: Dr. Javier Páez, Dra. Ma. Emilia Caballero, Dr. Jesús López y Dr. Sergio Macías que hicieron del presente escrito una realidad.

No puedo dejar de mencionar a los amigos con los que sufrí y gocé esta aventura llamada Matemáticas, los mejores deseos para Berta, Gerardo " el doctor", Ricardo, Lulú, Vera, Susana, así como a Pablo y Alvaro, por su alegría e interés en la última etapa del trabajo.

A todos ustedes, no encuentro palabras para decirles lo que al momento de escribir estas líneas siento al ver sus nombres reunidos en estas páginas. Pero de algo sí pueden estar seguros que cada uno forma parte de mí.

¡Mil gracias!
Julio César Cedillo Sánchez
Mayo, 1996.

INDICE

CAPITULO I. Teorema de aproximación de Weierstrass

CAPITULO II. Aplicaciones del Teorema de Weierstrass

CAPITULO III. Mejor aproximador polinomial

CAPITULO IV. Polinomios de Chebyshev

CAPITULO V. Polinomios Trigonométricos

CAPITULO VI. Teoremas de Jackson

CAPITULO VII. Teoremas trigonométricos de Jackson

CAPITULO VIII. Inversos a los Teoremas de Jackson

CAPITULO IX. Mejor aproximador con otras familias

CAPITULO I

TEOREMA DE APROXIMACION DE WEIERSTRASS

En este capítulo, nos vamos a enfocar al estudio de un teorema fundamental en la Teoría de Aproximación, el Teorema de Aproximación de Weierstrass (1885), daremos varias demostraciones de este teorema, una de las cuáles, utiliza una clase de polinomios debida a Bernstein y a la que dedicaremos un examen cuidadoso, así como otra demostración del Teorema de Weierstrass debida a Lebesgue y la prueba original de Weierstrass.

§1. El Teorema de Aproximación de Weierstrass

El Teorema dice lo siguiente:

(I.1) Teorema de Weierstrass (1885). *Sea $f \in C([a,b])$. Dada $\varepsilon > 0$ existe un polinomio $p = p(\varepsilon)$ (de grado suficientemente grande) para el cuál $\|f - p\|_{[a,b]} =$ $= \max\{|f(x) - p(x)| : \ x \in [a,b]\} < \varepsilon$.*

En otras palabras, el Teorema de Weierstrass asegura la posibilidad de aproximar *uniformemente* funciones continuas mediante polinomios sobre intervalos cerrados y acotados.

Cabe hacer la siguiente aclaración:

El Teorema de Taylor para funciones analíticas reales y el Teorema de Weierstrass no deben ser confundidos. El Teorema de Taylor garantiza que para una función analítica real en el intervalo $(x_0 - r, x_0 + r)$ tenemos que $f(x) = \sum_{k=1}^{\infty} a_k(x - x_0)^k$ y converge uniformemente en $|x - x_0| \leq r'$ con $r' \in (0, r)$, de donde dada $\varepsilon > 0$ existe $N = N(\varepsilon) \in \mathbb{N}$ tal que para $p_N(x) = \sum_{k=1}^{N} a_k(x - x_0)^k$ tenemos que $|f(x) - p_N(x)| < \varepsilon$ para $|x - x_0| < r'$, por lo anterior tenemos convergencia uniforme en $x_0 - r' < x <$ $< x_0 + r'$. Sin embargo existen funciones infinitamente diferenciables para las cuáles no hay expansión en series de potencias, pero esto no afecta al Teorema de Weierstrass ya que nos garantiza que podemos aproximarnos uniformemente a funciones que sean sólo continuas.

Así, dada una sucesión de números positivos $\{\varepsilon_n\} \to 0$, existen polinomios $p_{m_n}, \forall n \in \mathbb{N}$ tales que $\|f - p_{m_n}\|_{[a,b]} < \varepsilon_n$. Consecuentemente $f = \lim_{n\to\infty} p_{m_n}$ uniforme en $[a,b]$.

De este modo tenemos que:

$$S_k(x) = p_{n_1}(x) + (p_{n_2}(x) - p_{n_1}(x)) + \ldots + (p_{n_k}(x) - p_{n_{k-1}}(x))\ \forall k \in \mathbb{N}.$$

converge uniformemente a $f(x)\ \forall x \in [a,b]$, es decir,

$$f(x) = \sum_{k=1}^{\infty} [p_{n_{k+1}}(x) - p_{n_k}(x)] + p_{n_1}(x)$$

donde la serie converge uniformemente a $f(x)$.

Para concluir, podemos decir que una función analítica real puede expanderse en una ***serie de potencias*** uniformemente convergente y una función continua, no analítica, puede expanderse en una "***serie de polinomios***" uniformemente convergente.

Reproduciremos la demostración original debida a Weierstrass. Por el momento supondremos que:

(i) $[a,b] = [0,1]$

(ii) $f(0) = f(1) = 0$,

Si definimos $f(x) = 0\ \forall x \notin [0,1]$, entonces f es uniformemente continua en todo $\mathbb{R}$.

A continuación daremos algunas definiciones, así como resultados previos a la demostración de (**I.1**).

(I.2) Definición. Sean Q_n los polinomios de grado $\leq 2n$ en $[-1,1]$ definidos por $Q_n(x) = c_n(1-x^2)^n\ \forall n \in \mathbb{N}, x \in [-1,1]$ donde c_n es elegida de tal manera que $\int_{-1}^{1} Q_n(x)dx = 1\ \forall n \in \mathbb{N}$ y $Q_n(x) \geq 0$.

Ahora obtendremos información sobre el orden de magnitud de las constantes c_n de la definición anterior.

(I.3) Lema. $\int_{-1}^{1}(1-x^2)^n dx > \frac{1}{\sqrt{n}}\ \forall n \in \mathbb{N}$ de donde $c_n < \sqrt{n}\ \forall n \in \mathbb{N}$.

Demostración. Como $(1-x^2)^n$ es una función par y positiva en $[0,1]$ entonces:

$$\int_{-1}^{1}(1-x^2)^n dx = 2\int_0^1 (1-x^2)^n dx \geq 2\int_0^{\frac{1}{\sqrt{n}}} (1-x^2)^n dx.$$

Por otro lado es inmediato de la desigualdad de Bernoulli ver que $(1-x^2)^n \geq 1-nx^2$ $\forall n \in \mathbb{N}$, $x \in [0,1]$. Por lo anterior se sigue:

$$\int_{-1}^{1}(1-x^2)^n dx \geq 2\int_0^{\frac{1}{\sqrt{n}}}(1-nx^2)dx \ \forall n \in \mathbb{N},\ x \in [0,1]$$

pero $2\int_0^{\frac{1}{\sqrt{n}}}(1-nx^2)dx = \frac{4}{3\sqrt{n}} > \frac{1}{\sqrt{n}}$ $\forall n \in \mathbb{N}$, por lo que $\displaystyle\int_{-1}^{1}(1-x^2)^n dx > \frac{1}{\sqrt{n}}$. Por la definición (I.2) $c_n \neq 0$ $\forall n \in \mathbb{N}$ entonces:

$$\frac{1}{c_n} = \frac{1}{c_n}\int_{-1}^{1} Q_n(x)dx = \int_{-1}^{1}(1-x^2)^n dx > \frac{1}{\sqrt{n}}$$

y así $c_n < \sqrt{n}$ $\forall n$.

Q.E.D.

Probaremos el siguiente:

(I.4) Lema. $Q_n(x) \longrightarrow 0$ *cuando* $n \to \infty$ *uniformemente en* $\delta \leq |x| \leq 1$.

Demostración. Dada $\delta > 0$ por (I.3) tenemos que: $Q_n(x) = c_n(1-x^2)^n$ entonces $Q_n'(x) = -2c_n nx(1-x^2)^{n-1} \leq -2n\sqrt{n}x(1-x^2)^{n-1}$ y observemos que $Q_n'(x) \leq 0$ si $\delta \leq x \leq 1$, mientras que $Q_n'(x) \geq 0$ si $-1 \leq x \leq -\delta$ pero en ambos casos

$$Q_n(x) \leq \sqrt{n}(1-\delta^2)^n \quad (1).$$

si $\delta \leq |x| \leq 1$ y como $\sqrt{n}c^n \longrightarrow 0$ con $n \to \infty$ si $|c| < 1$ tenemos que $Q_n(x) \longrightarrow 0$ con $n \to \infty$ uniformemente en $\delta \leq |x| \leq 1$.

Q.E.D.

(I.5) Definición. Definimos:

$$p_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)dt,\ x \in [0,1].$$

De la definición anterior y dado que $f(x) = 0$ si $x \notin [0,1]$, tenemos que $p_n(x) = $ $\displaystyle = \int_{-x}^{1-x} f(x+t)Q_n(t)dt$, si hacemos el cambio de variable $\varphi(t) = t + x$ obtenemos:

$$p_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)dt = \int_0^1 f(t)Q_n(t-x)dt \ \forall n \in \mathbb{N}.$$

Claramente la última integral es un polinomio en la indeterminada x. Por lo tanto $\{p_n\}$ es una sucesión de polinomios con coeficientes reales.

A continuación daremos la demostración al Teorema de Weierstrass.

(I.6) Teorema. *Los polinomios p_n definidos en (I.5) convergen a f uniformemente en $x \in [0,1]$.*

Demostración. Dada $\varepsilon > 0$, existe $\delta(\varepsilon) > 0$ con $\delta < 1$ tal que $|x-y| < \delta$ entonces $|f(y) - f(x)| < \frac{\varepsilon}{2}$ ya que f es uniformemente continua en $\mathbb{R}$. Por (I.2) y **(1)** de (I.4) tenemos: $|p_n(x) - f(x)| = \left|\int_{-1}^{1} \{f(x+t)Q_n(t) - f(x)Q_n(t)\}dt\right| \leq$

$$\leq \int_{-1}^{1} |f(x+t)Q_n(t) - f(x)Q_n(t)|dt = \int_{-1}^{-\delta} |f(x+t)Q_n(t) - f(x)Q_n(t)|dt +$$

$$\int_{-\delta}^{\delta} |f(x+t)Q_n(t) - f(x)Q_n(t)|dt + \int_{\delta}^{1} |f(x+t)Q_n(t) - f(x)Q_n(t)|dt \leq$$

$$\leq 2||f||_{[0,1]} \int_{-1}^{-\delta} Q_n(t)dt + \frac{\varepsilon}{2}\int_{-\delta}^{\delta} Q_n(t)dt + 2||f||_{[0,1]}\int_{\delta}^{1} Q_n(t)dt \leq$$

$$\leq 4||f||_{[0,1]}\sqrt{n}(1-\delta^2)^n(1-\delta) + \tfrac{\varepsilon}{2}\int_{-1}^{1} Q_n(t)dt < 4||f||_{[0,1]}\sqrt{n}(1-\delta^2)^n + \frac{\varepsilon}{2} < \varepsilon$$

si n es suficientemente grande.

Q.E.D.

(I.7) Observaciones.

(i) Recordemos que al principio de la sección se supuso que $f(0) = f(1) = 0$. Dada cualquier función $f: [0,1] \to \mathbb{R}$ continua podemos definir $g: [0,1] \to \mathbb{R}$ como $g(x) = f(x) - f(0) - x(f(1) - f(0))$, de donde es claro que $g \in C([0,1])$ y $g(0) = g(1) = 0$. La interpretación geométrica de lo anterior es simplemente "restarle" a la función f la recta ℓ que pasa por los puntos $(0, f(0))$ y $(1, f(1))$, entonces por la prueba anterior existe una sucesión de polinomios (p_n) tal que $||g - p_n||_{[0,1]} \to 0$ pero es claro que $(p_n + \ell)$ es también una sucesión de polinomios que satisface: $||f - p_n - \ell||_{[0,1]} = ||g - p_n||_{[0,1]} \to 0$. Por lo tanto toda $f \in C([0,1])$ puede aproximarse uniformemente por polinomios.

(ii) Si $f \in C([a,b])$, consideramos el cambio lineal de variable $\varphi: [0,1] \to [a,b]$ dado por $\varphi(t) = (1-t)a + tb$ entonces $\varphi^{-1}(t) = \frac{t-a}{b-a}$. Consideramos a la función $g = f \circ \varphi \in C([0,1])$ entonces por el Teorema de Weierstrass existe una sucesión de polinomios (p_n) tal que $||g - p_n||_{[0,1]} \to 0$, es decir, $||f \circ \varphi - p_n||_{[0,1]} \to 0$. Dado que φ es lineal es fácil ver que $p_n \circ \varphi^{-1}$ es también un polinomio (del mismo grado que p_n), así tenemos: $||f \circ \varphi \circ \varphi^{-1} - p_n \circ \varphi^{-1}||_{[a,b]} \to 0$, por lo tanto $||f - p_n \circ \varphi^{-1}||_{[a,b]} \to 0$. Por lo que podemos concluir que toda función continua f en $[a,b]$ se puede aproximar uniformemente por polinomios.

(iii) Si f esta definida en un intervalo que no sea acotado ó cerrado, el Teorema de Weierstrass puede fallar, veamos los siguientes ejemplos:

(a) Sea $f(x) = e^x$, definida en $[0,\infty)$, si f se aproxima uniformemente con polinomios en dicho intervalo, entonces dada $\varepsilon = 1$ existe $p(x) = a_0 + a_1x + \ldots + a_nx^n$ $(a_n \neq 0)$ tal que $|e^x - p(x)| < 1$ $\forall x \in [0,\infty)$, si dividimos entre x^n, $x > 0$ entonces $|\frac{e^x}{x^n} - \frac{a_0}{x^n} - \ldots - a_n| < \frac{1}{x^n}$, lo cuál no es posible ya que el lado derecho de la desigualdad anterior tiende a cero conforme x tiende a ∞ mientras que por otro lado $\frac{e^x}{x^n}$ tiende a $+\infty$ conforme x lo hace también.

(b) Sea $f(x) = \ln(x)$, definida en $(0,1)$, si f se aproxima con polinomios uniformemente en dicho intervalo, entonces dada $\varepsilon = 1$ existe $p(x) = a_0 + a_1x + \ldots + a_nx^n$ $(a_n \neq 0)$ tal que $|\ln(x) - p(x)| < 1$ $\forall x \in (0,1)$ haciendo $x \to 0^+$ la desigualdad no se cumple.

(iv) Por último no podemos omitir la condición de que f sea continua en $[a,b]$ debidó al hecho de que f es el límite uniforme de una sucesión de funciones continuas y al teorema siguiente:

(I.8) Teorema. *Sea (f_n) una sucesión de funciones de un dominio $D \subset \mathbb{R}^n$ en $\mathbb{R}^m$ tal que $f_n \to f$ uniformemente en D. Si f_n es continua en $x_0 \in D$ para toda $n \in \mathbb{N}$ entonces f es continua en x_0.*

Demostración. Sea $\varepsilon > 0$ por hipótesis existe $N = N(\varepsilon) \in \mathbb{N}$ tal que $||f_n - f||_D < \frac{\varepsilon}{3}$ $\forall n \geq N$ y como f_N es continua en x_0 existe $\delta_1(\varepsilon, x_0) > 0$ tal que si $|x - x_0| < \delta_1$ entonces $|f_N(x) - f_N(x_0)| \leq \frac{\varepsilon}{3}$. De esta manera si $|x - x_0| < \delta_1$, $x \in D$ entonces $|f(x) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \leq$ $\leq 2||f_n - f||_D + |f_N(x) - f_N(x_0)| < \varepsilon$. Por lo que f es continua en x_0.

Q.E.D.

§2 Polinomios de Bernstein

Como se mencionó al principio del capítulo hay varias demostraciones del Teorema de Weierstrass, la que veremos a continuación es una prueba debida a S. N. Bernstein (1912). Para comenzar daremos la siguiente:

(I.9) Definición. Sea $f\colon [0,1] \to \mathbb{R}$ acotada. Entonces el n-ésimo *polinomio de Bernstein para f* se define por:

$$B_n(f)(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}.$$

Es claro que $B_n \in \mathcal{P}_n$. A partir de la definición se pueden deducir algunas propiedades básicas de esta clase de polinomios, como:

(i) B_n es un operador lineal y acotado de $A([0,1])$ en $\mathcal{P}_n$.

(ii) $\|B_n(f)\|_{[0,1]} \leq \|f\|_{[0,1]}$. Si $f(x) = x$ entonces $B_n(f)(x) = x$ por lo que $\|B_n(f)\| = 1$.

(iii) $B_n(f)(0) = f(0)$; $B_n(f)(1) = f(1)$ (observación: Adoptamos la convención que $0^0 = 1$).

(iv) Si $m \leq f(x) \leq M$ entonces $m \leq B_n(f) \leq M$.

(v) Si $f(x) \leq g(x)$ $\forall x \in [0,1]$ entonces $B_n(f) \leq B_n(g)$

(vi) Si $f\colon [a,b] \to \mathbb{R}$ es acotada entonces:

$$B_n(f)(x) = \frac{1}{(b-a)^n} \sum_{k=0}^{n} f\left(a + \frac{k(b-a)}{n}\right)\binom{n}{k}(x-a)^k(b-x)^{n-k}.$$

$\forall x \in [a,b]$, usando el cambio de variable $\varphi\colon [0,1] \to [a,b]$ dado por
$\varphi(x) = (1-x)a + xb$.

De las propiedades $(iii), (iv), (v)$ y (vi) puede concluirse que los polinomios de Bernstein se tratan de "parecer" a la función f, lo cuál se reafirmará más adelante al obtener algunos resultados teóricos de interés. Antes de probar el Teorema de Bernstein el cuál nos dará otra demostración del Teorema de Weierstrass, un lema previo nos premitirá calcular explícitamente algunos polinomios de Bernstein para algunas funciones f, así cómo para simplificar algunas expresiones, para lo cuál introducimos la siguiente:

(I.10) Definición. Sea $(y_n)_{n=0}^{\infty}$ una sucesión de números reales. Las *diferencias sucesivas de valores adyacentes*, Δy_k, se definen por: $\Delta y_k = y_{k+1} - y_k$ para toda $k \in \mathbb{N} \cup \{0\}$. Las *diferencias superiores* son definidas de manera similar, es decir, $\Delta^2 y_k = \Delta(\Delta y_k) = \Delta(y_{k+1} - y_k) = (y_{k+2} - y_{k+1}) - (y_{k+1} - y_k)$. En general: $\Delta^{n+1}(y_k) = \Delta(\Delta^n y_k) = \Delta^n y_{k+1} - \Delta^n y_k$ y definimos $\Delta^0 y_k = y_k$.

Claramente: $\Delta(cy_k) = c(y_{k+1} - y_k) = c\Delta y_k$ y $\Delta(x_k + y_k) = \Delta x_k + \Delta y_k$ por lo que Δ es "lineal".

(I.11) Lema. *Tenemos que* $\Delta^n y_k = \sum_{r=0}^{n} (-1)^{n-r} \binom{n}{r} y_{k+r}$ $(*)$

Demostración. El resultado es claro para $n = 0, 1$, supongamos que $(*)$ se cumple para n, entonces:

$$\Delta^{n+1}(y_k) = \Delta(\Delta^n y_k) = \Delta\Big(\sum_{r=0}^{n}(-1)^{n-r}\binom{n}{r} y_{k+r}\Big) =$$

$$= \sum_{r=0}^{n}(-1)^{n-r}\binom{n}{r}\Delta y_{k+r} = \sum_{r=0}^{n}(-1)^{n-r}\binom{n}{r}(y_{k+r+1} - y_{k+r}) =$$

$$= \sum_{r=0}^{n}(-1)^{n-r}\binom{n}{r} y_{k+r+1} - \sum_{r=0}^{n}(-1)^{n-r}\binom{n}{r} y_{k+r}$$

haciendo $j = r + 1$ en la primera suma tenemos:

$$\sum_{j=1}^{n+1}(-1)^{n-j+1}\binom{n}{j-1} y_{k+j} - \sum_{r=0}^{n}(-1)^{n-r}\binom{n}{r} y_{k+r} =$$

$$= \sum_{j=1}^{n}\left[(-1)^{n+1-j}\left[\binom{n}{j-1} + \binom{n}{j}\right] y_{k+j}\right] + \binom{n}{n} y_{k+n+1} - (-1)^n \binom{n}{0} y_k$$

pero $\binom{n}{j-1} + \binom{n}{j} = \binom{n+1}{j}$ por lo que:

$$= \sum_{j=1}^{n}(-1)^{n+1-j}\binom{n+1}{j} y_{k+j} + \binom{n+1}{n+1} y_{k+n+1} + (-1)^{n+1}\binom{n+1}{0} y_k =$$

$$= \sum_{j=0}^{n+1}(-1)^{n+1-j}\binom{n+1}{j} y_{k+j}.$$

Q.E.D.

(I.12) Corolario. *Si $k = 0$ tenemos* $\Delta^n y_0 = \sum_{r=0}^{n}(-1)^{n-r}\binom{n}{r} y_r$.

Lo anterior nos permite obtener a $B_n(f)$ usando diferencias sucesivas.

(I.13) Lema. *Sea $f:[0,1]\to \mathbb{R}$ acotada, entonces:*

$$B_n(f)(x)=\sum_{t=0}^{n}\Delta^t f(0)\binom{n}{t}x^t$$

donde las diferencias sucesivas se calculan en $y_r=f(\frac{r}{n})$ con $r\in \mathbb{N}\cup\{0\}$.

Demostración.

$$B_n(f)(x)=\sum_{k=0}^{n}f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k}=$$

$$=\sum_{k=0}^{n}f\left(\frac{k}{n}\right)\binom{n}{k}x^k\sum_{j=0}^{n-k}\binom{n-k}{j}(-1)^{n-k-j}x^{n-k-j}=$$

$$=\sum_{k=0}^{n}\sum_{j=0}^{n-k}f\left(\frac{k}{n}\right)\binom{n}{k}\binom{n-k}{j}x^{n-j}(-1)^{n-k-j}=$$

haciendo $t=n-j$ entonces $j=n-t$ tenemos:

$$\sum_{k=0}^{n}\sum_{t=k}^{n}f\left(\frac{k}{n}\right)\binom{n}{k}\binom{n-k}{n-t}x^t(-1)^{t-k}\ \textbf{(1)}$$

observemos que en la expresión anterior el término general depende de los índices k y t, llamemos a cada término por a_{kt} y analicemos como se suman. Para visualizar con facilidad la suma coloquemos los términos como en una matriz de la siguiente manera:

$$\begin{pmatrix} a_{00} & a_{01} & a_{02} & \cdots & a_{0n} \\ & a_{11} & a_{12} & \cdots & a_{1n} \\ & & a_{22} & \cdots & a_{2n} \\ & & & & \cdot \\ & & & & \cdot \\ & & & & \cdot \\ & & & & a_{nn} \end{pmatrix}$$

sólo escribimos los términos que se encuentran arriba de la diagonal y en está. Observemos que **(1)** lo que nos indica es que los términos a_{kt} los sumamos renglón por renglón de izquierda a derecha, pues bien vamos a reordenar está suma de la siguiente manera: sumemos columna por columna pero de arriba para abajo, por lo que **(1)** coincide con:

$$\sum_{t=0}^{n}\sum_{k=0}^{t}f\left(\frac{k}{n}\right)\binom{n}{k}\binom{n-k}{n-t}x^t(-1)^{t-k}$$

pero $\binom{n}{k}\binom{n-k}{n-t} = \binom{n}{t}\binom{t}{k}$ entonces por (I.12) tenemos:

$$\sum_{t=0}^{n} \binom{n}{t} x^t \sum_{k=0}^{t} f\left(\frac{k}{n}\right) \binom{t}{k} (-1)^{t-k} = \sum_{t=0}^{n} \Delta^t f(0) \binom{n}{t} x^t$$

Q.E.D.

(I.14) Ejemplos.

(i) Si $f \in \mathcal{P}_m$ entonces $\Delta^t f(0) = 0$ para toda $t > m$ aplicando el TVM-D.

(ii) Si $f(x) = 1$, tenemos que: $\Delta^0 f(0) = 1$ y $\Delta^t f(0) = \sum_{k=0}^{t} \binom{t}{k} (-1)^{t-k} =$

$= (1 + (-1))^t = 0 \; \forall t > 0$. De esté modo $B_n(f)(x) = 1$.

(iii) Si $f(x) = x$, tenemos que: $\Delta^0 f(0) = f(0) = 0$ y

$$\Delta^1 f(0) = \sum_{k=0}^{1} f\left(\frac{k}{n}\right) \binom{1}{k} (-1)^{1-k} = \frac{1}{n}.$$

De (i) $\Delta^n f(0) = 0 \; \forall \, n \geq 2$. Entonces $B_n(f)(x) = \frac{1}{n} \binom{n}{1} x = x$.

(iv) Si $f(x) = x^2$, tenemos que: $\Delta^0 f(0) = f(0) = 0$, $\Delta^1 f(0) = \frac{1}{n^2}$ y $\Delta^2 f(0) = \frac{2}{n^2}$ entonces por (i) $\Delta^n f(0) = 0 \; \forall \, n \geq 3$. Por lo que $B_n(f)(x) = \frac{1}{n} x + \left(1 - \frac{1}{n}\right) x^2$.

(v) Si $f(x) = e^{\alpha x}$, tenemos que $B_n(f)(x) = \sum_{k=0}^{n} \binom{n}{k} e^{\frac{\alpha k}{n}} x^k (1-x)^{n-k} =$

$$= \sum_{k=0}^{n} \binom{n}{k} \left(e^{\frac{\alpha}{n}} x\right)^k (1-x)^{n-k} = \left(x e^{\frac{\alpha}{n}} + 1 - x\right)^n$$

Calcular explícitamente el polinomio de Bernstein para una función f puede resultar una tarea complicada aunque para lo que necesitamos basta con los polinomios de Bernstein para las funciones $1, x, x^2$. Obsérvese que lo realisado en los ejemplos anteriores nos da un camino a seguir para calcular los polinomios de Bernstein para funciones $f \in \mathcal{P}_n$. Hay otra forma de calcular estós polinomios la cuál damos a continuación:

El lector cuidadoso se habrá dado cuenta que en algunos de los ejemplos anteriores se utilizó implicitamente el Teorema del Binomio, el cual volveremos a usar aquí; es claro que $(1+t)^n = \sum_{k=0}^{n} \binom{n}{k} t^k$, derivando de ambos lados y multiplicando por t tenemos $nt(1+t)^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} t^k$, repitiendo lo anterior se obtiene

$nt(1+nt)(1+t)^{n-2} = \sum_{k=0}^{n} k^2 \binom{n}{k} t^k$. Haciendo el cambio de variable $t = \frac{x}{1-x}$ con $x \neq 1$, $x \in \mathbb{R}$, obtenemos que $1+t = \frac{1}{1-x}$, entonces multiplicando las tres identidades obtenidas anteriormente por $(1-x)^n$ obtenemos en el mismo orden las siguientes identidades:

$$1 = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}$$

$$nx = \sum_{k=0}^{n} k \binom{n}{k} x^k (1-x)^{n-k}$$

$$nx(1-x+nx) = \sum_{k=0}^{n} k^2 \binom{n}{k} x^k (1-x)^{n-k}$$

Por lo tanto a partir de estás también podemos obtener los resultados de (I.14) y esté método lo podemos repetir para cualquier $f \in \mathcal{P}_n$.

(I.15) Teorema de Bernstein (1912). *Sea $f \in A([0,1])$ entonces se tiene que $B_n(f)(x) \to f(x)$ $\forall x \in [0,1]$ en donde f es continua. Si $f \in C([0,1])$ entonces $||B_n(f) - f||_{[0,1]} \to 0$ si $n \to \infty$.*

Demostración.

Parte I. Calcularemos la siguiente suma: $s(x) = \sum_{k=0}^{n} (k-nx)^2 \binom{n}{k} x^k (1-x)^{n-k}$.

Desarrollando el binomio y por (I.14) tenemos que $s(x) = nx(1-x)$, pero esta última función alcanza su valor máximo en $x = \frac{1}{2}$ en todo el eje real, de donde se tiene que $s(x) \leq \frac{n}{4}$ $\forall x \in [0,1]$.

Parte 2. Dada $\delta > 0$ y $x \in [0,1]$ acotaremos la siguiente suma:

$\hat{s}(x) = \sum_{|\frac{k}{n}-x| \geq \delta} \binom{n}{k} x^k (1-x)^{n-k}$, donde la notación significa que la suma se extiende sobre todos los valores de $k \in \{1, \ldots, n\}$ para los cuáles $|\frac{k}{n} - x| \geq \delta$, de donde $\frac{1}{\delta^2}(\frac{k}{n} - x)^2 \geq 1$. Por lo que:

$$\hat{s}(x) \leq \frac{1}{\delta^2} \sum_{|\frac{k}{n}-x| \geq \delta} \left(\frac{k}{n} - x\right)^2 \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{1}{n^2\delta^2} s(x) \leq \frac{1}{4n\delta^2}$$ (por la parte 1).

Parte 3. Sabemos que $1 = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}$, de donde se sigue que $f(x) =$

$= \sum_{k=0}^{n} f(x)\binom{n}{k}x^k(1-x)^{n-k}$. Como f es acotada existe $M > 0$ tal que $|f(x)| \leq M$ $\forall\ x \in [0,1]$ de donde $|f(x) - f(y)| \leq 2M\ \forall\ x,y \in [0,1]$. Además, dada $\varepsilon > 0$ y $x \in [0,1]$ donde f es continua, existe $\delta'(\varepsilon,x) > 0$ tal que si $y \in [0,1], |y-x| < \delta'$ se tiene que $|f(x) - f(y)| < \frac{\varepsilon}{2}$.

Sea $\delta(\varepsilon,x) = \max\{\delta'(\varepsilon,x), \frac{1}{n^\alpha}\}$ con $\alpha \in (0,\frac{1}{2})$ dada, entonces:

$$|f(x) - B_n(f)(x)| = \left|\sum_{k=0}^{n}\left(f(x) - f\left(\frac{k}{n}\right)\right)\binom{n}{k}x^k(1-x)^{n-k}\right| \leq$$

$$\leq \sum_{k=0}^{n}\left|f(x) - f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k} =$$

$$= \sum_{|\frac{k}{n}-x|<\delta(\varepsilon,x)}\left|f(x) - f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k} +$$

$$+ \sum_{|\frac{k}{n}-x|\geq\delta(\varepsilon,x)}\left|f(x) - f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k}$$

$< \frac{\varepsilon}{2} + \frac{M}{2n^{1-2\alpha}} < \varepsilon$ si n es suficientemente grande, por lo tanto $B_n(f)(x) \to f(x)$ en los puntos de $[0,1]$ donde f es continua.

Para la segunda parte del teorema, si $f \in C([0,1])$ entonces f es uniformemente continua en $[0,1]$, es decir, existe $\delta(\varepsilon) > 0$ tal que $\forall x,y \in [0,1]$, $|x-y| < \delta$ entonces $|f(x) - f(y)| < \frac{\varepsilon}{2}$, así revisando el análisis anterior se observa que:

$|f(x) - B_n(f)(x)| < \varepsilon$ independientemente de $x \in [0,1]$ por lo que $B_n(f)(x) \to f(x)$ uniformemente en $[0,1]$.

Q.E.D.

(I.16) Corolario. (Teorema de Weierstrass). *Dada $f \in C([a,b])$ entonces existe un polinomio $p_n = p_n(\varepsilon)$ para la cuál $||f - p_n||_{[a,b]} < \varepsilon$.*

Demostración. Sea $h:[0,1] \to [a,b]$ dada por $h(x) = a + (b-a)x$ y consideremos $g = f \circ h \in C([0,1])$. Por el teorema anterior dada $\varepsilon > 0$ existe q un polinomio tal que $||g - q||_{[0,1]} < \varepsilon$, sea $p = q \circ h^{-1}$, observemos que p es un polinomio en x (del mismo grado que q) y $||f - p||_{[a,b]} = ||f - q \circ h^{-1}||_{[a,b]} = ||f \circ h - q||_{[0,1]} =$ $= ||g - q||_{[0,1]} < \varepsilon$

Q.E.D.

Para concluir con esta sección haremos la siguiente observación, el Teorema de Bernstein no sólo prueba la existencia de polinomios que aproximan uniformente a la función f dada, sino que también los exhibe explícitamente.

§3 Aproximación simultánea de funciones y sus derivadas

A diferencia de otras formas de aproximación los polinomios de Bernstein son aproximadores bastante "suaves", en el siguiente sentido: si $f \in C^1([0,1])$, nó sólo $B_n(f) \to f$ uniformemente en el intervalo, sino que también $B_n'(f) \to f'$ uniformemente. Así mismo un resultado para derivadas de orden superior es válido. Para ver que la convergencia uniforme en general no garantiza la convergencia de las derivadas veamos el siguiente:

(I.17) Ejemplo. Sea $f_n(x) = \frac{1}{n}\text{sen}(nx)$ con $x \in [0, 2\pi]$, como $|f_n(x)| \leq \frac{1}{n}$ se tiene que $f_n \to 0$ uniformemente en $[0, 2\pi]$ pero $f_n'(x) = \cos nx$ y $f'(x) = 0$ pero f_n' no converge uniformemente a $f = 0$ ya que si consideramos a $x = 0$ tenemos que $f_n'(0) = 1 \neq f'(0) = 0$.

Probaremos algunos resultados previos al teorema que garantiza la aproximación de las derivadas.

(I.18) Teorema del Valor medio extendido.

Sean $f \in C^n([a,b])$ y $x_0, x_1, \ldots, x_n$ $(n+1)$ puntos dados en $[a,b]$ tales que $x_0 = a$ y $x_i = a + ih$ $\forall\, i = 1, \ldots n$ entonces $\Delta^n f(x_0) = h^n f^{(n)}(\xi)$, para alguna $\xi \in (x_0, x_n)$.

Demostración. Si $n = 1$, es claro que el teorema no es más que el TVM-D. Para el caso de $n = 2$ tenemos:

$$\Delta^2 f(x_0) = \Delta\big(\Delta f(x_0)\big) = \Delta\big(f(x_0+h) - f(x_0)\big) = \Delta(f(x_0+h)) - \Delta(f(x_0)) =$$

$$= f(x_0+2h) - f(x_0+h) - (f(x_0+h) - f(x_0)) = \int_0^h f'(x_0+h+s) - f'(x_0+s)ds.$$

Por el TVM-I existe $0 < \xi_1 < h$ tal que:

$\int_0^h f'(x_0+h+s) - f'(x_0+s)ds = [f'(x_0+h+\xi_1) - f'(x_0+\xi_1)]\,h$ y ahora por el TVM-D se tiene que $\Delta^2 f(x_0) = f''(\xi_2)h^2$ con $x_0 < \xi_2 < x_2$. Repitiendo esté procedimiento un número finito de veces llegamos a que $\Delta^n f(x_0) = h^n f^{(n)}(\xi)$, para alguna $\xi \in (x_0, x_n)$.

Q.E.D.

(I.19) Lema. *Sea $p \in \mathbb{N} \cup \{0\}$ y $f\colon [a,b] \to \mathbb{R}$ acotada, entonces:*

$$B_{n+p}^{(p)}(f)(x) = \frac{(n+p)!}{n!}\sum_{t=0}^{n} \Delta^p f\left(\frac{t}{n+p}\right)\binom{n}{t} x^t (1-x)^{n-t} \;\forall\, n \in \mathbb{N}.$$

Demostración. Por la fórmula de Leibnitz sabemos que:

$(uv)^{(p)} = \sum_{j=0}^{p} \binom{p}{j} (u)^{(j)}(v)^{(n-j)}$ aplicandola a:

$B_{n+p}(f)(x) = \sum_{k=0}^{n+p} f\left(\frac{k}{n+p}\right)\binom{n+p}{k} x^k(1-x)^{n+p-k}$ obtenemos:

$$B_{n+p}^{(p)}(f)(x) = \sum_{k=0}^{n+p} f\left(\frac{k}{n+p}\right)\binom{n+p}{k}\sum_{j=0}^{p}\binom{p}{j}(x^k)^{(j)}((1-x)^{n+p-k})^{(n-j)} \quad (*)$$

por otro lado se tiene que:

$(x^k)^{(j)} = k(k-1)\dots(k-j+1)x^{k-j} = \frac{k!}{(k-j)!}x^{k-j}$ si $k-j \geq 0$ **(1)**

$\left[(1-x)^{n+p-k}\right]^{(p-j)} = (-1)^r\left[(x-1)^r\right]^{(s)}$ donde $r = n+p-k$ y $s = p-j$ por la identidad anterior tenemos:

$= (-1)^{p-j}\frac{(n+p-k)!}{(n+j-k)!}(1-x)^{n+j-k}$ con $n+p-k \geq p-j$ **(2)**.

Sustituyendo **(1)** y **(2)** en **(*)** tenemos:

$$B_{n+p}^{(p)}(f)(x) =$$

$$= \sum_{k=0}^{n+p}\sum_{j=0}^{p} f\left(\frac{k}{n+p}\right)\binom{p}{j}\frac{(n+p)!}{(k-j)!(n+j-k)!}x^{k-j}(1-x)^{n+j-k}(-1)^{p-j}$$

donde $0 \leq k-j \leq n$, haciendo $t = k-j$ entonces $k = t+j$, por lo que $0 \leq t \leq n$ y tenemos:

$$B_{n+p}^{(p)}(f)(x) = \sum_{t=0}^{n}\sum_{j=0}^{p} f\left(\frac{t+j}{n+p}\right)\binom{p}{j}\frac{(n+p)!}{t!(n-t)!}x^t(1-x)^{n-t}(-1)^{p-j}$$

$$= \sum_{t=0}^{n}\frac{(n+p)!}{t!(n-t)!}x^t(1-x)^{n-t}\sum_{j=0}^{p} f\left(\frac{t+j}{n+p}\right)\binom{p}{j}(-1)^{p-j}$$

$$= \frac{(n+p)!}{n!}\sum_{t=0}^{n}\Delta^p f\left(\frac{t}{n+p}\right)\binom{n}{t}x^t(1-x)^{n-t}.$$

Q.E.D.

Estamos listos para probar el primer teorema de está sección.

(I.20) Teorema. *Sea* $f \in C^p([0,1])$ *entonces* $B_n^{(p)}(f)(x) \to f^{(p)}(x)$ *uniformemente en* $[0,1]$.

Demostración. Por el lema (I.18) tenemos que $\Delta^p f\left(\frac{t}{n+p}\right) = \frac{1}{(n+p)^p}f^{(p)}(\xi_t)$ para algún $\xi_t \in \left(\frac{t}{n+p}, \frac{t+p}{n+p}\right)$ $\forall t \in \{0,\dots,n\}$. Por el lema (I.19) se tiene que:

$$B_{n+p}^{(p)}(f)(x) = \tfrac{(n+p)!}{n!(n+p)^p} \sum_{t=0}^{n} f^{(p)}(\xi_t) \binom{n}{t} x^t(1-x)^{n-t}$$

así pues:

$$\tfrac{n!(n+p)^p}{(n+p)!} B_{n+p}^{(p)}(f)(x) = \sum_{t=0}^{n} f^{(p)}(\xi_t) \binom{n}{t} x^t(1-x)^{n-t}$$

de donde:

$$\tfrac{n!(n+p)^p}{(n+p)!} B_{n+p}^{(p)}(f)(x) = \sum_{t=0}^{n} f^{(p)}\left(\frac{t}{n}\right) \binom{n}{t} x^t(1-x)^{n-t} +$$

$$+\sum_{t=0}^{n} \left\{ f^{(p)}(\xi_t) - f^{(p)}\left(\frac{t}{n}\right)\right\} \binom{n}{t} x^t(1-x)^{n-t} \qquad \textbf{(1)}$$

Pero como $\frac{t}{n+p} \leq \xi_t \leq \frac{t+p}{n+p}$ entonces $|\xi_t - \frac{t}{n}| \leq \frac{p}{n+p}$.

Dado que $f \in C^p([0,1])$ y por la continuidad uniforme de $f^{(p)}$, dada $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $\forall n \geq n_0$ se tiene que $|f^{(p)}(\xi_t) - f^{(p)}(\frac{t}{n})| < \varepsilon$.

Regresando a **(1)** la segunda suma es menor que ε, luego tomando el límite cuando $n \to \infty$ se tiene que $\lim\limits_{n\to\infty} \dfrac{n!(n+p)^p}{(n+p)!} = 1$, mientras que la primera suma de **(1)**(por el Teorema de Bernstein (I.15)) converge uniformemente en $[0,1]$ a $f^{(p)}$. Por lo tanto $B_{n+p}^{(p)}(f) \to f^{(p)}$ uniformemente en $[0,1]$.

Q.E.D.

Veremos una propiedad adicional de los polinomios de Bernstein.

(I.21) Teorema. *Sea f una función convexa en $[0,1]$. Si $n \geq 2$ entonces $B_{n-1}(f)(x) \geq B_n(f)(x)$, $x \in (0,1)$.*

Si $f \in C([0,1])$, f es lineal en cada intervalo de la forma $[\frac{j-1}{n}, \frac{j}{n}]$ $\forall$ $j \in \{1,\ldots,n-1\}$ si y sólo si $B_{n-1}(f)(x) = B_n(f)(x)$.

Demostración. Dado que $x \in (0,1)$, basta con demostrar que:

$$(1-x)^{-n}\left(B_{n-1}(f)(x) - B_n(f)(x)\right) \geq 0$$

si hacemos el cambio de variable $t = \frac{x}{1-x}$ tenemos $1+t = \frac{1}{1-x}$ y así:

$$(1-x)^{-n}\left(B_{n-1}(f)(x) - B_n(f)(x)\right) = (1+t)\sum_{k=0}^{n-1} f\left(\frac{k}{n-1}\right)\binom{n-1}{k} t^k -$$

$$-\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} t^k = \sum_{k=0}^{n-1} f\left(\frac{k}{n-1}\right)\binom{n-1}{k} t^k +$$

$$+\sum_{k=0}^{n-1} f\left(\frac{k}{n-1}\right)\binom{n-1}{k} t^{k+1} - \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} t^k$$

$$B^{(p)}_{n+p}(f)(x) = \frac{(n+p)!}{n!(n+p)^p} \sum_{t=0}^{n} f^{(p)}(\xi_t) \binom{n}{t} x^t(1-x)^{n-t}$$

así pues:

$$\frac{n!(n+p)^p}{(n+p)!} B^{(p)}_{n+p}(f)(x) = \sum_{t=0}^{n} f^{(p)}(\xi_t) \binom{n}{t} x^t(1-x)^{n-t}$$

de donde:

$$\frac{n!(n+p)^p}{(n+p)!} B^{(p)}_{n+p}(f)(x) = \sum_{t=0}^{n} f^{(p)}\left(\frac{t}{n}\right) \binom{n}{t} x^t(1-x)^{n-t} +$$

$$+ \sum_{t=0}^{n} \left\{ f^{(p)}(\xi_t) - f^{(p)}\left(\frac{t}{n}\right) \right\} \binom{n}{t} x^t(1-x)^{n-t} \qquad \textbf{(1)}$$

Pero como $\frac{t}{n+p} \leq \xi_t \leq \frac{t+p}{n+p}$ entonces $|\xi_t - \frac{t}{n}| \leq \frac{p}{n+p}$.

Dado que $f \in C^p([0,1])$ y por la continuidad uniforme de $f^{(p)}$, dada $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $\forall n \geq n_0$ se tiene que $|f^{(p)}(\xi_t) - f^{(p)}(\frac{t}{n})| < \varepsilon$.

Regresando a **(1)** la segunda suma es menor que ε, luego tomando el límite cuando $n \to \infty$ se tiene que $\lim\limits_{n\to\infty} \dfrac{n!(n+p)^p}{(n+p)!} = 1$, mientras que la primera suma de **(1)**(por el Teorema de Bernstein (I.15)) converge uniformemente en $[0,1]$ a $f^{(p)}$. Por lo tanto $B^{(p)}_{n+p}(f) \to f^{(p)}$ uniformemente en $[0,1]$.

Q.E.D.

Veremos una propiedad adicional de los polinomios de Bernstein.

(I.21) Teorema. *Sea f una función convexa en $[0,1]$. Si $n \geq 2$ entonces $B_{n-1}(f)(x) \geq B_n(f)(x)$, $x \in (0,1)$.*

Si $f \in C([0,1])$, f es lineal en cada intervalo de la forma $[\frac{j-1}{n}, \frac{j}{n}]$ $\forall\ j \in \{1,\ldots,n-1\}$ si y sólo si $B_{n-1}(f)(x) = B_n(f)(x)$.

Demostración. Dado que $x \in (0,1)$, basta con demostrar que:

$$(1-x)^{-n}\left(B_{n-1}(f)(x) - B_n(f)(x)\right) \geq 0$$

si hacemos el cambio de variable $t = \frac{x}{1-x}$ tenemos $1+t = \frac{1}{1-x}$ y así:

$$(1-x)^{-n}\left(B_{n-1}(f)(x) - B_n(f)(x)\right) = (1+t)\sum_{k=0}^{n-1} f\left(\frac{k}{n-1}\right)\binom{n-1}{k} t^k -$$

$$-\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} t^k = \sum_{k=0}^{n-1} f\left(\frac{k}{n-1}\right)\binom{n-1}{k} t^k +$$

$$+\sum_{k=0}^{n-1} f\left(\frac{k}{n-1}\right)\binom{n-1}{k} t^{k+1} - \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} t^k$$

para la segunda suma hacemos $j = k+1$ entonces $j = 1, \ldots, n$ y dejamos a las tres sumas con los mismos índices de 1 a $n-1$:

$$\sum_{k=1}^{n-1} f\left(\frac{k}{n-1}\right)\binom{n-1}{k} t^k +$$

$$+\sum_{j=1}^{n-1} f\left(\frac{j-1}{n-1}\right)\binom{n-1}{j-1} t^j - \sum_{k=1}^{n-1} f\left(\frac{k}{n}\right)\binom{n}{k} t^k +$$

$$+f(0) + f(1)t^n - f(0) - f(1)t^n =$$

$$= \sum_{k=1}^{n}\binom{n}{k}\left[\frac{n-k}{n} f\left(\frac{k}{n-1}\right) + \frac{k}{n} f\left(\frac{k-1}{n-1}\right) - f\left(\frac{k}{n}\right)\right] t^k \quad (*)$$

Observemos que $\frac{n-k}{n}\left(\frac{k}{n-1}\right) + \frac{k}{n}\left(\frac{k-1}{n-1}\right) = \frac{k}{n}$, $\frac{k-1}{n-1} < \frac{k}{n} < \frac{k}{n-1}$ y como f es convexa entonces la expresión dentro de los corchetes de (*) es no negativa por lo que se cumple la primera parte del teorema.

Para probar la última afirmación del teorema, si f es lineal en cada intervalo $[\frac{i-1}{n}, \frac{i}{n}]$ tenemos que se cumple la igualdad dentro de los corchetes de (*) lo cual hace que $B_{n-1}(f)(x) = B_n(f)(x)$.

Inversamente, si $B_{n-1}(f)(x) = B_n(f)(x)$ se tiene $\frac{n-k}{n} f(\frac{k}{n-1}) + \frac{k}{n} f(\frac{k-1}{n-1}) - f(\frac{k}{n}) = 0$ y como $f \in C([0,1])$ y convexa implica que f es lineal en cada intervalo.

Q.E.D.

Probaremos algunos resultados más generales sobre los polinomios de Bernstein.

(I.22) Teorema.

Sea $p \in \mathbb{Z}$ fijo tal que $0 \leq p \leq n$.

(i) Si $m \leq f^{(p)}(x) \leq M$ $\forall x \in [0,1]$ entonces

$m \leq \frac{n^p}{n(n-1)\ldots(n-p+1)} B_n^{(p)}(f)(x) \leq M \ \forall\, x \in [0,1]$

(ii) Si $f^{(p)}(x) \geq 0$ en $[0,1]$ entonces $B_n^{(p)}(f)(x) \geq 0$ en $[0,1]$.

(iii) Si f es no decreciente en $[0,1]$ entonces $B_n(f)(x)$ es no decreciente.

(iv) Si f es convexa en $[0,1]$ entonces $B_n(f)(x)$ es convexa.

Demostración.

(i) Por (I.19) se tiene que:

$$B_n^{(p)}(f)(x) =$$

$$= n(n-1)\ldots(n-p+1)\sum_{t=0}^{n-p} \Delta^p f\left(\frac{t}{n}\right)\binom{n-p}{t} x^t(1-x)^{n-p-t} \quad (*)$$

Por el Lema (I.18) tenemos que: $\Delta^p f(\frac{t}{n}) = \frac{1}{n^p} f^{(p)}(\xi_t)$ con $\frac{t}{n} < \xi_t < \frac{t+p}{n}$ y la cuál es clara para $p = 0$ tomando $\xi_t = \frac{t}{n}$.

Entonces, sustituyendo en $(*)$ se tiene:

$$B_n^{(p)}(f)(x) = \frac{n(n-1)\ldots(n-p+1)}{n^p} \sum_{t=0}^{n-p} f^{(p)}(\xi_t) \binom{n-p}{t} x^t(1-x)^{n-p-t} \quad (1)$$

de **(1)** observemos que el factor de la suma es menor que 1, de donde:

$$B_n^{(p)}(f)(x) \leq \sum_{t=0}^{n-p} f^{(p)}(\xi_t) \binom{n-p}{t} x^t(1-x)^{n-p-t}$$

pero $m \leq f^{(p)}(\xi_t) \leq M$ por hipótesis, de donde se sigue que: $B_n^{(p)}(f)(x) \leq M$.

De **(1)** obtenemos:

$$\frac{n^p}{n(n-1)\ldots(n-p+1)} B_n^{(p)}(f)(x) = \sum_{t=0}^{n-p} f^{(p)}(\xi_t) \binom{n-p}{t} x^t(1-x)^{n-p-t}$$

y dado que $m \leq f^{(p)}(\xi_t) \leq M$ y los factores que aparecen en la última suma son positivos tenemos que:

$$m \leq \frac{n^p}{n(n-1)\ldots(n-p+1)} B_n^{(p)}(f)(x) \leq M.$$

(ii) Cómo $f^{(p)}(x) \geq 0$ haciendo $m = 0$ en el inciso anterior se sigue:

$\frac{n^p}{n(n-1)\ldots(n-p+1)} B_n^{(p)}(f)(x) \geq 0$ de donde $B_n^{(p)}(f)(x) \geq 0$.

(iii) Si f es no decreciente en $[0,1]$ entonces $\Delta f(\frac{t}{n}) = f(\frac{t+1}{n}) - f(\frac{t}{n}) \geq 0$ de $(*)$ con $p = 1$ se obtiene que $B_n'(f)(x) \geq 0$ en $[0,1]$ lo cuál implica que $B_n(f)(x)$ es no decreciente.

(iv) Si f es una función convexa en $[0,1]$ entonces :

$\Delta^2 f(\frac{t}{n}) = f(\frac{t+2}{n}) - f(\frac{t+1}{n}) - (f(\frac{t+1}{n}) - f(\frac{t}{n})) \geq 0$ por $(*)$ con $p = 2$ se tiene que $B_n''(f)(x) \geq 0$ en cada subintervalo cerrado de $(0,1)$ y como $B_n(f)(x)$ es continua se concluye que es convexa en $[0,1]$.

Q.E.D.

§4 Algunas otras propiedades de los polinomios de Bernstein

Despúes de haber visto algunas propiedades de los polinomios de Bernstein, ahora nos dedicaremos a estudiar brevemente la rapidez con la cuál convergen a una función f acotada en un punto x_0, por lo que probaremos el siguiente:

(I.23) Lema. *Para toda $x \in [0,1]$ existe una constante positiva c independiente de $n \in \mathbb{N}$ tal que:*

$$\sum_{|\frac{k}{n}-x|\geq n^{-\alpha}} \binom{n}{k} x^k(1-x)^{n-k} \leq \frac{c}{n^{3(1-2\alpha)}}$$

$\forall\, \alpha \in [0,\frac{1}{2})$.

Demostración. Consideremos las sumas:

$$S_m(x) = \sum_{k=0}^{n}(k-nx)^m \binom{n}{k} x^k(1-x)^{n-k} \quad \mathbf{(1)}$$

Por lo hecho en los ejemplos de (I.14) tenemos que: $S_0(x)=1, S_1(x)=0$ y $S_2(x)=nx(1-x)$. Diferenciando **(1)** con respecto a x.

$$S'_m(x) = -nm\sum_{k=0}^{n}\binom{n}{k}(k-nx)^{m-1}x^k(1-x)^{n-k}+$$

$$+\sum_{k=0}^{n}\binom{n}{k}(k-nx)^{m+1}x^{k-1}(1-x)^{n-k-1}$$

$= -nmS_{m-1}(x)+\frac{S_{m+1}(x)}{x(1-x)}$, de donde se tiene que:

$$S_{m+1}(x) = x(1-x)[S'_m(x)+mnS_{m-1}(x)] \;\; \forall x \in [0,1] \quad \mathbf{(2)}.$$

Observemos de las expresiones anteriores para $S_0(x), S_1(x), S_2(x)$ y **(2)** que $S_m(x)$ es un polinomio en x y n. Así tenemos que $S_3(x)=nx(1-x)(1-2x)$, por lo que $S_3(x)$ es un polinomio de primer grado en n. De **(2)** con $m=3$ y de lo anterior podemos observar que $S_4(x)$ es un polinomio de grado 2 en n. Continuando de está manera para $S_5(x)$ y $S_6(x)$ se tiene que son polinomios de grado 2 y 3 en n respectivamente. Por lo tanto existe $c=cte$ tal que $|S_6(x)|\leq cn^3 \;\forall\, x\in[0,1]$ **(3)**.

Por otro lado como $|\frac{k}{n}-x|\geq n^{-\alpha}$ se sigue que $n^{6(\alpha-1)}(k-nx)^6\geq 1$ por lo que:

$$\sum_{|\frac{k}{n}-x|\geq n^{-\alpha}}\binom{n}{k}x^k(1-x)^{n-k}\leq S_0(x)\leq n^{6(\alpha-1)}S_6(x).$$

Por **(3)** se tiene que:

$$\sum_{|\frac{k}{n}-x|\geq n^{-\alpha}} \binom{n}{k} x^k(1-x)^{n-k} \leq \frac{c}{n^{3(1-2\alpha)}} \ \forall x \in [0,1], \text{ y } \alpha \in [0, \tfrac{1}{2}).$$

Q.E.D.

Probaremos el siguiente resultado que da una respuesta a la pregunta hecha al principio de esta sección:

(I.24) Teorema. (E. Voronosky) (1932) *Sea $f:[0,1] \to \mathbb{R}$ acotada y $x_0 \in [0,1]$ tal que $f''(x_0)$ existe. Entonces*

$$\lim_{n\to\infty} n\,[B_n(f)(x_0) - f(x_0)] = \frac{1}{2}x_0(1-x_0)f''(x_0).$$

Demostración. Por el Teorema de Taylor sabemos que:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + e(x)(x - x_0)^2$$

con $\lim_{x\to x_0} e(x) = 0$ **(1)**. Considerando $x = \frac{k}{n}$ entonces

$$f\left(\frac{k}{n}\right) = f(x_0) + f'(x_0)\left(\frac{k}{n} - x_0\right) + \frac{f''(x_0)}{2}\left(\frac{k}{n} - x_0\right)^2 + e\left(\frac{k}{n}\right)\left(\frac{k}{n} - x_0\right)^2$$

multiplicando la expresión anterior por $\binom{n}{k} x_0^k(1-x_0)^{n-k}$ y sumando desde $k = 0$ hasta $k = n$ tenemos:

$$B_n(f)(x) = f(x_0)S_0(x_0) + \tfrac{f'(x_0)}{n}S_1(x_0) + \tfrac{f''(x_0)}{2n^2}S_2(x_0) + E(x) \quad \textbf{(2)}$$

donde $S_m(x)$ con $m = 0, 1, 2$ son las funciones definidas en la demostración del lema (I.23) y

$$E(x) = \sum_{k=0}^{n} e\left(\frac{k}{n}\right)\left(\frac{k}{n} - x_0\right)^2 \binom{n}{k} x_0^k(1 - x_0)^{n-k}$$

Dada $\varepsilon > 0$, por **(1)**, existe $n = n(\varepsilon) \in \mathbb{N}$ tal que si $|x - x_0| < \frac{1}{n^\alpha}$ con $\alpha \in [0, \frac{1}{3})$ entonces $|e(x)| < \varepsilon$. Por lo que se tiene:

$$|E(x)| \leq \sum_{|\frac{k}{n}-x_0|<n^{-\alpha}} \left|e\left(\frac{k}{n}\right)\right| \left(\frac{k}{n} - x_0\right)^2 \binom{n}{k} x_0^k(1 - x_0)^{n-k} +$$

$$+ \sum_{|\frac{k}{n}-x_0|\geq n^{-\alpha}} \left|e\left(\frac{k}{n}\right)\right| \left(\frac{k}{n} - x_0\right)^2 \binom{n}{k} x_0^k(1 - x_0)^{n-k} \leq$$

$\leq \frac{\varepsilon}{n^2} S_2(x) + \frac{Mc}{n^{3(1-2\alpha)}}$, por el lema (I.23) y donde $M = \sup\{|e(x)|(x-x_0)^2 \mid x \in [0,1]\}$.
Pero $S_2(x_0) = nx_0(1-x_0)$, por lo que:

$$|E(x)| \leq \frac{\varepsilon}{n} x_0(1-x_0) + \frac{Mc}{n^{3(1-2\alpha)}}$$

Despejando de **(2)** y multiplicando por n tenemos:

$\left|n\left(B_n(f)(x_0) - f(x_0)\right) - \frac{1}{2} f''(x_0) x_0(1-x_0)\right| = |nE(x)|$

$\leq \varepsilon x_0(1-x_0) + \frac{ncM}{n^{3(1-2\alpha)}} \leq \frac{\varepsilon}{4} + \frac{ncM}{n^{3(1-2\alpha)}}$, de donde $\frac{ncM}{n^{3(1-2\alpha)}}$ tiende a cero conforme $n \to \infty$, y como ε es arbitraria entonces se sigue el resultado.

Q.E.D.

Vamos analizar también la rapidez con la cuál los polinomios de Bernstein se aproximan a las funciones de cierta clase, para lo cuál necesitamos la siguiente:

(I.25) Definición. Sea $f: [a,b] \to \mathbb{R}$. Si existen $M > 0$ y $\alpha > 0$ constantes tales que $|f(x) - f(y)| \leq M|x-y|^\alpha \ \forall x, y \in [a,b]$ entonces decimos que f satisface la ***condición de Lipschitz de orden α en $[a,b]$***. La clase de tales funciones se denotará por $\mathrm{Lip}_M(\alpha)$. Por comodidad cuando $M = 1$ simplemente se denotará por $\mathrm{Lip}(\alpha)$.

Ahora demostraremos el siguiente:

(I.26) Teorema. *Sea $f: [0,1] \to \mathbb{R}$ una función tal que $f \in \mathrm{Lip}_M(1)$ entonces* $\|B_n(f) - f\|_{[0,1]} \leq \frac{5M}{4\sqrt{n}} \forall\, n \in \mathbb{N}$.

Demostración. Sea $x_0 \in [0,1]$ fija y $x \in [0,1]$. Si $|x - x_0| > \frac{1}{\sqrt{n}}$ entonces $|x-x_0|^2 > \frac{1}{\sqrt{n}}|x-x_0|$ como $f \in \mathrm{Lip}_M(1)$, entonces $|f(x) - f(x_0)| \leq M|x-x_0| =$
$= \frac{1}{\sqrt{n}} M\sqrt{n}|x-x_0| < M|x-x_0|^2\sqrt{n} = \frac{1}{\sqrt{n}} Mn|x-x_0|^2 < \frac{M}{\sqrt{n}}\left(1 + n|x-x_0|^2\right)$.

Si $|x - x_0| \leq \frac{1}{\sqrt{n}}$ es claro que se cumple la desigualdad anterior. Por lo tanto se tiene que: $|B_n(f)(x_0) - f(x_0)| \leq M \sum_{k=0}^{n} \frac{1}{\sqrt{n}} \left(1 + n\left|\frac{k}{n} - x_0\right|^2\right) \binom{n}{k} x_0^k (1-x_0)^{n-k}$

$= M\left[\frac{1}{\sqrt{n}}\left(1 + \frac{S_2(x)}{n}\right)\right]$, por la parte 1 de (I.15) tenemos que
$|B_n(f)(x_0) - f(x_0)| \leq \frac{5M}{4\sqrt{n}}$.

Q.E.D.

Por el teorema anterior tenemos que el orden de magnitud con el cuál $B_n(f)(x)$ converge a $f(x)$ es $O(\frac{1}{\sqrt{n}})$, donde $O(\frac{1}{\sqrt{n}}) = B_n(f)(x) - f(x)$ significa que existe $M > 0$ tal que $\sqrt{n}|B_n(f)(x) - f(x)| \leq M \ \forall n \geq N(M)$. Para concluir con está sección veremos un ejemplo en donde áquel es el mejor orden de magnitud, para lo cuál probaremos el siguiente:

(I.27) Lema. *Sea $f(x) = |x - \frac{1}{2}|$ en $[0,1]$ entonces:*

$$B_{2n}(f)\left(\frac{1}{2}\right) = B_{2n+1}(f)\left(\frac{1}{2}\right) = \frac{1}{2^{2n+1}}\binom{2n}{n}.$$

Demostración.

Claramente: $\binom{n}{k} = \binom{n}{n-k}$ y $k\binom{n}{k} = (n-k+1)\binom{n}{n-k+1}$.

Por lo que:

$$B_{2n}(f)(\tfrac{1}{2}) = \sum_{k=0}^{2n}\left|\frac{k}{2n} - \frac{1}{2}\right|\binom{2n}{k}\frac{1}{2^{2n}} =$$

$$= \sum_{k=0}^{n}\left(\frac{1}{2} - \frac{k}{2n}\right)\binom{2n}{k}\frac{1}{2^{2n}} + \sum_{k=n+1}^{2n}\left(\frac{k}{2n} - \frac{1}{2}\right)\binom{2n}{k}\frac{1}{2^{2n}} =$$

$$= \frac{1}{2^{2n+1}}\left[\sum_{k=0}^{n}\binom{2n}{k} - \sum_{k=n+1}^{2n}\binom{2n}{k}\right] +$$

$$+ \frac{1}{n2^{2n+1}}\left[-\sum_{k=0}^{n}k\binom{2n}{k} + \sum_{k=n+1}^{2n}k\binom{2n}{k}\right] =$$

$$= \frac{1}{2^{2n+1}}\left[\sum_{k=0}^{n}\binom{2n}{k} - \sum_{k=n+1}^{2n}\binom{2n}{n-k}\right] +$$

$$+ \frac{1}{n2^{2n+1}}\left[-\sum_{k=0}^{n}k\binom{2n}{k} + \sum_{k=n+1}^{2n}(2n-k+1)\binom{2n}{2n-k+1}\right] \qquad (1).$$

Por otro lado considerando a:

$$\sum_{k=n+1}^{2n}\binom{2n}{2n-k} = \sum_{j=0}^{n-1}\binom{2n}{j} \text{ y } \sum_{k=n+1}^{2n}(2n-k+1)\binom{2n}{2n-k+1} = \sum_{j=1}^{n}j\binom{2n}{j}$$

Por lo tanto $B_{2n}(f)(\frac{1}{2}) = \frac{1}{2^{2n+1}}\binom{2n}{n}$ ya que el último sumando de **(1)** es cero.

Procediendo de manera análoga se demuestra que $B_{2n+1}(f)(\frac{1}{2}) = \frac{1}{2^{2n+1}}\binom{2n}{n}$.

Q.E.D.

(I.28)Ejemplo. Sea $f(x) = |x - \frac{1}{2}|$, es claro que $f \in \mathrm{Lip}_1(1)$, sea $x_0 = \frac{1}{2}$ entonces:

$$\sqrt{n}\left|B_n(f)\left(\frac{1}{2}\right) - f\left(\frac{1}{2}\right)\right| = \sqrt{n}B_n(f)\left(\frac{1}{2}\right)$$

el cuál es igual a $\frac{\sqrt{2M}}{2^{2M+1}}\binom{2M}{M}$ si $n = 2M$ ó a $\frac{\sqrt{2M+1}}{2^{2M+1}}\binom{2M}{M}$ si $n = 2M+1$ (ver (I.27)). Por la fórmula de Stirling se tiene:

$$\frac{\sqrt{2M}}{2^{2M+1}}\binom{2M}{M} > \frac{1}{2\sqrt{\pi}X_M^2}$$

Donde $\{X_M\}$ es la sucesión $\{\,(1+\frac{1}{4M})^2\}$, como $\{X_M\}$ es decreciente, se tiene que $x_M^2 \leq (\frac{5}{4})^2\ \forall\ M \in \mathbb{N}$, de lo anterior se sigue que:

$$\frac{\sqrt{2M}}{2^{2M+1}}\binom{2M}{M} > \frac{8}{25\sqrt{\pi}} > 0$$

dado que $\frac{\sqrt{2M+1}}{2^{2M+1}}\binom{2M}{M} > \frac{\sqrt{2M}}{2^{2M+1}}\binom{2M}{M}$ obtenemos la misma cota inferior para ambos casos de $\sqrt{n}B_n(f)(\frac{1}{2})$. Por lo tanto el orden de magnitud de convergencia de $B_n(f)$ a nuestra f no es menor que $\frac{1}{\sqrt{n}}$.

Q.E.D.

§5 Una vez más el Teorema de Weierstrass

En esta sección daremos otra prueba al Teorema de Weierstrass, debida a Lebesgue (1908).

(I.29) Lema. *Sea* $f:[-1,1]\to\mathbb{R}$ *dada por* $f(x)=|x|$ *entonces existe una sucesión* (p_n) *de polinomios tal que* $\|p_n - f\|_{[-1,1]}\to 0$.

Demostración. Definimos la sucesión (p_n) como sigue: $p_0(x)=0$ y
$p_{n+1}(x)=p_n(x)+\frac{1}{2}\left[x^2-p_n^2(x)\right]$, $\forall n\in\mathbb{N}$
Ahora veamos si $|x|\le 1$ entonces es fácil verificar (por inducción) que se cumplen:
(i) $0\le p_n(x)\le |x|$
(ii) $p_n(x)\le p_{n+1}(x)$
(iii) $|x|-p_{n+1}(x)\le |x|\left(1-\frac{|x|}{2}\right)^n$.

De (iii) tenemos que si $x\ne 0$ entonces:
$||x|-p_{n+1}(x)|=|x|-p_{n+1}(x)\le |x|\left(1-\frac{|x|}{2}\right)^n\le\left(1-\frac{|x|}{2}\right)^n$, como $|x|\le 1$ entonces $0\le 1-\frac{|x|}{2}<1$ y dado que $c^n\to 0$ cuando $n\to\infty$ para toda $|c|<1$, mientras que si $x=0$ por (i) se tiene que $p_n(x)=|x|=0$ por lo que $\|f-p_n\|_{[-1,1]}\to\infty$.

Q.E.D.

El Lema anterior nos garantiza que $|x|$ puede aproximarse uniformemente por polinomios en $[-1,1]$.

Antes de dar la prueba de Lebesgue, consideremos $f\in C([0,1])$ y $N\in\mathbb{N}$, definimos $L:[0,1]\to\mathbb{R}$, cómo sigue:
(i) $L(\frac{k}{N})=f(\frac{k}{N})$ $\forall k\in\{0,\dots N\}$
(ii) $L(x)$ es el valor correspondiente del segmento de recta si $x\in(\frac{k}{N},\frac{k+1}{N})$ **(1)**.
Ahora encontraremos una regla de correspondencia conveniente para nuestros propósitos. Primero consideremos a funciones del siguiente estilo: Sea $c\in[0,1]$ y $\phi:[0,1]\to\mathbb{R}$ definida como $\phi(x)=a+\frac{1}{2}\{x-c+|x-c|\}$ con $a\in\mathbb{R}$. Con ayuda de estas funciones se demuestra que $L(x)=\sum_{k=0}^{N-1}c_k\left|x-\frac{k}{N}\right|$, así mismo se obtiene que:

$$c_k=f\left(\frac{k+1}{N}\right)-2f\left(\frac{k}{N}\right)+f\left(\frac{k-1}{N}\right)\ \forall\, k\in\{1,\dots,N-1\}$$

y el coeficiente c del término independiente en L es:

$$c=\frac{1}{2}\left[f(0)+Nf\left(\frac{N-1}{N}\right)-(N-1)f(1)\right]$$

y calculando el coeficiente c_0 del término lineal en la misma suma se obtiene:

$$c = \frac{N}{2}\left[f\left(\frac{1}{N}\right) - f(0) - f\left(\frac{N-1}{N}\right) + f(1)\right]$$

Daremos la demostración de Lebesgue al Teorema de Weierstrass (1908):

Demostración.

Sin pérdida de generalidad supondremos que $f \in C([0,1])$, entonces dada $\varepsilon > 0$ existe $\delta(\varepsilon) > 0$ tal que $|x-y| \leq \delta$, $x, y \in [0,1]$ se tiene que $|f(x) - f(y)| < \frac{\varepsilon}{2}$. Sea $N \in \mathbb{N}$ tal que $\frac{1}{N} \leq \delta$ y consideremos a $L: [0,1] \to \mathbb{R}$ tal que cumpla con las condiciones de (1). Por lo que si $x \in [\frac{k}{N}, \frac{k+1}{N}]$ para alguna $k \in \{0, \ldots, N-1\}$ entonces:

$$|f(x) - L(x)| \leq |f(x) - f(\tfrac{k}{N})| + |f(\tfrac{k}{N}) - L(\tfrac{k}{N})| + |L(x) - L(\tfrac{k}{N})| \leq$$
$$\leq |f(x) - f(\tfrac{k}{N})| + |L(\tfrac{k+1}{N}) - L(\tfrac{k}{N})| < \varepsilon.$$

Por lo tanto $L \to f$ uniformemente en $[0,1]$.

Por lo anterior a la demostración tenemos que : $L(x) = \sum_{k=0}^{N-1} c_k|x - \frac{k}{N}| + c$, por el lema (1.29), $|x|$ se aproxima uniformemente con polinomios en $[-1,1]$, entonces $|x-\lambda|$ se aproxima uniformemente por polinomios en $[0,1]$ para toda $\lambda \in [0,1]$ ya que si $\big|p(x) - |x|\big| < \varepsilon$ en $[-1,1]$ tenemos que $\big|p(x-\lambda) - |x-\lambda|\big| < \varepsilon$ en $[0,1]$ por lo que cada término de $L(x)$ se puede aproximar uniformemente por polinomios en $[0,1]$ lo cuál junto con el hecho de que $L \to f$ uniformemente en $[0,1]$ implica que f puede aproximarse uniformemente por polinomios en $[0,1]$.

Q.E.D.

CAPITULO II

APLICACIONES DEL TEOREMA DE WEIERSTRASS

En esté capítulo veremos algunas extenciones del Teorema de Weierstrass, como el Teorema de Stone-Weierstrass y algunas de sus aplicaciones. Así como los Teoremas de Extensión de Tietze y un teorema de interpolación de Walsh.

§1 Teorema de Stone-Weierstrass

Para facilitarnos el análisis introduciremos algunas definiciones.

(II.1) Definición. Sea X un espacio vectorial sobre un campo $\mathcal{K}$. Decimos que $(X,\cdot)$ es un álgebra, donde $(\cdot)$ es una función de $X \times X$ a X, si se cumplen las siguientes condiciones:

(i) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$

(ii) $x \cdot (y + z) = x \cdot y + x \cdot z$ y $(x + y) \cdot z = x \cdot z + y \cdot z$

(iii) $\alpha(x \cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y)$

$\forall x, y, z \in X$ y $\alpha \in \mathcal{K}$. A la función $(\cdot)$ se le llama *multiplicación* en X.

(II.2) Ejemplos.

(i) Sea $X = \mathbf{F}_{p,1}(D)$ con las operaciones usuales.

(ii) Sea $X = \mathcal{M}_{n\times n}(\mathbb{R})$ con las operaciones usuales.

(II.3) Definición. Una familia $\Im \subset \mathbf{F}_{p,q}(D)$ es una *celosía (o lattice)* si para $f, g \in \Im$ entonces $f \vee g, f \wedge g \in \Im$ donde $(f \vee g)(x) = \max\{f(x), g(x)\}$ y $(f \wedge g)(x) = \min\{f(x), g(x)\}$. Si además $\Im$ es un espacio vectorial entonces diremos que $\Im$ es una *celosía vectorial.*

(II.4) Observación.

Es fácil verificar que $f \vee (g \vee h) = (f \vee g) \vee h$, así como para $\wedge$, por lo que tiene perfecto sentido hablar de $f \vee g \vee h$ y $f \wedge g \wedge h$.

A partir de (II.4) tenemos que si $\Im$ es una celosía y $f_1, \ldots, f_n \in \Im$ entonces $f_1 \vee \ldots \vee f_n \in \Im$ y $f_1 \wedge \ldots \wedge f_n \in \Im$ donde $f_1 \vee \ldots \vee f_n = \max\{f_1, \ldots, f_n\}$ y $f_1 \wedge \ldots \wedge f_n = \min\{f_1, \ldots, f_n\}$.

(II.5) Ejemplos.

Dado que $\max\{a,b\} = \frac{1}{2}\{a+b+|a-b|\}$ y $\min\{a,b\} = \frac{1}{2}\{a+b-|a-b|\}$ $\forall a,b \in \mathbb{R}$ tenemos que $\mathbb{F}_{p,1}(D)$ es un celosía vectorial, así como $C_{p,1}(D)$.

Ahora probaremos el siguiente resultado:

(II.6) Lema. *Sean* $\Im \subset C_{p,1}(K)$, $K \subset \mathbb{R}^p$ *compacto, tal que si* $f_1, f_2 \in \Im$ *entonces* $f_1 \vee f_2 \in \Im$ *y* $f_1 \wedge f_2 \in \Im$. *Dada* $f \in C_{p,1}(K)$ *entonces* f *se aproxima uniformemente por elementos de* $\Im$ *si y sólo si* $\forall x_1, x_2 \in K$ *y* $\varepsilon > 0$ *existe* $g(\varepsilon) \in \Im$ *tal que* $|f(x_i) - g(x_i)| < \varepsilon$ $i = 1,2$.

Demostración. Si f se aproxima uniformemente por elementos de $\Im$, dada $\varepsilon > 0$ es claro que $\forall x_1, x_2 \in K$ existe $g(\varepsilon) \in \Im$ tal que $|f(x_i) - g(x_i)| < \varepsilon$ $i = 1,2$.

Inversamente, sean $x \in K$ y $\varepsilon > 0$ fijas. Entonces por hipótesis para toda $y \in K$ existe $g_{xy} \in \Im$ tal que $|f(x) - g_{xy}(x)| < \varepsilon$ y $|f(y) - g_{xy}(y)| < \varepsilon$ de donde se tiene que $g_{xy}(y) < f(y) + \varepsilon$ **(1)**.

Por la continuidad de f y g_{xy} existe $B_{r_y}(y)$ tal que **(1)** se cumple, considerando la cubierta abierta $G = \{B_{r_y}(y)\}$, del compacto K, existen $y_1, \ldots, y_n \in K$ tales que $K \subset \bigcup_{i=1}^{n} B_{r_{y_i}}(y_i)$, así cómo $g_{xy_i} \in \Im$ y $g_{xy_i}(z) < f(z) + \varepsilon$ con $z \in B_{r_{y_i}}(y_i)$ **(2)**.

Sea $g_x^- = g_{xy_1} \wedge \ldots \wedge g_{xy_n} \in \Im$ por hipótesis, por **(2)** tenemos que $g_x^-(z) < f(z) + \varepsilon$, $z \in K$, por construcción $|f(x) - g_{xy_i}(x)| < \varepsilon$ para toda $i \in \{1, \ldots, n\}$ entonces $g_{xy_i}(x) > f(x) - \varepsilon$, por la definición de g_x^- tenemos que: $g_x^-(x) > f(x) - \varepsilon$ **(3)**. Por la continuidad de g_x^- y f existe $B_{r_x}(x)$ tal que $g_x^-(z) > f(z) - \varepsilon$, $z \in B_{r_x}(x)$.

Dado que $x \in K$ fue arbitrario, consideramos a $\bigcup_{x \in K} B_{r_x}(x)$ la cuál es una cubierta abierta de K entonces existen $x_1, \ldots, x_m \in K$ tales que $K \subset \bigcup_{i=1}^{m} B_{r_{x_i}}(x_i)$ donde $g_{x_i}^-(z) > f(z) - \varepsilon$ para toda $z \in B_{r_{x_i}}(x_i)$ por lo que está última desigualdad se cumple para alguna $i \in \{1, \ldots, m\}$.

Sea $g = g_{x_1}^- \vee \ldots \vee g_{x_m}^- \in \Im$ por hipótesis, por lo anterior tenemos que $g(z) > f(z) - \varepsilon$, $z \in K$, por **(3)** $g_{x_i}^-(z) < f(z) + \varepsilon$ $z \in K$ y para toda $i = 1, \ldots, m$ por lo que $g(z) < f(z) + \varepsilon$. Por lo tanto $|f(z) - g(z)| < \varepsilon$ con $z \in K$ y así $\|f - g\|_K < \varepsilon$.

Q.E.D.

(II.7) Observación.

Notemos que al final de la demostración anterior la desigualdad estricta se preserva debido a que $|f(z) - g(z)| < \varepsilon$ con $z \in K$ y a la continuidad de $f - g$ en el compacto.

(II.8) Teorema de Aproximación de M. Stone (1937)

Sea $K \subset \mathbb{R}^p$ y $\Im \subset C_{p,1}(K)$ con las siguientes propiedades:

(i) $\Im$ *es una celosía*

(ii) Si $a, b \in \mathbb{R}$ y $x, y \in K$, $x \neq y$ entonces existe $f \in \Im$ tal que $f(x) = a$ y $f(y) = b$ (diremos que $\Im$ **distingue puntos en** *K).*

entonces dadas $\varepsilon > 0$ y $F \in C_{p,1}(K)$ existe $h = h(\varepsilon) \in \Im$ tal que $||F - h||_K < \varepsilon$.

Demostración. Dadas $x, y \in K$ por (ii) existe $g_{xy} \in \Im$ tal que $g_{xy}(x) = F(x)$, $g_{xy}(y) = F(y)$, la cuál cumple que si $\varepsilon > 0$ y $\forall F \in C_{p,1}(K)$, $|F(x) - g_{xy}(x)| < \varepsilon$ y $|F(y) - g_{xy}(y)| < \varepsilon$ por (II.6) tenemos que $||F - g||_K < \varepsilon$ para alguna $g \in \Im$.

Q.E.D.

(II.9) Corolario. *Sea $F: [a, b] \to \mathbb{R}$ continua entonces dada $\varepsilon > 0$ existe $h: [a, b] \to \mathbb{R}$ continua y lineal por tramos tal que $||F - h||_{[a,b]} < \varepsilon$.*

Demostración. Sea $\Im = \{f: [a, b] \to \mathbb{R} \mid f \text{es continua y lineal por tramos}\}$. Es claro que $\Im \subset C([a, b])$, así como si $f \in \Im$ entonces $|f| \in \Im$, así mismo no es díficil ver que $\Im$ es un espacio vectorial, a partir de eso es claro que $f \vee g$ y $f \wedge g$ pertenecen a $\Im$ por lo que $\Im$ es una celosía.

Para ver que $\Im$ distingue puntos, consideramos lo siguiente:

dados $x, y \in [a, b]$, $x < y$ y $A, B \in \mathbb{R}$ definimos la siguiente función:

$$h(z) = \begin{cases} A & \text{si } z \in [a, x] \\ \frac{B-A}{y-x}(z - x) + A & \text{si } z \in [x, y] \\ B & \text{si } z \in [y, b]. \end{cases}$$

Es claro que $h \in \Im$. Por lo tanto el resultado se sigue de (II.8).

Q.E.D.

Para la demostración del Teorema de Stone-Weierstrass veremos el siguiente:

(II.10) Lema. *Sean $D \subset \mathbb{R}^p$ y $\mathcal{A} \subset C_{p,1}(D)$ con las siguientes propiedades:*

(i) $\mathcal{A}$ es un álgebra de funciones

(ii) Para toda $x \in D$ existe $g \in \mathcal{A}$ tal que $g: D \subset \mathbb{R}^p \to \mathbb{R}$ y $g(x) \neq 0$ (diremos que $\mathcal{A}$ **no se anula en** *D).*

(iii) Si $x \neq y$ son dos puntos de D, existe una función $f \in \mathcal{A}$ tal que $f(x) \neq f(y)$ (diremos que $\mathcal{A}$ **separa los puntos de** *D)*

entonces $\mathcal{A}$ distingue puntos en D.

Demostración. Por hipótesis dados $x, y \in D$ $x \neq y$ existen $g_1, g_2 \in \mathcal{A}$ tales que $g_1(x)$ y $g_2(y)$ son distintos de cero, así cómo $f \in \mathcal{A}$ tal que $f(x) \neq f(y)$ de donde $g_1(x)f(x) \neq g_1(x)f(y)$ y $g_2(y)f(x) \neq g_2(y)f(y)$ por lo que consideramos las funciones:

$$h_1(z) = g_1(z)f(z) - g_1(z)f(y) \in \mathcal{A}$$

$$h_2(z) = g_2(z)f(z) - g_2(z)f(x) \in \mathcal{A}$$

de donde es claro que $h_1(x) \neq 0 \neq h_2(y)$ mientras que $h_1(y) = 0 = h_2(x)$. Por lo que dados $a, b \in \mathbb{R}$ considerando la función:

$$H(z) = a\frac{h_1(z)}{h_1(x)} + b\frac{h_2(z)}{h_2(y)} \in \mathcal{A}$$

tenemos que $H(x) = a$ y $H(y) = b$ por lo que $\mathcal{A}$ distingue puntos en D.

Q.E.D.

Veremos la demostración del Teorema de Stone-Weierstras, el cuál difiere del Teorema de Stone, en pedir condiciones algebraicas a conjunto de funciones $\mathcal{A}$.

(II.11) Teorema de Stone-Weierstrass (1948). *Sea $K \subset \mathbb{R}^p$ compacto y $\mathcal{A} \subset C_{p,1}(K)$ con las siguientes propiedades:*

(i) $\mathcal{A}$ es un álgebra de funciones

(ii) $\mathcal{A}$ no se anula en K

(iii) $\mathcal{A}$ separa los puntos de K

entonces dadas $\varepsilon > 0$ y $F \in C_{p,1}(K)$ existe $h \in \mathcal{A}$ tal que $||F - h||_K < \varepsilon$.

Demostración. Sean $\varepsilon > 0$, $F \in C_{p,1}(K)$ y

$\mathfrak{I} = \{f \in C_{p,1}(K) \mid$ existe $(H_n) \subset \mathcal{A}$ tal que $H_n \to f$ unif. en $K\} \subset C_{p,1}(K)$.

Probaremos que si $h \in \mathfrak{I}$ entonces $|h| \in \mathfrak{I}$, para lo cuál procedemos de la siguiente manera:

Dado que $h \in \mathfrak{I}$ entonces:

(a) $h \in C_{p,1}(K)$

(b) existe $(h_n) \subset \mathcal{A}$ tal que $||h_n - h||_K \to 0$. Por (a) y del hecho que K es compacto se tiene que $||h||_K < \infty$. Considerando el intervalo $[-M, M]$ donde $M = ||h||_K$ y por el Teorema de Weierstrass existe $p = p(\varepsilon) \in \mathcal{P}$ tal que $|p(y) - |y|| < \varepsilon$ para toda $y \in [-M, M]$. Como $\mathcal{A}$ es un álgebra, la función $g = p(h) \in \mathcal{A}$. Por definición de M se sigue que $|g(x) - |h(x)|| < \varepsilon \ \forall x \in K$ de donde $|h| \in \mathfrak{I}$.

De lo anterior se sigue que $f \vee g, f \wedge g \in \mathfrak{I} \ \forall f, g \in \mathfrak{I}$, por lo que $\mathfrak{I}$ es una celosía. Por (II.10) $\mathcal{A}$ distingue puntos en K y por otro lado $\mathcal{A} \subset \mathfrak{I}$, por lo tanto $\mathfrak{I}$ también distingue los puntos de K. Aplicando a $\mathfrak{I}$ el teorema (II.8), entonces existe $h \in \mathfrak{I}$ tal que $||h - F||_K < \frac{\varepsilon}{2}$ y por (b) existe $h_n \in \mathcal{A}$ tal que $||h - h_n||_K < \frac{\varepsilon}{2}$ y así se tiene que $||F - h_n||_K < \varepsilon$.

Q.E.D.

A continuación daremos algunas observaciones a la demostración anterior.

(II.12) Observaciones.

(i) Algunas versiones del teorema anterior sustituyen la condición de que $\mathcal{A}$ no se anula en K por el de que la función $e: K \to \mathbb{R}$, dada por $e(x) = 1$ para toda $x \in K$, pertenece a $\mathcal{A}$.

(ii) Es claro que $\Im \subset C_{p,1}(K)$ en (II.11), pero por lo expuesto al final de la demostración del teorema se concluye que $\Im = C_{p,1}(K)$, o dicho de otra manera $\overline{\mathcal{A}} = C_{p,1}(K)$.

(iii) Observemos que en la demostración de (II.11) utilizamos el Teorema de Weierstrass para aproximar uniformemente al valor absoluto por polinomios para garantizar que si $h \in \Im$ entonces $|h| \in \Im$, vamos a probar que si $f \in \mathcal{A}$ entonces $|f| \in \Im$. Si $||f||_K = 0$ entonces $f = 0$ y es claro que $|f| \in \Im$, por lo que supondremos que $||f||_K \neq 0$ entonces $\frac{f(x)}{||f||} \in [-1,1], \forall\, x \in K$, por otro lado sabemos que existe una sucesión (P_n) de polinomios tal que $P_n(t) \to |t|$ uniformemente en $[-1,1]$ (I.29), entonces $P_n(\frac{f(x)}{||f||}) \to \frac{|f(x)|}{||f||}$ uniformemente en K pero $P_n(\frac{f(x)}{||f||}) \in \mathcal{A}$ ya que $\mathcal{A}$ cumple con los dos primeros incisos de (II.11) por lo tanto $(\frac{|f(x)|}{||f||}) \in \Im$, así $|f| \in \mathcal{A}$. De donde es claro que si $h \in \Im$ entonces $|h|$ también.

(iv) Del inciso anterior y la demostración de (II.11) se concluye que el Teorema de Stone-Weierstrass y el Teorema de Weierstrass son equivalentes.

(v) Es fácil probar que $\Im \subset C_{p,1}(K)$ también es un álgebra de funciones.

Los siguientes incisos son sobre la necesidad de las hipótesis del Teorema anterior:

(vi) Si K no es compacto el teorema falla, por ejemplo $K = [0,\infty)$ y consideramos $f(x) = e^x$, así cómo $\mathcal{A} = \mathcal{P}$. (ver (I.7)).

(vii) Si $\mathcal{A}$ no contiene a la función $e(x) = 1$, sean $K = [0,1]$, $f \equiv 1$ y $\mathcal{A} = x\mathcal{P} =$

$= \{xp(x) \mid p \in \mathcal{P}\}$, es claro que $\mathcal{A}$ cumple con las condiciones (i) y (iii) del (II.11). pero F no es aproximable por elementos de $\mathcal{A}$ ya que:

$||1 - p||_{[0,1]} \geq |1 - p(0)| = 1$ para toda $p \in \mathcal{A}$.

(viii) Si $\mathcal{A}$ no separa puntos también falla el teorema (II.11), por ejemplo sean $K = [0.1]$, $f(x) = x$ y $\mathcal{A}$ el conjunto de todas las funciones constantes, es claro que $\mathcal{A}$ cumple con (i) y (ii) pero no (iii) de (II.11), pero si $a \in \mathcal{A}, a \neq 0$ entonces $||f - a||_{[0,1]} \geq |a|$ y si $a = 0$ entonces $||f - a||_{[0,1]} \geq |1 - a|$, de lo anterior se dedeuce que f no se aproxima por elementos de $\mathcal{A}$.

(ix) Si $\mathcal{A}$ no es un álgebra tenemos cómo contraejemplo el siguiente: $K = [0,1]$, $f(x) = x^2$ y $\mathcal{A} = \widehat{\mathcal{P}}_1$ =polinomios de grado 1, entonces $||x^2 - (ax+b)|| \geq |b| > 0$ si $b \neq 0$ y si $b = 0$ $||x^2 - (ax+b)|| = ||x^2 - ax|| \geq |1 - a| > 0$ si $a \neq 1$, si $a = 1$ entonces $||x^2 - ax|| = \frac{1}{4}$.

Veremos algunos corolarios del Teorema de Stone-Weierstrass:

Para lo cúal daremos la siguiente:

(II.13) Definición. *Un polinomio P de $\mathbb{R}^p$ a $\mathbb{R}^q$* es una función de $\mathbb{R}^p$ a $\mathbb{R}^q$ de la forma $P(\overline{x}) = \big(p_1(\overline{x}),\ldots,p_q(\overline{x})\big)$ donde cada $p_i(\overline{x}) \in \mathbb{R}[x_1,\ldots,x_p]$ para toda $i \in \{1,\ldots,q\}$.

(II.14) Corolario. (Teorema de Weierstrass en $\mathbb{R}^p$) *Sean $K \subset \mathbb{R}^p$ compacto y $f \in C_{p,q}(K)$. Entonces dada $\varepsilon > 0$ existe un polinomio P de $\mathbb{R}^p$ a $\mathbb{R}^q$ tal que $\|f - P\|_K < \varepsilon$.*

Demostración. Observemos que no podemos aplicar directamente (II.11) a la familia de polinomios de $\mathbb{R}^p$ a $\mathbb{R}^q$ ya que no es un álgebra por no estar definida una operación de multiplicación en $\mathbb{R}^p$, por lo que procedemos cómo sigue:

Dada $f \in C_{p,q}(K)$ escribimos $f = (f_1,\ldots,f_q)$ donde cada $f_i \in C_{p,1}(K)$. Sea $\mathcal{A} = \mathbb{R}[x_1,\ldots,x_p]$, es claro que $\mathcal{A}$ es un álgebra de funciones de $C_{p,1}(K)$. Por otro lado $e\colon K \to \mathbb{R}$ definida por $e(x_1,\ldots,x_p) = 1$ para todo $(x_1,\ldots,x_p) \in \mathbb{R}^p$ pertenece a $\mathcal{A}$, y si $\overline{x} \neq \overline{y}$, $\overline{x},\overline{y} \in K$ entonces $x_i \neq y_i$ para alguna $i \in \{1,\ldots,p\}$, consideremos $p_i(x_1,\ldots,x_p) = x_i, p_i \in \mathcal{A}$ y además $p_i(\overline{x}) \neq p_i(\overline{y})$ entonces por (II.11): dada $\varepsilon > 0$ existe $Q_i = Q_i(\varepsilon) \in \mathcal{A}$ tal que $\|f_i - Q_i\|_K < \frac{\varepsilon}{\sqrt{q}}$.

Sea $P = (Q_1,\ldots,Q_n)$. Es claro que P es un polinomio de $\mathbb{R}^p$ a $\mathbb{R}^q$ y cómo $|f(x)| \leq \sqrt{q}\max\{|f_i(\overline{x})| \mid i \in (1,\ldots,q\}\}$ entonces se tiene que $\|f - P\|_K < \varepsilon$.

Q.E.D.

(II.15) Corolario. *Sea $f \in C([0,\pi])$ entonces dada $\varepsilon > 0$ existe un polinomio P_{n_ε} en la variable $t = \cos x$ tal que $\|f - P_{n_\varepsilon}\|_{[0,\pi]} < \varepsilon$.*

Demostración. Sean $f \in C([0,\pi])$ y $\varepsilon > 0$, definimos:

$$\mathcal{A} = \{P_n(\cos x) \mid p_n \in \mathcal{P}, n \in \mathbb{N}\} \subset C([0,\pi])$$

es fácil ver que $\mathcal{A}$ es un álgebra de funciones, $e(x) = 1 = P_0(\cos x)$ con $P_0(x) = 1$, por lo que $e \in \mathcal{A}$ y si $x, y \in [0,\pi]$, $x \neq y$ sea $P_1(\cos x) = \cos x, P_1 \in \mathcal{A}$ y es claro que $P_1(\cos x) \neq P_1(\cos y)$, entonces por (II.11) existe $P_{n(\varepsilon)}(\cos x) \in \mathcal{A}$ tal que $\|f - P_{n(\varepsilon)}\|_{[0,\pi]} < \varepsilon$.

Q.E.D.

(II.16) Observación.

El corolario anterior no es cierto si elegimos $\mathcal{A} = \{P_n(\operatorname{sen} x) \mid p_n \in \mathcal{P}, n \in \mathbb{N}\}$ ya que $f(x) = x \in C([0,\pi])$ no se puede aproximar uniformemente con elementos en $\mathcal{A}$ pues $P_n(\operatorname{sen} 0) = P_n(\operatorname{sen}\pi) = a_0$. Si $a_0 \neq 0$ entonces $\|f - P_n(\operatorname{sen} x)\|_{[0,\pi]} \geq$
$\geq |f(0) - P_n(\operatorname{sen} 0)| = |a_0| > 0$ y en caso de que $a_0 = 0$ entonces $\|f - P_n(\operatorname{sen} x)\|_{[0,\pi]} \geq$

$\geq |f(\pi) - P_n(\text{sen}\pi)| = |\pi - 1| > 0.$

Ahora probaremos la versión compleja:

(II.17) Teorema de Stone-Weierstrass Complejo. *Sean $K \subset \mathbb{R}^p$ compacto, $F \in C_p^{\mathbb{C}}(K)$ y $\mathcal{A} \subset C_p^{\mathbb{C}}(K)$ tal que:*

(i) $\mathcal{A}$ es un álgebra de funciones complejas.

(ii) La función constante $e: K \subset \mathbb{R}^p \to \mathbb{C}$ dada por $e(x) = 1$ para toda $x \in K$, pertenece a $\mathcal{A}$.

(iii) Si $x, y \in K$ con $x \neq y$ entonces existe una función $f \in \mathcal{A}$ tal que $f(x) \neq f(y)$.

(iv) Si $f \in \mathcal{A}$ entonces $\overline{f} \in \mathcal{A}$ ($\overline{f}$ =la función conjugada compleja de f)

entonces dada $\varepsilon > 0$ existe $f = f(\varepsilon) \in \mathcal{A}$ tal que $||F - f||_K < \varepsilon$.

Demostración. Si $f \in \mathcal{A}$ entonces $Re(f) = \frac{1}{2}(f + \overline{f}) \in \mathcal{A}$ y $Im(f) = = \frac{1}{2i}(f - \overline{f}) \in \mathcal{A}$. Sea $\mathcal{A}' = \{f \in \mathcal{A} \mid f(K) \subset \mathbb{R}\} \subset C_{p,1}^{\mathbb{R}}(K)$, $\mathcal{A}' \neq \emptyset$ ya que $Re(f), Im(f) \in \mathcal{A}'$ y además $e(z) = 1, \forall\, z \in K$ pertenece a $\mathcal{A}'$ y es claro que $\mathcal{A}'$ es un álgebra real, así cómo que separa los puntos de K entonces por (II.11) dada $\varepsilon > 0$ existen $f_1, f_2 \in \mathcal{A}'$ tales que $||Re(F) - f_1||_K < \frac{\varepsilon}{2}$ y $||Im(F) - f_2||_K < \frac{\varepsilon}{2}$. Dado que $|w| \leq |Re(w)| + |Im(w)|$ entonces se tiene que: $f = f(\varepsilon) = f_1 + i f_2$ y $||F - f||_K \leq ||Re(F) - f_1||_K + ||Im(F) - f_2||_K < \varepsilon$.

Q.E.D.

(II.18) Observaciones.

(i) Para el teorema (II.17) si la condición (iv) no se cumple entonces el teorema falla. Consideramos a $K = \{z \in \mathbb{C} \mid |z| = 1\}$ (círculo unitario), $\mathcal{A}$ es el álgebra de las funciones $f(z) = \sum_{n=0}^{N} c_n z^n$, con $z = e^{i\theta}$ con $\theta \in [0, 2\pi)$. Es claro que $\mathcal{A}$ separa puntos en K y $f \equiv 1 \in \mathcal{A}$. Pero si $f \in \mathcal{A}$ entonces $\int_0^{2\pi} f\left(e^{i\theta}\right) e^{i\theta} d\theta = 0$, si consideramos $\widehat{f}(z) = \overline{z}$ entonces: $\int_0^{2\pi} \widehat{f}\left(e^{i\theta}\right) e^{i\theta} d\theta = 2\pi$ y si $\widehat{f}$ fuese el límite uniforme de funciones en $\mathcal{A}$ se tendría que:

$$\int_0^{2\pi} \widehat{f}\left(e^{i\theta}\right) e^{i\theta} d\theta = \lim_{n \to \infty} \int_0^{2\pi} f_n\left(e^{i\theta}\right) e^{i\theta} d\theta = 0$$

con $f_n \in \mathcal{A}$, lo cuál es imposible.

§2 Teorema de Extensión de Tietze y Teorema de Walsh

TEOREMA DE EXTENSION DE TIETZE

Ahora resolveremos un problema clásico de análisis, el Teorema de Extensión de Tietze en $\mathbb{R}^p$ usando el Teorema de Stone-Weierstrass, para lo cuál probaremos antes el siguiente resultado más débil.

(II.19) Lema. *Sean $K \subset \mathbb{R}^p$ compacto, no vacío, y $f \in C_{p,1}(K)$ entonces f puede ser extendida a una función h que es continua en $\mathbb{R}^p$ y $||h||_{\mathbb{R}^p} = ||f||_K$.*

Demostración. Sea $\varepsilon \in (0,1)$, por (II.11) existe $p_n(x)$ un polinomio de $\mathbb{R}^p$ a $\mathbb{R}$ tal que $||f - p_n||_K < \varepsilon^n, \forall n \in \mathbb{N}$ **(1)**.

Sea $q_1 = p_1$ y $q_n = p_n - p_{n-1}$ $\forall\ n \geq 2$, entonces $p_n = \sum_{j=1}^{n} q_j$, y por **(1)**, $||f - \sum_{j=1}^{n} q_j|| < \varepsilon^n$ **(2)**.

Por otro lado cómo $f \in C_{p,1}(K)$ entonces $||f||_K < \infty$ y se sigue de **(1)** que:

$|p_1(x)| < \varepsilon + ||f||_K\ \forall x \in K$ **(3)**

mientras que para toda $x \in K$ y $n \geq 2$ se tiene:

$|q_n(x)| = |p_n(x) - p_{n-1}(x)| \leq |f(x) - p_n(x)| + |f(x) - p_{n-1}(x)| < 2\varepsilon^{n-1}$ **(4)**.

Elegimos:

$h_1 = (-\varepsilon - ||f||_K) \vee ((\varepsilon + ||f||_K) \wedge q_1)$

y $h_n = (-2\varepsilon^{n-1}) \vee (q_n \wedge 2\varepsilon^{n-1})\ \forall n \geq 2$

entonces si $x \in K$ tenemos por **(2)** $h_1(x) = q_1(x)$, mientras que por **(4)** se tiene que: $h_n(x) = q_n(x)$ para toda $n \geq 2$, además es fácil ver que h_n es continua en $\mathbb{R}^p$ y $|h_n(x)| < 2\varepsilon^{n-1}\ \forall n \geq 2$ y $x \in \mathbb{R}^p$.

Por el criterio M de Weierstrass $\sum h_n$ converge uniformemente en $\mathbb{R}^p$. Sea $h = \sum_{n=0}^{\infty} h_n$, la cuál es continua por (I.8) y si $x \in K$ entonces por **(1)** y **(2)**

$h(x) = \lim_{k\to\infty} \sum_{n=0}^{k} h_n(x) = \lim_{m\to\infty} \sum_{n=1}^{m} q_n(x) = \lim_{m\to\infty} p_m(x) = f(x)$, además se tiene que:

$$|h(x)| \leq \sum_{n=1}^{\infty} |h_n(x)| = |h_1(x)| + \sum_{n=2}^{\infty} |h_n(x)| < \varepsilon + ||f||_K + 2\sum_{n=2}^{\infty} \varepsilon^{n-1}$$

de donde:

$|h(x)| \leq ||f||_K + \varepsilon + \frac{2\varepsilon}{1-\varepsilon}\ \forall x \in \mathbb{R}^p$ y así se sigue:

$||h||_{\mathbb{R}^p} \le ||f||_K$. Por lo tanto h es la función buscada.

Q.E.D.

(II.20) Teorema de Extensión de Tietze. *Sean $E \subset \mathbb{R}^p$ cerrado no vacío y $f \in C_{p,1}(E)$ entonces existe una función h continua en $\mathbb{R}^p$ que es una extensión de f y $||h||_{\mathbb{R}^p} = ||f||_E$.*

Demostración.

Caso 1. Si $E \subset \mathbb{R}^p$ cerrado y acotado entonces E es compacto y el resultado se sigue de (II.19).

Caso 2. Si $E \subset \mathbb{R}^p$ cerrado, no es acotado procedemos de la siguiente manera: Sea $m \in \mathbb{N}$ tal que $E_m = E \cap \overline{B}_m(0) \neq \emptyset$. Es claro que $E \cap \overline{B}_m(0)$ es compacto y definimos $f_m = f|_{E\cap\overline{B}_m(0)}$. Por (II.19) f_m tiene una extensión continua h_m en $\mathbb{R}^p$ tal que $||h_m||_{\mathbb{R}^p} = ||f_m||_{E_m}$, consideramos ahora a la función:

$$g_m(x) = \begin{cases} h_m(x) & \text{si } x \in \overline{B}_m(0) \\ f(x) & \text{si } x \in E_{m+1} \end{cases}$$

por construcción g_m es continua en $D_{g_m} = \overline{B}_m(0) \cup \left(E \cap \overline{B}_{m+1}(0)\right)$, así cómo $||g_m||_{D_{g_m}} = ||f||_{D_{g_m}}$ aplicando de nuevo (II.19) g_m tiene una extensión continua h_{m+1} en $\mathbb{R}^p$ tal que $||h_{m+1}||_{\mathbb{R}^p} = ||g_m||_{D_{g_m}}$, definimos:

$$g_{m+1}(x) = \begin{cases} h_{m+1}(x) & \text{si } x \in \overline{B}_{m+1}(0) \\ f(x) & \text{si } x \in E_{m+2} \end{cases}$$

continuando de este modo obtenemos una sucesión (g_k) de funciones continuas tales que:

(1) $D_{g_k} \subset D_{g_{k+1}}$ $\forall k \ge m$

(2) $||g_k|| = ||f||_{D_{g_k}}$ $\forall k \ge m$.

Sea $h: \mathbb{R}^p \to \mathbb{R}$ dada por $g(x) = g_k(x)$ donde $x \in \overline{B}_k(0)$, observemos que g está bien definida por (1) y la definición de cada g_k. Por otro lado g es continua en $\mathbb{R}^p$ y $g|_E = f$ de nuevo por la definición de cada g_k, $\forall k \ge m$ y por último $||g||_{\mathbb{R}^p} = ||f||_E$ ya que $||g_k||_E \le ||f||_E$ $\forall k \ge m$. Por lo tanto h es la extensión continua de f que se busca.

Q.E.D.

(I.21) Corolario. *Sean $E \subset \mathbb{R}^p$ cerrado y $f \in C_{p,q}(E)$, entonces existe una extensión continua $h: \mathbb{R}^p \to \mathbb{R}^q$ de f tal que $||h||_{\mathbb{R}^p} \le \sqrt{q}||f||_E$.*

Demostración.

Recordemos que $f = (f_1, \ldots, f_q)$ donde cada $f_j: E \subset \mathbb{R}^p \to \mathbb{R}$ es continua para cada $j \in \{1, \ldots, q\}$ entonces por (II.20) existe $h_j: \mathbb{R}^p \to \mathbb{R}$ extensión continua de f_j para cada $j \in \{1, \ldots, q\}$ y $\|h_j\|_{\mathbb{R}^p} = \|f_j\|_E$. Sea $h = (h_1, \ldots, h_q)$ entonces es claro que $h \in C_{p,q}(\mathbb{R}^p)$ y además $h|_E = f$ y $\|h\|_{\mathbb{R}^p} \leq \sqrt{q}\|f_j\|_E \leq \sqrt{q}\|f\|_E$.

Q.E.D.

TEOREMA DE WALSH

Para concluir con las aplicaciones del Teorema de Weierstrass resolveremos el siguiente problema:

Sean $f \in C([a,b]), x_1, \ldots, x_n \in [a,b]$, n puntos y $\varepsilon > 0$. ¿Existirá un polinomio $p(x)$ tal que $\|f - p\|_{[a,b]} < \varepsilon$ y $p(x_i) = f(x_i)$ $\forall i = 1, \ldots, n$?. Para contestar está pregunta probaremos el Teorema de Walsh, pero antes de enunciarlo definiremos importantes conceptos:

Dados $x_0, \ldots, x_n \in [a,b]$ $n+1$ puntos distintos definimos:

$$\ell_k(x) = \frac{\prod_{i \neq k}(x - x_i)}{\prod_{i \neq k}(x_k - x_i)}$$

observemos que $\ell_k \in \mathcal{P}_n$ y $\ell_k(x_j) = \delta_{kj}$ (delta de Kronecker).

Entonces dados los valores $y_0, \ldots, y_n$ el polinomio definido por:

$p_n(x) = \sum_{k=0}^{n} y_k \ell_k(x) \in \mathcal{P}_n$ y $p_n(x_k) = y_k$ para toda $k = 0, \ldots, n$. A esta expresión se le conoce como la ***Fórmula de interpolación de Lagrange***.

Otra forma alternativa muy útil es la siguiente:

Sea $w(x) = \prod_{i=0}^{n}(x - x_i)$ entonces $w'(x_k) = \prod_{i \neq k}(x_k - x_i)$ entonces:

$$\ell_k(x) = \frac{w(x)}{(x - x_k)w'(x_k)}$$

así tenemos que: $p_n(x) = \sum_{k=0}^{n} y_k \frac{w(x)}{(x - x_k)w'(x_k)}$ observemos que $p_n \in \mathcal{P}_n$ es el único polinomio en $\mathcal{P}_n$ tal que $p_n(x_k) = y_k$ para toda $k = 0, \ldots, n$ ya que si $q \in \mathcal{P}_n$ es tal que $q_n(x_k) = y_k = p_n(x_k)$ entonces $r_n = p_n - q_n \in \mathcal{P}_n$ y se anula en al menos $n+1$ puntos por lo que $p_n = q_n$.

(II.22) Teorema de Walsh. *Sean $f \in C([a,b])$ y $x_0, \ldots, x_n \in [a,b]$ $n+1$ puntos distintos y $y_k = f(x_k)$ $\forall k = 0, \ldots, n$ dados entonces dada $\varepsilon > 0$ existe un polinomio $p = p(\varepsilon)$ tal que $\|f - p\|_{[a,b]} < (1 + M)\varepsilon$ con $f(x_k) = p(x_k)$ y M una constante que depende sólo de $[a,b]$ y $x_0, \ldots, x_n$.*

Demostración.

Sea $\varepsilon > 0$ por (I.1) existe $Q = Q(\varepsilon) \in \mathcal{P}$ tal que $||f - Q||_{[a,b]} < \varepsilon$.

Consideremos $q(x) = \sum_{k=0}^{n} (f(x_k) - Q(x_k))\ell_k(x)$, como $q(x_k) = f(x_k) - Q(x_k)$ para toda $k \in \{0, \ldots, n\}$ **(1)**.

Además tenemos que $||q||_{[a,b]} \leq \sum_{k=0}^{n} |f(x_k) - Q(x_k)| ||\ell_k||_{[a,b]}$ entonces $||q||_{[a,b]} < M\varepsilon$ donde $M = \sum_{k=0}^{n} ||\ell_k||[a,b]$ **(2)** y donde M sólo depende de $[a,b]$ y $x_0, \ldots, x_n$.

Sea $p(x) = Q(x) + q(x)$ por **(1)** $p(x_k) = f(x_k)$ para toda $k \in \{0, \ldots, n\}$, mientras que por **(2)** y la elección de Q tenemos: $||f - p||_{[a,b]} \leq ||f - Q||_{[a,b]} + ||q||_{[a,b]} < (M + 1)\varepsilon$.

Q.E.D.

(II.23) Observación. Revisando la demostración de (II.22) se observa que ésta sigue siendo válida para toda función $f: [a,b] \to \mathbb{R}$ que se pueda aproximar uniformemente por polinomios en $[a,b]$.

Usando la densidad de $\mathbb{Q}$ en $\mathbb{R}$ es fácil comprobar que toda función $f \in C([a,b])$ se puede aproximar uniformemente con polinomios con coeficientes racionales. Así tenemos el siguiente:

(II.24) Corolario. *Sea $f \in C([a,b])$ entonces para toda $\varepsilon > 0$ existe $q = q(\varepsilon) \in \mathbb{Q}[x]$ tal que $||f - q||_{[a,b]} < \varepsilon$.*

(II.25) Observación.

En la conclusión del corolario anterior no se puede sustituir $\mathbb{Q}[x]$ por $\mathbb{Z}[x]$ como se muestra en el siguiente ejemplo: sea $[a,b] = [0,1]$ y $f(x) = r$ donde $r \notin \mathbb{Q}$ y $q_n \in \mathbb{Z}[x]$ entonces:

$||r - q_n||_{[0,1]} \geq |r - q_n(0)| = |r - a| \geq \max\{|[r] - r|, |[r] + 1 - r|\} > 0$ para toda $a \in \mathbb{Z}$ por lo tanto no toda función continua se puede aproximar uniformemente con polinomios con coeficientes enteros.

CAPITULO III

MEJOR APROXIMADOR POLINOMIAL

En este capítulo nos dedicaremos a responder algunas preguntas sobre el mejor aproximador polinomial como son su: existencia, unicidad y caracterización.

§1 Existencia

Partiremos del hecho de que conocemos lo que significa que un conjunto $X \neq \emptyset$ sea un espacio vectorial sobre un campo K.

(III.1) Definición. Sea X un espacio vectorial sobre $\mathbb{R}$. Entonces una ***norma en*** X es una función $\|\cdot\|$, de X a $\mathbb{R}$, que satisface las siguientes propiedades:

(i) $\|x\| \geq 0$, $\|x\| = 0$ si y sólo si $x = 0$ $\forall x \in X$.

(ii) $\|\alpha x\| = |\alpha|\|x\|$ $\forall \alpha \in \mathbb{R}$ (homógeneidad).

(iii) $\|x + y\| \leq \|x\| + \|y\|$ $\forall x, y \in X$ (desigualdad del triángulo).

A la pareja $(X, \|\cdot\|)$ se le conoce como un ***espacio vectorial normado***, (e.v.n). Si la función $\|\cdot\|$ satisface que $\|0\| = 0$, pero de $\|x\| = 0$ no se sigue que $x = 0$, entonces $\|\cdot\|$ es una ***seminorma***.

Sean $(X, \|\cdot\|)$ un e.v.n. y $y \in X$ fijo junto con $x_1, \ldots, x_n$ n vectores linealmente independientes en X. Nos planteamos el siguiente problema: Dada $y \in X$ y $W = \langle x_1, \ldots, x_n \rangle \subset X$ un subespacio de dimensión finita, ¿Existe $w^* \in W$ tal que $\|y - w^*\| \leq \|y - w\|$ $\forall w \in W$?. Para resolver esta pregunta empezaremos con la siguiente:

(III.2) Definición. Sea $(X, \|\cdot\|)$ un e.v.n., $W \subset X$ de dimensión finita y $x \in X$, diremos que $w^* \in W$ es ***un mejor aproximador*** a x:

si $\|x - w^*\| \leq \|x - w\|$ $\forall w \in W$.

A $x - w^*$ se le llama el ***error ó discrepancia***. Observemos que un mejor aproximador resuelve el problema de minimizar la norma del error.

(III.3) Ejemplos.

(i) Sean $X = C([a,b])$, $||f|| = ||f||_{[a,b]}$, $W =< 1 >$=subespacio vectorial de las constantes y $x = x(t) \in X$. Entonces deseamos encontrar $w^* \in W$ tal que $||x - w^*||_{[a,b]}$ sea mínima. Definimos $M = \max\{x(t) \mid t \in [a,b]\}$ y

$m = \min\{x(t) \mid t \in [a,b]\}$, dado que $x(t) \in X$ existen $t_1, t_2 \in [a,b]$ tales que $x(t_1) = M$ y $x(t_2) = m$. Sea $c \in \mathbb{R}$ tal que:

(a) $c \geq M$ entonces $||x - c||_{[a,b]} = c - m \geq M - c$

(b) $c \leq m$ entonces $||x - c||_{[a,b]} = M - c \geq c - m$

(c) $c \in [m, M]$ entonces $||x - c||_{[a,b]} = \max\{c - m, M - c\}$

de todos los casos anteriores el valor mínimo de $||x - c||_{[a,b]}$ se alcanza cuando minimizamos a $\max\{c - m, M - c\}$ y esto se da cuando $c - m = M - c$ de donde $c = \frac{M+m}{2}$, es decir $w^* = \frac{M+m}{2}$ y por el inciso (c) $||x - w^*||_{[a,b]} = \frac{M-m}{2}$.

(ii) **Sean** $X = C([a,b])$, $W =< 1 >$, $x = x(t) \in X$ y $||x||_2 = (\int_0^1 x^2(t)dt)^{\frac{1}{2}}$

hacemos $< x, x >= \int_0^1 x^2(t)dt$ entonces $||x||_2^2 =< x, x >$, **por lo que buscamos minimizar a:** $||x - c||_2^2 = c^2 - 2c < x, 1 > + < x, x >= (c - < x, 1 >)^2 +$
$+ < x, x > - < x, 1 >^2$, **de donde el valor mínimo de** $||x - c||_2^2$ es $< x, x > -$
$- < x, 1 >^2$ **cuando** $c =< x, 1 >$, **por lo que** $w^* =< x, 1 >$ y $||x - w^*||_{[a,b]} =$

$$= \left\{ \int_a^b x^2(t)dt - \left[\int_a^b x(t)dt \right]^2 \right\}.$$

(iii) **Sean** $X = \mathbb{R}^3$, $||x|| = \max\{|x_i| \mid i = 1,2,3\}$, $W =< e_1, e_2 >$ donde:
$e_1 = (1,0,0), e_2 = (0,1,0), x = (x_0, y_0, z_0)$. **Es fácil ver que el error mínimo es** $|z_0|$ **y es alcanzado por** $w^* = (w_1, w_2) \in W$ si $|w_1 - x_0| \leq |z_0|$ y $|w_2 - y_0| \leq |z_0|$.

(iv) **Sea** $X = \widehat{\mathcal{P}}_1$, $||x|| = |x(0)| + |x(1)|$. **¿Cuál es la mejor constante que aproxima a la función identidad? En nuestro ejemplo tenemos a** W **el mismo que el del inciso (i), entonces hay que resolver** $\min\{||x - w|| \mid w \in W\}$ **pero con** $||x - w|| = |w| + |1 - w|$, **por lo que:**

(a) si $w > 1$ **entonces** $||x - w|| = 2w - 1$

(b) si $w < 0$ **entonces** $||x - w|| = 1 - 2w$

(c) si $w \in [0,1]$ **entonces** $||x - w|| = 1$

De lo anterior se concluye que $w^* \in [0,1]$ **y** $||x - w^*|| = 1$.

Observemos por último en los ejemplos (iii) y (iv) el mejor aproximador no es único a diferencia de los otros ejemplos. Ahora probaremos el siguiente resultado fundamental que garantiza la existencia de mejores aproximadores.

(III.4) Teorema. *Sean $(X, ||\cdot||)$ un e.v.n. sobre un campo K, $W \subset X$ un subespacio vectorial de dimensión finita y $x \in X$ entonces existe $w^* \in W$ tal que* $\min_{w\in W} ||x - w|| = ||x - w^*||$.

Demostración. Supongamos que W es dimensión n y sea $\{w_1, \ldots, w_n\} \subset \subset W$ una base para W. Consideremos la siguiente función $d: K^n \to \mathbb{R}$ definida como $d(a_1, \ldots, a_n) = ||x - \sum_{i=1}^{n} a_i w_i||$ entonces d es uniformemente continua ya que $|d(a_1, \ldots, a_n) - d(b_1, \ldots, b_n)| \leq ||\sum_{i=1}^{n}(a_i - b_i)w_i|| \leq M \sum_{i=1}^{n} |a_i - b_i|$ con $M = \max\{||w_i|| \mid i = 1, \ldots, n\}$, para el caso particular en que $x = 0$ entonces la función $h(a_1, \ldots, a_n) = ||\sum_{i=1}^{n} a_i w_i||$ es continua **(1)**.

Consideremos $S = \{\overline{a} = (a_1, \ldots, a_n) \in K^n \mid \sum_{i=1}^{n} |a_i|^2 = 1\}$ es claro que S es compacto en K^n entonces por **(1)**, $h|_S$ alcanza su mínimo m en S, y dado que $\{w_1, \ldots, w_n\}$ es linealmente independiente entonces $m > 0$.

Por otro lado sea $r = \left(\sum_{i=1}^{n} |a_i|^2\right)^{\frac{1}{2}} \neq 0$ entonces $h(a_1, \ldots, a_n) = r||\sum_{i=1}^{n} \frac{a_i}{r} x_i||$ pero $\sum_{i=1}^{n} \left|\frac{a_i}{r}\right|^2 = 1$, por lo que $h(a_1, \ldots, a_n) \geq rm$. De donde tenemos la siguiente desigualdad $d(a_1, \ldots, a_n) \geq ||\sum_{i=1}^{n} a_i w_i|| - ||x|| \geq rm - ||x||$ **(2)**. Recordando la definición de r, se observa que d crece conforme r lo hace también. Lo anterior da la pauta para buscar el valor mínimo de d en una bola de K^n, por lo que consideremos $\rho = \inf\{d(a_1, \ldots, a_n)\}$, $(a_1, \ldots, a_n) \in K^n$, sea $R = \frac{1+\rho+||x||}{m}$, si $\sum_{i=1}^{n} |a_i|^2 > R^2$ entonces de **(2)** se tiene que:

$d(a_1, \ldots, a_n) > Rm - ||x|| = 1 + \rho > \rho$, de lo cuál se sigue que $\inf\{d(a_1, \ldots, a_n) \mid (a_1, \ldots, a_n) \in K^n\} = \inf\{d(a_1, \ldots, a_n) \mid (a_1, \ldots, a_n) \in \overline{B}_R(0)\}$. Pero como d es continua y $\overline{B}_R(0)$ es compacto entonces d alcanza su mínimo en $\overline{B}_R(0)$ por lo que d alcanza su mínimo en K^n.

Q.E.D.

(III.5) Observaciones.

(i) El requisito de que W sea un subespacio vectorial de dimensión finita es necesario, para lo cuál consideremos lo siguiente: sean $X = C([0, \frac{1}{2}])$, $W = \mathcal{P}$ y $f(x) = \frac{1}{1-x}$, por la expansión en series de Taylor de f tenemos que:

$f(x) = \sum_{i=1}^{\infty} x^i \ \forall\ x \in [0, \frac{1}{2}]$, así dada $\varepsilon > 0$ existe $N(\varepsilon) \in \mathbb{N}$ tal que $|f(x) - (1 + x + \ldots + x^N)| < \frac{1}{2^N} < \varepsilon$ si N es suficientemente grande para toda $x \in [0, \frac{1}{2}]$. Ahora si existiera $w^* \in W$ un mejor aproximador a f entonces $||f - w^*|| = \frac{1}{2^N} < \varepsilon$ de donde se tiene que $f = w^*$ lo cuál no es posible.

(ii) El teorema anterior permanece válido si $||\cdot||$ es sólo una seminorma.

(III.6) Corolario. *Sea $X = C([a,b])$ con la norma uniforme y $W = \mathcal{P}_n$ para alguna $n \in \mathbb{N}$ fija, entonces $\min\limits_{p \in \mathcal{P}_n} \{||f - p||_{[a,b]}\}$ tiene solución para toda $f \in X$.*

(III.7) Corolario. *Sea $X = C([a,b])$ con la norma p, $||f||_p = \left(\int_a^b |f|^p\right)^{\frac{1}{p}}$ y $W = \mathcal{P}_n$ para alguna $n \in \mathbb{N}$ fija, entonces $\min\limits_{p \in \mathcal{P}_n} \{||f - p||_p\}$ tiene solución para toda $f \in X$.*

(III.8) Corolario. *Sea $X = PC([-\pi, \pi])$ con la norma uniforme y $W = T_n$ para alguna $n \in \mathbb{N}$ fija, entonces existe $T_n(x) = \sum_{k=0}^{n}(a_k \cos kx + b_k \operatorname{sen} kx)$ tal que $||f - p||_{[-\pi,\pi]}$ es mínimo para toda $f \in X$.*

Para los siguientes corolarios necesitamos considerar la segunda observación de (III.5).

(III.9) Corolario. *Sean $x_0, \ldots, x_k$ $k+1$ números distintos con $k \geq n$ entonces el problema $\min\limits_{a_0,\ldots,a_k} \max\{|f(x_i) - (a_0 + \ldots + a_n x_i^n)|$ tiene solución, con f una función.*

Aquí consideramos a $X = \{(f(x_0), \ldots, f(x_k)) \mid f \text{ es función}\}$, $W =$ subespacio de X de dimensión $(n+1)$ consistente de todos los vectores $p = (p(x_0), \ldots, p(x_k))$ donde $p \in \mathcal{P}_n$ y $||f|| = \max\{|f(x_i)| \mid i = 0, \ldots, k\}$.

(III.10) Corolario. *Sean $x_0, \ldots, x_k$ $k+1$ números distintos con $k \geq n$ entonces el problema $\min\limits_{a_0,\ldots,a_k} \left(\sum_{i=0}^{k} |f(x_i) - (a_0 + \ldots + a_n x_i^n|^p\right)^{\frac{1}{p}}$ $p \geq 1$ tiene solución.*

Aquí X y W son iguales a los de (III.9) y $||f|| = \left[\sum_{i=0}^{k} |f(x_i)|^p\right]^{\frac{1}{p}}$

(III.11) Corolario. *Sean a_{ij}, y_i números reales dados para $1 \leq i \leq p$,*

$1 \leq j \leq n$ *y* $p > n$ *entonces* $\min\limits_{x_j} \max\{|y_i - (a_{i1}x_1 + \ldots + a_{in}x_n)| \mid 1 \leq i \leq p\}$ *tiene solución.*

Aquí el problema se puede ver como: $\min\limits_{\overline{x} \in \mathbb{R}^n} \|\overline{y} - A\overline{x}\|_\infty$ donde $\overline{y} = (y_1, \ldots, y_p)$, $\overline{x} = (x_1, \ldots, x_n)$ y $A \in \mathcal{M}_{p \times n}(\mathbb{R})$ y consideramos a $X = \mathbb{R}^p$, $W = A(\mathbb{R}^n) \subset \mathbb{R}^p$ y $\|x\| = \|x\|_\infty$.

(III.12) Definición. *Sean* $(X, \|\cdot\|)$ *un e.v.n.,* $x \in X$ *y* $W \subset X$ *subespacio de dimensión finita* n*. Definimos* $E_n(x) = \|x - w^*\|$ *donde* w^* *es un mejor aproximador de* x *en* W*.*

(III.13) Observaciones.

(i) **A partir de (III.12) se tiene que** $\{E_n(x)\}_{n \in \mathbb{N}}$ **es una sucesión no creciente de números reales no negativos, cuando los subespacios vectoriales forman una sucesión anidada creciente.**

(ii) **Por (I.1)** $\lim\limits_{n \to \infty} E_n(f) = 0 \;\forall f \in C([a, b])$.

(iii) **Dados** $x \in X$, $w^{**} \in W$ **(subespacio de dimensión finita), donde** w^{**} **es un mejor aproximador de** $x - w$ **de** W**, sabemos que** $E_n(x) = \|x - w^*\|$ **con** w^* **es un mejor aproximador de** x **en** W**, entonces** $w^* + w^{**} \in W$ **por lo que** $\|x - w^*\| \leq \|x - (w^* + w^{**})\|$ **así se sigue que** $E_n(x) \leq E_n(x - w)$. **Por otro lado** $E_n(x) = \|x - w^*\| = \|x - w - (w^* - w)\| \geq \|x - w - w^{**}\| = E_n(x - w)$ **por lo que** $E_n(x) = E_n(x - w) \;\forall w \in W$.

Hemos probado que si w^* **es un mejor aproximador a** x **de** W **entonces** $w^* - w$ **es un mejor aproximador a** $x - w$ **de** W**.**

§2 Unicidad

En la sección anterior encontramos condiciones suficientes bajo las cuáles garantizamos la existencia de un mejor aproximador, ahora examinaremos el problema de la unicidad.

(III.14) Definición. Si $(X, ||\cdot||)$ es un e.v.n., $x \in X$ y $W \subset X$ (un subespacio de X), denotaremos por $W^* = \{w^* \in W \mid w^*$es un mejor aproximador de x en $W\}$.

Por (III.4) sabemos que $W^* \neq \emptyset$ si W es de dimensión finita y si W es de dimensión infinita puede suceder que $W^* = \emptyset$.

(III.15) Definición. Sean $(X, ||\cdot||)$ un e.v.n. y $E \subset X$. Decimos que E *es convexo* si $tx + (1-t)y \in E \; \forall t \in (0,1)$ y $x, y \in E$.

(III.16) Teorema. *Sean $(X, ||\cdot||)$ un e.v.n, $x \in X$ y $W \subset X$ un subespacio de X entonces W^* es convexo y acotado. Además si W es cerrado entonces W^* lo es también.*

Demostración. Si $W^* = \emptyset$ no hay nada que hacer. Si $W^* \neq \emptyset$, como $0 \in W$ entonces $||x - w^*|| \leq ||x||$ para toda $w^* \in W^*$ de donde $||w^*|| \leq 2||x|| \; \forall \, w^* \in W^*$ por lo que W^* es acotado.

Sean $w_1^*, w_2^* \in W^*$ entonces $||x - w_1^*|| = ||x - w_2^*|| = r$. Para $t \in (0,1)$ tenemos: $||x-(tw_1^*+(1-t)w_2^*)|| = ||t(x-w_1^*)+(1-t)(x-w_2^*|| \leq t||x-w_1^*||+(1-t)||x-w_2^*|| = r$, pero $tw_1^* + (1-t)w_2^* \in W$ por lo que W^* es convexo.

Ahora supongamos que W es cerrado. Sea $w \in \overline{W^*}$ entonces existe una sucesión (w_n^*) tal que $w_n^* \to w$ cuando $n \to \infty$, $w_n^* \in W^*$ para toda $n \in \mathbb{N}$ pero $w_n^* \in W$ y como W es cerrado entonces $w \in W$. Dada $\varepsilon > 0$ existe $N(\varepsilon) \in \mathbb{N}$ tal que $||w - w_N^*|| < \varepsilon$ de donde tenemos que:

$r \leq ||x - w|| \leq ||x - w_N^*|| + ||w_N^* - w|| < r + \varepsilon$ para toda $\varepsilon > 0$, por lo tanto $||x - w|| = r$ y así tenemos que $w \in W^*$.

Q.E.D.

(III.17) Definición. Sea $(X, ||\cdot||)$ un e.v.n. Decimos que X es *estrictamente convexo* si siempre que $x \neq y$, $||x|| \leq r$, $||y|| \leq r$ entonces $||x + y|| < 2r$. En tal caso, también decimos que $||\cdot||$ es *estrictamente convexa*.

(III.18) Ejemplos.

(i) Sea $(X, <,>)$ un espacio vectorial con producto interior entonces el e.v.n. asociado $(X, ||\cdot||)$ es estrictamente convexo ya que por la desigualdad del triángulo si $x \neq y$ y $||x|| \leq r$ y $||y|| \leq r$ tenemos que $||x+y|| \leq ||x|| + ||y|| \leq 2r$ de donde

$||x+y|| < 2r$, ya que si se cumple que $||x+y|| = 2r$ entonces $||x||+||y|| = ||x+y||$ pero como la norma proviene de un producto interior necesariamente $x = y$, lo cuál es una contradicción.

(ii) La norma uniforme en $X = C([-1,1])$ no es estrictamente convexa para lo cuál consideramos a $f(x) = 1 - x^2$, $g(x) = 1 - x^4$, de donde $||f|| = ||g|| = 1$ y $||f + g|| = 2$ pero $f \neq g$.

(iii) Si $X = \mathbb{R}^2$ con $||(x,y)|| = ||\overline{x}||_\infty$ no es estrictamente convexo ya que

$||(1,0) + (1,1)||_\infty = ||(1,0)||_\infty + ||(1,1)||_\infty$.

(iv) Los espacios lineales normados ℓ^p y $\mathcal{L}^p([a,b])$, $p \in (1,\infty)$ son estrictamente convexos con la norma $||\ ||_p$, donde:

$$\ell^p = \{x_n : \mathbb{N} \to \mathbb{R} \mid \sum_{n=1}^{\infty} |x_n|^p < +\infty\}$$

$$\mathcal{L}^p([a,b]) = \{f : [a,b] \to \mathbb{R} \mid \int_a^b |f(t)|^p dt < +\infty\}$$

y $||x||_p = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}}$ ó $\left(\int_a^b |f(t)|^p dt\right)^{\frac{1}{p}}$ respectivamente.

Para probar que $||\ ||_p$ es estrictamente convexa, veremos la siguiente versión de la desigualdad de Hölder:

(III.19) Lema. *Sean $a, b > 0$ y $t \in [0,1]$ entonces $a^t b^{1-t} \leq ta + (1-t)b$. La igualdad se cumple si $t = 0, t = 1$ ó $a = b$.*

Demostración. Consideremos la función $f(x) = -\ln(x)$ entonces $f''(x) > 0$ en $(0,\infty)$ por lo que f es convexa en dicho intervalo.

De donde $f(ta+(1-t)b) \leq tf(a)+(1-t)f(b)$ $\forall t \in [0,1], a, b > 0$ y con igualdad cuando $t = 0, t = 1$ ó $a = b$, por lo tanto $-\ln(ta + (1-t)b) \leq -t\ln(a) - (1-t)\ln(b)$, aplicando la exponencial tenemos la desigualdad buscada.

Q.E.D.

Ahora veremos el siguiente:

(III.20) Teorema. *Sean $f_1, f_2 \in \mathcal{L}^p$ con $p \in (1,\infty)$, $f_1 \neq f_2$ tales que $||f_1||_p = ||f_2||_p = r$ entonces $||tf_1 + (1-t)f_2||_p < r$, $t \in (0,1)$.*

Demostración. Por la desigualdad del triángulo sabemos que:

$||tf_1 + (1-t)f_2||_p \leq r$. Supongamos que $||tf_1 + (1-t)f_2||_p = r$ entonces tenemos que:

$|tf_1 + (1-t)f_2|^p = |tf_1 + (1-t)f_2||tf_1 + (1-t)f_2|^{p-1} \leq t|f_1||tf_1 + +(1-t)f_2|^{p-1} +$

$+(1-t)|f_2||tf_1+(1-t)f_2|^{p-1}$ y la igualdad es posible cuando $f_1 f_2 > 0$ en $[a,b]$ **(1)**. Haciendo $a = |f_1|^p$, $b = |tf_1+(1-t)f_2|^p$ y $t = \frac{1}{p}$ por el lema (III.19) se sigue:

$$|f_1||tf_1+(1-t)f_2|^{p-1} \leq \frac{|f_1|^p}{p} + (1-\frac{1}{p})|tf_1+(1-t)f_2|^p \quad (2)$$

y la igualdad se da si $|f_1| = |tf_1+(1-t)f_2|$. Análogamente:

$$|f_2||tf_1+(1-t)f_2|^{p-1} \leq \frac{|f_2|^p}{p} + (1-\frac{1}{p})|tf_1+(1-t)f_2|^p \quad \mathbf{(3)}$$

y la igualdad se da si $|f_2| = |tf_1+(1-t)f_2|$. Por lo que de **(1)**, **(2)** y **(3)** se tiene que:

$$r^p = \int_a^b |tf_1+(1-t)f_2|^p dt \leq \frac{t}{p}\int_a^b |f_1|^p dt + \frac{(1-t)}{p}\int_a^b |f_2|^p dt +$$

$+(1-\frac{1}{p})\int_a^b |tf_1+(1-t)f_2|^p = r^p$, por lo que las igualdades se dan en **(1)**, **(2)** y **(3)**, así se tiene que $|f_1| = |f_2|$ y $f_1 f_2 > 0$ entonces $f_1 = f_2$ lo cuál no es posible.

Q.E.D.

(III.21) Observación. El teorema anterior falla si $p = 1$. Para lo cuál consideramos por ejemplo a $f_1(x) = \frac{3}{2}x^2$ y $f_2(x) = \frac{3}{4}(1-x^2)$.

A continuación daremos algunas definiciones alternativas al concepto de norma estrictamente convexa que junto con el teorema (III.20) prueban que ℓ^p y $\mathcal{L}^p$ son espacios estrictamente convexos $(1 < p < \infty)$.

(III.22) Lema. *Sea* $(X, \|\cdot\|)$ *un e.v.n. entonces las siguientes condiciones son equivalentes:*

(i) $(X, \|\cdot\|)$ *es estrictamente convexo*

(ii) Si $\|x\| = \|y\| = r$, $x \neq y$ *entonces:* $\|tx+(1-t)y\| < r$ *para toda* $t \in (0,1)$

(iii) Si $\|x\| = \|y\| = 1$, $x \neq y$ *entonces:* $\|tx+(1-t)y\| < 1$ *para toda* $t \in (0,1)$.

Demostración.

(i)⇒(ii) Sean $\|x\| = r = \|y\|, x \neq y$, y $t \in (0, \frac{1}{2}]$ entonces $\|tx+(1-t)y\| \leq t\|x+y\| + |1-2t|\|y\|$ por hipótesis tenemos $\|tx+(1-t)y\| < < 2tr + (1-2t)r = r$.

Si $t \in [\frac{1}{2}, 1)$ entonces $\|tx+(1-t)y\| \leq (1-t)\|x+y\| + |2t-1|\|x\|$ y por hipótesis tenemos $\|tx+(1-t)y\| < 2r(1-t) + (2t-1)r = r$.

(ii)⇒(iii) es claro, haciendo $r = 1$

(iii)⇒(i) Sean $x, y \in X, x \neq y$ tal que $\|x\| \leq r$ y $\|y\| \leq r$. Si $\|x\| < r$ ó $\|y\| < r$, de la desigualdad del triángulo se sigue que $\|x+y\| < 2r$. Por lo que supondremos que $\|x\| = r = \|y\|$, como $x \neq y$ entonces $r \neq 0$ y consideremos $\widehat{x} = \frac{x}{\|x\|}$, $\widehat{y} = \frac{y}{\|y\|}$. Por hipótesis sabemos que $\|t\widehat{x} + (1-t)\widehat{y}\| < 1$ para toda $t \in (0,1)$, en particular si $t = \frac{1}{2}$ tenemos $\|x+y\| < 2r$.

Q.E.D.

(III.23) Teorema. *Sean $(X, \|\cdot\|)$ un e.v.n. con norma estrictamente convexa, $x \in X$ y $W \subset X$ un subespacio de dimensión finita, entonces existe un único $w^* \in W$ tal que $\|x - w^*\|$ es mínima.*

Demostración. Supongamos que existen $w_1^*, w_2^* \in W^*$, $w_1^* \neq w_2^*$ entonces $E_n(x) = \|x - w_1^*\| = \|x - w_2^*\|$ y además $x - w_1^* \neq x - w_2^*$ por hipótesis

$\|x - w_1^* + x - w_2^*\| < 2E_n(x)$ de donde $\|x - \frac{1}{2}(w_1^* + w_2^*)\| < E_n(x)$ pero $\frac{1}{2}(w_1^* + w_2^*) \in$ $\in W^*$ por (III.16), lo cual no es posible.

Q.E.D.

(III.24) Corolario. *Sea $X = \mathcal{L}^p([a,b])$, con la norma $p \in (1, \infty)$ y $W = \mathcal{P}_n$ para alguna $n \in \mathbb{N}$ fija, entonces el mejor aproximador $w^* \in W$ es único para toda $f \in X$.*

§3 Caracterización

En está sección buscaremos una caracterización del mejor aproximador polinomial $p_n \in \mathcal{P}_n$ para una función $f \in C([a,b])$.

(III.25) Definición. Sean $f \in C([a,b])$ y $p_n^* \in \mathcal{P}_n$ un mejor aproximador de f entonces definimos $e_n(x) = f(x) - p_n^*(x)$ **(La función error)**.

Por la definición anterior y (III.12) tenemos que $E_n(f) = \|e_n\|_{[a,b]}$.

Dada $f \in C([a,b])$ de (III.3) (i), se tiene que la constante $c^* = \frac{M+m}{2}$ donde $M = \max\{f(x) \,:\, x \in [a,b]\}$ y $m = \min\{f(x) \,:\, x \in [a,b]\}$ hace que $E_0(f) =$
$= \|f - c^*\|_{[a,b]} = \frac{M-m}{2}$. Por la continuidad de f existen al menos dos puntos distintos $x_1, x_2 \in [a,b]$ en los que $|e_0(x_1)| = |e_0(x_2)| = E_0(f)$ y $e_0(x_1) = -e_0(x_2)$, está propiedad la cumplen en general los mejores aproximadores como lo muestra el siguiente:

(III.26) Lema. *Dada $f \in C([a,b])$ y $p_n^* \in \mathcal{P}_n$ un mejor aproximador a f entonces existen $x_1, x_2 \in [a,b]$ puntos distintos tales que:*

(i) $E_n(f) = |e_n(x_i)|, i = 1,2$

(ii) $e_n(x_1) = -e_n(x_2)$.

Demostración. Por definición $e_n \in C([a,b])$ y $|e_n(x)| \leq E_n(f)$ para toda $x \in [a,b]$, por lo que existe $x_1 \in [a,b]$ tal que $|e_n(x_1)| = E_n(f)$. Sin pérdida de generalidad supongamos que $e_n(x_1) = E_n(f)$ por lo que nos resta ver que existe $x_2 \in [a,b]$ tal que $e_n(x_2) = -E_n(f)$, si $e_n(x) > -E_n(f)$ en $[a,b]$ entonces
$m = \min\{e_n(x) \mid x \in [a,b]\} > -E_n(f)$. Sea $c \in (0, \frac{E_n(f)+m}{2}]$ y consideremos a $q_n = p_n^* + c \in \mathcal{P}_n$ entonces $f(x) - q(x) = e_n(x) - c$ pero $-(E_n(f) - c) \leq m - c \leq$
$\leq E_n(f) - c$ de donde $\|f - q_n\|_{[a,b]} \leq E_n(f) - c < E_n(f)$ lo cuál no es posible. Por lo tanto existe $x_2 \in [a,b]$, $x_1 \neq x_2$ tal que cumple con las condiciones del lema.

Q.E.D.

(III.27) Discusión. Ahora caracterizaremos a $p_1^*(x) = a_0 + a_1 x \in \mathcal{P}_1$ un mejor aproximador lineal a $f \in C([a,b])$. Por (III.26) existen $x_1, x_2 \in [a,b]$ distintos tales que cumplen con las condiciones del lema anterior. Supongamos que x_1 y x_2 son los únicos puntos en $[a,b]$ tales que $|e_1(x_i)| = E_1(f), i = 1,2$. Para fijar ideas digamos que $e_1(x_1) = E_1(f)$ y $e_1(x_2) = -E_1(f)$ con $x_1 < x_2$. Por continuidad encontramos dos intervalos cerrados I_1, I_2 tales que $x_1 \in I_1$, $x_2 \in I_2$ donde $e_1(x) > \frac{E_1}{2}$ $\forall\; x \in I_1$ y $e_1(x) < -\frac{E_1}{2}$ $\forall\; x \in I_2$, es claro que $I_1 \cap I_2 = \emptyset$ y $e_1(x) = f(x) - p_1^*(x)$ tiene signos opuestos en esos intervalos. Sea $x_0 \in [x_1, x_2]$ pero exterior a los intervalos cerrados I_1, I_2, consideramos una recta ℓ_1 que pasa por x_0 y tal que tiene el mismo signo que $e_1(x)$ en I_1 e I_2. Sea $J = \overline{[a,b] \setminus I_1 \setminus I_2}$ (cerradura de $[a,b] \setminus I_1 \setminus I_2$), escribimos $\widehat{E}_1(f) = ||f - p_1^*||_J$. Dado que $x_1, x_2 \notin J$ entonces $\widehat{E}_1(f) < E_1(f)$.

Por último elegimos $c \in (0, \frac{E_1(f) - \widehat{E}_1(f)}{2||\ell_1||_{[a,b]}})$ **(1)** y consideramos a $\widehat{p}_1 = p_1^* + c\ell_1 \in \mathcal{P}_1$. Si $x \in I_1$, entonces por **(1)**:

$0 < c\ell_1(x) < \ell_1(x)\frac{E_1 - \widehat{E}_1}{2||\ell_1||_{[a,b]}} \leq \frac{E_1 - \widehat{E}_1}{2} \leq \frac{E_1}{2} < e_1(x)$.

así tenemos:

$|e_1(x) - c\ell_1(x)| = e_1(x) - c\ell_1(x) \leq E_1 - \min_{x \in I_1}\{c\ell_1(x)\} < E_1$, análogamente

$|e_1(x) - c\ell_1(x)| < E_1$ en I_2. Ahora si $x \in J$, entonces por **(1)**:

$$||f - p_1^* - c\ell_1||_J \leq \widehat{E}_1 + c||\ell_1||_J < \widehat{E}_1 + \frac{E_1 - \widehat{E}_1}{2}\frac{||\ell_1||_J}{||\ell_1||_{[a,b]}} < \widehat{E}_1 + \frac{E_1 - \widehat{E}_1}{2} < E_1.$$

Por lo tanto $||f - p_1^* - c\ell_1||_J < E_1$, lo cuál no es posible, por lo que existe $x_3 \in [a,b]$ distinto de x_1 y x_2 tal que $|e_1(x_3)| = E_1(f)$ y además el razonamiento anterior implica que $e_1(x_3) = -e_1(x_2) = e_1(x_1)$, es decir, existen al menos tres puntos distintos en $[a,b]$ para los cuales el error E_1 se alcanza alternando de signo.

Lo anterior motiva la siguiente:

(III.28) Definición. Un conjunto $\{x_0, \ldots, x_k\}$ de $k+1$ puntos distintos tales que $a \leq x_0 < x_1 < \ldots < x_k \leq b$ se llama un ***k*-conjunto alternante** para una función $e \in C([a,b])$ si:

(i) $|e(x_i)| = ||e||_{[a,b]}$ para toda $i = 0, \ldots, k$

(ii) $e(x_i) = -e(x_{i+1})$ para toda $i = 0, \ldots, k$.

(III.29) Observación.

Si $e \in C([a,b])$ tiene un n-conjunto alternante ($n > 2$) entonces tiene un m-conjunto alternante para toda $m \in \{2, 3, \ldots, n\}$, por lo que si e *no* tiene un n-conjunto alternante entonces *no* tiene un m-conjunto alternante si $m \geq n$.

(III.30) Teorema de Equioscilación de Chebyshev (1859).

Sea $f \in C([a,b])$ entonces $p_n \in \mathcal{P}_n$ es un mejor aproximador uniforme (de grado n) para f si y sólo si existe un m-conjunto alternante con $m \geq n+1$ puntos para $e_n(f)$.

Demostración. Sea $p_n^* \in \mathcal{P}_n$ un mejor aproximador uniforme para f, como $e_n \in C([a,b])$, entonces e_n es uniformemente continua de donde dada $\varepsilon' = \frac{E_n}{2}$ existe $\delta(\varepsilon') > 0$ tal que $|e_n(x) - e_n(y)| < \varepsilon'$, si $|x - y| < \delta$.

Ahora dividimos al intervalo $[a,b]$ en subintervalos cerrados de longitud menor ó igual que δ, llamamos a los intervalos en los cuales $|e_n(x)|$ alcanza su máximo valor como $I_1, \ldots, I_m$. Dado que $e_n(x)$ puede variar a lo más $\frac{E_n}{2}$ en cualquiera de esos intervalos, se tiene que $e_n(x) > \varepsilon'$ ó $e_n(x) < -\varepsilon'$ **(1)**. Sea $u_1, \ldots, u_m (= \pm 1)$ el signo de $e_n(x)$ sobre esos intervalos. Lo que deseamos demostrar es que esta lista tiene al menos $n+1$ cambios de signo, lo cuál haremos por *reducción al absurdo*, es decir: Si existen menos de $n+1$ cambios de signo encontraremos un polinomio cuyo error E_n' es menor que el de p^*.

Por el lema (III.26) sabemos que al menos debe de haber cambios de signo a través de los intervalos $I_1, \ldots, I_m$, por lo cuál procedemos a clasificarlos por grupos consecutivos donde los signos sean el mismo, para fijar ideas supondremos que el primer grupo tiene signo positivo:

Primer grupo	$I^1, \ldots, I^{j_1}$	signo	$(+)$
Segundo grupo	$I^{j_1+1}, \ldots, I^{j_2}$	signo	$(-)$
Tercer grupo	$I^{j_2+1}, \ldots, I^{j_3}$	signo	$(+)$
K-ésimo grupo	$I^{j_{k-1}+1}, \ldots, I^{j_k}$	signo	$(-)^{k-1}$

donde $I^{(1)} = I_1, I^{(j_m)} = I_m$. El ordenamiento nos muestra $(k-1)$ cambios de signo, ***supongamos*** que $k-1 < n+1$, es decir, $k < n+2$. Consideremos $I^{(j_i)}$ e $I^{(j_i+1)}$, estos intervalos cerrados no pueden ser adyacentes por **(1)**, entonces existe $x_i \in [a,b]$ tal que $I^{(j_i)} < x_i < I^{(j_i+1)}$ para toda $i = 1, \ldots k-1$. Definimos $p(x) = \prod_{i=1}^{k-1}(x_i - x)$, dado que $k < n+2$ entonces $p \in \mathcal{P}_n$ y es claro que $p(x)$ sólo se anula en x_i, como cada x_i esta entre los intervalos $I_1, \ldots, I_m$ entonces $p(x)$ debe tener signo constante en cada intervalo ya que sobre el primer grupo cada factor de $p(x)$ es positivo, en el segundo grupo todos salvo el primer factor son negativos por lo que $p(x)$ es negativo y así sucesivamente $p(x)$ coincide en signo con $e(x)$ sobre los intervalos $I_1, \ldots, I_m$.

Sean $J = \overline{[a,b] \setminus I_1 \setminus \ldots \setminus I_m}$ y $E_n' = \|e_n\|_J$. Cómo el máximo de e_n sólo se alcanza en $I_1, \ldots, I_m$ entonces $E_n' < E_n$, sea $c \in \mathbb{R}$ tal que $0 < c < \frac{E_n - E_n'}{2\|p\|_{[a,b]}}$ **(2)**, y consideremos $p_n = p_n^* + cp$. Si $x \in J$ entonces por **(2)** $\|e_n - cp\|_J \le \|e_n\|_J + c\|p\|_J <$
$< E_n' + \frac{E_n - E_n'}{2} < \frac{E_n}{2} < e_n(x)$. Si $x \in I_j$ para alguna $j = 1, \ldots, m$ y $p(x) > 0$ entonces por **(1)** y **(2)** se tiene que:
$0 < cp(x) < \frac{E_n - E_n'}{2\|p\|_{[a,b]}} p(x) < \frac{E_n}{2} < e_n(x)$ por lo tanto tenemos que $|e_n(x) - cp(x)| =$
$= e_n(x) - cp(x) \le E_n - c \min_{x \in I_j} p(x) < E_n$. Para el caso en que $p(x) < 0$ procedemos de la siguiente manera: $\frac{p(x)}{\|p\|_{[a,b]}} \ge -1$ por **(1)** y **(2)** $cp(x) > \frac{E_n - E_n'}{2\|p\|_{[a,b]}} p(x) \ge \frac{E_n' - E_n}{2} \ge$
$\ge -\frac{E_n}{2} > e_n(x)$, de donde: $|e_n(x) - cp(x)| = cp(x) - e_n(x) \le E_n + c \max_{x \in I_j} p(x) < E_n$ ya que $p(x) < 0$. Por lo anterior $\|f - p_n\| < E_n$ lo cuál no es posible.

Para el recíproco, sea $a \le x_0 < x_1 < \ldots < x_{n+1} \le b$ un $(n+2)$-conjunto alternante para $f - p_n^*$ y $E = \|f - p_n^*\|$ entonces $E_n(f) \le E$. Supongamos que $E_n(f) < E$. Sea p_n un mejor aproximador de grado $\le n$ para f por (III.6), entonces $\|f - p_n\|_{[a,b]} = E_n(f) < E$ **(3)**.

Por otro lado $p_n^*(x_i) - p(x_i) = p_n^*(x_i) - f(x_i) - (p_n(x_i) - f(x_i))$ haciendo $t_i = p_n^*(x_i) - f(x_i)$ de **(3)** se tiene que $|t_i| < E$ pero $p_n^*(x_i) - p_n(x_i) = \pm E - t_i$

alternando de signo $(n+2)$ veces por lo que $p_n^* - p_n \in \mathcal{P}_n$ tiene $(n+1)$ raíces por lo tanto $p_n^* = p$.

Q.E.D.

(III.31) Observaciones.

(i) La demostración anterior es válida para el caso en que el primer grupo de intervalos tenga signo negativo simplemente considerando a $-p(x)$ en lugar de $p(x)$, para que e_n y $-p$ tengan el mismo signo en cada grupo de intervalos.

(ii) Observemos que el teorema (III.30) *no* garantiza la unicidad del conjunto alternante de $e_n(x)$, por ejemplo para $f(x) = \text{sen}4x$ en $[-\pi, \pi]$ la mejor constante que aproxima a f es la constante 0 y hay 16 diferentes conjuntos de alternanacia de 2 puntos en $[-\pi, \pi]$.

(iii) Tampoco el teorema *dice* que existe un conjunto alternante que tenga exactamente $(n+2)$ puntos. Consideresé el ejemplo del inciso anterior. El teorema nos garantiza que $p_0^* = p_1^* = \ldots = p_6^* = 0$ ya que $||\text{sen}4x||_{[-\pi,\pi]} = 1$ y sen$4x$ " alterna" 8 veces en $[-\pi, \pi]$ y $p_7 = 0$ no es un mejor aproximador.

(III.32) Corolario. *Si $f \in C([a,b])$ tiene un conjunto alternante de (n+2) puntos entonces $p_k^* = 0$ y $E_k(f) = ||f||$ para toda $k = 0, \ldots, n$.*

(III.33) Corolario. *Sea $f \in C^2([a,b])$ y $f''(x) > 0$ para toda $x \in [a,b]$. Si $p_1^*(x) = a_0 + a_1 x$ es un mejor aproximador lineal uniforme de f entonces:*
$a_1 = \frac{f(b)-f(a)}{b-a}$ y $a_0 = \frac{1}{2}(f(a) + f(c)) - a_1\frac{a+c}{2}$, donde c es la única solución de $f'(c) = \frac{f(b)-f(a)}{b-a}$ y $E_1(f) = \frac{f(a)(b-c)+f(b)(c-a)-f(c)(b-a)}{2(b-a)}$.

Demostración. Por el TVM-D existe $c \in (a,b)$ tal que $f'(c) = \frac{f(b)-f(a)}{b-a}$ como $f'' > 0$ entonces f' es creciente en $[a,b]$ por lo que c es única. Sea $e_1(x) =$ $= f(x) - p_1^*(x)$ por (III.30) existen al menos tres puntos distintos $a \leq x_0 < x_1 <$ $< x_2 \leq b$ tales que $|e_1(x_i)| = E_1(f)$. Es claro que $x_1 \in (a,b)$ entonces $e_1'(x_1) = 0$, pero $e_1'(x) = f'(x) - a_1$ y $e_1''(x) = f''(x) > 0$ por lo que $e_1'(x)$ es creciente y como $\{x_0, x_1, x_2\}$ es un conjunto alternante entonces $x_0 = a$ y $x_2 = b$ y $x_1 = c$. Como $e_1'(x_1) = 0$ con $a_1 = \frac{f(b)-f(a)}{b-a}$ se tiene:

$$f(a) - (a_0 + a_1 a) = -\{f(c) - (a_0 + a_1 c)\} = f(b) - (a_0 + a_1 b)$$

y $E_1(f) = f(a) - (a_0 + a_1 a)$, de donde se sigue que:
$a_0 = \frac{1}{2}(f(a) + f(c)) - a_1\frac{a+c}{2}$ y que $E_1(f) = \frac{f(a)(b-c)+f(b)(c-a)-f(c)(b-a)}{2(b-a)}$, (lo cuál es positivo por la convexidad de la función f).

Hasta el momento no hemos probado la unicidad del mejor aproximador uniforme de grado menor ó igual que n de una función $f \in C([a,b])$, lo cuál se hará en el siguiente:

(III.34) Teorema. *Dada $f \in C([a,b])$ el mejor aproximador uniforme $p_n^* \in \mathcal{P}_n$ es único.*

Demostración. Sean p_n^* y p_n dos mejores aproximadores de grado$\leq n$ para f entonces $E_n(f) = \|f - p_n\| = \|f - p_n^*\|$. Por (III.16) tenemos que $\frac{1}{2}(p_n + p_n^*)$ es un mejor aproximador de grado $\leq n$ para f, de donde por (III.30) existen $x_1, \ldots, x_{n+2}$ $(n+2)$ puntos en $[a,b]$ tales que $E_n(f) = |f(x_i) - \frac{1}{2}(p_n(x_i) + p_n^*(x_i))|$ para toda $i = 1, \ldots, n+2$. Por lo que se tiene que:

$$E_n(f) = \left|\tfrac{1}{2}\left(f(x_i) - p_n^*(x_i)\right) + \tfrac{1}{2}\left(f(x_i) - p_n(x_i)\right)\right| \leq \tfrac{1}{2}|f(x_i) - p_n^*(x_i)| + \\ + \tfrac{1}{2}|f(x_i) - p_n(x_i)| \leq \tfrac{1}{2}E_n(f) + \tfrac{1}{2}E_n(f) = E_n(f).$$

de donde se tiene que en las dos desigualdades anteriores necesariamente hay igualdad. Así pues:

(i) $|f(x_i) - p_n^*(x_i)| = |f(x_i) - p_n(x_i)| = E_n(f)$ para toda $i = 1, \ldots, n+2$

(ii) $sign\,[f(x_i) - p_n^*(x_i)] = sign\,[f(x_i) - p_n(x_i)]$ para toda $i = 1, \ldots, n+2$

por lo que $p_n^*(x_i) = p_n(x_i)$ para toda $i = 1, \ldots, n+2$, por lo tanto $p_n^* = p_n$.

Q.E.D.

CAPITULO IV

Polinomios de Chebyshev

Uno de los pocos casos en donde es posible calcular exactamente el mejor aproximador uniforme es el siguiente: sea $f(x) = x^{n+1}$ en $[-1,1]$ y por (III.6) existe $p_n^* \in \mathcal{P}_n$ tal que $E_n(f) = ||f - p_n^*||_{[-1,1]} = r$ por lo que $e_n(x) = x^{n+1} - p_n^*(x)$ satisface que $r = ||e_n||_{[-1,1]} = |e_n(x_i)|$ **(1)** para $i = 0, \ldots, n+1$ donde $-1 \leq x_0 <$ $< \ldots < x_{n+1} \leq 1$ (por el teorema de Equioscilación). Cómo $e_n \in \mathcal{P}_{n+1}$, $r^2 - e_n^2 \in$ $\in \mathcal{P}_{2n+2}$ y por **(1)** $r^2 - e_n^2(x_j) = 0$ para toda $j = 0, \ldots, n+1$. Cómo $e_n^2(x) \leq r^2$ en $[-1,1]$, e_n^2 tiene un máximo relativo en x_j para cada $j = 0, \ldots, n+1$, de donde se sigue que $\frac{d}{dx}(e_n^2(x)) = 0$ en $x = x_j$ $j = 1, \ldots, n$ (puntos críticos interiores), entonces $\frac{d}{dx}(r^2 - e_n^2(x)) = 0$ en $x = x_j$. Por lo tanto $r^2 - e_n^2(x)$ tiene un cero de multiplicidad no menor que 2 en $x = x_j$ $j = 1, \ldots, n$ y como $r^2 - e_n^2(x)$ es de grado $2n+2$ se sigue que $r^2 - e_n^2(x)$ tiene *exactamente* $2n+2$ raíces en $[-1,1]$.

De lo anterior x_0 y x_{n+1} no pueden ser puntos interiores por lo que concluimos que $x_0 = -1$ y $x_{n+1} = 1$ son raíces simples y $x_1, \ldots, x_n$ son raíces dobles.

Ahora consideremos el polinomio: $(1 - x^2)\,[e_n'(x)]^2 \in \mathcal{P}_{2n+2}$ **(2)**, es claro que tiene raíces simples en -1 y 1, así cómo raíces dobles en $x_1, \ldots, x_n$ por lo que **(2)** y $r^2 - e_n^2(x)$ tienen las mismas raíces, de donde existe k una constante tal que: $(1-x^2)\,[e_n'(x)]^2 = k\,(r^2 - e_n^2(x))$ **(3)**. Claramente k es el coeficiente del término de grado mayor el cuál vamos a calcular de la siguiente manera: $e_n(x) = x^{n+1} - p_n^*(x)$ y $e_n'(x) = (n+1)x^n - (p_n^*)'(x)$ analizando **(3)** se obtiene que $k = (n+1)^2$, por lo que llegamos a la siguiente ecuación diferencial de la función error: $(1 - x^2)\,[e_n'(x)]^2 =$ $= (n+1)^2\,(r^2 - e_n^2(x))$, como $e_n'(x) \in \mathcal{P}_n$ y $e_n'(x_j) = 0$ para cada $j = 1, \ldots, n$ entonces $e_n'(x)$ no cambia de signo en $[-1, x_1]$ ya que de no ser así $e'(x)$ tendría $(n+1)$ raíces, por lo que supondremos que $e_n'(x) \geq 0$ en $[-1, x_1]$ y así tenemos: $\frac{e_n'(x)}{\sqrt{r^2 - e_n^2(x)}} =$ $= \frac{(n+1)}{\sqrt{1-x^2}}$ **(4)**, resolviendo ésta ecuación tenemos $\arccos\left(\frac{e_n(x)}{r}\right) = (n+1)\arccos x + c$ haciendo $\theta = \arccos x$ entonces $x = \cos\theta$ y ya que $\cos\theta$ es una biyección entre $[0,\pi]$ y $[-1,1]$ tenemos : $e_n(x) = r(\cos((n+1)\theta) + c)$ con $0 \leq \theta_1 \leq \theta \leq \pi$ donde $x_1 = \cos\theta_1$ **(5)**. Por otro lado $e_n(-1) = -r$ ya que $e_n'(-1) > 0$, así tenemos que $\cos((n+1)\pi + c) = -1$ ya que $\pi = \arccos(-1)$ por lo que $c = m\pi$ $m \in \mathbb{Z}$ y $(n+m+1)$ es impar. Sustituyendo en **(5)**: $e_n(x) = \pm r\cos((n+1)\theta)$ **(6)**, de donde $\cos((n+1)\theta) \in \mathcal{P}_{n+1}$ en la indeterminada $x = \cos\theta$ con $\theta \in [0,\pi]$. Cabe observar que si $e_n'(x) \leq 0$ en $[-1, x_1]$ elegimos la raíz negativa en el lado derecho de **(4)** y el razonamiento se sigue igual que antes.

Por lo que daremos la siguiente:

(IV.1) Definición. (Polinomios de Chebyshev).

Denotaremos por $T_n(x) = \cos n\theta$ con $x = \cos\theta$, $\theta \in [0,\pi]$ para toda $n \in \mathbb{N}\cup\{0\}$ el cuál definimos cómo el *n-ésimo polinomio de Chebyshev* .

Veremos algunas propiedades de los polinomios de Chebyshev:

(IV.2) Teorema.

(i) $T_{n+1}(x) = 2xT_n - T_{n-1}(x)$ *para toda* $n \in \mathbb{N}$.

(ii) Si n es par (impar) entonces T_n es una función par(impar) y el grado de T_n es n.

(iii) El coeficiente del término principal de T_n es 2^{n-1}.

(iv) $T_n(x)$ tiene raíces simples en $x_k = \cos(\frac{2k-1}{2n}\pi)$ para $k = 1,\ldots,n$.

(v) $T_n(x)$ tiene un (n+1)-conjunto alternante $\{y_0,\ldots,y_n\}$ en $[-1,1]$ donde $y_k = \cos\frac{k}{n}\pi$.

(vi) $T_n(T_m(x)) = T_m(T_n(x)) = T_{nm}(x)$ (propiedad de semigrupo).

(vii) $T_m(x)T_n(x) = \frac{1}{2}\left[T_{n+m}(x) + T_{m-n}(x)\right]$ con $m > n$.

(viii) $T_n(x) = \frac{1}{2}\{(x+\sqrt{x^2-1})^n + (x-\sqrt{x^2-1})^n\}$.

(ix) $(1-x^2)T_n''(x) - xT_n'(x) + n^2T_n(x) = 0$.

Demostración.

(i) Sabemos que: $\cos(n+1)\theta = \cos(n\theta)\cos(\theta) - \operatorname{sen}(n\theta)\operatorname{sen}(\theta)$ y $\cos(n-1)\theta =$
$= \cos(n\theta)\cos(\theta) + \operatorname{sen}(n\theta)\operatorname{sen}(\theta)$ de donde se tiene la igualdad

$2\cos(n\theta)\cos(\theta) = \cos((n+1)\theta) + \cos((n-1)\theta)$ y por la definición de $T_n(x)$ entonces se tiene que $T_{n+1}(x) = 2xT_n - T_{n-1}(x)$ para toda $n \in \mathbb{N}$.

(ii) Observemos que $T_0(x) = 1$ y $T_1(x) = x$, así cómo si $T_n(x)$ es par (impar) entonces $xT_n(x)$ es impar (par) y por el inciso anterior se sigue el resultado.

(iii) Se sigue por inducción y aplicando (i).

(iv) $T_n(x_k) = \cos\left(\frac{2k-1}{2}\pi\right) = 0$ para toda $k = 1,\ldots,n$ y $T_n'(x_k) = \frac{n\operatorname{sen}(n\arccos x)}{\sqrt{1-x^2}}$ de donde $T_n'(x_k) \neq 0$ por lo tanto $x_k = \cos(\frac{2k-1}{2n}\pi)$ son ceros simples de $T_n(x)$.

(v) Del inciso anterior tenemos: $T_n'(y_k) = 0$ para cada $k = 0,\ldots,n$. Ahora por otro lado:$T_n(y_k) = \cos(k\pi) = (-1)^k$ pero $x \in [-1,1]$ y por la definición de $T_n(x)$ se tiene que $|T_n(x)| \leq 1$, por lo que $\{y_0,\ldots,y_n\}$ es un (n+1)-conjunto alternante.

(vi) $T_m(T_n(x)) = T_m(\cos n(\cos^{-1}x) = \cos(m\cos^{-1}(n\cos n(\cos^{-1}x))) =$
$= \cos(mn(\cos^{-1}x)) = T_{nm}(x)$.

(vii) Es clara utilizando la identidad trigonométrica:
$\cos x \cos y = \frac{1}{2}\{\cos(x+y) - \cos(x-y)\}$ y la definición de $T_n(x)$.

(viii) $T_n(x) = Re(\cos\theta + \mathrm{sen}\theta)^n = \frac{1}{2}\left[(\cos\theta + i\mathrm{sen}\theta)^n + (\cos\theta - i\mathrm{sen}\theta)^n\right]$ pero
$\mathrm{sen}\theta = \sqrt{1-x^2}$, $x = \cos\theta$ por lo que $i\mathrm{sen}\theta = \sqrt{x^2-1}$, por lo tanto tenemos que:
$T_n(x) = \frac{1}{2}\{(x+\sqrt{x^2-1})^n + (x-\sqrt{x^2-1})^n\}$.

(ix) Como $T_n(x) = \cos(n\theta)$ con $\theta = \arccos x$ entonces $T_n'(x) = \frac{n\mathrm{sen}(n\theta)}{\sqrt{1-x^2}}$ y así
$T_n''(x) = \frac{1}{1-x^2}\{-n^2T_n(x) + xT_n'(x)\}$ por lo que $(1-x^2)T_n''(x) - xT_n'(x) +$
$+ n^2T_n(x) = 0$.

Resumiremos lo hecho al principio de está sección en el siguiente:

(IV.3) Teorema. *Sea $p_n^* \in \mathcal{P}_n$ el mejor aproximador de x^{n+1} en $[-1,1]$ entonces $x^{n+1} - p_n^*(x) = 2^{-n}T_{n+1}(x)$ y $E_n(x^{n+1}) = 2^{-n}$.*

Demostración. Por **(6)** sabemos que $e(x) = \pm rT_{n+1}(x)$ pero $e_n(x) = x^{n+1} - p_n^*(x)$ por (IV.2)(iii) tenemos que $r = 2^{-n}$ por lo que $x^{n+1} - p_n^*(x) = 2^{-n}T_{n+1}(x)$ y como $e_n(x) = 2^{-n}T_{n+1}(x)$ es válido para $x \in [-1, x_1]$ entonces es cierto para $x \in [-1,1]$ y por (IV.2)(v) $E_n(x^{n+1}) = 2^{-n}$.

Q.E.D.

Ahora veremos algunas propiedades más sobre el polinomio mónico $\widehat{T}_{n+1}(x) = 2^{-n}T_{n+1}(x)$, el cuál es conocido como el **polinomio de Chebyshev normalizado.**

(IV.4) Teorema. *Sea $\check{\mathcal{P}}_n$ la clase de todos los polinomios mónicos de grado menor ó igual que n entonces: $\|p\|_{[-1,1]} \geq \frac{1}{2^{n-1}}$ para todo $p \in \check{\mathcal{P}}_n$. La igualdad ocurre si y sólo si $p = \widehat{T}_n$.*

Demostración. Por (IV.2)(v), sabemos que $|\widehat{T}_n|$ asume su valor máximo $\frac{1}{2^{n-1}}$ en (n+1) puntos. Supongamos que existe $p_0 \in \check{\mathcal{P}}_n$ tal que $\|p_0\|_{[-1,1]} < \frac{1}{2^{n-1}}$, consideremos a $Q(x) = \widehat{T}_n(x) - p_0(x)$, es claro que $Q \in \mathcal{P}_{n-1}$, por otro lado por (IV.2)(v) tenemos que $Q(y_k) = \frac{(-1)^k}{2^{n-1}} - p_0(y_k)$ para toda $k = 0,\ldots,n$ donde $y_k = \cos\frac{k}{n}\pi$ pero $|p_0(y_k)| < \frac{1}{2^{n-1}}$, de lo anterior tenemos que $Q(x)$ tiene un (n+1)-conjunto alternante por lo tanto $Q(x)$ tiene n raíces de donde se sigue que $Q \equiv 0$, y así $p_0 \equiv \widehat{T}_n$ pero $\|p_0\|_{[-1,1]} = \|\widehat{T}_n\|_{[-1,1]} = \frac{1}{2^{n-1}}$, lo cuál no es posible.

Q.E.D.

(IV.5) Corolario. *Sea $p \in \mathcal{P}_n$, $p(x) = a_nx^n + \ldots + a_0, a_n \neq 0$, entonces $\|p\|_{[a,b]} \geq |a_n|\frac{(b-a)^n}{2^{2n-1}}$. La igualdad ocurre si y sólo si $p(x) = c\widehat{T}_n\left(\ell^{-1}(x)\right)$, donde $c = a_n\frac{(b-a)^n}{2^n}$ y $\ell(t) = \frac{a+b}{2} + \frac{b-a}{2}t$.*

Demostración. Consideremos $\ell:[-1,1]\to[a,b]$ definida por $x=\ell(t)=$ $=\frac{a+b}{2}+\frac{b-a}{2}t$ y definimos $\widehat{p}(t)=p\circ\ell\in\mathcal{P}_n$ cuyo coeficiente principal es c. Aplicando (IV.4) a $\frac{1}{c}\widehat{p}(t)$ obtenemos la primera parte del resultado. Para la igualdad tenemos que $\widehat{p}=\widehat{T}_n$, de donde se tiene que $p=c\widehat{T}_n(\ell^{-1})$.

Q.E.D.

Con el siguiente resultado también encontramos el mejor aproximador a $f(x)=x^{n+2}$ con polinomios en $\mathcal{P}_n$.

(IV.6) Lema. *Sea $f(x)=x^{n+2}$ en $[-1,1]$ entonces el mejor aproximador $p_n^*\in\mathcal{P}_n$ a f es $x^{n+2}-\frac{1}{2^{n+1}}T_{n+2}(x)$ y $E_n=(x^{n+2})=\frac{1}{2^{n+1}}$.*

Demostración. Por el Teorema de equioscilación de Chebyshev es fácil comprobar que $T_{n+2}(x)=2^{n+1}x^{n+2}+b_nx^n+$términos de grado menor, por lo que $x^{n+2}-\frac{1}{2^{n+1}}T_{n+2}(x)\in\mathcal{P}_n$ es el mejor aproximador y además $E_n=(x^{n+2})=\frac{1}{2^{n+1}}$.

Ahora regresemos a ver algunas otras propiedades de los polinomios de Chebyshev.

(IV.7) Teorema.

(i) $\displaystyle\int_{-1}^{1}T_k(x)T_m(x)\frac{dx}{\sqrt{1-x^2}}=0$ *si* $m\neq k$.

(ii) $\displaystyle\int_{-1}^{1}T_k^2(x)\frac{dx}{\sqrt{1-x^2}}=\begin{cases}\frac{\pi}{2} & si\ k>0\\ \pi & si\ k=0\end{cases}.$

(iii) El conjunto $\{T_0,\ldots,T_n\}$ es una base para $\mathcal{P}_n$.

(iv) Sea $p\in\mathcal{P}_n$ no constante tal que $||p||_{[-1,1]}=1$. Si existen exactamente (n+2) puntos $x_0<x_1<\ldots<x_{n+1}$, en $[-1,1]$ tales que $|p(x_i)|=1$ entonces $p=\pm T_n$.

Demostración. Para los incisos (i) y (ii) con $k>0$ hay que integrar por partes haciendo $u=T_k(x)$ y $dv=\frac{T_m(x)dx}{\sqrt{1-x^2}}$, para el caso en que $k=0$ del inciso (ii) es fácil directamente verificar el resultado. Por último el inciso (iii) es inmediato de los incisos anteriores. La independencia lineal se sigue de (ii) y dado $p\in\mathcal{P}_n$ se tiene que $p=A_0T_0+\ldots+A_nT_n$ donde $A_m=\displaystyle\int_{-1}^{1}p(x)T_m(x)\frac{dx}{\sqrt{1-x^2}}$.

Para (iv) procedemos de la siguiente manera: como p es derivable entonces $p'\in\mathcal{P}_{n-1}\subset\mathcal{P}_n$ y además $x_1<\ldots<x_n$ son puntos interiores de $[-1,1]$ de donde $p'(x_i)=0$ para toda i $=1,\ldots,n$, por lo tanto $x_1,\ldots,x_n$ son todas las raíces de p'. El conjunto $\{x_0,\ldots,x_{n+1}\}$ es un conjunto alternante de p ya que $|p(x_0)|=||p||_{[-1,1]}$ y $p(x_{i+1})=-p(x_i)$, de no ser así por el Teorema de Rolle existe $\xi\in(x_i,x_{i+1})$ tal que $p'(\xi)=0$ lo cuál no es posible por lo anterior.

Por otro lado como $p\in\mathcal{P}_n$ entonces $p(x)=\displaystyle\sum_{k=0}^{n}a_kx^k$ con $a_n\neq0$ de donde

$\frac{p}{a_n} \in \mathcal{P}_n$ es el mejor aproximador mónico de grado$\leq n$ a $f \equiv 0$ en $[-1,1]$ pero por el Teorema de Equioscilación $\widehat{T}_n = \frac{1}{2^{n-1}}T_n$ es el mejor aproximador mónico y por unicidad se tiene que $a_n = 2^{n-1}$ y además $||p||_{[-1,1]} = ||T_n||_{[-1,1]}$ de donde $|p| = |T_n|$ y por lo tanto $p = \pm T_n$.

Q.E.D.

(IV.8) Observación. Dada $\varepsilon > 0$ existe un polinomio mónico $p = p(\varepsilon) \in \mathcal{P}_n$ tal que $||p||_{[a,b]} < \varepsilon$ si y sólo si $b-a < 4$, ya que por (IV.5) $||p||_{[a,b]} \geq \frac{(b-a)^n}{2^{2n-1}}$, por lo que si $b-a \geq 4$ entonces $||p||_{[a,b]} \geq 2\left(\frac{b-a}{4}\right)^n \geq 2$, lo cuál no es posible ya que $||p||_{[a,b]} < \varepsilon$ por lo que $b-a<4$.

Inversamente si $b-a<4$ y dada $\varepsilon > 0$, se sigue de (IV.5) que $||p||_{[a,b]} \geq \frac{(b-a)^n}{2^{2n-1}}$ para todo polinomio mónico, sea $p = p(\varepsilon)$ el mejor aproximador mónico de $f \equiv 0$ en $[a,b]$ entonces $||p||_{[a,b]} = 2\left(\frac{b-a}{4}\right)^n = k$ como $\frac{b-a}{4} < 1$ entonces $k \to 0$ cuando $n \to \infty$ por lo que dada $\varepsilon > 0$ se tiene que $||p||_{[a,b]} < \varepsilon$.

Continuando con el estudio de los polinomios de Chebyshev vamos a obtener una función generatriz la cual nos dara los coeficientes de los polinomios, para lo cuál introduciremos el **núcleo de Poisson.**

Sabemos que si $r \in (0,1)$ y $w \in \mathbb{C}$ es tal que $|w| < 1$ entonces $\sum_{k=0}^{\infty} r^k w^k =$

$= \frac{1}{1-rw}, |rw| < 1$ por lo que si $w = e^{i\theta}$ tenemos que: $\sum_{k=0}^{\infty} r^k e^{ik\theta} = \frac{1}{1-re^{i\theta}}$ y

de manera análoga se tiene que $\sum_{k=0}^{\infty} r^k e^{-i\theta} = \frac{1}{1-re^{-i\theta}}$. Por lo que tenemos que:

$$\sum_{k=0}^{\infty} r^k e^{ik\theta} + \sum_{k=0}^{\infty} r^k e^{-ik\theta} - 1 = \frac{1}{1-re^{i\theta}} + \frac{1}{1-re^{-i\theta}} - 1 = \frac{1-r^2}{1+r^2-2r\cos\theta}$$ **(Núcleo de Poisson)**. De donde se sigue que $\frac{1-r^2}{1+r^2-2r\cos\theta} = 2Re\left(\sum_{k=0}^{\infty} r^k e^{ik\theta}\right) - 1$ por lo que tenemos que:

$$\frac{1-r^2}{1+r^2-2r\cos\theta} = 1 + 2\sum_{n=0}^{\infty} r^n \cos(n\theta) \ \textbf{(1)}.$$

Recordemos que si n es par ó impar entonces T_n es par ó impar respectivamente por lo que si $k \leq n$ y deseamos hallar el coeficiente de x^k en $T_n(x)$ sólo debemos considerar los casos en los que k y n tienen la misma paridad.

Así tenemos que por **(1)**, el k-ésimo coeficiente de $T_n(x)$ es el doble del coeficiente de $r^n x^k$ en la expresión izquierda de **(1)**. Por lo que encontraremos prim el coeficiente de x^k (en términos de r) y entonces encontraremos el coeficiente de en la expansión del coeficiente de x^k.

$$\frac{1-r^2}{1-2rx+r^2}=\frac{1-r^2}{1+r^2}\frac{1}{1-\frac{2r}{1+r^2}}=\left(\frac{1-r^2}{1+r^2}\right)\sum_{j=0}^{\infty}\left(\frac{2r}{1+r^2}\right)^j x^j$$

ya que $|x|<1$ mientras que $(r-1)^2$ implica que $\frac{2r}{1+r^2}<1$.

De la expresión anterior tenemos que el coeficiente de x^k es:

$$\left(\frac{1-r^2}{1+r^2}\right)\left(\frac{2r}{1+r^2}\right)^k=(1-r^2)(2r)^k\sum_{j=0}^{\infty}\binom{j+k}{k}(-r^2)^j=$$

$$=2^k(1-r^2)\sum_{j=0}^{\infty}\binom{j+k}{k}(-1)^j r^{2j+k}.$$

Así pues se tiene que:

$$\left(\frac{1-r^2}{1+r^2}\right)\left(\frac{2r}{1+r^2}\right)^k=2^k\sum_{j=0}^{\infty}\binom{j+k}{k}(-1)^j r^{2j+k}-2^k\sum_{j=0}^{\infty}\binom{j+k}{k}(-1)^j r^{2j+k+2}$$

de donde cuando $j=\frac{n-k}{2}$ y $j=\frac{n-k}{2}-1$ en la primera y segunda serie respectivamente obtenemos los coeficientes de r^n, los cuales son:

$$2^k\left[(-1)^{\frac{n-k}{2}}\binom{\frac{n+k}{2}}{k}-(-1)^{\frac{n-k}{2}-1}\binom{\frac{n+k}{2}-1}{k}\right]=$$

$$=2^k(-1)^{\frac{n-k}{2}}\frac{2n}{n+k}\binom{\frac{n+k}{2}}{k}\ \mathbf{(2)}.$$

Por lo tanto el coeficiente de x^k en $T_n(x)$ se obtiene dividiendo la expresión (2) entre 2 por lo que obtenemos:

$$2^k(-1)^{\frac{n-k}{2}}\frac{n}{n+k}\binom{\frac{n+k}{2}}{k}$$

Veremos una prueba más del Teorema polinomial de Weierstrass, para el cuál necesitamos algunos resultados previos:

(IV.9) Lema. *Si $f:[-A,A]\to\mathbb{R}$ es una función par acotada (impar) con un mejor aproximador uniforme entonces tiene un mejor aproximador uniforme par(impar).*

Demostración.

Sea $f:[-A,A]\to\mathbb{R}$ una función par acotada, por hipótesis existe un mejor aproximador uniforme h de f, por definición $\|f-h\|_{[-A,A]}\le\|f-g\|_{[-A,A]}$ para toda función g acotada, definimos $h_1(x)=\frac{h(x)+h(-x)}{2}$, es claro que h_1 es par y $\|f-h_1\|_{[-A,A]}\le\|f-g\|_{[-A,A]}$, de donde h_1 es un mejor aproximador uniforme para f.

Analogámente para el caso en que f sea impar considerando $h_2(x) = \frac{h(x)-h(-x)}{2}$.

Q.E.D.

(IV.10) Definición. Un *polinomio trigonométrico T de grado$\leq n$* es una expresión de la forma $T(x) = \alpha_0 + \sum_{k=1}^{n} \alpha_k \cos(kx) + \beta_k \mathrm{sen}(kx)$ donde $\alpha_0, \alpha_k, \beta_k \in \mathbb{R}$ para toda $k = 1, \ldots, n$ y $\alpha_n^2 + \beta_n^2 \neq 0$.
Decimos que T es un *polinomio trigonométrico par* si $\beta_k = 0$ para toda $k = 1, \ldots, n$ y T es un *polinomio trigonométrico impar* si $\alpha_k = 0$ para toda $k = 0, \ldots, n$.

Denotaremos por $\mathcal{T}$ a la familia de los polinomios trigonométricos. Es claro que para todo $T \in \mathcal{T}$ existen $T_i \in \mathcal{T}$, $i = 1, 2$ con T_1 par mientras que T_2 es impar tales que $T = T_1 + T_2$.

(IV.11) Lema. $T(x) = \cos^k x$ *es un polinomio trigonométrico par de grado k para toda $k \in \mathbb{N}$.*

Demostración. Si $k = 1$ entonces $T(x) = \cos x$ y es claro que es par de grado 1. Supongamos que $T(x) = \cos^k x$ es par de grado k. Entonces $T(x) =$
$= \alpha_0 + \sum_{j=1}^{k} \alpha_j \cos(jx)$, $\alpha_k \neq 0$ por lo que $T(x) = \cos^{k+1} x = \cos^k x \cos x = \alpha_0 \cos x +$
$+ \sum_{j=1}^{k} \frac{1}{2} \alpha_j \{\cos(j+1)x - \cos(j-1)x\}$ de donde se sigue que $T(x) = \cos^{k+1} x$ es par y de grado (k+1) ya que $\frac{1}{2}\alpha_k \neq 0$.

Q.E.D.

(IV.12) Observación. El lema anterior es falso para potencias de $\mathrm{sen} x$

Recordando (II.15) y por (IV.11), toda $f \in C([0, \pi])$ se puede aproximar uniformemente por polinomios trigonométricos pares. Observando que si $f: [-\pi, \pi] \to \mathbb{R}$ es par entonces para $\varepsilon > 0$ dada y para $f_1 = f|_{[0,\pi]}: [0, \pi] \to \mathbb{R}$ existe $T_1 \in \mathcal{T}$ par tal que $||f_1 - T_1||_{[0,\pi]} < \varepsilon$, como f y T_1 son pares entonces $||f - T_1||_{[-\pi,\pi]} < \varepsilon$ por lo que tenemos el siguiente:

(IV.13) Teorema. *Sea $f \in PC([0, 2\pi])$. Dada $\varepsilon > 0$ existe $T \in \mathcal{T}$ par tal que $||f - T||_{[-\pi,\pi]} < \varepsilon$.*

Ahora daremos otra prueba del Teorema de Weierstrass (polinomial) ahora en $[-1, 1]$ (ver: (I.1)).

Demostración. Sea $f \in C([-1,1])$ definimos $g:[0,2\pi] \to \mathbb{R}$ cómo sigue: $g(\theta) = f(\cos\theta)$. Es claro que $g \in PC([0,2\pi])$ y es par. Por (IV.13) existe $T(\theta) = \alpha_0 + \sum_{k=1}^{N} \alpha_k \cos(k\theta)$ tal que $||g - T||_{[0,2\pi]} < \varepsilon$. Haciendo el cambio de variable $\theta = \arccos x$ entonces $g(\theta) = f(x)$ y $T_k(x) = \cos(k\theta)$ por lo que $T(x)$ es un polinomio en x y además tenemos que $||f - T||_{[-1,1]} < \varepsilon$.

Q.E.D.

Para concluir con el capítulo veremos algunos ejemplos donde aplicaremos algunos resultados.

(IV.14) Ejemplos.

(i) Sea $X = A([a,b])$ el espacio de las funciones acotadas en $[a,b]$ y sea $W = \mathcal{P}_n$. Por (III.4) se tiene que existe $p^* \in \mathcal{P}_n$ tal que $||f - p^*||_{[a,b]} = \min_{p \in \mathcal{P}_n} ||f - p||_{[a,b]}$ para $f \in X$.

Pero el mejor aproximador *no* es único ya que si consideramos a $[a,b] = [-1,1]$ y $f:[-1,1] \to \mathbb{R}$ definida por:

$$f(x) = \begin{cases} -1 & \text{si } x \in [-1,0] \\ 1 & \text{si } x \in (0,1] \end{cases}$$

es claro que $f \in X$ pero tiene al menos dos mejores aproximadores lineales ya que si consideramos a $p^* \equiv 0$ ó $p^*(x) = 2x$ tenemos que $||f - p^*||_{[-1,1]} = 1$. Este ejemplo muestra que no podemos omitir la hipótesis de que f sea continua en (III.34) para la unicidad del mejor aproximador uniforme polinomial.

(ii) **(Aproximación Pitágorica.)** Sea $f(x) = \sqrt{1+x^2}$, es fácil ver que f es convexa por lo que de (II.31) el mejor aproximador lineal uniforme a f en $[a,b] = [0,1]$ es $a_1 x + a_0$ donde: $a_1 = \frac{f(b)-f(a)}{b-a} = \sqrt{2} - 1$ y $a_0 = \frac{1}{2}(f(a)+$ $+f(c)) - a_1 \frac{a+c}{2}$ donde $f'(c) = \sqrt{2} - 1$ y c es la *única* solución, por lo que $f'(c) = \frac{c}{\sqrt{1+c^2}} = \sqrt{2} - 1$ y así $c = \left(\frac{\sqrt{2}-1}{2}\right)^{\frac{1}{2}}$. Por lo tanto el mejor aproximador está dado por: $a_1 x + a_0$ con $a_0 = \frac{2c+1}{2} + (\sqrt{2} - 1)$ y $c = \left(\frac{\sqrt{2}-1}{2}\right)^{\frac{1}{2}}$ por lo que un aproximador lineal a: $\sqrt{1+x^2}$ es $.955 + .414x$. De lo anterior si $x \geq y > 0$ entonces $z = \frac{y}{x} \in (0,1]$ por lo que $\sqrt{1+z^2} \approx .955 + .414z$ de donde $\sqrt{x^2+y^2} \approx .955x + .414y$ la cuál es conocida cómo la aproximación pitágorica ya que nos ayuda a conocer aproximadamente la distancia de un punto $P(x,y)$, con $y \leq x$, al origen.

(iii) Sean $w, f \in C([-1,1])$, $w(x) > 0$ fijas entonces existe un *único* $p \in \mathcal{P}_n$ tal que $\min\{||w(f-q)||_{[-1,1]}\} = ||w(f-p)||_{[-1,1]}$, aquí consideramos a $W = \mathcal{P}_n$, $|||f||| =$ $= ||wf||$, la cuál es fácil de verificar que es una norma en $C([-1,1])$ y aplicamos (III.4) par la existencia y la unicidad se sigue de (III.34).

CAPITULO V

POLINOMIOS TRIGONOMETRICOS

En este capítulo vamos a generalizar algunos resultados de los capítulos anteriores a polinomios trigonométricos, de los cuáles ya se habló un poco al final del capítulo anterior.

§1. Teorema Trigonométrico de Weierstrass

Para empezar, daremos la versión trigonométrica del Teorema de Weierstrass (I.1), para lo cuál necesitaremos algunos resultados previos y recordemos que el teorema (IV.13) nos garantiza que dada $f \in PC([0,2\pi])$ existe $T \in \mathcal{T}$ par tal que $||f-T||_{[-\pi,\pi]} < \varepsilon$ para $\varepsilon > 0$ dada.

(V.1) Definición. Denotaremos por $\mathcal{T}_n$ a la familia de todos los polinomios trigonométricos de grado$\leq n$

(V.2) Lema. *Sean $T \in \mathcal{T}_n$ y $S \in \mathcal{T}_m$ entonces:*

(i) $T+S \in \mathcal{T}_N$ con $N \leq \max\{n,m\}$

(ii) $TS \in \mathcal{T}_N$ con $N = nm$.

(iii) Dada $\theta \in \mathbb{R}$ entonces $T(x+\theta) \in \mathcal{T}_n$.

Demostración.

(i) Sean $T(x) = \alpha_0 + \sum_{j=1}^{n} \alpha_j \cos(jx) + \beta_j \text{sen}(jx)$, $\alpha_n^2 + \beta_n^2 > 0$ y $S(x) = \alpha_0' +$

$$+\sum_{k=1}^{m} \alpha_k' \cos(kx) + \beta_k' \text{sen}(kx),\ (\alpha_m')^2 + (\beta_m')^2 > 0.$$

Supongamos que $n > m$, entonces $T(x)+S(x) = \alpha_0'' + \sum_{k=1}^{n} \alpha_k'' \cos(kx) + \beta_k'' \text{sen}(kx)$ con:

$\alpha_0'' = \alpha_0 + \alpha_0'$

$\alpha_k'' = \alpha_k + \alpha_k'$ si $k = 1,\ldots,m$ y $\alpha_k'' = \alpha_k$ si $k = m+1,\ldots,n$ y

$\beta_k'' = \beta_k + \beta_k'$ si $k = 1,\ldots,m$ y $\beta_k'' = \beta_k$ si $k = m+1,\ldots,n$

y como $\alpha_n^2 + \beta_n^2 > 0$ tenemos que $(\alpha_n'')^2 + (\beta_n'')^2 > 0$ de donde se tiene que $T + S \in \mathcal{T}_N$ con $N = n$. Si $n = m$ entonces es claro que $T + S \in \mathcal{P}_n$.

(ii) Para el producto TS, observemos que nos quedan términos de la siguiente forma:

$\cos(kx)\cos(jx), \cos(kx)\text{sen}(jx), \text{sen}(kx)\cos(jx)$ ó $\text{sen}(kx)\text{sen}(jx)$. Pero por las siguientes identidades:

$$\text{sen}(kx)\cos(jx) = \tfrac{1}{2}\{\text{sen}(k+j)x + \text{sen}(k-j)x\}$$
$$\cos(kx)\cos(jx) = \tfrac{1}{2}\{\cos(k+j)x + \cos(k-j)x\}$$
$$\text{sen}(kx)\text{sen}(jx) = \tfrac{1}{2}\{\cos(k-j)x - \cos(k+j)x\}$$

TS puede escribirse de la forma siguiente:

$\alpha_0'' + \sum_{k=1}^{nm} \alpha_k'' \cos(kx) + \beta_k''\text{sen}(kx) \in \mathcal{T}_{nm}$ con $(\alpha_n'')^2 + (\beta_n'')^2 > 0$ dado que $\alpha_n^2 + \beta_n^2 > 0$ y $(\alpha_n')^2 + (\beta_n')^2 > 0$.

(iii) $T(x+\theta) = \alpha_0 + \sum_{j=1}^{n} \alpha_j \cos(jx + j\theta) + \beta_j\text{sen}(jx + j\theta)$ con $\alpha_n^2 + \beta_n^2 > 0$, desarrollando se tiene que:

$$T(x+\theta) = \alpha_0 + \sum_{j=1}^{n}(\alpha_j \cos(j\theta) + \beta_j\text{sen}(j\theta))\cos(jx) + (\beta_j \cos(j\theta) - \alpha_j\text{sen}(j\theta))\text{sen}(jx)$$

entonces $A_n = \alpha_n \cos(n\theta) + \beta_n\text{sen}(n\theta)$ mientras que $B_n = \beta_n \cos(n\theta) - \alpha_n\text{sen}(n\theta)$ de donde:

$$\begin{pmatrix} A_n \\ B_n \end{pmatrix} = \begin{pmatrix} \cos(n\theta) & \text{sen}(n\theta) \\ -\text{sen}(n\theta) & \cos(n\theta) \end{pmatrix} = M \begin{pmatrix} \alpha_n \\ \beta_n \end{pmatrix}$$

entonces tenemos que $det|M| = 1$ por lo que $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(\overline{x}) = M\overline{x}$ es inyectiva, por lo tanto $(A_n, B_n) \neq \overline{0}$ ya que $\alpha_n, \beta_n \neq 0$. Por lo que se cumple el resultado.

Q.E.D.

(V.3) Teorema trigonométrico de Weierstrass (1885).

Sea $f \in PC([0, 2\pi])$ entonces dada $\varepsilon > 0$ existe $T = T(\varepsilon) \in \mathcal{T}$ tal que $||f - T||_{[-\pi,\pi]} < \varepsilon$.

Demostración.

Consideremos $F_1(x) = f(x) + f(-x)$ y $F_2(x) = (f(x) - f(-x))\text{sen}x$, es fácil ver que $F_1, F_2 \in PC([0, 2\pi])$ y son pares. Por (IV.13) existen $T_1, T_2 \in \mathcal{T}$ pares tales que $||F_1 - T_1||_{[-\pi,\pi]} < \frac{\varepsilon}{2}$ y $||F_2 - T_2||_{[-\pi,\pi]} < \frac{\varepsilon}{2}$.

Por otro lado tenemos que: $f(x)\text{sen}^2x = \frac{1}{2}\left\{F_1(x)\text{sen}^2x + F_2(x)\text{sen}x\right\}$ y si consideramos a: $T_3(x) = \frac{1}{2}\{T_1(x)\text{sen}^2x + T_2(x)\text{sen}x\} \in \mathcal{T}_n$, entonces $|f(x)\text{sen}^2x - T_3(x)| \leq$

$\leq \frac{1}{2}|F_1(x) - T_1(x)| + \frac{1}{2}|F_2(x) - T_2(x)| < \frac{\epsilon}{2}$ para toda $x \in [-\pi, \pi]$, de donde se tiene que $||f(x)\text{sen}^2 x - T_3||_{[-\pi,\pi]} < \frac{\epsilon}{2}$.

Si lo anterior lo hacemos para $f(x - \frac{\pi}{2})$ en lugar de $f(x)$ obtenemos que: $||f(x - \frac{\pi}{2})\text{sen}^2 x - T_3||_{[-\pi,\pi]} < \frac{\epsilon}{2}$ **(1)**, sustituyendo $y = x - \frac{\pi}{2}$ en T_3 tenemos que $T_4 = T_3(y + \frac{\pi}{2}) \in \mathcal{T}_n$. De la desigualdad **(1)** y el hecho de que $\text{sen}^2(x - \frac{\pi}{2}) = \cos^2 x$ tenemos que: $||f(y)\cos^2 y - T_4||_{[-\pi,\pi]} < \frac{\epsilon}{2}$ **(2)**. Sea $T = T_3 + T_4$ entonces por **(1)** y **(2)** se tiene que $||f - T||_{[-\pi,\pi]} < \epsilon$.

Q.E.D.

(V.4) Observación. El teorema anterior es válido para toda $f: \mathbb{R} \to \mathbb{R}$ continua y de período 2π ya que nos restringimos a un intervalo de longitud 2π, aplicamos (V.3) y luego extendemos periodicamente a todos los reales.

Vamos a dar la versión trigonométrica al Teorema de Walsh (II.22), para lo cuál necesitamos algunos resultados previos.

(V.5) Lema.
Sea $T \in \mathcal{T}_n$ tal que T tiene (2n+1) raíces no congruentes (mod 2π) entonces $T \equiv 0$.

Demostración.

Sea $T(x) = \sum_{k=0}^{n} A_k \cos(kx) + B_k \text{sen}(kx)$, sabemos que $\cos(kx) = Re(e^{ikx}) = \frac{1}{2}\{e^{ikx} + e^{-ikx}\}$ y $\text{sen}(kx) = Im(e^{ikx}) = \frac{1}{2i}\{e^{ikx} - e^{-ikx}\}$ entonces tenemos que $T(x) =$

$$= \sum_{k=0}^{n} \left(\frac{A_k - iB_k}{2}\right) e^{ikx} + \left(\frac{A_k + iB_k}{2}\right) e^{-ikx},$$ por lo tanto: $$T(x) = \sum_{k=-n}^{k=n} c_k e^{ikx}$$

donde $c_0 = A_0$, $c_k = \frac{A_k - iB_k}{2}$ y $c_{-k} = \frac{A_k + iB_k}{2}$ **(forma compleja de un polinomio trigonométrico)**.

Así tenemos que: $T(x)e^{inx} = \sum_{j=0}^{2n} d_j e^{ijx}$ con $j = k + n$ y donde $d_j = c_{j-n}$, haciendo $z = e^{ix}$ en la última suma tenemos que: $p(z) = T(x)z^n = \sum_{j=0}^{2n} d_j z^j \in \mathcal{P}_{2n}[\mathbb{C}]$, pero por hipótesis T se anula en (2n+1) puntos por lo que $p \equiv 0$ y cómo $z^n \neq 0$ se tiene que $T \equiv 0$.

Q.E.D.

(V.6) Corolario. *Todo $T \in \mathcal{T}_n$, $T \not\equiv 0$ tiene 2n raíces contando su multiplicidad en $[-\pi, \pi)$.*

(V.7) Discusión.

Sean $x_0, \ldots, x_{2n}$ (2n+1) puntos en $\mathbb{R}$ no congruentes (mod 2π) y $w_0, \ldots, w_{2n}$ puntos en $\mathbb{R}$. Deseamos $T \in \mathcal{T}$ tal que $T(x_i) = w_i$ para toda $i = 0, \ldots, 2n$. Primero hallamos $t_k \in \mathcal{T}_n$ tal que $t_k(x_j) = \delta_{jk}$.

Sea:

$$t_k(x) = \frac{\prod_{j \neq k} \operatorname{sen}\left(\frac{x - x_j}{2}\right)}{\prod_{j \neq k} \operatorname{sen}\left(\frac{x_k - x_j}{2}\right)}$$

el cuál es llamado **polinomio trigonométrico fundamental de Lagrange**, es claro que $t_k(x_j) = \delta_{jk}$. Nos resta ver que $t_k(x) \in \mathcal{T}_n$ para lo cuál observemos que $t_k(x)$ tiene la siguiente forma $\frac{1}{a} \prod_{2n\text{términos}} \operatorname{sen}\left(\frac{x - x_j}{2}\right)$ con $a \in \mathbb{R}$ **(1)**.

Pero $\operatorname{sen}\left(\frac{x-a}{2}\right)\operatorname{sen}\left(\frac{x-b}{2}\right)$ es un polinomio trigonométrico en $\mathcal{T}_1$ debido a que: $\operatorname{sen} A \operatorname{sen} B = \frac{1}{2}\{\cos(A - B) + \cos(A + B)\}$. Ahora como en **(1)** hay 2n factores tenemos n pares y por (V.2) tenemos que $t_k \in \mathcal{T}_n$. Por lo que definimos:

$$T(x) = \sum_{k=0}^{2n} w_k t_k(x).$$

Es claro que $T(x_k) = w_k$ y T es único por (V.7).

Enunciaremos el siguiente:

(V.8) Teorema trigonométrico Interpolador de Walsh.

Sean $f \in PC([0, 2\pi])$ $\varepsilon > 0$ y $x_0, \ldots, x_{2n} \in \mathbb{R}$ (2n+1) puntos incongruentes (mod 2π) entonces existe un polinomio $T = T(\varepsilon) \in \mathcal{T}$ tal que $\|f - T\|_{\mathbb{R}} < (1 + N)\varepsilon$ con N una constante que depende sólo de $[0, 2\pi]$ y $x_0, \ldots, x_{2n}$ y $f(x_k) = T(x_k)$ para toda $k = 0, \ldots, 2n$.

Demostración.

Sea $\varepsilon > 0$ por (V.4) existe $T_1 = T_1(\varepsilon) \in \mathcal{T}$ tal que $\|f - T_1\|_{[0,2\pi]} < \varepsilon$. Consideremos el polinomio trigonométrico $\widehat{T}(x) = \sum_{k=0}^{2n} (f(x_k) - T_1(x_k)) t_k(x)$, como $\widehat{T}(x_k) = f(x_k) - T_1(x_k)$ para toda $k \in \{0, \ldots, 2n\}$ **(1)** y además tenemos que $\|\widehat{T}\|_{[0,2\pi]} \leq \sum_{k=0}^{2n} |f(x_k) - T_1(x_k)| \|t_k\|_{[0,2\pi]}$ entonces $\|\widehat{T}\|_{[0,2\pi]} < N\varepsilon$ donde $N =$

$= \sum_{k=0}^{2n} ||t_k||_{[0,2\pi]}$ **(2)** y donde N sólo depende de $[0,2\pi]$ y $x_0,\ldots,x_{2n}$.

Sea $T(x) = T_1(x) + \widehat{T}(x)$ por **(1)** $T(x_k) = f(x_k)$ para toda $k \in \{0,\ldots,2n\}$, mientras que por **(2)** y la elección de T_1 tenemos: $||f - T||_{[-0,2\pi]} \leq ||f - T_1||_{[0,2\pi]}$ $+||\widehat{T}||_{[0,2\pi]} < (N+1)\varepsilon$.

Q.E.D.

(V.9) Observación. Revisando la demostración de (V.8) se observa que ésta sigue siendo válida para toda función $f:[0,2\pi] \to \mathbb{R}$ que se puede aproximar uniformemente por polinomios trigonométricos en $[0,2\pi]$.

§2. Teorema de Equioscilación Trigonométrico de Chebyshev

En esta sección revisaremos algunos resultados del capítulo III que podemos extender a polinomios trigonométricos.

Para empezar recordemos que por (III.8), dada función $f \in PC([-\pi,\pi])$ existe $T^* \in \mathcal{T}_n$, el cuál es "un" mejor aproximador trigonométrico uniforme de grado $\leq n$, en otras palabras, $\min_{T\in\mathcal{T}_n}\{\|f-T\|_{[-\pi,\pi]}\} = \|f-T^*\|_{[-\pi,\pi]} = \widehat{E}_n(f)$.

Vamos a caracterizar a un mejor aproximador trigonométrico T_n^* y veremos que es único cómo se hizo en el capítulo anterior, para lo cuál introducimos la siguiente:

(V.10) Definición. Sean $f \in PC([-\pi,\pi])$ y $T^* \in \mathcal{T}_n$ un mejor aproximador trigonométrico uniforme a f entonces definimos $\widehat{e}_n(f) = f - T_n^*$. Observemos que $\|\widehat{e}_n(f)\|_{[0,2\pi]} = \widehat{E}_n(f)$.

Enunciaremos el resultado análogo a (III.27) y cuya demostración es la misma.

(V.11) Lema. *Dada $f \in PC([-\pi,\pi])$ y $T^* \in \mathcal{T}_n$ un mejor aproximador trigonométrico uniforme a f entonces existen $x_1, x_2 \in [-\pi,\pi]$ puntos distintos tales que:*

(i) $\widehat{E}_n(f) = |\widehat{e}_n(x_i)|$ $i = 1,2$

(ii) $\widehat{e}_n(x_1) = -\widehat{e}_n(x_2)$.

Ahora probaremos el siguiente:

(V.12) Teorema Trigonométrico de Equioscilación (Chebyshev 1859). *Sea $f \in PC([-\pi,\pi])$ entonces $T_n^* \in \mathcal{T}_n$ es un mejor aproximador uniforme (de grado n) para f si y sólo si existe un m-conjunto alternante con $m \geq 2n+1$ puntos para $\widehat{e}_n(f)$.*

Demostración. Sea $T_n^* \in \mathcal{T}_n$ un mejor aproximador uniforme para f, como $\widehat{e}_n$ es uniformemente continua dada $\varepsilon = \frac{\widehat{E}_n}{2}$ existe $\delta(\varepsilon) > 0$ tal que $|\widehat{e}_n(x) - \widehat{e}_n(y)| < \varepsilon$, si $|x-y| < \delta$.

Dividimos al intervalo $[-\pi,\pi]$ en subintervalos cerrados de longitud menor ó igual que δ y enumeramos a los intervalos en los cuales $|\widehat{e}_n(x)|$ alcanza su máximo valor como sigue: $I_1,\ldots,I_m$. Dado que $\widehat{e}_n(x)$ puede variar a lo más $\frac{\widehat{E}_n}{2}$ en cualquiera de esos intervalos, se tiene que $\widehat{e}_n(x) > \varepsilon$ ó $\widehat{e}_n(x) < -\varepsilon$ **(1)**. Sea $u_1,\ldots,u_m (=\pm 1)$ el signo de $\widehat{e}_n(x)$ sobre esos intervalos, lo que deseamos demostrar es que esta sucesión tenga al menos 2n+1 cambios de signo, para lo cuál lo ***probaremos por reducción al absurdo***, es decir: Si existen menos de 2n+1 cambios de signo encontraremos un polinomio trigonométrico cuyo error $\widehat{E}_n'$ es menor que el de T_n^*.

Por el lema (V.11) sabemos que debe de haber cambios de signo a través de los intervalos $I_1, \ldots, I_m$, por lo cuál procedemos a clasificarlos por grupos consecutivos donde los signos sean iguales, para fijar ideas supondremos que el primer grupo tiene signo positivo:

Primer grupo I^1,	$\ldots, I^{j_1}$	signo	$(+)$
Segundo grupo I^{j_1+1},	$\ldots, I^{j_2}$	signo	$(-)$
Tercer grupo I^{j_2+1},	$\ldots, I^{j_3}$	signo	$(+)$
K-ésimo grupo $I^{j_{k-1}+1}$,	$\ldots, I^{j_k}$	signo	$(-)^{k-1}$

donde $I^{(1)} = I_1, I^{(j_m)} = I_m$. El ordenamiento nos muestra $(k-1)$ cambios de signo. *Supongamos* que $k - 1 < 2n + 1$, es decir, $k < 2n + 2$. Consideremos $I^{(j_i)}$ e $I^{(j_i+1)}$, estos intervalos cerrados no pueden ser adyacentes por **(1)**, entonces existe $x_i \in [-\pi, \pi]$ tal que $I^{(j_i)} < x_i < I^{(j_i+1)}$ para toda $i = 1, \ldots k-1$. Definimos $T(x) = \prod_{i=1}^{k-1} \operatorname{sen}(\frac{x_i - x}{2})$, dado que $k < 2n + 2$ entonces $k - 1 \leq 2n$. Si $k = 2s + 1$ entonces $k - 1 = 2s \leq 2n$ **(2)**, de donde podemos formar s pares de la forma $\operatorname{sen}(\frac{x-a}{2})\operatorname{sen}(\frac{x-b}{2}) \in \mathcal{T}_1$ por lo que de **(2)** se tiene que $T \in \mathcal{T}_s \subset \mathcal{T}_n$.

Si $k = 2s$ tenemos que $2s \leq 2n + 1$ de donde $2s - 1 \leq 2n$ y $2s \leq 2n$, pero aquí hay un problema ya que $T(x)$ tendría un número impar de factores de la forma $\operatorname{sen}(\frac{x_i - x}{2})$ y nosotros necesitamos un número par para poder garantizar que $T \in \mathcal{T}_s$, para lo cuál consideramos a $T(x) = \operatorname{sen}(\frac{2\pi - x}{2}) \prod_{i=1}^{k-1} \operatorname{sen}(\frac{x_i - x}{2})$ de donde T tiene k factores y sabemos que $k = 2s \leq 2n$ por lo que $T \in \mathcal{T}_n$ y además $\operatorname{sen}(\frac{2\pi - x}{2}) \geq 0$ en $x \in [-\pi, \pi)$.

De este modo en ambos casos cuando k es par o impar $T \in \mathcal{T}_n$. Así mismo T se anula sólo en los puntos x_i y como cada x_i esta entre los intervalos $I_1, \ldots, I_m$ entonces $T(x)$ debe tener signo constante en cada intervalo. Consideramos $uT(x)$ con $u = \pm 1$ para que $uT(x)$ coincida en signo con $\widehat{e}(x)$ sobre los intervalos $I_1, \ldots, I_m$.

Sean $J = \overline{[a,b] \setminus I_1 \setminus \ldots \setminus I_m}$ y $\widehat{E}'_n = \|\widehat{e}_n\|_J$, cómo el máximo de $\widehat{e}_n$ sólo se alcanza en $I_1, \ldots, I_m$ entonces $\widehat{E}'_n < \widehat{E}_n$, sea $c \in \mathbb{R}$ tal que $0 < c < \frac{\widehat{E}_n - \widehat{E}'_n}{2\|uT\|_{[-\pi,\pi]}}$ **(3)**, y consideremos $T_n = T_n^* + cuT \in \mathcal{T}_n$. Si $x \in J$ entonces por **(1)** y **(3)**: $\|\widehat{e}_n - cuT\|_J \leq \|\widehat{e}_n\|_J + c\|uT\|_J < \widehat{E}'_n + \frac{\widehat{E}_n - \widehat{E}'_n}{2} < \frac{\widehat{E}_n}{2} < \widehat{E}_n(x)$.

Si $x \in I_j$ para alguna $j = 1, \ldots, m$ y $uT(x) > 0$ entonces por **(1)** y **(3)** se tiene que:

$0 < cuT(x) < \frac{\widehat{E}_n - \widehat{E}'_n}{2\|uT\|_{[-\pi,\pi]}} uT(x) < \frac{\widehat{E}_n - \widehat{E}'_n}{2} < \widehat{e}_n(x)$ por lo tanto tenemos que

$|\widehat{e}_n(x) - cuT(x)| = \widehat{e}_n(x) - cT(x) \leq \widehat{E}_n - c \min_{x \in I_j} uT(x) < \widehat{E}_n$, para el caso en que $uT(x) < 0$ procedemos de la siguiente manera: $\frac{uT(x)}{\|uT\|_{[-\pi,\pi]}} \geq -1$ por **(1)** y **(3)**

$cuT(x) > \frac{\widehat{E}_n - \widehat{E}'_n}{2\|uT\|_{[-\pi,\pi]}} uT(x) \geq \frac{\widehat{E}'_n - \widehat{E}_n}{2} \geq -\frac{\widehat{E}_n}{2} > \widehat{e}_n(x)$, de donde:

$|\widehat{e}_n(x) - cuT(x)| = cuT(x) - \widehat{e}_n(x) < \dot{E}_n + c \max_{x \in I_j} uT(x) < \widehat{E}_n$ ya que $uT(x) < 0$.

Por lo anterior $\|f - T_n\| < \widehat{E}_n$ lo cuál no es posible.

Para el recíproco, sea $-\pi \leq x_0 < x_1 < \ldots < x_{2n} \leq \pi$ un $(2n + 2)$-conjunto alternante para $f - T_n^*$ y $E = \|f - T_n^*\|$ entonces $\widehat{E}_n(f) \leq E$. Supongamos que $\widehat{E}_n(f) < E$, por (III.8) sea T_n un mejor aproximador trigonométrico de grado $\leq n$ para f, por lo que $\|f - T_n\|_{[-\pi,\pi]} = \widehat{E}_n(f) < E$ **(4)**.

Por otro lado $T_n^*(x_i) - T(x_i) = T_n^*(x_i) - f(x_i) - (T_n(x_i) - f(x_i))$ haciendo $t_i = T_n^*(x_i) - f(x_i)$ de **(4)** se tiene que $|t_i| < E$ pero $T_n^*(x_i) - T(x_i) = \pm E - t_i$ alternando de signo (2n+2) veces por lo que $T_n^* - T \in \mathcal{T}_n$ tiene (2n+1) raíces por (V.5) se tiene que $T_n^* = T$.

Q.E.D.

(V.13) Teorema. *Dada $f \in PC([0, 2\pi])$ el mejor aproximador uniforme $T_n^* \in \mathcal{T}_n$ es único.*

Demostración.

Sean T_n^* y T_n dos mejores aproximadores de grado$\leq n$ para f entonces $\widehat{E}_n(f) =$ $= \|f - T_n\| = \|f - T_n^*\|$ por (III.16) tenemos que $\frac{1}{2}(T_n + T_n^*)$ es un mejor aproximador de grado $\leq n$ para f, de donde por (V.12) se sigue que existen $x_1, \ldots, x_{2n+2}$ (2n+2) puntos en $[-\pi, \pi]$ tales que $\widehat{E}_n(f) = |f(x_i) - \frac{1}{2}(T_n(x_i) + T_n^*(x_i))|$ para toda $i = 1, \ldots, 2n + 2$. Por lo que se tiene que:

$$\widehat{E}_n(f) = \left|\tfrac{1}{2}\left(f(x_i) - T_n^*(x_i)\right) + \tfrac{1}{2}\left(f(x_i) - T_n(x_i)\right)\right| \leq \tfrac{1}{2}|f(x_i) - T_n^*(x_i)| +$$
$$+ \tfrac{1}{2}|f(x_i) - T_n(x_i)| \leq \tfrac{1}{2}\widehat{E}_n(f) + \tfrac{1}{2}\widehat{E}_n(f) = \widehat{E}_n(f).$$

de donde se tiene que en las dos desigualdades anteriores necesariamente hay igualdad. Así pues:

(i) $|f(x_i) - T_n^*(x_i)| = |f(x_i) - T_n(x_i)| = \widehat{E}_n(f)$ para toda $i = 1, \ldots, 2n + 2$

(ii) $sign\,[f(x_i) - T_n^*(x_i)] = sign\,[f(x_i) - T_n(x_i)]$ para toda $i = 1, \ldots, 2n + 2$

por lo que $T_n^*(x_i) = T_n(x_i)$ para toda $i = 1, \ldots, 2n + 2$, por (V.5) se tiene que $T_n^* = T_n$.

Q.E.D.

(V.14) Corolario. *Sea $f \in PC([0, 2\pi])$. Si f tiene un conjunto alternante de (2n+2) puntos entonces $T_n^* \equiv 0$ y $\widehat{E}_k(f) = \|f\|$ para toda $k = 0, \ldots, n$.*

Para terminar con este capítulo probaremos un resultado análogo al del Teorema (IV.7)(iv)

(V.15) Lema. *Sea $T \in \mathcal{T}_n$ tal que $\|T\|_{[0,2\pi]} = 1$. Si existen $2n$ puntos distintos $\theta_1 < \ldots < \theta_{2n}$ en $[0, 2\pi)$ tales que $|T(\theta_i)| = 1$ entonces $T(\theta) = \cos(n\theta + \alpha)$ para alguna $\alpha \in \mathbb{R}$.*

Demostración. Sea $T \in \mathcal{T}_n$ entonces $T(\theta) = \alpha_0 + \sum_{k=1}^{n} \alpha_k \cos(k\theta) + \beta_k \mathrm{sen}(k\theta)$ con $\alpha_k^2 + \beta_k^2 > 0$ de donde al examinar al término de grado mayor $T'(\theta) =$
$= \sum_{k=1}^{n} -k\alpha_k \mathrm{sen}(k\theta) + k\beta_k \mathrm{sen}(k\theta)$, por hipótesis $|T(\theta_i)| = 1$ para toda $i = 1, \ldots, 2n$ y $1 - T^2(\theta_i) = 0$ por lo que $(T'(\theta))^2 = c(1 - T^2(\theta))$ de donde tenemos que: $(-n\alpha_n \cos(n\theta) + n\beta_n \mathrm{sen}(n\theta))^2 = c(\alpha_n \cos(n\theta) + \beta_n \mathrm{sen}(n\theta))^2$ por lo tanto $c = n^2$ tenemos $(T'(\theta))^2 = n^2(1 - T^2(\theta))$ entonces $\frac{(T'(\theta))^2}{1-T^2(\theta)} = n^2$ integrando se tiene que $-\arccos(T(\theta)) = n\theta + \alpha$ y entonces se tiene que $T(\theta) = \cos(n\theta + \alpha)$.

Q.E.D.

CAPITULO VI

TEOREMAS DE JACKSON

En este capítulo nos dedicaremos a estudiar la rapidez con la cuál los aproximadores considerados en los capítulos anteriores tienden a las funciones que aproximan. Introduciremos algunos conceptos.

§1. Módulo de Continuidad

(VI.1) Definición. Sea $\mathcal{I}$ un intervalo y $f\colon \mathcal{I} \to \mathbb{R}$. El *módulo de continuidad* de f en $\mathcal{I}$ evaluado en $\delta > 0$, denotado por $w(f,\mathcal{I},\delta)$ es igual a:

$$w(f,\mathcal{I},\delta) = \sup\{|f(x) - f(y)| \mid |x-y| \leq \delta, x, y \in \mathcal{I}\}.$$

Convenimos en definir $w(f,\mathcal{I},0) = 0$.

(VI.2) Observaciones.

(i) Observemos que el módulo de continuidad depende de la función f, de δ y del intervalo.

(ii) Utilizaremos $w(\delta)$ para denotar a $w(f,\mathcal{I},\delta)$ si f e $\mathcal{I}$ son conocidos.

(VI.3) Ejemplos.

(i) f es constante si y sólo si $w(\delta) = 0$.

(ii) Sea $f(x) = x$ entonces $w(f,\delta) = \delta$.

(iii) Sea $f(x) = \frac{1}{x}$ en $(0,1)$, entonces $w(\delta) = +\infty$.

(iv) Sea $f(x) = \mathrm{sen}(\frac{1}{x})$ en $(0,1)$, entonces $w(\delta) = 2$.

A continuación calcularemos algunos módulos de continuidad para algunas funciones:

(v) Sea $f \in C^2([a,b])$ tal que $f' > 0$ y es creciente, entonces f es creciente. Además $f'' > 0$, es decir, f es convexa en $[a,b]$. Sean $s,t \in [a,b]$ tales que $|s-t| \leq \delta$ si $\delta \geq b-a$ entonces $w(\delta) = f(b) - f(a)$, por lo que supondremos que $\delta < b-a$.

Caso 1. Sean $t < s$, $t,s \in [b-\delta,b]$ entonces como f es una función creciente $f(s) - f(t) \leq f(b) - f(b-\delta)$ de donde $|f(s) - f(t)| \leq f(b) - f(b-\delta)$.

Caso 2. Sean $t < s$, $t,s \in [a,b-\delta]$ por TVM-D existe $c \in (t,s)$ tal que

$f(s) - f(t) = f'(c)(s-t) \leq f'(c)\delta \leq f'(c')\delta$ donde por el TVM-D se tiene que $f'(c') = f(b) - f(b-\delta)$ para alguna $c' \in (b-\delta, b)$ y f' es creciente, por lo que $|f(s) - f(t)| = f(s) - f(t) \leq f(b) - f(b-\delta)$.

Caso 3. Sean $t \in [a, b-\delta]$ y $s \in [b-\delta, b]$ entonces:

$f(b-\delta) - f(t) = f'(c)(b-\delta-t) \leq f'(c)(b-s) \leq f'(c')(b-s) = f(b) - f(s)$ para alguna $c' \in (s, b)$, por lo que $|f(s) - f(t)| \leq f(b) - f(b-\delta)$. En conclusión tenemos que si $\delta \in (0, b-a)$ entonces $w(\delta) = f(b) - f(b-\delta)$.

De manera análoga se puede probar que:

(vi) Sea $f \in C^2([a,b])$ tal que $f' > 0$ y decreciente entonces $w(\delta) = f(a+\delta) - f(a)$.

(vii) Sea $f \in C^2([a,b])$ tal que $f' < 0$ y decreciente entonces $w(\delta) = f(b-\delta) - f(b)$.

(viii) Sea $f \in C^2([a,b])$ tal que $f' < 0$ y creciente entonces $w(\delta) = f(a) - f(a+\delta)$.

Ahora veremos algunas propiedades sobre el módulo de continuidad que nos serán útiles.

(VI.4) Teorema. *Sean $\mathcal{I} = [a,b] \subset \mathbb{R}$ un intervalo y $f: \mathcal{I} \to \mathbb{R}$ una función entonces dadas $\delta_1, \delta_2 > 0$ tenemos que:*

(i) $w(\delta_1) \geq 0$.

(ii) Si $\delta_1 < \delta_2$ entonces $w(\delta_1) \leq w(\delta_2)$.

(iii) $w(\delta_1 + \delta_2) \leq w(\delta_1) + w(\delta_2)$.

(iv) $w(n\delta) \leq n w(\delta)$ $n \in \mathbb{N}$.

(v) $w(\lambda\delta) \leq (\lambda + 1) w(\delta)$ $(\lambda \in \mathbb{R}^+)$.

(vi) $f: \mathcal{I} \to \mathbb{R}$ es uniformemente continua si y sólo si $w(\delta) \to 0$ cuando $\delta \to 0$.

(vii) Si $\frac{w(\delta)}{\delta} \to 0$ cuando $\delta \to 0^+$ entonces $f \equiv c$, con c una constante.

Demostración.

(i) Es claro a partir de la definición.

(ii) A partir de la definición de $w(\delta)$ y del hecho de que $|s-t| \leq \delta_1 < \delta_2$ se tiene el resultado.

(iii) Sean $s, t \in \mathcal{I}$ tales que $|s-t| \leq \delta_1 + \delta_2$. Si $|s-t| \leq \delta_1$ entonces por (i) $|f(s) - f(t)| \leq w(\delta_1) \leq w(\delta_1) + w(\delta_2)$ de donde $w(\delta_1 + \delta_2) \leq w(\delta_1) + w(\delta_2)$. Ahora supongamos que $\delta_1 < s - t \leq \delta_1 + \delta_2$ entonces $\delta_1 + t < s$ y $s - (t + \delta_1) \leq \delta_2$ de donde se tiene que $|f(s) - f(t)| \leq |f(s) - f(t+\delta_1)| + |f(t+\delta_1) - f(t)| \leq$

$\leq |f(s) - f(t+\delta_1)| + w(\delta_1) \leq w(\delta_1) + w(\delta_2)$.

(iv) La prueba se sigue por inducción.

(v) Sea $\lambda > 0$ hallamos $n \in \mathbb{N} \cup \{0\}$ tal que $n \leq \lambda < n+1$ entonces

$w(\lambda\delta) \leq w((n+1)\delta) \leq (n+1) w(\delta) \leq (\lambda+1) w(\delta)$.

(vi) Supongamos que $w(\delta) \to 0$, $\delta \to 0^+$ entonces dada $\varepsilon > 0$ existe $\eta(\varepsilon) = \eta$ tal que si $\delta > 0, \delta < \eta$ tenemos que $w(f,\delta) < \varepsilon$, proponemos $\delta(\varepsilon) = \eta$ entonces si $|s-t| < \delta$, $s,t \in \mathcal{I}$ tenemos que $|f(s) - f(t)| \leq w(f,\delta) < \varepsilon$ por lo que f es uniformemente continua en $\mathcal{I}$.

Ahora si f es uniformemente continua en $\mathcal{I}$ dada $\varepsilon > 0$ existe $\delta^*(\varepsilon) > 0$ tal que para toda $s,t \in \mathcal{I}$ tales que $|s-t| \leq \delta$ entonces $|f(s) - f(t)| < \varepsilon$. Sea $\eta = \delta^*$ por lo que si $\delta \leq \eta$ tenemos que $|f(s) - f(t)| < \varepsilon$ para toda $s,t \in \mathcal{I}$ con $|s-t| < \eta$ de donde $w(\delta) \leq \varepsilon$ y así tenemos que $w(\delta) \to 0$.

(vii) Sea $t \in [a,b]$, $t \neq a,b$ elegimos $0 < \delta < b-t$, entonces se tiene que

$|f(t+\delta) - f(t)| \leq w(\delta)$ de donde $|\frac{f(t+\delta)-f(t)}{\delta}| \leq \frac{w(\delta)}{\delta}$ por lo que $|f'(t)| = 0$ y así $f \equiv c$.

Q.E.D.

Ahora veremos con que rapidez se aproximan los polinomios de Bernstein a una función acotada $f\colon [0,1] \to \mathbb{R}$.

(VI.5) Teorema. *Sea $f\colon [0,1] \to \mathbb{R}$ acotada entonces $\|f - B_n(f)\| \leq \frac{3}{2} w(\frac{1}{\sqrt{n}})$.*

Demostración. Tenemos que:

$$|f(x) - B_n(f)(x)| = \left|\sum_{k=0}^{n} \left[f(x) - f\left(\frac{k}{n}\right)\right] \binom{n}{k} x^k (1-x)^{n-k}\right| \leq$$

$$\leq \sum_{k=0}^{n} \left|f(x) - f\left(\frac{k}{n}\right)\right| \binom{n}{k} x^k (1-x)^{n-k} \leq$$

$$\leq \sum_{k=0}^{n} w\left(\left|x - \frac{k}{n}\right|\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

Por otro lado tenemos por (VI.4)(v) que $w(|x - \frac{k}{n}|) = w(\sqrt{n}|x - \frac{k}{n}|\frac{1}{\sqrt{n}}) \leq$

$\leq \left(1 + \sqrt{n}\,|x - \frac{k}{n}|\right) w\left(\frac{1}{\sqrt{n}}\right)$

por lo que:

$$|f(x) - B_n(f)(x)| \leq \sum_{k=0}^{n} \left(1 + \sqrt{n}\left|x - \frac{k}{n}\right|\right) w\left(\frac{1}{\sqrt{n}}\right) \binom{n}{k} x^k (1-x)^{n-k} =$$

$$= w\left(\frac{1}{\sqrt{n}}\right) \left[1 + \sqrt{n} \sum_{k=0}^{n} \left|x - \frac{k}{n}\right| \binom{n}{k} x^k (1-x)^{n-k}\right] \quad \textbf{(1)}.$$

Por la desigualdad de Cauchy, **(CBS)**, tenemos que:

$$\sum_{k=0}^{n} \left|x - \frac{k}{n}\right| \binom{n}{k} x^k (1-x)^{n-k} =$$

$$= \sum_{k=0}^{n} \left(\left| x - \frac{k}{n} \right| \left(\binom{n}{k} x^k (1-x)^{n-k} \right)^{\frac{1}{2}} \right) \left(\binom{n}{k} x^k (1-x)^{n-k} \right)^{\frac{1}{2}} \leq$$

$$\leq \left\{ \sum_{k=0}^{n} \left| x - \frac{k}{n} \right|^2 \binom{n}{k} x^k (1-x)^{n-k} \right\}^{\frac{1}{2}}.$$

Pero por (I.15) Parte 1 se tiene que $\sum_{k=0}^{n} (k-nx)^2 \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{n}{4}$ de donde se sigue que:

$$\sum_{k=0}^{n} \left(x - \frac{k}{n} \right)^2 \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{1}{4n}.$$

De (**1**) se obtiene que $|f(x) - B_n(f)(x)| \leq \frac{3}{2} w(\frac{1}{\sqrt{n}})$.

Q.E.D.

(VI.6) Observaciones.

(i) Si $f \in C([0,1])$, por (VI.4)(vi) tenemos que $w(\frac{1}{\sqrt{n}}) \to 0$ cuando $n \to \infty$ por lo que obtenemos una prueba más para el Teorema polinomial de Weierstrass.

(ii) Si $f \in \mathrm{Lip}_M(\alpha)$, $\alpha \in (0,1]$ tenemos que $\|f - B_n(f)\| \leq \frac{3}{2} \frac{M}{n^{\frac{\alpha}{2}}}$.

Probaremos algunas otras propiedades del módulo de continuidad:

(VI.7) Teorema.

(i) *Si $f \in \mathrm{Lip}_M(\alpha)$, $\alpha \in (0,1]$. Entonces $w(\delta) \leq M\delta^\alpha$ en particular se tiene que si: $|f'(x)| \leq M$ en $[a,b]$ entonces $w(f,[a,b],\delta) \leq M\delta$.*

(ii) *Sea $f\colon [0,a] \to \mathbb{R}$ una función tal que:*

(a) $f(x) \leq f(y)$ si $x < y$, $x, y \in [0,a]$.

(b) $f(x+y) \leq f(x) + f(y)$.

(c) $f(0) = 0$

entonces $f(x) = w(f,x)$ en $[0,a]$.

(iii) *Si $[c,d] \subset [a,b]$ entonces $w(f,[c,d],\delta) \leq w(f,[a,b],\delta)$.*

(iv) *Si f es una función par en $[-a,a]$ entonces $w(f,[-a,a],) = w(f,[0,a],\delta)$.*

(v) *Si $x = \cos\theta$, $x \in [-1,1]$ y $g(\theta) = f(\cos\theta)$ entonces*

$w(g;[-\pi,\pi],\delta) = w(g;[0,\pi],\delta) \leq w(f,[-1,1],\delta)$.

(vi) *Si f tiene período T entonces $w(f,[a,a+T],\delta)$ es independiente de a.*

(vii) *Si $g(x) = f(Ax+B)$, $A > 0$ con $x \in [c,d]$ entonces*

$w(g, [c, d], \delta) = w(f, [Ac + B, Ad + B], A\delta)$.

Demostración.

(i) Es claro a partir de la definición.

(ii) Debemos probar que $f(\delta) = \sup\limits_{|x-y|\leq\delta} \{|f(x) - f(y)| \mid x, y \in [0, a]\}$ para $\delta \in [0, a]$.

Si hacemos $x = 0$ y $y = \delta$ entonces por hipótesis

$f(\delta) = |f(y) - f(x)| \leq \sup\limits_{x,y\in[0,a]} \{|f(x) - f(y)| \mid |x - y| \leq \delta\}$.

Ahora supongamos que $x, y \in [0, a]$ son tales que $x < y$ y $|x - y| \leq \delta$, es decir, $x < y \leq x + \delta$ entonces por hipótesis $f(x) \leq f(y)$ y $f(y) \leq f(y - x) + f(x)$ por lo que $0 \leq f(y) - f(x) \leq f(y - x) \leq f(\delta)$ de donde $|f(y) - f(x)| \leq f(\delta)$ con $|x - y| \leq \delta$ por lo que se tiene la otra desigualdad.

(iii) Del hecho de que si $x, y \in [c, d]$ tales que $|x - y| \leq \delta$ implica que $x, y \in [a, b]$ y $|x - y| \leq \delta$ se sigue que $w(f, [c, d], \delta) \leq w(f, [a, b], \delta)$.

(iv) Por el inciso anterior tenemos que $w(f, [0, a], \delta) \leq w(f, [-a, a], \delta)$. Para la otra desigualdad analizaremos diversos casos:

Caso 1. Si $s, t \in [0, a]$, $s < t$ entonces $|f(s) - f(t)| \leq w(f, [0, a], \delta)$ con $|s - t| \leq \delta$.

Caso 2. Si $s < t$, $s < 0$ y $t \geq 0$, $s, t \in [-a, a]$, $|s - t| \leq \delta$ entonces $|f(s) - f(t)| =$ $= |f(-s) - f(t)| \leq w(f, [0, a], \delta)$ ya que $|-s - t| \leq |s - t| \leq \delta$.

Caso 3. Si $s, t \in [-a, 0]$, $|s - t| \leq \delta$ entonces $|f(s) - f(t)| = |f(-s) - f(-t)| \leq$ $\leq w(f, [0, a], \delta)$ ya que $-s, -t \in [0, a]$ por lo que se cumple que

$w(f, [-a, a], \delta) \leq w(f, [0, a], \delta)$.

(v) Por el inciso anterior tenemos la primer igualdad, para la desigualdad procedemos como sigue: $w(g, [0, \pi], \delta) = \sup\limits_{|\theta_1-\theta_2|\leq\delta} \{|f(\cos\theta_1) - f(\cos\theta_2)|\}$ si hacemos $x_1 = \cos\theta_1$ y $x_2 = \cos\theta_2$, $|x_1 - x_2| \leq |\theta_1 - \theta_2| \leq \delta$ por el TVM-D, por lo que $w(g, [0, \pi], \delta) \leq w(f, [-1, 1], \delta)$.

(vi) Es clara despúes de observar que para cualesquiera dos intervalos de longitud T hay una isometría entre ellos y que la función f preserva los valores en puntos correspondientes.

(vii) Sean $s, t \in [c, d]$, $|s - t| \leq \delta$ entonces $|g(s) - g(t)| = |f(As + B) - f(At + B)|$, pero tenemos que $As + B, At + B \in [Ac + B, Ad + B]$ y $|As + B - (At + B)| \leq A\delta$ por lo que $w(g; [c, d], \delta) \leq w(f; [Ac + B, Ad + B], A\delta)$.

Para la otra desigualdad procedemos así: sean $s, t \in [Ac + B, Ad + B]$, $|s - t| \leq$ $\leq A\delta$ entonces $|f(s) - f(t)| = |g(\frac{s-B}{A}) - g(\frac{t-B}{A})| \leq w(g; [c, d], \delta)$ ya que del hecho que si $t \in [Ac + B, Ad + B]$ tenemos que $\frac{s-B}{A}, \frac{t-B}{A} \in [c, d]$ y $|\frac{s-B}{A} - \frac{t-B}{A}| \leq \delta$, por lo tanto se cumple la otra desigualdad.

Q.E.D.

(VI.8) Observaciones.

(i) El inciso (ii) de (VI.7) es válido también en cualquier intervalo $[b, c]$ simplemente definiendo a $\widehat{w}(x) = w(x + b)$ con $x \in [0, c - b]$ y con ayuda de (VI.7)(vii).

(ii) Una función $f(x)$ definida en $[0, a]$ es módulo de continuidad de una función en $[0, a]$ si y sólo si $f(x)$ satisface las condiciones de (VI.7)(ii).

§2. Teoremas Polinomiales de Jackson

Los Teoremas polinomiales de Jackson nos dan una respuesta sobre la conducta que tiene el error $E_n(f)$ cuando $n \to \infty$, para establecerlos necesitaremos algunos temas introductorios.

Dada $g \in PC([-\pi,\pi])$ le asociamos un polinomio $t_n \in \mathcal{T}_n$ dado por $t_n(g)(\theta) =$ $= \frac{a_0}{2} + \sum_{k=1}^{n}(a_k \cos(k\theta) + b_x \mathrm{sen}(k\theta))$ donde los coeficientes estan definidos por: $a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\cos(kt)dt$ con $k = 0,\ldots,n$ y $b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\mathrm{sen}(kt)dt$ con $k = 1,\ldots,n$ los cuáles son llamados los **coeficientes de Fourier clásicos de la función** g. Para nuestro propósito introduciremos una generalización dada por la siguiente:

(VI.9) Definición. Sea $f \in PC([0,2\pi])$ definimos:

$$q_n(f;x) = \frac{a_0}{2} + \sum_{k=1}^{n} r_{k,n}(a_k \cos(kx) + b_k \mathrm{sen}(kx))$$

donde a_0, a_k, b_k, $k = 1,\ldots,n$ son los coeficientes de Fourier de la función f y $r_{k,n} \in \mathbf{R}$ para $k = 1,\ldots,n$.

(VI.10) Lema. *Sea $f \in PC([-\pi,\pi])$ entonces*

$$q_n(f;x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)U_n(t)dt$$

donde $U_n(t) = \frac{1}{2} + \sum_{k=1}^{n} r_{k,n}\cos(kt)$.

Demostración. Por definición de $q_n(f;x)$ y de los coeficientes de Fourier y recordando que $\cos(A-B) = \cos(A)\cos(B) + \mathrm{sen}(A)\mathrm{sen}(B)$ tenemos que:

$$q_n(f;x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left\{\frac{1}{2} + \sum_{k=1}^{n} r_{k,n}\cos(k(t-x))\right\}dt =$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)U_n(t-x)dt.$$

haciendo el cambio de variable $y = t - x$ tenemos que:

$$q_n(f;x) = \frac{1}{\pi}\int_{-\pi-x}^{\pi-x} f(x+y)U_n(y)dy.$$

Por otro lado f y U_n son funciones de período 2π entonces

$$\tfrac{1}{\pi}\int_{-\pi-x}^{-\pi} f(x+y)U_n(y)dy = \frac{1}{\pi}\int_{\pi-x}^{\pi} f(x+y)U_n(y)dy \ \textbf{(1)}$$

por lo que $q_n(f;x) = \frac{1}{\pi}\displaystyle\int_{-\pi-x}^{-\pi} f(x+y)U_n(y)dy + \frac{1}{\pi}\int_{-\pi}^{\pi-x} f(x+y)U_n(y)dy$, por **(1)** tenemos que $q_n(f;x) = \frac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} f(x+y)U_n(y)dy$.

Q.E.D.

Los siguientes lemas nos ayudarán a rpobar el resultado (VI.13).

(VI.11) Lema. $\frac{2}{\pi}|x| \leq |\mathrm{sen}x|$ *si* $|x| \leq \frac{\pi}{2}$.

Demostración. Sea $x \in (0, \frac{\pi}{2}]$ y consideremos a la función $f(x) = \mathrm{sen}x$ entonces $f''(x) < 0$ en $[0, \frac{\pi}{2}]$, de donde f es concáva, por lo que $f(x) = \mathrm{sen}x \geq \frac{\pi}{2}x$ en $[0, \frac{\pi}{2}]$, ahora si $x \in [-\frac{\pi}{2}, 0]$ entonces $-x \in [0, \frac{\pi}{2}]$ por lo que $\mathrm{sen}(-x) \geq \frac{-2x}{\pi}$, es decir, $-\mathrm{sen}x \geq -\frac{2}{\pi}x$ por lo tanto se cumple la desigualdad del lema.

Q.E.D.

(VI.12) Lema. *Sea* $x \in [0, \infty)$ *entonces* $\mathrm{sen}x \leq x$.

Demostración. Consideremos a la función $h(x) = x - \mathrm{sen}x$ y apliquemos el TVM-D.

Q.E.D.

(VI.13) Teorema. P.P. Korovkin.

Sean $r_{1,n}, \ldots, r_{n,n} \in \mathbb{R}$ *tales que* $U_n(t) \geq 0$ *en* $[-\pi, \pi]$, *con* $U_n(t)$ *como en (VI.10). Entonces para cualquier* $m > 0$:

$$||f - q_n||_{[-\pi,\pi]} \leq w(\frac{1}{\sqrt{m}})\left[1 + \frac{m\pi}{\sqrt{2}}(1 - r_{1,n})^{\frac{1}{2}}\right].$$

con q_n *como en (VI.10).*

Demostración.

De la definición de $U_n(t)$, (VI.10), es fácil verificar que $\frac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} U_n(t)dt = 1$, por lo que tenemos:

$$|f(x) - q_n(f;x)| = \left|\tfrac{1}{\pi}\int_{-\pi}^{\pi}(f(x) - f(x+t))U_n(t)dt\right| \leq \tfrac{1}{\pi}\int_{-\pi}^{\pi} w(|t|)U_n(t)dt \ \textbf{(1)}.$$

Por otro lado se tiene que: $|f(x)-f(x+t)| \leq w(f,[-\pi+x,\pi+x],|t|)$ ya que $t\in[-\pi,\pi]$, por (VI.7)(vi) se tiene que $|f(x)-f(x+t)|\leq w(f,[-\pi,\pi],|t|)$.

Además por (VI.4)(v) se tiene que $w(|t|)=w(m|t|\frac{1}{m})\leq(1+m|t|)w(\frac{1}{m})$, así tenemos que $|f(x)-f(x+t)|\leq(1+m|t|)w(\frac{1}{m})$, regresando a **(1)** obtenemos:

$$|f(x)-q_n(f;x)|\leq w(\tfrac{1}{m})\left[1+\tfrac{m}{\pi}\int_{-\pi}^{\pi}|t|U_n(t)dt\right]\ \textbf{(2)}.$$

Ahora acotaremos la última integral. Por la desigualdad **CBS** tenemos:

$$\tfrac{1}{\pi}\int_{-\pi}^{\pi}|t|U_n(t)dt=\frac{1}{\pi}\int_{-\pi}^{\pi}|t|\,[U_n(t)]^{\frac{1}{2}}\,[U_n(t)]^{\frac{1}{2}}\,dt\leq\left[\frac{1}{\pi}\int_{-\pi}^{\pi}|t|^2U_n(t)dt\right]^{\frac{1}{2}}\ \textbf{(3)}.$$

Por (VI.11), como $|t|\leq\pi$ entonces $\left(\frac{t}{2}\right)^2\leq\frac{\pi^2}{4}\mathrm{sen}^2(\frac{t}{2})=\frac{1}{8}\pi^2(1-\cos t)$ de aquí se sigue que $t^2\leq\frac{\pi^2}{2}(1-\cos t)$, de **(3)** se tiene:

$$\tfrac{1}{\pi}\int_{-\pi}^{\pi}|t|U_n(t)dt\leq\left[\frac{\pi}{2}\int_{-\pi}^{\pi}U_n(t)(1-\cos t)dt\right]^{\frac{1}{2}}$$

pero $\int_{-\pi}^{\pi}\cos(kt)\cos(nt)dt=0$ si $k\neq n$, $k,n\in\mathbb{Z}$ y $\int_{-\pi}^{\pi}\cos^2(kt)dt=\pi$, de donde:

$\left[\frac{\pi}{2}\int_{-\pi}^{\pi}U_n(t)(1-\cos t)dt\right]^{\frac{1}{2}}=\frac{\pi}{\sqrt{2}}(1-r_{1,n})^{\frac{1}{2}}$, lo cuál al sustituirlo en **(2)** nos da el resultado.

Q.E.D.

El Teorema anterior presupone la existencia de números $r_{1,n},\ldots,r_{n,n}\in\mathbb{R}$ tales que $U_n(t)\geq 0$. Una forma de producir polinomios trigonométricos pares no negativos esta contenida en la siguiente:

(VI.14) Discusión.

Sabemos que:

$$0\leq\left|\sum_{k=0}^{n}c_ke^{ik\theta}\right|^2=\left(\sum_{k=0}^{n}c_ke^{ik\theta}\right)\left(\sum_{k=0}^{n}c_ke^{-ik\theta}\right)$$

para cada $c_k\in\mathbb{R}$ $k=0,\ldots,n$. Así tenemos que:

$$\left|\sum_{k=0}^{n}c_ke^{ik\theta}\right|^2=\sum_{k=0}^{n}c_k^2+\left(2\sum_{k=0}^{n-1}c_kc_{k+1}\right)\cos(\theta)+$$

$$+\left(2\sum_{k=0}^{n-2}c_kc_{k+2}\right)\cos(2\theta)+\ldots+2c_0c_n\cos(n\theta)\ \textbf{(1)}.$$

Si deseamos que **(1)** sea una $U_n(t)$ como en (VI.10), debemos tener:

$\sum_{k=0}^{n} c_k^2 = \frac{1}{2}$ para lo cuál elegimos a $c_k = c\,\mathrm{sen}\left(\frac{k+1}{n+2}\pi\right)$ para cada $k = 0, \ldots, n$ donde $c^2 = \left(2\sum_{k=0}^{n} \mathrm{sen}^2\left(\frac{k+1}{n+2}\pi\right)\right)^{-1}$ (2), y así tenemos que $\sum_{k=0}^{n} c_k^2 = \frac{1}{2}$ por lo que obtenemos que $U_n(t) \geq 0$.

Para aplicar el teorema (VI.13) encontraremos una expresión más simple para:

$$r_{1,n} = 2c^2 \sum_{k=0}^{n-1} \left(\mathrm{sen}\left(\frac{k+1}{n+2}\pi\right)\right)\left(\mathrm{sen}\left(\frac{k+2}{n+2}\pi\right)\right)$$

observemos que:

$$\sum_{k=0}^{n-1} \left(\mathrm{sen}\left(\frac{k+1}{n+2}\pi\right)\right)\left(\mathrm{sen}\left(\frac{k+2}{n+2}\pi\right)\right) = \sum_{k=0}^{n} \left(\mathrm{sen}\left(\frac{k+1}{n+2}\pi\right)\right)\left(\mathrm{sen}\left(\frac{k+2}{n+2}\pi\right)\right)$$

ya que sólo agregamos un cero a la última suma. Mientras que por otro lado tenemos:

$$\sum_{k=0}^{n-1} \left(\mathrm{sen}\left(\frac{k+1}{n+2}\pi\right)\right)\left(\mathrm{sen}\left(\frac{k+2}{n+2}\pi\right)\right) = \sum_{k=0}^{n} \left(\mathrm{sen}\left(\frac{k}{n+2}\pi\right)\right)\left(\mathrm{sen}\left(\frac{k+1}{n+2}\pi\right)\right)$$

por un simple corrimiento de índices y agregar un cero al principio, por lo anterior se tiene que:

$$\sum_{k=0}^{n-1} \left(\mathrm{sen}\left(\frac{k+1}{n+2}\pi\right)\right)\left(\mathrm{sen}\left(\frac{k+2}{n+2}\pi\right)\right) =$$

$$= \frac{1}{2}\sum_{k=0}^{n} \left\{\mathrm{sen}\left(\frac{k}{n+2}\pi\right) + \mathrm{sen}\left(\frac{k+2}{n+2}\pi\right)\right\} \mathrm{sen}\left(\frac{k+1}{n+2}\pi\right) \text{ (3).}$$

Aquí necesitamos recordar que: $\mathrm{sen}A\cos B = \frac{1}{2}\{\mathrm{sen}(A+B) + \mathrm{sen}(A-B)\}$, para nuestro caso partícular $A + B = \frac{k}{n+2}\pi$ y $A - B = \frac{k+2}{n+2}\pi$, de donde $A = \frac{k+1}{n+2}\pi$ y $B = -\frac{\pi}{n+2}$. Por lo tanto de (3) se tiene que:

$$\sum_{k=0}^{n-1} \left(\mathrm{sen}\left(\frac{k+1}{n+2}\pi\right)\right)\left(\mathrm{sen}\left(\frac{k+2}{n+2}\pi\right)\right) =$$

$$= \cos\left(\tfrac{\pi}{n+2}\right)\sum_{k=0}^{n} \mathrm{sen}^2\left(\frac{k+1}{n+2}\pi\right) = \frac{1}{2c^2}\cos\left(\frac{\pi}{n+2}\right).$$

Por lo tanto $r_{1,n} = \cos\left(\frac{\pi}{n+2}\right)$. Por (VI.12) y la identidad $\mathrm{sen}^2(\theta) = \frac{1}{2}(1 - \cos(2\theta))$, se sigue que: $(1 - r_{1,n})^{\frac{1}{2}} = \left(1 - \cos\left(\frac{\pi}{n+2}\right)\right)^{\frac{1}{2}} = \sqrt{2}\mathrm{sen}\left(\frac{\pi}{2(n+2)}\right) \leq \frac{\sqrt{2}\pi}{2(n+2)}$, de donde se tiene que: $1 + \frac{n\pi}{\sqrt{2}}(1 - r_{1,n})^{\frac{1}{2}} \leq 1 + \frac{n\pi^2}{2(n+2)} < 6$.

De (VI.13) concluimos que $|g(x) - q_n(g;x)| \leq 6w(\frac{1}{n})$. Por (III.8) sabemos que existe $T_n^* \in \mathcal{T}$ tal que $\widehat{E}_n(f) = ||f - T_n^*||$ por lo que hemos probado el siguiente:

(VI.15) Teorema de Jackson Trigonométrico. *Sea $f \in PC([-\pi,\pi])$ entonces $||f - T_n^*|| \leq 6w(\frac{1}{n})$.*

Ahora supongamos que $f \in PC([-1,1])$, entonces definimos $g \in C([0,\pi])$ por $g(t) = f(\cos t)$ la extendemos, a $[-\pi,\pi]$ poniendo $g(-t) = g(t)$ si $t \in [-\pi, 0)$ y así obtenemos una función par, continua y de período 2π, por lo que si q_n^* es el mejor aproximador trigonométrico de grado$\leq$n de g, entonces q_n^* es par también (ver (IV.9)). Así pues $q_n^*(t) = \frac{a_0}{2} + a_1\cos(t) + \ldots + a_n\cos(nt)$. Dado que $\cos(nt)$ es un polinomio en $\cos(t)$ de grado $\leq$n, tenemos que $q_n^*(t) = d_0 + d_1\cos^2(t) + \ldots + d_n\cos^n(t)$ haciendo $x = \cos(t)$ se tiene que $p_n^*(x) = d_0 + d_1x + \ldots + d_nx^n \in \mathcal{P}_n$.

Por (VI.7)(v) y (VI.15) hemos probado el siguiente:

(VI.16) Teorema de Jackson Polinomial. *Sea $f \in C([-1,1])$ entonces $E_n(f,[-1,1]) \leq 6w(\frac{1}{n})$.*

(VI.17) Corolario. *Sea $f \in C([a,b])$ entonces $E_n(f,[a,b]) \leq 6w(\frac{b-a}{2n})$.*

Demostración. Sea $\ell: [-1,1] \to [a,b]$ definida por $\ell(x) = \frac{b-a}{2}(x-1) + b$, consideramos $g(x) = (f \circ \ell)(x)$, por (VI.16) tenemos $E_n(g,[-1,1]) \leq 6w(\frac{1}{n})$ y como ℓ es una biyección entonces $E_n(f,[a,b]) \leq 6w(\frac{1}{n})$ y por (VI.7)(vii) tenemos $E_n(f,[a,b]) \leq 6w(\frac{b-a}{2n})$.

Q.E.D.

(VI.18) Corolario. *Si $f \in \mathrm{Lip}_k(\alpha)$ en el intervalo $[a,b]$ entonces tenemos que: $E_n(f,[-1,1]) \leq 6k\left(\frac{b-a}{2n}\right)^\alpha$.*

Demostración. El resultado se sigue de (VI.7)(i).

Q.E.D.

(VI.19) Corolario. *Si $|f'(x)| \leq M$ para toda $x \in [a,b]$ entonces:*

$E_n(f,[a,b]) \leq 3M\frac{(b-a)}{n}$.

Demostración. El resultado se sigue de (VI.7)(i).

Q.E.D.

(VI.20) Corolario. *Si $f' \in C([a,b])$, entonces se tiene que:*

$E_n(f;[a,b]) \leq \frac{3}{n}(b-a)E_{n-1}(f';[a,b])$.

Demostración. Sea $E_{n-1}(f';[a,b]) = ||f' - p_{n-1}^*||$, consideremos $p_n(x) = \int_0^x p_{n-1}^*(t)dt$ por (III.13)(iii) tenemos que $E_n(f;[a,b]) = E_n(f - p_n;[a,b])$ pero

$p'_n(x) = p^*_{n-1}(x)$ por lo que $|f' - p'| = |f' - p^*_{n-1}| \leq E_{n-1}(f') = M$, por (VI.19) se sigue el resultado.

Q.E.D.

(VI.21) Corolario. *Sea $f \in C^{(k)}([a,b])$ entonces si $n > k$ se tiene que $E_n(f,[a,b]) \leq \frac{c}{n^k} w_k\left(\frac{b-a}{2(n-k)}\right)$, donde w_k es el módulo de continuidad de $f^{(k)}$ y $c = 2 \cdot 3^{K+1}(b-a)^k e^k (1+k)^{-1}$.*

Demostración.

Por (VI.20) tenemos que $E_{n-j}(f^{(j)}) \leq 3(b-a)E_{n-j-1}(f^{(j+1)})(n-j)^{-1}$ **(1)**, entonces multiplicando obtenemos:

$E_n(f')E_{n-1}(f'') \cdots E_{n-k+1}(f^{(k-1)}) \leq$

$\leq \{3(b-a)\}^k \left[n(n-1)\cdots(n-k+1)\right]^{-1} E_{n-1}(f')E_{n-2}(f'') \cdots E_{n-k}(f^{(k)})$ **(2)**.

Si $E_{n-j}(f^{(j)}) = 0$ para alguna j entonces de (1) $E_n(f) = 0$ y el resultado es trivial. Por lo que supondremos lo contrario. De **(2)** se sigue:

$E_n(f) \leq \{3(b-a)\}^k \left[n(n-1)\cdots(n-k+1)\right]^{-1} E_{n-k}(f^{(k)})$, por (VI.17) tenemos que:

$E_n(f) \leq$

$6\{3(b-a)\}^k \left[n(n-1)\cdots(n-k+1)\right]^{-1} w_k\left(\frac{b-a}{2(n-k)}\right)$ **(3)**.

Por otro lado se tiene que:

$$\ln\left[n(n-1)\cdots(n-k+1)\right] = \sum_{j=0}^{k-1} \ln(n-j) \geq \int_{n-k}^{n} \ln(t)dt$$

pero es fácil verificar que:

$$\int_{n-k}^{n} \ln(t)dt = \ln\left(\frac{n^n}{(n-k)^{n-k}}\right) - k$$

Por lo que aplicando la exponencial en ambos lados de la desigualdad anterior obtenemos que: $n(n-1)\cdots(n-k+1) \geq n^k e^{-k}\left(1+\frac{k}{(n-k)}\right)^{n-k}$.

Por la desigualdad de Bernoulli tenemos que: $n(n-1)\cdots(n-k+1) \geq n^k e^{-k}(1+k)$ tomando recíprocos se tiene: $\left[n(n-1)\cdots(n-k+1)\right]^{-1} \leq n^{-k}e^k(1+k)^{-1}$, por lo tanto de **(3)** se sigue el resultado.

Q.E.D.

CAPITULO VII

TEOREMAS TRIGONOMETRICOS DE JACKSON

En este capítulo estudiaremos lo desarrollado en el anterior pero ahora para polinomios trigonométricos.

§1. Introducción

(VII.1) Definición. Sea $f \in PC([0,2\pi])$ entonces *la serie de Fourier* de f es la serie: $\frac{1}{2}a_0 + \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \operatorname{sen}(kx)$ donde a_k y b_k son los coeficientes de Fourier de la función f. Para indicar la relación de la serie de Fourier con la función, se denotará: $f(x) \sim \frac{1}{2}a_0 + \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \operatorname{sen}(kx)$.

La notación anterior no significa que la serie de Fourier converja a $f(x)$ en algún punto específico. Para mostrar este fenómeno veremos el siguiente:

(VII.2) Ejemplo. Sea $f\colon[-\pi,\pi] \to \mathbb{R}$ definida por:

$$f(x) = \begin{cases} -1 & \text{si } x \in (-\pi, 0) \\ 1 & \text{si } x = -\pi, x \in [0,\pi] \end{cases}$$

es fácil ver que la serie de Fourier de f es $\frac{4}{\pi}\sum_{k=1}^{\infty} \frac{\operatorname{sen}(kx)}{k}$ y es claro que no converge a $f(x)$ en $x = 0, \pm\pi$, aunque si lo hace en los demás puntos del intervalo.

Denotaremos por $S_n(f)(x)$ la *n-ésima suma parcial de la serie de Fourier de* f. Encontraremos una expresión cerrada para $S_n(f)$.

(VII.3) Lema. *Sea $f \in PC([-\pi,\pi])$ entonces $S_n(f)$ está dada por:*

$$S_n(f)(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x \pm t) D_n(t)\,dt$$

donde $D_n(t)$ es el **n-ésimo núcleo de Dirichlet** *definido por*

$D_n(t) = \frac{1}{2} + \sum_{k=1}^{n} \cos(kt).$

Demostración. Por definición de $S_n(f)$,a_k y b_k tenemos que:

$S_n(f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)D_n(x-t)dt$ haciendo el cambio de variable $y = x - t$ obtenemos $S_n(f)(x) = \frac{1}{\pi} \int_{x-\pi}^{x+\pi} f(x-y)D_n(y)dy$, como f y $D_n \in PC([-\pi,\pi])$ se sigue:

$S_n(f)(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x-y)D_n(y)dy.$

Por otro lado si hacemos el cambio de variable $t = x + s$ y observando que a partir de la definición de D_n tenemos que D_n es par, obtenemos que

$S_n(f)(x) = \frac{1}{\pi} \int_{x-\pi}^{x+\pi} f(x+s)D_n(s)ds.$

Q.E.D.

Enunciaremos algunas propiedades del núcleo de Dirichlet.

(VII.4) Lema. *Sea $D_n(t)$ el n-ésimo núcleo de Dirichlet entonces:*

(i)

$$D_n(t) = \begin{cases} \frac{\operatorname{sen}(n+\frac{1}{2})t}{2\operatorname{sen}(\frac{1}{2}t)} & 0 < |t| \leq \pi \\ n + \frac{1}{2} & t = 0 \end{cases}.$$

(ii) $D_n(t) = \sum_{j=-n}^{n} e^{ijt}$. *(forma compleja)*

(iii) $D_n \in \mathcal{T}_n$, *es par y continua.*

(iv) $\frac{1}{\pi} \int_{-\pi}^{\pi} D_n(t)dt = \frac{2}{\pi} \int_{0}^{\pi} D_n(t)dt = 1.$

Demostración.

(i) Multiplicando $D_n(t)$ por $\operatorname{sen}(\frac{t}{2})$ y usando la siguiente identidad trigonométrica $\operatorname{sen} x \cos y = \frac{1}{2}\{\operatorname{sen}(x+y) - \operatorname{sen}(x-y)\}$, obtenemos la expresión buscada.

(ii) Recordemos que $e^{ikt} = \cos(kt) + i\operatorname{sen}(kt)$ por lo que $\cos(kt) = \frac{1}{2}\{e^{ikt} + e^{-ikt}\}$ para toda $k = 1, \ldots, n$.

(iii) Se sigue a partir de la definición de $D_n(t)$.

(iv) Observemos que $\int_{-\pi}^{\pi} \cos(kt) = 0$ si $k = 1, \ldots, n$, por otro lado $\frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{2} dt = 1$

y así tenemos que $\frac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} D_n(t)dt = 1.$

Q.E.D.

A pesar de que $S_n(f) \in \mathcal{T}_n$ el ejemplo (VII.2) muestra que $S_n(f)$ no es una elección adecuada para aproximar uniformente a $f \in PC([-\pi,\pi])$. Para este fin consideramos las medias aritméticas de $S_n(f)$:

$\sigma_n(f)(x) = \frac{1}{n}\{S_0(f)(x) + \ldots + S_{n-1}(f)(x)\} =$

$= \frac{a_0}{2} + \displaystyle\sum_{k=1}^{n} \frac{n-k}{n}(a_k\cos(kx) + b_k\operatorname{sen}(kx))$ **(sumas de Fejer) (1).**

Por (VII.3) y (VII.4)(i) obtenemos que:

$$\sigma_n(f) = \frac{1}{n\pi}\int_{-\pi}^{\pi} f(x \pm t)K_n(t)dt = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x\pm t)\frac{\operatorname{sen}(\frac{t}{2}) + \ldots + \operatorname{sen}(n-\frac{1}{2})t}{2n\operatorname{sen}\frac{t}{2}}dt \quad (2)$$

a $K_n(t) = \frac{1}{n}\{D_0(t) + \ldots + D_{n-1}(t)\}$ es llamado el **Núcleo de Fejer de n-ésimo orden.**

Pero si $\widehat{K}(t) = \operatorname{sen}\frac{t}{2} + \operatorname{sen}(\frac{3}{2}t) + \ldots + \operatorname{sen}(n-\frac{1}{2})t$ entonces

$\operatorname{sen}(\frac{t}{2})\widehat{K}(t) = \frac{1}{2}(1-\cos(nt)) = \operatorname{sen}^2(\frac{n}{2}t)$, utilizando la siguiente identidad

$\operatorname{sen}\alpha\operatorname{sen}\beta = \frac{1}{2}\{\cos(\alpha-\beta) - \cos(\alpha+\beta)\}$, por lo que:

$\widehat{K}(t) = \frac{\operatorname{sen}^2(\frac{n}{2}t)}{\operatorname{sen}\frac{t}{2}}$ de **(2)** se sigue: $K_n(t) = \frac{1}{2n}\frac{\widehat{K}(t)}{\operatorname{sen}\frac{1}{2}t}$ si $t \neq 0$ y si $t = 0$ tenemos que,

$K_n(0) = \frac{1}{n}\displaystyle\sum_{k=0}^{n-1} D_k(0) = \frac{1}{n}\sum_{k=0}^{n-1}(k+\frac{1}{2})$ por (VII.4)(i), de donde $K_n(0) = \frac{n}{2}$. Por otro lado debidó a las propiedades que cumple $D_n(t)$ (ver (VII.4)), obtenemos el siguiente:

(VII.5) Teorema. *Sea $K_n(t)$ el n-ésimo núcleo de Fejer entonces:*

(i)

$$K_n(t) = \begin{cases} \dfrac{1}{2n}\left(\dfrac{\operatorname{sen}(\frac{n}{2}t)}{\operatorname{sen}(\frac{1}{2}t)}\right)^2 & 0 < |t| \leq \pi \\ \frac{n}{2} & t = 0 \end{cases}.$$

(ii) $K_n \in \mathcal{T}_n$ *es par ,* $K_n(t) \geq 0$ *y continua.*

(iii) $\frac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} K_n(t)dt = \frac{2}{\pi}\int_0^{\pi} K_n(t)dt = 1.$

(iv) $K_n(t) = \displaystyle\sum_{j=-n+1}^{n-1}\left(1 - \frac{|j|}{n}\right)e^{ikt}$. *(forma compleja)*

Demostración.

(i) Es la discusión previa al teorema.

(ii) $K_n \geq 0$ es clara del inciso anterior y el resto del inciso es claro a partir de la propiedad de $D_n(t)$ ((VII.4)(iii)).

(iii) Es obvio de (VII.4)(iv).

(iv) Dado que $K_n(t) = \frac{1}{n}\sum_{k=0}^{n-1} D_k(t)$ y $D_k(t) = \sum_{j=-k}^{k} e^{ijt} = \sum_{j=0}^{k} e^{ijt} + \sum_{j=1}^{k} e^{-ijt}$ tenemos que:

$$K_n(t) = \frac{1}{n}\sum_{k=0}^{n-1}\left[\sum_{j=0}^{k} e^{ijt} + \sum_{j=1}^{k} e^{-ijt}\right] =$$

$$= \frac{1}{n}\left\{\sum_{j=0}^{n-1}(n-j)e^{ijt} + \sum_{j=1}^{n-1}(n-j)e^{-ijt}\right\} =$$

$$= \frac{1}{n}\left\{\sum_{j=-n+1}^{n-1}(n-|j|)e^{ijt}\right\} = \sum_{j=-n+1}^{n-1}\left(1-\frac{|j|}{n}\right)e^{ijt}.$$

Q.E.D.

Ahora probaremos el siguiente célebre resultado:

(VII.6) Teorema (Fejer) (1904). ***Sea $f \in PC([-\pi,\pi])$ entonces la sucesión de sus promedios de Fejer $\{\sigma_n(f)\}$ converge uniformemente a f.***

Demostración.

Tenemos que: $\sigma_n \in T_n$. Sustituyendo los coeficientes de Fourier se sigue que:

$$\sigma_n(f)(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left\{\frac{1}{2} + \sum_{k=1}^{n}\frac{n-k}{n}\cos(k(t-x))\right\}dx =$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(y+x)\left\{\frac{1}{2} + \sum_{k=1}^{n}\frac{n-k}{n}\cos(ky)\right\}dy, \text{ debido al cambio de variable}$$

$y = t - x$. **Haciendo** $U_n(t) = \frac{1}{2} + \sum_{k=1}^{n}\frac{n-k}{n}\cos(kt)$, **tenemos que** $U_n(t) \geq 0$ **ya que** $U_n(t) = K_n(t) \geq 0$ **de donde por (VI.13):**

$|f(x) - \sigma_n(f)(x)| \leq w(\frac{1}{m})\left[1 + \frac{m\pi}{\sqrt{2}}(1-r_{1,n})^{\frac{1}{2}}\right]$ con $m \in \mathbb{R}^+$ y $r_{1,n} = \frac{n-1}{n}$ por lo que si $m = \sqrt{n}$ se tiene que:

$|f(x) - \sigma_n(f)(x)| \leq w(\frac{1}{\sqrt{n}})\left[1 + \frac{\pi}{\sqrt{2}}\right]$. Como f es uniformemente continua entonces $w(\delta) \to 0$ cuando $\delta \to 0$ por lo que $||f - \sigma_n||_{[-\pi,\pi]} \to 0$.

Q.E.D.

(VII.7) Observación.

El teorema anterior da otra demostración al Teorema Trigonométrico de Weierstrass, en la que los polinomios trigonométricos aproximadores se dan *explícitamente*.

Daremos otra prueba al Teorema polinomial de Weierstrass, usando el Teorema trigonométrico, con lo cuál concluimos con la equivalencia entre los dos resultados. Sin pérdida de generalidad nos restringimos a $[0,\pi]$.

(VII.8) Teorema polinomial de Weierstrass (1885) *Sea* $f \in C([0,\pi])$ *entonces dada* $\varepsilon > 0$ *existe* $p \in \mathcal{P}$ *tal que* $||f - p||_{[0,\pi]} < \varepsilon$.

Demostración. Extendemos f a una función par en $[-\pi,\pi]$ de tal manera que $f \in PC([-\pi,\pi])$ entonces por el Teorema Trigonométrico de Weierstrass, dada $\varepsilon > 0$ existe $T \in \mathcal{T}$ tal que $||f - T||_{[-\pi,\pi]} < \frac{\varepsilon}{2}$. Como T es una función analítica en $\mathbb{R}$, su serie de potencias converge uniformemente en subconjuntos compactos de $\mathbb{R}$, en partícular en $[-\pi,\pi]$.

Por lo tanto existe un polinomio $p \in \mathcal{P}$ tal que $||T - p||_{[-\pi,\pi]} < \frac{\varepsilon}{2}$. Por lo que se sigue que $||f - p||_{[-\pi,\pi]} < \varepsilon$, así se tiene que $||f - p||_{[0,\pi]} < \varepsilon$.

Q.E.D.

§2. Núcleo de Jackson

Estudiaremos con que rapidez tienden los polinomos trigonométricos a una función $f \in PC([0,2\pi])$, para lo cuál introduciremos el **núcleo de Jackson**.

Primero motivaremos su definición. Observemos que en cierto sentido el núcleo de Fejer es esencialmente el cuadrado del núcleo de Dirichlet. Los únicos dos cambios son:

(i) $\frac{n}{2}$ por $n+\frac{1}{2}$ para que $K_n(t)$ sea un polinomio trigonométrico de grado n.

(ii) $\frac{1}{2n}$ por $\frac{1}{2}$ para que $\frac{1}{\pi}\int_{-\pi}^{\pi} K_n(t)dt = 1$.

La idea de Jackson es tomar el cuadrado del núcleo de Fejer con las mismas modificaciones para lo cuál consideramos:

$$\left(\frac{\operatorname{sen}\left(\frac{nt}{2}\right)}{\operatorname{sen}\left(\frac{t}{2}\right)}\right)^4$$

Dado que $K_n \in \mathcal{T}_n$ entonces $K_n^2 \in \mathcal{T}_{2n}$ por lo que vamos a sustituir a n por $\left[\frac{n}{2}\right]+1$ para que:

$$\left(\frac{\operatorname{sen}\left(\left(\left[\frac{n}{2}\right]+1\right)\frac{t}{2}\right)}{\operatorname{sen}\left(\frac{t}{2}\right)}\right)^4 \in \mathcal{T}_n.$$

Sea $c_n \in \mathbb{R}^+$ tal que:

$$J_n(x) = \frac{1}{c_n}\left(\frac{\operatorname{sen}\left(\left(\left[\frac{n}{2}\right]+1\right)\frac{t}{2}\right)}{\operatorname{sen}\left(\frac{t}{2}\right)}\right)^4$$

cumpla con $\int_{-\pi}^{\pi} J_n(x)dx = 1$, $n \geq 1$, esto nos lleva a la siguiente:

(VII.9) Definición. El n-ésimo núcleo de Jackson $J_n(x)$ se define por:

$$J_n(x) = \frac{1}{c_n}\left(\frac{\operatorname{sen}\left(\left(\left[\frac{n}{2}\right]+1\right)\frac{t}{2}\right)}{\operatorname{sen}\left(\frac{t}{2}\right)}\right)^4$$

a $J_n^*(f)(x) = \int_{-\pi}^{\pi} f(t)J_n(x-t)dt$ se le conoce como el **operador de Jackson** y c_n es una constante tal que $\int_{-\pi}^{\pi} J_n(x)dx = 1$.

(VII.10) Teorema.

Sea $J_n(x)$ el n-ésimo núcleo de Jackson entonces:

(i) $J_n(x) \geq 0$ *en* $[-\pi, \pi]$ *y es par.*

(ii) $\int_{-\pi}^{\pi} J_n(x)dx = 2\int_0^{\pi} J_n(x)dx = 1.$

(iii) $c_n = \frac{2\pi}{3}m(2m^2+1)$ *con* $m = \left[\frac{n}{2}\right] + 1$ *para toda* $n \geq 2$.

(iv) $J_n(0) = \frac{m^4}{c_n}$.

Demostración.

(i) y (ii) se siguen de la definición.

(iii) Sea $n \in \mathbb{N}$. Por (VII.10)(ii) $c_n = \int_{-\pi}^{\pi}\left(\frac{\operatorname{sen}\left(m\frac{x}{2}\right)}{\operatorname{sen}\left(\frac{x}{2}\right)}\right)^4 dx$ con $m = \left[\frac{n}{2}\right] + 1$. Por la definición del núcleo de Fejer sabemos que si $0 < |x| \leq \pi$ entonces:

$$\left(\frac{\operatorname{sen}(\frac{mx}{2})}{\operatorname{sen}(\frac{x}{2})}\right)^2 =$$
$$= 2\sum_{j=0}^{m-1} D_j(x)$$
$$= m + 2\left[(m-1)\cos(x) + (m-2)\cos(2x) + \cdots + \cos(m-1)x\right] \quad (1).$$

Si $x = 0$ por un lado tenemos $J_m(0) = \frac{m}{2}$ de donde $2mJ_m(0) = m^2$ mientras que en **(1)** se obtiene m^2, por lo que podemos concluir que:

$$\left(\frac{\operatorname{sen}(\frac{mx}{2})}{\operatorname{sen}(\frac{x}{2})}\right)^2 = m + 2\left[(m-1)\cos(x) + (m-2)\cos(2x) + \cdots + \cos(m-1)x\right]$$

para toda $x \in [-\pi, \pi]$.

Por otro lado dado que $\int_{-\pi}^{\pi} \cos(kt)\cos(jt) = \pi\delta_{ij}$ para toda $i, j \in \mathbb{N} \cup \{0\}$, tenemos que:

$$c_n = \int_{-\pi}^{\pi} \left(m + 2\left[(m-1)\cos(x) + (m-2)\cos(2x) + \cdots + \cos(m-1)x\right]\right)^2 dx =$$
$$= 2\pi m^2 + 4\pi\left[(m-1)^2 + \ldots + 1^2\right] = \frac{2\pi}{3}m(2m^2+1).$$

(iv) Por **(1)** del inciso anterior se tiene que $J_n(0) = \frac{m^4}{c_n}$

Q.E.D.

Veremos de nuevo el siguiente:

(VII.11) Teorema Trigonométrico de Jackson. *Sea $f \in PC([0, 2\pi])$ entonces $\widehat{E}_n(f) \leq c\omega(\frac{1}{n})$ $n > 0$ y con $c > 0$ constante.*

Demostración. Sea $x \in [-\pi, \pi]$ entonces

$$|f(x)-J_n^*(f)(x)| = |f(x)-\int_{-\pi}^{\pi} f(t)J_n(x-t)dt| = |f(x)-\int_{-\pi}^{\pi} f(x-t)J_n(t)dt| =$$

$$= \left|\int_{-\pi}^{\pi}(f(x)-f(x-t))J_n(t)dt\right| \leq \int_{-\pi}^{\pi}|f(x)-f(x-t)|J_n(t)dt \leq$$

$$\leq \int_{-\pi}^{\pi} w(f,|t|)J_n(t)dt.$$

por (VI.4) incisos (iii), (iv) y (v) tenemos que:

$$w(f,|t|) = w(f,\tfrac{n|t|}{n}) \leq w(f,\tfrac{[n|t|]+1}{n}) \leq ([n|t|]+1)w(f,\tfrac{1}{n})$$
$$\leq (n|t|+1)w(f,\tfrac{1}{n}).$$

por lo anterior se tiene:

$$|f(x)-J_n^*(x)| \leq \int_{-\pi}^{\pi}(n|t|+1)w\left(f,\frac{1}{n}\right)J_n(t)dt =$$

$$= 2nw(f,\tfrac{1}{n})\int_0^{\pi}|t|J_n(t)dt + 2w\left(f,\frac{1}{n}\right) \text{ (1)}$$

pues $|t|J_n(t)$ es una función par. Resta acotar a la integral $\int_0^{\pi}|t|Jn(t)dt$ para lo cual procedemos como sigue: haciendo el cambio de variable $t = 2y$ tenemos que $\int_0^{\pi}|t|J_n(t)dt = \int_0^{\pi} tJ_n(t) = 4\int_0^{\frac{\pi}{2}} yJ_n(2y)dy$, por definición del núcleo de Jackson y eligiendo $m = [\frac{n}{2}]+1$ suficientemente grande tenemos que:

$$\int_0^{\pi} tJ_n(t) = \frac{4}{c_n}\int_0^{\frac{1}{m}} t\left(\frac{\operatorname{sen}(mt)}{\operatorname{sen}(t)}\right)^4 dt + \frac{4}{c_n}\int_{\frac{1}{m}}^{\frac{\pi}{2}} t\left(\frac{\operatorname{sen}(mt)}{\operatorname{sen}(t)}\right)^4 dt \text{ (2)}.$$

por (VI.11) y del hecho que $\operatorname{sen}(t) \leq t$ en $[0,2\pi]$ se tiene que:

$\left(\frac{\operatorname{sen}(mt)}{\operatorname{sen}(t)}\right)^4 \leq \left(\frac{\pi}{2t}\right)^4$ y es fácil ver que $|\operatorname{sen}(nt)| \leq n|\operatorname{sen}(t)|$ si $t \in [0,\frac{\pi}{2}]$ entonces de **(2)** se sigue que:

$$\int_0^{\pi} tJ_n(t) \leq \frac{4}{c_n}\int_0^{\frac{1}{m}} tm^4dt + \frac{4}{c_n}\int_{\frac{1}{m}}^{\frac{\pi}{2}} t\left(\frac{\pi}{2t}\right)^4 dt.$$

evaluando las integrales anteriores se obtiene que $\int_0^{\pi} tJ_n(t) \leq \frac{2}{c_n}m^2 + \frac{\pi^4}{8c_n}m^2$, pero $c_n \geq \frac{4\pi}{3}m^3$ entonces se tiene que $\frac{2}{c_n}m^2 \leq \frac{3}{2\pi m}$ pero $n \leq 2m$, ya que $\frac{n}{2}-1 \leq [\frac{n}{2}]$, por lo que $\frac{2}{c_n}m^2 \leq \frac{3}{\pi n} < \frac{1}{n}$, mientras que $\frac{\pi^4}{8c_n}m^2 \leq \frac{3\pi^3}{16m} < \frac{6}{n}$, por lo tanto se tiene que $\int_0^{\pi} tJ_n(t) \leq \frac{7}{n}$, por lo que de **(1)** se sigue que $|f(x)-J_n^*(f)(x)| \leq 16w(f,\frac{1}{n})$, y a partir de la definición de $J_n^*(f)(x)$ y del hecho que al integrar un polinomio trigonométrico su grado no crece entonces $J_n^*(f) \in \mathcal{T}_n$ por lo que $\widehat{E}_n(f) \leq 16w(f,\frac{1}{n})$.

Q.E.D.

(VII.12) Corolario. *Sea* $\mathcal{L}^* = \{f \in PC([0,2\pi]) \mid w(f,\delta) \leq \delta\}$ *entonces* $\widehat{E}_n(f) \leq \frac{16}{n}$ *para toda* $f \in \mathcal{L}^*$.

(VII.13) Corolario. *Sea* $f \in \mathrm{Lip}_M(\alpha) \cap PC([0,2\pi])$, $\alpha \in (0,1)$ *entonces* $\widehat{E}_n(f) \leq \frac{16c}{n^\alpha}$ *para alguna constante* c *y para toda* $n \in \mathbb{N}$.

(VII.14) Definición. Definimos a $PC^p([0,2\pi]$ cómo el siguiente conjunto: $\{f \in PC([0,2\pi]) \mid f^{(j)}$ existe y $f^{(j)} \in PC([0,2\pi])$ para cada $j = 1,\ldots,p\}$.

Entonces tenemos:

(VII.15) Corolario. *Sea* $f \in PC^1([0,2\pi])$, *entonces* $\widehat{E}_n(f) \leq \frac{16}{n}M$ *donde* $|f'(x)| \leq M$ *para toda* $x \in [-\pi,\pi]$.

(VII.16) Teorema. *Sea* $f \in PC^1([-\pi,\pi])$, *entonces* $\widehat{E}_n(f) \leq \frac{c}{n}\widehat{E}_n(f')$ *para alguna* $c \in \mathbb{R}$.

Demostración. Sea T_n^* el mejor aproximador trigonométrico de grado$\leq$n para f' entonces $\widehat{E}_n(f') = \|f'-T_n^*\|_{[-\pi,\pi]}$. Sea $T_n(x) = \displaystyle\int_0^x T_n^*(t)dt \in \mathcal{T}_n$, por (III.13)(iii) tenemos que $\widehat{E}_n(f,[-\pi,\pi]) = \widehat{E}_n(f - T_n,[-\pi,\pi])$ pero $T_n'(x) = T_n^*(x) \in \mathcal{T}_n$ por lo que $|f' - T_n'| = |f' - T_n^*| \leq \widehat{E}_n(f') = M$ por (VII.15) se sigue que $\widehat{E}_n(f,[-\pi,\pi]) \leq$ $\leq \frac{16}{n}\widehat{E}_n(f')$.

Q.E.D.

(VII.17)Teorema. *Sea* $f \in PC^p([-\pi,\pi])$, *entonces* $\widehat{E}_n(f) \leq \frac{k}{n^p}w(f^{(p)},\frac{1}{n})$ *con* $k > 0$ *constante que sólo depende de* p.

Demostración. Por el Teorema anterior tenemos que $\widehat{E}_n(f) \leq \frac{16}{n}\widehat{E}_n(f') \leq$ $\leq \left(\frac{16}{n^2}\right)\widehat{E}_n(f'')$ por lo que $\widehat{E}_n(f) \leq \left(\frac{16}{n}\right)^p \widehat{E}_n(f^{(p)}) \leq \frac{16^{p+1}}{n^p}w(f^{(p)},\frac{1}{n})$.

Q.E.D.

CAPITULO VIII

INVERSOS A LOS TEOREMAS DE JACKSON

En este capítulo nos dedicaremos a estudiar los inversos a los resultados en el capítulo anterior, así como a estudiar una clase importante de funciones llamada la clase de Zygmund.

§1. Teoremas de Bernstein

Empezamos primero con **las desigualdades de Bernstein y Markoff**. Sea $f \in C([-1,1])$ dada, sea $p \in \mathcal{P}_n$ el polinomio interpolador de Lagrange construido con los ceros de $T_n(x)$ (polinomio de Chebyshev de grado$= n$), los cuales son: $x_i = \cos(\frac{(2i-1)\pi}{2n})$ con $i = 1, \ldots, n$ (IV.2)(iv) y en los que $p(x_i) = f(x_i)$. Por la discusión previa al teorema de Walsh (II.22) sabemos que:

$$\ell_k(x) = \frac{w(x)}{(x - x_i)w'(x_i)}$$

donde $w(x) = 2^{n-1}\prod_{i=1}^{n}(x - x_i) = T_n(x)$ (1), con $\theta = \arccos x$, $\theta \in [0,\pi]$, $x \in [-1,1]$, entonces:

$$p(x) = \sum_{k=1}^{n} f(x_i)\frac{T_n(x)}{(x - x_i)w'(x_i)}$$

Por otro lado $w'(x_i) = \frac{n\operatorname{sen}(n\theta_i)}{\sqrt{1-x_i^2}}$ donde $\theta_i = \arccos x_i$, por lo que

$\operatorname{sen}(n\theta_i) = (-1)^{i-1}$ y como $\operatorname{sen}(\theta_i) = \sqrt{1 - x_i^2}$, de lo anterior se sigue que:

$$p(x) = \frac{1}{n}\sum_{k=1}^{n} f(x_i)(-1)^{i-1}\frac{T_n(x)\operatorname{sen}(\theta_i)}{(x - x_i)} \quad (1).$$

(VIII.1) Lema. *Sea $p \in \mathcal{P}_{n-1}$ entonces $||p||_{[-1,1]} \le ||n\sqrt{1-x^2}p||_{[-1,1]}$.*

Demostración. Sea $M = ||n\sqrt{1-x^2}p||_{[-1,1]}$. Basta con demostrar que $|p(x)| \le M$ para toda $x \in [-1,1]$. Sabemos que los ceros del polinomio de Chebyshev ocurren en los puntos $x_i = \cos(\frac{(2i-1)\pi}{2n})$ con $i = 1, \ldots, n$.

Si $x \in [x_n, x_1]$ como $|x| \leq |x_1|$ entonces $\sqrt{1-x^2} \geq \sqrt{1-x_1^2}$, y además por (VI.11) $\sqrt{1-x_1^2} = \text{sen}(\frac{\pi}{2n}) \geq \frac{1}{n}$ se tiene que $1 \leq n\sqrt{1-x^2}$ para toda $x \in [x_n, x_1]$.

Así tenemos que $|p(x)| \leq n\sqrt{1-x^2}|p(x)| \leq M$ para toda $x \in [x_n, x_1]$.

Para $x \in [-1, x_n] \cup [x_1, 1]$ aplicaremos **(1)**, observemos que los números $x - x_i$ tienen el mismo signo, por lo que de **(1)** se tiene que:

$$|p(x)| \leq \frac{1}{n}\sum_{k=1}^{n} |p(x_i)| \frac{|T_n(x)||\sqrt{1-x_i^2}|}{|x-x_i|} \leq$$

$$\leq \frac{M}{n^2}\sum_{i=1}^{n}\left|\frac{T_n(x)}{x-x_i}\right| = \frac{M}{n^2}\left|\sum_{i=1}^{n}\frac{T_n(x)}{x-x_i}\right|.$$

Por otro lado como $T_n(x) = 2^{n-1}\prod_{i=1}^{n}(x-x_i)$ se tiene que $T'(x) = \sum_{i=1}^{n}\frac{T_n(x)}{x-x_i}$ y por lo anterior se tiene que: $|p(x)| \leq \frac{M}{n^2}|T_n'(x)| = \frac{M}{n}\left|\frac{\text{sen}(n\theta)}{\text{sen}(\theta)}\right| \leq M$.

Q.E.D.

(VIII.2) Lema. *Sea $T \in \mathcal{T}_n$ impar, entonces:* $||\frac{T}{\text{sen}\theta}||_{[-\pi,\pi]} \leq n||T||_{[-\pi,\pi]}$.

Demostración.

Por inducción y usando la identidad $\text{sen}(k+1)\theta = \text{sen}(k\theta)\cos(\theta) + \text{sen}(\theta)\cos(k\theta)$ se tiene que $\frac{\text{sen}(k\theta)}{\text{sen}(\theta)}$ es un polinomio de grado$\leq k-1$ en $\cos(\theta)$. Por lo que existe $p \in \mathcal{P}_{n-1}$ tal que $p(\cos(\theta)) = \frac{T(\theta)}{\text{sen}(\theta)}$ y por (VIII.1) se tiene:

$||p(\cos(\theta))||_{[-\pi,\pi]} = ||p(x)||_{[-1,1]} \leq ||n\sqrt{1-x^2}p||_{[-1,1]}$.

Por otro lado se tiene que $n\sqrt{1-x^2}p(x) = n\text{sen}(\theta)p(\cos(\theta)) = nT(\theta)$. Sustituyendo en la desigualdad anterior se tiene que:

$||\frac{T}{\text{sen}\theta}||_{[-\pi,\pi]} \leq n||T||_{[-\pi,\pi]}$, para todo $T \in \mathcal{T}_n$ impar.

Q.E.D.

(VIII.3) Teorema. Desigualdad de Bernstein (1912). *Sea $T \in \mathcal{T}_n$, entonces:* $||T'||_{[-\pi,\pi]} \leq n||T||_{[-\pi,\pi]}$.

Demostración. Definimos $f(\alpha,\theta) = \frac{1}{2}[T(\alpha+\theta) - T(\alpha-\theta)]$. Para α fija en $[-\pi,\pi]$ tenemos que $f(\alpha,\theta) \in \mathcal{T}_n$ y es impar. Sea $M = ||T||$ entonces $|f(\alpha,\theta)| \leq M$ y por (VIII.2) tenemos que: $||\frac{f(\alpha,\theta)}{\text{sen}(\theta)}|| \leq n||f(\alpha,\theta)|| \leq nM$.

Pero $T'(\alpha) = \lim_{\theta\to\infty}\frac{T(\alpha+\theta)-T(\alpha-\theta)}{2\theta} = \lim_{\theta\to\infty}\frac{f(\alpha,\theta)}{\text{sen}\theta}\frac{\text{sen}\theta}{\theta} \leq$

$\leq nM \lim_{\theta\to\infty}\frac{\text{sen}\theta}{\theta} = nM$, por lo tanto se tiene que:

$|T'(\alpha)| \leq n||T||$ para toda $\alpha \in [-\pi, \pi]$ de donde $||T'||_{[-\pi,\pi]} \leq n||T||_{[-\pi,\pi]}$.

Q.E.D.

(VIII.4) Teorema. Desigualdad de Markoff (1889). *Sea $p \in \mathcal{P}_n$, entonces:* $||p'||_{[-1,1]} \leq n^2||p||_{[-1,1]}$.

Demostración. Sea $M = ||p||_{[-1,1]}$ entonces $||p(\cos(\theta))||_{[-\pi,\pi]} = ||p||_{[-1,1]}$ pero $p(\cos\theta) \in \mathcal{T}_n$ y es par (IV.11). Por (VIII.3) tenemos que:

$|p'(\cos\theta)\text{sen}\theta| \leq n||p||_{[-1,1]}$, en otras palabras se tiene que:

$|p'(x)\sqrt{1-x^2}| \leq n||p||_{[-1,1]}$ en $[-1,1]$ (1).

Aplicando (VIII.1) a $p'(x)$ obtenemos que: $||p'||_{[-1,1]} \leq ||n\sqrt{1-x^2}p'||$ y de **(1)** se sigue que $||p'||_{[-1,1]} \leq n^2||p||_{[-1,1]}$.

Q.E.D.

(VIII.5) Observación. Si $p = T_n$ (el n-ésimo polinomio de Chebyshev) entonces $||p'||_{[-1,1]} = n^2||p||_{[-1,1]}$, por lo que no se puede mejorar la cota n^2 en el teorema anterior.

Estamos listos para estudiar los Teoremas de Bernstein. Recordando uno de los Teoremas de Jackson (VII.13), el inverso a esté resultado es:

(VIII.6) Teorema. *Dada $f \in PC([0, 2\pi])$ entonces:*

(i) Sea $\alpha \in (0,1)$. Si $\hat{E}_n(f) \leq \frac{c}{n^\alpha}$ para toda $n \in \mathbb{N}$, entonces: $f \in \text{Lip}_M(\alpha)$.

(ii) Si $\hat{E}_n(f) \leq \frac{c}{n}$ para cada $n \in \mathbb{N}$, entonces: $w(f,\varepsilon) \leq c_2\varepsilon\ln(\frac{1}{\varepsilon})$ si $\varepsilon \in (0, \frac{1}{e}]$.

Demostración. Sea $\alpha \in (0,1]$ y $T_n \in \mathcal{T}_n$ tal que $\hat{E}_n(f) = ||f - T_n|| \leq \frac{c}{n^\alpha}$ para toda $n \in \mathbb{N}$. Consideramos $Q_n(x) = T_{2^n}(x) - T_{2^{n-1}}(x)$ para toda $n \in \mathbb{N}$ y $Q_0(x) = T_1(x)$ entonces $||Q_n|| \leq ||T_{2^n} - f|| + ||f - T_{2^{n-1}}|| \leq \frac{c_1}{2^{n\alpha}}$ con $c_1 = c(1+2^\alpha)$.

Hasta el momento los métodos usados para esta desigualdad no son aplicables para $n = 0$, pero para este caso especial argumentaremos de la siguiente manera: $||Q_0|| = ||T_1|| \leq c < (1+2^\alpha)c$, por lo que $||Q_n|| \leq \frac{c_1}{2^{n\alpha}}$ para toda $n \in \mathbb{N}\cup\{0\}$ **(1)**.

Sea $S_n(x) = \sum_{k=0}^{n} Q_k(x) = T_{2^n}(x)$. Como $||f - T_{2^n}|| \leq \frac{c}{2^{n\alpha}}$ se sigue que $S_n \to f$ cuando $n \to \infty$ uniformemente en $[0, 2\pi]$, por lo tanto $f(x) = \sum_{k=0}^{\infty} Q_k(x)$ (uniformemente en $[0, 2\pi]$).

Por otro lado $Q_n \in \mathcal{T}_{2^n}$. Aplicando la desigualdad de Bernstein y **(1)** se tiene que $||Q_n'|| \leq c_1 2^{n(1-\alpha)}$ **(2)**.

Sea $\varepsilon \in (0,1)$ y $n \in \mathbb{N}$ fija entonces:

$$|f(x+\varepsilon)-f(x)| = \left|\sum_{k=0}^{\infty} Q_k(x+\varepsilon)-Q_k(x)\right| \leq$$

$$\leq \left|\sum_{k=0}^{n-1} Q_k(x+\varepsilon)-Q_k(x)\right| + \left|\sum_{k=n}^{\infty} Q_k(x+\varepsilon)-Q_k(x)\right| \leq$$

$$\leq \varepsilon \sum_{k=0}^{n-1} |Q_k'(\xi)| + 2\sum_{k=n}^{\infty} \|Q_k\|$$

$$\leq \varepsilon \sum_{k=0}^{n-1} \|Q_k'\| + 2\sum_{k=n}^{\infty} \|Q_k\| \leq c_1\varepsilon \sum_{k=0}^{n-1} 2^{k(1-\alpha)} + 2c_1 \sum_{k=n}^{\infty} \frac{1}{2^{k\alpha}} \ \textbf{(3)} \ (\text{por } \textbf{(1)} \text{ y } \textbf{(2)}).$$

(i) Por lo anterior dada $\alpha \in (0,1)$ tenemos que:
$|f(x+\varepsilon)-f(x)| \leq \frac{c_1}{2^{(1-\alpha)}-1}(\varepsilon 2^{n(1-\alpha)}) + \frac{2c_1}{1-2^{-\alpha}}(2^{-n\alpha})$.
Sea $c_2 = \max\{\frac{c_1}{2^{(1-\alpha)}-1}, \frac{2c_1}{1-2^{-\alpha}}\}$ entonces se tiene que:

$|f(x+\varepsilon)-f(x)| \leq c_2(\varepsilon 2^{n(1-\alpha)} + 2^{-n\alpha})$ **(4)**, y ahora consideramos $n \in \mathbb{N}$ tal que $\frac{1}{\varepsilon} < 2^n \leq \frac{2}{\varepsilon}$ por lo que de **(4)** se sigue que: $|f(x+\varepsilon)-f(x)| \leq M\varepsilon^{\alpha}$ con $M = c_2 2^{(1-\alpha)}$ por lo tanto $f \in \mathrm{Lip}_M(\alpha)$.

(ii) Si $\alpha = 1$ de **(3)** obtenemos que: $|f(x+\varepsilon)-f(x)| \leq c_1(\varepsilon n + 4\cdot 2^{-n})$ elegimos $n \in \mathbb{N}$ tal que $2^n < \frac{1}{\varepsilon} \leq 2^{n+1}$ de lo anterior tenemos que: $|f(x+\varepsilon)-f(x)| \leq$

$\leq c_1(\varepsilon \log_2 \frac{1}{\varepsilon} + 8\varepsilon)$ pero $\log_a(x) = \frac{\ln(x)}{\ln(a)}$, por lo que se sigue:

$|f(x+\varepsilon)-f(x)| \leq c_1(\frac{\varepsilon}{\ln(2)}\ln(\frac{1}{\varepsilon}) + 8\varepsilon)$ y como $\varepsilon \in (0, \frac{1}{e}]$ entonces $\varepsilon \leq \varepsilon\ln(\frac{1}{\varepsilon})$ se tiene que $|f(x+\varepsilon)-f(x)| \leq c_2\varepsilon\ln(\frac{1}{\varepsilon})$ con $c_2 = \frac{c_1}{\ln(2)} + 8c_1$, de donde se concluye $w(f,\varepsilon) \leq c_2\varepsilon\ln(\frac{1}{\varepsilon})$.

Q.E.D.

En lo que respecta a las derivadas el corolario (VII.17) junto con el (VII.13) nos dicen:

Dada $f \in PC^p([-\pi,\pi])$ y $f^{(p)} \in \mathrm{Lip}(\alpha)$ $\alpha \in (0,1]$ entonces $\hat{E}_n(f) \leq \frac{c}{n^{\alpha+p}}$, Bernstein probó el inverso dado en el siguiente resultado.

(VIII.7) Teorema. *Sean $p \in \mathbb{N}$ y $f \in PC([-\pi,\pi])$ entonces:*

(i) Si $\alpha \in (0,1)$ y $\hat{E}_n(f) \leq \frac{c}{n^{p+\alpha}}$ para toda $n \in \mathbb{N}$ entonces $f \in PC^p([-\pi,\pi])$ y $f^{(p)} \in \mathrm{Lip}(\alpha)$.

(ii) Si $\hat{E}_n(f) \leq \frac{c}{n^{p+1}}$ para toda $n \in \mathbb{N}$ entonces $f \in PC^p([-\pi,\pi])$ y
$w(f^{(p)},h) \leq kh\ln(h)$ con k una constante y $w(f^{(p)},h) \leq kh\ln(h)$ $h \in (0,\frac{1}{e}]$.

Demostración. Procediendo igual que en la demostración anterior sabemos que $f(x) = \sum_{k=0}^{\infty} Q_k(x)$ y que $||Q_k|| \leq ||T_{2^k} - f|| + ||f - T_{2^{k-1}}|| \leq \frac{c_1}{2^{k(p+\alpha)}}$ con $c_1 = c(1 + 2^{p+\alpha})$. **(1)**

Por la desigualdad de Bernstein $||Q'_k|| \leq 2^k ||Q_k||$, aplicando. Como $\sum_{k=0}^{\infty} \frac{1}{2^{k\alpha}}$ converge entonces tenemos que $\sum_{k=0}^{\infty} Q_k^{(p)}$ converge uniformemente pero $\sum_{k=0}^{\infty} Q_k \to f$ uniformemente entonces f es p-derivable, por lo que:

$$\hat{E}^*_{2^n}(f^{(p)}) \leq ||f^{(p)} - \sum_{k=0}^{n} Q_k^{(p)}|| = || \sum_{k=n+1}^{\infty} Q_k^{(p)} || \leq$$

$$\leq c_1 \sum_{k=n+1}^{\infty} \frac{1}{2^{k\alpha}} \leq \frac{c_2}{2^{\alpha n}} \text{ con } c_2 = \tfrac{c_1}{2^\alpha - 1} \ \mathbf{(2)}.$$

(i) **Si $\alpha \in (0,1)$ y para toda $k > 0$, elegimos $n \in \mathbb{N}$ tal que $2^n \leq k < 2^{n+1}$ entonces:** $\hat{E}^*_k(f^{(p)}) \leq \hat{E}^*_{2^n}(f^{(p)}) \leq \frac{c_2}{2^{\alpha n}} < \frac{2^\alpha c_2}{k^\alpha}$, **por (VIII.6)(i)** $f \in PC^p([-\pi, \pi])$ **y**

$f^{(p)} \in$ **Lip(α) ya que f es el límite uniforme de una serie de funciones de período 2π.**

(ii) **Si $\alpha = 1$ tenemos que $\hat{E}^*_k(f^{(p)}) \leq \frac{c_2}{2^k}$ por (2). Ahora elegimos $n \in \mathbb{N}$ tal que $2^n \leq k < 2^{n+1}$ con $k > 0$ entonces $\hat{E}^*_k(f^{(p)}) \leq \hat{E}^*_{2^n}(f^{(p)}) \leq \frac{c_2}{2^k}$ por (VIII.6)(ii) se tiene que $w(f^{(p)}, \varepsilon) \leq k\varepsilon \ln(\varepsilon)$ $w(f^{(p)}, \varepsilon) \leq k\varepsilon \ln(\varepsilon)$ para $\varepsilon \in (0, \frac{1}{e}]$.**

Q.E.D.

Recordemos que si $f \in \mathrm{Lip}_M(1)$ entonces por el Teorema trigonométrico de Jackson tenemos que: $\hat{E}_n(f) \leq \frac{c}{n}$. Veremos que el recíproco no es cierto, para esto consideramos el siguiente:

(VIII.8) Ejemplo. **Sea $f(x) = \sum_{k=1}^{\infty} \frac{\mathrm{sen}(kx)}{k^2}$, es claro que $f \in PC([-\pi, \pi])$ ya que $\sum_{k=1}^{\infty} \frac{1}{k^2}$ es convergente.**

Además $\hat{E}_n(f) \leq ||f(x) - \sum_{k=1}^{n} \frac{\mathrm{sen}(kx)}{k^2}|| = || \sum_{k=n+1}^{\infty} \frac{\mathrm{sen}(kx)}{k^2} || \leq \sum_{k=1}^{\infty} \frac{1}{k^2} < \frac{1}{n}$.

Pero $f \notin$ Lip(1) pues probaremos que $\left| \frac{f(h) - f(0)}{h} \right|$ **no es acotada cuando** $h \to 0^+$. Procedemos como sigue:

$$\left| \frac{f(h) - f(0)}{h} \right| = \left| \sum_{n=1}^{\infty} \frac{\mathrm{sen}(nh)}{hn^2} \right| \geq$$

$$\geq \left| \sum_{n=1}^{\left[\frac{\pi}{2h}\right]-1} \frac{\operatorname{sen}(nh)}{hn^2} \right| - \left| \sum_{n=\left[\frac{\pi}{2h}\right]}^{\infty} \frac{\operatorname{sen}(nh)}{hn^2} \right|$$

por (VI.11) se tiene que $|\operatorname{sen}(nh)| \geq \frac{2}{\pi}nh$ si $n \leq \frac{\pi}{2h}$ por lo que:

$$\left|\frac{f(h)-f(0)}{h}\right| \geq \sum_{n=1}^{\left[\frac{\pi}{2h}\right]-1} \frac{2}{\pi h} - \frac{1}{h} \sum_{n=\left[\frac{\pi}{2h}\right]}^{\infty} \frac{1}{n^2}$$

pero la primera suma de la última desigualdad es mayor ó igual que $\frac{2}{\pi} \int_1^{\left[\frac{\pi}{2h}\right]} \frac{dx}{x}$ y la segunda suma es menor ó igual que $2 \int_{\left[\frac{\pi}{2h}\right]}^{\infty} \frac{dx}{x^2}$ por lo que se tiene que:

$\left|\frac{f(h)-f(0)}{h}\right| \geq \frac{2}{\pi}\ln(\left[\frac{\pi}{2h}\right]) - \frac{2}{h}\frac{1}{\left[\frac{\pi}{2h}\right]}$ pero $\ln(\left[\frac{\pi}{2h}\right])$ no esta acotada cuando $h \to 0^+$.

Q.E.D.

2. Clase de Zygmund.

En está sección veremos condiciones necesarias y suficientes para que $\widehat{E}_n(f) \leq \leq \frac{c}{n}$ con c una constante, para lo cuál introducimos la siguiente:

(VIII.9) Definición. Decimos que $f \in \mathcal{Z}$,donde $\mathcal{Z}$ es la **clase de Zygmund,** si existe $c > 0$ constante tal que $\|f(x+h) + f(x-h) - 2f(x)\|_1 \leq ch$ para toda $h \geq 0$.

(VIII.10) Observaciones.

(i) Es claro que si $f \in \text{Lip}(1)$, entonces $f \in \mathcal{Z}$.

(ii) El recíproco del inciso anterior es *falso* como lo muestra el siguiente ejemplo: Sea $f(x) = x\ln|x|$ en $[-a, a]$ con $a > 0$ entonces $f'(x) = \ln(x) + 1$ la cuál no es acotada en $[-a, a]$; por lo que $f \notin \text{Lip}(1)$. Para ver que $f \in \mathcal{Z}$ primero se hará un análisis de la función:

$\varphi(x) = \frac{1}{h}\{f(x+h) + f(x-h) - 2f(x)\}$ **(1)**

es fácil ver que φ es una función impar por lo que reduciremos nuestro análisis a $[0, \infty)$.

Ahora veremos el comportamiento de φ cuando $x \to \infty$ para lo cuál **(1)** la escribimos como:

$$\varphi(x) = \frac{1}{h}\left\{\ln\left((1 - \frac{h^2}{x^2}\right)^x\right\} + \ln\frac{x+h}{x-h}$$

y así se tiene que:

$$\lim_{x\to\infty} \varphi(x) = \frac{1}{h} \lim_{x\to\infty} \ln\left[\left((1 - \frac{h^2}{x^2}\right)^{-\frac{x^2}{h^2}}\right]^{-\frac{h^2}{x}}$$

pero $e = \lim\limits_{x\to\infty}\left(1 + \frac{1}{x}\right)^x$ y por la continuidad de $\ln(x)$ se tiene que $\lim\limits_{x\to\infty}\varphi(x) = 0$, por lo que φ está acotada cuando $x \to \infty$ y así para encontrar su valor máximo, encontraremos sus puntos críticos estacionarios, pero $\varphi'(x) = \frac{1}{h}\ln\frac{|x^2-h^2|}{x^2}$ de donde, no existe la derivada cuando $x = 0, h$ con $h > 0$ y $\varphi'(x) = 0$ si y sólo si $|x^2 - h^2| = x^2$, de donde si $x \geq h$ con $h > 0$ no hay solución, y si $0 < x < h$ entonces $h^2 - x^2 = x^2$ de donde $x = \frac{\sqrt{2}h}{2}$ pero como $\varphi(x) = 0$ cuando $x = 0$ y $x \to \infty$ se tiene que el valor máximo de φ es $\varphi\left(\frac{\sqrt{2}h}{2}\right) = \ln(3 + 2\sqrt{2}) \approx 1.8 < 2$ y así tenemos que $|\varphi(x)| \leq 2$ para toda $x \in \mathbb{R}$ de donde se sigue que $f \in \mathcal{Z}$.

Ahora probaremos el siguiente:

(VIII.11) Teorema. (A. Zygmund (1945))*Sea $f \in PC([-\pi,\pi])$ entonces $f \in \mathcal{Z}$ si y sólo si $\widehat{E}_n(f) \leq \frac{c}{n}$ para toda $n \in \mathbb{N}$.*

Demostración. Sea $f \in \mathcal{Z}$, $J_n(\theta)$ el núcleo de Jackson. Es fácil verificar que $\displaystyle\int_{-\pi}^{\pi} f(x-\theta)J_n(\theta)d\theta = \int_{-\pi}^{\pi} f(x+\theta)J_n(\theta)d\theta$, por lo que se tiene que:

$f(x) - \displaystyle\int_{-\pi}^{\pi} f(x-\theta)J_n(\theta)d\theta = f(x) - \int_{-\pi}^{\pi} f(x+\theta)J_n(\theta)$, entonces se sigue que:

$$|f(x) - \int_{-\pi}^{\pi} f(x-\theta)J_n(\theta)d\theta| = \frac{1}{2}|\int_{-\pi}^{\pi}(2f(x) - f(x-\theta) - f(x+\theta))J_n(\theta)d\theta|$$

$$\leq \tfrac{c}{2}\int_{-\pi}^{\pi}|\theta|J_n(\theta)d\theta = c\int_{0}^{\pi}\theta J_n(\theta)d\theta \leq \frac{c_2}{n}$$

ya que $\displaystyle\int_{-\pi}^{\pi}\theta J_n(\theta)d\theta \leq \frac{k}{2}$ con k una constante (ver (VII.11)) y

$\displaystyle\int_{-\pi}^{\pi} f(x-\theta)J_n(\theta)d\theta \in \mathcal{T}_n$ por lo tanto $\widehat{E}_n(f) \leq \frac{c}{n}$.

Para el recíproco supongamos que $\widehat{E}_n(f) \leq \frac{c}{n}$. Procediendo como antes sabemos que $f(x) = \displaystyle\sum_{n=0}^{\infty} Q_n(x)$ donde Q_n es de grado$\leq 2^n$, por lo que tenemos que $||Q_n|| \leq \frac{c}{2^n}$ **(1)**.

Por lo tanto $|Q_n(x+h) - 2Q_n(x) + Q_n(x-h)| = |h||Q_n'(\xi_1) - Q_n'(\xi_2)|$ con $|\xi_1|, |\xi_2| \leq h$, de donde $|\xi_1 - \xi_2| \leq 2h$. Por la desigualdad de Bernstein junto con **(1)** se sigue que:

$$|Q_n(x+h) - 2Q_n(x) + Q_n(x-h)| \leq 2h^2||Q''|| \leq 2h^2(2^n)^2||Q|| < 2h^2c2^n$$

entonces dada $n \in \mathbb{N}$ se tiene que:

$$|f(x+h) - 2f(x) + f(x-h)| \leq \left|\sum_{n=0}^{\infty} Q_k(x+h) - 2Q_k(x) + Q_k(x-h)\right| \leq$$

$$\leq \sum_{k=0}^{n}|Q_k(x+h) - 2Q_k(x) + Q_k(x-h)| + 4\sum_{k=n+1}^{\infty}||Q_k|| \leq 2h^2c\sum_{k=0}^{n}2^k +$$

$$+4c\sum_{k=n+1}^{\infty}\frac{1}{2^k} \leq 2h^2c2^{n+1} + \frac{4c}{2^n}.$$

eligiendo $h \in (0,1]$, escogemos n tal que $\frac{1}{h} \leq 2^n \leq \frac{2}{h}$ por lo que se tiene que $|f(x+h) - 2f(x) + f(x-h)| \leq 12ch$, por lo tanto $f \in \mathcal{Z}$.

Q.E.D.

3. Aproximación en Lip(α)

Sea $\alpha \in (0,1)$ y $f \in \mathrm{Lip}(\alpha)$, $f\colon [a,b] \to \mathbb{R}$ si $\alpha' < \alpha$ entonces $f \in \mathrm{Lip}(\alpha')$. Por otro lado sabemos que si $f \in \mathrm{Lip}(\alpha)$ tenemos que $\widehat{E}_n(f) \leq \frac{c}{n^\alpha}$ de donde se tiene que $n^{\alpha'}\widehat{E}_n(f) \leq \frac{n^{\alpha'}c}{n^\alpha}$ y cuando $n \to \infty$ tenemos que para toda $\alpha' \in (0,\alpha)$ existe $f \in \mathrm{Lip}(\alpha')$ tal que $n^{\alpha'}\widehat{E}_n(f) \to \infty$ cuando $n \to \infty$., por lo anterior nos preguntamos: ¿Si $\frac{c}{n^\alpha}$ es el mejor orden de aproximación para toda $f \in \mathrm{Lip}(\alpha)$?, cuya respuesta es **si**, para lo cuál probaremos el siguiente:

(VIII.12) Teorema. *Sea $\alpha \in (0,1)$ entonces existe una función $f \in \mathrm{Lip}(\alpha)$ tal que $n^\alpha \widehat{E}_n(f) \geq c(\alpha) > 0$ con $c(\alpha)$ una constante.*

Demostración. Sea $f(x) = \displaystyle\sum_{k=1}^{\infty} \frac{\cos(A^k x)}{A^{\alpha k}}$ donde $A \in \mathbb{N}$ es tal que $A^\alpha > 2$, primero probaremos que $f \in \mathrm{Lip}(\alpha)$, dada $h > 0$ se tiene que:

$\cos\left(A^k(x+h)\right) - \cos(A^k x) = -A^k h \operatorname{sen}\xi$, para alguna ξ entre x y $x+h$, de donde $|\cos\left(A^k(x+h)\right) - \cos(A^k x)| \leq A^k h$.

Sea $N \in \mathbb{N}$ entonces:

$$|f(x+h) - f(x)| = \left|\sum_{k=1}^{\infty} \frac{\cos\left(A^k(x+h)\right) - \cos(A^k x)}{A^{\alpha k}}\right| \leq$$

$$\leq \sum_{k=1}^{N} \frac{|\cos\left(A^k(x+h)\right) - \cos(A^k x)|}{A^{\alpha k}} + \sum_{k=N+1}^{\infty} \frac{|\cos\left(A^k(x+h)\right) - \cos(A^k x)|}{A^{\alpha k}} \leq$$

$$\leq \sum_{k=1}^{N} \frac{A^k h}{A^{\alpha k}} + \sum_{k=N+1}^{\infty} \frac{2}{A^{\alpha k}} = h\left(\frac{A^{(N+1)(1-\alpha)}}{A^{(1-\alpha)}-1} - A^{1-\alpha}\right) + A^{-N\alpha}\left(\frac{2}{A^\alpha - 1}\right) \leq$$

$$\leq \tfrac{A^{(1-\alpha)}}{A^{(1-\alpha)}-1} h A^{N(1-\alpha)} + A^{-N\alpha}\left(\tfrac{2}{A^\alpha-1}\right) \leq$$

$$\leq c_1(hA^{N(1-\alpha)} + A^{-Nk}) \text{ donde } c_1 = \max\{\tfrac{A^{(1-\alpha)}}{A^{(1-\alpha)}-1}, \tfrac{2}{A^\alpha - 1}\}.$$

Elegimos $N \in \mathbb{N}$ tal que $\frac{1}{h} \leq A^N < \frac{A}{h}$ por lo que tenemos que:

$A^{N(1-\alpha)} < \left(\frac{A}{h}\right)^{1-\alpha}$ y $A^{-N} < h$ entonces $A^{-N\alpha} < h^\alpha$ por lo tanto:

$|f(x+h) - f(x)| \leq (c_1 A^{1-\alpha} + 1)h^\alpha$ por lo que $f \in \mathrm{Lip}(\alpha)$.

Ahora probaremos que f no puede aproximarse mejor que $\frac{c}{n^\alpha}$. Sea $T_{A^n-1} \in \mathcal{T}_{A^n-1}$ entonces:

$$f(x) - T_{A^n-1}(x) = Q(x) + \tfrac{\cos(A^n x)}{A^{\alpha n}} + \sum_{k=n+1}^{\infty} \frac{\cos(A^k x)}{A^{k\alpha}}$$

donde $Q(x) \in \mathcal{T}_{A^n-1}$, por otro lado observemos que:

$$\left|\sum_{k=n+1}^{\infty}\frac{\cos(A^k x)}{A^{k\alpha}}\right| \leq \sum_{k=n+1}^{\infty}\frac{1}{A^{k\alpha}} = \frac{1}{A^{\alpha n}(A^{\alpha}-1)}$$

mientras que $\left|\frac{\cos(A^k x)}{A^{k\alpha}}\right| = \frac{1}{A^{\alpha n}}$ en $2A^n$ distintos puntos de $[0, 2\pi)$ **(1)**.

En esos puntos tenemos:

$$\left|\frac{\cos(A^n x)}{A^{n\alpha}} + \sum_{k=n+1}^{\infty}\frac{\cos(A^k x)}{A^{k\alpha}}\right| \geq \frac{1}{A^{\alpha n}} - \frac{1}{A^{\alpha n}(A^{\alpha}-1)} = \frac{c}{A^{\alpha n}}$$

donde $c = \frac{A^{\alpha}-2}{A^{\alpha}-1}$ **(2)**.

Dado que $Q(x) \in \mathcal{T}_{A^n-1}$, entonces $Q(x)$ tiene a lo más $2(A^n - 1)$ ceros en $[0, 2\pi]$. Por lo tanto, en uno de lo lugares donde se cumple **(1)**, $Q(x)$ debe tener el mismo signo de $\frac{\cos(A^k x)}{A^{k\alpha}}$. Por lo tanto para todo $T \in \mathcal{T}_{A^n-1}$ en ese punto se tiene que:

$$\|f - T_{A^n-1}\| \geq \left|Q(x) + \frac{\cos(A^n x)}{A^{\alpha n}} + \sum_{k=n+1}^{\infty}\frac{\cos(A^k x)}{A^{k\alpha}}\right| \geq$$

$$\geq \left|Q(x) + \frac{\cos(A^n x)}{A^{\alpha n}}\right| - \left|\sum_{k=n+1}^{\infty}\frac{\cos(A^k x)}{A^{k\alpha}}\right| \geq$$

$$\geq \left|\frac{\cos(A^n x)}{A^{\alpha n}}\right| - \left|\sum_{k=n+1}^{\infty}\frac{\cos(A^k x)}{A^{k\alpha}}\right| \geq \frac{1}{A^{\alpha n}}c.$$

Ahora sea $m \in \mathbb{N}$ tal que $A^{n-1} \leq m \leq A^n - 1$ entonces $\widehat{E}_m(f) \geq \widehat{E}_{A^n-1}(f) \geq$ $\geq \frac{c}{(Am)^{\alpha}} \geq \frac{c(\alpha)}{m^{\alpha}}$ donde $c(\alpha) = \frac{c}{A^{\alpha}}$. Por lo que se tiene que $\widehat{E}_m(f) \geq \frac{c(\alpha)}{m^{\alpha}}$.

Q.E.D.

CAPITULO IX

MEJOR APROXIMADOR CON OTRAS FAMILIAS

En esté capítulo haremos un breve estudio de la Teoría de aproximación con familias. Probaremos los teoremas de Müntz (1914), los cuales nos garantizan bajo que condiciones se puede aproximar uniformemente a funciones en $C([-1,1])$ mediante combinaciones lineales finitas de la forma $\sum \lambda_i x^{p_i}$ con $p_i > -\frac{1}{2}$.

§1. Teoremas de Müntz

Los Teoremas de Müntz son de gran interés por que relacionan dos hechos que aparentan no estarló, es decir, que el conjunto $< 1, x, x^2 \ldots >$ sea denso en $C([0,1])$ y que la serie de los exponentes recíprocos, $1 + \frac{1}{2} + \frac{1}{3} + \cdots$ diverja. Esta es la propiedad que debe tener cualquier subconjunto $\{x^{p_1}, x^{p_2}, \ldots\}$ con $p_i > 1\frac{1}{2}$ para que sea denso en $C([0,1])$.

Sea $(V, < >)$ un espacio vectorial sobre $\mathbb{R}$ o $\mathbb{C}$ con producto interior (e.v.i.), de aquí en adelante.

(IX.1) Teorema. *Sea $\{x_1, \ldots, x_n\}$ un conjunto linealmente independiente de V y $\{x_1^*, x_2^*, \ldots, x_n^*\}$ el conjunto ortonormal asociado a $\{x_1, \ldots, x_n\}$. Entonces para toda $y \in V$ se tiene que:*

$$\left\langle y - \sum_{k=1}^{n} < y, x_k^* > x_k^*, x_j^* \right\rangle = 0 \text{ para toda } j = 1, \ldots, n.$$

Demostración.

$$\left\langle y - \sum_{k=1}^{n} < y, x_k^* > x_k^* , x_j^* \right\rangle =< y, x_j^* > - < y, x_j^* >< x_j^*, x_j^* >= 0.$$

Q.E.D.

Si consideramos el espacio vectorial de dimensión finita $W = \langle\{x_1, \ldots, x_n\}\rangle$ entonces dado $y \in V$ existe $w^* \in W$ único tal que $\min_{w \in W}\{||y-w||\} = ||y-w^*||$. Es fácil ver que $w^* = \sum_{k=1}^{n} < y, x_k^* > x_k^*$, así como obtener el error mínimo, concretamente tenemos el siguiente:

(IX.2) Teorema. *Sea* $\{x_1^*, x_2^*, \ldots, x_n^*\}$ *un sistema ortonormal y* $y \in V$ *entonces* $\|y - \sum_{i=1}^{n} < y, x_i^* > x_i^*\| \leq \|y - \sum_{i=1}^{n} a_i x_i^*\|$ *para cualesquiera* $a_1, \ldots, a_n \in \mathbb{C}$.

Demostración.

$$\|y - \sum_{i=1}^{n} a_i x_i^*\|^2 = \left\langle y - \sum_{i=1}^{n} a_i x_i^*, y - \sum_{i=1}^{n} a_i x_i^* \right\rangle$$

$$=< y, y > - \sum_{i=1}^{n} a_i < x_i^*, y > - \sum_{i=1}^{n} \overline{a}_i < x_i^*, y > + \sum_{i=1}^{n} |a_i|^2 =$$

$$=< y, y > - \sum_{i=1}^{n} < x_i^*, y >< y, x_i^* > + \sum_{i=1}^{n} < x_i^*, y >< y, x_i^* > -$$

$$- \sum_{i=1}^{n} a_i < x_i^*, y > - \sum_{i=1}^{n} \overline{a}_i < x_i^*, y > + \sum_{i=1}^{n} |a_i|^2 =$$

$$=< y, y > - \sum_{i=1}^{n} | < x_i^*, y > |^2 + \sum_{i=1}^{n} |a_i - < x_i^*, y > |^2.$$

Como los dos primeros términos son independientes de las a_i's entonces el mínimo valor se alcanza si y sólo si $a_i =< y, x_i^* >$.

Q.E.D.

(IX.3) Corolario.

$$\min_{a_i} \|y - \sum_{i=1}^{n} a_i x_i\|^2 = \|y\|^2 - \sum_{i=1}^{n} | < y, x_1 > |^2.$$

(IX.4) Teorema. *Sea* $w^* = \sum_{i=1}^{n} a_i x_i$ *el mejor aproximador de* $y \in V$ *en* $\langle \{x_1, \ldots, x_n\} \rangle$ *con* $\{x_1, \ldots, x_n\} \subset V$, *linealmente independiente entonces los coeficientes de* w^* *son solución al siguiente sistema:*

$$\begin{array}{l} a_1 < x_1, x_1 > + a_2 < x_2, x_1 > + \ldots + a_n < x_n, x_1 >=< y, x_1 > \\ a_1 < x_1, x_2 > + a_2 < x_2, x_2 > + \ldots + a_n < x_n, x_2 >=< y, x_2 > \\ \\ a_1 < x_1, x_n > + a_2 < x_2, x_n > + \ldots + a_n < x_n, x_n >=< y, x_n > \end{array}$$

A estás ecuaciones se les conoce como las **ecuaciones normales.**

Demostración. Por (IX.1) tenemos que $y - w^*$ es ortogonal a cada x_i por lo que $< y - w^*, x_i >= 0$ de donde se obtiene la i-ésima ecuación del sistema de ecuaciones normales.

Q.E.D.

Ahora introducimos la siguiente:

(IX.5) Definición. Sean $\{x_1, \ldots, x_n\} \subset V$. La matriz de tamaño $n \times n$ definida por:

$$G = (< x_i, x_j >) = \begin{pmatrix} < x_1, x_1 > & \ldots & < x_1, x_n > \\ < x_2, x_1 > & \ldots & < x_2, x_n > \\ & & \\ < x_n, x_1 > & \ldots & < x_n, x_n > \end{pmatrix}$$

es conocida como la *matriz de Gram*. Denotaremos a su determinante como $g = g(x_1, \ldots, x_n) = |< x_i, x_j >|$, el cuál se conoce cómo el *determinante de Gram*.

(IX.6) Teorema. *Sea $\{x_1, \ldots, x_n\} \subset V$ linealmente independiente. Si $\delta = \min_{a_i}\{||y - \sum_{i=1}^{n} a_i x_i||\}$ entonces $\delta^2 = \frac{g(x_1,\ldots,x_n,y)}{g(x_1,\ldots,x_n)}$.*

Demostración. Sea $w^* = \sum_{i=1}^{n} a_i x_i$ el vector que minimiza a $||y - \sum_{i=1}^{n} a_i x_i||$ entonces $\delta^2 = ||y - w^*||^2 =< y - w^*, y > - < y - w^*, w^* >$ por (IX.1) se tiene que: $\delta^2 =< y - w^*, y >=< y, y > - < w^*, y >$ de donde $< w^*, y >=< y, y > -\delta^2$ **(1)**, ahora escribiendo las ecuaciones normales y agregando **(1)** tenemos:

$$\begin{array}{c} a_1 < x_1, x_1 > + a_2 < x_2, x_1 > + \ldots + a_n < x_n, x_1 > - < y, x_1 >= 0 \\ a_1 < x_1, x_2 > + a_2 < x_2, x_2 > + \ldots + a_n < x_n, x_2 > - < y, x_2 >= 0 \\ \\ a_1 < x_1, x_n > + a_2 < x_2, x_n > + \ldots + a_n < x_n, x_n > - < y, x_n >= 0 \\ a_1 < x_1, y > + a_2 < x_2, y > + \ldots + a_n < x_n, y > + \delta^2 - < y, y >= 0 \end{array}$$

Si hacemos $a_{n+1} = 1$ el coeficiente de la última columna, entonces el sistema anterior es un sistema homógeneo con (n+1) ecuaciones con (n+1) variables $(a_1, \ldots, a_n, a_{n+1} = 1)$, el cuál posee una solución *no trivial* (que son los coeficientes de w^*) por lo que el determinante del sistema se anula:

$$0 = \begin{vmatrix} < x_1, x_1 > & \ldots & < x_n, x_1 > & 0 - < y, x_1 > \\ < x_1, x_2 > & \ldots & < x_n, x_2 > & 0 - < y, x_2 > \\ & & & \\ < x_1, x_n > & \ldots & < x_n, x_n > & 0 - < y, x_n > \\ < x_1, y > & \ldots & < x_n, y > & \delta^2 - < y, y > \end{vmatrix}$$

entonces se tiene que:

$$\begin{vmatrix} <x_1,x_1> & \dots & <x_n,x_1> & 0 \\ <x_1,x_2> & \dots & <x_n,x_2> & 0 \\ \\ <x_1,x_n> & \dots & <x_n,x_n> & 0 \\ <x_1,y> & \dots & <x_n,y> & \delta^2 \end{vmatrix} =$$

$$\begin{vmatrix} <x_1,x_1> & \dots & <x_n,x_1> & <y,x_1> \\ <x_1,x_2> & \dots & <x_n,x_2> & <y,x_2> \\ \\ <x_1,x_n> & \dots & <x_n,x_n> & <y,x_n> \\ <x_1,y> & \dots & <x_n,y> & <y,y> \end{vmatrix}$$

por lo que $\delta^2 g(x_1,\dots,x_n) = g(x_1,\dots,x_n,y)$ y como $\{x_1,\dots,x_n\}$ es linealmente independiente se tiene que $g(x_1,\dots,x_n) \neq 0$ de donde se sigue el resultado.

Q.E.D.

Probaremos el siguiente resultado que nos será útil en el cálculo de un error mínimo

(IX.7) Lema. (Cauchy (1841)) *Sean $a_1,\dots,a_n$ y $b_1,\dots,b_n$ números dados tales que $a_i + b_j \neq 0$ para toda i, j entonces si:*

$$D_n = \begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \dots & \frac{1}{a_1+b_n} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \dots & \frac{1}{a_2+b_n} \\ \\ \frac{1}{a_n+b_1} & \frac{1}{a_n+b_2} & \dots & \frac{1}{a_n+b_n} \end{vmatrix}$$

se tiene que:

$$D_n = \frac{\prod\limits_{i>j}^{n}(a_i - a_j)(b_i - b_j)}{\prod\limits_{i,j=1}^{n}(a_i + b_j)}$$

Demostración. Considerando a las a_i's y b_j's como 2n variables independientes y pensando a D_n desarrollado, encontramos como común denominador a $\prod\limits_{i,j=1}^{n}(a_i + b_j)$.

Cada término individual es de la forma $\frac{1}{a_i+b_j}$ y es de grado -1 y así D_n es de grado $-n$. Observemos que el común denominador tiene n^2 factores por lo que es de grado n^2 y por lo anterior se sigue que el denominador debe ser un polinomio de grado $n^2 - n$ en las variables a_i's y b_j's. Notemos que si $a_i = a_j$ o $b_i = b_j$ entonces el i-ésimo y j-ésimo renglón son iguales de donde $D_n = 0$, de manera que el numerador debe tener factores de la forma $\prod_{i>j}^{n}(a_i - a_j)\prod_{i>j}^{n}(b_i - b_j)$.

Es fácil ver que cada producto tiene $1 + \cdots + (n-1) = \frac{n(n-1)}{2}$ factores por lo que producto completo tiene $n(n-1)$, por lo tanto tenemos que:

$$D_n = c_n \frac{\prod_{i>j}^{n}(a_i - a_j)(b_i - b_j)}{\prod_{i,j=1}^{n}(a_i + b_j)} \quad \textbf{(1)}$$

donde c_n es una constante independiente de las a_i's y b_j's. Por lo que sólo resta probar que $c_n = 1$ para toda n.

Observemos que:

$$a_n D_n = \begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n} \\ \frac{a_n}{a_n+b_1} & \frac{a_n}{a_n+b_2} & \cdots & \frac{a_n}{a_n+b_n} \end{vmatrix}$$

entonces:

$$\lim_{a_n \to \infty} a_n D_n = \begin{vmatrix} \frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 1 & 1 & \cdots & 1 \end{vmatrix}$$

y así tenemos que:

$$\lim_{b_n \to \infty} \lim_{a_n \to \infty} a_n D_n = \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_{n-1}} & 0 \\ \frac{1}{a_{n-1}+b_1} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} & 0 \\ \cdot & \cdots & \cdot & \cdot \\ 1 & \cdots & 1 & 1 \end{vmatrix} = D_{n-1}$$

por lo tanto $\lim_{b_n \to \infty} \lim_{a_n \to \infty} \frac{a_n D_n}{D_{n-1}} = 1$. Por **(1)** se tiene que:

$$\frac{a_nD_n}{D_{n-1}}=\frac{a_nc_n}{c_{n-1}}\frac{\prod\limits_{i>j}^{n}(a_i-a_j)(b_i-b_j)}{\prod\limits_{i,j=1}^{n}(a_i+b_j)}\frac{\prod\limits_{i,j=1}^{n-1}(a_i+b_j)}{\prod\limits_{i>j}^{n-1}(a_i-a_j)(b_i-b_j)}=$$

$$=\frac{a_nc_n}{c_{n-1}}\frac{\prod\limits_{j=1}^{n-1}(a_n-a_j)\prod\limits_{j=1}^{n-1}(b_n-b_j)}{\prod\limits_{j=1}^{n}(a_n+b_j)\prod\limits_{j=1}^{n-1}(b_n+a_j)}$$

de donde se sigue que $\lim\limits_{b_n\to\infty}\lim\limits_{a_n\to\infty}\dfrac{a_nD_n}{D_{n-1}}=1=\dfrac{c_n}{c_{n-1}}$.

Por lo que $c_n=c_{n-1}$ y como $c_1=1$ se tiene que $c_n=1$ para toda $n\in\mathbb{N}$.

Q.E.D.

(IX.8) Lema. *Sean $q,p_1,p_2,\ldots,p_n\in\mathbb{R}$, $p_i\neq p_j$ si $i\neq j$ y $p_i,q>-\frac{1}{2}$ entonces:*

$$\delta^2=\min_{a_k}\int_0^1|x^q-(a_1x^{p_1}+\ldots+a_nx^{p_n})|^2dx=\frac{1}{2q+1}\prod_{i=1}^{n}\left\{\frac{p_i-q}{p_i+q+1}\right\}^2.$$

Demostración. Por (IX.6) $\delta^2=\frac{g(x^{p_1},\ldots,x^{p_n},x^q)}{g(x^{p_1},\ldots,x^{p_n})}$ y por definición:

$$g(x^{p_1},\ldots,x^{p_n})=\begin{vmatrix}<x^{p_1},x^{p_1}> & \ldots & <x^{p_1},x^{p_n}>\\ <x^{p_n},x^{p_1}> & \ldots & <x^{p_n},x^{p_n}>\end{vmatrix}$$

donde $<f,g>=\int_0^1 fg$, por lo que $<x^{p_i},x^{p_j}>=\frac{1}{p_i+p_j+1}$ y así tenemos que:

$$g(x^{p_1},\ldots,x^{p_n})=\begin{vmatrix}\frac{1}{p_1+p_1+1} & \cdots & \frac{1}{p_1+p_n+1}\\ \frac{1}{p_n+p_1+1} & \cdots & \frac{1}{p_n+p_n+1}\end{vmatrix}$$

por (IX.7) con $a_i=p_i$ y $b_j=p_j+1$ se tiene que:

$$g(x^{p_1},\ldots,x^{p_n},x^q)=\frac{\prod\limits_{i>j}^{n}(p_i-p_j)^2}{\prod\limits_{i,j=1}^{n}(p_i+p_j+1)}\ \textbf{(1)}.$$

De manera análoga pero ahora considerando a $a_i=p_i$ y $b_i=p_i+1$ para toda $i=1,\ldots,n$ y además de $a_{n+1}=q$ y $b_{n+1}=q+1$ se tiene que:

$$g(x^{p_1},\ldots,x^{p_n},x^q)=\frac{\prod_{i>j}^{n+1}(a_i-b_j)(b_i-b_j)}{\prod_{i,j=1}^{n+1}(a_i+b_j)}$$

de donde:

$$g(x^{p_1},\ldots,x^{p_n},x^q)=\frac{\prod_{i>j}^{n}(a_i-a_j)(b_i-b_j)}{\prod_{i,j=1}^{n}(p_i+p_j+1)}\frac{\prod_{j=1}^{n}(a_{n+1}-a_j)(b_{n+1}-b_j)}{\prod_{j=1}^{n+1}(a_{n+1}+b_j)\prod_{j=1}^{n+1}(b_{n+1}+a_j)}=$$

$$=\frac{\prod_{i>j}^{n}(p_i-p_j)^2}{\prod_{i,j=1}^{n}(p_i+p_j+1)}\frac{\prod_{j=1}^{n}(q-p_j)^2}{\prod_{j=1}^{n+1}(q+b_j)\prod_{j=1}^{n+1}(q+1+a_j)}.$$

De aquí aplicando **(1)** se obtiene el resultado.

Q.E.D.

El siguiente resultado nos es útil para estudiar la convergencia de productos infinitos.

(IX.9) Lema. *Sea* $a_n>0$, $a_n\neq 1$ *y* $a_n\to 0$ *entonces* $\prod_{n=1}^{\infty}(1-a_n)=0$ *si y sólo si* $\sum_{n=1}^{\infty}a_n=+\infty$.

Demostración. Como $a_n\to 0$ existe $m(\varepsilon)\in\mathbf{N}$ tal que $a_n<\frac{1}{2}$ para toda $n\geq m$, para estas n's se sigue de la expansión en serie de Taylor para $\ln(1-x)$ que:

$\left|\frac{\ln(1-a_n)}{a_n}+1\right|=\left|\frac{a_n}{2}+\frac{a_n^2}{3}+\cdots\right|<$

$<\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^3+\cdots=\frac{1}{2}$,

así para $n\geq m$ entonces se tiene que $-\frac{3}{2}\leq\frac{\ln(1-a_n)}{a_n}\leq-\frac{1}{2}$ de donde se sigue que las series $\sum_{n=1}^{\infty}a_n$ y $\sum_{n=1}^{\infty}\ln(1-a_n)$ convergen o divergen simúltaneamente. Por lo tanto $\sum_{n=1}^{\infty}a_n=+\infty$ si y sólo si $\sum_{n=1}^{\infty}\ln(1-a_n)=\ln\left(\prod_{n=1}^{\infty}(1-a_n)\right)=-\infty$ de donde se sigue el resultado.

Q.E.D.

(IX.10) Definición. Un conjunto finito o infinito de elementos $\{x_1, x_2, \ldots\}$ en un espacio vectorial normado X es *cerrado ó fundamental* en $W \subset X$, si para toda $x \in W$ y para toda $\varepsilon > 0$ existen $a_i \in \mathbb{R}$ $i = 1, \ldots, n$ tales que $\|x - \sum_{i=1}^{n} a_i x_i\| < \varepsilon$.

(IX.11) Teorema. (Lauricella) *Sea X un e.v.n. y $\{x_n\}$ un conjunto cerrado en X entonces un segundo conjunto $\{y_n\}$ es cerrado en X si y sólo si $\{y_n\}$ es cerrado en $\{x_n\}$.*

Demostración. Supongamos que $\{y_n\}$ es cerrado en X y $x_n \in X$, entonces x_n se puede aproximar por una combinación lineal de y_n's tan cercana cómo se quiera, por lo que $\{y_n\}$ es cerrada en $\{x_n\}$.

Inversamente, sean $x \in X$ y $\varepsilon > 0$. Como $\{x_n\}$ es cerrado existen $a_1, \ldots, a_N$ tales que $\|x - \sum_{i=1}^{N} a_i x_i\| < \frac{\varepsilon}{2}$, supongamos que $a_i \neq 0$ para toda $i = 1, \ldots, N$ (si no es así simplemente los ignoramos). Dado que $\{y_n\}$ es cerrada en $\{x_n\}$ existen $b_{i1}, \ldots, b_{1N_i}$ tales que $\|x_i - (b_{i1}y_1 + \ldots + b_{iN_i}y_{N_i})\| < \frac{\varepsilon}{2N|a_i|}$ para toda $i = 1, \ldots N$.

Sea $w_i = a_i(b_{i1}y_1 + \ldots + b_{iN_i}y_{N_i})$ $i = 1, \ldots, N$ entonces:

$\|a_i x_i - w_i\| \leq \frac{\varepsilon}{2N}$ y por la desigualdad del triángulo tenemos que:

$$\|x - \sum_{i=1}^{n} w_i\| \leq \|x - \sum_{i=1}^{N} a_i x_i\| + \sum_{i=1}^{N} \|a_i x_i - w_i\| \leq \frac{\varepsilon}{2} + \sum_{i=1}^{N} \frac{\varepsilon}{2N}$$

por lo tanto $\|x - \sum_{i=1}^{N} w_i\| < \varepsilon$.

Q.E.D.

(IX.12) Definición. Una sucesión de elementos $\{x_k\}_{k=1}^{\infty}$ es *completa en V* si la única funcional lineal continua tal que $L(x_k) = 0$ para toda $k \in \mathbb{N}$, entonces $L \equiv 0$.

(IX.13) Teorema. *Sea $(V, \|\ \|)$ un e.v.n. Entonces $\{x_k\}$ es cerrado en V si y sólo si $\{x_k\}$ es completa en V.*

Demostración. Supongamos que $\{x_k\}$ es cerrada. Sea $L: V \to \mathbb{R}$ una funcional lineal, continua tal que $L(x_k) = 0$ para toda $k \in \mathbb{N}$. Dado $v \in V$ y $\varepsilon > 0$ existen $a_1, \ldots, a_n \in \mathbb{R}$ tales que $\|x - \sum_{i=1}^{n} a_i x_i\| < \varepsilon$, entonces:

$|L(x)| = |L(x - \sum_{i=1}^{n} a_i x_i)| \leq \|L\|\ \|x - \sum_{i=1}^{n} a_i x_i\| < \|L\|\varepsilon$ para toda $\varepsilon > 0$, por lo que $L(x) = 0$ y así $L \equiv 0$.

Inversamente si $L(x_k) = 0$ para toda $k \in \mathbb{N}$ implica que $L \equiv 0$. Sea $x_0 \in V$, $W = \langle \{x_n\} \rangle$ y $d = \inf_{y \in W} \|x_0 - y\|$, supongamos que $d > 0$ entonces por el teorema de Hahn-Banach (ver:[BN]) existe una funcional lineal, continua $L: X \to \mathbb{R}$ tal que $L(x_0) = 1$ y $L(y) = 0$ para toda $y \in W$, en partícular $L(x_k) = 0$ para toda $k \in \mathbb{N}$ de donde $L \equiv 0$ lo cuál no es posible.

Q.E.D.

(IX.14) Lema. *El conjunto* $\{1, x, x^2, \ldots\}$ *es cerrado en* $\mathcal{L}^2([a,b])$.

Demostración. Sea $f \in \mathcal{L}^2([a,b])$. Definimos $L: \mathcal{L}^2([a,b]) \to \mathbb{R}$ por $L(g) = \int_a^b f(x)g(x)dx$. Supongamos que $L(x^k) = 0$ para toda $k \in \mathbb{N} \cup \{0\}$. Por (IX.13) basta demostrar que $L \equiv 0$, lo cuál se hará probando que $f \equiv 0$ (c.d).

Consideremos $F(x) = \int_a^x f(t)dt$ entonces $F \in C([a,b])$ y $F(a) = F(b) = 0$, integrando por partes se obtiene que:

$$0 = \int_a^b x^n f(x)dx = x^n F(x)|_a^b - n\int_a^b x^{n-1}F(x)dx \text{ por lo que:}$$

$\int_a^b x^{n-1}F(x)dx = 0$, para toda $n \in \mathbb{N}$. Por el Teorema de Weierstrass tenemos que $\{1, x, x^2 \ldots\}$ es cerrado en $C([a,b])$ con la norma uniforme por lo que $F(x) \equiv 0$ y por lo tanto $f \equiv 0$ (c.d) y así $L \equiv 0$.

Q.E.D.

(IX.15) Observación. De acuerdo con el Teorema de Representación de Riesz *toda* funcional lineal L en $\mathcal{L}^2$ es de la forma indicada en el lema anterior.

(IX.16) Teorema. (Primer Teorema de Müntz en $\mathcal{L}^2$) (1914).

Sea $A = \{x^{p_1}, x^{p_2}, \ldots\}$ *un conjunto de funciones tales que* $p_n > -\frac{1}{2}$ $p_i \neq p_j$ *y* $\lim_{n\to\infty} p_n = +\infty$. *Entonces* A *es fundamental en* $\mathcal{L}^2([0,1])$ *si y sólo si* $\sum_{p_n \neq 0} \frac{1}{p_n} = +\infty$.

Demostración. Por (IX.11) y (IX.13) para probar que A es cerrado en $\mathcal{L}^2([0,1])$ basta demostrar que x^q es aproximable por $\{x^{p_1}, x^{p_2}, \ldots\}$ para toda $q \in \mathbb{N} \cup \{0\}$. Por lo que debemos encontrar $p_1, p_2, \ldots$ tales que:

$$\lim_{n\to\infty} \min_{a_i} \int_0^1 |x^q - \sum_{i=1}^n a_i x^{p_i}|^2 dx = 0$$

y por (IX.8) basta con demostrar que:

$$\lim_{n\to\infty} \prod_{i=1}^n \left\{ \frac{p_i - q}{p_i + q + 1} \right\}^2 = 0 \text{ } \mathbf{(1)}.$$

Supongamos que $\sum_{p_n \neq 0} \frac{1}{p_n} = +\infty$ entonces:

$$\prod_{i=1}^{n} \left\{ \frac{p_i - q}{p_i + q + 1} \right\}^2 = \frac{\prod_{i=1}^{n} \left\{ 1 - \frac{q}{p_i} \right\}^2}{\prod_{i=1}^{n} \left\{ 1 + \frac{q+1}{p_i} \right\}^2} \quad \textbf{(2)}$$

Por hipótesis $\lim_{n\to\infty} p_n = +\infty$ entonces $\frac{q}{p_i} \to 0$ y como $\sum_{p_n \neq 0} \frac{1}{p_n} = +\infty$ se sigue que $\sum_{p_n \neq 0} \frac{q}{p_n} = +\infty$ y por (IX.9) $\prod_{i=1}^{n} \left\{ 1 - \frac{q}{p_i} \right\}^2 = 0$.

Por otro lado $\prod_{i=1}^{n} \left\{ 1 + \frac{q+1}{p_i} \right\}^2 = +\infty$ ya que $\prod_{i=1}^{\infty}(1 + a_n)$ converge si y sólo si $\sum_{i=1}^{\infty} a_n$ converge también, por lo que **(1)** se cumple.

Ahora supongamos que A es cerrado en $\mathcal{L}^2([0,1])$ pero que $\sum_{p_n \neq 0} \frac{1}{p_n} < \infty$. Elegimos $q \neq p_1, p_2, \ldots$ entonces el numerador y denominador en **(2)** convergen a un valor distinto de cero, lo que significa que x^q no puede aproximarse por $x^{p_1}, x^{p_2}, \ldots$ contrario a nuestra suposición.

Q.E.D.

Otra versión del Teorema de Müntz es :

(IX.17) Teorema de Müntz

Sea $A = \{x^{p_1}, x^{p_2}, \ldots\}$ *un conjunto de funciones tales que* $p_n > -\frac{1}{2}$ $p_i \neq p_j$ *y* $\lim_{n\to\infty} p_n = -\frac{1}{2}$. *Entonces* A *es fundamental en* $\mathcal{L}^2([0,1])$ *si y sólo si* $\sum_{p_n \neq 0} (p_i + \frac{1}{2}) = +\infty$.

Demostración. Procediendo de manera análoga a (IX.16), si $\lim_{n\to\infty} p_i = -\frac{1}{2}$ y $\sum_{p_n \neq 0} (p_i + \frac{1}{2}) = +\infty$ tenemos que:

$$\prod_{i=1}^{n}\left\{\frac{p_i-q}{p_i+q+1}\right\}^2=\frac{\prod_{i=1}^{n}\left\{1-\frac{p_i+\frac{1}{2}}{q+\frac{1}{2}}\right\}^2}{\prod_{i=1}^{n}\left\{1+\frac{p_i+\frac{1}{2}}{q+\frac{1}{2}}\right\}^2}\quad(1)$$

para i suficientemente grande tenemos que: $0<\frac{p_i+\frac{1}{2}}{q+\frac{1}{2}}<1$ por lo que

$0<1-\frac{p_i+\frac{1}{2}}{q+\frac{1}{2}}<1$ de donde $\prod_{i=1}^{n}\left\{1-\frac{p_i+\frac{1}{2}}{q+\frac{1}{2}}\right\}^2$ permanece acotado y

$\prod_{i=1}^{n}\left\{1+\frac{p_i+\frac{1}{2}}{q+\frac{1}{2}}\right\}^2$ diverge a $+\infty$ por lo que se tiene que:

$$\lim_{n\to\infty}\prod_{i=1}^{n}\left\{\frac{p_i-q}{p_i+q+1}\right\}^2=0.$$

Ahora si A es cerrado y $\sum_{p_n\neq 0}(p_i+\frac{1}{2})<+\infty$ con $\lim_{n\to\infty}p_i=-\frac{1}{2}$.

Elegimos $q\neq p_1,p_2,\ldots$ entonces el numerador en (1) no converge a cero ya que ningún factor se anula pero $\sum_{p_n\neq 0}\frac{(p_i+\frac{1}{2})}{q+\frac{1}{2}}<+\infty$ entonces el numerador converge a un número finito positivo, de la misma forma el denominador converge a un número positivo por lo que x^q no se puede aproximar por $x^{p_1},x^{p_2},\ldots$ contrario a nuestra suposición.

Q.E.D.

(IX.18) Segundo Teorema de Münts. (Norma uniforme)

Sea A un conjunto de números reales **no negativos**. *Si $0\in A$ y A contiene un sucesión $\{p_i\}$ tal que $\lim_{i\to\infty}p_i=\infty$ y $\sum_{i=1}^{\infty}\frac{1}{p_i}=\infty$ entonces $\{x^p\}$ tal que $p\in A$ es cerrada en $C([0,1])$ con la norma uniforme.*

Demostración. Sea $n>0$, $p_i>0$ entonces:

$$\left|x^n-\sum_{i=1}^{N}a_ix^{p_i}\right|=n\left|\int_0^x\left(t^{n-1}-\sum_{i=1}^{N}\frac{a_ip_i}{n}t^{p_i-1}\right)dt\right|\leq$$

$$\leq n\int_0^x\left|t^{n-1}-\sum_{i=1}^{N}\frac{a_ip_i}{n}t^{p_i-1}\right|dt$$

$$\leq n\int_0^1\left|t^{n-1}-\sum_{i=1}^{N}\frac{a_ip_i}{n}t^{p_i-1}\right|dt$$

$$\leq n\left(\int_0^1\left|t^{n-1}-\sum_{i=1}^{N}\frac{a_ip_i}{n}t^{p_i-1}\right|^2dt\right)^{\frac{1}{2}}$$

por hipótesis $p_i \to \infty$ cuando $i \to \infty$ y $\frac{1}{2}\sum_{i=1}^{\infty}\frac{1}{p_i} < \sum_{i=1}^{\infty}\frac{1}{p_i-1}$ de donde se tiene que $\sum_{i=1}^{\infty}\frac{1}{p_i-1} = \infty$ entonces por (IX.17) $\{x^{p_1-1}, x^{p_2-1}, \ldots\}$ es cerrado en $\mathcal{L}^2([0,1])$.

Por la desigualdad anterior para $n-1 = 0,1,2,\ldots$ $\|x^n - \sum_{i=1}^{N} a_i x^{p_i}\|$ se puede hacer tan pequeña como se quiera para una elección apropiada de a_i's. Por lo que el conjunto $\{x^{p_1-1}, x^{p_2-1}, \ldots\}$ es cerrado en $\{x, x^2, \ldots\}$ y agregando 1, el conjunto $\{1, x^{p_1-1}, x^{p_2-1}, \ldots\}$ será cerrado en $\{1, x, x^2, \ldots\}$ y por (IX.11) se sigue el resultado.

Q.E.D.

§2 Aproximación Racional

En esta sección haremos un breve estudio, a la Teoría de aproximación con otras familias , muy especialmente con funciones racionales. Cómo es de esperar la situación en estos casos es más complicada que en los casos considerados con anterioridad. Para lo cuál veremos los siguientes:

(IX.19) Ejemplos.

(i) Resolvamos el problema de $\min_a \max\{|0 - e^{ax}| \mid 1 \leq x \leq 2\}$. Dada $a \in \mathbb{R}$ se tiene que:

$$\max_{1\leq x\leq 2} |e^{ax}| = \begin{cases} e^{a} & \text{si } a > 0 \\ 1 & \text{si } a = 0 \\ e^{2a} & \text{si } a < 0 \end{cases}$$

de donde $\min_a \max_{1\leq x\leq 2} |e^{ax}| = e^{2a}$ y observamos que tomando a valores cada vez más pequeños se observa que el valor mínimo se hace tan pequeño como se quiera, por lo que el valor mínimo *no* se alcanza, es decir el problema no tiene solución en $\mathbb{R}$.

(ii) Ahora consideremos $\min_a ||\frac{1}{2} - x^{a^2}||_{[0,1]}$, hacemos $f(x) = \frac{1}{2} - x^{a^2}$, f es creciente para toda $a \in \mathbb{R} \setminus \{0\}$ por lo tanto $||\frac{1}{2} - x^{a^2}||_{[a,b]} = f(0) = \frac{1}{2}$ y si $a = 0$, es claro que $||\frac{1}{2} - x^{a^2}|| = \frac{1}{2}$ de donde $\min_a ||\frac{1}{2} - x^{a^2}||_{[0,1]} = \frac{1}{2}$ y cualquier valor resuelve el problema.

De los ejemplos anteriores se observa que para algunas familias el mejor aproximador puede existir ó no y además no tiene por que ser único para una función continua.

Ahora veremos que dada una función $f \in C([a,b])$ se puede aproximar por medio de funciones racionales para lo cuál daremos antes la siguiente:

(IX.20) Definición. Se dice que un conjunto $\Im$ en $C_{p,q}(D)$ es *acotado uniformemente* en D si existe $M > 0$ tal que $||f||_D \leq M$ para toda $f \in \Im$.

(IX.21) Lema. *Sea $\Im$ una familia de polinomios de grados acotados. Si $\Im$ es acotada uniformemente en $[a,b]$ entonces sus coeficientes son acotados.*

Demostración. Supongamos lo contrario entonces existe $(p_k) \subset \Im$ una sucesión, tal que para algún $a_i^{(k)}$ coeficiente de p_k, $i \in \{0, \ldots, n\}$ se tiene que $a_i^{(k)} \to \infty$, cuando $k \to \infty$. Como los grados son acotados entonces existe $i_0 \in \{0, \ldots, n\}$ tal que $a_{i_0}^{(k)} \to \infty$, cuando $k \to \infty$ a partir de alguna $k_0 \in \mathbb{N}$. Sin pérdida de generalidad suponemos que $p_k(x) = a_n^{(k)}x^n + \ldots + a_0^{(k)}$, por hipótesis $|p_k(x)| \leq M$ con $M > 0$, de donde $|\frac{1}{a_{i_0}^{(k)}}p_k(x)| \leq \frac{M}{a_{i_0}^{(k)}}$ para toda $k \geq k_0$ (1). Pero $\frac{1}{a_{i_0}^{(k)}}p_k(x) =$

$= |b_n^{(k)}x^n + \ldots + 1 \cdot x^{i_0} + \ldots + b_0^{(k)}|$ y cuando $k \to \infty$ se tiene que (1) no es cierta.

Q.E.D.

(IX.22) Teorema. *Sean $f \in C([a,b])$ y $n, m \in \mathbb{N} \cup \{0\}$. Entonces el problema $\min_{a_i, b_i} ||f - R||_{[a,b]}$ tiene solución, donde $R(x) = \frac{p(x)}{q(x)}$ con $p(x) = \sum_{i=0}^{n} a_i x^{n-i}$ y $q(x) = \sum_{j=0}^{m} b_j x^{m-j}$.*

Demostración.

Supondremos que:

(i) $q(x)$ no se anula en $[a,b]$, lo cuál se puede realizar ya que si $q(x)$ se anula en un punto de $[a,b]$ entonces $||f - R||_{[a,b]}$ no esta acotada y esté caso quedaría descartado para resolver el problema.

(ii) Haremos que $b_0^2 + \ldots + b_m^2 = 1$ ya que cómo $q \neq 0$ se puede considerar a $\widehat{q} = \frac{1}{B}q(x)$ con $B = \sum(b_j)^2 \neq 0$.

Por lo que ahora sea $\Delta = \inf_{a_i, b_j} ||f - R||_{[a,b]}$, por lo tanto existe una sucesión de funciones racionales $R_k(x) = \frac{A_k(x)}{B_k(x)}$ con $A_k(x) = \sum_{i=0}^{n} a_i^{(k)} x^{n-i}$, $B_k(x) = \sum_{j=0}^{m} b_j^{(k)} x^{m-j}$ tales que $\Delta_k = ||f - R_k||_{[a,b]}$ y además $\lim_{k\to\infty} \Delta_k = \Delta$.

Por la condición (ii) los coeficientes $b_j^{(k)}$ son acotados uniformemente lo cuál implica que los coeficientes $a_i^{(k)}$ son acotados debido a que: $|f(x) - R_k(x)| \leq \Delta_k$ por lo que tenemos $|R_k(x)| \leq \Delta_k + |f(x)| \leq \Delta_k + ||f||_{[a,b]} \leq M$, $M > 0$ (ya que Δ_k es acotada por ser convergente), por lo tanto $|R_k(x)| \leq M$ para toda $k \in \mathbb{N}$. De la definición de R_k se tiene que: $|A_k(x)| \leq M|B_k(x)|$ (1).

$$\text{Pero } |B_k(x)| = \sum_{j=0}^{m} b_j^{(k)} x^{m-j} \leq \left(\sum_{j=0}^{m}\left(b_j^{(k)}\right)^2\right)^{\frac{1}{2}} \left(\sum_{j=0}^{m} x^{2(m-j)}\right)^{\frac{1}{2}} \leq$$

$$\leq \left(\sum_{j=0}^{m} x^{2(m-j)}\right)^{\frac{1}{2}} \leq ||F||_{[a,b]} \text{ con } F(x) = \left(\sum_{j=0}^{m} x^{2(m-j)}\right)^{\frac{1}{2}}$$

de donde $B_k(x)$ esta acotada uniformemente en $[a,b]$ para toda $k \in \mathbb{N}$ y de (1) se tiene que $A_k(x)$ también lo esta, por (IX.21) se tiene que los $a_i^{(k)}$ y $b_j^{(k)}$ son acotados. Consideremos los puntos $p_k = (a_0^{(k)}, a_1^{(k)}, \ldots, a_n^{(k)}, b_0^{(k)}, \ldots b_m^{(k)})$ en $\mathbb{R}^{n+m+2}$, los cuales son acotados, entonces por **TWB** existe una subsucesión (p_{k_i}) que converge a un punto

$p' = (a'_0, \ldots, a'_n, b'_0, \ldots, b'_m)$.

Considerando a R_{k_i} las funciones racionales correspondientes a los subíndices (p_{k_i}), y reordenando los índices de la subsucesión tenemos:

$\lim\limits_{k\to\infty} a_i^{(k)} = a'_i$ para $i = 0, \ldots, n$ y $\lim\limits_{k\to\infty} b_j^{(k)} = b'_j$ para $j = 0, \ldots, m$.

Sea $R'(x) = \frac{p'(x)}{q'(x)}$ donde $p'(x) = \sum\limits_{i=0}^{n} a'_i x^{n-i}$ y $q'(x) = \sum\limits_{j=0}^{m} b'_i x^{m-j}$, probaremos que $||f - R'|| = \Delta$, dado que R' tiene a lo más un número finito de ceros en el denominador, dada $x \in [a,b]$ tal que $q'(x) \neq 0$ tenemos que $R_k(x) \to R'(x)$ cuando $k \to \infty$ por lo que $R'(x) = f(x) + R_k(x) - f(x) + R'(x) - R_k(x)$ y así $|R'(x)| \leq$ $\leq |f(x)| + |R_k(x) - f(x)| + |R'(x) - R_k(x)|$ de donde se tiene que $|R'(x)| \leq ||f||_{[a,b]} +$ $+ \Delta_k + |R'(x) - R_k(x)|$. Como (Δ_k) converge existe $M = \sup \Delta_k$ y haciendo $k \to \infty$ tenemos que $|R'(x)| \leq M + ||f||_{[a,b]}$ por lo que $|f(x) - R'(x)| \leq |f(x) - R_k(x)| +$ $+ |R_k(x) - R'(x)| \leq \Delta_k + |R_k(x) - R'(x)|$, haciendo $k \to \infty$ tenemos que $||f(x) - R'(x)||_{[a,b]} \leq \Delta$ con $x \in [a,b]$ tal que $q'(x) \neq 0$.

Ahora si $q'(x) = 0$ encontramos una sucesión de puntos (x_n) en $[a,b]$ tales que $x_n \to x$ y $q'(x_i) \neq 0$ entonces $|f(x_i) - R'(x_i)| \leq \Delta$ y por continuidad $|f(x) - R(x)| \leq$ $\leq \Delta$. Por lo tanto $||f - R'||_{[a,b]} \leq \Delta$.

Q.E.D.

LISTA DE SIMBOLOS

$C([a,b])$ conjunto de funciones continuas del intervalo $[a,b]$ en $\mathbb{R}$.

$||f||_D$ norma uniforme sobre el conjunto D.

$\mathcal{P}_n$ polinomios de grado $\leq$n.

$A(D)$ conjunto de funciones acotadas de D en $\mathbb{R}$.

TVM-D Teorema del valor medio para la derivada.

TVM-I Teorema del valor medio para la integral.

$C^p([a,b])$ conjunto de funciones p-derivables en $[a,b]$ y con cada una de sus derivadas continuas.

$\mathbf{F}_{p,q}(D)$ conjunto de funciones sobre el conjunto D.

$C_{p,q}(D)$ conjunto de funciones continuas de $D \subset \mathbb{R}^p$ a $\mathbb{R}^q$.

$\mathcal{M}_{m\times n}(R)$ conjunto de matrices de tamaño $m \times n$ con coeficientes en R.

$B_r(x)$ Bola abierta en $\mathbb{R}^n$.

$\overline{B}_r(x)$ Bola cerrada en $\mathbb{R}^n$.

$\mathbb{R}[x]$ polinomios con coeficientes reales en la indeterminada x.

$\mathbb{Q}[x]$ polinomios con coeficientes racionales en la indeterminada x.

$\mathbb{R}[x_1,\ldots,x_n]$ polinomios con coeficientes reales en las indeterminadas $x_1,\ldots,x_n$.

$C^K_{p,1}(D)$ conjunto de funciones continuas sobre D con valores en K.

$\widehat{\mathcal{P}}_n$ conjunto de polinomios de grado n.

TWB Teorema de Bolzano-Weierstrass.

$\mathcal{T}_n$ conjunto de polinomios trigonométricos de grado $\leq$n.

BIBLIOGRAFIA

1. [BN] Bachman G. y Narici L. Functional Analysis, Academic Press, 1966.

2. [Ba] Bartle, R.G. Introducción al Análisis Matématico, Limusa, 1989.

3. [Ch] Cheney, E.W. Introduction to Approximation Theory, Chelsea, 1981.

4. [Da] Davis, P.J. Interpolation and Approximation, Dover, 1975.

5. [FD] Feinerman, R.P. Newman, D.J. Polynomial Approximation,
The Williams and Wilkins Company, 1974.

6. [Ha] Hasser, B.N. Real Analysis, Dover,1991.

7. [Ko] Korovkin, P.P. Linear Operators and Approximation Theory,
Hindustan Publishing Corp. (India), 1960.

8. [Kn] Knopp, K. Theory and application of infinite Series, Dover, 1990.

9. [Ku] Kudriávstev, L.D. Curso de Análisis Matématico, Vol. 2,
Editorial Mir, 1984.

10. [Lo] Lorentz, G.G. Bernstein polynomials, Toronto, Chelsea, 1953.

11. [Ri1] Rivlin, T.J. Chebyshev Polynomials, 2nd. edición, John Wiley
and sons, 1990.

12. [Ri2] Rivlin, T.J. An introduction to Approximation of Funtions,
Dover, 1969.

13. [Ru] Rudin. W. Principles of mathematical Analysis, 2nd. edición,
Mc. Graw Hill,1964.