

Capítulo 5

Teoría de Aproximación.

5.1 Introducción. Mejor aproximación

En este capítulo pretendemos generalizar lo estudiado en el capítulo anterior ya que la interpolación es una forma especial de aproximación.

Planteamiento del problema general de aproximación de elementos de un conjunto, E , mediante elementos de un subconjunto, $H \subset E$.

Dado $f \in E$, hallar $p \in H$ tal que *aproxime* a f lo mejor posible.

En este momento, hay varias consideraciones a tener presentes:

- ¿qué se entiende por *aproxime lo mejor posible*?
- ¿qué espacios se pueden usar?
- ¿admite solución el problema?
- ¿hay unicidad? y, en caso afirmativo, ¿cómo se obtiene?

Para dar respuesta a todas estas cuestiones hemos de introducir algunos conceptos: norma, distancia, métrica, etc ; en definitiva necesitamos una herramienta que permita medir la proximidad entre dos elementos del espacio E .

Nosotros vamos a considerar métricas asociadas a normas definidas sobre el espacio de trabajo E . Más concretamente,

Definición 5.1 Una norma sobre el conjunto E es una aplicación, $\|\bullet\| : E \rightarrow \mathbb{R}$ que cumple:

1. $\|x\| > 0, \forall x \in E \quad x \neq 0$
2. $\|\lambda x\| = \lambda \|x\|, \forall x \in E \quad \forall \lambda \in \mathbb{R}$
3. $\|x + y\| \leq \|x\| + \|y\|, \forall x, y \in E$

El par $(E, \|\bullet\|)$ se llama espacio normado.

Se dice que la aplicación es una seminorma, si **1** se sustituye por : $\|x\| \geq 0, \forall x \in E$.

5.1.1 Problema general de aproximación en espacios normados.

En un espacio vectorial normado podemos entender el concepto de proximidad; es decir, podemos medir la separación entre dos elementos dados. Así, ésta se mide por:

$$\text{dist}(f, g) = \|f - g\|$$

Por lo tanto, el problema general de aproximación es:

Dado $(\mathbb{E}, \|\bullet\|)$ espacio vectorial normado y $H \subset \mathbb{E}$ un subconjunto;

Para $f \in \mathbb{E}$, hallar $u \in H$ verificando:

$$\|f - u\| = \text{dist}(f, H) = \inf_{h \in H} \{\|f - h\|\}$$

Según sea el subconjunto H , se pueden presentar los caso siguientes:

- Si $\text{dist}(f, H) = 0$, entonces existe una sucesión de elementos de H que converge al elemento f ; es decir, nos podemos aproximar a f tanto como queramos.
- Si $\text{dist}(f, H) = d > 0$, entonces el problema planteado se conoce como problema de mejor aproximación (**m.a.**)

Si existe un elemento $u \in H$ con $\text{dist}(f, u) = d$ se llama **mejor aproximación de f en H**

Ejemplos

1. Sea $\mathbb{E} = C([0, 1])$ y $\|f\| = \max_{0 \leq x \leq 1} \{|f(x)|\}$ consideramos el subespacio, $H = \text{gen}\{x\}$; entonces, los elementos $h = \lambda x$ con $0 \leq \lambda \leq 2$ son m.a. de $f(x) = 1$.
2. Sea $\mathbb{E} = \mathbb{R}$ con $\|x\| = |x|$ y sea $H =]0, 1[$, entonces no existe m.a. para todo $x > 1$ ó $x < 0$

Ahora, enumeramos los resultados básicos de existencia y unicidad de la m.a.

Teorema 5.1 (EXISTENCIA) *Sea C un compacto no vacío de \mathbb{E} (e.v.n.). Entonces, el problema de m.a. en $H = C$ admite solución.*

Definición 5.2 *Se dice que C es convexo si, y sólo si,*

$$\forall x, y \in C \Rightarrow \lambda x + (1 - \lambda)y \in C \quad \forall \lambda \in [0, 1]$$

Teorema 5.2 (Espacios finito-dimensionales) *Si H es un subespacio vectorial de \mathbb{E} con $\dim(H) = \text{finita}$, entonces:*

- $\forall f \in \mathbb{E}$ existe **m.a.** de f en H
- Si $H_f = \{h \in H : \text{es m.a. de } f\}$ es un subconjunto convexo de H .

Ahora que sabemos cuando hay solución es de interés saber si ésta es única. Para ello, introducimos el concepto de espacio estrictamente convexo; es decir,

Definición 5.3 *Se dice que \mathbb{E} es un espacio normado estrictamente convexo si, y sólo si,*

$$\forall x, y \in \mathbb{E} \quad (\text{distintos}) \text{ tales que } \|x\| = \|y\| = 1 \Rightarrow \left\| \frac{x + y}{2} \right\| < 1$$

Teorema 5.3 (UNICIDAD) Si \mathbb{E} es un espacio normado estrictamente convexo y H subespacio de dimensión finita; entonces para cada $f \in \mathbb{E}$ existe una única m.a. en H

¿Qué espacios son estrictamente convexos?

Definición 5.4 Un e.v. real \mathbb{E} es un espacio **prehilbertiano** si hay definido un producto escalar o interno; es decir, una aplicación, $\langle \bullet, \bullet \rangle : \mathbb{E} \rightarrow \mathbb{R}$ que cumple:

1. $\langle x, x \rangle > 0, \forall x \in \mathbb{E} \quad x \neq 0$
2. $\langle x, y \rangle = \langle y, x \rangle, \forall x, y \in \mathbb{E}$
3. $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle, \forall x, y, z \in \mathbb{E} \quad \forall \lambda, \mu \in \mathbb{R}$

Desde este concepto es fácil entender que en todo espacio preHilbert se tiene definida la norma natural siguiente:

$$\|x\| = \sqrt{\langle x, x \rangle}$$

Además, se verifica la desigualdad de Cauchy-Schwarz:

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

(la desigualdad es estricta si, y sólo si, f y g son linealmente independientes)

Así, se tienen los resultados de interés siguientes:

Teorema 5.4 Si \mathbb{E} es un espacio pre-hilbertiano, entonces es estrictamente convexo.

Demostración

Sean $f, g \in \mathbb{E}$, distintos y unitarios; entonces,

$$\left\| \frac{f+g}{2} \right\|^2 = \left\langle \frac{f+g}{2}, \frac{f+g}{2} \right\rangle = \frac{1}{4} \left(\|f\|^2 + \|g\|^2 + 2 \langle f, g \rangle \right)$$

Puesto que, se cumple la desigualdad de Cauchy-Schwarz, se tendrá:

- Para el caso de independencia entre f y g :

$$\left\| \frac{f+g}{2} \right\|^2 = \frac{1}{4} \left(\|f\|^2 + \|g\|^2 + 2 \langle f, g \rangle \right) < \frac{1}{4} \left(\|f\|^2 + \|g\|^2 + 2\|f\|\|g\| \right) = 1$$

- Para el caso de dependencia sólo se puede dar cuando $f = -g$; entonces $f + g = 0 \Rightarrow \left\| \frac{f+g}{2} \right\| = \|0\| = 0 < 1$

Por lo tanto, siempre se tiene: $\left\| \frac{f+g}{2} \right\| < 1$; es decir, el espacio es estrictamente convexo.

Corolario 5.1 Si \mathbb{E} es un espacio prehilbertiano y H un subespacio de dimensión finita; entonces, el problema de **m.a.** tiene solución única.

5.2 Aproximación en espacios vectoriales dotados de un producto escalar

En esta sección nos centramos en el análisis del problema de m.a. en un espacio con producto escalar o interno.

5.2.1 Caracterización de la mejor aproximación.

Sea \mathbb{E} un espacio vectorial real pre-Hilbert con producto escalar, $\langle \bullet, \bullet \rangle$ y sea H un subespacio con $\dim(H) = n$; el problema de m.a. puede expresarse como:

Dado $f \in \mathbb{E}$, hallar $u \in H$ verificando:

$$\|f - u\|^2 = \langle f - u, f - u \rangle = \min_{h \in H} \|f - h\|^2 \quad (5.1)$$

A veces, el problema (5.1) se conoce como problema de **mínimos cuadrados** (aunque esta denominación se reserva cuando trabajamos con espacios de funciones).

Pues bien, el resultado que permite acceder a una forma de construir la **m.a.** de f en H es el siguiente:

Teorema 5.5 (Caracterización) $u \in H$ es la m.a. de $f \in \mathbb{E}$ si, y sólo si,

$$\langle f - u, h \rangle = 0 \quad \forall h \in H \quad (5.2)$$

5.2.2 Construcción de la m.a.

Es sencillo darse cuenta que la condición (5.2) puede sustituirse por:

$$\langle f - u, v_j \rangle = 0 \quad j = 1, \dots, n \quad (5.3)$$

donde $\{v_1, v_2, \dots, v_n\}$ es una base de H .

Ahora, usando (5.3), si suponemos que $u = a_1 v_1 + a_2 v_2 + \dots + a_n v_n$, entonces calcular la **m.a.** de f en H equivale a resolver el sistema $n \times n$:

$$\sum_{i=1}^n \langle v_j, v_i \rangle a_i = \langle f, v_j \rangle \quad j = 1, \dots, n$$

que en formato matricial es:

$$\begin{pmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \cdots & \langle v_1, v_n \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \cdots & \langle v_2, v_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \cdots & \langle v_n, v_n \rangle \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \langle f, v_1 \rangle \\ \langle f, v_2 \rangle \\ \vdots \\ \langle f, v_n \rangle \end{pmatrix} \quad (5.4)$$

La matriz de coeficientes del sistema (5.4) se conoce como matriz de Gram y la notamos por $G(v_1, v_2, \dots, v_n)$. Además, puede probarse directamente que $\det(G) \neq 0$ para todo conjunto de vectores l.i. y todo producto escalar.

Mejor aproximación en el trasladado de un subespacio.

Supongamos que $C = \{\phi + h \in \mathbb{E} : h \in H\} = \phi + H$ el trasladado mediante ϕ del subespacio vectorial H . ¿Cómo se puede calcular la m.a. de $f \in \mathbb{E}$ en C ? La respuesta la damos en resultado siguiente:

Teorema 5.6 Sea \mathbb{E} un espacio pre-Hilbert y H un subespacio vectorial de dimensión finita; entonces, son equivalentes:

1. u^* es la m.a. de $f \in \mathbb{E}$ en $C = \phi + H$
2. $u = u^* - \phi$ es la m.a. de $f - \phi$ en H

Por lo tanto, desde este resultado, la m.a. en $C = \phi + H$ se puede obtener calculando la m.a. (u) de $f - \phi$ en H y tomar como m.a. $u^* = u + \phi$

Ejemplo 1:

Hallar la m.a. de $f(x) = 1$ $x \in [0, 1]$ en $C = \{g(x) = x^2 + \lambda x : \lambda \in \mathbb{R}\} \subset C([0, 1])$ usando el producto escalar: $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$

En efecto, el conjunto de funciones consideradas se puede escribir como: $C = x^2 + H$ donde H es el un subespacio de $\mathbb{E} = C([0, 1])$ de dimensión 1 con base: $\{v_1(x) = x\}$.

En primer lugar calculamos la **m.a.** de la función $g(x) = f(x) - x^2 = 1 - x^2$ en H .

En este caso el sistema es 1×1 para $u(x) = a_1 v_1(x) = a_1 x$; a saber,

$$a_1 \langle x, x \rangle = \langle 1 - x^2, x \rangle \Leftrightarrow a_1 = \frac{\langle 1 - x^2, x \rangle}{\langle x, x \rangle} \Leftrightarrow \frac{\int_0^1 (1 - x^2) x dx}{\int_0^1 x^2 dx} = \frac{1/4}{1/3} = \frac{3}{4}$$

Así, la **m.a.** de $1 - x^2$ en H es $u(x) = \frac{3}{4}x$ y, en consecuencia, la **m.a.** de $f(x) = 1$ en C es: $v(x) = x^2 + \frac{3}{4}x$. Además, la distancia mínima es:

$$\text{dist} \left(1, x^2 + \frac{3}{4}x \right) = \frac{1}{4} \sqrt{\frac{83}{15}}$$

Ejemplo 2: distancia de un punto de \mathbb{R}^3 a un plano Π

Vamos a calcular la distancia del punto $P = (1, -2, 1) \in \mathbb{R}^3$ al plano $\Pi := 2x + y - z = 3$

En primer lugar, hemos de interpretar este problema como un problema de m.a. en un espacio vectorial (\mathbb{R}^3) con producto escalar (en este caso el euclideo) y $H =$ plano vectorial asociado a Π ; es decir,

$$H = \{(x, y, z) \in \mathbb{R}^3 : 2x + y - z = 0\} = \text{gen}\{v_1 = (0, 1, 1), v_2 = (1, 0, 2)\}$$

¿Quién es el elemento f a aproximar con elementos de H ? Basta considerar un punto cualquiera del plano Π (p.e. $Q = (1, 0, -1)$) y formamos el vector $f = \overrightarrow{QP} = (0, -2, 2)$. Por lo tanto, el problema original se plantearía como sigue:

Hallar $u = a_1 v_1 + a_2 v_2 = a_1(0, 1, 1) + a_2(1, 0, 2)$ verificando:

$$\|f - u\|^2 = \min_{a_1, a_2} \left\{ \|f - a_1 v_1 - a_2 v_2\|^2 \right\} \equiv (\text{dist}(f, H))^2 = (\text{dist}(P, \Pi))^2$$

Como estamos usando la norma euclídea, que proviene del producto escalar euclídeo, podemos aplicar lo visto en la caracterización de m.a.; a saber,

- Matriz de Gram: $\begin{pmatrix} v_1 \bullet v_1 & v_1 \bullet v_2 \\ v_2 \bullet v_1 & v_2 \bullet v_2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}$
- Términos independientes: $\begin{pmatrix} f \bullet v_1 \\ f \bullet v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$
- El sistema a resolver es: $\begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$
cuya solución es: $a_1 = -\frac{4}{3}$, $a_2 = \frac{4}{3}$

Por lo tanto, la m.a. de f en H es: $u = \frac{4}{3}(-v_1 + v_2) = \frac{4}{3}(1, -1, 1)$ y la distancia pedida será:

$$\begin{aligned} \text{dist}((1, -2, 1), \Pi) &= \|f - u\| = \sqrt{(f - u) \bullet (f - u)} = \\ &= \sqrt{\frac{2}{3}(-2, -1, 1) \bullet \frac{2}{3}(-2, -1, 1)} = \frac{2}{3}\sqrt{6} \end{aligned}$$

5.2.3 Bases especiales.

A veces, el sistema (5.4) puede simplificarse si la base utilizada cumple ciertas propiedades respecto del producto escalar usado.

Más concretamente, si conocemos una base de H , $\{u_1, u_2, \dots, u_n\}$, ortogonal; es decir, para $i \neq j$ $\langle u_i, u_j \rangle = 0$; entonces la m.a. de f en H es:

$$u = \frac{\langle f, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \frac{\langle f, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 + \dots + \frac{\langle f, u_n \rangle}{\langle u_n, u_n \rangle} u_n$$

Además, si los vectores de la base son unitarios ($\|u_i\| = 1$); es decir, la base es ortonormal, entonces:

$$u = \langle f, u_1 \rangle u_1 + \langle f, u_2 \rangle u_2 + \dots + \langle f, u_n \rangle u_n$$

A los coeficientes de la m.a. de f en H respecto de una base ortonormal se conocen como coeficientes de Fourier de f y u es la suma de Fourier de f respecto de la base.

En esta situación, es interesante verificar la igualdad siguiente:

$$\|f - u\|^2 = \|f\|^2 - \sum_{i=1}^n (\langle f, u_i \rangle)^2$$

Ahora bien, ¿cómo podemos conocer este tipo de bases? La respuesta es sencilla, mediante un proceso que permite generar una base ortogonal a partir de una base cualquiera del espacio de trabajo. Dicho proceso es conocido como: **ortonormalización de Gram-Schmidt**.

ALGORITMO.

Sea $\{v_1, v_2, \dots, v_n\}$ una base de H ; entonces,

- Tomamos $u_1 = v_1$
- Para $j = 2, \dots, n$
 - Para $k = 1, \dots, j - 1$ calcular $\lambda_k = \frac{\langle v_j, u_k \rangle}{\langle u_k, u_k \rangle}$
 - Tomamos $u_j = v_j - \lambda_1 u_1 - \dots - \lambda_{j-1} u_{j-1}$

Entonces, $\{u_1, u_2, \dots, u_n\}$ es una base ortogonal de H . Desde esta base se puede conseguir una base ortonormal sin más que dividir cada vector por su norma; es decir,

$$\left\{ \frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}, \dots, \frac{u_n}{\|u_n\|} \right\}$$

es base ortonormal de H .

5.3 Aproximación por mínimos cuadrados: continua y discreta.

En esta sección nos ocupamos de dos problemas particulares de **m.a.** con funciones continuas y con funciones evaluadas en un número finito de nodos. Son los problemas **m.a. mínimos cuadrados continua** y **discreta** respectivamente.

5.3.1 Problema 1: mejor aproximación mínimos cuadrados continua.

Consideramos el espacio de funciones continuas en $[a, b]$, $\mathbb{E} = C([a, b])$ con el producto escalar:

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx \tag{5.5}$$

cuya norma asociada es: $\|f\| = \sqrt{\int_a^b f(x)^2 dx}$.

Así, el problema de mejor aproximación mínimos cuadrados continua se describe como sigue:

Dado $f \in \mathbb{E}$, hallar $u \in H$ verificando:

$$\|f - u\|^2 = \int_a^b (f(x) - u(x))^2 dx = \min_{h \in H} \left\{ \int_a^b (f(x) - h(x))^2 dx \right\}$$

Para resolver este problema basta utilizar las caracterizaciones descritas para espacios pre-hilbertianos. Por lo tanto, teniendo en cuenta el producto escalar particular, si $\{v_1(x), \dots, v_n(x)\}$ es una base de H , entonces:

$$u(x) = a_1 v_1(x) + a_2 v_2(x) + \dots + a_n v_n(x)$$

es la m.a.m.c. continua de $f(x)$ en H si, y sólo si, los coeficientes son la solución del sistema:

$$\begin{pmatrix} \int_a^b v_1(x)^2 dx & \int_a^b v_1(x)v_2(x) dx & \dots & \int_a^b v_1(x)v_n(x) dx \\ \int_a^b v_2(x)v_1(x) dx & \int_a^b v_2(x)^2 dx & \dots & \int_a^b v_2(x)v_n(x) dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_a^b v_n(x)v_1(x) dx & \int_a^b v_n(x)v_2(x) dx & \dots & \int_a^b v_n(x)^2 dx \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \int_a^b v_1(x)f(x) dx \\ \int_a^b v_2(x)f(x) dx \\ \vdots \\ \int_a^b v_n(x)f(x) dx \end{pmatrix}$$

Un caso particular es el de **m.a.m.c.c.** con polinomios de grado $\leq n$. En esta situación si se considera la base de \mathbb{P}_n , $\{1, x, \dots, x^n\}$, entonces la m.a. de $f(x)$ en \mathbb{P}_n es $p(x) = a_0 + a_1x + \dots + a_nx^n$ con a_i $i = 0, \dots, n$ solución del sistema:

$$\left(\int_a^b x^j dx\right) a_0 + \left(\int_0^1 x^{j+1} dx\right) a_1 + \dots + \left(\int_0^1 x^{n+j} dx\right) a_n = \int_0^1 x^j f(x) dx \quad j = 0, 1, \dots, n$$

Este sistema presenta diversos inconvenientes entre los que se encuentra que pequeños errores de redondeo pueden producir grandes desviaciones en la solución (debido al mal condicionamiento)

Así, es de sumo interés poder contar con bases de polinomios ortogonales u ortonormales que analizamos en la siguiente sección.

Si en lugar de utilizar el producto (5.5) usamos:

$$\langle f, g \rangle = \int_a^b w(x) f(x) g(x) dx \quad (5.6)$$

donde $w(x)$ (llamada función peso) es positiva salvo en un número finito de puntos en $]a, b[$, y existe $\int_a^b w(x) f(x) dx \quad \forall f$.

el problema se conoce como **m.am.m.c. continua ponderada** cuya solución se obtiene de forma idéntica al caso anterior pero introduciendo en cada integración el factor o peso $w(x)$.

5.3.2 Problema 2: mejor aproximación mínimos cuadrados discreta.

Consideramos el espacio vectorial, $\mathbb{E} = \mathbb{R}^m$ con el producto escalar euclideo:

$$\langle f, g \rangle = \sum_{i=1}^m f_i g_i$$

cuya norma asociada es:

$$\|f\| = \sqrt{\sum_{i=1}^m f_i^2}$$

Así, el problema de mejor aproximación mínimos cuadrados discreta (**m.a.m.c.d.**) se describe como sigue:

Dado $f \in \mathbb{E}$, y $H \subset \mathbb{E}$ con $\dim H = n < m$, hallar $u \in H$ verificando:

$$\|f - u\|^2 = \sqrt{\sum_{i=1}^m (f_i - u_i)^2} = \min_{h \in H} \left\{ \sqrt{\sum_{i=1}^m (f_i - h_i)^2} \right\}$$

Un caso particular de esta situación es la siguiente:

Dados los valores $\{y_1, y_2, \dots, y_m\}$ en los nodos $\{x_1, x_2, \dots, x_m\}$ (p.e. $y_i = f(x_i) \quad i = 1, \dots, m$),

Hallar $u \in H$ subespacio de funciones (p.e. polinómicas) de $\dim H = n \ll m$ verificando:

$$\sum_{i=1}^m (y_i - u(x_i))^2 = \min$$

Este problema es conocido como ajuste de datos por mínimos cuadrados y es muy usado en las Ciencias Experimentales.

Si consideramos el caso de $H = \mathbb{P}_n$ con $n \leq m - 1$, entonces calcular la **m.a.m.c. discreta** de $y = (y_1, y_2, \dots, y_m)$ en \mathbb{P}_n equivale a obtener el polinomio $u = p(x) = a_0 + a_1x + \dots + a_nx^n$ cuyos coeficientes son la solución del sistema lineal:

$$M^t \cdot M \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = M^t \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix} \quad (5.7)$$

donde M es la matriz asociada a una base de \mathbb{P}_n , p.e. $\{1, x, x^2, \dots, x^n\}$, evaluada en los nodos $\{x_1, x_2, \dots, x_m\}$. Más concretamente,

$$M = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{pmatrix}$$

Observaciones.

1. La matriz $M^t M$ es la matriz de Gram asociada al problema de m.a. en este caso concreto.
2. El sistema lineal admite solución única si, y sólo si, $rg(M) = n + 1$
3. $rg(M) = n + 1$ si hay, al menos, $n + 1$ nodos distintos en el conjunto $\{x_1, x_2, \dots, x_m\}$.
4. Si $m=n+1$, entonces el problema de m.a. es el de interpolación polinomial.
5. En la práctica se usan valores pequeños de n (p.e. $n = 1, 2$, ó 3) dando lugar a los llamados: ajuste lineal, cuadrático ó cúbico respectivamente.

5.4 Polinomios ortogonales

Como ya hemos comentado en la sección anterior el uso de polinomios para aproximar funciones continuas en cierto intervalo es muy frecuente. También es conveniente el uso de bases ortogonales u ortonormales para facilitar la obtención de la mejor aproximación. Ahora bien, en el caso de los polinomios la obtención de tales bases se reduce de manera significativa como veremos a continuación.

Así, en lo que sigue, suponemos que el producto escalar usado es del tipo (5.5) o (5.6) y obtendremos una sucesión de polinomios ortogonales que permitirán resolver el problema de un modo más simple.

Definición 5.5 Diremos que $\{p_0, p_1, \dots, p_n, \dots\}$ es una sucesión de polinomios ortogonales respecto de los productos (5.5) o (5.6) si:

- son no nulos;
- $grado(p_j) = j \quad j = 0, 1, \dots$
- $\langle p_i, p_j \rangle = 0 \quad \forall i \neq j$

¿Cómo calculamos tales polinomios? El siguiente resultado nos proporciona como hacerlo además de mostrar algunas propiedades de gran interés.

Teorema 5.7 Sea $\{p_0, p_1, \dots, p_n, \dots\}$ una sucesión de polinomios ortogonales respecto de (5.5) o (5.6); entonces:

1. Si p_n tiene coeficiente líder 1 (es mónico), entonces, $\|p_n\| \leq \|p(x)\|$ para cualquier otro polinomio mónico $p(x)$ de grado n
2. La sucesión satisface una relación de recurrencia a tres términos del tipo:

$$p_{n+1} = A_n(x - B_n)p_n - A_n C_n p_{n-1} \quad (5.8)$$

3. Si tomamos los polinomios mónicos todos ellos, entonces (5.8) queda como:

$$p_{n+1} = (x - B_n)p_n - C_n p_{n-1} \quad (5.9)$$

Además, los valores de A_n, B_n , y C_n son:

$$A_n = \frac{a_{n+1}}{a_n}, \quad B_n = \frac{\langle xp_n, p_n \rangle}{\langle p_n, p_n \rangle}, \quad C_n = \frac{\langle p_n, p_n \rangle}{A_{n-1} \langle p_{n-1}, p_{n-1} \rangle}$$

y a_n es el coeficiente principal de p_n .

Polinomios de Legendre

Son los polinomios ortogonales en el $[-1, 1]$ y peso $w = 1$. Cumplen la recurrencia:

$$p_{n+1} = \frac{2n+1}{n+1}xp_n - \frac{n}{n+1}p_{n-1}$$

Polinomios de Chebyshev de 1ª especie.

Son los polinomios ortogonales en el $[-1, 1]$ y peso $w = \frac{1}{\sqrt{1-x^2}}$. Cumplen la recurrencia:

$$p_{n+1} = 2xp_n - p_{n-1}$$

Polinomios de Laguerre

Son los polinomios ortogonales en el $]0, \infty[$ y peso $w = e^{-x}$.

Polinomios de Hermite

Son los polinomios ortogonales en el $] - \infty, \infty[$ y peso $w = e^{-x^2}$.

5.5 Aproximación uniforme. Teorema de Weierstrass. Polinomios de Bernstein. Curvas de Bézier.

En esta sección tratamos de resolver el problema de m.a. de una función continua en $[a, b]$ usando polinomios de grado fijo y con la norma uniforme o del máximo; es decir:

$$\left. \begin{array}{l} \text{Hallar } p(x) \in \mathbb{P}_n \text{ verificando:} \\ \|f - p\| = \max_{x \in [a, b]} \{|f(x) - p(x)|\} = \min_{q \in \mathbb{P}_n} \|f - q\| \end{array} \right\} \quad (5.10)$$

Ahora bien, es conocido (Teorema de Weierstrass) que dada una función continua en un intervalo $[a, b]$ podemos aproximarla tanto como queramos mediante polinomios de grado apropiado. En este resultado tiene especial relevancia los llamados polinomios de Bernstein que introduciremos más adelante.

5.5.1 Aproximación con polinomios de grado fijo.

Nos ocupamos, ahora, del problema (5.10). En este caso no podemos usar los resultados dados para el caso de mínimos cuadrados ya que el espacio de funciones continuas con la norma uniforme o del máximo no es un espacio estrictamente convexo; por lo tanto, la existencia de solución está asegurada desde lo visto en el teorema 5.2 pero no la unicidad.

En esta situación se hace necesario un tratamiento específico; más concretamente, primero presentamos una caracterización de la m.a., luego probaremos la unicidad y, por último daremos un método de construcción de la m.a. uniforme.

Teorema 5.8 (Caracterización) $p(x) \in \mathbb{P}_n$ es m.a. uniforme de $f(x) \in C([a, b])$ si, y sólo si, $f(x) - p(x)$ alcanza su valor absoluto **máximo** al menos $n+2$ -veces con signos alternos; es decir, existen nodos $a \leq x_0 < x_1 < \dots < x_n < x_{n+1} \leq b$ tales que:

$$f(x_j) - p(x_j) = (-1)^j M \quad j = 0, 1, \dots, n+1 \quad (5.11)$$

donde $M = \pm \|f - p\|$.

Este resultado nos indica que la mejor aproximación uniforme en \mathbb{P}_n es un interpolante de la función en, al menos, $n+1$ nodos del intervalo $[a, b]$. Con este hecho es relativamente sencillo probar la UNICIDAD de la m.a.u.

Teorema 5.9 (UNICIDAD) Dada $f \in C([a, b])$ existe un único polinomio en \mathbb{P}_n verificando (5.11).

Ahora, la cuestión es: ¿cómo construir la mejor aproximación uniforme de una función continua en $[a, b]$? La respuesta, en general, no es tan sencilla como con productos escalares ya que la búsqueda de los nodos con la condición de alternancia (5.11) es bastante difícil salvo situaciones muy particulares.

Así, un método general de búsqueda es debido a Rémèz que propone una discretización del problema y mejora sucesiva de los nodos usados hasta conseguir un polinomio lo suficientemente próximo al buscado. Antes de describir este procedimiento analizamos dos situaciones particulares en las que los nodos pueden conseguirse en forma exacta y, por tanto, la m.a.u.

Mejor aproximación uniforme en \mathbb{P}_0

La m.a.u. de una función continua mediante ctes es muy sencilla ya que se puede probar el resultado siguiente:

Proposición 5.1 Sea $f \in C([a, b])$ y sean $m = \min_{x \in [a, b]} \{f(x)\}$, $M = \max_{x \in [a, b]} \{f(x)\}$; entonces, $p = \frac{m+M}{2}$ es la m.a.u. de f en \mathbb{P}_0 .

Ejemplo.

La mejor aproximación uniforme de $f(x) = (x + 1)^2$ $x \in [-1, 2]$ en \mathbb{P}_0 es $p = 4.5$ como puede verse gráficamente:

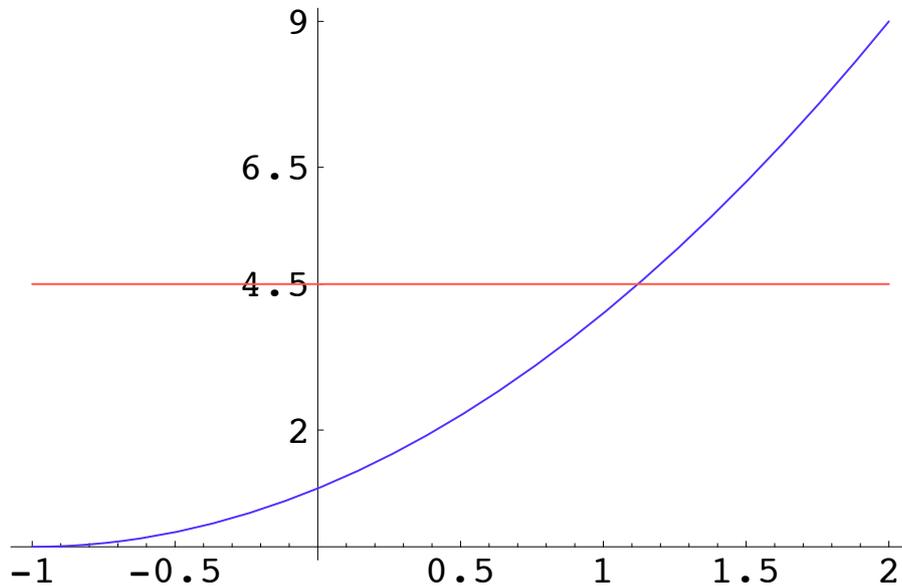


Figura 5.1: m.a.u. de f en \mathbb{P}_0

Mejor aproximación uniforme en \mathbb{P}_1

Una situación particular de m.a.u. es la siguiente. Supongamos que $f(x)$ es una función derivable en $]a, b[$ y de concavidad estricta constante en todo el intervalo (la derivada es estrictamente monótona); entonces, la m.a.u. puede obtenerse sin más que resolver el sistema siguiente:

$$\left. \begin{aligned} f(a) - p(a) &= M \\ f(\xi) - p(\xi) &= -M \\ f(b) - p(b) &= M \\ f'(\xi) - p'(\xi) &= 0 \end{aligned} \right\} \quad (5.12)$$

Esto se deduce del hecho conocido que una función continua en $[a, b]$ alcanza sus extremos en el interior y/o en $x = a, b$. Así, como la función $f - p$ es derivable y hay, al menos tres nodos donde toma valores alternos $M, -M, M$; entonces, existe un punto ξ donde $f'(\xi) - p'(\xi) = 0$ que, además, será único por el crecimiento o decrecimiento estricto de $f'(x)$.

En consecuencia, el valor absoluto máximo de $f - p$ se alcanzará en los nodos $x_0 = a, x_1 = \xi, x_2 = b$ como se indica en (5.12).

Ejemplo.

La mejor aproximación uniforme de $f(x) = (x + 1)^2$, $x \in [-1, 2]$ en \mathbb{P}_1 es $p = a_0 + a_1x$ donde los coeficientes se obtienen resolviendo (5.12) que en este caso es:

$$\left. \begin{aligned} f(-1) - p(-1) &= 0 - a_0 + a_1 = M \\ f(\xi) - p(\xi) &= (\xi + 1)^2 - a_0 - a_1\xi = -M \\ f(2) - p(2) &= 9 - a_0 - 2a_1 = M \\ f'(\xi) - p'(\xi) &= 2(\xi + 1) - a_1 = 0 \end{aligned} \right\} \quad (5.13)$$

cuya solución es: $a_0 = \frac{15}{8}$, $a_1 = 3$, $\xi = \frac{1}{2}$, $M = \frac{9}{8}$; por lo tanto la m.a.u. es: $p(x) = \frac{15}{8} + 3x$.

Ahora la gráfica comparada es:

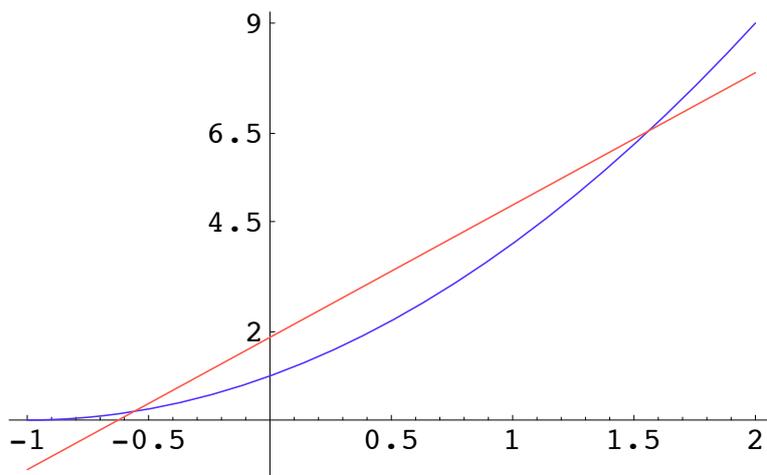


Figura 5.2: m.a.u. de f en \mathbb{P}_1

Aproximación en \mathbb{P}_n : Algoritmo de Rémèz.

En una situación más general la construcción de la m.a.u. de $f \in C([a, b])$ en \mathbb{P}_n no es tan simple y requiere de un procedimiento iterado como se propone con el ALG que explicamos a continuación. La idea está en resolver el problema discretizado siguiente:

$$\text{Hallar } p(x) \in \mathbb{P}_n \text{ verificando :} \quad (5.14)$$

$$\max_{x \in \{z_1, \dots, z_m\}} \{|f(x) - p(x)|\} = \min_{q \in \mathbb{P}_n} \left\{ \max_{x \in \{z_1, \dots, z_m\}} \{|f(x) - q(x)|\} \right\}$$

donde $\{z_1, \dots, z_m\}$ es un conjunto de $m \geq n + 2$ nodos distintos en el intervalo $[a, b]$.

Así, en particular, si tomamos $m = n + 2$; entonces, podemos calcular la solución exacta de (5.14) sin más que resolver el sistema (5.11) con nodos de alternancia $\{z_1, \dots, z_{n+2}\}$.

El polinomio obtenido de esta manera lo llamamos $p_{n,1}(x)$.

Si $\|f(x) - p_{n,1}(x)\| \equiv \max_{x \in \{z_1, \dots, z_m\}} \{|f(x) - p_{n,1}(x)|\} = |r_1|$, entonces $p_{n,1}(x)$ es la m.a.u. de f en todo el intervalo $[a, b]$. Sin embargo, esto no será lo habitual (salvo casos excepcionales), por lo que tendremos que $|f(x^*) - p_{n,1}(x^*)| > |r_1|$ para algún $x^* \in [a, b]$.

Ahora, renovamos los nodos sustituyendo uno de ellos por el nodo x^* según sea el signo de $f(x^*) - p_{n,1}(x^*)$; a saber,

- Si $x^* < z_1 \Rightarrow$
 - sustituimos z_1 por el nodo x^* si $\text{sign}(f(x^*) - p_{n,1}(x^*)) = \text{sign}(f(z_1) - p_{n,1}(z_1))$; o
 - eliminamos z_{n+2} en caso contrario.
- Si $x^* > z_{n+2} \Rightarrow$
 - sustituimos z_{n+2} por el nodo x^* si $\text{sign}(f(x^*) - p_{n,1}(x^*)) = \text{sign}(f(z_{n+2}) - p_{n,1}(z_{n+2}))$; o
 - eliminamos z_1 en caso contrario.

- Si $z_i < x^* < z_{i+1} \Rightarrow$
 - Si $\text{sign}(f(x^*) - p_{n,1}(x^*)) = \text{sign}(f(z_i) - p_{n,1}(z_i))$, sustituimos z_i por el nodo x^*
 - Si $\text{sign}(f(x^*) - p_{n,1}(x^*)) = \text{sign}(f(z_{i+1}) - p_{n,1}(z_{i+1}))$, sustituimos z_{i+1} por el nodo x^*

A continuación se calcula un nuevo polinomio $p_{n,2}(x)$ m.a.u. para los nuevos nodos y se repite el razonamiento anterior de modo iterado.

Por lo tanto, en general, se obtendrá una sucesión de polinomios $\{p_{n,k}(x)\}_{k=0}^{\infty}$ con la propiedad:

$$\lim_{k \rightarrow \infty} p_{n,k}(x) = p(x)$$

donde $p(x)$ es la m.a.u. de $f \in C([a, b])$ en \mathbb{P}_n .

En el paso k -ésimo, se resuelve un sistema lineal del tipo:

$$f(z_j^{(k)}) - p_{n,k}(z_j^{(k)}) = (-1)^j r_k \quad j = 1, \dots, n+2 \quad (5.15)$$

con incógnitas, $a_{0,k}, a_{1,k}, \dots, a_{n,k}, r_k$ y los nodos $\{z_j^{(k)}\}$ son conocidos

En la práctica, el proceso debe detenerse tras un número finito de etapas introduciendo para ello un criterio de parada; por ejemplo,

$$|M_k - M_{d,k}| < tol$$

donde $M_k = \max_{x \in [a,b]} \{|f(x) - p_{n,k}(x)|\}$ y $M_{d,k} = |r_k|$.

Polinomios de Chebyshev y aproximación uniforme.

En el capítulo de interpolación se introdujeron los polinomios de Chebyshev para elegir los nodos óptimos de interpolación. Ahora, vemos que también son de utilidad en el cálculo de aproximaciones uniformes.

Así, supongamos que pretendemos obtener el polinomio de grado n m.a.u. de $f(x) = x^{n+1}$ en $[-1, 1]$; entonces,

Teorema 5.10 $p_n(x)$ es la m.a.u de x^{n+1} en \mathbb{P}_n si, y sólo si, $p(x) = x^{n+1} - p_n(x)$ es la m.a.u. de $f(x) = 0$ en el espacio de polinomios mónicos de grado $n + 1$.

Por lo tanto, a tenor de la caracterización de la mejor a.u. (teorema de alternancia) buscamos un polinomio mónico de grado $n + 1$ que alcance su valor absoluto máximo en $n + 2$ -nodos del intervalo $[-1, 1]$ con signos alternos.

Es conocido que la función $\cos(\theta)$ toma su valor absoluto máximo= 1 como sigue:

$$\cos(\theta) = \begin{cases} 1 & \text{si } \theta = 0 \\ -1 & \text{si } \theta = \pi \end{cases}$$

luego $\cos((n+1)\theta_j) = (-1)^j$ para $j = 0, 1, \dots, n+1$ con $\theta_j = \frac{j\pi}{n+1}$; es decir, alcanza su valor absoluto máximo en $n + 2$ puntos del intervalo $[0, \pi]$.

Ahora bien, si usamos el cambio de variable $\theta = \arccos(x)$ y definimos la función:

$$T_{n+1}(x) = \cos[(n+1) \arccos(x)]$$

Se verifican las siguientes propiedades:

- Para $j = 0, 1, \dots, n+1$ y $\theta_j = \frac{j\pi}{n+1}$; $T_{n+1}(x_j) = (-1)^j$
- Para cada n , $T_n(x) \in \mathbb{P}_n$ pues, $T_0(x) = 1$, $T_1(x) = x$ y

$$T_{n+1}(x) = 2xT_n - T_{n-1} \quad (5.16)$$

- El coeficiente principal de $T_n = \frac{1}{2^{n-1}}$
- Forman un sistema ortogonal de polinomios; a saber,

$$\int_{-1}^1 \frac{T_n T_m}{\sqrt{1-x^2}} dx = \begin{cases} 0 & \forall n \neq m \\ \frac{\pi}{2} & n = m \end{cases} \quad (5.17)$$

Con todo lo anterior es sencillo probar el siguiente resultado.

Teorema 5.11 La mejor aprox. uniforme de $f(x) = 0$ en $[-1, 1]$ en el espacio de polinomios mónicos de grado $n + 1$ es $\frac{T_{n+1}(x)}{2^n}$

Con este resultado concluimos nuestra búsqueda; es decir,

Corolario 5.2 La mejor aprox. uniforme de x^{n+1} en \mathbb{P}_n (en $[-1, 1]$) es $p(x) = x^{n+1} - \frac{T_{n+1}(x)}{2^n}$ y los nodos de alternancia son $x_j = \cos(\frac{j\pi}{n+1})$

Ejemplo: mejor aproximación uniforme de x^3 en $\mathbb{P}_2([-1, 1])$

Para obtener la m.a.u. necesitamos los polinomios de Chebyshev, T_0, T_1, T_2 y T_3 ; es decir,

$$T_0 = 1, \quad T_1 = x, \quad T_2 = 2xT_1 - T_0 = 2x^2 - 1, \quad T_3 = 2xT_2 - T_1 = 4x^3 - 3x$$

Por lo tanto, la m.a.u. será: $p(x) = x^3 - \frac{T_3}{4} = \frac{3}{4}x$

¿Qué ocurre si se desea la m.a.u. en el intervalo $[a, b]$?

En este caso, se procede a un cambio de variable que transforme el intervalo $[a, b]$ en el intervalo $[-1, 1]$ para, posteriormente, usar los polinomios de Chebyshev descritos anteriormente. El cambio es:

Si $t \in [-1, 1]$, entonces, $x = \frac{b-a}{2}t + \frac{a+b}{2} \in [a, b]$. Por lo tanto,

$$x^{n+1} = \left(\frac{b-a}{2}t + \frac{a+b}{2} \right)^{n+1} = \left(\frac{b-a}{2} \right)^{n+1} t^{n+1} + \dots$$

es un polinomio en $t \in [-1, 1]$ de grado $n+1$ y coeficiente líder $\left(\frac{b-a}{2}\right)^{n+1}$. Así, bastará darse cuenta que la m.a.u. de $f(t) = 0$ en el espacio de polinomios de grado $n+1$ y coeficiente principal $\left(\frac{b-a}{2}\right)^{n+1}$ es, $\left(\frac{b-a}{2}\right)^{n+1} \frac{T_{n+1}(t)}{2^n}$ para obtener el resultado siguiente:

Proposición 5.2 *En el intervalo $[a, b]$,*

- La m.a.u. de x^{n+1} en \mathbb{P}_n es el polinomio:

$$p_n(x) = x^{n+1} - \left(\frac{b-a}{2} \right)^{n+1} \frac{1}{2^n} T_{n+1} \left(\frac{a+b-2x}{a-b} \right) \quad (5.18)$$

- Los nodos de alternancia son:

$$x_j = \frac{b-a}{2} \cos\left(\frac{j\pi}{n+1}\right) + \frac{a+b}{2} \quad j = 0, 1, \dots, n+1$$

Los nodos anteriores son buenos nodos de inicio para el algoritmo de Remez.

Aproximación uniforme con polinomios de grado bajo: aproximación telescópica.

Dada una función $f(x) \in C([a, b])$ y un polinomio $q(x)$ que aproxima la función con error menor que ϵ ; es decir,

$$\|f - q\| < \epsilon$$

¿Es posible obtener otro polinomio de grado inferior que siga cumpliendo que el error sea $< \epsilon$?

Para conseguirlo nos basamos en:

Proposición 5.3 1. Todo polinomio $p(x) \in \mathbb{P}_m$ puede escribirse como: $p(x) = \sum_{j=0}^m \alpha_j T_j(x)$

2. La m.a.u. de $p(x)$ en \mathbb{P}_{m-1} es: $p_{m-1}(x) = \sum_{j=0}^{m-1} \alpha_j T_j(x)$ ó $p(x) - \alpha_m T_m(x)$

3. El error máximo es: $\|p(x) - p_{m-1}(x)\| = |\alpha_m|$

El proceso seguido para llegar al polinomio de grado lo más bajo posible con error máximo $< \epsilon$ se conoce con el nombre de **aproximación telescópica**.

Mas concretamente, supongamos que hemos conseguido un polinomio de grado m , p_m que cumple: $\|f - p_m\| < \epsilon_0 < \epsilon$ (p.e. mediante desarrollo de Taylor si $f(x)$ es suficientemente derivable).

Ahora, podemos calcular la m.a.u. de p_m en \mathbb{P}_{m-1} como se hace en la proposición anterior. Entonces, se tendrá:

- Si $\|p_m(x) - p_{m-1}(x)\| = \epsilon_1$ y $\epsilon_0 + \epsilon_1 < \epsilon$, tomamos el polinomio p_{m-1} como aproximación y repetimos el proceso.
- Si $\epsilon_0 + \epsilon_1 > \epsilon$ nos quedaremos con p_m y el proceso finaliza.

Ejemplo.

Hallar un polinomio de grado lo más bajo posible que aproxime $f(x) = e^x$, $x \in [-1, 1]$ con error menor que 0.2

Solución.

El polinomio de Taylor de e^x en $x=0$, de grado 3

$$p_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

tiene error máximo: $\|e^x - p_3(x)\| = \|\frac{e^\theta}{24}x^4\| < 0.125 = \epsilon_0 < 0.2$. Ahora, vemos si podemos conseguir un polinomio de grado dos y error menor que 0.2 (con los polinomios de Chebyshev). Como, $T_3(x) = 4x^3 - 3x$, entonces, la m.a.u. de $p_3(x)$ en \mathbb{P}_2 es:

$$p_2(x) = p_3(x) - \frac{1}{24}T_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} - \frac{1}{24}(4x^3 - 3x) = 1 + \frac{9}{8}x + \frac{x^2}{2}$$

¿Seguirá cumpliendo p_2 la propiedad $\|e^x - p_2\| < 0.2$?

En efecto, como $\|p_3 - p_2\| = \|\frac{1}{24}T_3(x)\| = \frac{1}{24} \simeq 0.041 = \epsilon_1$; entonces,

$$\epsilon_0 + \epsilon_1 = 0.166 < 0.2$$

y, por tanto, tomaríamos el nuevo polinomio como aproximación.

¿Se conseguirá rebajar aún más el grado? Compruébalo.

Para finalizar, esta sección se dan sendos resultados que permiten estudiar el problema de m.a.u de una función continua en un espacio vectorial de dimensión finita. Se trata de generalizaciones de lo ocurrido con \mathbb{P}_n .

Definición 5.6 Sea $H \in C([a, b])$ un subespacio de dimensión finita, n . Se dice que cumple la **condición de Haar** o es un **subespacio de Chebyshev** si, y sólo si, todo elemento **no nulo** de H admite, **a lo sumo, $(n - 1)$ ceros** en $[a, b]$. Además, en tal caso, una base de H se llama **sistema de Chebyshev**

Teorema 5.12 Sea $H \subset C([a, b])$ un subespacio de dimensión n . Entonces,

1. **[Teorema de Haar]** $f \in C([a, b])$ admite una **única m.a.u.** en H si, y sólo si, H es un subespacio de Chebyshev.
2. **[Teorema de Chebysev]** Si H es subespacio de Chebyshev, entonces $\varphi \in H$ es la **m.a.u.** de $f \in C([a, b])$ en H si, y sólo si, existen $n + 1$ -nodos, $x_0 < x_1 < \dots < x_n$, en $[a, b]$ verificando:

$$f(x_j) - \varphi(x_j) = (-1)^j M \quad j = 0, 1, \dots, n \tag{5.19}$$

donde $M = \pm \|f - \varphi\|$

5.5.2 Teorema de Weierstrass. Polinomios de Bernstein.

Aquí, se presenta una demostración constructiva del teorema de Weierstrass mencionado al inicio de la sección. En ella tiene una relevancia capital los llamados polinomios de Bernstein que describimos a continuación junto con alguna de sus propiedades más relevantes.

El conjunto de polinomios de grado n , $\left\{ b_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k} : k = 0, 1, \dots, n \right\}$ forman una base de \mathbb{P}_n llamada base de Bernstein que posee propiedades de interés como:

Proposición 5.4 En el intervalo $[0, 1]$, los polinomios básicos de Bernstein de grado n satisfacen:

$$0 \leq b_{n,k}(x) \leq 1 \quad (5.20)$$

$$1 = \sum_{k=0}^n b_{n,k}(x) \quad (5.21)$$

$$x = \sum_{k=0}^n \frac{k}{n} b_{n,k}(x) \quad (5.22)$$

$$\frac{x}{n} + \left(1 - \frac{1}{n}\right) x^2 = \sum_{k=0}^n \frac{k^2}{n^2} b_{n,k}(x) \quad (5.23)$$

$$\text{Recurrencia:} \quad b_{n,k}(x) = (1-x)b_{n-1,k}(x) + xb_{n-1,k-1}(x) \quad (5.24)$$

$$\text{Simetría:} \quad b_{n,k}(x) = b_{n,n-k}(1-x) \quad (5.25)$$

Observando detenidamente las propiedades anteriores cobra sentido el concepto siguiente:

Definición 5.7 Sea $f(x) \in C([0, 1])$ y $\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$ una partición uniforme del intervalo; entonces,

$$B_{n,f}(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) b_{n,k}(x) \quad (5.26)$$

es el n -ésimo polinomio de Bernstein asociado a f en $[0, 1]$

Ejemplo. Las gráficas de los polinomios de Bernstein de grados sucesivos ($n=1,2,4,8$) para la función $f(x) = \text{sen}(\pi x)$ en $[0, 1]$ son las de la figura siguiente:

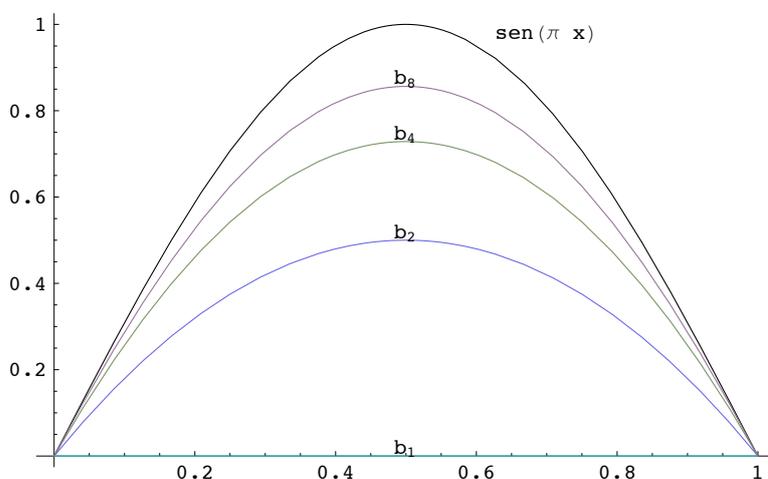


Figura 5.3: curva de f y polinomios de Bernstein.

En el ejemplo anterior se puede observar que los polinomios de Bernstein se aproximan a la función cosa que se describe en el teorema siguiente:

Teorema 5.13 *Sea f una función continua en $[0, 1]$; entonces, para cada $\epsilon > 0$ existe n_0 tal que*

$$\|f - B_{n,f}\| < \epsilon \quad \forall n \geq n_0$$

Este resultado permite demostrar el teorema de Weierstrass para una función continua en $[a, b]$ sin más que aplicar un cambio de variable adecuado para utilizar los polinomios de Bernstein y sus propiedades.

Más concretamente, se puede definir el **n-ésimo polinomio de Bernstein** asociado a f en $[a, b]$ como:

$$B_{n,f}(x) = \frac{1}{(b-a)^n} \sum_{k=0}^n f\left(a + \frac{k}{n}(b-a)\right) \binom{n}{k} (x-a)^k (b-x)^{n-k} \quad (5.27)$$

5.6 Curvas Bézier.

En esta sección se describen las propiedades básicas de las llamadas curvas Bézier que se fundamentan en los polinomios de Bernstein introducidos anteriormente y cuyo campo de aplicación principal está en el diseño gráfico con ordenador.

Definición 5.8 *Sea $\{p_0, p_1, \dots, p_n\}$ un conjunto de puntos de \mathbb{R}^2 ó \mathbb{R}^3 que llamaremos **red de control**; se llama *curva de Bézier de grado n asociada a la red a la curva paramétrica definida por:**

$$B_n(t) = \sum_{k=0}^n p_k b_{n,k}(t) \quad t \in [0, 1] \quad (5.28)$$

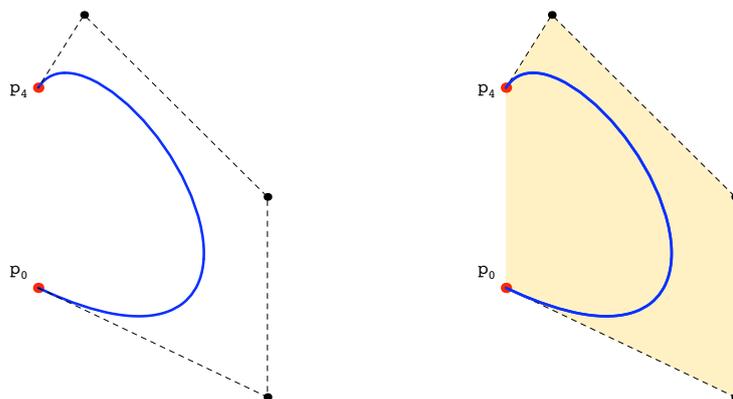


Figura 5.4: Red de control y curva Bézier de grado 4. Envoltente convexa y curva Bézier.

Algunas propiedades de interés.

1. La curva interpola p_0 y p_n ; es decir, $B_n(0) = p_0$, y $B_n(1) = p_n$ (ver figura 5.4)
2. Dado el conjunto $EC_n = \{p \in \mathbb{R}^k : p = \sum_{i=0}^n \lambda_i p_i \text{ con } \lambda_i \geq 0, \sum \lambda_i = 1\}$ (llamado **envoltente convexa** de la red de control); entonces, $B_n(t) \in EC_n \quad \forall t \in [0, 1]$ (figura 5.4)
3. Si $\bar{B}_n(t) = \sum_{k=0}^n p_{n-k} b_{n,k}(t)$; entonces, $B_n(t) = \bar{B}_n(1-t)$

4. Para cada recta del plano, el número de cortes con la curva Bézier es, a lo sumo, igual al número de cortes con el polígono de control de la curva.

5. La derivada paramétrica de la curva viene dada por:

$$B'_n(t) = n \sum_{k=0}^{n-1} (p_{k+1} - p_k) b_{n-1,k}(t) \tag{5.29}$$

y, en particular, $B'_n(0) = n(p_1 - p_0)$, $B'_n(1) = n(p_n - p_{n-1})$

Construcción de la curva Bézier: ALGORITMO DE CASTELJAU.

Una forma sencilla y eficiente de construir los puntos de la curva Bézier es debida a Casteljau que se explicita en el **algoritmo** siguiente:

- ENTRADA:
 - valor $t \in [0, 1]$,
 - Red de control: $p_{i,0} = p_i \quad i = 0, 1, \dots, n$
- Para $j = 1, \dots, n$
 - para $i = 0, \dots, n - j$
 - calcula $p_{i,j} = (1 - t)p_{i,j-1} + tp_{i+1,j-1} \quad i = 0, 1, \dots, n$
- SALIDA: el punto de la curva obtenido es: $B_n(t) = p_{0,n}$

En forma esquemática el algoritmo "desciende" desde los $(n + 1)$ - puntos iniciales hasta el punto de la curva tras n -etapas o niveles disminuyendo un punto en cada nivel. Esto se resume en la tabla siguiente:

Red=Nivel 0	Nivel 1	Nivel 2	...	Nivel n
$p_{00} = p_0$	p_{01}	p_{02}	...	$p_{0,n}$
$p_{10} = p_1$	p_{11}	p_{12}	...	
$p_{20} = p_2$	p_{21}	\vdots	\ddots	
$p_{30} = p_3$	\vdots	$p_{n-2,2}$		
\vdots	$p_{n-1,1}$			
$p_{n0} = p_n$				

Además, los puntos $p_{i,j}$ están situados en los segmentos que unen $p_{i,j-1}$ con $p_{i+1,j-1}$ como puede verse en la figura 5.5 donde, cada nivel, se identifica con un color diferente.

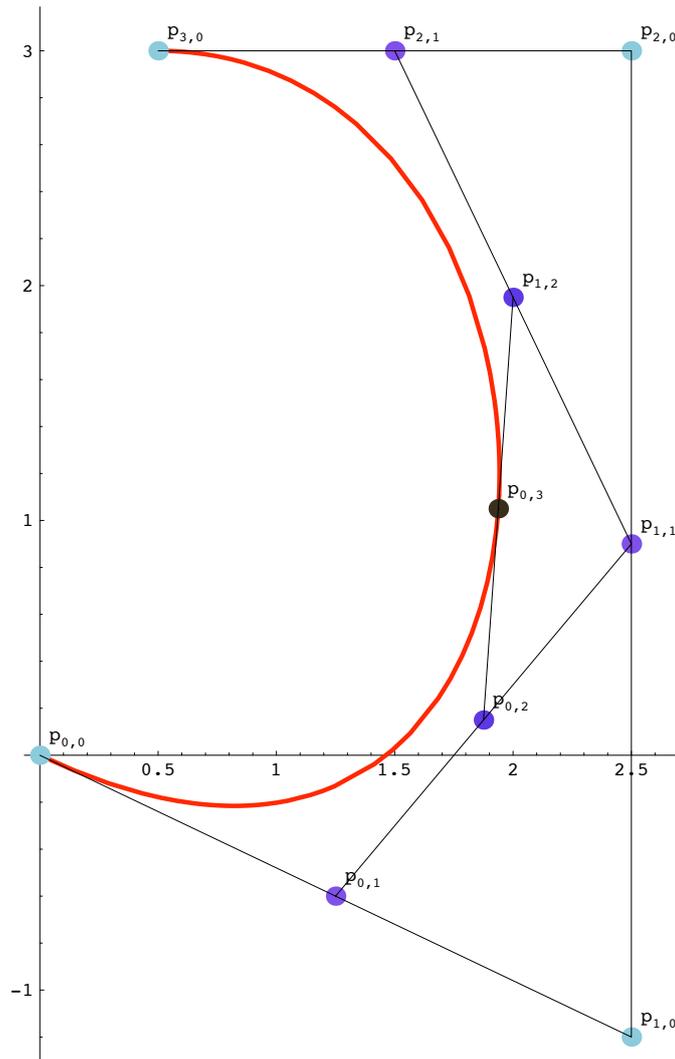


Figura 5.5: Evolución del algoritmo de Casteljau con $t = \frac{1}{2}$ y curva completa.

El algoritmo anterior permite construir la expresión analítica de la curva en forma iterada.

Ejemplo: Curva con red de control es: $\{(-1, 2), (1, -2), (0, 1)\}$

- En la tabla que sigue se puede observar cómo se construye la curva paso a paso.

Red=Nivel 0	Nivel 1	Nivel 2
$(-1, 2)$	$(-1 + 2t, 2 - 4t)$	$B_2(t) = (-1 + 4t - 3t^2, 2 - 8t + 7t^2)$
$(1, -2)$	$(1 - t, -2 + 3t)$	
$(0, 1)$		

- Ahora usamos la definición dada por 5.28; es decir,

$$B_2(t) = (-1, 2)b_{2,0}(t) + (1, -2)b_{2,1}(t) + (0, 1)b_{2,2}(t) = (x(t), y(t))$$

donde

$$x(t) = (-1)(1 - t)^2 + (1)2(1 - t)t + (0)t^2 = -3t^2 + 4t - 1$$

$$y(t) = (2)(1 - t)^2 + (-2)2(1 - t)t + (1)t^2 = 7t^2 - 8t + 2$$

por lo tanto,

$$B_2(t) = (-3t^2 + 4t - 1, 7t^2 - 8t + 2)$$

que, evidentemente, proporciona la misma curva que el algoritmo.

5.6.1 Curvas de Bézier compuestas.

Anteriormente se han descrito las curvas de Bézier para una red de $n+1$ puntos cuyo grado es n (en general); ahora deseamos obtener curvas del tipo anterior pero con grado limitado (el más popular es 3) pero que pueda servir como interpolante de un conjunto arbitrario de puntos en el plano.

Así, queremos ir desde un punto al siguiente con una curva Bézier de grado fijo (p.e. 3) y construir la curva por unión de todas ellas. En primer lugar, puesto que cada curva individual se construye desde una red de control con 4 puntos (grado $n = 3$) tendremos que introducir 2 puntos adicionales para cada pareja de puntos, consecutivos, iniciales.

Más precisamente, si inicialmente tenemos m -puntos, $\{p_0, p_1, \dots, p_m\}$, entonces para cada par de puntos p_i, p_{i+1} asociamos una red de control del tipo:

$$\{p_i = q_{0,i}, q_{1,i}, q_{2,i}, q_{3,i} = p_{i+1}\}$$

Ahora, se puede definir la curva Bézier, que notaremos por $C_i(t)$, como antes; es decir,

$$C_i(t) = q_{0,i}b_{3,0}(t) + q_{1,i}b_{3,1}(t) + q_{2,i}b_{3,2}(t) + q_{3,i}b_{3,3}(t) \quad t \in [0, 1] \quad (5.30)$$

Finalmente, la curva total o curva de Bézier compuesta para los m -puntos iniciales se puede definir como sigue:

$$BC(s) = \begin{cases} C_0(s) & \text{si } 0 \leq s \leq 1 \\ C_1(s-1) & \text{si } 1 \leq s \leq 2 \\ \vdots & \\ C_{m-1}(s-m+1) & \text{si } m-1 \leq s \leq m \end{cases}$$

Ejemplo.

Curva Bézier compuesta de grado 3 para los puntos $\{(0, 2), (1, 0.3), (2.5, 2.3), (4.5, 2), (5, 0.3)\}$.

Para estos puntos consideramos la red de control compuesta siguiente:

$$\{(0, 2), (-0.5, 1.), (0, 0), (1, 0.3), (2, 0.6), (1.3, 1.1), (2.5, 2.3), (2.8, 2.6), (4, 2.5), (4.5, 2), (5, 1.5), (5.3, 0.5), (5, 0.3)\}$$

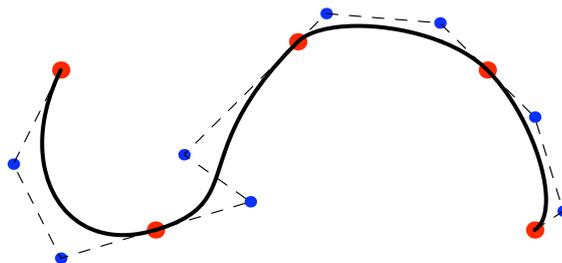


Figura 5.6: Curva de Bézier compuesta de grado 3.

Los puntos de color azul son los que completan, en cada caso, la red de control (figura 5.6). Además la situación relativa de éstos en trozos adyacentes incide sobre la suavidad global de la curva.

La continuidad de la curva está asegurada por la forma de elegir las redes de control; más precisamente,

$$q_{n,i-1} = p_i = q_{0,i} \quad i = 1, \dots, m - 1 \quad (5.31)$$

Por último, ¿que condición han de cumplir los puntos de control introducidos para asegurar la "suavidad" de la curva?

Si se observa detenidamente la curva se puede comprobar que, en cada punto rojo unión de dos trozos, los puntos azules adyacentes parecen alineados con él lo que se traduce en tangente única a izquierda y derecha del mismo.

Definición 5.9 Sea $BC(s)$ la curva de Bézier compuesta definida por (5.6); entonces,

- la curva tiene continuidad C^1 si, para $i = 0, \dots, m - 1$,

$$C'_{i-1}(t)|_{t=1} = C'_i(t)|_{t=0} \quad (5.32)$$

- la curva tiene continuidad geométrica G^1 si la igualdad (5.32) se transforma en una proporcionalidad; es decir,

$$C'_{i-1}(t)|_{t=1} = \lambda C'_i(t)|_{t=0} \quad (5.33)$$

donde t es el parámetro local de cada trozo de curva.

La definición anterior se traduce respecto de los puntos de control como sigue:

Proposición 5.5 Sean $\{p_0, p_1, \dots, p_m\}$ los puntos de interpolación de la curva y

$$\{q_{0,0}, \dots, q_{n-1,0}, q_{n,0}, \dots, q_{0,i}, \dots, q_{n-1,i}, q_{n,i}, \dots, q_{0,m-1}, \dots, q_{n-1,m-1}, q_{n,m-1}\}$$

con $q_{0,0} = p_0$, y $q_{n,i-1} = p_i = q_{0,i}$ $i = 1, \dots, m - 1$ una red de control compuesta asociada. Entonces, la curva tiene continuidad C^1 si los puntos de control adyacentes a cada punto intermedio de interpolación satisfacen:

$$\left. \begin{aligned} q_{n,i-1} - q_{n-1,i-1} &= q_{1,i} - q_{0,i} \\ &\text{ó} \\ q_{n-1,i-1} - 2p_i + q_{1,i} &= 0 \end{aligned} \right\} i = 1, \dots, m - 1 \quad (5.34)$$

Observaciones:

- Debe hacerse notar que aún siendo la tangente a la curva única, los vectores de dirección a izquierda y derecha pueden ser iguales (condición (5.34) o proporcionales (condición llamada continuidad geométrica G^1))
- La curva de la figura 5.6 tiene continuidad C^1 en las uniones primera y tercera pero continuidad G^1 en la segunda.