

Aproximación funcional e interpolación

Laboratori de Càlcul Numèric (LaCàN)

10 de abril de 2012

1. Introducción

Una función $p(x)$ definida en un intervalo $[a, b]$ aproxima a una función $f(x)$ en el mismo intervalo si

$$p(x) \simeq f(x) \tag{1}$$

Para describir un método de aproximación funcional, son necesarios dos ingredientes fundamentales. En primer lugar, hay que definir el espacio de funciones donde se elige el aproximante $p(x)$. Es decir, hay que elegir el *tipo de aproximación*.

En segundo lugar, es necesario formalizar matemáticamente el significado de “es aproximadamente igual a”, \simeq . Es decir, hay que definir el *criterio de aproximación*.

Ambos aspectos serán discutidos a continuación. Comentemos antes la utilidad de la aproximación funcional. Se utiliza en dos situaciones típicas: con la función dato $f(x)$ conocida a partir de su expresión analítica, o con $f(x)$ conocida de forma discreta, únicamente en unos puntos base x_i ($i = 0, \dots, n$).

Si se dispone de la expresión de $f(x)$, el objetivo es aproximarla por una función $p(x)$ más manejable y computable, que habitualmente nos permitirá además aproximar integrales o derivadas de la función original.

Si sólo se dispone de valores discretos $f(x_i)$, además del objetivo reseñado, nos interesa obtener $p(x)$ para poder evaluarlo en puntos distintos a los puntos base. Las funciones discretas son comunes en muchas áreas, desde el diseño asistido por ordenador (CAD) al tratamiento de señales, pasando por las medidas experimentales.

1.1. Tipo de aproximación

En la gran mayoría de aplicaciones prácticas, se trabaja con espacios de funciones de dimensión finita, y se expresa el aproximante $p(x)$ como combinación lineal de los términos de una base. De esta forma, una vez elegido el espacio de funciones y el criterio de aproximación, la determinación del aproximante $p(x)$ se reduce a la obtención de los coeficientes de combinación lineal.

Los espacios de funciones más utilizados son los siguientes:

1. Polinomios

$$p(x) = \sum_{i=0}^n a_i x^i \quad (2)$$

2. Funciones trigonométricas (series de Fourier):

$$p(x) = \sum_{i=0}^n a_i \cos(ix) + b_i \sin(ix) \quad (3)$$

3. Funciones exponenciales

$$p(x) = \sum_{i=0}^n a_i \exp(b_i x) \quad (4)$$

4. Funciones racionales

$$p(x) = \frac{\sum_{i=0}^n a_i x^i}{\sum_{j=0}^m b_j x^j} \quad (5)$$

5. Funciones definidas a trozos

Los polinomios son, con diferencia, las funciones más utilizadas en la aproximación funcional. Tienen dos propiedades muy atractivas: son fáciles de calcular, y sus derivadas e integrales son también polinomios.

Las funciones trigonométricas y las exponenciales tienen propiedades interesantes para el tratamiento de señales y en la resolución numérica de ecuaciones en derivadas parciales. Son algo más difíciles de tratar que los polinomios. Sus derivadas e integrales son funciones del mismo tipo.

Las funciones racionales son fáciles de calcular, pero sus derivadas e integrales son complicadas. Su principal interés es que, a diferencia de las demás funciones comentadas, permiten obtener aproximantes $p(x)$ con asíntotas verticales.

Para cualquiera de estas funciones, lo más habitual es aproximar la función $f(x)$ por una única función $p(x)$ en todo el intervalo $[a, b]$. En algunas aplicaciones, sin embargo, resultará más conveniente definir el aproximante a trozos, dividiendo el intervalo $[a, b]$ en subintervalos.

1.2. Criterios de aproximación

Una vez definido el tipo de aproximación, tres métodos permiten escoger las constantes que determinan el aproximante.

Interpolación

Una vez fijados $n + 1$ puntos base $\{x_0, x_1, \dots, x_n\}$, se exige que

$$p(x_i) = f(x_i) \quad \text{para } i = 0, \dots, n \quad (6)$$

Esta idea se puede generalizar al caso en que, además de los valores de la función $f(x)$ en los puntos base, se conozcan algunas de sus derivadas. Así pues, el desarrollo en serie de Taylor puede considerarse un caso particular de la interpolación.

Aproximación por mínimos cuadrados

Definida la norma \mathcal{L}_2 (euclídea) de forma discreta o continua, se minimiza la distancia entre el aproximante $p(x)$ y la función $f(x)$:

$$\min_{p(x)} (\|f(x) - p(x)\|_2)^2 \quad (7)$$

La norma \mathcal{L}_2 de una función $g(x)$ se define como

$$\|g(x)\|_2 = \left\{ \int_a^b [g(x)]^2 dx \right\}^{1/2} \quad (8)$$

en forma continua y como

$$\|g(x)\|_2 = \left\{ \sum_{i=0}^n [g(x_i)]^2 \right\}^{1/2} \quad (9)$$

en forma discreta (asociada a los $n + 1$ puntos base $\{x_0, x_1, \dots, x_n\}$).

Tal y como discutiremos más adelante, en algunos casos puede resultar interesante incluir una función de peso $w(x)$ en la definición de la norma \mathcal{L}_2 .

Aproximación mini-max

La idea es parecida a la aproximación por mínimos cuadrados, pero se trabaja con la norma \mathcal{L}_∞ (norma del máximo), que se define como

$$\|g(x)\|_\infty = \max_{x \in [a,b]} |g(x)| \quad (10)$$

en forma continua y como

$$\|g(x)\|_\infty = \max_{i=0,\dots,n} |g(x_i)| \quad (11)$$

en forma discreta.

El objetivo es, de nuevo, minimizar la distancia entre la función $f(x)$ y el aproximante $p(x)$:

$$\min_{p(x)} \|f(x) - p(x)\|_\infty = \min_{p(x)} \{\max |f(x) - p(x)|\} \quad (12)$$

1.3. Teorema de Weierstrass

Para terminar este apartado de introducción, enunciaremos el teorema de Weierstrass, que establece la “bondad” de los polinomios como funciones de aproximación.

Teorema 1. *Sea $f(x)$ continua en $[a, b]$. Para cualquier $\varepsilon > 0$, existe un entero n que depende de ε , $n(\varepsilon)$, tal que*

$$|f(x) - p_n(x)| < \varepsilon \quad (13)$$

para cualquier $x \in [a, b]$, donde $p_n(x)$ es un polinomio de grado n .

Es decir, toda función continua puede aproximarse uniformemente hasta una precisión arbitraria deseada mediante un polinomio. Conviene observar que el teorema no indica el grado del polinomio.

2. Interpolación polinómica

Uno de los métodos de aproximación funcional es la interpolación polinómica. Nótese que el nombre indica tanto el tipo de aproximación (polinomios) como el criterio de aproximación (interpolación).

2.1. Propiedades básicas de los polinomios

Empezaremos recordando algunas propiedades básicas de los polinomios, que nos resultarán útiles al estudiar las distintas técnicas de obtención del polinomio interpolador $p(x)$.

Estructura de espacio vectorial de los polinomios

Sea P_n el conjunto de los polinomios de grado menor o igual a n . Este conjunto tiene estructura de espacio vectorial. Esto significa que cualquier polinomio de P_n puede escribirse como combinación lineal de los elementos de una base. De hecho, la forma habitual de escribir un polinomio,

$$p_n(x) = \sum_{i=0}^n a_i x^i = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \quad (14)$$

corresponde a expresarlo en la llamada *base trivial* de los polinomios,

$$\{1, x, x^2, \dots, x^n\} \quad (15)$$

El número de elementos de la base ($n + 1$) indica la dimensión del espacio P_n . Nótese que los coeficientes a_i del polinomio son, por lo tanto, sus componentes en la base trivial.

La base trivial de polinomios es una manera muy cómoda y natural de expresar un polinomio. Sin embargo, en algunos casos resultará preferible trabajar con otras bases, tal y como discutiremos más adelante.

Evaluación de polinomios: regla de Horner

¿Cuántas operaciones son necesarias para evaluar un polinomio $p_n(x)$ en un cierto punto \hat{x} ? Depende de cómo se organicen las operaciones (sumas y productos) de la ecuación (14).

Si se efectúan $i - 1$ productos para calcular \hat{x}^i , se realizan $n(n + 1)/2$ productos y n sumas (compruébese). Está claro que esta manera de evaluar el polinomio es poco eficiente: si se utiliza una variable auxiliar para almacenar la potencia \hat{x}^{i-1} , entonces el cálculo de \hat{x}^i implica un único producto (en lugar de $i - 1$). De esta forma se consigue reducir el número de operaciones a $2n - 1$ productos y n sumas (verifíquese).

Este número de operaciones puede rebajarse más utilizando la llamada *regla de Horner*, que consiste en evaluar $p_n(\hat{x})$ como

$$p_n(\hat{x}) = a_0 + \hat{x} \left(a_1 + \hat{x} \left(a_2 + \hat{x} \left(\cdots + \hat{x} (a_{n-1} + \hat{x} a_n) \right) \right) \right) \quad (16)$$

La expresión (16) contiene únicamente n productos y n sumas. Esta reducción en el número de productos es un aspecto relevante siempre que sea necesario evaluar un polinomio muchas veces durante un cálculo.

Desarrollo en serie de Taylor

Una función $f(x)$ suficientemente continua puede aproximarse por un polinomio $p(x)$ en el entorno de un punto mediante un desarrollo en serie de Taylor.

Si $f(x)$ es de clase \mathcal{C}^{n+1} (continua y con las $n + 1$ primeras derivadas continuas) en el entorno de un punto x_0 , puede expresarse como

$$f(x) = p_n(x) + R_n(x) \quad (17)$$

con

$$p_n(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2!}f''(x_0) + \cdots + \frac{(x - x_0)^n}{n!}f^{(n)}(x_0) \quad (18)$$

$$R_n(x) = \frac{(x - x_0)^{n+1}}{(n + 1)!}f^{(n+1)}(\mu(x)) \quad \text{con } \mu \in [x_0, x] \quad (19)$$

La diferencia entre la función $f(x)$ y el polinomio aproximante $p_n(x)$ es el resto de Lagrange $R_n(x)$, que puede interpretarse como un caso particular de error de truncamiento: el desarrollo en serie de Taylor se trunca en un número finito de términos. Dicho error de truncamiento depende de la derivada $(n + 1)$ -ésima de $f(x)$ en un cierto punto $\mu(x)$ desconocido del intervalo $[x_0, x]$.

Observación 1. Se denota por $[x_0, x]$ el intervalo de extremos x_0 y x , ya sea con $x_0 < x$ o con $x < x_0$.

Puesto que $\mu(x)$ es desconocido, no podremos, en general, evaluar el resto de Lagrange. A efectos prácticos, sin embargo, será suficiente obtener una cota de dicho error:

$$|R_n(x)| \leq \frac{|x - x_0|^{n+1}}{(n + 1)!} \max_{\tau \in [x_0, x]} |f^{(n+1)}(\tau)| \quad (20)$$

2.2. Teorema fundamental del álgebra

¿En qué condiciones puede garantizarse la existencia y unicidad del polinomio interpolador? El siguiente teorema aborda esta cuestión.

Teorema 2. Sea $(x_i, f(x_i))$ con $i = 0, \dots, n$ un conjunto de $n + 1$ puntos tales que $x_i \neq x_j$ para $i \neq j$ (abscisas diferentes). Entonces, existe un único polinomio $p_n(x)$ de grado menor o igual que n tal que $p(x_i) = f(x_i)$ para $i = 0, \dots, n$.

Demostración

La demostración es constructiva: además de demostrar que el polinomio existe y es único, veremos un posible procedimiento para obtenerlo.

Para ello, resulta útil expresar el polinomio $p_n(x)$ en la base trivial, ecuación (14):

$$p_n(x) = \sum_{j=0}^n a_j x^j$$

Se trata simplemente de obtener los coeficientes a_j . Impongamos el criterio de interpolación. Para $i = 0, \dots, n$, debe verificarse que $p_n(x_i) = f(x_i)$ y, por tanto:

$$\sum_{j=0}^n a_j x_i^j = f(x_i) \quad (21)$$

Tenemos pues un sistema lineal de orden $n + 1$:

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix} \quad (22)$$

La matriz del sistema (22) es una *matriz de Vandermonde*, cuyo determinante es distinto de cero si todas las x_i son distintas. Así pues, el sistema lineal a resolver para obtener los coeficientes a_i es compatible y determinado y, por lo tanto, el polinomio interpolador $p_n(x)$ existe y es único.

Observación 2. *Al resolver el sistema (22), puede ocurrir que el coeficiente a_n (y otros) sea nulo. El polinomio $p_n(x)$ es entonces de grado menor que n . En cualquier caso, el teorema garantiza que se trata del único polinomio de grado menor o igual que n que interpola los $n + 1$ puntos dados.*

A pesar de responder a una idea muy natural, el procedimiento descrito plantea dos inconvenientes: la matriz de Vandermonde es mal condicionada y, además, no se dispone de ningún tipo de información sobre el resto de Lagrange.

En la práctica se utilizan otros métodos para obtener el polinomio interpolador, que corresponden a trabajar con bases de polinomios distintas a la trivial.

2.3. Interpolación de Lagrange

En la interpolación de Lagrange, el polinomio $p_n(x)$ se expresa como

$$p_n(x) = \sum_{i=0}^n a_i L_i(x) \quad (23)$$

donde

$$\{L_0(x), L_1(x), L_2(x), \dots, L_n(x)\} \quad (24)$$

es la base de polinomios de Lagrange, que enseguida definiremos.

Al imponer el criterio de interpolación, se obtiene el sistema lineal

$$\begin{bmatrix} L_0(x_0) & L_1(x_0) & \cdots & L_n(x_0) \\ L_0(x_1) & L_1(x_1) & \cdots & L_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ L_0(x_n) & L_1(x_n) & \cdots & L_n(x_n) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} \quad (25)$$

La idea de la interpolación de Lagrange es simplificar al máximo el sistema lineal a resolver. De hecho, se exige que la matriz sea la identidad. Ello equivale a imponer que los polinomios de Lagrange verifiquen

$$L_i(x_j) = \delta_{ij} \quad (26)$$

donde δ_{ij} es la delta de Kronecker ($\delta_{ij} = 1$ para $i = j$, $\delta_{ij} = 0$ para $i \neq j$). Esto se consigue definiendo los polinomios de Lagrange como

$$L_i(x) = \frac{\prod_{\substack{j=0 \\ j \neq i}}^n (x - x_j)}{\prod_{\substack{j=0 \\ j \neq i}}^n (x_i - x_j)} = \frac{(x - x_0) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n)}{(x_i - x_0) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)} \quad (27)$$

Nótese que todos los polinomios $L_i(x)$ son de grado n .

Puesto que la matriz del sistema (25) es la identidad, los coeficientes a_i son simplemente las ordenadas $f(x_i)$, y el polinomio interpolador se escribe como

$$p_n(x) = \sum_{i=0}^n f(x_i) L_i(x) \quad (28)$$

Tal y como comentábamos en el apartado 2.1, trabajar con una base de polinomios distinta a la trivial resulta claramente ventajoso, por que los datos $f(x_i)$ son directamente los coeficientes de combinación lineal.

Otra ventaja de la interpolación de Lagrange es que nos proporciona una expresión del resto de Lagrange.

El resto de Lagrange

La ecuación (19) proporciona el resto de Lagrange para el desarrollo en serie de Taylor, en el que el polinomio aproximante se obtiene a partir del valor de la función y sus n primeras derivadas en un mismo punto x_0 .

¿Cómo es el resto de Lagrange para la interpolación polinómica a partir de $n + 1$ puntos?

Proposición 1. Sea $f(x)$ una función de clase \mathcal{C}^{n+1} y sea $p_n(x)$ el polinomio interpolador puro que verifica $p_n(x_i) = f(x_i)$ para $i = 0, \dots, n$. Sea $R_n(x)$ el error de $p_n(x)$ como aproximante de $f(x)$, $f(x) = p_n(x) + R_n(x)$. Dicho resto de Lagrange $R_n(x)$ puede expresarse como

$$R_n(x) = L(x) \frac{f^{(n+1)}(\mu)}{(n+1)!} \quad (29)$$

con

$$\mu \in [x_0, x_1, \dots, x_n, x] \quad (30)$$

$$L(x) = \prod_{i=0}^n (x - x_i) = (x - x_0)(x - x_1) \cdots (x - x_n) \quad (31)$$

Demostración

Comprobemos en primer lugar que la expresión (29) es válida para los puntos base de la interpolación x_i . Para ello, basta notar que $L(x_i) = 0$ ($L(x)$ es precisamente un polinomio de grado $n + 1$ que tiene como ceros los puntos base x_i) y, en consecuencia, $R_n(x_i) = 0$, tal como corresponde al criterio de interpolación, $p_n(x_i) = f(x_i)$.

Pasemos ahora al caso no trivial: $x = \hat{x} \neq x_i$. Empezaremos definiendo una función auxiliar $g(x)$ como

$$g(x) = f(x) - p_n(x) - KL(x) \quad (32)$$

Puesto que $L(\hat{x}) \neq 0$, la constante K puede definirse de tal manera que la función $g(x)$ se anule en el punto \hat{x} que estamos considerando:

$$g(\hat{x}) = 0 \Leftrightarrow K = \frac{f(\hat{x}) - p_n(\hat{x})}{L(\hat{x})} \quad (33)$$

De esta forma, podemos asegurar que la función $g(x)$ tiene, como mínimo, $n + 2$ ceros en el intervalo $[x_0, x_1, \dots, x_n, \hat{x}]$ (es decir, el intervalo que contiene a todos estos puntos): los $n + 1$ puntos base x_i y el punto \hat{x} .

Si aplicamos reiteradamente el teorema de Rolle, podemos asegurar que, en el intervalo $[x_0, x_1, \dots, x_n, \hat{x}]$, la función $g'(x)$ tiene, como mínimo, $n + 1$ ceros; la función $g''(x)$, n ceros; \dots ; la función $g^{(n+1)}(x)$, un cero μ .

Recordando que $p_n(x)$ y $L(x)$ son polinomios de grado n y $n + 1$ respectivamente, la derivada $n + 1$ -ésima de $g(x)$ es

$$g^{(n+1)}(x) = f^{(n+1)}(x) - K(n + 1)! \quad (34)$$

Puesto que $g^{(n+1)}(\mu) = 0$, podemos obtener una nueva expresión para K :

$$g^{(n+1)}(\mu) = f^{(n+1)}(\mu) - K(n + 1)! = 0 \Rightarrow K = \frac{f^{(n+1)}(\mu)}{(n + 1)!} \quad (35)$$

Finalmente, basta igualar las dos expresiones de K , ecuaciones (33) y (35) para obtener que

$$f(\hat{x}) = p_n(\hat{x}) + L(\hat{x}) \frac{f^{(n+1)}(\mu)}{(n + 1)!} \quad (36)$$

y verificar que la expresión del resto de Lagrange dada por la ecuación (29) es efectivamente válida.

Observación 3. *Nótese que μ depende de x : al variar el punto \hat{x} , varía también la función auxiliar $g(x)$ y, en consecuencia, también lo hace el cero μ de su derivada $(n + 1)$ -ésima.*

3. Aproximación por mínimos cuadrados

3.1. Motivación

En algunas aplicaciones, el criterio de interpolación (polinómica o de otro tipo) no resultará satisfactorio. Puede ocurrir, por ejemplo, que los datos $f(x_i)$ sean medidas experimentales afectadas por un cierto error inherente; en ese caso, no tiene demasiado sentido exigir que la función aproximante $p(x)$ pase exactamente por unos puntos que contienen errores.

Por otra parte, al aumentar el número de puntos base x_i , el polinomio interpolador puro presenta muchas oscilaciones ya que aumenta su grado n , y esto dificulta o impide su utilización en muchas aplicaciones: dibujo por ordenador, cálculo de derivadas, integración numérica, etc. Resultará preferible utilizar una función más suave (por ejemplo, un polinomio de grado $m < n$) y renunciar al criterio de interpolación.

Comentemos, para terminar, una aplicación muy habitual de la aproximación por mínimos cuadrados en la que se combinan estas dos ideas. Supongamos que disponemos de un cierto modelo teórico de un fenómeno físico. Por ejemplo, una relación cuadrática entre la abscisa x y la ordenada y , $y = f(x) = ax^2 + bx + c$. Tenemos también medidas experimentales, $(x_i, f(x_i))$ con $i = 0, \dots, n$. Nuestro objetivo es determinar el polinomio $p_2(x)$ (es decir, los coeficientes a , b y c) óptimo en el sentido de mínimos cuadrados. No nos interesa obtener el polinomio interpolador puro $p_n(x)$ por tres motivos: (1) los datos experimentales están afectados por errores de medida y (2) $p_n(x)$ presenta oscilaciones importantes que no se ajustan a la física del problema que, según hemos supuesto, (3) se describe mejor por un polinomio de grado 2.

3.2. Planteamiento general

Recordemos (véase el apartado 1.2) que la aproximación por mínimos cuadrados consiste en minimizar $\|f(x) - p(x)\|$, es decir la distancia (medida en norma euclídea) entre la función $f(x)$ y el aproximante $p(x)$. Esta norma está inducida por un producto escalar, $\langle \cdot, \cdot \rangle$: $\|f - p\|^2 = \langle f - p, f - p \rangle$.

Recordemos en primer lugar las propiedades de las normas y los productos escalares:

Producto escalar

$\langle \cdot, \cdot \rangle$ es un producto escalar si es una forma bilineal simétrica y definida positiva:

1. Linealidad: $\langle \alpha f + \beta g, h \rangle = \alpha \langle f, h \rangle + \beta \langle g, h \rangle$ para escalares α , β y funciones f , g y h cualesquiera.
2. Simetría: $\langle f, g \rangle = \langle g, f \rangle$ para funciones f y g cualesquiera.
3. Definición positiva: $\langle f, f \rangle \geq 0$; $\langle f, f \rangle = 0 \Leftrightarrow f = 0$

Norma

$\|\cdot\|$ es una norma si, para funciones f y g y escalar α cualesquiera,

1. $\|f\| \geq 0$; $\|f\| = 0 \Leftrightarrow f = 0$
2. $\|\alpha f\| = |\alpha| \|f\|$
3. $\|f + g\| \leq \|f\| + \|g\|$

Es importante observar que disponemos, tanto para el producto escalar como para la correspondiente norma euclídea inducida, de dos versiones, la continua y la discreta:

$$\|f\|^2 = \langle f, f \rangle \text{ con } \langle f, g \rangle = \begin{cases} \int_a^b f(x)g(x)dx & \text{caso continuo} \\ \sum_{i=0}^n f(x_i)g(x_i) & \text{caso discreto} \end{cases} \quad (37)$$

En ambos casos, puede resultar útil incluir una función de peso positiva $w(x) > 0$ en la definición del producto escalar:

$$\langle f, g \rangle = \begin{cases} \int_a^b w(x)f(x)g(x)dx & \text{caso continuo} \\ \sum_{i=0}^n w_i f(x_i)g(x_i) & \text{caso discreto} \end{cases} \quad (38)$$

Antes de abordar, en el apartado siguiente, la obtención del aproximante, hagamos dos consideraciones fundamentales. En primer lugar, resulta equivalente y más cómodo minimizar *el cuadrado* de la distancia entre la función $f(x)$ y el aproximante $p(x)$ que la distancia,

$$E = (\|f(x) - p(x)\|)^2 = \langle f(x) - p(x), f(x) - p(x) \rangle \quad (39)$$

En segundo lugar, nos restringiremos al caso en que el aproximante $p(x)$ puede escribirse como combinación lineal de las funciones de una base

$$\{\psi_0(x), \psi_1(x), \dots, \psi_m(x)\} \quad (40)$$

es decir,

$$p(x) = \sum_{i=0}^m c_i \psi_i(x) \quad (41)$$

Así pues, determinar el aproximante $p(x)$ se reduce a determinar los correspondientes coeficientes c_i .

El análisis de combinaciones no lineales de funciones de aproximación es considerablemente más complejo y menos utilizado que el caso lineal que tratamos aquí.

3.3. Las ecuaciones normales

Gracias a la expresión (41), hemos modificado profundamente el problema a resolver. Ya no se trata de minimizar el funcional $E(f)$, ecuación (39), sino únicamente de minimizar una función E de $m + 1$ variables reales:

$$\begin{aligned} \min_{\mathbf{c} \in \mathbb{R}^{m+1}} E(c_0, c_1, \dots, c_m) \\ = \min_{\mathbf{c} \in \mathbb{R}^{m+1}} \left\{ \left\langle f(x) - \sum_{i=0}^m c_i \psi_i(x), f(x) - \sum_{j=0}^m c_j \psi_j(x) \right\rangle \right\} \end{aligned} \quad (42)$$

La función $E(c_0, c_1, \dots, c_m)$ es una función continua de los coeficientes (c_0, c_1, \dots, c_m) , véase [IK94]. Por consiguiente, se puede imponer la condición necesaria de extremo, es decir

$$\frac{\partial E}{\partial c_i} = 0 \quad \text{para } i = 0, \dots, m \quad (43)$$

Para ello, resulta útil reescribir la ecuación (39) como

$$E = \langle f(x), f(x) \rangle - 2 \langle p(x), f(x) \rangle + \langle p(x), p(x) \rangle \quad (44)$$

Derivemos ahora respecto de c_i . Puesto que $\partial p(x)/\partial c_i = \psi_i(x)$ se llega a

$$\frac{\partial E}{\partial c_i} = -2 \langle \psi_i(x), f(x) \rangle + 2 \langle \psi_i(x), p(x) \rangle = 0 \quad \text{para } i = 0, \dots, m \quad (45)$$

Para la obtención de las ecuaciones (44) y (45) se ha tenido en cuenta la linealidad y la simetría del producto escalar.

La ecuación (45) proporciona dos resultados de interés. En primer lugar,

$$\langle \psi_i(x), f(x) - p(x) \rangle = 0 \quad \text{para } i = 0, \dots, m \quad (46)$$

que establece la ortogonalidad entre la diferencia $f(x) - p(x)$ (o sea, el error de $p(x)$ como aproximación a $f(x)$) y cada una de las funciones $\psi_i(x)$. Así pues, la aproximación por mínimos cuadrados puede interpretarse como la

proyección ortogonal (según el producto escalar, continuo o discreto, de trabajo) de la función $f(x)$ sobre el espacio de aproximación generado por las funciones $\psi_i(x)$.

En segundo lugar, expresando $p(x)$ como combinación lineal de las $\psi_i(x)$ llegamos a

$$\sum_{j=0}^m \langle \psi_i(x), \psi_j(x) \rangle c_j = \langle \psi_i(x), f(x) \rangle \quad \text{para } i = 0, \dots, m \quad (47)$$

o, más explícitamente, a

$$\begin{bmatrix} \langle \psi_0, \psi_0 \rangle & \langle \psi_0, \psi_1 \rangle & \cdots & \langle \psi_0, \psi_m \rangle \\ \langle \psi_1, \psi_0 \rangle & \langle \psi_1, \psi_1 \rangle & \cdots & \langle \psi_1, \psi_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \psi_m, \psi_0 \rangle & \langle \psi_m, \psi_1 \rangle & \cdots & \langle \psi_m, \psi_m \rangle \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} \langle \psi_0, f \rangle \\ \langle \psi_1, f \rangle \\ \vdots \\ \langle \psi_m, f \rangle \end{bmatrix} \quad (48)$$

Se trata de un sistema lineal de ecuaciones de orden $m+1$ cuyas incógnitas son los coeficientes c_j denominado *ecuaciones normales*.

Teorema 3. *Si las funciones $\psi_0(x), \psi_1(x), \dots, \psi_m(x)$ son linealmente independientes, el problema de mínimos cuadrados (42) tiene una solución única,*

$$p(x) = \sum_{i=0}^m c_i \psi_i(x)$$

donde los coeficientes c_i verifican las ecuaciones normales (47).

Demostración

Basta con comprobar que la matriz de las ecuaciones normales es regular (en otras palabras, que el sistema lineal es compatible determinado). Para ello, supondremos que la matriz es singular con funciones $\psi_i(x)$ linealmente independientes y llegaremos a una contradicción.

Si la matriz es singular, el sistema lineal homogéneo tiene soluciones distintas a la trivial,

$$\sum_{j=0}^m \langle \psi_i, \psi_j \rangle \hat{c}_j = 0 \quad \text{para } i = 0, 1, \dots, m \text{ con algún } \hat{c}_j \neq 0 \quad (49)$$

Calculemos ahora la norma al cuadrado de la función $\sum_{j=0}^m \hat{c}_j \psi_j(x)$:

$$\begin{aligned} & \left\| \sum_{i=0}^m \hat{c}_i \psi_i(x) \right\|^2 \\ &= \left\langle \sum_{i=0}^m \hat{c}_i \psi_i(x), \sum_{j=0}^m \hat{c}_j \psi_j(x) \right\rangle = \sum_{i=0}^m \underbrace{\left[\sum_{j=0}^m \langle \psi_i, \psi_j \rangle \hat{c}_j \right]}_{=0} \hat{c}_i = 0 \quad (50) \end{aligned}$$

Hemos llegado pues a la conclusión de que la función $\sum_{j=0}^m \hat{c}_j \psi_j(x)$ es nula (la norma sólo se anula para la función cero). Puesto que las funciones $\psi_j(x)$ son linealmente independientes, los coeficientes \hat{c}_j deben ser todos nulos. Esto lleva a una contradicción con la hipótesis realizada en la ecuación (49).

La regresión lineal

Hasta ahora hemos discutido el *criterio* de aproximación de mínimos cuadrados, pero no hemos tratado el *tipo* de aproximación, véanse los apartados 1.1 y 1.2.

Veamos un caso concreto de aproximación por mínimos cuadrados: la regresión lineal (ajuste de una recta a medidas experimentales). Corresponde a elegir P_1 (el espacio de polinomios de grado menor o igual que uno) como espacio de aproximación, la base trivial $\{\psi_0(x) = 1, \psi_1(x) = x\}$ y el producto escalar discreto, ecuación (37).

Las ecuaciones normales son, en este caso,

$$\begin{aligned} & \begin{bmatrix} \langle \psi_0, \psi_0 \rangle & \langle \psi_0, \psi_1 \rangle \\ \langle \psi_1, \psi_0 \rangle & \langle \psi_1, \psi_1 \rangle \end{bmatrix} \begin{Bmatrix} c_0 \\ c_1 \end{Bmatrix} = \begin{Bmatrix} \langle \psi_0, f \rangle \\ \langle \psi_1, f \rangle \end{Bmatrix} \\ & \implies \begin{bmatrix} m+1 & \sum_{i=0}^m x_i \\ \sum_{i=0}^m x_i & \sum_{i=0}^m x_i^2 \end{bmatrix} \begin{Bmatrix} c_0 \\ c_1 \end{Bmatrix} = \begin{Bmatrix} \sum_{i=0}^m f(x_i) \\ \sum_{i=0}^m x_i f(x_i) \end{Bmatrix} \quad (51) \end{aligned}$$

y la recta de regresión es $p_1(x) = c_0 + c_1 x$.

Malcondicionamiento de las ecuaciones normales

La matriz de las ecuaciones normales (48) puede ser muy mal condicionada. Para ilustrar este punto, supongamos que el espacio de aproximación en P_m (polinomios de grado menor o igual que m), que se trabaja con la base trivial,

$$\psi_0(x) = 1 \quad ; \quad \psi_1(x) = x \quad ; \quad \psi_2(x) = x^2 \quad ; \dots ; \quad \psi_m(x) = x^m \quad (52)$$

Cuadro 1: Las matrices de Hilbert son muy mal condicionadas

Orden ($m + 1$)	Número de condición
2	$1,9 \times 10^1$
3	$5,2 \times 10^2$
5	$4,8 \times 10^5$
10	$1,6 \times 10^{13}$
15	$6,1 \times 10^{20}$

y que el producto escalar es

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx \quad (53)$$

En estas condiciones, la matriz del sistema es la llamada *matriz de Hilbert* de orden $m + 1$,

$$\begin{bmatrix} 1 & 1/2 & 1/3 & \cdots & 1/(m+1) \\ 1/2 & 1/3 & & & \vdots \\ 1/3 & & & & \vdots \\ \vdots & & & & \vdots \\ 1/(m+1) & \cdots & \cdots & \cdots & 1/(2m+1) \end{bmatrix} \quad (54)$$

El cuadro 1 muestra la relación entre el orden y el número de condición de la matriz de Hilbert.

3.4. Funciones ortogonales

Para eludir los problemas numéricos asociados al malcondicionamiento de las ecuaciones normales, puede optarse por trabajar con una base de funciones ortogonales,

$$\langle \psi_i(x), \psi_j(x) \rangle = 0 \quad \text{para } i \neq j \quad (55)$$

Se tiene entonces un sistema lineal de ecuaciones con matriz diagonal,

$$\begin{bmatrix} \langle \psi_0, \psi_0 \rangle & & & \\ & \langle \psi_1, \psi_1 \rangle & & \\ & & \cdots & \\ & & & \langle \psi_m, \psi_m \rangle \end{bmatrix} \begin{Bmatrix} c_0 \\ c_1 \\ \vdots \\ c_m \end{Bmatrix} = \begin{Bmatrix} \langle \psi_0, f \rangle \\ \langle \psi_1, f \rangle \\ \vdots \\ \langle \psi_m, f \rangle \end{Bmatrix} \quad (56)$$

que puede resolverse explícitamente, obteniendo los llamados coeficientes de Fourier c_i ,

$$c_i = \frac{\langle \psi_i, f \rangle}{\langle \psi_i, \psi_i \rangle} \quad \text{para } i = 0, \dots, m \quad (57)$$

Para los aproximantes y productos escalares más habituales, las correspondientes funciones ortogonales pueden encontrarse tabuladas en manuales de fórmulas matemáticas o en la biblioteca de funciones de distintos paquetes de software matemático.

Si se trabaja con aproximación polinómica y el producto escalar continuo en el intervalo $[-1, 1]$, por ejemplo, las funciones ortogonales son los llamados *polinomios de Legendre*; si el producto escalar es discreto con puntos equiespaciados en el intervalo $[-1, 1]$, los *polinomios de Gram*.

Referencias

- [IK94] Eugene Isaacson and Herbert Bishop Keller. *Analysis of numerical methods*. Dover Publications, New York, 1994. Corrected reprint of the 1966 original [John Wiley & Sons, New York].