Aproximación e interpolación mediante polinomios

por

Miguel Marano y Marta Marcolini

En este trabajo se muestra una relación entre los conceptos de interpolación y mejor aproximación de una función continua mediante polinomios. Para ello se recurre a una definición de multiplicidad de un cero de una función continua.

Introducción

La aproximación y la interpolación de una función continua definida en un intervalo compacto de la recta mediante polinomios de grado a lo sumo n son dos problemas clásicos en matemáticas que tienen varias aplicaciones en las ciencias experimentales. Nuestro objetivo es mostrar una relación entre ambos. Un importante antecedente histórico es el clásico resultado de Chebyshev [1; ver también 2, Ch. 3, §4] que caracteriza al mejor polinomio de aproximación, en norma uniforme, mediante una condición de alternancia. No obstante, nuestra idea no es trabajar con un determinado criterio de mejor aproximación, como el de la norma uniforme, sino en partir sólo de una propiedad aproximativa del polinomio, que —creemos— éste debería cumplir para que sea considerado razonablemente bueno, al menos si asociamos la idea de aproximación con el concepto euclídeo de distancia entre dos puntos. La propiedad es la siguiente: no debe existir ningún otro polinomio de grado a lo sumo n que esté uniformemente más cerca de la función aproximada. Se prueba, en el Teorema 1, resultado principal de este trabajo, que esta propiedad es equivalente a que el polinomio sea interpolador de la función, en un sentido que se precisará después y que nos conduce a dar una definición de *multiplicidad* de un cero de una función continua. Según nuestro conocimiento, esta definición es nueva en la literatura. En la parte final del artículo se muestra qué criterios clásicos de mejor aproximación producen correspondientes polinomios que cumplen la propiedad susodicha.

Una propiedad de aproximación

Sea f una función continua definida en un intervalo $[a, b]$, $a < b$. Para un entero $n \geq 0$, fijo pero arbitrario, llamamos $\mathbb{P}_n$ al espacio lineal de los polinomios de grado a lo sumo n.

Definición 1. Se dice de un polinomio P en $\mathbb{P}_n$ que es una *buena aproximación de f sobre $\mathbb{P}_n$* si cada vez que Q en $\mathbb{P}_n$ satisface

$$|f - Q| \leq |f - P| \quad \text{en } [a, b],$$

resulta que $Q = P$.

Ejemplo 1. Consideremos la función $f(x) = |x|$ en el intervalo $[-1, 1]$. $P \equiv 1/2$ es el polinomio de mejor aproximación de f sobre $\mathbb{P}_1$ con respecto a la norma uniforme. En -1, 0 y 1 se produce la alternancia mencionada en la Introducción. P es también una buena aproximación de f sobre $\mathbb{P}_1$. Pero lo mismo es cierto para cada P_α, donde $P_\alpha(x) = \alpha x$, $-1 \leq \alpha \leq 1$.

Sea h una función continua definida en $[a, b]$. Las siguientes definiciones son usuales en el cálculo numérico. Un punto $z \in [a, b]$ se dice un *cero de h* si $h(z) = 0$. El punto $z \in [a, b]$ es un cero *aislado* de h si existe $\delta > 0$ tal que z es el único cero de h en $(z - \delta, z + \delta) \cap [a, b]$. Si $z \in (a, b)$ es un cero aislado de h, entonces existe un $\delta > 0$ tal que h tiene un signo determinado, tanto en $(z - \delta, z) \cap [a, b]$ como en $(z, z + \delta) \cap [a, b]$. Ahora queda completamente claro lo que significa que *h cambie su signo en z* (*no cambie su signo en z*). Si el cero aislado z de h es a, o b, entonces se entiende que h no cambia su signo en z. Observar que estas definiciones de cambio de signo en $z \in [a, b]$ son también (localmente) válidas para el cociente h_1/h_2 si h_1 y h_2 son continuas en $[a, b]$ y z es un cero aislado de ambas.

Sean h_1 y h_2 continuas en $[a, b]$. Si $z \in [a, b]$ es un cero aislado de ambas, diremos que $h_1(x) \succeq h_2(x)$ en $x = z$ si $h_2(x)/h_1(x)$ es acotada en un entorno *reducido* de z, es decir, en $[(z - \delta, z + \delta) \setminus \{z\}] \cap [a, b]$ para algún $\delta > 0$, y no cambia su signo en z.

Definición 2. Sea $z \in [a, b]$ un cero de la función continua h. Si z es un cero aislado, entonces se define la *multiplicidad* de z como el entero dado por

$$m_h(z) := \min \left\{ k \in \mathbb{N} : h(x) \succeq (x - z)^k \text{ en } x = z \right\}.$$

Si el conjunto de enteros positivos de arriba es vacío, se define $m_h(z) := +\infty$. Si z es un cero no aislado, entonces se define $m_h(z) := +\infty$.
Además, definimos Z_h de la siguiente manera: Si h no tiene ningún cero en $[a, b]$, entonces $Z_h := 0$; $Z_h := +\infty$ si h tiene infinitos ceros en $[a, b]$. Finalmente, $Z_h := m_h(z_1) + \cdots + m_h(z_i)$ si h sólo tiene los ceros (aislados) $z_1, \cdots, z_i$ en $[a, b]$, con multiplicidades $m_h(z_1), \cdots, m_h(z_i)$, respectivamente.

Si h es derivable r veces en el cero $z \in [a, b]$, $r \geq 1$ (derivable lateralmente en $z = a$ o $z = b$), entonces

$$h(x) = h'(z)(x - z) + \cdots + \frac{h^{(r)}(z)}{r!}(x - z)^r + R(x) \quad \text{para todo } x \in [a, b],$$

donde $(x-z)^{-r}R(x) \to 0$ cuando $x \to z$. De aquí se deduce fácilmente que la definición anterior de $m_h(z)$ coincide, para este caso y $m_h(z) = 1, 2, \cdots, r$, con la definición usual, por todos conocida.

Lema. *Sean h y g funciones continuas en $[a,b]$, y sea $z \in [a,b]$ un cero aislado de ambas. Entonces*

$$h(x) \succeq g(x) \text{ en } x = z \Longrightarrow m_h(z) \le m_g(z).$$

Demostración. La prueba es obvia si $m := m_g(z) = +\infty$. Luego supongamos $m < +\infty$. Así $(x-z)^m/g(x)$ es acotada en un entorno reducido de z, y no cambia su signo en z. Usando la hipótesis, se sigue que existe un entorno reducido de z donde las mismas condiciones son ciertas para $(x-z)^m/h(x) = [(x-z)^m/g(x)][g(x)/h(x)]$. Luego $m_h(z) \le m$. ■

Teorema 1. *Sea f una función continua definida en $[a,b]$ y sea $P \in \mathbb{P}_n$. Entonces las siguientes afirmaciones son equivalentes:*

(a) $Z_{f-P} \ge n+1$.

(b) *P es una buena aproximación de f sobre $\mathbb{P}_n$.*

Demostración.

(a) $\Longrightarrow$ (b) Supongamos que (b) no vale. Luego existe $Q \in \mathbb{P}_n$, $Q \ne P$, tal que

$$|f-Q| = |f-P-(Q-P)| \le |f-P| \quad \text{en } [a,b]. \tag{1}$$

Bajo esta suposición, probaremos que (a) no vale. Sigue de (1) que

$$|Q-P| \le |f-Q| + |f-P| \le 2\,|f-P| \quad \text{en } [a,b]. \tag{2}$$

Se deduce de aquí que los ceros de $f-P$, si los hay, son también ceros de $Q-P$ y por tanto son ceros aislados de $f-P$. Si $f-P$ no tiene ceros, entonces (a) no vale. Supongamos ahora que z es un cero (aislado) de $f-P$ de multiplicidad m, donde $1 \le m \le +\infty$. Se desprende de (2) y de la desigualdad en (1) que $(f-P)(x) \succeq (Q-P)(x)$ en $x = z$. Por el Lema, $m_{f-P}(z) \le m_{Q-P}(z)$. Se concluye de aquí que $Z_{f-P} \le Z_{Q-P} \le n$, que es la negación de (a). Esto prueba que (a) $\Longrightarrow$ (b).

(b) $\Longrightarrow$ (a) Neguemos (a) suponiendo que $Z_{f-P} =: r \le n$. Probaremos que P no es una buena aproximación de f sobre $\mathbb{P}_n$, que es negar (b). Si $r = 0$, entonces (b) es falso, ya que $|f-P-C| \le |f-P|$ para algún $C \in \mathbb{P}_0 \setminus \{0\}$, y $P + C \in \mathbb{P}_n$. Luego supongamos $r > 0$ y sean $m_1, \cdots, m_i$ las multiplicidades de todos y cada uno de los ceros $z_1, \cdots, z_i$ de $f-P$, respectivamente, donde $m_1 + \cdots + m_i = r$. Consideremos $Q \in \mathbb{P}_n$ dado por

$$Q(x) = (x-z_1)^{m_1} \ldots (x-z_i)^{m_i}.$$

Para cualquier entero j satisfaciendo $1 \leq j \leq i$, se tiene por definición de m_j que $(f-P)(x) \succeq Q(x)$ en $x = z_j$. De aquí sigue que $Q(x)/(f-P)(x)$ es acotada en un entorno reducido D_j de z_j, y no cambia su signo en z_j. Como además $Q/(f-P)$ es continua en el conjunto compacto $[a,b] \setminus \cup_{j=1}^{i} D_j$, y es continua y sin ceros en $[a,b] \setminus \{z_1, \cdots, z_i\}$, se concluye que existe un $\epsilon \in (-1,0) \cup (0,1)$ tal que se cumplen las dos siguientes condiciones:

- $|f(x) - P(x)| > |\epsilon Q(x)|$ para $x \in [a,b] \setminus \{z_1, \cdots, z_i\}$;
- El signo de $f - P$ coincide con el signo de ϵQ en $[a,b]$.

De estas dos condiciones se obtiene que

$$|f(x) - P(x) - \epsilon Q(x)| < |f(x) - P(x)| \quad \text{para } x \in [a,b] \setminus \{z_1, \cdots, z_i\},$$

lo que implica obviamente que P no es una buena aproximación de f sobre $\mathbb{P}_n$, por lo que (b) no se cumple. ■

Un criterio de mejor aproximación de una función continua f en un intervalo $[a,b]$ es un mecanismo bien definido de selección de un conjunto no vacío de polinomios en $\mathbb{P}_n$, los llamados polinomios de mejor aproximación de f (con respecto al criterio dado). Ahora podemos decir que un criterio de mejor aproximación es *bueno* si para toda función continua f cada polinomio P de mejor aproximación de f es una buena aproximación de f sobre $\mathbb{P}_n$.

Criterios clásicos de mejor aproximación

Los siguientes criterios de mejor aproximación son ampliamente conocidos: son aquéllos que determinan P como consecuencia de que se cumpla

$$\int_a^b |f - P|^p \, d\mu \leq \int_a^b |f - S|^p \, d\mu \quad \text{para todo } S \in \mathbb{P}_n,$$

donde p es un número real positivo, fijo pero arbitrario, y μ es una medida boreliana no nula. Denotemos $\mathcal{C}_{n,p}(\mu)$ a estos criterios de aproximación. Para evitar que toda función μ-medible definida en $[a,b]$ se identifique, fuera de un conjunto de medida nula, con infinitos polinomios en $\mathbb{P}_n$, pidámosle a μ el único requisito de no estar concentrada en n o menos puntos.

Teorema 2. *El criterio de mejor aproximación $\mathcal{C}_{n,p}(\mu)$ es bueno, siendo μ una medida boreliana no nula en $[a,b]$, no concentrada en n o menos puntos.*

Demostración. Sea $P \in \mathbb{P}_n$ un polinomio de mejor aproximación de f con respecto a $\mathcal{C}_{n,p}(\mu)$. Supongamos que $Q \in \mathbb{P}_n$ satisface

$$|f - Q| \leq |f - P| \quad \text{en } [a,b]. \tag{3}$$

El teorema quedará probado si mostramos que $Q = P$. Sea $T := (P + Q)/2$, de modo que $T \in \mathbb{P}_n$. Como consecuencia de (3) es fácil ver que

$$|f(x) - T(x)| < |f(x) - P(x)| \quad \text{si } Q(x) \neq P(x).$$

Ahora bien, si $Q \neq P$, entonces $Q(x) \neq P(x)$ para todo $x \in [a, b]$ excepto para, a lo sumo, n valores distintos de x. Como μ es una medida no concentrada en n o menos puntos, se concluye que $|f - T|^p < |f - P|^p$ en un conjunto de medida μ positiva en $[a, b]$, y por lo tanto

$$\int_a^b |f - T|^p \, d\mu < \int_a^b |f - P|^p \, d\mu,$$

lo cual contradice que P es un polinomio en $\mathbb{P}_n$ de mejor aproximación de f. De esta manera, Q debe coincidir con P. ■

Para el criterio particular de mejor aproximación $\mathcal{C}_{n,p}(\mu)$ es de esperar resultados más precisos en cuanto a la condición interpoladora de P. Esto es así en muchos casos, pero no en todos, como veremos más adelante. Por ejemplo, cuando $p > 1$ se sabe que $f - P$ debe tener al menos $n + 1$ ceros distintos. Aparte del trabajo pionero de Chebyshev, mencionado en la Introducción, esta conexión entre mejor aproximación e interpolación se ha investigado en los artículos [3] y [4]. En [3] se considera una medida μ concentrada en al menos $n + 2$ puntos, mientras que en [4] μ es la medida de Lebesgue. También en [5] se analizan estas cuestiones, pero bajo un enfoque más general, extendiéndose tanto los criterios como las funciones de aproximación.

Ahora se da un ejemplo simple con el fin de mostrar que para un polinomio P de mejor aproximación de f, aun con respecto a $\mathcal{C}_{n,1}(\mu)$, no es posible obtener un resultado más fino que el dado por (b) $\Longrightarrow$ (a) del Teorema 1.

Ejemplo 2. Sea f dada por $f(x) = |x|$ para $x \in [a, b] = [-1, 1]$. Calcularemos los polinomios en $\mathbb{P}_1$ de mejor aproximación de f con respecto a $\mathcal{C}_{1,1}(\mu)$, donde μ es la medida concentrada en -1, 0 y 1, con $\mu(\{-1\}) = \mu(\{1\}) = 1$ y $\mu(\{0\}) = 3$. Luego, para toda función continua h definida en $[-1, 1]$ tenemos

$$\int_{-1}^1 h \, d\mu = h(-1) + h(1) + 3h(0).$$

Sea P un polinomio en $\mathbb{P}_1$ de mejor aproximación de f. Si $T \in \mathbb{P}_1$ está dado por $T(x) = P(-x)$, sigue que

$$\int_{-1}^1 |f - T| \, d\mu = \int_{-1}^1 |f - P| \, d\mu.$$

Luego

$$\int_{-1}^1 |f - (P + T)/2| \, d\mu \leq \frac{1}{2} \int_{-1}^1 |f - P| \, d\mu + \frac{1}{2} \int_{-1}^1 |f - T| \, d\mu = \int_{-1}^1 |f - P| \, d\mu,$$

por lo que $\int_{-1}^{1} |f-(P+T)/2|\,d\mu = \int_{-1}^{1} |f-P|\,d\mu$ y de aquí resulta que $(P+T)/2$ es también un polinomio de mejor aproximación de f. Como $(P+T)/2$ es una función par, se deduce que $(P+T)/2 \in \mathbb{P}_0$. Así acabamos de probar que un polinomio en $\mathbb{P}_0$ de mejor aproximación de f con respecto a $\mathcal{C}_{0,1}(\mu)$ es también un polinomio de mejor aproximación de f con respecto a $\mathcal{C}_{1,1}(\mu)$. Basta pues encontrar una constante C que haga cumplir

$$\int_{-1}^{1} |f-C|\,d\mu \leq \int_{-1}^{1} |f-S|\,d\mu \quad \text{para toda constante } S.$$

Es fácil ver que $C = 0$ es la única solución de este problema de mínimos. Usando estos resultados, ahora no cuesta trabajo hallar todos los polinomios de mejor aproximación de f con respecto a $\mathcal{C}_{1,1}(\mu)$. Son aquellos P_λ dados por $P_\lambda(x) = \lambda x$, con $-1 \leq \lambda \leq 1$. De aquí, $f - P_\lambda$, con $-1 < \lambda < 1$, tiene sólo un cero, de multiplicidad 2, en $x = 0$. Ejemplos similares a éste pueden ser construidos para $\mathcal{C}_{n,1}$ con $n > 1$.

Finalmente consideramos otro criterio clásico de mejor aproximación de una función continua f, precisamente el de la norma uniforme, que produce su polinomio de mejor aproximación minimizando

$$\max_{x\in[a,b]} |f(x) - S(x)| \quad \text{sobre } S \in \mathbb{P}_n.$$

¿Es bueno este criterio? La respuesta es afirmativa, si bien no por una propiedad intrínseca del criterio sino debido a que en este caso el polinomio P de mejor aproximación de f es único, resultado este último muy conocido en teoría de mejor aproximación. De aquí la demostración de que P es una buena aproximación de f sobre $\mathbb{P}_n$ es inmediata.

Agradecimientos

Deseamos agradecer a los revisores del artículo por sus interesantes comentarios. En especial, la formulación de la Definición 2, equivalente a la dada originalmente por los autores pero menos complicada, así como una simplificación en la demostración del Teorema 1, fueron sugerencias de uno de los revisores.

Referencias

[1] P. L. Chebyshev, Théorie des mécanismes connus sous le nom de parallélogrammes, *Mém. Acad. Sci. Pétersb.* vii (1854).

[2] E. W. Cheney, *Introduction to Approximation Theory*, Chelsea Publishing Company, New York (1982).

[3] T. S. Motzkin, J. L. Walsh, Polynomials of best approximation on a real finite point set. I, *Trans. Amer. Math. Soc* **91** (1959), 231–245.

[4] T. S. Motzkin, J. L. Walsh, Polynomials of best approximation on an interval, *Proc. Nat. Acad. Sci. U.S.A.* **45** (1959), 1523–1528.

[5] J. R. Rice, Best approximations and interpolating functions, *Trans. Amer. Math. Soc.* **101** (1961), 477–498.

Miguel Marano y Marta Marcolini
Departamento de Matemáticas
Universidad de Jaén
23071 Jaén
correo electrónico: mmarano@ujaen.es
mmarcoli@ujaen.es