3ª PARTE: APROXIMACION EN ESPACIOS NORMADOS.

TEMA 7º: APROXIMACION OPTIMA: CARACTERIZACION; EXISTENCIA Y UNICIDAD.

1. INTRODUCCION. DEFINICION. PROPIEDADES.

A) El propósito de este tema es dar los conceptos básicos sobre la aproximación de funciones en $C(I)$ mediante funciones conocidas, que en general serán polinomios.

A lo largo de este capítulo N será un espacio normado sobre el cuerpo $\mathbb{K}$ y L será un subespacio de N.

B) DEFINICION: Se dirá que $g_0 \in L$ es una aproximación óptima de $f \in N$ si se verifica que

$$\|f-g_0\| = \inf_{g\in L} \|f-g\|$$

Denotaremos $\inf_{g\in L} \|f-g\|$ por $e_L(f)$.

Denotaremos por $\mathcal{P}_L(f)$ el conjunto de todas las aproximaciones óptimas de f relativas a L, es decir $\mathcal{P}_L(f) = \{g_0 \in L \,/\, \|f-g_0\| = e_L(f)\}$.

C) PROPIEDADES:

P.1| Si $f \in L$ entonces $\mathcal{P}_L(f) = \{f\}$.

P.2| $e_L(f) \leq \|f\|$.

P.3| $\mathcal{P}_L(f)$ es convexo.

P.4| Condición necesaria para que $\mathcal{P}_L(f) \neq \emptyset$, $\forall f \in N$ es que L sea cerrado.

P.5| Condición suficiente para que $\mathcal{P}_L(f) \neq \emptyset$, $\forall f \in N$ es que L sea de dimensión finita.

DEMOSTR.: P.1| Evidente pues $e_L(f) = 0$ y $\{g_0 \in L \,/\, \|f-g_0\| = 0\} = \{f\}$.

P.2| Siendo L subespacio, $0 \in L$. Por tanto

$$e_L(f) = \inf_{g\in L} \|f-g\| \leq \|f-0\| = \|f\|.$$

P.3| Se trata de ver que dados $g_1, g_2 \in \mathcal{P}_L(f)$ y $t \in]0,1[$, entonces $tg_1 + (1-t)g_2 \in \mathcal{P}_L(f)$.

$$\|f-(tg_1+(1-t)g_2)\| = \|tf+(1-t)f-(tg_1+(1-t)g_2)\| \leq$$
$$\leq t\|f-g_1\| + (1-t)\|f-g_2\| = t\,e_L(f) + (1-t)\,e_L(f) = e_L(f)$$

Luego $\|f-(tg_1+(1-t)g_2)\| \leq e_L(f)$.

Por otra parte, como L es subespacio $tg_1+(1-t)g_2 \in L$ y por tanto $e_L(f) \leq \|f-[tg_1+(1-t)g_2]\|$.

En definitiva se da la igualdad y $tg_1+(1-t)g_2 \in \mathcal{P}_L(f)$.

P.4 Supongamos que L no es cerrado. Existiría entonces $\bar{f}$ tal que $\bar{f} = \lim_n g_n$, con $g_n \in L$ y tal que $\bar{f} \notin L$. Entonces será

$$\|\bar{f} - g_n\| \geq \inf_{g \in L} \|\bar{f} - g\|, \ \forall n \in \mathbb{N}$$

de donde $0 = \inf_{n \in \mathbb{N}} \|\bar{f} - g_n\| \geq \inf_{g \in L} \|\bar{f} - g\| \geq 0$.

Luego $e_L(\bar{f}) = \inf_{g \in L} \|\bar{f} - g\| = 0$.

Si existiese una aproximación óptima de $\bar{f}$ relativa L, es decir, si existiese $g_0 \in L$ tal que $\|\bar{f} - g_0\| = e_L(\bar{f})$ se tendría que $\bar{f} = g_0$, pues $e_L(\bar{f}) = 0$, y por tanto, $\bar{f} \in L$, en contra de la hipótesis. Deberá ser entonces $P_L(\bar{f}) = \emptyset$.

Por tanto, si $P_L(f) \neq \emptyset, \forall f \in N$, L debe ser cerrado.

P.5 Si $g \in L$ es tal que $\|g\| > 2\|f\|$ entonces

$$\|f - g\| \geq \|g\| - \|f\| > 2\|f\| - \|f\| \geq e_L(f)$$

es decir, si $g \in L$ es tal que $\|g\| > 2\|f\|$ entonces $\|f - g\| > e_L(f)$.

Por tanto, de alcanzarse $e_L(f) = \inf_{g \in L} \|f - g\|$ ha de ser en los $g \in L$ tales que $\|g\| \leq 2\|f\|$.

Sea $\bar{B}_L = \bar{B}_L(0, 2\|f\|)$ la bola cerrada de centro 0 y radio $2\|f\|$ en el subespacio L. Consideremos la aplicación

$$\phi : g \in \bar{B}_L \longmapsto \|f - g\| \in \mathbb{R}.$$

$\bar{B}_L$ es cerrado y acotado en L, que es de dimensión finita por hipótesis; luego $\bar{B}_L$ es un compacto. Como ϕ es continua en el compacto $\bar{B}_L$, alcanzará un mínimo en $\bar{B}_L$, es decir, existirá $g_0 \in \bar{B}_L$ tal que

$$\|f - g_0\| = \inf_{g \in \bar{B}_L} \|f - g\| = \inf_{g \in L} \|f - g\|$$

y, por tanto, $g_0 \in P_L(f)$. c.s.q.d.

1.1. TEOREMA: Sea $f \in N$. Entonces la aplicación

$$\phi : g \in L \longmapsto \phi(g) = \|f - g\| \in \mathbb{R}$$

no admite mínimos locales distintos de $e_L(f)$, que puede no ser alcanzado.

Demostr.: Supongamos que existe $g_1 \in L$ en el que ϕ alcanza un mínimo local distinto de $e_L(f)$, es decir, supongamos que existe $B(g_1, r)$ tal que

$$\phi(g_1) \leq \phi(g), \ \forall g \in B(g_1, r)$$

y $e_L(f) = \inf_{g \in L} \phi(g) < \phi(g_1)$.

Existe por tanto $g_2 \in L$ tal que $\phi(g_2) < \phi(g_1)$.

Sea $L[g_1, g_2] = \{ t g_1 + (1-t) g_2 \,/\, t \in [0,1] \}$

Es trivial que $L[g_1,g_2] \cap B(g_1,r) \neq \emptyset$, es ~~más~~, existe $\lambda \in]0,1[$ tal que

$g_\lambda = \lambda g_1 + (1-\lambda) g_2 \in B(g_1, r)$

y, por tanto, según lo supuesto, será $\phi(g_1) \leq \phi(g_\lambda)$. (1)

Por otra parte

$\phi(g_\lambda) = \| f - \lambda g_1 - (1-\lambda) g_2 \| = \| \lambda f + (1-\lambda) f - \lambda g_1 - (1-\lambda) g_2 \| \leq$

$\leq \lambda \|f - g_1\| + (1-\lambda)\|f - g_2\| = \lambda \phi(g_1) + (1-\lambda)\phi(g_2) < \lambda \phi(g_1) + (1-\lambda)\phi(g_1) = \phi(g_1)$

es decir, $\phi(g_\lambda) < \phi(g_1)$.

Hemos llegado a una contradicción con (1). Por tanto, ϕ no admitirá mínimos locales distintos de $e_L(f)$. csqd.

DEFINICION: (Subespacio proximinal)

Diremos que L es proximinal cuando $\forall f \in N$, $P_L(f) \neq \emptyset$.

Por tanto, según P.4 y P.5, una condición necesaria para que L sea proximinal es que L sea cerrado, y una condición suficiente para que L sea proximinal es que L sea de dimensión finita.

D) DEFINICIONES.

DEFINICION a) Diremos que L es semi-Chebyshev si $\forall f \in N$, $P_L(f)$ tiene a lo sumo un elemento.

b) Diremos que L es Chebyshev si es proximinal y semi-Chebyshev, es decir, si $\forall f \in N$, $P_L(f)$ tiene uno y solo un elemento.

DEFINICION: N es estrictamente convexo si para cualesquiera $f, g \in N$, $f \neq g$, tales que $\|f\| = \|g\|$ se verifica que $\|f+g\| < \|f\| + \|g\|$.

DEFINICION: Un conjunto $C \subset N$ es estrictamente convexo si

$\forall f, g \in C$, $\lambda f + (1-\lambda) g \in \overset{\circ}{C}$, $\forall \lambda \in]0,1[$.

1.2. PROPOSICION: N es estrictamente convexo si y solo si las bolas de N son conjuntos estrictamente convexos.

1.3. PROPOSICION: N es estrictamente convexo si y solo si cualesquiera que sean $f, g \in N$ distintos y de norma 1 se verifica que $\|f+g\| < 2$.

Demostr: $\Rightarrow$] Trivial

$\Leftarrow$] Sean $f, g \in N$ tales que $f \neq g$ y $\|f\| = \|g\|$.

Los vectores $\frac{f}{\|f\|}$ y $\frac{g}{\|g\|}$ son de norma 1 y distintos. Luego

$\left\|\frac{f}{\|f\|}+\frac{g}{\|g\|}\right\|<2$.

Como $\|f\|=\|g\|$ se deduce que $\|f+g\|<2\|f\|=\|f\|+\|g\|$. c.s.q.d.

1.4. TEOREMA: Las proposiciones siguientes son equivalentes:

a) Todo subespacio L de N es semi-Chebyshev.

b) N es estrictamente convexo.

Demostr.: $\Rightarrow$ (Caso real) (*)

Supongamos que todo subespacio L de N es semi-Chebyshev.

Si N no es estrictamente convexo, existirán $f,g\in N$ distintos tales que $\|f\|=\|g\|$ y $\|f+g\|=\|f\|+\|g\|$

Consideramos el subespacio de N engendrado por $f-g$: $L=\langle f-g\rangle$.

Vamos a probar que L tiene mas de una aproximación óptima.

Consideremos el conjunto $\{\lambda(f-g)\,/\,\lambda\in[0,1]\}$ y la función

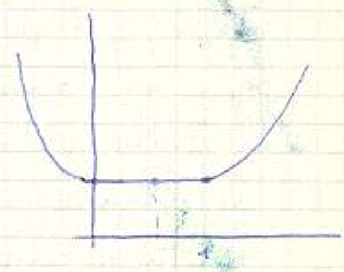
$$\varphi:\lambda\in\mathbb{R}\longmapsto\varphi(\lambda)=\|f-\lambda(f-g)\|\in\mathbb{R}.$$

Se verifica que

$$\varphi(1)=\|g\|,\ \varphi(0)=\|f\|,\ \varphi(\tfrac{1}{2})=\tfrac{1}{2}\|f+g\|=\|f\|$$

Ademas φ es una función convexa, es decir, dados $\rho,\mu\in\mathbb{R}$ y $\lambda\in[0,1]$, se verifica que $\varphi[\lambda\rho+(1-\lambda)\mu]\leq\lambda\varphi(\rho)+(1-\lambda)\varphi(\mu)$. (**)

Por tanto, siendo φ convexa y tomando el mismo valor en 0, $1/2$ y 1 debe ser φ constante en $[0,1]$. Además, por ser φ convexa tomará en $[0,1]$ el valor mínimo (su gráfica será de la forma de la figura). Luego, $e_L(f)$ se alcanza en todos los puntos de $[0,1]$, en contra de que L es semi-Chebyshev.



$\Leftarrow$ Suponemos que N es estrictamente convexo.

Sea L un subespacio de N y $f\in N$ arbitrario.

Se trata de probar que $P_L(f)$ tiene a lo sumo un elemento.

Si $f\in L$, entonces $P_L(f)=\{f\}$ y no habría nada que probar.

Supongamos que $f\notin L$ y que existen $g_1,g_2\in L$ distintos tales que $g_1,g_2\in P_L(f)$

Entonces $\|f-g_1\|=\|f-g_2\|=e_L(f)$.

Luego $\|f-\frac{1}{2}(g_1+g_2)\|=\|\frac{1}{2}f+\frac{1}{2}f-\frac{1}{2}g_1-\frac{1}{2}g_2\|<\|\frac{1}{2}(f-g_1)\|+\|\frac{1}{2}(f-g_2)\|=$

$=\frac{1}{2}\|f-g_1\|+\frac{1}{2}\|f-g_2\|$, por ser N estrictamente convexo y $\|\frac{1}{2}(f-g_1)\|=$ $=\|\frac{1}{2}(f-g_2)\|$.

Luego $\|f-\frac{1}{2}(g_1+g_2)\|<\frac{1}{2}e_L(f)+\frac{1}{2}e_L(f)=e_L(f)$. Llegamos entonces a una contradicción, pues $e_L(f) = \inf_{g\in L} \|f-g\|$ y $\frac{1}{2}(g_1+g_2) \in L$ y dista de f menos que $e_L(f)$.

Por tanto, L es Semi-Chebyshev. csqd.

2. Espacios normados uniformemente convexos.

DEFINICION: N es uniformemente convexo si para cualquier $\varepsilon>0$ existe $\delta>0$ tal que cualesquiera que sean $f,g\in N$ de norma uno verificando que $\|\frac{1}{2}(f+g)\| > 1-\delta$ entonces $\|f-g\|<\varepsilon$.

2.1. TEOREMA: a) Si N es uniformemente convexo, entonces N es estrictamente convexo.

b) Si N es de dimensión finita, entonces N es uniformemente convexo si, y sólo si, es estrictamente convexo.

Demostr.: a) Sea N un espacio normado uniformemente convexo. Supongamos que no es estrictamente convexo. Existen entonces $f,g\in N$ unitarios y distintos tales que $\|f+g\| = \|f\|+\|g\| = 2$.

Se tiene entonces que

$\|\frac{1}{2}(f+g)\| = 1 > 1-\delta, \ \forall \delta>0,$

con lo cual se llega a una contradicción sin más que tomar $\varepsilon < \|f-g\|$, es decir, se ha probado que

$\exists \varepsilon>0 (\varepsilon<\|f-g\|) / \forall \delta>0, \exists f,g\in N / \|f\|=\|g\|=1, \ \|\frac{1}{2}(f+g)\|>1-\delta$ y $\|f-g\| \geq \varepsilon$

que es la negación de que N sea uniformemente convexo.

b) Falta probar que es cierto el recíproco de a) supuesto que N es de dimensión finita. Supongamos entonces que N es de dimensión finita y estrictamente convexo. Veamos que es uniformemente convexo. Dado $\varepsilon>0$ definimos el conjunto

$$S_\varepsilon = \{(f,g)\in N\times N / \|f\|=\|g\|=1, \ \|f-g\|\geq\varepsilon\}.$$

Consideremos la aplicación

$$\Psi: (f,g)\in S_\varepsilon \longmapsto 1-\frac{1}{2}\|f+g\| \in \mathbb{R}.$$

Si $(f,g)\in S_\varepsilon$, por ser N estrictamente convexo se tendrá que

$$\|f+g\| < \|f\|+\|g\| = 2$$

con lo cual $1-\frac{1}{2}\|f+g\|>0$.

S_ε es cerrado y acotado en $N\times N$, que es de dimensión finita; por tanto, S_ε es compacto. Como la aplicación Ψ es continua y estrictamente positiva(*) alcanzará un mínimo en S_ε, es decir

$\exists (f_0, g_0) \in S_\varepsilon \,/\, 1 - \frac{1}{2}\|f_0 + g_0\| = \min\limits_{(f,g)\in S_\varepsilon} \left(1 - \frac{1}{2}\|f+g\|\right)$

con lo cual, haciendo $\delta = 1 - \frac{1}{2}\|f_0 + g_0\|$ se tiene que

$$1 - \frac{1}{2}\|f+g\| \geq \delta \quad , \forall (f,g) \in S_\varepsilon .$$

Existe entonces $\delta > 0$ tal que si $\|f\| = \|g\| = 1$ y $1 - \frac{1}{2}\|f+g\| < \delta$ entonces $\|f-g\| < \varepsilon$, pues si $\|f-g\|$ fuese mayor o igual que ε, (f,g) pertenecería a S_ε y se tendría que $1 - \frac{1}{2}\|f+g\| \geq \delta$, contra lo supuesto. Luego N es uniformemente convexo. **csqd.**

2.2. TEOREMA: Sea L un subespacio cerrado de un espacio de Banach N. Si N es uniformemente convexo entonces L es de Chebyshev.

Demostr.: Se trata de probar que $\forall f \in N$, $\mathcal{P}_L(f)$ tiene uno y solo un elemento.

i) Supongamos que $e_L(f) = 0$, es decir, $\inf\limits_{g\in L}\|f-g\| = 0$. Existirá entonces una sucesión $\{g_n\}_n$ en L tal que $\lim\limits_n \|f-g_n\| = 0$, es decir, tal que $f = \lim\limits_n g_n$. Siendo L cerrado, se tendrá que $f \in L$ y, por tanto, $\mathcal{P}_L(f) = \{f\}$.

ii) Supongamos entonces que $e_L(f) \neq 0$. Estudiemos en primer lugar el caso en que $e_L(f) = 1$. Siendo $\inf\limits_{g\in L}\|f-g\| = 1$, existe una sucesión $\{g_n\}_n$ en L tal que $\lim\limits_n \|f-g_n\| = 1$ y $\|f-g_n\| \geq 1, \forall n \in \mathbb{N}$. (I)

Por hipótesis, N es uniformemente convexo, es decir, que

$$\forall \varepsilon > 0, \exists \delta > 0 \,/\, \|f\| = \|g\| = 1 \wedge \left\|\tfrac{1}{2}(f+g)\right\| > 1 - \delta \Rightarrow \|f-g\| < \varepsilon. \quad \text{(II)}$$

De (I) se deduce que dado $\delta > 0$

$$\exists \nu \in \mathbb{N} \,/\, n \geq \nu \Rightarrow \|f-g_n\| - 1 < \delta . \qquad \text{(III)}$$

Dados $p, q \geq \nu$ se tiene que

$$\frac{1}{2}\left\|\frac{f-g_p}{\|f-g_p\|} + \frac{f-g_q}{\|f-g_q\|}\right\| = \frac{1}{2}\left\|(f-g_p) + (f-g_q) - \left(1 - \frac{1}{\|f-g_p\|}\right)(f-g_p) - \left(1 - \frac{1}{\|f-g_q\|}\right)(f-g_q)\right\| \geq$$

$$\geq \frac{1}{2}\|f-g_p+f-g_q\| - \frac{1}{2}\left(1 - \frac{1}{\|f-g_p\|}\right)\|f-g_p\| - \frac{1}{2}\left(1 - \frac{1}{\|f-g_q\|}\right)\|f-g_q\| =$$

$$= \frac{1}{2}\|f-g_p+f-g_q\| - \frac{1}{2}(\|f-g_p\| - 1) - \frac{1}{2}(\|f-g_q\| - 1) >$$

$$\overset{\text{(III)}}{>} \frac{1}{2}\|f-g_p+f-g_q\| - \delta = \left\|f - \frac{1}{2}(g_p+g_q)\right\| - \delta \geq 1 - \delta$$

pues $\frac{1}{2}(g_p+g_q) \in L$ y $\|f-g\| \geq 1, \forall g \in L$, pues $\inf\limits_{g\in L}\|f-g\| = 1$.

Entonces, por (II), tenemos que

$$\left\|\frac{f-g_p}{\|f-g_p\|} - \frac{f-g_q}{\|f-g_q\|}\right\| < \varepsilon, \text{ si } p,q \geq \nu.$$

Se tiene por tanto [illegible] $\{f-g_n\}$ [illegible]

y siendo N de Banach, existirá $h_0 \in N$ tal que
$$\lim_n \frac{f-g_n}{\|f-g_n\|} = h_0 .$$

Veamos ahora que $\lim_n (f-g_n) = h_0$.

$$\|f-g_n-h_0\| \leq \left\|f-g_n - \frac{f-g_n}{\|f-g_n\|}\right\| + \left\|\frac{f-g_n}{\|f-g_n\|} - h_0\right\| =$$

$$= \|f-g_n\| \left(1 - \frac{1}{\|f-g_n\|}\right) + \left\|\frac{f-g_n}{\|f-g_n\|} - h_0\right\|$$

Como $\|f-g_n\|$ está acotado, $\lim_n \left(1 - \frac{1}{\|f-g_n\|}\right) = 0$ y $\lim_n \left\|\frac{f-g_n}{\|f-g_n\|} - h_0\right\| = 0$

se tiene que $\lim_n \|f-g_n-h_0\| = 0$.

Por tanto, $\lim_n (f-g_n) = h_0$, o bien $\lim_n g_n = f - h_0$.

Si $g_0 = f - h_0$, se tiene que $g_0 \in L$, pues L es cerrado y $\lim_n g_n = g_0$, siendo $g_n \in L$, $\forall n \in \mathbb{N}$.

Se tiene además que $\|f-g_0\| = \lim_n \|f-g_n\| = 1$.

Es decir, $\|f-g_0\| = e_L(f)$, o también, $g_0 \in \mathcal{P}_L(f)$.

Además g_0 es único, pues siendo N uniformemente convexo es estrictamente convexo y, por tanto, L es semi-Chebyshev. (Teoremas 2.1. y 1.4.).

iii) Queda estudiar el caso en que $e_L(f) = \alpha$, $\alpha \neq 1$, $\alpha \neq 0$.

Estudiemos $e_L(\frac{1}{\alpha} f)$ y $\mathcal{P}_L(\frac{1}{\alpha} f)$:

$$e_L\left(\frac{1}{\alpha}f\right) = \inf_{g\in L} \left\|\frac{1}{\alpha}f - g\right\| = \inf_{g\in L} \left\|\frac{1}{\alpha}f - \frac{1}{\alpha}g\right\| = \inf_{g\in L} \frac{1}{\alpha}\|f-g\| = \frac{1}{\alpha} e_L(f) = 1 .$$

Aplicando el caso (i) podemos afirmar que existe una aproximación óptima g_0', y sólo una, para $\frac{1}{\alpha}f$, es decir, $\mathcal{P}_L(\frac{1}{\alpha}f) = \{g_0'\}$.

Esto quiere decir que

$$\left\|\frac{1}{\alpha}f - g_0'\right\| = \inf_{g\in L} \left\|\frac{1}{\alpha}f - g\right\| \quad \text{y por tanto}$$

$$\alpha \left\|\frac{1}{\alpha}f - g_0'\right\| = \alpha \inf_{g\in L} \left\|\frac{1}{\alpha}f - g\right\| = \inf_{g\in L} \|f - \alpha g\| = \inf_{g\in L} \|f-g\| = e_L(f)$$

Luego $\|f - \alpha g_0'\| = e_L(f)$, y por tanto, $\alpha g_0' \in \mathcal{P}_L(f)$.

Además, $\alpha g_0'$ es única, trivialmente. c.s.q.d.

TEMA 8º: ESPACIOS DE HILBERT.

A lo largo de este tema V será un espacio vectorial sobre $\mathbb{K}$ ($\mathbb{R}$ ó $\mathbb{C}$).

1. DEFINICION

DEFINICION: (Producto escalar)

Se denomina producto escalar (o interior) a una aplicación

$$\langle\,,\,\rangle : V\times V \longrightarrow \mathbb{K}$$
$$(x,y) \longmapsto \langle x,y\rangle$$

que satisface las siguientes propiedades

i) $\langle \alpha x+\beta y, z\rangle = \alpha\langle x,z\rangle + \beta\langle y,z\rangle$, $\forall x,y,z\in V$, $\forall \alpha,\beta\in\mathbb{K}$.

ii) $\langle x,y\rangle = \overline{\langle y,x\rangle}$, $\forall x,y\in V$.

iii) $\langle x,x\rangle \geq 0$, $\forall x\in V$.

iv) $\langle x,x\rangle = 0 \Leftrightarrow x=0$.

Una aplicación que verifique i) e ii) se denomina una forma hermitiana. Si verifica iii) se dice que es semidefinida positiva. Si además verifica iv) se dice que es definida positiva.

Así pues un producto escalar es una forma hermitiana definida positiva.

Consecuencia de los cuatro axiomas precedentes son las siguientes propiedades:

v) $\langle\,,\,\rangle$ es antilineal en la segunda componente, es decir
$\langle x, \alpha y+\beta z\rangle = \bar{\alpha}\langle x,y\rangle + \bar{\beta}\langle x,z\rangle$, $\forall x,y,z\in V$, $\forall\alpha,\beta\in\mathbb{K}$.

vi) $|\langle x,y\rangle|^2 \leq \langle x,x\rangle\cdot\langle y,y\rangle$, $\forall x,y\in V$. (DESIGUALDAD DE CAUCHY-SCHWARTZ).

vii) $(\langle x+y,x+y\rangle)^{1/2} \leq \langle x,x\rangle^{1/2} + \langle y,y\rangle^{1/2}$. (DESIGUALDAD DE MINKOWSKY).

1.1. TEOREMA: La aplicación $\|\cdot\| : x\in V \longmapsto \|x\| = \langle x,x\rangle^{1/2}\in\mathbb{R}$ es una norma sobre V.

DEFINICION: (Espacio prehilbertiano)

Se llama espacio prehilbertiano a un par $H=(V;\langle\,,\,\rangle)$ donde V es un espacio vectorial sobre $\mathbb{K}$ y $\langle\,,\,\rangle$ un producto escalar en V.

La norma $\|\cdot\|$ inducida en V por el producto escalar (en el sentido del teorema 1.1) induce una topología en V. Si V es completo con esta topología diremos que $H=(V;\langle\,,\,\rangle)$ es un ESPACIO DE HILBERT.

Supongamos que $\langle\,,\,\rangle : V\times V\longrightarrow\mathbb{R}$ es una forma hermitiana semidefinida positiva. A partir de ella podemos definir la aplicación

$$P: x\in V \longrightarrow P(x) = \langle x,x\rangle^{1/2}$$

Esta aplicación verifica las siguientes propiedades:

1) $P(x) \geq 0$, $\forall x \in V$.

2) $P(\alpha x) = |\alpha| P(x)$, $\forall x \in V$, $\forall \alpha \in \mathbb{K}$.

3) $P(x+y) \leq P(x) + P(y)$, $\forall x, y \in V$.

A P (~~x~~ cualquier aplicación que verifique 1), 2) y 3)) se le llama una Seminorma sobre V.

1.2. TEOREMA: (de Jordan - Von Newmann)

La norma de un espacio normado N esta inducida por un producto interior si, y solo si, verifica la "igualdad del paralelogramo":

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right), \quad \forall x, y \in N$$

2. ORTOGONALIDAD.

DEFINICION: Sea H un espacio prehilbertiano. Se dice que $x \in H$ es ortogonal a $y \in H$, y lo representaremos por $x \perp y$, si $\langle x, y \rangle = 0$.

Se verifican las siguientes propiedades:

i) $x \perp y \Rightarrow y \perp x$. (Simetría)

ii) $x \perp y \wedge x \perp z \Rightarrow x \perp (y+z)$. (Aditividad).

iii) $x \perp y \Rightarrow \lambda x \perp \mu y$, $\forall \lambda, \mu \in \mathbb{K}$. (Homogeneidad).

iv) Teorema de Pitágoras

$$x \perp y \Rightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2$$

OBSERVACIONES: a) Si $x \perp y$ entonces $\|x-y\|^2 = \|x\|^2 + \|y\|^2$, pues $\|x-y\|^2 = \|x+(-y)\|^2 = \|x\|^2 + \|-y\|^2 = \|x\|^2 + \|y\|^2$.

b) En el caso $\mathbb{K} = \mathbb{R}$ se verifica el recíproco del teorema de Pitágoras, es decir, $\|x+y\|^2 = \|x\|^2 + \|y\|^2 \Rightarrow x \perp y$.

c) La ortogonalidad no es transitiva, es decir, no es cierta, en general, la implicación: $x \perp y \wedge y \perp z \Rightarrow x \perp z$.

2.1. TEOREMA: Dados $x, y \in H$, con $y \neq 0$, existe $c \in \mathbb{K}$ tal que

$$x - cy \perp y.$$

Demostr.: Si se ha de verificar que $\langle x - cy, y \rangle = 0$ deberá ser $\langle x, y \rangle - c\langle y, y \rangle = 0$, o bien $\langle x, y \rangle - c\|y\|^2 = 0$.

Basta entonces tomar $c = \frac{\langle x, y \rangle}{\|y\|^2}$. ■

Nota: El escalar c anterior está unívocamente determinado por x e y y se le denomina coeficiente de Fourier de x relativo a y.

2.2. TEOREMA: Sean $x_1, x_2, \ldots, x_n \in H - \{0\}$ tales que $x_i \perp x_j$ si $i \neq j$. Sea c_i el coeficiente de Fourier de $x \in H$ relativo a x_i. Entonces

$$(x - c_1 x_1 - \cdots - c_n x_n) \perp x_j, \quad j = 1, \ldots, n.$$

$$\langle x - c_1 x_1 - \dots - c_n x_n , x_j \rangle = \langle x - c_j x_j , x_j \rangle - \sum_{\substack{i=1 \\ i\neq j}}^{n} c_i \langle x_i , x_j \rangle .$$

DEFINICION: (Ortogonalidad en el sentido de Birkhoff ó B-ortogonalidad)
Sea N un espacio normado y $f, g \in N$. Se dice que f es ortogonal a g en el sentido Birkhoff, y lo denotaremos por $f \perp_B g$ si
$$\|f + \lambda g\| \geq \|f\| , \quad \forall \lambda \in \mathbb{K}.$$

Propiedades: a) No es simétrica: $f \perp_B g \not\Rightarrow g \perp_B f$.
b) $0 \perp_B g$ y $g \perp_B 0$, $\forall g \in N$
c) $\perp_B$ es homogénea: $f \perp_B g \Rightarrow \lambda f \perp_B \mu g$, $\forall \lambda, \mu \in \mathbb{K}$.

2.3. TEOREMA: Sea N un espacio normado y L un subespacio de N. Sea $f \in N$. Entonces $g_0 \in \mathcal{P}_L(f)$ si, y solo si, $f - g_0 \perp_B L$.

Demostr.: $\Rightarrow$] Si $g_0 \in \mathcal{P}_L(f)$, entonces $\|f - g_0\| \leq \|f - g\|$, $\forall g \in L$.
O tambien $\|f - g_0\| \leq \|f - \lambda g\|$, $\forall g \in L$, $\forall \lambda \in \mathbb{K}$.
O tambien $\|f - g_0\| \leq \|f - g_0 + g_0 - \lambda g\|$, $\forall g \in L$, $\forall \lambda \in \mathbb{K}$.
Cuando g recorre L y λ recorre $\mathbb{K}$, $g_0 - \lambda g$ recorre L. Luego
$\|f - g_0\| \leq \|f - g_0 + g\|$, $\forall g \in L$, o tambien $\|f - g_0\| \leq \|f - g_0 + \lambda g\|$, $\forall \lambda \in \mathbb{K}$, $\forall g \in L$,
es decir $f - g_0 \perp_B L$.
$\Leftarrow$] Supongamos que $f - g_0 \perp_B L$, es decir, $\forall g^* \in L$, $f - g_0 \perp_B g^*$.
Entonces, $\forall g^* \in L$, $\|f - g_0\| \leq \|f - g_0 + \mu g^*\|$, $\forall \mu \in \mathbb{K}$.
Cuando $\mu \in \mathbb{K}$ y $g^* \in L$ se tiene que $g_0 - \mu g^*$ recorre L y por tanto
$$\|f - g_0\| \leq \|f - g\|, \ \forall g \in L,$$
es decir, $g_0 \in \mathcal{P}_L(f)$. csqd.

2.4. TEOREMA: Sea H un espacio prehilbertiano. Entonces
$$f \perp g \iff f \perp_B g .$$

Demostr.: $\Rightarrow$] Sean $f, g \in H$ tales que $f \perp g$.
Dado $\lambda \in \mathbb{K}$ se verifica que
$$\|f + \lambda g\|^2 = \langle f + \lambda g , f + \lambda g \rangle = \langle f, f \rangle + \lambda \langle g, f \rangle + \bar{\lambda} \langle f, g \rangle + \lambda \bar{\lambda} \langle g, g \rangle =$$
$$= \|f\|^2 + |\lambda|^2 \|g\|^2 \geq \|f\|^2$$
Por tanto, $\|f + \lambda g\| \geq \|f\|$, $\forall \lambda \in \mathbb{K}$, es decir $f \perp_B g$.
$\Leftarrow$] Supongamos que $f \perp_B g$. Supongamos que f no es ortogonal a g y probemos que f no es B-ortogonal a g, contra lo supuesto.
Si f no es ortogonal a g, $\langle f, g \rangle \neq 0$ y $g \neq 0$.
Sea $\lambda = - \dfrac{\langle f, g \rangle}{\|g\|^2}$. Entonces
$$\|f + \lambda g\|^2 = \langle f - \frac{\langle f, g \rangle}{\dots} g , f - \frac{\langle f, g \rangle}{\dots} \dots \rangle =$$

$$= \|f\|^2 - \frac{<f,g>}{\|g\|^2}<g,f> - \frac{\overline{<f,g>}}{\|g\|^2}<f,g> + \frac{<f,g>\overline{<f,g>}}{\|g\|^2} =$$

$$= \|f\|^2 - \frac{|<f,g>|^2}{\|g\|^2} < \|f\|^2$$

Luego $\|f+\lambda g\| < \|f\|$, en contra de que $f \perp_B g$.

Debe ser entonces $f \perp g$. csqd.

3. Algunas propiedades de espacios prehilbertianos.

3.1. Teorema: Todo espacio prehilbertiano es uniformemente convexo.

Demostr.: Se trata de probar que

$$\forall \varepsilon>0, \exists \delta>0 \,/\, \underset{f,g\in H}{\|f\|=\|g\|=1} \wedge \|\tfrac{1}{2}(f+g)\|>1-\delta \Rightarrow \|f-g\|<\varepsilon.$$

Sea $\varepsilon>0$ y $f,g\in H$ tales que $\|f\|=\|g\|=1$.

Por la igualdad del paralelogramo

$$\|f+g\|^2+\|f-g\|^2 = 2(\|f\|^2+\|g\|^2)=4 \qquad (1)$$

Se trata de hallar $\delta>0$ tal que si $\|\frac{1}{2}(f+g)\|>1-\delta$ entonces $\|f-g\|<\varepsilon$, o bien, $\|f-g\|^2<\varepsilon^2$, o bien, según (1), que $4-\|f+g\|^2<\varepsilon^2$, es decir, $\|f+g\|^2+\varepsilon^2>4$.

Basta entonces tomar δ de forma que $[2(1-\delta)]^2+\varepsilon^2=4$, es decir, $\delta = 1-\frac{\sqrt{4-\varepsilon^2}}{2}$, pues si $\frac{1}{2}\|f+g\|>1-\delta=\frac{\sqrt{4-\varepsilon^2}}{2}$ se tiene que $\|f+g\|>\sqrt{4-\varepsilon^2}$ y por tanto $\|f+g\|^2>4-\varepsilon^2$, o bien $\|f+g\|^2+\varepsilon^2>4$ de donde se deduce que $\|f-g\|<\varepsilon$, según se indicó anteriormente. ∎

En relación con la teoría de aproximación óptima tenemos los siguientes resultados que recogemos en el

3.2. Teorema: Sea H un espacio prehilbertiano. Entonces

a) Todo subespacio de dimensión finita es Chebyshev.

b) Todo subespacio de H es semi-Chebyshev.

c) Si H es de Hilbert, todo subespacio cerrado de H es Chebyshev.

3.3. Teorema: Si L es un subespacio cerrado de un espacio de Hilbert H, y f es un elemento de H, entonces existe un único $g_0\in L$ tal que

i) $\|f-g_0\| = \inf_{g\in L}\|f-g\|$.

ii) $f-g_0 \perp L$.

Definición: Al elemento g_0 a que hace referencia el teorema anterior se le llama proyección ortogonal de f sobre L.

TEMA 9º: EL ESPACIO DE FUNCIONES CONTINUAS C(I).

1. PRODUCTO ESCALAR Y NORMA SOBRE C(I).

A) Sea $I=[a,b]$ un intervalo compacto de $\mathbb{R}$. Se define $C(I)$ como el conjunto de las funciones reales continuas definidas en I.

La aplicación $\|\cdot\|: f\in C(I) \longmapsto \|f\| = \sup_{x\in I}|f(x)| \in \mathbb{R}$

es una norma sobre $C(I)$ llamada norma del supremo o norma infinito.

Siendo I compacto, toda funcion real continua definida en I es uniformemente continua.

Se prueba que el espacio $C(I)$ dotado de la norma infinito es un espacio normado completo (de Banach).

B) DEFINICION: Dadas $f,g\in C(I)$, $I=[a,b]$, se define

$$\langle f,g\rangle = \int_a^b f(x)g(x)\,dx \in \mathbb{R}$$

La aplicación $\langle\,,\rangle$ así definida constituye un producto escalar sobre $C(I)$.
La norma inducida por este producto interior se representa por $\|\cdot\|_2$ y viene definida por

$$\|f\|_2 = \langle f,f\rangle^{1/2} = \left(\int_a^b f^2(x)\,dx\right)^{1/2}$$

$C(I)$ dotado con esta norma no es completo.

2. SISTEMA ORTOGONAL.

DEFINICION: Sea H un espacio prehilbertiano y $\{f_n\}_n$ una sucesión en H. Se dice que $\{f_n\}_n$ es un sistema ortogonal si $f_i\perp f_j$, $\forall i\neq j$ (se supone que $f_n\neq 0, \forall n$).
Si además se verifica que $\|f_n\|=1, \forall n$, se dice que $\{f_n\}_n$ es un sistema ortonormal.

De otra forma, $\{f_n\}_n$ es ortonormal si

$$\langle f_i, f_j\rangle = \delta_{ij} = \begin{cases} =1 & \text{si } i=j \\ =0 & \text{si } i\neq j \end{cases}$$

DEFINICION: (Sistema total)

Sea H un espacio prehilbertiano y $\{v_n\}$ una sucesión ortonormal. Sea F el conjunto de las combinaciones lineales finitas de elementos de la sucesión. Se dice que $\{v_n\}_n$ es un sistema total si la adherencia de F es H, o de otra forma, si F es denso en H.

2.1. PROPOSICION: Sea $\{v_n\}_n$ una familia de elementos ortogonales. Entonces los v_n son linealmente independientes.

Demostr.: Se trata de probar que es libre cualquier subfamilia finita. Sea $\{i_k\}_{k=1}^n$ una colección finita de números naturales. Sean $\lambda_1, \dots, \lambda_n$ escalares tales que $\lambda_1 v_{i_1} + \lambda_2 v_{i_2} + \dots + \lambda_n v_{i_n} = 0$. Se trata de probar que $\lambda_k = 0$, $k = 1, \dots, n$.

Se tiene que $\langle \lambda_1 v_{i_1} + \dots + \lambda_n v_{i_n}, v_{i_k} \rangle = 0$, $k = 1, \dots, n$.
Luego $\lambda_1 \langle v_{i_1}, v_{i_k} \rangle + \dots + \lambda_n \langle v_{i_n}, v_{i_k} \rangle = 0$, es decir

$$\lambda_k \langle v_{i_k}, v_{i_k} \rangle = 0$$

o tambien $\lambda_k \|v_{i_k}\|^2 = 0$, y, por tanto, $\lambda_k = 0$, pues suponemos que $v_n \neq 0$, $\forall n \in \mathbb{N}$. csqd.

2.2. TEOREMA: Sea H un espacio prehilbertiano, $\{v_n\}_n$ una sucesión ortogonal en H. Sea $f \in H$ y $\{c_n\}_n$ los coeficientes de Fourier de f relativos a cada v_n, es decir, $c_n = \dfrac{\langle f, v_n \rangle}{\langle v_n, v_n \rangle}$, $\forall n \in \mathbb{N}$.
Si $\{a_n\}_n$ es una sucesión de complejos arbitrarios se verifica que

$$\left\| f - \sum_{i=1}^n c_i v_i \right\| \leq \left\| f - \sum_{i=1}^n a_i v_i \right\|, \ \forall n \in \mathbb{N}.$$

Demostr.:

$$\left\| f - \sum_{i=1}^n a_i v_i \right\|^2 = \left\| f - \sum_{i=1}^n c_i v_i + \sum_{i=1}^n (c_i - a_i) v_i \right\|^2 \stackrel{(1)}{=} \left\| f - \sum_{i=1}^n c_i v_i \right\|^2 + \left\| \sum_{i=1}^n (c_i - a_i) v_i \right\|^2$$

y por tanto $\left\| f - \sum_{i=1}^n a_i v_i \right\| \geq \left\| f - \sum_{i=1}^n c_i v_i \right\|$.
La igualdad (1) es cierta en virtud del teorema de Pitágoras, pues

$$\left\langle f - \sum_{i=1}^n c_i v_i, \sum_{i=1}^n (c_i - a_i) v_i \right\rangle = \left\langle f, \sum_{i=1}^n (c_i - a_i) v_i \right\rangle - \left\langle \sum_{i=1}^n c_i v_i, \sum_{j=1}^n (c_j - a_j) v_j \right\rangle =$$

$$= \sum_{i=1}^n (\overline{c_i} - \overline{a_i}) \langle f, v_i \rangle - \sum_{i=1}^n c_i \left\langle v_i, \sum_{j=1}^n (c_j - a_j) v_j \right\rangle =$$

$$= \sum_{i=1}^n (\overline{c_i} - \overline{a_i}) \langle f, v_i \rangle - \sum_{i=1}^n \sum_{j=1}^n c_i (\overline{c_j} - \overline{a_j}) \langle v_i, v_j \rangle =$$

$$= \sum_{i=1}^n (\overline{c_i} - \overline{a_i}) \langle f, v_i \rangle - \sum_{i=1}^n c_i (\overline{c_i} - \overline{a_i}) \langle v_i, v_i \rangle =$$

$$= \sum_{i=1}^n (\overline{c_i} - \overline{a_i}) \langle f - c_i v_i, v_i \rangle = 0,$$ pues $f - c_i v_i \perp v_i$, $\forall i \in \mathbb{N}$. csqd.

OBSERVACION: La igualdad $\left\| f - \sum_{i=1}^n c_i v_i \right\| = \left\| f - \sum_{i=1}^n a_i v_i \right\|$ se da si, y solo si, $c_i = a_i$, $\forall i \in \{1, \dots, n\}$.

3. PROCESO DE ORTONORMALIZACION DE GRAM-SCHMIDT.

3.1. TEOREMA: Sea H un espacio prehilbertiano y $\{f_n\}_n$ una sucesión de elementos linealmente independientes de H. Existe una sucesión $\{g_n\}_n$ que verifica lo siguiente:
a) $\{g_n\}_n$ es ortonormal.
b) Cada g_n es combinación lineal de $f_1, \dots, f_n$ y, por tanto, el espacio generado por $\{f_i\}_i$ coincide con el generado por $\{g_i\}$

Demostr.: Construiremos la sucesión $\{g_n\}_n$ por recurrencia. Construiremos primero una sucesión ortogonal $\{g_n^*\}_n$ que verifica b).

Tomamos $g_1^* = f_1$ y $g_2^* = f_2 - c_1^2 g_1^*$ donde c_1^2 es el coeficiente de Fourier de f_2 relativo a g_1^*. Por tanto $g_2^* \perp g_1^*$.

Supongamos que $g_1^*, \dots, g_{n-1}^*$ han sido construidos verificando b) y que son ortogonales dos a dos.

Definimos entonces

$$g_n^* = f_n - c_1^n g_1^* - \dots - c_{n-1}^n g_{n-1}^*$$

donde c_j^n es el coeficiente de Fourier de f_n relativo a g_j^*.

Entonces $g_n^* \perp g_j^*$, $\forall j \in \{1, \dots, n-1\}$.

La sucesión $\{g_n\}_n$ definida por $g_n = \dfrac{g_n^*}{\|g_n^*\|}$ verifica a) y b). csqd.

4. Aproximacion por minimos cuadrados en $C(I)$.

A) Sea $\omega \in C(I)$ verificando las condiciones

a) $\omega(x) \geq 0$, $\forall x \in I$

b) $\forall f \in C(I)$, existe $\int_I f(x)\omega(x)\,dx$.

c) $\int_J \omega(x)\,dx > 0$, para todo J subintervalo de I.

La aplicación $\langle \, , \rangle_\omega : (f,g) \in C(I) \times C(I) \longmapsto \langle f,g \rangle_\omega = \int_I f(x)g(x)\omega(x)\,dx \in \mathbb{R}$ es un producto interior en $C(I)$.

La aproximación en el espacio prehilbertiano $(C(I), \langle \, , \rangle_\omega)$ se denomina de mínimos cuadrados.

Sea $\{g_1, \dots, g_n\}$ una colección finita de elementos de $C(I)$, y $f \in C(I)$. Si pretendemos aproximar f mediante combinaciones lineales de $g_1, \dots, g_n$, interesará saber si existen $\tilde{c}_1, \dots, \tilde{c}_n \in \mathbb{R}$ tales que

$$\|f - (\tilde{c}_1 g_1 + \dots + \tilde{c}_n g_n)\|_\omega \leq \|f - (c_1 g_1 + \dots + c_n g_n)\|_\omega, \ \forall c_1, \dots, c_n \in \mathbb{R}. \quad (I)$$

Consideremos dos casos:

i) $\{g_1, \dots, g_n\}$ es una familia ortonormal. Entonces, según el teorema 2.2, se verifica (I) para $\tilde{c}_i = \dfrac{\langle f, g_i \rangle_\omega}{\|g_i\|_\omega^2} = \langle f, g_i \rangle_\omega$.

ii) Supongamos simplemente que $g_1, \dots, g_n$ son linealmente independientes. Sea $L = \langle \{g_1, \dots, g_n\} \rangle$. Entonces $\dim L = n$.

Puesto que $(C(I), \langle \, , \rangle_\omega)$ es un espacio prehilbertiano y L es de dimensión finita, se verifica que L es Chebyshev, es decir, dado $f \in C(I)$, existe g_0 tal que $\mathcal{P}_L(f) = \{g_0\}$. Como $g_0 \in L$, existen unos únicos escalares $\tilde{c}_1, \dots, \tilde{c}_n$ tales que

$$g_0 = \tilde{c}_1 g_1 + \dots + \tilde{c}_n g_n.$$

Veamos ahora la forma de encontrar estos $\tilde{c}_i$.

$$P_L(f) = \{g_0\} \Leftrightarrow f - g_0 \perp L \Leftrightarrow \begin{cases} \langle f-g_0, g_1\rangle = 0 \\ \dots\dots \\ \langle f-g_0, g_n\rangle = 0 \end{cases} \Leftrightarrow \begin{cases} \langle f, g_1\rangle = \langle g_0, g_1\rangle \\ \dots\dots \\ \langle f, g_n\rangle = \langle g_0, g_n\rangle \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \langle f, g_1\rangle = \tilde{c}_1 \langle g_1, g_1\rangle + \dots + \tilde{c}_n \langle g_n, g_1\rangle \\ \dots\dots \\ \langle f, g_n\rangle = \tilde{c}_1 \langle g_1, g_n\rangle + \dots + \tilde{c}_n \langle g_n, g_n\rangle \end{cases}$$

Como este sistema admite solución única $\tilde{c}_1, \dots, \tilde{c}_n$ se verificará que

$$\Delta = \begin{vmatrix} \langle g_1, g_1\rangle & \dots & \langle g_n, g_1\rangle \\ \dots & & \dots \\ \langle g_1, g_n\rangle & \dots & \langle g_n, g_n\rangle \end{vmatrix} \neq 0.$$

Este determinante se denomina determinante de Gram.
Los valores $\tilde{c}_i$ se calcularan mediante

$$\tilde{c}_i = \frac{\begin{vmatrix} \langle g_1, g_1\rangle & \dots & \langle f, g_1\rangle & \dots & \langle g_n, g_1\rangle \\ \dots & & & & \dots \\ \langle g_1, g_n\rangle & \dots & \langle f, g_n\rangle & \dots & \langle g_n, g_n\rangle \end{vmatrix}}{\Delta}$$

OBSERVACIONES: a) Si $\{g_1, \dots, g_n\}$ es ortonormal, $\Delta = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & 1 \end{vmatrix} = 1$

y $\tilde{c}_i = \langle f, g_i\rangle$.

b) $\| f - \tilde{c}_1 g_1 - \dots - \tilde{c}_n g_n \| = 0$ si, y solo si, $f \in L\langle \{g_1, \dots, g_n\}\rangle$.

c) Sea $g_0 = \tilde{c}_1 g_1 + \dots + \tilde{c}_n g_n$ la aproximación óptima de f relativa a $L = \langle \{g_1, \dots, g_n\}\rangle$.
El error que se cometerá en la aproximación será

$$\|f - g_0\| = \inf_{(a_1, \dots, a_n)} \|f - a_1 g_1 - \dots - a_n g_n\|$$

Supuesto que $\{g_1, \dots, g_n\}$ sea ortonormal

$$\|f - g_0\|^2 = \langle f - g_0, f - g_0\rangle = \langle f, f\rangle - \langle f, g_0\rangle - \langle g_0, f\rangle + \langle g_0, g_0\rangle =$$
$$= \|f\|^2 - 2\langle f, g_0\rangle + \langle g_0, g_0\rangle =$$
$$= \|f\|^2 - 2\langle f, (\langle f, g_1\rangle g_1 + \dots + \langle f, g_n\rangle g_n)\rangle + \langle (\langle f, g_1\rangle g_1 + \dots + \langle f, g_n\rangle g_n), (\langle f, g_1\rangle g_1 + \dots + \langle f, g_n\rangle g_n)\rangle =$$
$$= \|f\|^2 - 2\sum_{i=1}^{n} |\langle f, g_i\rangle|^2 + \sum_{i=1}^{n} |\langle f, g_i\rangle|^2 =$$
$$= \|f\|^2 - \sum_{i=1}^{n} |\langle f, g_i\rangle|^2.$$

B) En lo visto anteriormente queda reflejada la utilidad de las bases ortonormales, o al menos ortogonales, para la determinación de la aproximación óptima. Por este motivo, en adelante se estudiaran propiedades de sucesiones de polinomios ortogonales, pues los subespacios de $C(J)$ respecto de los cuales se aproximarán funciones continuas serán los P_n (conjunto de polinomios de grado menor o igual que n).

4.1. TEOREMA: Una sucesión de polinomios $\{Q_0, Q_1, \dots, Q_n, \dots\}$ tal que grado $Q_n = n$, $\forall n \in \mathbb{N}$, verifica lo siguiente:

i) Es linealmente independiente.

ii) Todo polinomio de grado menor o igual que K se expresa de forma única como combinación lineal de $Q_0, Q_1, \dots, Q_K$.

Demostr.: Como el ser linealmente independientes equivale a que solo la combinación lineal con coeficientes nulos es igual a cero, la segunda afirmación implica la primera.

Demostraremos ii) por inducción en n.

- Para $n=0$ es trivial, pues los polinomios de grado cero son las constantes.
- Supuesto que ii) se verifica para todo polinomio de grado menor o igual que $n-1$, probemos que también lo es para todo polinomio de grado n.

Veamos en primer lugar que un polinomio $P(x) = \sum_{K=0}^{n} a_K x^K$, $a_n \neq 0$, es combinación lineal de $Q_0, Q_1, \dots, Q_n$.

Q_n es de la forma $Q_n(x) = \sum_{K=0}^{n} c_K x^K$, $c_n \neq 0$.

Sea $\lambda_n \in \mathbb{R}$ tal que $a_n = \lambda_n c_n$.

Entonces $P(x) - \lambda_n Q_n(x) = (a_{n-1} - \lambda_n c_{n-1}) x^{n-1} + \dots + (a_1 - \lambda_n c_1) x + (a_0 - \lambda_n c_0)$ que es de grado menor o igual que $n-1$. Por la hipótesis de inducción existen $\lambda_0, \lambda_1, \dots, \lambda_{n-1} \in \mathbb{R}$ tales que

$$P(x) - \lambda_n Q_n(x) = \sum_{K=0}^{n-1} \lambda_K Q_K(x)$$

y, por tanto, $P(x) = \sum_{K=0}^{n} \lambda_K Q_K(x)$.

Además $\lambda_0, \lambda_1, \dots, \lambda_{n-1}$ son únicos.

Entonces, si $P(x) = \sum_{K=0}^{n} \mu_K Q_K(x)$ se verifica que $a_n = \mu_n c_n = \lambda_n c_n$, y, por tanto, $\mu_n = \lambda_n$.

Luego $\sum_{K=0}^{n-1} \mu_K Q_K(x) = \sum_{K=0}^{n-1} \lambda_K Q_K(x)$

y por la hipótesis de inducción será $\lambda_K = \mu_K$, $K = 0, 1, \dots, n-1$. c.s.q.d.

4.2. TEOREMA: Sea $\{\Psi_0, \Psi_1, \dots, \Psi_n, \dots\}$ una sucesión de polinomios ortogonales verificando que grado $\Psi_K = K$, $\forall K \in \mathbb{N}$. Entonces no existe ningún polinomio de grado K ortogonal a Ψ_K, $K = 0, 1, \dots$ Además, Ψ_K es ortogonal a todo polinomio de grado menor que K.

Demostr.: Sea $P(x)$ un polinomio de grado K. Existen entonces unos escalares

$a_0, a_1, \dots, a_K$, con $a_K \neq 0$ de forma que

$$P = \sum_{i=0}^{K} a_i \Psi_i .$$

Entonces $\langle P, \Psi_K \rangle_\omega = \sum_{i=0}^{K} a_i \langle \Psi_i, \Psi_K \rangle_\omega = a_K \langle \Psi_K, \Psi_K \rangle_\omega$, que es distinto de cero.

Por otra parte, si grado $P \leq K-1$, $P = \sum_{i=0}^{K-1} a_i \Psi_i$ y por consiguiente

$$\langle P, \Psi_K \rangle_\omega = \sum_{i=0}^{K-1} a_i \langle \Psi_i, \Psi_K \rangle_\omega = 0 \text{ . csqd.}$$

Este resultado se utilizará en la demostración del teorema siguiente, que establece la unicidad, salvo signo, de sucesiones de polinomios ortonormales y, como corolario, la unicidad salvo constantes de sucesiones de polinomios ortogonales.

4.3. TEOREMA: Si $\{\phi_0, \phi_1, \dots, \phi_n, \dots\}$ y $\{\Psi_0, \Psi_1, \dots, \Psi_n, \dots\}$ son dos sucesiones de polinomios ortonormales tales que grado ϕ_K = grado $\Psi_K = K$, $(K=0,1,2,\dots)$, entonces

$$\phi_K = \Psi_K \quad \text{ó} \quad \phi_K = -\Psi_K \quad , K=0,1,2,\dots$$

Demostr.: Sean $\phi_K = \sum_{i=0}^{K} a_i x^i, a_K \neq 0$ y $\Psi_K = \sum_{i=0}^{K} b_i x^i, b_K \neq 0$.

Supongamos que $a_K \cdot b_K > 0$ y probemos que $\phi_K = \Psi_K$.

En primer lugar

$$\|\phi_K - \Psi_K\|^2 = \langle \phi_K - \Psi_K, \phi_K - \Psi_K \rangle_\omega = \|\phi_K\|^2 - 2\langle \phi_K, \Psi_K \rangle_\omega + \|\Psi_K\|^2 = 2 - 2\langle \phi_K, \Psi_K \rangle_\omega \quad (1)$$

Por otra parte

$$\langle \phi_K, \Psi_K \rangle_\omega = \int_I \phi_K(x)\Psi_K(x)\omega(x)\,dx = \int_I \Big(\sum_{i=0}^{K} a_i x^i\Big)\Psi_K(x)\omega(x)\,dx =$$

$$= a_K \int_I x^K \Psi_K(x)\omega(x)\,dx + \int_I \Big(\sum_{i=0}^{K-1} a_i x^i\Big)\Psi_K(x)\omega(x)\,dx = a_K \int_I x^K \Psi_K(x)\omega(x)\,dx \quad (2)$$

pues por el teorema anterior $\langle \sum_{i=0}^{K-1} a_i x^K, \Psi_K(x) \rangle_\omega = 0$.

Analogamente

$$\langle \phi_K, \Psi_K \rangle_\omega = b_K \int_I x^K \phi_K(x)\omega(x)\,dx + \int_I \phi_K(x)\Big(\sum_{i=0}^{K-1} b_i x^i\Big)\omega(x)\,dx = b_K \int_I x^K \phi_K(x)\omega(x)\,dx . \quad (3)$$

También por el teorema anterior se tiene que

$$1 = \|\phi_K\|^2 = \langle \phi_K, \phi_K \rangle_\omega = a_K \int_I x^K \phi_K(x)\omega(x)\,dx . \quad (4)$$

$$\text{y} \quad 1 = \|\Psi_K\|^2 = \langle \Psi_K, \Psi_K \rangle_\omega = b_K \int_I x^K \Psi_K(x)\omega(x)\,dx . \quad (5)$$

Se tiene por tanto, de (2) y (5), que

$$\langle \phi_K, \Psi_K \rangle_\omega = \frac{a_K}{b_K}$$

y de (3) y (4) que

$$\langle \phi_K, \Psi_K \rangle_\omega = \frac{b_K}{a_K}$$

Por tanto, $\frac{a_K}{b_K} = \frac{b_K}{a_K}$, de donde $a_K = b_K$. Además $\langle \phi_K, \Psi_K \rangle_\omega = 1$.

Entonces por (1) se tiene que $\|\phi_k - \psi_k\|^2 = 0$, de donde se sigue que $\phi_k = \psi_k$.
Si hubiésemos supuesto que $a_k b_k < 0$ hubiéramos llegado a que $\phi_k = -\psi_k$, c.s.q.

El proceso de ortonormalización de Gram-Schmidt ya estudiado nos permitirá construir una base ortonormal de P_n partiendo, por ejemplo, de la base

$$x \in I \longmapsto 1$$
$$x \in I \longmapsto x$$
$$x \in I \longmapsto x^2$$
$$\text{- - - - - - -}$$
$$x \in I \longmapsto x^n$$
$$\text{- - - - - - -}$$

Sin embargo, en este caso particular existe un procedimiento más sencillo para construir una base ortogonal de P_n, que se expone en el siguiente teorema:

4.4. TEOREMA: La sucesión de polinomios $\{\psi_n\}_n$ definida por

$$\psi_n = (x - a_n)\psi_{n-1} - b_n \psi_{n-2}$$

siendo

$$\psi_0 = 1, \quad \psi_1 = x - a_1, \quad a_n = \frac{\langle x\psi_{n-1}, \psi_{n-1}\rangle}{\langle \psi_{n-1}, \psi_{n-1}\rangle}, \quad b_n = \frac{\langle x\psi_{n-1}, \psi_{n-2}\rangle}{\langle \psi_{n-2}, \psi_{n-2}\rangle}$$

es ortogonal.

Demostr.: Se ve fácilmente por inducción que para cada $n \in \mathbb{N}$, el grado de ψ_n es n y el coeficiente principal uno. Por consiguiente, $\langle \psi_{n-1}, \psi_{n-1}\rangle \neq 0$ y $\langle \psi_{n-2}, \psi_{n-2}\rangle \neq 0$, y a_n y b_n están bien definidos.
Probaremos ahora por inducción que

$$\langle \psi_n, \psi_i\rangle_\omega = 0, \quad \forall i < n.$$

- Para $n = 1$:

$$\langle \psi_1, \psi_0\rangle_\omega = \langle x - a_1, \psi_0\rangle_\omega = \langle x\psi_0 - a_1\psi_0, \psi_0\rangle_\omega =$$
$$= \langle x\psi_0, \psi_0\rangle_\omega - \langle a_1\psi_0, \psi_0\rangle_\omega = \langle x\psi_0, \psi_0\rangle_\omega - a_1\langle \psi_0, \psi_0\rangle_\omega =$$
$$= \langle x\psi_0, \psi_0\rangle_\omega - \frac{\langle x\psi_0, \psi_0\rangle_\omega}{\langle \psi_0, \psi_0\rangle_\omega}\langle \psi_0, \psi_0\rangle_\omega = \langle x\psi_0, \psi_0\rangle_\omega - \langle x\psi_0, \psi_0\rangle_\omega = 0$$

- Supongamos el resultado cierto para $n-1$, es decir

$$\langle \psi_{n-1}, \psi_i\rangle_\omega = 0, \quad \forall i < n-1.$$

Entonces

$$\langle \psi_n, \psi_{n-1}\rangle_\omega = \langle x\psi_{n-1} - a_n\psi_{n-1} - b_n\psi_{n-2}, \psi_{n-1}\rangle_\omega =$$
$$= \langle x\psi_{n-1}, \psi_{n-1}\rangle_\omega - a_n\langle \psi_{n-1}, \psi_{n-1}\rangle_\omega - b_n\langle \psi_{n-2}, \psi_{n-1}\rangle_\omega =$$
$$= \langle x\psi_{n-1}, \psi_{n-1}\rangle_\omega - \frac{\langle x\psi_{n-1}, \psi_{n-1}\rangle_\omega}{\langle \psi_{n-1}, \psi_{n-1}\rangle_\omega}\langle \psi_{n-1}, \psi_{n-1}\rangle_\omega - b_n \cdot 0 = 0$$

Analogamente

$\langle\Psi_n,\Psi_{n-2}\rangle_w=\langle x\Psi_{n-1},\Psi_{n-2}\rangle_w-a_n\langle\Psi_{n-1},\Psi_{n-2}\rangle_w-b_n\langle\Psi_{n-2},\Psi_{n-2}\rangle_w=$

$=\langle x\Psi_{n-1},\Psi_{n-2}\rangle_w-b_n\langle\Psi_{n-2},\Psi_{n-2}\rangle_w=$

$=\langle x\Psi_{n-1},\Psi_{n-2}\rangle_w-\dfrac{\langle x\Psi_{n-1},\Psi_{n-2}\rangle_w}{\langle\Psi_{n-2},\Psi_{n-2}\rangle_w}\langle\Psi_{n-2},\Psi_{n-2}\rangle_w=0.$

Y si $i<n-2$

$\langle\Psi_n,\Psi_i\rangle_w=\langle x\Psi_{n-1},\Psi_i\rangle_w-a_n\langle\Psi_{n-1},\Psi_i\rangle-b_n\langle\Psi_{n-2},\Psi_i\rangle_w=$

$=\langle x\Psi_{n-1},\Psi_i\rangle_w=\langle\Psi_{n-1},x\Psi_i\rangle_w=$

$=\langle\Psi_{n-1},\Psi_{i+1}+a_{i+1}\Psi_i+b_{i+1}\Psi_{i-1}\rangle=0$

por la hipotesis de inducción. csqd.

Si se quiere obtener una sucesión ortonormal basta dividir cada Ψ_k por $\|\Psi_k\|_w$. Los polinomios Ψ_k tienen una notable propiedad que queda reflejada en el siguiente teorema:

4.5. TEOREMA: Los polinomios Ψ_k construidos en el teorema anterior son los mónicos de norma mínima entre los mónicos de grado k. (*)

Demostr.: Se trata de probar que $\|\Psi_k\|_w\leq\|f\|_w$, $\forall f\in\mathbb{P}_k^*$, siendo $\mathbb{P}_k^*$ el conjunto de los polinomios mónicos de grado k.

Todo polinomio $f\in\mathbb{P}_k^*$ se puede expresar en la forma

$$f=\Psi_k+\sum_{i=0}^{k-1}a_i\Psi_i$$

por ser $\{\Psi_i\}_{i=0}^k$ "base" de $\mathbb{P}_k^*$. Se trata entonces de probar que

$$\|\Psi_k\|\leq\|\Psi_k+\sum_{i=0}^{k-1}a_i\Psi_i\|,\quad\forall\sum_{i=0}^{k-1}a_i\Psi_i\in\mathbb{P}_{k-1}\qquad(\mathrm{I})$$

Será suficiente probar que 0 es la aproximación óptima de Ψ_k relativa a $\mathbb{P}_{k-1}$, pues en este caso se tendría

$$\|\Psi_k\|=\inf_{g\in\mathbb{P}_{k-1}}\|\Psi_k-g\|$$

y, en particular, se tendría (I)

Basta entonces probar que $\Psi_k-0\perp\mathbb{P}_{k-1}$, lo cual es cierto pues Ψ_k es ortogonal a todo polinomio de grado menor que k, en virtud del TEOREMA 4.2. csqd.

El interés del teorema que se prueba a continuación es el de facilitar el cálculo del valor numérico de un polinomio, cuando está escrito como combinación lineal de polinomios mónicos ortogonales.

4.6. TEOREMA: Sea $\{\Psi_0, \Psi_1, \ldots, \Psi_n\}$ un conjunto de polinomios mónicos ortogonales tal que grado $\Psi_k = k$. Si $P \in \mathbb{P}_n$, entonces $P = \sum_{k=0}^{n} c_k \Psi_k$ y $P(x)$ se calcula con $2n-1$ multiplicaciones mediante el algoritmo:

$$d_{n+2} = d_{n+1} = 0$$

$$d_k = c_k + (x - a_{k+1}) d_{k+1} - b_{k+2} d_{k+2}, \quad k \leq n$$

siendo a_k y b_k los coeficientes definidos en el teorema 4.4. Se tiene entonces que $P(x) = d_0$.

Demostr.:

$$P(x) = \sum_{k=0}^{n} c_k \Psi_k(x) = \sum_{k=0}^{n} [d_k - (x - a_{k+1}) d_{k+1} + b_{k+2} d_{k+2}] \Psi_k(x) =$$

$$\overset{(1)}{=} d_0 \Psi_0(x) + d_1 [\Psi_1(x) - (x - a_1)\Psi_0(x)] + \sum_{k=2}^{n} d_k [\Psi_k(x) - (x - a_k)\Psi_{k-1}(x) + b_k \Psi_{k-2}(x)]$$

Basta para probar (1) con desarrollar el sumatorio anterior y sacar factor común los d_k.

Pero $\Psi_1(x) - (x - a_1)\Psi_0(x) = 0$, por definición de Ψ_1 y $\Psi_k(x) - (x - a_k)\Psi_{k-1}(x) + b_k \Psi_{k-2}(x) = 0$ por definición de Ψ_k.

Luego $P(x) = d_0 \Psi_0(x) = d_0$, pues $\Psi_0(x) = 1$. csqd.

4.7. TEOREMA: Sea $\{\Psi_0, \ldots, \Psi_n, \ldots\}$ una sucesión de polinomios ortogonales con grado $\Psi_k = k$. Si $f \in \mathcal{C}(I)$ es ortogonal a Ψ_k para $k = 0, 1, \ldots, n-1$ entonces f debe cambiar de signo al menos n veces en $\mathring{I} =]a,b[$ o anularse en todo punto de I.

Demostr.: Como $f \perp \Psi_0$ se tiene que $\int_I f(x)\Psi_0(x)\omega(x)\,dx = 0$ y siendo Ψ_0 constante debe ser $\int_I f(x)\omega(x)\,dx = 0$. De las propiedades de ω se deduce que f debe cambiar de signo al menos una vez en $\mathring{I} =]a,b[$, supuesto $f \neq 0$.

Supongamos que f cambia de signo menos de n veces en $]a,b[$, es decir, supondremos que f sólo cambia de signo en $r_1, \ldots, r_k \in]a,b[$ con $k < n$. Supongamos que $r_1 < \ldots < r_k$. Por supuesto $k \geq 1$.

Entonces f en los intervalos $]a, r_1[,]r_1, r_2[, \ldots,]r_k, b[$ tiene signo constante de forma que en intervalos adyacentes tiene signos opuestos. De esta propiedad goza también el polinomio

$$P(x) = \prod_{i=1}^{k} (x - r_i).$$

Por consiguiente, $\int_I f(x)P(x)\omega(x)\,dx \neq 0$, pues $f(x)P(x)$ tiene signo constante en $]a,b[$. Pero esto es una contradicción, pues siendo P un polinomio de grado k es combinación lineal de

$\Psi_0, \ldots, \Psi_k$ y, por tanto, ortogonal f. Luego f debe cambiar de signo al menos n veces. c.s.q.d.

Este teorema tiene dos corolarios importantes. El primero establece una propiedad de interpolación de aproximación óptima.

4.8. COROLARIO: Sea $P = \sum_{i=0}^{n} c_i \Psi_i$ el polinomio de grado menor o igual que n que es aproximación óptima de $f \in \mathcal{C}(I)$. Entonces P interpola a f en al menos $n+1$ puntos de $\mathring{I} =]a,b[$, es decir, existen $r_1, \ldots, r_{n+1} \in]a,b[$ tales que $f(r_i) = P(r_i)$, $i = 1, \ldots, n+1$.

Demostr.: P es aproximación óptima de f relativa a $\mathbb{P}_n$ si y solo si $f - P \perp \mathbb{P}_n$, o lo que es equivalente, si y solo si

$$f - P \perp \Psi_k, \quad k = 0, 1, \ldots, n$$

El teorema anterior asegura que $f-P$ cambia de signo al menos $n+1$ veces, y, por tanto, siendo $f-P$ continua, debe anularse al menos $n+1$ veces en $]a,b[$. c.s.q.d.

4.9 COROLARIO: Sea $\{\Psi_0, \Psi_1, \ldots, \Psi_k, \ldots\}$ una sucesión de polinomios ortogonales siendo k el grado de Ψ_k, es decir, $\int_I \Psi_i(x)\Psi_j(x)w(x)dx = 0$, si $i \neq j$. Entonces las k raíces de Ψ_k son simples, reales y están en $\mathring{I} =]a,b[$.

Demostr.: Puesto que Ψ_k es ortogonal a todo polinomio de grado menor que k, es decir, Ψ_k es ortogonal a $\Psi_0, \Psi_1, \ldots, \Psi_{k-1}$, se tiene por TEOREMA 4.7 que Ψ_k cambia de signo al menos k veces en $\mathring{I} =]a,b[$. Como grado $\Psi_k = k$, como máximo cambiará k veces de signo. Por tanto, Ψ_k cambia de signo exactamente k veces en $\mathring{I}$ de donde se deduce que Ψ_k tendrá todas sus raíces simples y reales, y están en $\mathring{I}$. c.s.q.d.

C) Estudiamos a continuación las sucesiones de polinomios ortogonales que se obtienen para elecciones concretas del peso $w(x)$.

C.1. POLINOMIOS DE LEGENDRE: Consideremos el intervalo $I = [a,b] = [-1,1]$ y como función peso $w(x) = 1$, $\forall x \in I$. Los polinomios mónicos ortogonales son en este caso

$$\Psi_0(x) = 1$$

$$\Psi_1(x) = x$$

$$\Psi_2(x) = x^2 - \frac{1}{3}$$

$$\Psi_3(x) = x^3 - \frac{3}{5}x$$

$$\cdots\cdots$$

Los polinomios P_n definidos por

$$P_n(x) = \frac{1}{n!\,2^n}\,\frac{(2n)!}{n!}\,\psi_n(x)$$

son clásicos en Análisis y se llaman polinomios de Legendre. Naturalmente, $\{P_0, P_1, \dots, P_n, \dots\}$ constituye también un sistema ortogonal con grado $P_n = n$.

Con respecto a estos polinomios se pueden demostrar los siguientes resultados:

i) $P_n(x) = \dfrac{1}{n!\,2^n}\,\dfrac{d^n}{dx^n}(x^2-1)^n$

ii) $\|P_n\|_\omega = \left(n+\frac{1}{2}\right)^{-1/2}$

iii) P_n es una función par (impar) si n es par (impar).

iv) $P_n(1)=1$, $P_n(-1)=(-1)^n$, $|P_n(x)| \leq 1$.

v) $(1-x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x) = 0$.

vi) $P_{n+1}(x) = \dfrac{2n+1}{n+1}\,x\,P_n(x) - \dfrac{n}{n+1}\,P_{n-1}(x)$, $n \geq 1$.

Esta última relación permite obtener la sucesión $\{P_0, P_1, \dots, P_n, \dots\}$ si se conocen previamente P_0 y P_1 ($P_0(x)=1$, $P_1(x)=x$)

Considerar esta sucesión de polinomios ortogonales presenta claras ventajas sobre el método iterativo válido con caracter general (para cualquier ω) que se expuso anteriormente.

Nótese tambien que dividiendo P_n por su norma, que es $\left(n+\frac{1}{2}\right)^{-1/2}$ se obtiene un sistema ortonormal.

<u>C.2. POLINOMIOS DE Chebyshev (o Tchebycheff)</u>: Sea $I=[a,b]=[-1,1]$ y como función paso $\omega(x) = \dfrac{1}{\sqrt{1-x^2}}$

Los polinomios $T_n(x) = 2^{n-1}\psi_n(x)$ se llaman polinomios de Chebyshev y verifican las siguientes propiedades:

i) $T_n(x) = \cos(n \arccos x) = 2^{n-1}x^n + \dots$, o bien $T_n(x) = \cos n\theta$ si $x = \cos\theta$.

ii) $-1 \leq T_n(x) \leq 1$.

iii) T_n es una funcion par (impar) si n es par (impar).

iv) $T_0(x)=1$, $T_1(x)=x$, $T_{n+1}(x) = 2x\,T_n(x) - T_{n-1}(x)$.

v) $T_n(x) = 0 \iff x = \cos\left[(2k+1)\dfrac{\pi}{2n}\right]$ $k=0,1,\dots,n-1$. (*)

vi) $(1-x^2)T_n''(x) - x\,T_n'(x) + n^2 T_n(x) = 0$.

vii) $T_n(x) = \left(\sqrt{1-x^2}\,\dfrac{d^n}{dx^n}\dfrac{(1-x^2)^n}{\sqrt{1-x^2}}\right)\dfrac{(-2)^n n!}{(2n)!}$.

C.3. POLINOMIOS DE JACOBI: Sea $I = [a,b] = [-1,1]$ y consideramos como función paso $\omega(x) = (1-x)^{\alpha}(1+x)^{\beta}$, $\alpha > -1$, $\beta > -1$.

Si $\alpha = \beta = 0$, $\omega(x) = 1$ y obtenemos los polinomios de Legendre.

Si $\alpha = \beta = -1/2$, $\omega(x) = \dfrac{1}{\sqrt{1-x^2}}$ y obtenemos los polinomios de Chebyshev.

En general, los polinomios

$$J_n(x) = \frac{1}{2^{2n} n!}\binom{2n+\alpha+\beta}{n}\Psi_n(x)$$

Se llaman polinomios de Jacobi. Se verifica que

$$J_n(x) = \frac{(-1)^n}{2^n n!}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}\left[(1-x)^{\alpha+n}(1+x)^{\beta+n}\right].$$

C.4. POLINOMIOS DE LAGUERRE: Sea $I = [0, +\infty[$ y consideramos como función paso $\omega(x) = e^{-\alpha x}$, $\alpha > 0$.

Si $\alpha = 1$, los polinomios $L_n(x) = (-1)^n \Psi_n$ se llaman polinomios de Laguerre. Se verifica lo siguiente:

i) $L_n(x) = e^x \dfrac{d^n}{dx^n}(x^n e^{-x})$

ii) $L_{n+1}(x) = (1+2n-x)L_n(x) - n^2 L_{n-1}(x)$, $L_0(x) = 1$, $L_1(x) = 1-x$.

C.5. POLINOMIOS DE HERMITE: Sea $I =]-\infty, \infty[$ y la función paso $\omega(x) = e^{-\alpha^2 x^2}$. Si $\alpha = 1$, los polinomios $H_n(x) = 2^n \Psi_n(x)$ se llaman polinomios de Hermite. Se verifican las siguientes propiedades

i) $H_n(x) = (-1)^n e^{+x^2}\dfrac{d^n}{dx^n}(e^{-x^2})$

ii) $H_0(x) = 1$, $H_1(x) = 2x$, $H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x)$.

5. APROXIMACION UNIFORME EN $\mathcal{C}(I)$.

A) Recibe el nombre de aproximación uniforme en $\mathcal{C}(I)$ la aproximación en el espacio $\mathcal{C}(I)$ con la norma del supremo

$$\|f\| = \max_{x \in I} |f(x)|.$$

$\mathcal{C}(I)$ con esta norma es un espacio de Banach no estrictamente convexo. Por consiguiente, no todos sus subespacios son semi-Chebyshev. Probaremos, sin embargo, que sí lo son una clase de subespacios, llamados de Haar, entre los que se encuentran los polinomios de grado menor o igual que n, $\forall n \in \mathbb{N}$.

Veamos a continuación una serie de definiciones y resultados que utilizaremos más adelante. Sea V un espacio vectorial.

B) <u>DEFINICION:</u> Un subconjunto C de V se dice que es convexo si $\forall x,y \in C$, $\{\lambda x + (1-\lambda)y \,/\, \lambda \in [0,1]\} \subset C$.

<u>5.1. PROPOSICION:</u> $C \subset V$ es convexo si y solo si $\sum_{i=1}^{m} \theta_i x_i \in C$ siempre que $x_i \in C$ y θ_i son números reales positivos tales que $\sum_{i=1}^{m} \theta_i = 1$.

<u>DEFINICION:</u> Sea A un subconjunto cualquiera de V. Se denomina envolvente convexa de A al subconjunto $C(A)$ de V definido por

$$C(A) = \{\sum_{i=1}^{m} \theta_i x_i \,/\, x_i \in A,\ \theta_i \in \mathbb{R},\ \theta_i \geq 0,\ \sum_{i=1}^{m} \theta_i = 1,\ m \in \mathbb{N}\}.$$

<u>5.2. PROPOSICION:</u> a) $A \subset C(A)$.

b) $C(A)$ es convexo.

c) $C(A)$ es el menor convexo que contiene a A.

Por la definición de $C(A)$ se ve que cualquier elemento es de la forma $\sum_{i=1}^{m} \theta_i x_i$, $\theta_i \geq 0$, $\sum_{i=1}^{m} \theta_i = 1$, es decir, lo que se llama una combinación lineal convexa de elementos de A. El número de sumandos es finito, pero arbitrario. Sin embargo, si $\dim V = n$, entonces $n+1$ puntos de A son suficientes para obtener cualquier punto de $C(A)$, como se prueba en el siguiente teorema.

<u>5.3. TEOREMA:</u> (de Caratheodory)

Si $\dim V = n$ y $A \subset V$, entonces cada punto de $C(A)$ es una combinación lineal convexa de $n+1$ o menos elementos de A.

Demostr.: Si $y \in C(A)$, entonces $y = \sum_{i=0}^{k} \theta_i x_i$ con $\theta_i \geq 0$ y $\sum_{i=0}^{k} \theta_i = 1$, $x_i \in A$. Supongamos que k es mínimo en el sentido de que $\theta_i > 0$, $i = 0,1,\dots,k$. Puesto que $(\sum_{i=0}^{k} \theta_i x_i) - y = (\sum_{i=0}^{k} \theta_i x_i) - y \sum_{i=0}^{k} \theta_i = \sum_{i=0}^{k} \theta_i (x_i - y) = 0$, el conjunto de vectores $y_i = x_i - y$, $i = 0,\dots,k$ es linealmente dependiente. Si fuese $k > n$, el conjunto $\{y_1,\dots,y_k\}$ será linealmente dependiente. Existen entonces $\alpha_1,\dots,\alpha_k$ no todos nulos tales que $\sum_{i=1}^{k} \alpha_i y_i = 0$. Tomando $\alpha_0 = 0$ se tiene que para cada λ

$$\sum_{i=0}^{k} (\theta_i + \lambda \alpha_i) y_i = 0 \qquad (I)$$

Entre todos los λ que anulan algún coeficiente $\theta_i + \lambda\alpha_i$ (*) elegimos el de módulo más pequeño. Entonces los restantes coeficientes son mayores o iguales que cero, y no todos cero, pues

$$\theta_0 + \lambda \alpha_0 = \theta_0 > 0.$$

Para ese λ, el sumatorio $\sum_{i=0}^{k}(\theta_i+\lambda\alpha_i)y_i$ tiene a lo sumo K sumandos no nulos.

Por (I), $\sum_{i=0}^{k}(\theta_i+\lambda\alpha_i)y_i=0$, o bien, sustituyendo y_i por x_i-y obtenemos

$$y\sum_{i=0}^{k}(\theta_i+\lambda\alpha_i)=\sum_{i=0}^{k}(\theta_i+\lambda\alpha_i)x_i$$

y por tanto $y=\sum_{i=0}^{k}\frac{1}{\sum_{i=0}^{k}(\theta_i+\lambda\alpha_i)}(\theta_i+\lambda\alpha_i)x_i$, pues $\sum_{i=0}^{k}(\theta_i+\lambda\alpha_i)>0$.

Tenemos así una combinación lineal convexa de y de, a lo sumo, k sumandos, lo cual es una contradicción con la hipótesis de que K era mínimo (en el sentido de que toda otra combinación lineal convexa de y, tiene K+1 sumandos o más). Debe ser entonces $K\leq n$. csqd.

OBSERVACION: Los n+1 puntos cuya existencia asegura el teorema anterior son distintos, en general, para cada $y\in C(A)$.

5.4. COROLARIO: En un espacio de dimensión finita la envolvente convexa de un conjunto compacto es compacto.

Demostr.: Sea N un espacio normado de dimensión finita y $X\subset N$ un conjunto compacto. Probaremos que C(X) verifica la propiedad de Bolzano-Weierstrass, es decir, que toda sucesión en C(X) admite una sucesión parcial convergente. (*)

Sea $\{v_K\}_K$ una sucesión en C(X). Por el teorema anterior, para cada $K\in\mathbb{N}$, $v_K=\sum_{i=0}^{n}\theta_{Ki}x_{Ki}$ donde $\theta_{Ki}\geq 0$, $\sum_{i=0}^{n}\theta_{Ki}=1$ y $x_{Ki}\in X$.

Consideremos las sucesiones

$\{x_{K0}\}_K, \{x_{K1}\}_K, \ldots, \{x_{Kn}\}_K, \{\theta_{K0}\}_K, \ldots, \{\theta_{Kn}\}_K$.

Como X es compacto, y el conjunto

$\{(\theta_0,\theta_1,\ldots,\theta_n)\in\mathbb{R}^{n+1} / \theta_i\geq 0, \sum_{i=0}^{n}\theta_i=1\}$

también es compacto, deberá existir una sucesión de naturales $\{K_j\}_j$ tal que

$\lim_{j\to\infty} x_{K_j i}=x_i\in X$, $i=0,1,\ldots,n$

$\lim_{j\to\infty}\theta_{K_j i}=\theta_i$, $i=0,1,\ldots,n$

Se comprueba fácilmente que $\theta_i\geq 0$, $\sum_{i=0}^{n}\theta_i=1$ y que

$$\lim_{j\to\infty} v_{K_j}=\sum_{i=0}^{n}\theta_i x_i\in C(X).\ \text{csqd}$$

Por tanto, C(X) es compacto. csqd

c) Sabemos que en un espacio de Banach uniformemente convexo, todo subespacio cerrado es Chebyshev. Este resultado también es válido para convexos cerrados; demostraremos esto en el caso de convexos cerrados en espacios de Hilbert.

5.5. TEOREMA: Sea H un espacio de Hilbert y $K\subset H$ un convexo cerrado. Existe entonces un único punto en K en el que se alcanza la norma mínima.

Demostr.: Existencia: Sea $d=\inf_{x\in K}\|x\|$.

Existirá entonces una sucesión $\{x_n\}_n$ en K tal que $\lim_{n\to\infty}\|x_n\|=d$.

Por la igualdad del paralelogramo

$$\|x_i-x_j\|^2=2\|x_i\|^2+2\|x_j\|^2-4\|\tfrac{1}{2}(x_i+x_j)\|^2,\ \forall i,j\in\mathbb{N}.$$

Como K es convexo, $\frac{1}{2}(x_i+x_j)\in K$ y por tanto $\|\frac{1}{2}(x_i+x_j)\|\geq d$.

Por tanto $\|x_i-x_j\|^2\leq 2\|x_i\|^2+2\|x_j\|^2-4d^2$. (1)

Tomando límites en (1) cuando $i,j\to\infty$ se ve que $\{x_n\}_n$ es una sucesión de Cauchy.

La completitud del espacio asegura la convergencia de $\{x_n\}_n$ a un cierto x_0, que pertenecerá a K por ser K cerrado.

Por la continuidad de la norma tendremos que $\lim_n\|x_n\|=\|x_0\|$ y, por tanto, $\|x_0\|=d$.

Unicidad: Supongamos que existen $x,y\in K$, $x\neq y$, tales que $\|x\|=\|y\|=d$.

Entonces, por la igualdad del paralelogramo

$$\|x-y\|^2=2\|x\|^2+2\|y\|^2-4\|\tfrac{1}{2}(x+y)\|^2=4d^2-4\|\tfrac{1}{2}(x+y)\|^2.\quad(2)$$

Pero $\|\frac{1}{2}(x+y)\|\leq\frac{1}{2}\|x\|+\frac{1}{2}\|y\|=d$

y por otra parte, puesto que $\frac{1}{2}(x+y)\in K$, $\|\frac{1}{2}(x+y)\|\geq d$.

Luego $\|\frac{1}{2}(x+y)\|=d$. De (2) se deduce que $\|x-y\|^2=0$ y, por tanto, $x=y$ contra lo supuesto. Luego, la norma mínima se alcanza en un solo punto. csqd.

D) En adelante emplearemos la notación $\langle u,z\rangle$ para representar el producto escalar ordinario en $\mathbb{R}^n$: $\langle u,z\rangle=\sum_{i=1}^{n}u_i\cdot z_i$.

DEFINICION: Sea $\mathcal{U}$ un subconjunto de $\mathbb{R}^n$. Diremos que el conjunto de desigualdades

$$\langle u,z\rangle>0\ ,\quad u\in\mathcal{U}$$

es inconsistente cuando no existe $z\in\mathbb{R}^n$ que las verifique simultáneamente.

Veamos en relación con esto el siguiente teorema.

5.6. TEOREMA: Sea $\mathcal{U}$ un subconjunto compacto de $\mathbb{R}^n$. Condición necesaria y suficiente para que el sistema de desigualdades

$$<u, z> \, > 0 \ , \quad u \in \mathcal{U}$$

sea inconsistente es que $0 \in \mathbb{C}(\mathcal{U})$.

Demostr.: Suficiencia: Si $0 \in C(\mathcal{U})$ existen $u_1, \ldots, u_m \in \mathcal{U}$ y $\lambda_1, \ldots, \lambda_m \geq 0$ con $\sum_{i=1}^{m} \lambda_i = 1$ tales que $\sum_{i=1}^{m} \lambda_i u_i = 0$.

Por consiguiente, $\sum_{i=1}^{m} \lambda_i <u_i, z> = 0$, $\forall z \in \mathbb{R}^n$. (I)

No existirá entonces $z \in \mathbb{R}^n$ tal que $<u, z> \, > 0$, $\forall u \in \mathcal{U}$, pues de existir no se verificaría (I).

Necesidad: Supongamos que $0 \notin \mathbb{C}(\mathcal{U})$. $C(\mathcal{U})$ es convexo y cerrado, pues $C(\mathcal{U})$ es compacto, por serlo $\mathcal{U}$. Entonces por el teorema anterior existe un único $z \in C(\mathcal{U})$ de norma mínima. Como $0 \notin C(\mathcal{U})$, $\|z\| > 0$.

Dado $u \in \mathcal{U}$, por ser $C(\mathcal{U})$ convexo, $\rho u + (1-\rho) z \in C(\mathcal{U})$ si $\rho \in [0,1]$.

Por tanto, $\|z\| \leq \|\rho u + (1-\rho) z\|$.

Luego $0 \leq \|\rho u + (1-\rho) z\|^2 - \|z\|^2$. (II)

Pero $\|\rho u + (1-\rho) z\|^2 - \|z\|^2 = \|\rho(u-z) + z\|^2 - \|z\|^2 =$

$= <\rho(u-z) + z, \rho(u-z) + z> - <z, z> = \rho^2 <u-z, u-z> + 2\rho <u-z, z> =$

$= \rho^2 \|u-z\|^2 + 2\rho <u-z, z>$.

De (II) se deduce entonces que

$$0 \leq \rho^2 \|u-z\|^2 + 2\rho <u-z, z>.$$

Debe ser entonces $<u-z, z> \, \geq 0$, pues si $<u-z, z> \, < 0$, tomando ρ suficientemente pequeño podemos conseguir que

$$\rho \|u-z\|^2 < 2 |<u-z, z>| = -2 <u-z, z>$$

y se tendría entonces

$$0 \leq \rho^2 \|u-z\|^2 + 2\rho <u-z, z> \, < \rho(-2 <u-z, z>) + 2\rho <u-z, z> = 0$$

lo cual es absurdo.

Por tanto, $<u-z, z> \, \geq 0$, $\forall u \in \mathcal{U}$.

Luego $<u, z> \, \geq \, <z, z> = \|z\|^2 > 0$, $\forall u \in \mathcal{U}$

y esto contradice que el sistema $<u, t> \, > 0$, $u \in \mathcal{U}$ es inconsistente.

Debe ser entonces $0 \in \mathbb{C}(\mathcal{U})$. c.s.q.d.

E) Volvamos ahora al problema de la aproximación uniforme en $C(I)$.

Si L es un subespacio de dimensión finita toda función $f \in C(I)$ tiene al menos una aproximación óptima relativa a L (Record[illegible] ción suficiente para que L sea proximinal ($P_L(f) \neq \emptyset$, $\forall f \in C(I)$) [illegible] que son aproximación óptima de una $f\in C(I)$ dada.

5.7 TEOREMA: Sea $\{g_1,\dots,g_n\}$ una base de L. Entonces $g_0=\sum_{i=1}^n c_ig_i$ es aproximación óptima uniforme de $f\in C(I)$ relativa a L si y solo si

$$0\in C(\{r(x)\hat{x}\in\mathbb{R}^n / |r(x)|=\|r\|\})$$

donde $\hat{x}=(g_1(x),\dots,g_n(x))$ y $r=g_0-f$.

Demostr.:

$\Leftarrow$) Supongamos que g_0 no es aproximación óptima. Esto implica la existencia de $\tilde{q}=\sum_{i=1}^n e_ig_i$ tal que $\|\sum_{i=1}^n e_ig_i-f\|<\|\sum_{i=1}^n c_ig_i-f\|$

Sea $d_i=c_i-e_i$, $i=1,\dots,n$, y $X_0=\{x\in I / |r(x)|=\|r\|\}$.

Si $x\in X_0$, entonces

$$(r(x)-\sum_{i=1}^n d_ig_i(x))^2=(\sum_{i=1}^n c_ig_i(x)-f(x)-\sum_{i=1}^n d_ig_i(x))^2=$$

$$=(\sum_{i=1}^n e_ig_i(x)-f(x))^2\leq\|\sum_{i=1}^n e_ig_i-f\|^2<\|\sum_{i=1}^n c_ig_i-f\|^2=\|r\|^2=(r(x))^2$$

de donde se obtiene que

$$r(x)\cdot\sum_{i=1}^n d_ig_i(x)=r(x)<d,\hat{x}>\ >0,\quad \text{si } x\in X_0,\quad (1)$$

donde $d=(d_1,\dots,d_n)$, pues $r(x)^2-2r(x)<d,\hat{x}>+<d,\hat{x}>^2<(r(x))^2$, según (1) y por tanto $<d,\hat{x}>^2<2r(x)<d,\hat{x}>$, que prueba que $r(x)<d,\hat{x}>\ >0$, si $x\in X_0$

Entonces, como $r(x)<d,\hat{x}>\ =\ <d,r(x)\hat{x}>$, resulta que el conjunto de desigualdades $<z,r(x)\hat{x}>\ >0$, $x\in X_0$ no son inconsistentes.

Por otra parte, el conjunto $\{r(x)\hat{x}/x\in X_0\}$ es compacto.

En efecto, vamos a probarlo. Como $X_0\subset I$ e I es compacto es suficiente ver que X_0 es cerrado. Para probar esto demostraremos que X_0 contiene los límites de todas las sucesiones convergentes de puntos de X_0. Sea $\{x_n\}\subset X_0$ tal que $\{x_n\}\to\bar{x}$. Como $|r(x_n)|=\|r\|$ se verifica que $|r(\bar{x})|=\|r\|$ y, por tanto, $\bar{x}\in X_0$. Luego X_0 es compacto. Consideremos la aplicación continua

$$r(\cdot)g(\cdot): x\in X_0\longmapsto(r(x)g_1(x),\dots,r(x)g_n(x))\in\mathbb{R}^n$$

Por tanto, $\{r(x)\hat{x}/x\in X_0\}$ es compacto como imagen continua de un compacto.

El teorema 5.6. asegura en estas hipótesis que $0\notin C(\{r(x)\hat{x}/x\in X_0\})$, lo cual contradice la hipótesis. Por tanto, g_0 es aproximación óptima de f.

$\Rightarrow$) Supongamos ahora que $0\notin C(\{r(x)\hat{x}/x\in X_0\})$. Por el teorema anterior existe una n-tupla $d=(d_1,\dots,d_n)$ tal que $<d,r(x)\hat{x}>\ >0$, $\forall x\in X_0$.

Como X_0 es compacto, el número $\varepsilon=\min_{x\in X_0} r(x)<d,\hat{x}>$ es mayor que cero.

Definimos $X_1 = \{x \in I / r(x) < d, \hat{x} > \leq \varepsilon/2\}$. Supongamos $X_1 \neq \emptyset$.
Es fácil ver que X_1 es un cerrado que no contiene puntos de X_0. Como además es acotado, se sigue que $|r(x)|$ alcanza su supremo E sobre X_1 en un punto de X_1, es decir, puesto que X_1 es compacto y $|r(x)|$ continua en X_1 existe $x_1 \in X_1$ tal que $E = \sup_{x \in X_1} |r(x)| = |r(x_1)|$. Se verifica además que $|r(x_1)| < \|r\|$, pues $X_1 \cap X_0 = \emptyset$ y la igualdad $|r(x)| = \|r\|$ se daba solo en los puntos de X_0.

Si probamos que existe $\lambda > 0$ tal que $\|r - \lambda \sum_{i=1}^n d_i g_i\| < \|r\|$ se tendría que g_0 no es aproximación óptima uniforme de f pues

$$\|r - \lambda \sum_{i=1}^n d_i g_i\| = \|\sum_{i=1}^n c_i g_i - f - \lambda \sum_{i=1}^n d_i g_i\| = \|\sum_{i=1}^n (c_i - \lambda d_i) g_i - f\| < \|g_0 - f\| = \|r\|$$

lo cual significa que existe $\sum_{i=1}^n (c_i - \lambda d_i) g_i (\neq g_0)$ que dista de f menos que g_0.

Probemos que existe un $\lambda > 0$ como el descrito.
Sea $x \in X_1$ y $\lambda \in \mathbb{R}$ tal que $0 < \lambda < \dfrac{\|r\| - E}{\|\sum_{i=1}^n d_i g_i\|}$

Entonces $|r(x) - \lambda \sum_{i=1}^n d_i g_i(x)| \leq |r(x)| + \lambda |\sum_{i=1}^n d_i g_i(x)| \leq E + \lambda \|\sum_{i=1}^n d_i g_i\| < \|r\|$

Sea ahora $x \notin X_1$ y λ tal que $0 < \lambda < \dfrac{\varepsilon}{\|\sum_{i=1}^n d_i g_i\|^2}$

Entonces

$$(r(x) - \lambda \sum_{i=1}^n d_i g_i(x))^2 = (r(x))^2 - 2\lambda r(x) < d, \hat{x} > + \lambda^2 (\sum_{i=1}^n d_i g_i(x))^2 <$$

$$\overset{(*)}{<} \|r\|^2 + \varepsilon \lambda + \lambda^2 \|\sum_{i=1}^n d_i g_i\|^2 = \|r\|^2 + \lambda(-\varepsilon + \lambda \|\sum_{i=1}^n d_i g_i\|^2) \overset{(*)}{<} \|r\|^2$$

Por tanto, tomando λ tal que $0 < \lambda < \min\left\{\dfrac{\|r\| - E}{\|\sum_{i=1}^n d_i g_i\|}, \dfrac{\varepsilon}{\|\sum_{i=1}^n d_i g_i\|^2}\right\}$

queda visto que $\|r - \lambda \sum_{i=1}^n d_i g_i\| < \|r\|$ y por tanto que g_0 no es aproximación óptima uniforme de f relativa a L, contra lo supuesto.
Se verificará entonces que $0 \in C(\{r(x)\hat{x} / x \in X_0\})$. c.s.q.d.

6. Subespacios de Haar en $C(I)$.

A) Definiremos en este apartado ciertos subespacios de $C(I)$, entre los que se encuentran los P_n, $n \in \mathbb{N}$, para los que esta caracterización (Teorema 5.7) se da de una forma más conveniente. Por otra parte, respecto a esos subespacios se dará la unicidad de aproximación óptima, como veremos más adelante.

DEFINICION: Un subespacio de $C(I)$ de dimensión n se dice que es de Haar si toda función distinta de cero se anula en, a lo sumo, $n-1$ puntos de I.

6.1. TEOREMA: Son equivalentes las siguientes proposiciones:

1) L es un subespacio de Haar de $C(I)$ de dimensión n.

2) Para toda base de L $\{\varphi_1,\dots,\varphi_n\}$ y todo sistema de n puntos distintos $\{x_1,\dots,x_n\}\subset I$ los vectores $(\varphi_1(x_1),\dots,\varphi_n(x_1)),\dots,(\varphi_1(x_n),\dots,\varphi_n(x_n))$ son independientes, es decir,

$$D(x_1,\dots,x_n)=\det\begin{pmatrix}\varphi_1(x_1) & \dots & \varphi_n(x_1)\\ - & - & -\\ \varphi_1(x_n) & \dots & \varphi_n(x_n)\end{pmatrix}\neq 0 .$$

Demostr.: 1) ⇒ 2) Supongamos que 2) no es cierta, es decir, que existen una base $\{\varphi_1,\dots,\varphi_n\}$ de L y $\{x_1,\dots,x_n\}\subset I$, $x_i\neq x_j$ si $i\neq j$ tal que

$$\det\begin{pmatrix}\varphi_1(x_1) & \dots & \varphi_n(x_1)\\ - & - & -\\ \varphi_1(x_n) & \dots & \varphi_n(x_n)\end{pmatrix}=0$$

Entonces el sistema homogéneo en las incógnitas $c_1,\dots,c_n$

$$\begin{cases} c_1\varphi_1(x_1)+\dots+c_n\varphi_n(x_1)=0\\ - - - - - - -\\ c_1\varphi_1(x_n)+\dots+c_n\varphi_n(x_n)=0\end{cases}$$

tiene una solución $\bar c_1,\dots,\bar c_n$ no trivial.

Como $\{\varphi_1,\dots,\varphi_n\}$ es una base de L, la función $f=\sum_{i=1}^n \bar c_i\varphi_i$ es no nula. Se deduce entonces que L no es de Haar, pues f es una función no nula de L que tiene al menos n raíces, contra la hipótesis.

2) ⇒ 1) Si L no es un subespacio de Haar, existirá $\varphi\in L-\{0\}$ tal que φ se anula en n puntos $x_1,\dots,x_n$ distintos de I.

Sea $\{\varphi_1,\dots,\varphi_n\}$ una base de L. Existen entonces $\bar c_1,\dots,\bar c_n\in\mathbb{R}$ tales que $\varphi=\sum_{i=1}^n \bar c_i\varphi_i$. Además los $\bar c_i$ no son todos nulos, pues $\varphi\neq 0$.

Entonces el sistema

$$\begin{cases} c_1\varphi_1(x_1)+\dots+c_n\varphi_n(x_1)=0\\ - - - - - -\\ c_1\varphi_1(x_n)+\dots+c_n\varphi_n(x_n)=0\end{cases}$$

tiene soluciones distintas de la trivial y, por tanto

$$\det\begin{pmatrix}\varphi_1(x_1) & \dots & \varphi_n(x_1)\\ - & - & -\\ \varphi_1(x_n) & \dots & \varphi_n(x_n)\end{pmatrix}=0$$

en contra de la hipótesis.

L debe ser, entonces, un subespacio de Haar. c.q.d.

6.2. COROLARIO: Las proposiciones 1) y 2) del teorema anterior son equivalentes a la proposición siguiente

3) Existe al menos una base de L $\{\varphi_1,\dots,\varphi_n\}$ tal que para todo sistema de n puntos distintos $\{x_1,\dots,x_n\}\subset I$ los vectores $(\varphi_1(x_1),\dots,\varphi_n(x_1)),\dots,(\varphi_1(x_n),\dots,\varphi_n(x_n))$ son independientes.

La demostración es trivial.

DEFINICION: Un conjunto $\{\varphi_1,\dots,\varphi_n\}\subset \mathcal{C}(I)$ se llama un sistema de Haar si para todo sistema de n puntos distintos $\{x_1,\dots,x_n\}\subset I$ los vectores $(\varphi_1(x_1),\dots,\varphi_n(x_1)),\dots,(\varphi_1(x_n),\dots,\varphi_n(x_n))$ son linealmente independientes.

6.3. TEOREMA: Si $\{\varphi_1,\dots,\varphi_n\}\subset \mathcal{C}(I)$ es un sistema de Haar, entonces es un conjunto de vectores linealmente independientes en $C(I)$.

Demostr.: Supongamos que la tesis es falsa. Existirán entonces $\bar{c}_1,\dots,\bar{c}_n$ no todos nulos tales que $\bar{c}_1\varphi_1+\dots+\bar{c}_n\varphi_n=0$, y para cada sistema $\{x_1,\dots,x_n\}$ de n puntos distintos dos a dos de I el sistema homogeneo

$$\begin{cases} c_1\varphi_1(x_1)+\dots+c_n\varphi_n(x_1)=0 \\ \dots\dots\dots \\ c_1\varphi_1(x_n)+\dots+c_n\varphi_n(x_n)=0 \end{cases}$$

tendrá la solución no trivial $\bar{c}_1,\dots,\bar{c}_n$, para lo cual es necesario que

$$\det\begin{pmatrix}\varphi_1(x_1) & \dots & \varphi_n(x_1)\\ \dots & & \dots \\ \varphi_1(x_n) & \dots & \varphi_n(x_n)\end{pmatrix}=0$$

Entonces las filas no serán independientes y $\{\varphi_1,\dots,\varphi_n\}$ no sería un sistema de Haar, contra lo supuesto. ∎

6.4. COROLARIO: Un subespacio de $C(I)$ de dimensión n es de Haar si, y sólo si, admite un sistema de Haar con n elementos.

B) Para obtener un teorema de caracterización de aproximaciones óptimas uniformes, respecto de subespacios de Haar, damos el siguiente

6.5. LEMA: Sea $\{g_1,\dots,g_n\}$ un sistema de Haar. Sean $a\le x_0<x_1<\dots<x_n\le b$ un sistema de $n+1$ puntos y sean $\lambda_0,\dots,\lambda_n$ constantes no nulas. Entonces $0\in C(\{\lambda_0\hat{x}_0,\dots,\lambda_n\hat{x}_n\})$ si, y solo si, $\lambda_i\lambda_{i-1}<0$, $i=1,\dots,n$.

Demostr.: Notamos en primer lugar que los determinantes $D(x_1,\dots,x_n)=$
$$=\det\begin{pmatrix} g_1(x_1) & \dots & g_n(x_1)\\ \dots & & \dots \\ g_1(x_n) & \dots & g_n(x_n)\end{pmatrix}$$
tienen signo constante siempre que $x_1<\dots<x_n$. Además no se anulan por ser $\{g_1,\dots,g_n\}$ un sistema de Haar. En efecto, tenemos

mismo signo, pues si existiesen $x_1<\dots<x_n$ e $y_1<\dots<y_n$ tales que

$$D(x_1,\dots,x_n)<0<D(y_1,\dots,y_n)$$

entonces la función real de variable real continua

$$\lambda\in[0,1]\longmapsto D(\lambda x_1+(1-\lambda)y_1,\dots,\lambda x_n+(1-\lambda)y_n)\in\mathbb{R}.$$

tomaría signos opuestos en los extremos del intervalo $[0,1]$ con lo cual existirá $\bar\lambda\in]0,1[$ tal que

$$D(\bar\lambda x_1+(1-\bar\lambda)y_1,\dots,\bar\lambda x_n+(1-\bar\lambda)y_n)=0$$

en contra de que $\{g_1,\dots,g_n\}$ es un sistema de Haar.

$\Longrightarrow$ | $0\in C(\{\lambda_0\hat x_0,\dots,\lambda_n\hat x_n\})$ si, y solo si, existen $\theta_0\geq 0,\dots,\theta_n\geq 0$ con $\sum_{i=0}^n\theta_i=1$ tales que $\sum_{i=0}^n\theta_i\lambda_i\hat x_i=0$, que, supuesto $\theta_0\neq 0$, se puede escribir en la forma

$$\hat x_0=\sum_{i=1}^n\left(-\frac{\theta_i\lambda_i}{\theta_0\lambda_0}\hat x_i\right)$$

Como $\hat x_i=(g_1(x_i),\dots,g_n(x_i))$, $i=0,1,\dots,n$ se tiene que

$$\begin{cases}-\dfrac{\theta_1\lambda_1}{\theta_0\lambda_0}g_1(x_1)-\cdots-\dfrac{\theta_n\lambda_n}{\theta_0\lambda_0}g_1(x_n)=g_1(x_0)\\ \text{---------------------}\\ -\dfrac{\theta_1\lambda_1}{\theta_0\lambda_0}g_n(x_1)-\cdots-\dfrac{\theta_n\lambda_n}{\theta_0\lambda_0}g_n(x_n)=g_n(x_0)\end{cases}$$

Luego $\exists\,\theta_0\geq 0,\dots,\theta_n\geq 0$ con $\sum_{i=0}^n\theta_i=1$ tales que $a_1=-\dfrac{\theta_1\lambda_1}{\theta_0\lambda_0},\dots,a_n=-\dfrac{\theta_n\lambda_n}{\theta_0\lambda_0}$

son soluciones del sistema

$$\begin{cases}a_1g_1(x_1)+\cdots+a_ng_1(x_n)=g_1(x_0)\\ \text{---------------}\\ a_1g_n(x_1)+\cdots+a_ng_n(x_n)=g_n(x_0)\end{cases}$$

El determinante de este sistema es $D(x_1,\dots,x_n)$ que es no nulo por ser $\{g_i\}_{i=1}^n$ un sistema de Haar. Por tanto, por la regla de Cramer

$$a_i=-\frac{\theta_i\lambda_i}{\theta_0\lambda_0}=\frac{\begin{vmatrix}g_1(x_1)\cdots \overset{i}{g_1(x_0)}\cdots g_1(x_n)\\ \text{--------------}\\ g_n(x_1)\cdots g_n(x_0)\cdots g_n(x_n)\end{vmatrix}}{D(x_1,\dots,x_n)}=\frac{D(x_1,\dots,x_{i-1},x_0,x_{i+1},\dots,x_n)}{D(x_1,\dots,x_n)},\quad i=1,\dots,n$$

Cambiando en el numerador las columnas para que los x_j aparezcan ordenados obtenemos que

$$-\frac{\theta_i\lambda_i}{\theta_0\lambda_0}=a_i=\frac{(-1)^{i-1}D(x_0,x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)}{D(x_1,\dots,x_n)}$$

Como los determinantes del numerador y denominador tienen el mismo signo según se indicó al principio se tiene que

$$\text{signo}\left(-\frac{\theta_i\lambda_i}{\theta_0\lambda_0}\right)=(-1)^{i-1}$$

y por tanto $\operatorname{sig} \lambda_i = (-1)^i \operatorname{sig} \lambda_0$

Se deduce ya de aquí que $\lambda_i \lambda_{i-1} < 0$, $i=1,\dots,n$.

$\Leftarrow$] Por hipótesis $\lambda_i \lambda_{i-1} < 0$, $i=1,\dots,n$, o lo que es equivalente $\operatorname{sign} \frac{\lambda_i}{\lambda_0} = (-1)^i$.

Queremos probar que $0 \in C(\{\lambda_0 \hat{x}_0, \dots, \lambda_n \hat{x}_n\})$, o bien que existen $\theta_0 \geq 0, \dots, \theta_n \geq 0$, con $\sum_{i=0}^{n} \theta_i = 1$ tales que

$$-\frac{\theta_i \lambda_i}{\theta_0 \lambda_0} = (-1)^{i-1} \frac{D(x_0, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)}{D(x_1, \dots, x_n)}, \quad i=1,\dots,n$$

Si hacemos $a_i = (-1)^{i-1} \frac{D(x_0, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)}{D(x_1, \dots, x_n)}$, queremos probar que existen los $\theta_i \geq 0$ tales que $\sum_{i=0}^{n} \theta_i = 1$ y $\theta_i = -\theta_0 \frac{\lambda_0}{\lambda_i} a_i$.

Los λ_i, por hipótesis, son no nulos.

Definimos $\theta_0 = 1$ y $\theta_i = -\frac{\lambda_0}{\lambda_i} a_i$.

Evidentemente verifican que $\theta_i = -\theta_0 \frac{\lambda_0}{\lambda_i} a_i$. Además son estrictamente positivos pues

$$\operatorname{sig} \theta_i = \operatorname{sig}\left(-\frac{\lambda_0}{\lambda_i} a_i\right) = \operatorname{sig}\left(-\frac{\lambda_0}{\lambda_i}\right) \cdot \operatorname{sig} a_i = (-1)^{i-1} \cdot (-1)^{i-1} = 1$$

No verifican, sin embargo, que $\sum_{i=0}^{n} \theta_i = 1$.

Consideremos entonces

$$\theta_i' = \frac{\theta_i}{\sum_{i=0}^{n} \theta_i} > 0, \quad i=0,\dots,n.$$

Son todos positivos, su suma vale 1 y además $a_i = -\frac{\theta_i' \lambda_i}{\theta_0' \lambda_0}$, $i=1,\dots,n$

lo cual equivale, como sabemos, a que $0 \in C(\{\lambda_0 \hat{x}_0, \dots, \lambda_n \hat{x}_n\})$. c.s.q.d.

El teorema fundamental de los dos últimos apartados caracteriza una aproximación óptima respecto de un subespacio de Haar. Es el siguiente:

6.6. TEOREMA: Sea $\{g_1, \dots, g_n\}$ un sistema de Haar en $\mathcal{C}(I)$. Entonces $g_0 = \sum_{i=1}^{n} c_i g_i$ es aproximación óptima de f relativa a $L = \langle g_1, \dots, g_n \rangle$ si, y solo si, la función error

$$r : x \in I \longmapsto r(x) = g_0(x) - f(x)$$

tiene en I al menos $n+1$ alternancias, es decir, existen $x_0 < x_1 < \dots < x_n$ tales que $r(x_i) = -r(x_{i-1}) = \pm \|r\|$, $i=1,\dots,n$.

Demostr.: Por el teorema 5.7. g_0 es aproximación óptima si, y solo si, el cero está en la envolvente convexa del conjunto $\{r(x)\hat{x} / |r(x)| = \|r\|\}$. Por el teorema de Caratheodory existen $K \leq n$ y unos puntos $x_0, \dots, x_K$ y unos escalares $\lambda_0, \dots, \lambda_K$ positivos con $\sum_{i=0}^{K} \lambda_i = 1$ tales que $0 = \sum_{i=0}^{K} \lambda_i r(x_i) \hat{x}_i$, y $|r(x_i)| = \|r\|$, $i=0,\dots,K$

Supondremos $\|r\|\neq 0$, pues de lo contrario sería $f=g_0$ y no sería preciso aproximar. Entonces $r(x_i)\neq 0$, $i=0,\dots,k$. Como los λ_i no serán todos nulos, el sistema de vectores $\{\hat{x}_0,\hat{x}_1,\dots,\hat{x}_k\}$ es ligado.
Como $\hat{x}_i=(g_1(x_i),\dots,g_n(x_i))$, al ser $\{g_i\}_{i=1}^n$ sistema de Haar

$$D(x_0,\dots,x_{n-1})\neq 0$$

Como $\hat{x}_0,\dots,\hat{x}_{n-1}$ son linealmente independientes y $\hat{x}_0,\dots,\hat{x}_k$ son linealmente dependientes debe ser $k>n-1$ y, por tanto, $k\geq n$.

Será entonces $k=n$.

Por el lema 6.5. $0\in C(r(x_0)\hat{x}_0,\dots,r(x_n)\hat{x}_n)$ si y solo si, los números $r(x_i)$ alternan en signo. c.s.q.d.

Un teorema de sencilla demostración es el siguiente:

6.7. TEOREMA: (de Vallee-Poussin)

Sea $\{g_1,\dots,g_n\}$ un sistema de Haar. Sea $P\in L=\langle\{g_1,\dots,g_n\}\rangle$ tal que $f-P$ toma alternativamente valores positivos y negativos en $n+1$ puntos $x_0<x_1<\dots<x_n$ de $I=[a,b]$. Entonces

$$e_L(f)=\inf_{g\in L}\|f-g\|\geq \min_{0\leq i\leq n}|f(x_i)-P(x_i)|$$

Demostr.: Si la tesis fuese falsa existiría $P_0\in L$ tal que $\|f-P_0\|<\min_i|f(x_i)-P(x_i)|$
Consideremos el elemento de L, P_0-P. Este elemento se puede poner en la forma

$$P_0-P=(f-P)-(f-P_0)$$

Veamos que P_0-P cambia de signo $n+1$ veces.
Puesto que $\|f-P_0\|<\min_{0\leq i\leq n}|f(x_i)-P(x_i)|$ y $\|f-P_0\|=\max_{x\in I}|f(x)-P_0(x)|$ se deduce que

$$|f(x_i)-P_0(x_i)|<|f(x_i)-P(x_i)|,\quad i=0,1,\dots,n.$$

Como $f-P$ cambia de signo alternativamente $n+1$ veces en los puntos $x_0,\dots,x_n$ y en esos puntos el módulo de $f-P_0$ es menor que el de $f-P$ se deduce que el signo de P_0-P en estos puntos es el de $f-P$ y, por tanto, P_0-P cambia de signo $n+1$ veces en $x_0,\dots,x_n$. Entonces, P_0-P se anula por lo menos en n puntos de I.

Será entonces $P_0-P=0$, pues siendo L subespacio de Haar de dimensión n toda función no nula se anula a lo sumo en $n-1$ puntos.
Por tanto, se tendría $P_0=P$, lo cual contradice que $\|f-P_0\|<\min|f(x_i)-P(x_i)|$
Debe ser entonces $e_L(f)\geq\min_i|f(x_i)-P(x_i)|$. c.s.q.d.

OBSERVACION: Supongamos que $f-P$ se anula en $n+1$ puntos y va tomando valores es que la menor de estas alturas $(f-P)(x_j)$ es menor o igual que el error que se va a cometer en este subespacio de Haar.

El siguiente teorema dice que subespacios de dimensión finita de $C(I)$ son semi-Chebyshev, y, por tanto, Chebyshev.

6.8. TEOREMA: Existe una única aproximación óptima uniforme relativa al subespacio $L = \langle\{g_1, \dots, g_n\}\rangle$, donde $g_1, \dots, g_n$ son linealmente independientes, para cada $f \in C(I)$ si, y solo si, $\{g_1, \dots, g_n\}$ es un sistema de Haar.

Demostr.: $\Leftarrow$ Sea $f \in C(I)$ y $\{g_1, \dots, g_n\}$ un sistema de Haar. Siendo $L = \langle\{g_1, \dots, g_n\}\rangle$ de dimensión finita es proximinal (Tema 7º), es decir, $P_L(f) \neq \emptyset$, $\forall f \in C(I)$. Probemos que la aproximación óptima es única.

Supongamos que p y q son aproximaciones óptimas uniformes distintas de f relativas a L, es decir

$$\|f-p\| = \|f-q\| = e_L(f).$$

Siendo $P_L(f)$ convexo $\frac{1}{2}(p+q) \in P_L(f)$. Por el teorema 6.6 existen $x_0 < x_1 < \dots < x_n$ puntos de I tales que

$$f(x_i) - \frac{1}{2}(p+q)(x_i) = (-1)^i \varepsilon, \quad i = 0, 1, \dots, n$$

donde $|\varepsilon| = \|f - \frac{1}{2}(p+q)\| = \|f-p\| = \|f-q\|$

Luego

$$\frac{1}{2}(f(x_i) - p(x_i)) + \frac{1}{2}(f(x_i) - q(x_i)) = (-1)^i \varepsilon, \quad i = 0, 1, \dots, n.$$

Pero, por tratarse de la norma del máximo

$$|f(x_i) - p(x_i)| \leq \|f-p\| = |\varepsilon|$$

y $|f(x_i) - q(x_i)| \leq \|f-q\| = |\varepsilon|$

Por consiguiente

$$|\varepsilon| = |(-1)^i \varepsilon| = \left|\frac{1}{2}(f(x_i) - p(x_i)) + \frac{1}{2}(f(x_i) - q(x_i))\right| \leq \frac{1}{2}|f(x_i) - p(x_i)| + \frac{1}{2}|f(x_i) - q(x_i)| \leq$$
$$\leq \frac{1}{2}|\varepsilon| + \frac{1}{2}|\varepsilon| = |\varepsilon|$$

Luego $|f(x_i) - p(x_i)| = |f(x_i) - q(x_i)| = |\varepsilon|$

y por tanto $f(x_i) - p(x_i) = f(x_i) - q(x_i) = (-1)^i \varepsilon$, $i = 0, 1, \dots, n$.

Entonces $p(x_i) = q(x_i)$, $i = 0, 1, \dots, n$

o también $(p-q)(x_i) = 0$, $i = 0, 1, \dots, n$

Siendo $\{g_1, \dots, g_n\}$ un sistema de Haar y anulándose $p-q$ en $n+1$ puntos distintos de I necesariamente será $p-q=0$, es decir, $p=q$ lo cual contradice la hipótesis de que p y q eran distintos.

Por tanto, la aproximación óptima uniforme es única.

$\Rightarrow$ Supongamos que la aproximación óptima uniforme de todo $f \in C(I)$

relativa a L es única. Supongamos que $\{g_1,\dots,g_n\}$ no es un sistema de Haar. Existen, entonces, $x_1,\dots,x_n \in I$ tales que $D(x_1,\dots,x_n)=0$, es decir

$$\begin{vmatrix} g_1(x_1) & \cdots & g_n(x_1) \\ - & - & - \\ g_1(x_n) & \cdots & g_n(x_n) \end{vmatrix} = 0$$

Por tanto, las filas y columnas de la matriz $(g_i(x_j))_{i,j=1,\dots,n}$ son linealmente dependientes. Existirán entonces $(a_1,\dots,a_n)$ y $(b_1,\dots,b_n)$ tales que

$$\sum_{i=1}^n a_i g_i(x_j) = 0 \ , \quad j=1,\dots,n$$

$$\sum_{j=1}^n b_j g_i(x_j) = 0 \ , \quad i=1,\dots,n \qquad (1)$$

Sea $q = \sum_{i=1}^n a_i g_i \in L = \langle\{g_1,\dots,g_n\}\rangle$.

Entonces $q(x_j)=0$, $j=1,\dots,n$.

Podemos suponer que $\|q\|<1$, pues en caso contrario, dividiendo los a_i por una constante conveniente podemos conseguir que $\|q\|<1$ y se sigue verificando que $q\in L$ y $q(x_j)=0$, $j=1,\dots,n$.

Sea $f\in C(I)$ tal que $\|f\|=1$ y $f(x_j)=1\cdot \operatorname{sig} b_j$, $j=1,\dots,n$. (*)

Consideremos la función

$$F(x) = f(x)\,[1-|q(x)|]$$

Esta función F verifica que

$$F(x_j) = f(x_j) = 1\cdot \operatorname{sig} b_j \ , \quad j=1,\dots,n.$$

Además, probemos que si $P\in L$ entonces $\|F-P\|\geq 1$.

En efecto, si fuese $\|F-P\|<1$ debería ocurrir que $\operatorname{sig} P(x_j) = \operatorname{sig} F(x_j) = \operatorname{sig} b_j$ ya que si $\operatorname{sig} P(x_j) \neq \operatorname{sig} F(x_j)$ entonces

$$\|F-P\| \geq |F(x_j)-P(x_j)| > |F(x_j)| = |f(x_j)| = 1 \ \text{ y sería } \|F-P\|\geq 1.$$

Luego $\|F-P\|<1 \Rightarrow \operatorname{sig} P(x_j) = \operatorname{sig} F(x_j)$, $j=1,\dots,n$

Por tanto, como $\operatorname{sig} P(x_j) = \operatorname{sig} b_j$ será $b_j P(x_j) \geq 0$, y mayor estrictamente para algún j, pues $P(x_j)\neq 0$ y $(b_1,\dots,b_n)\neq(0,\dots,0)$.

Por tanto, $\sum_{j=1}^n b_j P(x_j) > 0$. (2)

Por otra parte, como $P\in L$, existen $c_1,\dots,c_n$ tales que $P = \sum_{i=1}^n c_i g_i$.

Entonces $\sum_{j=1}^n b_j P(x_j) = \sum_{j=1}^n b_j \left(\sum_{i=1}^n c_i g_i(x_j)\right) = \sum_{i=1}^n c_i \left(\sum_{j=1}^n b_j g_i(x_j)\right) = 0$, en virtud de (1). Hemos llegado así a una contradicción con (2).

Debe ser entonces $\|F-P\|\geq 1$, $\forall P\in L$. (3)

Veamos ahora que $\forall \lambda \in [0,1]$, λq es aproximación óptima uniforme de F relativa a L, con lo cual la aproximación óptima no será única.

$\forall x \in I$, $|F(x) - \lambda q(x)| \leq |F(x)| + \lambda |q(x)| = |f(x)| \cdot |1 - |q(x)|| + \lambda |q(x)| \leq$
$\leq |1 - |q(x)|| + \lambda |q(x)| = 1 - |q(x)| + \lambda |q(x)|$
pues $|q(x)| < 1$, $|f(x)| \leq 1$.
Como $\lambda \in [0,1]$, $1 - |q(x)| + \lambda |q(x)| = 1 + (\lambda - 1)|q(x)| < 1 + \lambda - 1 = \lambda \leq 1$
Luego $\forall x \in I$, $|F(x) - \lambda q(x)| \leq 1$. (4)
y por tanto $\|F - \lambda q\| \leq 1$.
Por otra parte $\|F - P\| \geq 1$, $\forall P \in L$. Como $\lambda q \in L$, se tendrá que $\|F - \lambda q\| \geq 1$, y teniendo en cuenta (4) sera $\|F - \lambda q\| = 1$, $\forall \lambda \in [0,1]$.
Trivialmente $e_L(F) = 1$, pues $\|F - P\| \geq 1$, $\forall P \in L$ y $\|F - \lambda q\| = 1$, si $\lambda \in [0,1]$.
Luego, $\forall \lambda \in [0,1]$, λq es aproximación óptima uniforme de F relativa a L, en contra de lo supuesto.
Por tanto, $\{q_1, \dots, q_n\}$ es un sistema de Haar. c.s.q.d.