

6	Aproximación de funciones e interpolación	109
6.1	Aproximación de funciones e interpolación en problemas físicos	109
6.2	Interpolación polinómica global	109
6.2.1	Interpolación de Lagrange	110
6.2.2	Fenómeno de Runge para la interpolación polinómica global	113
6.2.3	Errores de redondeo	116
6.2.4	Diferencias divididas de Newton	117
6.2.5	Interpolación en tablas y diferencias finitas	121
6.2.6	Errores en la interpolación en tablas (no entra en examen)	123
6.2.7	Interpolación osculatoria o de Hermite	125
6.2.8	Estimación de las derivadas de Bessel	129
6.3	Interpolación polinómica trozos	130
6.3.1	Interpolación nodal por polinomios a trozos discontinuos	130
6.3.2	Interpolación nodal por polinomios a trozos continuos	135
6.3.3	Interpolación nodal por polinomios a trozos con derivada continua	137
6.3.4	Interpolación mediante splines cúbicas	141
6.4	Otra derivación de la interpolación con esplines cúbicas (no entra en examen)	146

6.5	Polinomios trigonométricos	148
6.5.1	Interpolación trigonométrica y transformada discreta de Fourier	152
6.5.2	Transformada rápida de Fourier (no entra en examen)	155
6.6	Aproximación de funciones	156
6.6.1	Aproximación por polinomios. Teorema de Weierstrass	156
6.6.2	Teoría de Sturm-Liouville	157
6.6.3	Funciones trigonométricas de Legendre, Bessel, Hermite y Chebyshev	162
6.6.4	Aproximaciones continuas mediante mínimos cuadrados	166
6.6.5	*Aproximación polinómica uniforme ¡¡¡¡NO!!!	167

Bibliografía**169**

19 de noviembre de 2001

© Francisco R. Villatoro, Carmen M. García, Juan I. Ramos. Estas notas están protegidas por derechos de copyright y pueden ser distribuidas libremente sólo con propósitos educativos sin ánimo de lucro. *These notes are copyright-protected but may be freely distributed for instructional nonprofit purposes.*

CAPÍTULO 6

APROXIMACIÓN DE FUNCIONES E INTERPOLACIÓN

6.1 Aproximación de funciones e interpolación en problemas físicos

Problema de interpolación de una función $f(x)$: Dado un conjunto de nodos (abcisas) $\{x_j\}$, con $j = 0, 1, \dots, n$, y sean los valores de la función en dichos nodos, $\{f(x_j)\}$, encuentre una función p_n (normalmente, un polinomio) que coincida con la función en todos los nodos, $p_n(x_j) = f(x_j)$.

Problema de interpolación en tablas: Dada una tabla de puntos (pares)

$$\{(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)\},$$

encuentre una función p_n (polinomio) que pase por todos los puntos, $p_n(x_j) = y_j$.

Tipos de interpolación: Depende del tipo de la función, puede ser global (si es una sola función para todo el intervalo) o local (si se divide el intervalo en subintervalos y se utiliza una función por cada subintervalo). Además influye el tipo de la función: polinomio, polinomio trigonométrico (series de Fourier), etc.

La interpolación sirve para aproximar funciones, como base para las técnicas de integración, de resolución de ecuaciones diferenciales ordinarias (y en derivadas parciales), de resolución de ecuaciones no lineales, etc. y para calcular valores de funciones descritas mediante tablas.

6.2 Interpolación polinómica global

Dados $(n + 1)$ pares de puntos $(x_0, y_0), \dots, (x_n, y_n)$ con $x_i \neq x_j$ si $i \neq j$, podemos determinar un polinomio (de interpolación) $p_n(x)$ de grado n tal que $p_n(x_i) = y_i$. Note que dada $f(x)$,

podemos hacer $y_j = f(x_j)$.

Escribiendo el polinomio de interpolación en forma estándar

$$p_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

y aplicando la condición de interpolación, obtenemos un sistema lineal de $(n + 1)$ ecuaciones para los coeficientes a_i de la forma

$$a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_n x_i^n = y_i,$$

o en forma matricial

$$A a = \begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix} = b.$$

La matriz de coeficientes de este sistema A es una matriz de Vandermonde. Una matriz de Vandermonde tiene como determinante

$$\text{Det}(A) = \prod_{j>i=0}^n (x_j - x_i) \neq 0,$$

cuya demostración se puede encontrar en libros de álgebra lineal, por lo que siempre tiene inversa, por lo que existe una solución única del sistema lineal considerado y por tanto existe un único polinomio que pasa por los $(n + 1)$ puntos. Sin embargo, esta matriz de Vandermonde es una matriz mal condicionada (especialmente si al menos dos x_i están próximos entre sí) y el procedimiento de calcular el polinomio de interpolación mediante la solución de este sistema lineal también está mal condicionado.

Es muy importante notar que el polinomio de interpolación es único, aunque ello no quita que se pueda escribir de diferentes formas. En las siguientes secciones veremos su cálculo mediante bases de Lagrange y de Newton.

6.2.1 Interpolación de Lagrange

El polinomio $p_n(x)$ se puede escribir

$$p_n(x) = \sum_{j=0}^n y_j l_j(x),$$

donde $l_j(x)$ es un polinomio de Lagrange para la secuencia de puntos $(x_0, y_0), \dots, (x_n, y_n)$ con $x_i \neq x_j$, que cumple con la propiedad $l_j(x_i) = \delta_{ij}$, es decir, una delta de Kronecker. Estos polinomios se pueden escribir fácilmente como

$$l_j(x) = \frac{\prod_{i=0, j \neq i}^n (x - x_i)}{\prod_{i=0, j \neq i}^n (x_j - x_i)} = \frac{(x - x_0)(x - x_1) \cdots (x - x_{j-1})(x - x_{j+1}) \cdots (x - x_n)}{(x_j - x_0)(x_j - x_1) \cdots (x_j - x_{j-1})(x_j - x_{j+1}) \cdots (x_j - x_n)}.$$

Ejemplo. Tomando solamente dos puntos $(x_0, y_0), (x_1, y_1)$,

$$p_1(x) = a_0 + a_1 x = y_0 \frac{x - x_1}{x_0 - x_1} + y_1 \frac{x - x_0}{x_1 - x_0}.$$

Este polinomio coincide con la solución del sistema lineal de Vandermonde

$$\begin{pmatrix} 1 & x_0 \\ 1 & x_1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix},$$

que es

$$\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \frac{1}{x_1 - x_0} \begin{pmatrix} x_1 & -x_0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_1 y_0 - x_0 y_1 \\ y_1 - y_0 \end{pmatrix},$$

con lo que

$$\begin{aligned} p_1(x) &= \frac{1}{x_1 - x_0} (x_1 y_0 - y_1 x_0 + (y_1 - y_0) x) \\ &= \frac{1}{x_1 - x_0} (y_0 (x_1 - x) + y_1 (x - x_0)), \end{aligned}$$

que coincide con el polinomio de Lagrange. Fin de ejemplo.

El polinomio de Lagrange se puede escribir de otras formas. Por ejemplo, denotando

$$\varphi(x) = (x - x_0) \cdots (x - x_n) = \prod_{i=0}^n (x - x_i),$$

cuya derivada es

$$\frac{d\varphi(x)}{dx} = \varphi'(x) = \sum_{i=0}^n (x - x_0) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n),$$

que evaluada en $x = x_i$ nos da

$$\varphi'(x_i) = (x - x_0)(x - x_1) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n),$$

por lo que el polinomio interpolador de Lagrange se puede escribir como

$$p_n(x) = \sum_{i=0}^n f(x_i) l_i(x) = \sum_{i=0}^n f(x_i) \frac{\varphi(x)}{(x - x_i) \varphi'(x_i)}.$$

Teorema. El error de la interpolación lagrangiana

$$E(x) = f(x) - p_n(x) = f(x) - \sum_{j=0}^n f(x_j) l_j(x),$$

es exactamente igual a

$$E(x) = \frac{(x - x_0)(x - x_1) \cdots (x - x_n)}{(n + 1)!} f^{(n+1)}(\xi),$$

donde $\xi \in I \equiv [x, x_0, \dots, x_n] \equiv [\min\{x, x_j\}, \max\{x, x_j\}]$, es decir, el corchete es el intervalo más pequeño que contiene a todos los números $\{x, x_j\}$.

Demostración. La fórmula es trivialmente cierta para $x \equiv x_i$, ya que $E(x_i) = 0$. Sea x un valor fijado a partir de este momento. Definamos la función

$$G(y) = E(y) - \frac{\varphi(y)}{\varphi(x)} E(x),$$

donde

$$\varphi(y) = \prod_{j=0}^n (y - x_j).$$

Si $f \in \mathbf{C}^{n+1}$ entonces $G \in \mathbf{C}^{n+1}$,

$$G(x_i) = E(x_i) - \frac{\varphi(x_i)}{\varphi(x)} E(x) = 0, \quad i = 0, \dots, n,$$

y además $G(x) = 0$, luego G tiene $(n + 2)$ ceros en el intervalo I , y por el teorema del valor medio, G' tiene $(n + 1)$ ceros en I , G'' tiene n ceros en I , e iterativamente, $G^{(j)}$ tiene $(n + 2 - j)$ ceros en I con $j = 0, \dots, n + 1$, por lo que $G^{(n+1)}$ tiene exactamente un cero en I , sea éste ξ . Ahora bien, como $p^{(n+1)}(x) \equiv 0$,

$$E^{(n+1)}(\xi) = f^{(n+1)}(\xi), \quad \varphi^{(n+1)}(\xi) = (n + 1)!,$$

y

$$G^{(n+1)}(y) = f^{(n+1)}(y) - \frac{(n + 1)!}{\varphi(x)} E(x),$$

tenemos que

$$G^{(n+1)}(\xi) = 0 = f^{(n+1)}(\xi) - \frac{(n + 1)!}{\varphi(x)} E(x),$$

luego

$$E(\xi) = \frac{\varphi(t)}{(n+1)!} f^{(n+1)}(\xi),$$

cqd.

Ejemplo. Para el caso de sólo dos puntos,

$$E(x) = f(x) - \frac{(x_1 - x)f(x_0) + (x - x_0)f(x_1)}{x_1 - x_0} = \frac{(x_0 - x)(x_1 - x)}{2!} f''(\xi),$$

donde $\xi \in [x_0, x_1, x]$. Fin de ejemplo.

Con la fórmula de error de interpolación podemos acotar el error puntual en la forma

$$|f(x) - p_n(x)| \leq \frac{|x - x_0| \cdots |x - x_n|}{(n+1)!} \max_{\xi \in I} |f^{(n+1)}(\xi)|,$$

que introduciendo la norma infinito

$$\|f\|_\infty = \max_{x \in I} |f(x)|,$$

nos conduce a

$$\|f - p_n\|_\infty \leq \frac{\|\varphi\|_\infty}{(n+1)!} \|f^{(n+1)}\|_\infty.$$

Tomando $[a, b] = I \equiv [x_0, x_1, \dots, x_n, x]$, podemos relajar esta cota a

$$\|f(x) - p_n(x)\|_\infty \leq \frac{(b-a)^{n+1}}{(n+1)!} \|f^{(n+1)}(x)\|_\infty.$$

6.2.2 Fenómeno de Runge para la interpolación polinómica global

Aparentemente, la cota del error de interpolación

$$\|f(x) - p_n(x)\|_\infty \leq \frac{(b-a)^{n+1}}{(n+1)!} \|f^{(n+1)}(x)\|_\infty,$$

parece indicar que el error decrecerá conforme se introduzcan más nodos de interpolación ($n \rightarrow \infty$). Sin embargo, es importante notar que eso no es cierto, es decir, no siempre se cumple que

$$\lim_{n \rightarrow \infty} \|f(x) - p_n(x)\| = 0,$$

ya que la cota depende de que las derivadas de la función $f(x)$ no crezcan demasiado rápido. Este hecho, se expresa en la existencia funciones para las que la sucesión $\{p_n(x)\}$ diverge conforme n crece [2]. Por ejemplo, Runge demostró en 1901 que para la función (llamada de Runge)

$$f(x) = \frac{1}{1+x^2},$$

la sucesión $\{p_n(x)\}$, de polinomios que la interpolan con $n+1$ puntos distribuidos uniformemente en el intervalo $[-5, 5]$, para $|x| > c$, con $c \approx 3.63$, diverge conforme n crece indefinidamente. Bernstein probó en 1912 que para la función $f(x) = |x|$, interpolada en el intervalo $[-1, 1]$, la sucesión $\{p_n(x)\}$ diverge para todo $0 < |x| < 1$, es decir, en casi todo punto del intervalo.

Sin embargo, Krylov probó en 1962 que si la función $f(z)$, de variable compleja, es analítica en la región definida por la intersección de los dos círculos de radio $b - a$ centrados en $z = a$ y $z = b$, entonces la interpolación de la función $f(x)$, de variable real, en el intervalo $[a, b]$ converge conforme el número de nodos, n , en dicho intervalo tiende a infinito, es decir,

$$\lim_{n \rightarrow \infty} \|f(x) - p_n(x)\| = 0.$$

Recordemos que la función de Runge no es analítica ya que posee dos polos complejos, $z = \pm i$, y, por tanto, no cumple las hipótesis del teorema de Krylov en el intervalo $[-5, 5]$.

Estas dificultades de convergencia con el proceso de interpolación se conocen con el nombre de fenómeno de Runge. Desde un punto de vista práctico este fenómeno se refleja en grandes oscilaciones del polinomio interpolador, sobre todo en los extremos del intervalo de interpolación. Una justificación intuitiva de este comportamiento es que como el polinomio de interpolación $p_n(x)$, de grado n , tiene n ceros, que en muchos casos se encuentran dentro del intervalo de integración, incluso si la función original $f(x)$ no tiene ceros en el mismo, éste se ve obligado a realizar grandes oscilaciones cerca de los extremos del intervalo de interpolación.

Cuando se interpola una función suave que no tiene ceros pero que decrece suficientemente rápido a cero, por ejemplo, la función de Runge

$$f(x) = \frac{1}{1 + x^2},$$

mediante un polinomio de grado n , que tiene n ceros, es posible que el polinomio interpolador oscile excesivamente cerca de los extremos del intervalo. La figura 6.1 muestra la función de Runge y su polinomio de interpolación para los intervalos $[-5, 5]$ (parte superior) y $[-1, 1]$ (parte inferior) para diferente número de nodos.

En la parte superior izquierda de la figura 6.1 se observa las grandes oscilaciones que se producen en los extremos del intervalo para un polinomio de grado 10. En el centro del intervalo el polinomio de interpolación es bastante preciso. En la parte superior derecha se observa que estas oscilaciones en los extremos aumentan mucho conforme se aumenta el número de nodos.

En la parte inferior izquierda de la figura 6.1 se observa que las oscilaciones del polinomio interpolador de grado 10 se reducen grandemente cuando el intervalo de interpolación se reduce.

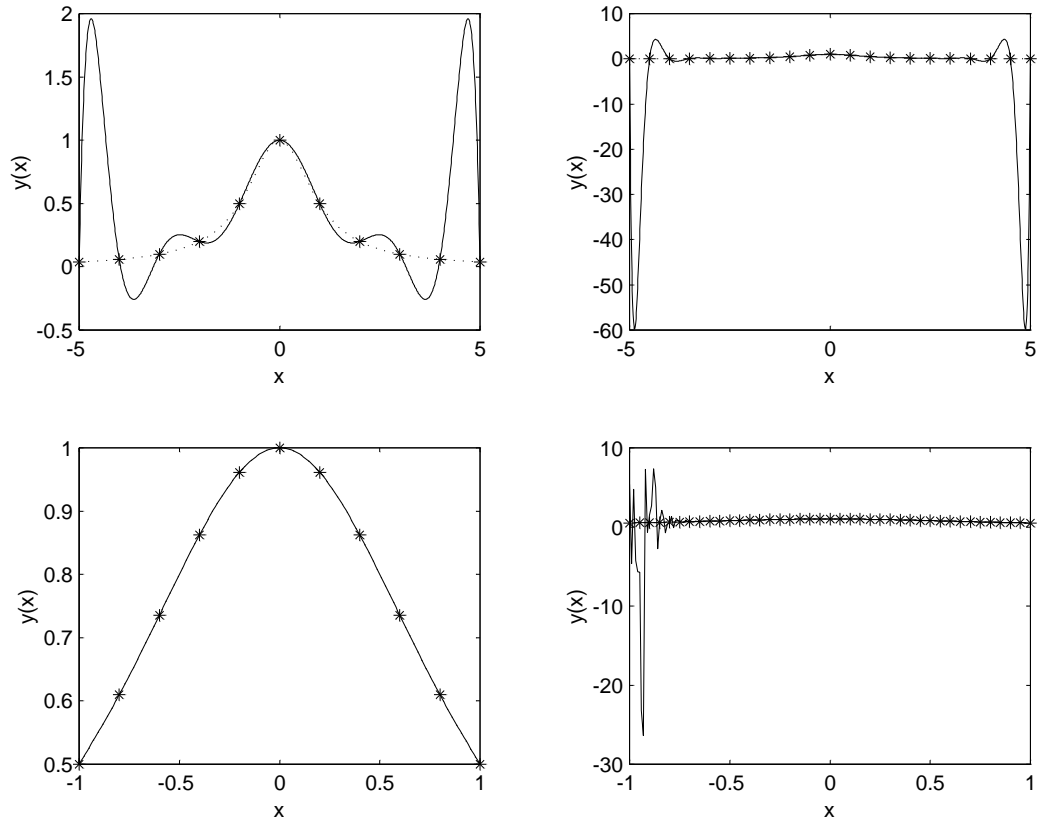


Figura 6.1. Ilustración del fenómeno de Runge. La línea discontinua muestra la función de Runge, la continua el polinomio interpolador de Lagrange y los asteriscos la función en los nodos de interpolación. Las figura superior izquierda y derecha utilizan 11 y 22 nodos de interpolación en el intervalo $[-5, 5]$, respectivamente, y las figuras inferior izquierda y derecha, 11 y 41 nodos en $[-1, 1]$, respectivamente.

Sin embargo, no desaparecen estas oscilaciones, como muestra la parte inferior derecha de dicha figura. Las oscilaciones que se producen en esta figura son debidas en gran medida a errores de redondeo por diferencias cancelativas en la evaluación del polinomio de Lagrange, para lo que se ha utilizado el algoritmo de Aitken (que no estudiaremos en este curso).

Para evitar estas oscilaciones se requiere un grado muy alto de dicho polinomio.

Con la fórmula de error de interpolación podemos acotar el error puntual en la forma

$$|f(x) - p_n(x)| \leq \frac{|x - x_0| \cdots |x - x_n|}{(n + 1)!} \max_{\xi \in I} |f^{(n+1)}(\xi)|,$$

que introduciendo la norma infinito

$$\|f\|_\infty = \max_{x \in I} |f(x)|,$$

nos conduce a

$$\|f(x) - p_n(x)\|_\infty \leq \frac{\|\varphi(x)\|_\infty}{(n + 1)!} \|f^{(n+1)}(x)\|_\infty,$$

que indica que el error se reduce conforme aumenta el número de nodos, ya que para $f(x) \in \mathbf{C}^{n+1}(I)$,

$$\lim_{n \rightarrow \infty} \|f(x) - p_n(x)\| = 0,$$

Sin embargo esta cota, en general, no es uniforme, es decir,

$$\lim_{n \rightarrow \infty} |f(x) - p_n(x)| \neq 0,$$

ya que, por ejemplo, $f(x) = e^x$ no tiene ningún cero y sin embargo $p_n(x)$ tiene n ceros, número que crece conforme n crece. A este comportamiento oscilatorio de la interpolación (global) por polinomios se le denomina fenómeno de Runge.

6.2.3 Errores de redondeo

En la expresión del error también podemos tener en cuenta los errores de redondeo. Para el caso de dos puntos, supondremos un error en los valores dados de la función f a interpolar,

$$f(x_0) = f_0 - \epsilon_0, \quad f(x_1) = f_1 - \epsilon_1,$$

por lo que el error total $\Sigma(x)$

$$\begin{aligned} \Sigma(x) &= f(x) - \frac{(x_1 - x)f(x_0) + (x - x_0)f(x_1)}{x_1 - x_0} + \frac{(x_1 - x)\epsilon_0 + (x - x_0)\epsilon_1}{x_1 - x_0} \\ &= E(x) + R(x). \end{aligned}$$

Podemos acotar fácilmente esta expresión para el error. Para los errores de redondeo que siguen una línea recta

$$R(x) = \frac{(x_1 - x) \epsilon_0 + (x - x_0) \epsilon_1}{x_1 - x_0},$$

los podemos acotar por

$$\max_{x_0 \leq x \leq x_1} |R(x)| = \max\{|\epsilon_0|, |\epsilon_1|\},$$

y por otro lado el error de interpolación es una parábola,

$$E(x) = \frac{(x_0 - x)(x_1 - x)}{2!} f''(\xi), \quad \xi \in [x_0, x_1, x]$$

por lo que, calculando su máximo,

$$\frac{dE}{dx} = \frac{x - x_1 + x - x_0}{2} f''(\xi) = 0, \quad x = \frac{x_0 + x_1}{2},$$

$$\max E(x) = -\frac{(x_1 - x_0)^2}{8} f''(\xi),$$

tenemos finalmente la siguiente cota del error total

$$|\Sigma(x)| \leq \frac{h^2}{8} \max |f''(\xi)| + \max\{|\epsilon_0|, |\epsilon_1|\}.$$

6.2.4 Diferencias divididas de Newton

Las desventajas fundamentales de la interpolación de Lagrange son (1) hay que recalcular los polinomios de Lagrange $l_i(x)$ cada vez que se añada o quite un punto, es decir, el polinomio obtenido interpolando $n - 1$ puntos no sirve para nada a la hora de calcular el polinomio interpolante con n puntos; y (2) requiere un gran número de operaciones aritméticas y es susceptible de producir diferencias cancelativas que son proclives a la propagación de los errores de redondeo.

Un método más eficiente de determinar el polinomio de Lagrange es utilizar el método de diferencias divididas de Newton. Dados $(n + 1)$ puntos $\{(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)\}$ con $x_i \neq x_j$, el método de Newton consiste en calcular un polinomio $p_n(x)$ de grado n tal que

$$p_n(x_i) = y_i, \quad p_n(x) = p_{n-1}(x) + C(x),$$

donde $p_{n-1}(x)$ es el polinomio interpolador para los n primeros puntos, $\{(x_0, y_0), \dots, (x_{n-1}, y_{n-1})\}$, y $C(x)$ es un polinomio de grado n . De esta forma, observamos que

$$C(x_i) = p_n(x_i) - p_{n-1}(x_i) = y_i - y_i = 0, \quad i = 0, 1, \dots, n - 1,$$

por lo que

$$C(x) = a_n (x - x_0) \cdots (x - x_{n-1}),$$

y para calcular a_n , aplicamos

$$p_n(x_n) = y_n = a_n (x_n - x_0) \cdots (x_n - x_{n-1}) + p_{n-1}(x_n),$$

$$a_n = \frac{y_n - p_{n-1}(x_n)}{\prod_{i=0}^{n-1} (x_n - x_i)} = \frac{p_n(x_n) - p_{n-1}(x_n)}{\prod_{i=0}^{n-1} (x_n - x_i)},$$

y a los coeficientes a_n se les denomina diferencias divididas de Newton y se denotan

$$a_n \equiv f[x_0, x_1, \dots, x_n],$$

por lo que el polinomio interpolador es

$$p_n(x) = p_{n-1}(x) + f[x_0, x_1, \dots, x_n] (x - x_0) \cdots (x - x_{n-1}),$$

y se puede escribir en la forma de Newton como

$$p_n(x) = f[x_0] + f[x_0, x_1] (x - x_0) + f[x_0, x_1, x_2] (x - x_0) (x - x_1) \\ + \cdots + f[x_0, x_1, \dots, x_n] (x - x_0) \cdots (x - x_{n-1}),$$

donde $f[x_0] \equiv f(x_0)$.

Ejemplo. Para dos puntos

$$p_0(x_0) = y_0, \quad p_1(x) = p_0(x) + C(x), \quad C = a_1 (x - x_0),$$

por lo que

$$a_1 = \frac{p_1(x_1) - p_0(x_1)}{x_1 - x_0} = \frac{y_1 - y_0}{x_1 - x_0} = f[x_0, x_1].$$

Fin de ejemplo.

Podemos relacionar la expresión de la diferencia dividida directamente con el polinomio de Lagrange. Recordando que

$$\varphi_n(x) = (x - x_0) \cdots (x - x_n) = \prod_{i=0}^n (x - x_i),$$

cuya derivada es

$$\frac{d\varphi_n(x)}{dx} = \varphi_n'(x) = \sum_{i=0}^n (x - x_0) \cdots (x - x_{i-1}) (x - x_{i+1}) \cdots (x - x_n),$$

por lo que

$$\varphi_n'(x_i) = (x_i - x_0) \cdots (x_i - x_{i-1}) (x_i - x_{i+1}) \cdots (x_i - x_n).$$

Así podemos escribir el polinomio de Newton (que coincide con el de Lagrange)

$$p_n(x) = p_{n-1}(x) + f[x_0, \dots, x_n] (x - x_0) \cdots (x - x_{n-1}),$$

de la forma

$$\sum_{i=0}^n f(x_i) \frac{\varphi_n(x)}{(x - x_i) \varphi'_n(x_i)} = \sum_{i=0}^{n-1} f(x_i) \frac{\varphi_{n-1}(x)}{(x - x_i) \varphi'_{n-1}(x_i)} + f[x_0, x_1, \dots, x_n] \varphi_{n-1}(x),$$

con lo que la diferencia dividida de orden n se escribe

$$f[x_0, x_1, \dots, x_n] = \sum_{i=0}^n f(x_i) \frac{x - x_n}{(x - x_i) \varphi'_n(x_i)} - \sum_{i=0}^{n-1} f(x_i) \frac{1}{(x - x_i) \varphi'_{n-1}(x_i)},$$

ya que $\varphi_n(x) = \varphi_{n-1}(x) (x - x_n)$, y como $\varphi'_n(x_i) = \varphi'_{n-1}(x_i) (x_i - x_n)$, obtenemos

$$f[x_0, x_1, \dots, x_n] = \sum_{i=0}^{n-1} \frac{f(x_i)}{(x - x_i) \varphi'_{n-1}(x_i)} \left(\frac{x - x_n}{x_i - x_n} - 1 \right) + \frac{f(x_n)}{\varphi'_n(x_n)},$$

y finalmente se tiene la siguiente expresión general para la diferencia dividida

$$f[x_0, x_1, \dots, x_n] = \sum_{i=0}^{n-1} \frac{f(x_i)}{(x_i - x_n) \varphi'_{n-1}(x_i)} + \frac{f(x_n)}{\varphi'_n(x_n)} = \sum_{i=0}^n \frac{f(x_i)}{\varphi'_n(x_i)}.$$

Sin embargo, las diferencias divididas se pueden obtener de una manera más rápida y sencilla utilizando una fórmula de recurrencia muy simple que vamos a derivar seguidamente. Para ello vamos a relacionar el polinomio interpolador de Newton (de grado a lo sumo n)

$$p_n(x) = f(x_0) + \sum_{i=1}^n f[x_0, x_1, \dots, x_i] (x - x_0) \cdots (x - x_{i-1}),$$

con dos polinomios $p_{n-1}(x)$ y $q(x)$ de grado a lo sumo $n - 1$, donde $p_{n-1}(x)$ es el polinomio interpolador de los n puntos $(x_0, y_0), \dots, (x_{n-1}, y_{n-1})$, es decir,

$$p_{n-1}(x) = f(x_0) + \sum_{i=1}^{n-1} f[x_0, x_1, \dots, x_i] (x - x_0) \cdots (x - x_{i-1}),$$

y $q(x)$ es el polinomio interpolador de los n puntos $(x_1, y_1), \dots, (x_n, y_n)$, sea

$$q(x) = f(x_1) + \sum_{i=2}^n f[x_1, x_2, \dots, x_i] (x - x_1) \cdots (x - x_{i-1}).$$

Podemos escribir el polinomio de Newton utilizando estos dos polinomios de la forma

$$p_n(x) = q(x) + \frac{x - x_n}{x_n - x_0} (q(x) - p_{n-1}(x)), \quad (6.1)$$

que se prueba fácilmente verificando las condiciones de interpolación

$$p_n(x_0) = q(x_0) + \frac{x_0 - x_n}{x_n - x_0} (q(x_0) - f(x_0)) = f(x_0),$$

$$p_n(x_i) = f(x_i) + \frac{x_i - x_n}{x_n - x_0} (f(x_i) - f(x_i)) = f(x_i),$$

$$p_n(x_n) = f(x_n) + \frac{x_n - x_n}{x_n - x_0} (f(x_n) - p_{n-1}(x_n)) = f(x_n).$$

Ahora bien, comparando el coeficiente de x^n en la expresión (6.1), con el obtenido con la expresión

$$p_n(x) = p_{n-1}(x) + f[x_0, \dots, x_n] (x - x_0) \cdots (x - x_{n-1}),$$

obtenemos directamente

$$f[x_0, \dots, x_n] = \frac{f[x_1, \dots, x_n] - f[x_0, \dots, x_{n-1}]}{x_n - x_0},$$

y en particular, definiendo $f[x_0] = f(x_0)$,

$$f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0},$$

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0},$$

y así sucesivamente. Para calcular las diferencias divididas se suele utilizar una tabla triangular como la siguiente

$$\begin{array}{l|lll} x_0 & f[x_0] & f[x_0, x_1] & f[x_0, x_1, x_2] & f[x_0, x_1, x_2, x_3] \\ x_1 & f[x_1] & f[x_1, x_2] & f[x_1, x_2, x_3] & \\ x_2 & f[x_2] & f[x_2, x_3] & & \\ x_3 & f[x_3] & & & \end{array}$$

en la que se calculan las cantidades de izquierda a derecha.

La fórmula del error para el polinomio interpolador de Newton es la misma que la del polinomio de Lagrange porque dichos polinomios son iguales.

Lema. Si $f \in \mathbf{C}^n$ en $I = [x_0, x_1, \dots, x_n]$, y los puntos x_i son distintos entre sí, entonces existe un punto $\xi \in I$ tal que

$$f[x_0, x_1, \dots, x_n] = \frac{f^{(n)}(\xi)}{n!}.$$

Demostración. Sea $p_{n-1}(x)$ el polinomio interpolador de $f(x)$ en los puntos $\{x_0, x_1, \dots, x_{n-1}\}$. Por la fórmula del error de Lagrange evaluada en $x = x_n$,

$$f(x_n) - p_{n-1}(x_n) = \frac{f^{(n)}(\xi)}{n!} \prod_{j=0}^{n-1} (x_n - x_j),$$

y por la definición de diferencia dividida,

$$f(x_n) - p_{n-1}(x_n) = f[x_0, x_1, \dots, x_n] \prod_{j=0}^{n-1} (x_n - x_j),$$

cqd.

Corolario. Si $f \in \mathbf{C}^n$ entonces

$$f[\overbrace{x_0, x_0, \dots, x_0}^{(n)}] = \frac{f^{(n)}(x_0)}{n!}.$$

6.2.5 Interpolación en tablas y diferencias finitas

Si tenemos una función dada por una tabla de valores $(x_i, y_i = f(x_i))$, podemos utilizar la expresión en tabla triangular que hemos presentado en la sección anterior. Sin embargo, si la tabla tiene valores equidistantes, es decir, si $h = x_{i+1} - x_i$ es constante, es mejor calcular las diferencias divididas utilizando diferencias finitas.

Definimos dos operadores en diferencias finitas, uno hacia adelante

$$\Delta f_i = \Delta^+ f(x_i) = f(x_{i+1}) - f(x_i),$$

y otro hacia atrás

$$\nabla f_i = \Delta^- f(x_i) = f(x_i) - f(x_{i-1}).$$

Teorema. Podemos calcular con diferencias hacia adelante

$$f[x_0, x_1, \dots, x_n] = \frac{1}{k! h^k} \Delta^k f(x_0).$$

Demostración. Lo probaremos por inducción,

$$k = 0, \quad f[x_0] = \frac{1}{0! h^0} f(x_0) = f(x_0),$$

$$k = 1, \quad f[x_0, x_1] = \frac{1}{1! h^1} (f(x_1) - f(x_0)) = \frac{f(x_1) - f(x_0)}{h},$$

y asumiendo que esta expresión es válida para $k = r$ como hipótesis de inducción,

$$\begin{aligned} f[x_0, x_1, \dots, x_{r+1}] &= \frac{f[x_1, \dots, x_{r+1}] - f[x_0, \dots, x_r]}{x_{r+1} - x_0} \\ &= \left(\frac{1}{r! h^r} \Delta^r f(x_1) - \frac{1}{r! h^r} \Delta^r f(x_0) \right) \frac{1}{(r+1)h} \\ &= \frac{1}{(r+1)! h^{r+1}} (\Delta^r f(x_1) - \Delta^r f(x_0)) \\ &= \frac{1}{(r+1)! h^{r+1}} \Delta^{r+1} f(x_0), \end{aligned}$$

cqd.

De esta forma se puede escribir el polinomio de interpolación utilizando diferencias finitas. Si tomamos

$$\mu = \frac{x - x_0}{h},$$

podemos escribir

$$x - x_j = x_0 + \mu h - x_j = x_0 + \mu h - (x_0 + j h) = (\mu - j) h,$$

con lo que

$$(x - x_0) \cdots (x - x_k) = \mu (\mu - 1) \cdots (\mu - k) h^{k+1},$$

y podemos escribir el polinomio interpolador

$$p_n(x) = f_0 + \mu h \frac{\Delta f_0}{h} + \mu (\mu - 1) h^2 \frac{\Delta^2 f_0}{2! h^2} + \cdots + \mu (\mu - 1) \cdots (\mu - n + 1) h^n \frac{\Delta^n f_0}{n! h^n},$$

y como los números binomiales

$$\binom{\mu}{k} = \frac{\mu!}{k! (\mu - k)!} = \frac{\mu (\mu - 1) \cdots (\mu - k + 1)}{k!},$$

obtenemos finalmente

$$p_n(x) = \sum_{j=0}^n \binom{\mu}{j} \Delta^j f(x_0).$$

Esta es la forma de Newton del polinomio interpolador utilizando diferencias hacia adelante.

Existen fórmulas análogas con diferencias hacia atrás. Se demuestra fácilmente que, por un lado,

$$f[x_0, x_1, \dots, x_n] = \frac{1}{k! h^k} \nabla^k f(x_n),$$

y por otro

$$p_n(x) = \sum_{j=0}^n (-1)^j \binom{\hat{\mu}}{j} \nabla^j f(x_0),$$

donde $\hat{\mu} = (x - x_n)/h$.

6.2.6 Errores en la interpolación en tablas (no entra en examen)

La interpolación mediante diferencias finitas tanto hacia adelante como hacia atrás se puede usar para detectar (y corregir) errores y ruido que aparecer en tablas de funciones.

Lema. Las diferencias finitas aproximan las derivadas, ya que existe $x_i \leq \xi_i \leq x_{i+r}$, tal que

$$\Delta^r f(x_i) = h^r f^{(r)}(\xi_i).$$

Demostración.

$$\Delta^r f(x_i) = r! h^r f[x_i, x_{i+1}, \dots, x_{i+r}] = r! h^r \frac{f^{(r)}(\xi_i)}{r!} = h^r f^{(r)}(\xi_i).$$

Lema. Linealidad del operador en diferencias

$$\Delta^r (\alpha f(x) + \beta g(x)) = \alpha \Delta^r f(x) + \beta \Delta^r g(x)$$

con α y β constantes.

Demostración. Lo demostraremos por inducción. El caso base es trivial

$$\Delta^0 (\alpha f(x) + \beta g(x)) = \alpha f(x) + \beta g(x).$$

Tomando como hipótesis de inducción que es cierto para r , tenemos

$$\begin{aligned} & \Delta^{r+1} (\alpha f(x) + \beta g(x)) \\ &= \Delta^r (\alpha f(x+h) + \beta g(x+h)) - \Delta^r (\alpha f(x) + \beta g(x)) \\ &= \alpha (\Delta^r f(x+h) - \Delta^r f(x)) + \beta (\Delta^r g(x+h) - \Delta^r g(x)) \\ &= \alpha \Delta^{r+1} f(x) + \beta \Delta^{r+1} g(x), \quad c.q.d. \end{aligned}$$

Asumamos ahora que $f(x_i) = \tilde{f}(x_i) + e_i$ donde $f(x_i)$ es el valor exacto, $\tilde{f}(x_i)$ es el valor obtenido interpolando en la tabla y e_i es el error de cometido, entonces

$$\Delta^r \tilde{f}_i = \Delta^r (f_i - e_i) = \Delta^r f_i - \Delta^r e_i = h^r f^{(r)}(\xi_i) - \Delta^r e_i,$$

y si asumimos ahora que $e(x_i) = \epsilon \delta_{ik}$, podemos ilustrar la propagación y crecimiento de los errores utilizando la siguiente tabla

x_i	$e(x_i)$	$\Delta e(x_i)$	$\Delta^2 e(x_i)$	$\Delta^3 e(x_i)$
x_{k-3}	0	0	0	ϵ
x_{k-2}	0	0	ϵ	-3ϵ
x_{k-1}	0	ϵ	-2ϵ	3ϵ
x_k	ϵ	$-\epsilon$	ϵ	$-\epsilon$
x_{k+1}	0	0	0	0.

Como $h^r f^{(r)}(\xi_i)$ disminuye normalmente conforme r aumenta, entonces $\Delta^r \tilde{f}_i$ también aumenta, por lo que se pueden estimar los errores

$$\Delta^r e_i = \Delta^r f_i - \Delta^r \tilde{f}_i \approx -\Delta^r \tilde{f}_i,$$

y utilizar estos resultados para corregir el valor interpolado. Veamos un ejemplo práctico,

\tilde{f}_i	$\Delta \tilde{f}_i$	$\Delta^2 \tilde{f}_i$	$\Delta^3 \tilde{f}_i$	$\Delta^r e_i$
0.10396	0.01700	-0.00014	-0.00003	0
0.12096	0.01686	-0.00017	-0.00002	0
0.13782	0.01669	-0.00019	0.00006	ϵ
0.15451	0.01650	-0.00013	-0.00025	-3ϵ
0.17101	0.01637	-0.00038	0.00021	3ϵ
0.18738	0.01559	-0.00017	-0.00010	$-\epsilon$
0.20337	0.01582	-0.00027		
0.21919	0.01555			
0.23474				

con lo que, para $r = 3$ podemos tomar

$$\epsilon = 0.00010, \quad f_i = \tilde{f}_i + e_i = 0.18738 + 0.00010 = 0.18748,$$

aunque es más preciso promediar en la forma

$$\epsilon = \frac{\epsilon - (-3\epsilon) + 3\epsilon - (-\epsilon)}{8} = \frac{0.00062}{8} \approx 0.000077,$$

con lo que $f_i = \tilde{f}_i + e_i = 0.18738 + 0.000077 = 0.187457$.

6.2.7 Interpolación osculatoria o de Hermite

La interpolación osculatoria intenta evitar el fenómeno de Runge, es decir, que el polinomio interpolador oscile demasiado por lo que interpola tanto los valores de la función como los valores de sus derivadas. Con ello el fenómeno de Runge no se evita completamente, pero sí se minimiza bastante en la práctica.

Por ejemplo, podemos interpolar los puntos (x_0, y_0) , (x_0, y'_0) y (x_1, y_1) , mediante un polinomio cuadrático (en forma de Newton)

$$\begin{aligned} p_2(x) &= f(x_0) + f[x_0, x_0](x - x_0) + f[x_0, x_0, x_1](x - x_0)^2 \\ &= y_0 + y'_0(x - x_0) + \frac{\frac{y_1 - y_0}{x_1 - x_0} - y'_0}{x_1 - x_0}(x - x_0)^2, \end{aligned}$$

donde hemos usado que

$$f[x_0, x_0] = \lim_{x_1 \rightarrow x_0} f[x_0, x_1] = \lim_{x_1 \rightarrow x_0} \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f'(x_0).$$

El tipo más simple es la interpolación de Hermite, que interpola la secuencia de puntos (x_i, y_i, y'_i) , $i = 1, \dots, n$. Esta secuencia nos da $2n$ condiciones y nos permite calcular un polinomio de grado $(2n - 1)$.

Como hicimos con la forma de Lagrange del polinomio interpolador, podemos definir una base nodal para los polinomios de Hermite, de forma que todo polinomio se puede escribir como

$$H_n(x) = \sum_{i=1}^n y_i h_i(x) + \sum_{i=1}^n y'_i \tilde{h}_i(x),$$

donde las funciones base $\{h_i(x), \tilde{h}_i(x)\}$ son tales que

$$h_i(x) \Rightarrow h_i(x_j) = \delta_{ij}, \quad h'_i(x_j) = 0,$$

$$\tilde{h}_i(x) \Rightarrow \tilde{h}_i(x_j) = 0, \quad \tilde{h}'_i(x_j) = \delta_{ij}.$$

Podemos calcular estas funciones base por inspección. Como los nodos $\{x_j\}$, con $j \neq i$, son ceros dobles, ya que son ceros de la función y su derivada, estas funciones serán el cuadrado ($l_i^2(x)$) del polinomio de Lagrange

$$l_i(x) \Rightarrow l_i(x_j) = \delta_{ij}, \quad l_i'(x_j) = \delta_{ij},$$

multiplicadas por otro polinomio. Como los polinomios Lagrange tienen grado $n - 1$ (hay n nodos), su cuadrado tiene grado $2n - 2$. El polinomio de Hermite, de grado $2n - 1$ requiere que multipliquemos éste por un polinomio de grado 1, que requiere determinar dos coeficientes.

$$h_i(x) = \alpha_i (x + \beta_i) l_i^2(x), \quad \tilde{h}_i(x) = \tilde{\alpha}_i (x + \tilde{\beta}_i) l_i^2(x).$$

De esta forma hemos cumplido con todas las condiciones de la interpolación de Hermite excepto dos, las que se cumplen para $j = i$. Aplicando estas condiciones podremos obtener un sistema de dos ecuaciones (lineales) con dos incógnitas. Hemos reducido el problema a uno muy fácil de resolver. Sin embargo, podemos simplificarlo aún más si volvemos a aplicar algo más de inspección.

En el caso de $\tilde{h}_i(x)$ las condiciones que nos faltan por cumplir son $\tilde{h}_i(x_i) = 0$ y $\tilde{h}_i'(x_i) = 1$, que se cumplen automáticamente para

$$\tilde{h}_i(x) = (x - x_i) l_i^2(x), \quad \tilde{h}_i(x_i) = 0,$$

ya que su derivada es

$$\tilde{h}_i'(x) = l_i^2(x) + 2(x - x_i) l_i(x), \quad \tilde{h}_i'(x_i) = 1.$$

En el caso de $h_i(x)$ nos faltan por cumplir las condiciones $h_i(x_i) = 1$ y $h_i'(x_i) = 0$, y podemos tantear el polinomio

$$h_i(x) = (1 - \alpha_i (x - x_i)) l_i^2(x), \quad h_i(x_i) = 1,$$

cuya derivada nos permitirá calcular α_i , dado que

$$h_i'(x) = (1 - \alpha_i (x - x_i)) 2 l_i(x) l_i'(x) - \alpha_i l_i^2(x),$$

y necesitamos que $h_i'(x_i) = 0$, obtenemos la ecuación

$$h_i'(x_i) = 2 l_i'(x_i) - \alpha_i = 0, \quad \alpha_i = 2 l_i'(x_i),$$

con lo que el polinomio (base de Hermite) buscado es

$$h_i(x) = (1 - 2 l_i'(x_i) (x - x_i)) l_i^2(x),$$

que junto a

$$\tilde{h}_i(x) = (x - x_i) l_i^2(x),$$

permiten formar una base nodal para desarrollar el polinomio interpolador de Hermite

$$H_n(x) = \sum_{i=1}^n y_i h_i(x) + \sum_{i=1}^n y'_i \tilde{h}_i(x).$$

Por construcción este polinomio de grado $(2n - 1)$ cumple con las condiciones

$$p(x_i) = y_i, \quad p'(x_i) = y'_i.$$

El polinomio de Hermite es único, ya que si existiera otro polinomio

$$G_n(x) = \sum_{i=1}^n y_i g_i(x) + \sum_{i=1}^n y'_i \tilde{g}_i(x),$$

que cumpliera que

$$G_n(x_i) = y_i, \quad G'_n(x_i) = y'_i,$$

entonces $R_n(x) = H_n(x) - G_n(x)$ sería un polinomio de grado $(2n - 1)$ tal que

$$R_n(x_i) = 0 = R'_n(x_i),$$

es decir, sería un polinomio cuyas raíces x_i son dobles,

$$R_n(x) = q(x) (x - x_1)^2 \cdots (x - x_n)^2,$$

que si $q(x) \not\equiv 0$ tiene un grado $\geq 2n$, lo que está en contradicción con el hecho de que tiene grado $(2n - 1)$ por construcción. Por tanto, $R_n(x) \equiv 0$, y el polinomio de Hermite es único.

Igual que con el polinomio de Lagrange podemos desarrollar el polinomio de Hermite en diferencias divididas (forma de Newton de dicho polinomio). Podemos representar cualquier polinomio, por ejemplo, el polinomio de Hermite, utilizando la forma de Newton con los centros $x_1, \dots, x_n, x_{n+1}, \dots, x_{2n-1}$, de la forma

$$\begin{aligned} H_n(x) = p_{2n-1}(x) &= f(x_1) + f[x_1, x_2] (x - x_1) + \cdots \\ &+ f[x_1, x_2, \dots, x_{2n}] (x - x_1) \cdots (x - x_{2n-1}), \end{aligned}$$

eligiendo adecuadamente los puntos x_{n+1}, \dots, x_{2n} . De esta forma obtenemos la siguiente forma para el error de interpolación (que ya estudiamos en una sección anterior)

$$f(x) - p_{2n-1}(x) = f[x_1, x_2, \dots, x_{2n}, x] (x - x_1) \cdots (x - x_{2n}).$$

Ahora podemos demostrar fácilmente que la forma de Newton del polinomio de Hermite se obtiene tomando los centros $x_{n+j} = x_j$, $j = 1, \dots, n$, es decir,

$$p_{2n-1}(x) = f(x_1) + f[x_1, x_1](x - x_1) + f[x_1, x_1, x_2](x - x_1)^2 + \dots \\ + f[x_1, x_1, \dots, x_n, x_n](x - x_1)^2 \cdots (x - x_{n-1})^2 (x - x_n),$$

que conduce a la siguiente fórmula de error

$$f(x) - p_{2n-1}(x) = f[x_1, x_1, \dots, x_n, x_n, x](x - x_1)^2 \cdots (x - x_n)^2.$$

Es fácil comprobar que $p_{2n-1}(x_i) = f(x_i)$ y, además, derivando esta expresión obtenemos

$$f'(x) - p'_{2n-1}(x) = (x - x_1)^2 \cdots (x - x_n)^2 \frac{d}{dx} f[x_1, x_1, \dots, x_n, x_n, x] \\ + 2 f[x_1, x_1, \dots, x_n, x_n, x] \sum_{i=1}^n \left((x - x_i) \prod_{i \neq j=1}^n (x - x_j)^2 \right),$$

que nos permite demostrar comprobar que también se cumple que

$$f'(x_i) - p'_{2n-1}(x_i) = 0,$$

con lo que hemos demostrado que $p_{2n-1}(x) \equiv H_n(x)$ es el polinomio de Hermite (que ya vimos que era único). Más aún, para todo $x \in [x_1, x_2, \dots, x_n]$ existe un $\xi \in [x_1, x_2, \dots, x_n]$ tal que

$$f(x) - p_{2n-1}(x) = f[x_1, x_1, \dots, x_n, x_n, x] \prod_{i=1}^n (x - x_i)^2 = \prod_{i=1}^n (x - x_i)^2 \frac{f^{(2n)}(\xi)}{(2n)!}.$$

Ejemplo. Calcularemos el polinomio de interpolación cúbica de Hermite, que cumple las condiciones

$$y(a) = f(a), \quad y(b) = f(b), \quad y'(a) = f'(a), \quad y'(b) = f'(b),$$

y que se escribe

$$p_3(x) = f(a) + f[a, a](x - a) + f[a, a, b](x - a)^2 + f[a, a, b, b](x - a)^2 (x - b),$$

donde

$$f[a, a] = f'(a), \quad f[b, b] = f'(b), \\ f[a, a, b] = \frac{f[a, b] - f'(a)}{b - a} = \frac{f(b) - f(a)}{(b - a)^2} - \frac{f'(a)}{b - a}, \\ f[a, b, b] = \frac{f'(b) - f[a, b]}{b - a} = \frac{f'(b)}{b - a} - \frac{f(b) - f(a)}{(b - a)^2},$$

$$f[a, a, b, b] = \frac{f[a, b, b] - f[a, a, b]}{b - a} = \frac{f'(b) + f'(a)}{(b - a)^2} - 2 \frac{f(b) - f(a)}{(b - a)^3},$$

y cuyo error de interpolación toma la forma

$$f(x) - p_3(x) = f[a, a, b, b, x] (x - a)^2 (x - b)^2 = \frac{f^{(4)}(\xi)}{4!} (x - a)^2 (x - b)^2,$$

donde $x, \xi \in [a, b]$. Además, si tomamos $x = (a + b)/2$,

$$\psi(x) = (x - a)^2 (x - b)^2, \quad \psi\left(\frac{a + b}{2}\right) = \frac{(b - a)^4}{16},$$

y por tanto el error se puede acotar como

$$\|f(x) - p_3(x)\|_\infty \leq \frac{(b - a)^4}{384} \|f^{(4)}(x)\|_\infty, \quad \|f^{(4)}(x)\|_\infty = \max_{x \in [a, b]} |f^{(4)}(x)|.$$

6.2.8 Estimación de las derivadas de Bessel

A la hora de utilizar la interpolación de Hermite es necesario conocer los valores de la derivada de la función. Sin embargo, también se puede obtener una estimación de dichos valores a partir de los valores nodales de la función. Suponiendo que ésta es suficientemente diferenciable, podemos utilizar el siguiente procedimiento debido a Bessel. A partir de las expresiones

$$f_{i+1} = f_i + f'_i h_i + f''_i \frac{h_i^2}{2} + f'''_i \frac{h_i^3}{3!} + O(h_i^4),$$

$$f_{i-1} = f_i - f'_i h_i + f''_i \frac{h_i^2}{2} - f'''_i \frac{h_i^3}{3!} + O(h_i^4),$$

donde $h_i = x_{i+1} - x_i$, y $f_i = f(x_i)$. Pasando todos los valores de la función al lado izquierdo de la igualdad y dejando las derivadas en el derecho podemos obtener

$$(f_{i+1} - f_i) h_{i-1}^2 = f'_i h_i h_{i-1}^2 + f''_i \frac{h_i^2 h_{i-1}^2}{2} + O(h_i^3 h_{i-1}^2),$$

$$(f_{i-1} - f_i) h_i^2 = -f'_i h_i^2 h_{i-1} + f''_i \frac{h_i^2 h_{i-1}^2}{2} + O(h_i^2 h_{i-1}^3),$$

que restando nos permite eliminar la segunda derivada, conduciendo a

$$(f_{i+1} - f_i) h_{i-1}^2 - (f_{i-1} - f_i) h_i^2 = f'_i h_i h_{i-1} (h_i + h_{i-1}) + O(h^5),$$

donde $h = \max\{h_i, h_{i-1}\}$, con lo que obtenemos la siguiente aproximación para las derivadas en los nodos

$$f'_i = \frac{f[x_i, x_{i+1}] h_{i-1} + f[x_{i-1}, x_i] h_i}{h_i + h_{i-1}} + O(h^2),$$

que es una aproximación de segundo orden y que es suficientemente precisa para ser aplicada a la interpolación de Hermite. De esta manera evitamos la necesidad de conocer, o calcular, las derivadas de la función a interpolar. El alumno avisado notará que tenemos un problema en el contorno, donde las fórmulas que hemos obtenido no son adecuadas. De la misma forma se pueden derivar fórmulas especiales para la evaluación en el contorno, dichas fórmulas serán estudiadas en el capítulo de derivación numérica.

6.3 Interpolación polinómica trozos

El problema fundamental de la interpolación polinómica global es que los polinomios que se obtienen oscilan excesivamente entre los nodos de interpolación conforme el grado del polinomio crece (fenómeno de Runge). Para evitar esto se pueden utilizar técnicas de interpolación polinómica a trozos, en las que se divide el conjunto de nodos de interpolación en subconjuntos disjuntos que se interpolan con un polinomio de grado bajo, con lo que la función interpolante está definida mediante trozos. Normalmente se imponen condiciones de continuidad del interpolante y de algunas de sus derivadas. De esta forma se define una función polinómica a trozos.

Este tipo de funciones se utilizan mucho en la teoría de la aproximación, en gráficos por ordenador, en el ajuste de datos, como base para las técnicas de integración y derivación de funciones y para la resolución numérica de ecuaciones diferenciales e integrales.

Las técnicas de interpolación polinómica a trozos se pueden clasificar según diferentes criterios. En función de si el cálculo de los coeficientes de los polinomios a trozos es explícito o implícito. Las técnicas de elementos finitos son explícitas, mientras que las esplines cúbicas son implícitas. También en función de si un cambio en un coeficiente afecta a toda la curva (técnicas no locales) o sólo a una parte de la misma (técnicas locales). Las esplines cúbicas son técnicas no locales, y las técnicas de elementos finitos, las esplines de Bézier, y las esplines básicas o B-esplines (*basis-splines*), son locales.

En este curso estudiaremos sólo las técnicas de interpolación explícita mediante polinomios a trozos, también llamadas de elementos finitos, y las técnicas de interpolación implícita mediante polinomios a trozos de tipo esplines, con énfasis en las esplines cúbicas.

6.3.1 Interpolación nodal por polinomios a trozos discontinuos

Podemos interpretar el espacio $\mathcal{P}^{(n)}(a, b)$ de los polinomios de grado a lo sumo n definidos en el intervalo (a, b) como un espacio vectorial respecto al cuerpo de los reales (como es fácil de

demostrar). La dimensión de este espacio es $n + 1$. Podemos escribir diferentes bases para este espacio. La base canónica

$$\{1, t, t^2, \dots, t^n\},$$

para la que $p(x) \in \mathcal{P}^{(n)}(a, b)$ se escribe

$$p(x) = \sum_{i=0}^n a_i t^i = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n,$$

la base de Taylor respecto al punto $x = a$,

$$\{1, t - a, (t - a)^2, \dots, (t - a)^n\},$$

para la que $p(x) \in \mathcal{P}^{(n)}(a, b)$ se escribe

$$p(x) = \sum_{i=0}^n \alpha_i (t - a)^i = p(a) + p'(a)(x - a) + \frac{p''(a)}{2}(x - a)^2 + \dots + \frac{p^{(n)}(a)}{n!}(x - a)^n,$$

la base nodal o de Lagrange sobre el conjunto de $n + 1$ nodos $\{x_j\}$, con $j = 0, 1, \dots, n$,

$$\{l_0(x), l_1(x), \dots, l_n(x)\},$$

para la que $p(x) \in \mathcal{P}^{(n)}(a, b)$ se escribe

$$p(x) = \sum_{i=0}^n f(x_i) l_i(x) = f(x_0) l_0(x) + f(x_1) l_1(x) + \dots + f(x_n) l_n(x),$$

o la base de Newton sobre los nodos $\{x_j\}$,

$$\{1, (x - x_0), (x - x_0)(x - x_1), \dots, (x - x_0)(x - x_1) \cdots (x - x_{n-1})\},$$

para la que $p(x) \in \mathcal{P}^{(n)}(a, b)$ se escribe

$$p(x) = \sum_{i=0}^n \beta_i \prod_{j=0}^{i-1} (x - x_j) = f[x_0] + f[x_0, x_1](x - x_0) + \dots + f[x_0, x_1, \dots, x_n](x - x_0) \cdots (x - x_{n-1}).$$

El alumno notará que el espacio vectorial de los polinomios es un espacio normado completo (espacio de Banach) con cualquier norma- q , sean

$$\|p(x)\|_q = \left(\int_a^b |p(x)|^q dx \right)^{1/q},$$

donde completo significa que todas las sucesiones de Cauchy de polinomios son convergentes. Además es un espacio con producto interior completo (espacio de Hilbert) con el producto interior euclídeo,

$$\langle p_1(x), p_2(x) \rangle = \left(\int_a^b p_1(x) p_2(x) dx \right)^{1/2},$$

que induce, como norma asociada, la norma-2 o euclídea,

$$\|p(x)\|_2 = \left(\int_a^b |p(x)|^2 dx \right)^{1/2}.$$

En un espacio de Hilbert todas sus bases se pueden ortogonalizar, por el procedimiento de Gram-Schmidt, y ortonormalizar. En esta y las próximas secciones no utilizaremos estas propiedades del espacio vectorial de polinomios directamente.

Podemos definir, de igual forma, el espacio vectorial de los polinomios a trozos $W_{\mathcal{T}}^{(n)}$. En general, estos polinomios a trozos serán discontinuos en los nodos. Definamos primero una malla \mathcal{T} adecuada del intervalo (a, b) . Sea ésta formada por el conjunto de $N + 1$ nodos $\{x_j\}$ dados por

$$a = x_0 < x_1 < x_2 < \cdots < x_N = b,$$

y donde denotaremos por $I_j = [x_{j-1}, x_j)$, para $j = 1, 2, \dots, N - 1$, e $I_N = [x_{N-1}, x_N]$.

Un polinomio a trozos discontinuos de grado a lo sumo n , sea $v(x) \in W_{\mathcal{T}}^{(n)}$, cumple que $v(x) \in \mathcal{P}^{(n)}(I_j)$, para $x \in I_j$, lo que denotaremos normalmente como $v|_{I_j}(x) \equiv v_j(x) \in \mathcal{P}^{(n)}(I_j)$. Y por tanto, $v(x) = v_j(x)$, para todo $x \in I_j$. En general, los valores $v_j(x_j) \neq v_{j+1}(x_j)$, es decir, el polinomio a trozos $v(x)$ es discontinuo en los nodos. Denotaremos los dos valores en la discontinuidad de la forma

$$v(x_j^+) \equiv v(x_j + 0) = v_{j+1}(x_j), \quad v(x_j^-) \equiv v(x_j - 0) = v_j(x_j),$$

donde los signos $+$ y $-$ denotan los valores por la derecha y por la izquierda, respectivamente, de la función en un nodo. En la figura ? aparece un polinomio a trozos lineales discontinuo.

Para construir una base del espacio $W_{\mathcal{T}}^{(n)}$ partiremos de una base nodal del espacio $\mathcal{P}^{(n)}(I_j)$, cuya dimensión es $n + 1$, y por tanto requiere $n + 1$ nodos en el intervalo I_j , sean $x_{jk} \in I_j$, con $k = 0, 1, \dots, n$ y donde

$$x_{j-1} \equiv x_{j0} < x_{j1} < \cdots < x_{jn} \equiv x_j.$$

Normalmente estos nodos se tomarán equi-espaciados, es decir,

$$x_{jk} = x_{j-1} + k \frac{(x_j - x_{j-1})}{n}.$$

Para obtener el polinomio (interpolador) nodal de la función $f(x)$ en el intervalo I_j será necesario conocer todos los valores (nodales) $f(x_{jk})$. Dicho polinomio tomará la forma

$$v|_{I_j}(x) \equiv v_j(x) = \sum_{k=0}^n f(x_{jk}) l_{jk}(x),$$

donde se han utilizado las bases (nodales) de Lagrange

$$l_{jk}(x) = \frac{\prod_{i=0, k \neq i}^n (x - x_{ji})}{\prod_{i=0, k \neq i}^n (x_{jk} - x_{ji})}.$$

Obtendremos una base del espacio $W_{\mathcal{T}}^{(n)}$ si prolongamos con cero las funciones base (nodales) de Lagrange a todo el intervalo (a, b) , sean éstas

$$\psi_{jk}(x) = \begin{cases} l_{jk}(x), & x \in I_j, \\ 0, & x \notin I_j. \end{cases}$$

De esta forma el conjunto de $(n+1)N$ funciones $\{\psi_{jk}(x)\}$, con $j = 1, 2, \dots, N$ y $k = 0, 1, \dots, n$, forman una base del espacio $W_{\mathcal{T}}^{(n)}$ de polinomios a trozos discontinuos de grado a lo sumo n . Una función de dicho espacio se escribe como

$$v(x) = \sum_{j=1}^N \sum_{k=0}^n f(x_{jk}) \psi_{jk}(x), \quad v(x) \in W_{\mathcal{T}}^{(n)}.$$

Por ejemplo, el espacio $W_{\mathcal{T}}^{(0)}$ de polinomios a trozos constantes en la malla $\mathcal{T} \equiv \{x_j\}$, con $0 \leq j \leq N$, tiene como base las funciones $\{\psi_{j0}(x)\}$, con $0 \leq j \leq N$, definidas de la forma

$$\psi_{j0}(x) = \begin{cases} 1, & x \in I_j, \\ 0, & x \notin I_j, \end{cases}$$

y toda función $v(x) \in W_{\mathcal{T}}^{(0)}$ se escribe como

$$v(x) = \sum_{j=1}^N f(x_{j-1}) \psi_{j0}(x).$$

Note que se utiliza el valor $f(x_{j-1})$ dado que el punto $x_{j-1} \in I_j$ pero $x_j \notin I_j$ (para $j < N$). En la figura ? se ve una función de este espacio.

El espacio $W_{\mathcal{T}}^{(1)}$ de polinomios a trozos lineales discontinuos en la malla $\mathcal{T} \equiv \{x_j\}$, con $0 \leq j \leq N$, tiene como base las funciones $\{\psi_{j0}(x), \psi_{j1}(x)\}$, con $0 \leq j \leq N$, definidas de la forma

$$\psi_{j0}(x) = \begin{cases} \frac{x - x_j}{x_{j-1} - x_j}, & x \in I_j, \\ 0, & x \notin I_j, \end{cases}$$

$$\psi_{j1}(x) = \begin{cases} \frac{x - x_{j-1}}{x_j - x_{j-1}}, & x \in I_j, \\ 0, & x \notin I_j, \end{cases}$$

y toda función $v(x) \in W_{\mathcal{T}}^{(1)}$ se escribe como

$$v(x) = \sum_{j=1}^N f(x_{j-1}^+) \psi_{j0}(x) + f(x_j^-) \psi_{j1}(x).$$

Como vemos se necesitan conocer $2N$ valores nodales de la función, dos valores para cada nodo interno $f(x_j^-)$ y $f(x_j^+)$, $0 < j < N$, y un valor para los extremos $f(x_0^+)$ y $f(x_N^-)$. En la figura ? se muestran las funciones base de este espacio así como una función del mismo.

El espacio $W_{\mathcal{T}}^{(2)}$ de polinomios a trozos cuadráticos discontinuos en la malla $\mathcal{T} \equiv \{x_j\}$, con $0 \leq j \leq N$ requiere la introducción de los nodos intermedios $\{x_{jk}\}$, $1 \leq j \leq N$ y $0 \leq k \leq 2$, tales que

$$x_{j0} = x_{j-1}, \quad x_{j1} = \hat{x}_j = \frac{x_{j-1} + x_j}{2}, \quad x_{j2} = x_j,$$

y tiene como base las funciones $\{\psi_{j0}(x), \psi_{j1}(x), \psi_{j2}(x)\}$, con $1 \leq j \leq N$, definidas de la forma

$$\psi_{j0}(x) = \begin{cases} \frac{(x - x_j)(x - \hat{x}_j)}{(x_{j-1} - x_j)(x_{j-1} - \hat{x}_j)}, & x \in I_j, \\ 0, & x \notin I_j, \end{cases}$$

$$\psi_{j1}(x) = \begin{cases} \frac{(x - x_{j-1})(x - x_j)}{(\hat{x}_j - x_{j-1})(\hat{x}_j - x_j)}, & x \in I_j, \\ 0, & x \notin I_j, \end{cases}$$

$$\psi_{j2}(x) = \begin{cases} \frac{(x - x_{j-1})(x - \hat{x}_j)}{(x_j - x_{j-1})(x_j - \hat{x}_j)}, & x \in I_j, \\ 0, & x \notin I_j, \end{cases}$$

y toda función $v(x) \in W_{\mathcal{T}}^{(2)}$ se escribe como

$$v(x) = \sum_{j=1}^N f(x_{j-1}^+) \psi_{j0}(x) + f(\hat{x}_j) \psi_{j1}(x) + f(x_j^-) \psi_{j2}(x).$$

Como vemos se necesitan conocer $3N$ valores nodales de la función. En la figura ? se muestran las funciones base de este espacio así como una función del mismo.

El espacio $W_{\mathcal{T}}^{(3)}$ de polinomios a trozos cúbicos discontinuos en la malla $\mathcal{T} \equiv \{x_j\}$, con $0 \leq j \leq N$ requiere la introducción de los nodos intermedios $\{x_{jk}\}$, $1 \leq j \leq N$ y $0 \leq k \leq 3$,

tales que

$$x_{j0} = x_{j-1}, \quad x_{j1} = \frac{2x_{j-1} + x_j}{3}, \quad x_{j2} = \frac{x_{j-1} + 2x_j}{3}, \quad x_{j3} = x_j,$$

tiene como base las funciones $\{\psi_{jk}(x)\}$, con $1 \leq j \leq N$ y $0 \leq k \leq 3$, y toda función $v(x) \in W_{\mathcal{T}}^{(3)}$ se escribe como

$$v(x) = \sum_{j=1}^N f(x_{j-1}^+) \psi_{j0}(x) + f(x_{j1}) \psi_{j1}(x) + f(x_{j2}) \psi_{j2}(x) + f(x_j^-) \psi_{j3}(x).$$

Como vemos se necesitan conocer $4N$ valores nodales de la función. En la figura ? se muestran las funciones base de este espacio así como una función del mismo.

6.3.2 Interpolación nodal por polinomios a trozos continuos

El espacio de polinomios a trozos continuos de grado a lo sumo $n > 0$, sea $V_{\mathcal{T}}^{(n)}$, es el subespacio vectorial de $W_{\mathcal{T}}^{(n)}$, denotado $V_{\mathcal{T}}^{(n)} \prec W_{\mathcal{T}}^{(n)}$, definido como

$$V_{\mathcal{T}}^{(n)} = \{v(x) \in W_{\mathcal{T}}^{(n)} : v(x) \in \mathbf{C}^0[a, b]\} = \{v(x) \in \mathbf{C}^0[a, b] : v|_{I_j}(x) \in \mathcal{P}^{(n)}(I_j)\}.$$

Como es obvio, $\dim\{V_{\mathcal{T}}^{(n)}\} < \dim\{W_{\mathcal{T}}^{(n)}\}$. Para una función $v(x) \in V_{\mathcal{T}}^{(n)}$ se cumplirá siempre que

$$v(x_j^+) = v(x_j^-) = v(x_j), \quad v(x_j) = v_{j+1}(x_j) = v_j(x_j),$$

Dada una base $\{\psi_{jk}(x)\}$ con $j = 1, 2, \dots, N$ y $k = 0, 1, \dots, n$, del espacio $W_{\mathcal{T}}^{(n)}$ podremos construir una base de su subespacio vectorial $V_{\mathcal{T}}^{(n)}$ buscando combinaciones lineales de estos vectores base que den funciones continuas. Obviamente, las funciones $\{\psi_{jk}(x)\}$ con $1 \leq j \leq N$ y $0 < k < n$, son continuas, luego formarán parte de la base de $V_{\mathcal{T}}^{(n)}$. También son continuas las funciones $\{\psi_{10}(x), \psi_{Nn}(x)\}$. Sin embargo, las funciones $\{\psi_{j0}(x)\}$, $1 < j \leq N$, y $\{\psi_{jn}(x)\}$, $1 \leq j < N$, son discontinuas en los nodos x_{j-1} y x_j , respectivamente. La combinación lineal más simple de estas funciones que es continua es la siguiente

$$\varphi_{j0}(x) = \begin{cases} \psi_{jn}(x), & x \in I_j, \\ \psi_{(j+1)0}(x), & x \in I_{j+1}, \\ 0, & x \notin I_j \cup I_{j+1}, \end{cases}$$

donde $1 < j < N$. Estas funciones junto con

$$\varphi_{00}(x) = \psi_{10}(x), \quad \varphi_{N0}(x) = \psi_{Nn}(x),$$

y

$$\varphi_{jk}(x) = \psi_{jk}(x), \quad j = 1, 2, \dots, N, \quad k = 1, 2, \dots, n-1,$$

constituyen una base $\{\varphi_{jk}(x)\}$, con $1 \leq j \leq N$ y $0 \leq k < n$, de $V_{\mathcal{T}}^{(n)}$. Por tanto, la dimensión de este espacio es $N + 1 + (n-1)N \equiv nN + 1$ y toda función de dicho espacio se escribe como

$$v(x) = \sum_{j=0}^N f(x_j) \varphi_{j0}(x) + \sum_{j=1}^N \sum_{k=1}^{n-1} f(x_{jk}) \varphi_{jk}(x), \quad v(x) \in V_{\mathcal{T}}^{(n)}.$$

Por ejemplo, el espacio $V_{\mathcal{T}}^{(1)}$ de polinomios a trozos lineales continuos en la malla $\mathcal{T} \equiv \{x_j\}$, con $0 \leq j \leq N$, tiene como base las funciones $\{\varphi_{j0}(x)\}$, con $0 \leq j \leq N$, definidas de la forma

$$\varphi_{00}(x) = \begin{cases} \frac{x - x_1}{x_0 - x_1}, & x \in I_1, \\ 0, & x \notin I_1, \end{cases}$$

$$\psi_{j0}(x) = \begin{cases} \psi_{jn}(x) \equiv \frac{x - x_{j-1}}{x_j - x_{j-1}}, & x \in I_j, \\ \psi_{(j+1)0}(x) \equiv \frac{x - x_{j+1}}{x_j - x_{j+1}}, & x \in I_{j+1}, \\ 0, & x \notin I_j, \end{cases}$$

$$\psi_{N0}(x) = \begin{cases} \frac{x - x_{N-1}}{x_N - x_{N-1}}, & x \in I_N, \\ 0, & x \notin I_N, \end{cases}$$

y toda función $v(x) \in V_{\mathcal{T}}^{(1)}$ se escribe como

$$v(x) = \sum_{j=0}^N f(x_j) \varphi_{j0}(x).$$

Como vemos se necesitan conocer $(N + 1)$ valores nodales de la función. En la figura ? se muestran las funciones base de este espacio así como una función del mismo.

El espacio $V_{\mathcal{T}}^{(2)}$ de polinomios a trozos cuadráticos discontinuos en la malla $\mathcal{T} \equiv \{x_j\}$, con $0 \leq j \leq N$ requiere la introducción de los nodos intermedios $\{\hat{x}_j\}$, $1 \leq j \leq N$, tales que

$$\hat{x}_j = \frac{x_{j-1} + x_j}{2},$$

y tiene como base las funciones $\{\varphi_{j0}(x)\}$, con $0 \leq j \leq N$, y $\{\varphi_{j1}(x)\}$, con $1 \leq j \leq N$, definidas de la forma

$$\varphi_{00}(x) = \begin{cases} \frac{(x - x_1)(x - \hat{x}_1)}{(x_0 - x_1)(x_0 - \hat{x}_1)}, & x \in I_1, \\ 0, & x \notin I_1, \end{cases}$$

$$\varphi_{j0}(x) = \begin{cases} \frac{(x - x_{j-1})(x - \hat{x}_j)}{(x_j - x_{j-1})(x_j - \hat{x}_j)}, & x \in I_j, \\ \frac{(x - x_{j+1})(x - \hat{x}_{j+1})}{(x_j - x_{j+1})(x_j - \hat{x}_{j+1})}, & x \in I_{j+1}, \\ 0, & x \notin I_j \cup I_{j+1}, \end{cases} \quad 1 \leq j \leq N-1,$$

$$\psi_{N0}(x) = \begin{cases} \frac{(x - x_{N-1})(x - \hat{x}_N)}{(x_N - x_{N-1})(x_N - \hat{x}_N)}, & x \in I_N, \\ 0, & x \notin I_N, \end{cases}$$

$$\varphi_{j1}(x) = \begin{cases} \frac{(x - x_{j-1})(x - x_j)}{(\hat{x}_j - x_{j-1})(\hat{x}_j - x_j)}, & x \in I_j, \\ 0, & x \notin I_j, \end{cases} \quad 1 \leq j \leq N,$$

y toda función $v(x) \in V_{\mathcal{T}}^{(2)}$ se escribe como

$$v(x) = \sum_{j=0}^N f(x_j) \varphi_{j0}(x) + \sum_{j=1}^N f(\hat{x}_j) \varphi_{j1}(x).$$

Como vemos se necesitan conocer $(2N + 1)$ valores nodales de la función. En la figura ? se muestran las funciones base de este espacio así como una función del mismo.

El espacio $V_{\mathcal{T}}^{(3)}$ de polinomios a trozos cúbicos continuos en la malla $\mathcal{T} \equiv \{x_j\}$, con $0 \leq j \leq N$ requiere la introducción de los nodos intermedios $\{x_{jk}\}$, $1 \leq j \leq N$ y $1 \leq k \leq 2$, tales que

$$x_{j1} = \frac{2x_{j-1} + x_j}{3}, \quad x_{j2} = \frac{x_{j-1} + 2x_j}{3},$$

y tiene como base las funciones $\{\varphi_{jk}(x)\}$, $0 \leq j \leq N$, y $\{\varphi_{jk}(x)\}$, $1 \leq j \leq N$ y $1 \leq k \leq 2$, con lo que toda función $v(x) \in V_{\mathcal{T}}^{(3)}$ se escribe como

$$v(x) = \sum_{j=0}^N f(x_j) \psi_{j0}(x) + \sum_{j=1}^N f(x_{j1}) \psi_{j1}(x) + \sum_{j=1}^N f(x_{j2}) \psi_{j2}(x).$$

Como vemos se necesitan conocer $(3N + 1)$ valores nodales de la función. En la figura ? se muestran las funciones base de este espacio así como una función del mismo.

6.3.3 Interpolación nodal por polinomios a trozos con derivada continua

El espacio de polinomios a trozos continuos y con derivada continua de grado a lo sumo $m \geq 3$ e impar, sea $m = 2n + 1$ con $n > 0$, denotado $V_{H,\mathcal{T}}^{(m)}$, es el subespacio vectorial de $V_{\mathcal{T}}^{(m)}$, definido como

$$V_{H,\mathcal{T}}^{(m)} = \{v(x) \in V_{\mathcal{T}}^{(m)} : v(x) \in \mathbf{C}^1(a, b)\} = \{v(x) \in \mathbf{C}^1(a, b) : v|_{I_j}(x) \in \mathcal{P}^{(m)}(I_j)\}.$$

Estos polinomios a trozos cumplirán

$$v(x_j) = v_{j+1}(x_j) = v_j(x_j), \quad v'(x_j) = v'_{j+1}(x_j) = v'_j(x_j).$$

Aunque podemos obtener una base del espacio $V_{H,T}^{(m)}$ formando combinaciones lineales adecuadas de los funciones base $\{\varphi_{jk}(x)\}$ del espacio $V_T^{(m)}$, este proceso es bastante complicado y requiere una selección adecuada de los nodos interiores a cada intervalo I_j que dependen de la derivada de la función vía la aplicación del teorema del valor medio. Es mucho más fácil utilizar las funciones base (nodales) de la interpolación de Hermite, por eso hemos requerido que $m - 1$ sea un número par, sea $m - 1 = 2n$.

Para construir una base del espacio $V_{H,T}^{(m)}$ partiremos de una base nodal de Hermite del espacio $\mathcal{P}^{(m)}(I_j)$, cuya dimensión es $m - 1 = 2n$, y por tanto requiere $n + 1$ nodos en el intervalo I_j , sean $x_{jk} \in I_j$, con $k = 0, 1, \dots, n$ y donde

$$x_{j-1} \equiv x_{j0} < x_{j1} < \dots < x_{jn} \equiv x_j.$$

Normalmente estos nodos se tomarán equi-espaciados, es decir,

$$x_{jk} = x_{j-1} + k \frac{(x_j - x_{j-1})}{n}.$$

Para obtener el polinomio (interpolador) nodal de Hermite de la función $f(x)$ en el intervalo I_j será necesario conocer todos los valores (nodales) $f(x_{jk})$ y $f'(x_{jk})$. Dicho polinomio tomará la forma

$$v|_{I_j}(x) \equiv v_j(x) = \sum_{k=0}^n y_i f(x_{jk}) h_{jk}(x) + \sum_{k=0}^n y_i f'(x_{jk}) \tilde{h}_{jk}(x),$$

donde se han utilizado las bases (nodales) de Hermite

$$h_{jk}(x) = (1 - 2l'_{jk}(x_{jk})(x - x_{jk})) l_{jk}^2(x),$$

$$\tilde{h}_{jk}(x) = (x - x_{jk}) l_{jk}^2(x),$$

donde las bases (nodales) de Lagrange son

$$l_{jk}(x) = \frac{\prod_{i=0, k \neq i}^n (x - x_{ji})}{\prod_{i=0, k \neq i}^n (x_{jk} - x_{ji})}.$$

Para obtener una base del espacio $V_{H,T}^{(m)}$, primero prolongaremos a todo el intervalo de interpolación las funciones base de hermite, para formar las funciones

$$\psi_{H,jk}(x) = \begin{cases} h_{jk}(x), & x \in I_j, \\ 0, & x \notin I_j, \end{cases}$$

$$\tilde{\psi}_{H,jk}(x) = \begin{cases} \tilde{h}_{jk}(x), & x \in I_j, \\ 0, & x \notin I_j, \end{cases}$$

donde $j = 0, 1, \dots, N$ y $k = 0, 1, \dots, n$.

Las funciones $\psi_{H,jk}(x)$ y $\tilde{\psi}_{H,jk}(x)$, para $0 < k < n$, son continuas y con derivada continua, por lo que formarán parte de la base del espacio $V_{H,T}^{(m)}$ y las llamaremos $\varphi_{H,jk}(x)$ y $\tilde{\varphi}_{H,jk}(x)$, respectivamente. Sin embargo, las funciones $\psi_{H,j0}(x)$ son discontinuas para $0 < j \leq n$, y las funciones $\psi_{H,jn}(x)$, lo son para $0 \leq j < n$, ya que

$$\psi_{H,j0}(x_{j-1}) = 1 = \psi_{H,jn}(x_j), \quad 0 < j < n,$$

aunque como

$$\psi'_{H,j0}(x_{j-1}) = 0 = \psi'_{H,jn}(x_j), \quad 0 < j < n,$$

su combinación lineal

$$\varphi_{H,j0}(x) = \begin{cases} \psi_{H,jn}(x), & x \in I_j, \\ \psi_{H,(j+1)0}(x), & x \in I_{j+1}, \\ 0, & x \notin I_j \cup I_{j+1}, \end{cases} \quad 1 < j < N,$$

es continua y con derivada continua, luego formará parte de la base de $V_{H,T}^{(m)}$ junto con

$$\varphi_{H,00}(x) = \psi_{10}(x), \quad \varphi_{H,N0}(x) = \psi_{Nn}(x).$$

Por otro lado, las funciones $\tilde{\psi}_{H,j0}(x)$ son continuas pero con derivada discontinua para $0 < j \leq n$, y las funciones $\tilde{\psi}_{H,jn}(x)$, lo son para $0 \leq j < n$, ya que

$$\tilde{\psi}_{H,j0}(x_{j-1}) = 0 = \tilde{\psi}_{H,jn}(x_j), \quad 0 < j < n,$$

aunque como

$$\tilde{\psi}'_{H,j0}(x_{j-1}) = 1 = \tilde{\psi}'_{H,jn}(x_j), \quad 0 < j < n,$$

luego su combinación lineal

$$\tilde{\varphi}_{H,j0}(x) = \begin{cases} \tilde{\psi}_{H,jn}(x), & x \in I_j, \\ \tilde{\psi}_{H,(j+1)0}(x), & x \in I_{j+1}, \\ 0, & x \notin I_j \cup I_{j+1}, \end{cases} \quad 1 < j < N,$$

es continua y con derivada continua, y formará parte de la base de $V_{H,\mathcal{T}}^{(m)}$ junto con

$$\tilde{\varphi}_{H,00}(x) = \tilde{\psi}_{10}(x), \quad \tilde{\varphi}_{H,N0}(x) = \tilde{\psi}_{Nn}(x).$$

Por tanto, hemos obtenido la base nodal de Hermite del espacio $V_{H,\mathcal{T}}^{(m)}$, formada por las funciones

$$\{\varphi_{H,j0}(x), \tilde{\varphi}_{H,j0}(x)\}_{j=0}^N \cup \{\varphi_{H,jk}(x), \tilde{\varphi}_{H,jk}(x)\}_{j=1, k=1}^{N, n-1},$$

por lo que la dimensión de este espacio es $2(N+1) + N(n-1)$.

Por ejemplo, el espacio $V_{H,\mathcal{T}}^{(3)}$ de polinomios a trozos cúbicos continuos y con derivada continua en la malla $\mathcal{T} \equiv \{x_j\}$, con $0 \leq j \leq N$, tiene como base las $2(N+1)$ funciones

$$\{\varphi_{H,j0}(x), \tilde{\varphi}_{H,j0}(x)\}_{j=0}^N.$$

Recuerde que $m = 3$ y $n = 1$ en este caso. Estas funciones toman la forma

$$\varphi_{H,00}(x) = \begin{cases} \left(1 - 2 \frac{x - x_0}{x_0 - x_1}\right) \frac{(x - x_1)^2}{(x_0 - x_1)^2}, & x \in I_1, \\ 0, & x \notin I_1, \end{cases}$$

$$\varphi_{H,j0}(x) = \begin{cases} \left(1 - 2 \frac{x - x_j}{x_j - x_{j-1}}\right) \frac{(x - x_{j-1})^2}{(x_j - x_{j-1})^2}, & x \in I_j, \\ \left(1 - 2 \frac{x - x_j}{x_j - x_{j+1}}\right) \frac{(x - x_{j+1})^2}{(x_j - x_{j+1})^2}, & x \in I_{j+1}, \\ 0, & x \notin I_j \cup I_{j+1}, \end{cases} \quad 0 < j < N,$$

$$\varphi_{H,N0}(x) = \begin{cases} \left(1 - 2 \frac{x - x_N}{x_N - x_{N-1}}\right) \frac{(x - x_{N-1})^2}{(x_N - x_{N-1})^2}, & x \in I_N, \\ 0, & x \notin I_N, \end{cases}$$

$$\tilde{\varphi}_{H,00}(x) = \begin{cases} \frac{(x - x_0)(x - x_1)^2}{(x_0 - x_1)^2}, & x \in I_1, \\ 0, & x \notin I_1, \end{cases}$$

$$\tilde{\varphi}_{H,j0}(x) = \begin{cases} \frac{(x - x_j)(x - x_{j-1})^2}{(x_j - x_{j-1})^2}, & x \in I_j, \\ \frac{(x - x_j)(x - x_{j+1})^2}{(x_j - x_{j+1})^2}, & x \in I_{j+1}, \\ 0, & x \notin I_j \cup I_{j+1}, \end{cases} \quad 0 < j < N,$$

$$\tilde{\varphi}_{H,N0}(x) = \begin{cases} \frac{(x - x_{j+1})(x - x_j)^2}{(x_{j+1} - x_j)^2}, & x \in I_N, \\ 0, & x \notin I_N, \end{cases}$$

y toda función $v(x) \in V_{H,\mathcal{T}}^{(3)}$ se escribe como

$$v(x) = \sum_{j=0}^N f(x_j) \varphi_{H,j0}(x) + \sum_{j=0}^N f'(x_j) \tilde{\varphi}_{H,j0}(x).$$

Como vemos se necesitan conocer $(N + 1)$ valores nodales de la función y otros tantos de su derivada. En la figura ? se muestran las funciones base de este espacio así como una función del mismo.

6.3.4 Interpolación mediante splines cúbicas

Utilizando las técnicas de interpolación explícita (nodal) mediante polinomios a trozos de grado a lo sumo n , estudiadas en las secciones anteriores, para $n > 1$, sólo podemos obtener funciones continuas y con derivada continua hasta grado $n - 2$. Por ejemplo, con trozos cuadráticos, $n = 2$, no podemos conseguir que la primera derivada sea continua; con trozos cúbicos, $n = 3$, no podemos obtener una segunda derivada continua. Ello no significa que no existan polinomios a trozos de grado $n > 0$ continuos y con derivada continua hasta grado $n - 1$. Estos polinomios a trozos se denominan *esplines* (*splines*) y se obtienen mediante técnicas de interpolación implícitas que estudiaremos en esta sección. El alumno notará que si un polinomio a trozos de grado $n > 0$ es continuo con derivada continua hasta grado n , entonces todos los trozos corresponden al mismo polinomio, y la función a trozos en realidad es el único polinomio interpolador (global) de la nube de puntos.

Definiremos una spline $s(x)$ como un polinomio a trozos, de grado a lo sumo n , continuo y con derivadas continuas $s^{(r)}(x)$ en todo el intervalo $x \in (a, b)$, hasta un orden menos que el suyo, es decir, para $0 \leq r \leq m - 1$. El espacio vectorial de estos polinomios se denomina $V_{S,\mathcal{T}}^{(n)}$, está descrito sobre la malla $\mathcal{T} \equiv \{x_j\}$ del intervalo $[a, b]$ dada por los nodos

$$a = x_0 < x_1 < x_2 < \cdots < x_N = b,$$

donde denotaremos por $I_j = [x_{j-1}, x_j]$ y por $h_j = x_j - x_{j-1}$, y se define como

$$V_{S,\mathcal{T}}^{(n)} = \{s(x) : \mathbf{C}^{n-1} \rightarrow \mathbb{R} : s|_{I_j}(x) \in \mathcal{P}^{(n)}(I_j)\}.$$

Normalmente llamaremos $s_j(x)$ a $s|_{I_j}(x)$, e igual para sus derivadas.

El ejemplo más importante de curvas esplines son las esplines cúbicas, que son las de menor grado con segunda derivada continua. Éstas definen curvas en el plano cuya curvatura es continua, que es un requisito muy importante en muchas aplicaciones prácticas. Además, producen una (función) interpolante muy suave y con pocas oscilaciones. Sin embargo, tienen un inconveniente, en los extremos necesitaremos condiciones adicionales (a las de interpolación) con objeto de obtener la unicidad, que se denominan condiciones de contorno. Como veremos, hay diferentes maneras de seleccionar estas condiciones de contorno, y en función de la aplicación conviene una elección u otra.

En general, una spline cúbica $s(x)$ toma la forma

$$s(x) = a_j + b_j x + c_j x^2 + d_j x^3, \quad x \in I_j \equiv [x_{j-1}, x_j],$$

por lo que tiene $4N$ incógnitas y cumple las siguientes condiciones de interpolación, continuidad y continuidad de las derivadas,

$$s(x_j) = f(x_j) \equiv y_j, \quad j = 0, 1, \dots, N, \quad r = 0, 1, 2,$$

$$s^{(r)}(x_j^+) \equiv s_{j+1}^{(r)}(x_j) = s_j^{(r)}(x_j) \equiv s^{(r)}(x_j^-), \quad j = 0, 1, \dots, N, \quad r = 0, 1, 2,$$

que constituyen $(N+1) + 3(N-1) = 4N - 2$ ecuaciones, por lo que nos quedan dos grados de libertad que aplicaremos como condiciones de contorno en los extremos del intervalo. Las $4n - 2$ ecuaciones para los coeficientes de la spline cúbica están acopladas entre sí y deberán ser obtenidos mediante la resolución de un sistema de ecuaciones lineales, que procederemos a derivar seguidamente. Por ello, a la interpolación con esplines se la denomina interpolación implícita.

Sean $M_j = s''(x_j)$, $j = 0, 1, \dots, N$, nuestros valores incógnita. Como $s(x)$ es un polinomio cúbico en $I_j \equiv [x_{j-1}, x_j]$, entonces $s'(x)$ es uno cuadrático y $s''(x)$ es lineal. La condición de continuidad para el polinomio a trozos lineales $s''(x)$, se puede escribir como

$$s_j''(x_j) = s_{j+1}''(x_j), \quad j = 1, 2, \dots, N-1,$$

y se cumple automáticamente, por construcción, si tomamos

$$s_j''(x) = \frac{(x - x_{j-1}) M_j + (x_j - x) M_{j-1}}{h_j},$$

recuerde que $s_j(x)$ está definida en $x \in I_j \equiv [x_{j-1}, x_j]$. Integrando esta expresión obtenemos la primera derivada,

$$s_j'(x) = \frac{(x - x_{j-1})^2 M_j - (x_j - x)^2 M_{j-1}}{2h_j} + E_j,$$

y volviendo a integrar, el polinomio cúbico salvo dos constantes de integración, que podemos reescribir por conveniencia como

$$s_j(x) = \frac{(x - x_{j-1})^3 M_j + (x_j - x)^3 M_{j-1}}{6 h_j} + C_j (x_j - x) + D_j (x - x_{j-1}).$$

Podemos aplicar la condición de interpolación y de continuidad directamente si hacemos

$$s_j(x_j) = f(x_j), \quad s_j(x_j - 1) = f(x_{j-1}),$$

ya que entonces $s_j(x_j) = f(x_j) = s_{j+1}(x_j)$, con lo que obtenemos el sistema de dos ecuaciones lineales con dos incógnitas (las constantes de integración) siguiente

$$s_j(x_j) = f(x_j) = \frac{h_j^2}{6} M_j + D_j h_j,$$

$$s(x_{j-1}) = f(x_{j-1}) = \frac{h_j^2}{6} M_{j-1} + C_j h_j,$$

cuya solución es fácil de obtener,

$$D_j = \frac{f(x_j)}{h_j} - \frac{h_j M_j}{6}, \quad C_j = \frac{f(x_{j-1})}{h_j} - \frac{h_j M_{j-1}}{6}.$$

De esta forma hemos obtenido el siguiente polinomio cúbico

$$\begin{aligned} s_j(x) &= \frac{(x - x_{j-1})^3 M_j + (x_j - x)^3 M_{j-1}}{6 h_j} \\ &+ \frac{f(x_{j-1})(x_j - x) + f(x_j)(x - x_{j-1})}{h_j} \\ &- \frac{h_j}{6} (M_{j-1}(x_j - x) + M_j(x - x_{j-1})), \end{aligned}$$

en el que faltan por determinar las segundas derivadas $\{M_j\}$.

Ahora sólo nos queda una condición por aplicar, la de la continuidad de la derivada en los nodos, $s'_j(x_j) = s'_{j+1}(x_j)$. Derivando $s_j(x)$ obtenemos que

$$s'_j(x) = \frac{(x - x_{j-1})^2 M_j - (x_j - x)^2 M_{j-1}}{2 h_j} + \frac{f(x_j) - f(x_{j-1})}{h_j} - \frac{M_j - M_{j-1}}{6} h_j,$$

y, derivando $s_{j+1}(x)$, que

$$s'_{j+1}(x) = \frac{(x - x_j)^2 M_{j+1} - (x_{j+1} - x)^2 M_j}{2 h_{j+1}} + \frac{f(x_{j+1}) - f(x_j)}{h_{j+1}} - \frac{M_{j+1} - M_j}{6} h_{j+1}.$$

Así, la condición de continuidad de la derivada, $s'_j(x_j) = s'_{j+1}(x_j)$, conduce a

$$\frac{h_j M_j}{2} + \frac{f(x_j) - f(x_{j-1})}{h_j} - \frac{M_j - M_{j-1}}{6} h_j = -\frac{h_{j+1} M_j}{2} + \frac{f(x_{j+1}) - f(x_j)}{h_{j+1}} - \frac{M_{j+1} - M_j}{6} h_{j+1},$$

que podemos reescribir como

$$\frac{h_j}{6} M_{j-1} + \left(\frac{h_j - h_{j+1}}{2} - \frac{h_j - h_{j+1}}{6} \right) M_j + \frac{h_{j+1}}{6} M_{j+1} = \frac{f(x_{j+1}) - f(x_j)}{h_{j+1}} - \frac{f(x_j) - f(x_{j-1}))}{h_j},$$

que conduce al siguiente sistema de ecuaciones lineales tridiagonal

$$\frac{h_j}{6} M_{j-1} + \left(\frac{h_j - h_{j+1}}{3} \right) M_j + \frac{h_{j+1}}{6} M_{j+1} = \frac{f(x_{j+1}) - f(x_j)}{h_{j+1}} - \frac{f(x_j) - f(x_{j-1}))}{h_j},$$

válido para $j = 1, 2, \dots, N-1$, es decir, que tenemos $(N-1)$ ecuaciones para $(N+1)$ incógnitas M_0, M_1, \dots, M_N . Necesitamos conocer los valores de M_0 y M_N . Para ello habrá que aportar condiciones de contorno adicionales que los determinen. Una vez especificadas estas condiciones de contorno, tendremos un sistema tridiagonal de ecuaciones lineales cuya matriz de coeficientes es diagonalmente dominante, luego no es singular, y por tanto el sistema tendrá solución única. Para resolver este sistema de ecuaciones se suele utilizar el algoritmo de Thomas para sistemas tridiagonales, cuya complejidad es lineal en el número de incógnitas.

Es importante notar que, dadas que fijadas las condiciones de contorno, la espline cúbica interpolante es única. Además, un cambio en las condiciones de contorno afecta a la curva entera, aunque afecta más cerca de los extremos de la misma y sólo ligeramente en su parte central. Esta propiedad de no-localidad de las esplines cúbicas es su mayor inconveniente.

Hay muchas formas de especificar los valores de las segundas derivadas M_0 y M_N en los extremos del intervalo de interpolación. Como ilustración de diferentes tipos de condiciones de contorno para las esplines cúbicas, presentaremos cinco posibilidades.

Caso 1. Si conocemos la segunda derivada de la función a interpolar, $f''(x)$, podemos especificar directamente las segundas derivadas de la espline en el contorno, es decir,

$$M_0 = s''(x_0) \equiv f''(a), \quad M_N = s''(x_N) \equiv f''(b).$$

Caso 2. Normalmente no conocemos la segunda derivada de $f(x)$, pero podemos especificar que los nodos en los extremos sean libres, es decir, que

$$M_0 = s''(a) = 0, \quad M_N = s''(b) = 0.$$

Estas condiciones son muy utilizadas por su simplicidad.

Caso 3. Si conocemos la primera derivada de la función $f(x)$ podemos especificar un valor dado para las derivadas de la espline en los puntos extremos, sean

$$s'(a) = f'(a), \quad s'(b) = f'(b),$$

que nos permiten obtener dos ecuaciones para M_0 y M_N . En el extremo derecho, tendremos

$$f'(a) = s'(x_0) = -\frac{h_1}{2} M_0 + \frac{f(x_1) - f(x_0)}{h_1} - \frac{h_1}{6} (M_1 - M_0),$$

de donde obtenemos la ecuación

$$\frac{h_1}{3} M_0 + \frac{h_1}{6} M_1 = \frac{f(x_1) - f(x_0)}{h_1} - f'(a),$$

y en el extremo izquierdo,

$$f'(b) = s'(x_N) = \frac{h_N}{2} M_N + \frac{f(x_N) - f(x_{N-1})}{h_N} - \frac{h_N}{6} (M_N - M_{N-1}),$$

de donde la otra ecuación

$$\frac{h_N}{6} M_{N-1} + \frac{h_N}{3} M_N = f'(b) - \frac{f(x_N) - f(x_{N-1})}{h_N}.$$

Añadiendo estas dos ecuaciones a las $(N - 1)$ que teníamos previamente, obtenemos un sistema lineal que sigue siendo tridiagonal y diagonalmente dominante, por lo que tiene solución única y se puede seguir utilizando el método de Thomas para su solución.

El alumno puede comprobar que para una malla \mathcal{T} uniforme, en la cual $h_j = h$ es constante, la matriz de coeficientes del sistema lineal con estas condiciones de contorno es, además de diagonalmente dominante, simétrica y definida positiva.

Caso 4. Otra posibilidad es especificar que la tercera derivada de la espline, $s^{(3)}(x)$, sea continua en los nodos x_1 y x_{N-1} . Como

$$s_j'''(x) = \frac{M_j - M_{j-1}}{h_j},$$

la continuidad en x_1 nos da

$$s_1'''(x_1) = \frac{M_1 - M_0}{h_1} = \frac{M_2 - M_1}{h_2} = s_2'''(x_1),$$

y la continuidad en x_N ,

$$s_N'''(x_{N-1}) = \frac{M_N - M_{N-1}}{h_N} = \frac{M_{N-1} - M_{N-2}}{h_{N-1}} = s_{N-1}'''(x_{N-1}),$$

que son dos ecuaciones lineales que relacionan M_0 y M_N con el resto de las incógnitas $\{M_j\}$.

Es posible demostrar que el sistema lineal de ecuaciones, que ya no es tridiagonal, aunque próximo a uno tridiagonal, tiene solución única. Estas condiciones de contorno son poco utilizadas.

Caso 5. Una última posibilidad, también utilizada algunas veces por su simplicidad, es imponer que $x_0 = x_3$ y $x_N = x_{N-2}$, que mantienen la tridiagonalidad del sistema de ecuaciones lineal resultante.

6.4 Otra derivación de la interpolación con esplines cúbicas (no entra en examen)

También se pueden derivar las ecuaciones para la interpolación con esplines cúbicas considerando como incógnitas las primeras derivadas $s'(x_j) = S_j$ en lugar de en función de las segundas derivadas $s''(x_j) = M_j$.

En esta derivación consideraremos la malla \mathcal{T} del intervalo $[a, b]$ definida por los nodos

$$a = x_0 < x_1 < x_2 < \cdots < x_N = b,$$

donde denotaremos por $I_j = [x_j, x_{j+1}]$ y por $h_j = x_{j+1} - x_j$. El lector debe notar la ligera diferencia de notación.

Denotaremos la curva spline cúbica como $p(x)$ y su restricción al intervalo I_j como $s_j(x)$. Buscaremos un polinomio cúbico en la base de Hermite,

$$\begin{aligned} p_j(x) &= f(x_j) + f[x_j, x_j](x - x_j) + f[x_j, x_j, x_{j+1}](x - x_j)^2 \\ &\quad + f[x_j, x_j, x_{j+1}, x_{j+1}](x - x_j)^2(x - x_{j+1}), \end{aligned}$$

donde las primeras derivadas $f[x_j, x_j] = f'(x_j) \equiv S_j$ serán las incógnitas del sistema de ecuaciones que obtendremos seguidamente. Introduciendo

$$x - x_{j+1} = x - x_j + x_j - x_{j+1},$$

en el polinomio de Hermite obtenemos

$$\begin{aligned} p_j(x) &= f(x_j) + f[x_j, x_j](x - x_j) \\ &\quad + (f[x_j, x_j, x_{j+1}] - f[x_j, x_j, x_{j+1}, x_{j+1}]h_j)(x - x_j)^2 \\ &\quad + f[x_j, x_j, x_{j+1}, x_{j+1}](x - x_j)^3, \end{aligned}$$

que podemos escribir en forma de Taylor como

$$p_j(x) = c_0 + c_1(x - x_j) + c_2(x - x_j)^2 + c_3(x - x_j)^3,$$

donde, sustituyendo las primeras derivadas (incógnitas) S_j y aplicando la definición de diferencias divididas, podemos escribir

$$c_0 = f(x_j), \quad c_1 = f[x_j, x_j] = S_j,$$

$$c_2 = \frac{f[x_j, x_{j+1}] - S_j}{h_j} - f[x_j, x_j, x_{j+1}, x_{j+1}] h_j = \frac{f[x_j, x_{j+1}] - S_j}{h_j} - c_3 h_j,$$

$$c_3 = f[x_j, x_j, x_{j+1}, x_{j+1}] = \frac{f[x_j, x_{j+1}, x_{j+1}] - f[x_j, x_j, x_{j+1}]}{h_j} = \frac{S_j + S_{j+1} - 2f[x_j, x_{j+1}]}{h_j^2}.$$

Las derivadas del polinomio cúbico son ahora

$$p'_j(x) = c_1 + 2c_2(x - x_j) + 3c_3(x - x_j)^2,$$

$$p''_j(x) = 2c_2 + 6c_3(x - x_j),$$

y para aplicar las condiciones de continuidad en las derivadas necesitamos también

$$p_{j-1}(x) = c'_0 + c'_1(x - x_{j-1}) + c'_2(x - x_{j-1})^2 + c'_3(x - x_{j-1})^3,$$

$$p'_{j-1}(x) = c'_1 + 2c'_2(x - x_{j-1}) + 3c'_3(x - x_{j-1})^2,$$

$$p''_{j-1}(x) = 2c'_2 + 6c'_3(x - x_{j-1}).$$

Ahora la continuidad de la segunda derivada

$$p''_j(x_j) = p''_{j-1}(x_j), \quad 2c_2 = 2c'_2 + 6c'_3 h_{j-1},$$

es decir,

$$2 \left(\frac{f[x_j, x_{j+1}] - S_j}{h_j} - c_3 h_j \right) = 2 \left(\frac{f[x_{j-1}, x_j] - S_{j-1}}{h_{j-1}} - c'_3 h_{j-1} \right) + 6c'_3 h_{j-1},$$

cuyo miembro izquierdo es, recordando la definición de c_3 ,

$$\begin{aligned} 2 \frac{f[x_j, x_{j+1}] - S_j}{h_j} - \frac{2}{h_j} (S_j + S_{j+1} - 2f[x_j, x_{j+1}]) \\ = 6 \frac{f[x_j, x_{j+1}]}{h_j} - \frac{4}{h_j} S_j - \frac{2}{h_j} S_{j+1}, \end{aligned}$$

y cuyo miembro derecho es, aplicando lo mismo para c'_3 ,

$$\begin{aligned} 2 \frac{f[x_{j-1}, x_j] - S_{j-1}}{h_{j-1}} + \frac{4}{h_{j-1}} (S_{j-1} + S_j - 2f[x_{j-1}, x_j]) \\ = -6 \frac{f[x_{j-1}, x_j]}{h_{j-1}} + \frac{2}{h_{j-1}} S_{j-1} + \frac{4}{h_{j-1}} S_j, \end{aligned}$$

e igualando estas expresiones obtenemos la ecuación

$$6 \left(\frac{f[x_j, x_{j+1}]}{h_j} + \frac{f[x_{j-1}, x_j]}{h_{j-1}} \right) = 4 \left(\frac{1}{h_j} + \frac{1}{h_{j-1}} \right) S_j + \frac{2}{h_{j-1}} S_{j-1} + \frac{2}{h_j} S_{j+1},$$

es decir, obtenemos para las pendientes S_j el sistema lineal de ecuaciones tridiagonal

$$2(h_j + h_{j-1}) S_j + h_j S_{j-1} + h_{j-1} S_{j+1} = 3(f[x_j, x_{j+1}] h_{j-1} + f[x_{j-1}, x_j] h_j),$$

que también es diagonalmente dominante por filas.

La aplicación de las condiciones de contorno a este sistema se deja como ejercicio para los alumnos interesados.

6.5 Polinomios trigonométricos

En la interpolación de funciones periódicas se recomienda el uso de técnicas basadas en desarrollo en serie de Fourier, es decir, en el uso de polinomios trigonométricos.

En lugar de considerar una función periódica de periodo cualquiera, consideraremos sólo funciones de periodo 2π , lo que se puede conseguir con un sencillo cambio de escala en la variable independiente. Sea una función periódica de periodo τ ,

$$\tilde{f}(t' + \tau) = \tilde{f}(t'),$$

reescalando la variable independiente podemos hacer que tenga un periodo 2π . Definiendo

$$t = \frac{2\pi t'}{\tau}, \quad f(t) = \tilde{f}\left(\frac{\tau}{2\pi} t\right) = \tilde{f}(t'),$$

obtenemos que $f(t)$ tiene periodo 2π , ya que

$$f(t) = \tilde{f}(t') = \tilde{f}(t' + \tau) = \tilde{f}\left(\frac{\tau}{2\pi} (t + 2\pi)\right) = f(t + 2\pi).$$

A partir de aquí consideraremos, sin pérdida de generalidad, sólo funciones periódicas de periodo 2π . Ello simplifica bastante las operaciones.

En las asignatura de cálculo de primer curso se presenta el teorema de Dirichlet, que dice que si $f(x)$ es continua y periódica de periodo 2π , entonces tiene desarrollo en serie de Fourier convergente, sea éste

$$f(t) = a_0 + \sum_{j=1}^{\infty} a_j \cos(jt) + b_j \sin(jt).$$

Truncando esta expresión, podemos obtener el polinomio trigonométrico

$$p_n(t) = a_0 + \sum_{j=1}^n a_j \cos(jt) + b_j \sin(jt),$$

que es de grado n si $|a_n| + |b_n| \neq 0$. Lo denominamos polinomio ya que aplicando la fórmula de Euler-de Moivre,

$$e^{inx} = \cos(nx) + i \sin(nx) = (\cos x + i \sin x)^n,$$

que es un polinomio en $\sin x$ y $\cos x$, podemos escribir el polinomio trigonométrico como

$$p_n(t) = a_0 + \sum_{j=1}^n \alpha_j \cos^j t + \beta_j \sin^j t.$$

Podemos ver más claramente que $p_n(t)$ es un polinomio si utilizamos las fórmulas de Euler

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}, \quad \sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i},$$

que nos permite escribir

$$\begin{aligned} p_n(t) &= \sum_{j=-n}^n c_j e^{ijt} \\ &= \sum_{j=1}^n c_{-j} e^{-ijt} + c_0 + \sum_{j=1}^n c_j e^{ijt} \\ &= c_0 + \sum_{j=1}^n (c_j + c_{-j}) \cos jt + i(c_j - c_{-j}) \sin jt, \end{aligned}$$

donde las constantes $\{c_j\}$ se calculan como la solución de

$$c_0 = a_0, \quad c_j + c_{-j} = a_j, \quad (c_j - c_{-j})i = b_j,$$

que nos da

$$c_j = \frac{a_j - ib_j}{2}, \quad c_{-j} = \frac{a_j + ib_j}{2}, \quad j = 1, \dots, n.$$

Escribiendo $z = e^{it}$, tenemos

$$p_n(t) = \sum_{j=-n}^n c_j z^j = P_n(z),$$

y multiplicando por z^n obtendremos

$$z^n p_n(t) = z^n P_n(z) = Q_n(z),$$

que es un polinomio de grado menor o igual que $2n$. De esta forma se ve claramente porque llamamos “polinomio” a un polinomio trigonométrico.

La analogía con los polinomios se puede llevar más lejos. Igual que el espacio $\mathcal{P}^{(n)}(0, 2\pi)$ es un espacio vectorial de dimensión finita $(n+1)$, el espacio de polinomios trigonométricos $\mathcal{F}^{(n)}(0, 2\pi)$ de grado a lo sumo n también es un espacio vectorial de dimensión finita, en concreto $(2n+1)$. Ya hemos presentado dos bases para este espacio, la de las funciones trigonométricas

$$\{1, \cos(jt), \sin(jt)\}, \quad j = 1, 2, \dots, n,$$

con lo que el cuerpo sobre el que está definido este espacio vectorial es el de los números reales, y la de exponenciales complejas

$$\{e^{-ij t}\}, \quad -n \leq j \leq n,$$

que utiliza los números complejos como cuerpo del espacio vectorial.

Como ocurría con el espacio vectorial de los polinomios, este espacio también es un espacio normado y completo (espacio de Banach) con las normas- q ,

$$\|p_n(x)\|_q = \left(\int_0^{2\pi} |p_n(x)|^q dx \right)^{1/q}.$$

Además es un espacio de Hilbert (espacio con producto interior completo) con el producto interior euclídeo, que en el caso de la base trigonométrica sobre el cuerpo de los números reales es

$$\langle p_{n,1}(x), p_{n,2}(x) \rangle = \left(\int_0^{2\pi} p_{n,1}(x) p_{n,2}(x) dx \right)^{1/2},$$

con norma asociada

$$\|p_n(x)\|_2 = \left(\int_0^{2\pi} p_n(x)^2 dx \right)^{1/2},$$

y que en el caso de la base de exponenciales complejas sobre el cuerpo de los números complejos es

$$\langle p_{n,1}(x), p_{n,2}(x) \rangle = \left(\int_0^{2\pi} \overline{p_{n,1}(x)} p_{n,2}(x) dx \right)^{1/2},$$

con norma asociada

$$\|p_n(x)\|_2 = \left(\int_0^{2\pi} |p_n(x)|^2 dx \right)^{1/2},$$

donde $|p_n(x)|^2 = \overline{p_n(x)} p_n(x)$.

En todo espacio de Hilbert se puede obtener una base ortogonal (y ortonormal). Las dos bases que hemos definido para $\mathcal{F}^{(n)}(0, 2\pi)$ son ortogonales.

Consideremos, primero, el caso de exponenciales complejas sobre el cuerpo de los números complejos. Dadas dos funciones de la base de este espacio tenemos

$$\langle e^{ij t}, e^{ik t} \rangle = \int_0^{2\pi} e^{i(k-j)t} dt = \frac{i}{k-j} \left(1 - e^{(k-j)2\pi i} \right) = 2\pi \delta_{jk} = \begin{cases} 2\pi, & j = k, \\ 0, & j \neq k. \end{cases}$$

Podemos definir una base ortonormal, que será,

$$\left\{ \frac{1}{2\pi} e^{ij t} \right\}, \quad -n \leq j \leq n.$$

Los coeficientes (de Fourier) $\{c_j\}$ del desarrollo en serie de Fourier de una función $f(x)$ se calculan fácilmente aprovechando esta base ortonormal, ya que

$$\langle e^{ij t}, f(t) \rangle = \sum_{k=-n}^n c_k \langle e^{ij t}, e^{ik t} \rangle = \sum_{k=-n}^n 2\pi c_k \delta_{jk} = 2\pi c_j,$$

con lo que (como es conocido de cálculo de primer curso)

$$c_j = \frac{\langle e^{ij t}, f(t) \rangle}{2\pi} = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ij t} dt.$$

Por otro lado, para la base de funciones trigonométricas sobre el cuerpo de los reales, el producto interior de todos los pares de funciones base nos da

$$\langle 1, 1 \rangle = \int_0^{2\pi} 1 dt = 2\pi,$$

$$\langle 1, \cos(j t) \rangle = \int_0^{2\pi} \cos(j t) dt = \left[\frac{\sin(j t)}{j} \right]_0^{2\pi} = 0, \quad 0 < j \leq n,$$

$$\langle 1, \sin(j t) \rangle = \int_0^{2\pi} \sin(j t) dt = \left[-\frac{\cos(j t)}{j} \right]_0^{2\pi} = 0, \quad 0 < j \leq n,$$

$$\begin{aligned} \langle \cos(j t), \sin(k t) \rangle &= \int_0^{2\pi} \cos(j t) \sin(k t) dt \\ &= \left[\frac{k \cos(j t) \cos(k t) + j \sin(j t) \sin(k t)}{j^2 - k^2} \right]_0^{2\pi} = 0, \quad 0 < j, k \leq n, \end{aligned}$$

$$\begin{aligned} \langle \cos(j t), \cos(k t) \rangle &= \int_0^{2\pi} \cos(j t) \cos(k t) dt \\ &= \left[\frac{j \cos(k t) \sin(j t) - k \cos(j t) \sin(k t)}{j^2 - k^2} \right]_0^{2\pi} \\ &= \frac{1}{2} \left(\frac{\sin(2(j-k)\pi)}{j-k} + \frac{\sin(2(j+k)\pi)}{j+k} \right) = \pi \delta_{jk}, \quad 0 < j, k \leq n, \end{aligned}$$

$$\begin{aligned} \langle \sin(j t), \sin(k t) \rangle &= \int_0^{2\pi} \sin(j t) \sin(k t) dt \\ &= \left[\frac{k \cos(k t) \sin(j t) - j \cos(j t) \sin(k t)}{j^2 - k^2} \right]_0^{2\pi} \\ &= \frac{1}{2} \left(\frac{\sin(2(j-k)\pi)}{j-k} - \frac{\sin(2(j+k)\pi)}{j+k} \right) = \pi \delta_{jk}, \quad 0 < j, k \leq n. \end{aligned}$$

Con ello se observa que los coeficientes (de Fourier) del desarrollo de Fourier de una función $f(t)$ periódica de periodo 2π , sea

$$f(t) = a_0 + \sum_{j=1}^n a_j \cos(jt) + b_j \sin(jt),$$

se calculan fácilmente como

$$a_0 = \frac{\langle 1, f(t) \rangle}{2\pi} = \frac{1}{2\pi} \int_0^{2\pi} f(t) dt,$$

$$a_j = \frac{\langle \cos(jt), f(t) \rangle}{2\pi} = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(jt) dt,$$

$$b_j = \frac{\langle \sin(jt), f(t) \rangle}{2\pi} = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(jt) dt,$$

resultado ya conocido de cálculo de primer curso.

Se puede acotar el error de la aproximación de una función $f(t) \in \mathbf{C}^0[0, 2\pi]$, periódica, tal que 2π es un entero múltiplo de su periodo, entonces

$$\max_{0 \leq t \leq 2\pi} |f(t) - p_n(t)| \leq C \log(n+2) \varrho_n(f), \quad n \geq 0,$$

donde c no es un función ni de f ni de n , y

$$\varrho_n(f) = \inf_{\text{grado}(q) \leq n} \max_{0 \leq t \leq 2\pi} |f(t) - q(t)|,$$

y $q(t)$ es un polinomio trigonométrico. No demostraremos este teorema [1].

Estos resultados serán generalizados a espacios de Hilbert más generales cuando estudiemos la teoría de la aproximación de Sturm-Liouville más adelante en este tema (a partir de la sección 6.6.2).

6.5.1 Interpolación trigonométrica y transformada discreta de Fourier

Aprovechando la analogía entre un polinomio trigonométrico y un polinomio podemos interpolar directamente los valores de una función $f(t)$ periódica de periodo 2π en el intervalo $[0, 2\pi)$ mediante un polinomio trigonométrico. Para ello tomaremos una malla de dicho intervalo formada por los $(2n+1)$ nodos $\{t_j\}$ tales que

$$0 \leq t_0 < t_1 < \cdots < t_{2n} < 2\pi.$$

La condición de interpolación aplicada al polinomio trigonométrico

$$p_n(t) = a_0 + \sum_{j=1}^n a_j \cos(jt) + b_j \sin(jt),$$

que tiene $(2n + 1)$ parámetros conduce al sistema de $(2n + 1)$ ecuaciones

$$p_n(t_i) = f(t_i).$$

Dado que las funciones base $\{1, \cos(jt), \sin(jt)\}$ son ortogonales entre sí (dos a dos), son también linealmente independientes dentro del espacio vectorial $\mathcal{F}^{(n)}(0, 2\pi)$, por lo que dicho sistema de ecuaciones tiene una matriz de coeficientes no singular y su solución existe y es única. Esta solución nos dará directamente los coeficientes del polinomio trigonométrico interpolador a_0 , a_j y b_j , $j = 1, \dots, n$. A veces se denomina a este polinomio trigonométrico como polinomio de colocación.

Los coeficientes del trigonométrico de interpolación se pueden calcular más fácilmente, de forma explícita, si los nodos de interpolación están distribuidos uniformemente por el intervalo de interpolación, es decir, nodos igualmente espaciados de la forma

$$t_j = j \frac{2\pi}{2n+1}, \quad j = 0, 1, 2, \dots, 2n.$$

El desarrollo se simplifica si utilizamos la base de exponenciales complejas del espacio $\mathcal{F}^{(n)}(0, 2\pi)$. En esta base el polinomio se escribe

$$p_n(t) = \sum_{j=-n}^n c_j e^{ijt} = \sum_{j=-n}^n c_j z^j = P_n(z),$$

y los nodos de interpolación se pueden considerar distribuidos en la circunferencia unidad del plano complejo, sean

$$z_j = e^{it_j}, \quad |z_j| = 1.$$

Para calcular los coeficientes c_j del polinomio de Fourier interpolante utilizaremos el siguiente lema.

Lema. $\forall k \in \mathbb{Z}$,

$$\sum_{j=0}^{2n} e^{ikt_j} = \begin{cases} 2n+1, & e^{it_k} = 1 \\ 0, & e^{it_k} \neq 1. \end{cases}$$

Además, $z_k = e^{it_k} = 1$ implica que t_k es un entero l múltiplo de 2π ,

$$t_k = l2\pi = j \frac{2\pi}{2n+1} = t_j, \quad j = l(2n+1),$$

con j un entero múltiplo de $(2n+1)$.

Demostración. Tomando

$$z = e^{it_k}, \quad t_j = j \frac{2\pi}{2n+1},$$

obtenemos

$$e^{ikt_j} = e^{ik2\pi j/(2n+1)} = e^{ijt_k} = z^j,$$

por lo que

$$s = \sum_{j=0}^{2n} e^{ikt_j} = \sum_{j=0}^{2n} e^{ijt_k} = \sum_{j=0}^{2n} z^j.$$

Esta expresión se suma fácilmente para $z = 1$, dando $s = 2n + 1$. Para $z \neq 1$ tenemos una progresión geométrica cuya suma es

$$s = \frac{z^{2n+1} - 1}{z - 1} = \frac{e^{(2n+1)it_k} - 1}{z - 1} = \frac{e^{ikt(2n+1)} - 1}{z - 1} = \frac{e^{i2\pi k} - 1}{z - 1} = 0,$$

cq.d.

Para calcular los coeficientes c_j del interpolante trigonométrico, aplicaremos la condición que lo define, $P_n(z_j) = f(t_j)$, que se escribe

$$\sum_{k=-n}^n c_k e^{ikt_j} = f(t_j),$$

y para todo $-n \leq l \leq n$, podemos sumarla de la forma

$$\sum_{j=0}^{2n} e^{-ilt_j} \left(\sum_{k=-n}^n c_k e^{ikt_j} \right) = \sum_{j=0}^{2n} e^{-ilt_j} f(t_j),$$

y como el lema anterior nos dice que

$$\sum_{j=0}^{2n} e^{it_j(k-l)} = \begin{cases} 2n+1, & e^{it_{k-l}} = 1, \\ 0, & e^{it_{k-l}} \neq 1, \end{cases}$$

es decir, que es distinto de cero si existe un entero m tal que

$$(k-l) \frac{2\pi}{2n+1} = m2\pi,$$

luego $k = l + (2n+1)m$, que sólo es cierto para $m = 0$, y por tanto,

$$\sum_{k=-n}^n c_k \sum_{j=0}^{2n} e^{it_j(k-l)} = \sum_{k=-n}^n c_k (2n+1) \delta_{kl},$$

con lo que obtenemos

$$c_l = \frac{1}{2n+1} \sum_{j=0}^{2n} f(t_j) e^{-ilt_j},$$

para $l = -n, \dots, n$. Estos coeficientes se denominan transformada discreta de Fourier de la función discreta $\{f_j\}$, con $j = 0, 1, \dots, 2n$, donde $f_j \equiv f(t_j)$. Normalmente se escribe la transformada de Fourier discreta como

$$\mathbb{F}\{f_j\} = c_k = \frac{1}{2n+1} \sum_{j=0}^{2n} f_j \exp(-2\pi i k j / (2n+1)), \quad -n \leq k \leq n,$$

y la transformada inversa de Fourier discreta como

$$\mathbb{F}^{-1}\{c_k\} = f_j = \sum_{k=-n}^n c_k \exp(2\pi i k j / (2n+1)), \quad 0 \leq j \leq 2n.$$

La razón por la que se denomina a estas expresiones transformadas de Fourier discretas es por analogía con la transformada de Fourier de una función continua, sea $f(t) : \mathbb{R} \rightarrow \mathbb{R}$, que el alumno ha estudiado en cursos anteriores y que es

$$\mathbb{F}[f(t)] = \hat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikt} f(t) dt, \quad k \in \mathbb{R},$$

cuya inversa, la transformada de Fourier inversa, es

$$\mathbb{F}^{-1}[\hat{f}(k)] = f(t) = \int_0^{2\pi} e^{ikt} \hat{f}(k) dt, \quad t \in \mathbb{R}.$$

De hecho la transformada discreta de Fourier se puede obtener a partir de la continua mediante la aplicación de la regla de integración del trapecio (que estudiaremos en el tema 9) usando $(2n+1)$ subdivisiones del intervalo $[0, 2\pi]$, con puntos igualmente espaciados entre sí y siendo $f(t)$ una función periódica en $[0, 2\pi]$.

La transformada de Fourier discreta y su inversa, que no son otra cosa que la interpolación mediante polinomios trigonométricos, son una herramienta indispensable en el análisis de señales digitales en el, así denominado, dominio de la frecuencia. El campo de la teoría de la señal se centra fundamentalmente en los sistemas lineales invariantes en el tiempo. En el caso continuo se uti

6.5.2 Transformada rápida de Fourier (no entra en examen)

Dada la importancia de la transformada discreta de Fourier en el análisis de señales, para su cálculo se utiliza un algoritmo que reduce su coste computacional de $O(n^2)$ a $O(n \log n)$, y que por ello se denomina transformada rápida de Fourier. Su importancia práctica en las aplicaciones de la interpolación trigonométrica nos obliga a introducir aquí las bases de este algoritmo, uno de los diez algoritmos más importantes de todo el análisis numérico.

TO BE WRITTEN...

6.6 Aproximación de funciones

La idea de la aproximación de funciones es aproximar una función complicada por una más simple, de tal forma que la función aproximada sea más fácil de evaluar, derivar, integrar, etc.

La aproximación de funciones está muy relacionada con la interpolación y, de hecho, todas las técnicas de aproximación se pueden aplicar como técnicas de interpolación y viceversa.

Como funciones simples se pueden utilizar polinomios, funciones racionales (cocientes de polinomios), polinomios trigonométricos, polinomios a trozos o, incluso, funciones más generales mediante técnicas de mínimos cuadrados.

6.6.1 Aproximación por polinomios. Teorema de Weierstrass

Teorema. Dada una función $f(x) \in \mathbf{C}^0[a, b]$ y $\epsilon > 0$, existe un polinomio $p(x)$ tal que

$$|f(x) - p(x)| < \epsilon, \quad \forall x \in [a, b].$$

Para aproximar una función mediante un polinomio podemos utilizar las técnicas de interpolación de Lagrange, Newton, Hermite, splines, etc.

También podemos utilizar una aproximación local de Taylor. Por ejemplo, para la exponencial

$$e^x, \quad x \in [-1, 1],$$

podemos aproximarla por un polinomio de Taylor de tercer grado

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

que por el teorema del resto de Taylor comete el error

$$e^x - p_3(x) = \frac{x^4}{4!} e^\xi, \quad \xi \in (0, x).$$

En el intervalo $0 \leq x \leq 1$ podemos acotar

$$\frac{x^4}{4!} \leq e^x - p_3(x) \leq \frac{e x^4}{4!},$$

y en el intervalo $0 \geq x \geq -1$ podemos acotar

$$\frac{x^4}{4!} \geq e^x - p_3(x) \geq \frac{x^4}{e 4!},$$

con lo que observamos que el error crece conforme lo hace $|x|$ y que no está uniformemente distribuido en todo el intervalo $[-1, 1]$.

Para obtener aproximaciones cuyo error se distribuya uniformemente en todo el intervalo de interés deberemos utilizar aproximaciones de tipo minimax que se basan en aproximaciones de mínimos cuadrados. Para introducir estas aproximaciones conviene introducir previamente la teoría de Sturm-Liouville para ecuaciones diferenciales de segundo orden, que es equivalente a la teoría de matrices hermitianas, y que permiten definir bases de funciones que permiten aproximar funciones consideradas como vectores en un espacio vectorial de funciones.

6.6.2 Teoría de Sturm-Liouville

Sea $x \in [a, b]$, estudiaremos la ecuación diferencial ordinaria de segundo orden

$$\mathcal{L}(y) \equiv p \frac{d^2 y}{dx^2} + q \frac{dy}{dx} + r y = f,$$

donde p , q y r son funciones de $x \in [a, b]$.

En nuestro análisis necesitaremos del operador autoadjunto de esta ecuación, que se obtiene integrando por partes

$$\begin{aligned} \int_a^b z \mathcal{L}(y) dx &= \int_a^b \left(p z \frac{d^2 y}{dx^2} + q z \frac{dy}{dx} + r z y \right) dx \\ &= p z \frac{dy}{dx} \Big|_a^b - \int_a^b \frac{d}{dx} (p z) \frac{dy}{dx} dx \\ &\quad + q z y \Big|_a^b - \int_a^b y \frac{d(q z)}{dx} dx + \int_a^b r z y dx \\ &= \left(p z \frac{dy}{dx} + q z y - y \frac{d}{dx} (p z) \right) \Big|_a^b \\ &\quad + \int_a^b y \left(\frac{d^2 (p z)}{dx^2} - \frac{d(q z)}{dx} + r z \right) dx \\ &= \left(p z \frac{dy}{dx} + q z y - y \frac{d}{dx} (p z) \right) \Big|_a^b + \int_a^b y \mathcal{L}^*(z) dx, \end{aligned}$$

y se denomina operador adjunto $\mathcal{L}^*(z)$,

$$\mathcal{L}^*(y) \equiv \frac{d^2 (p y)}{dx^2} - \frac{d(q y)}{dx} + r y$$

Vamos a demostrar que toda ecuación diferencial ordinaria lineal de segundo orden se puede escribir en forma de ecuación de Sturm-Liouville, sea,

$$\frac{d}{dx} \left(\alpha \frac{dy}{dx} \right) + \beta y = \gamma.$$

Es obvio, que podemos escribirla de la forma

$$\begin{aligned} \mathcal{L}(y) &= p \frac{d^2 y}{dx^2} + q \frac{dy}{dx} + r y = \frac{d}{dx} \left(\alpha \frac{dy}{dx} \right) + \beta \frac{dy}{dx} + r y \\ &= \alpha \frac{d^2 y}{dx^2} + \frac{d\alpha}{dx} \frac{dy}{dx} + \beta \frac{dy}{dx} + \gamma y, \end{aligned}$$

donde

$$\alpha = p, \quad \beta = q - \frac{d\alpha}{dx}, \quad \gamma = r.$$

Sin embargo, para escribir en la forma de Sturm-Liouville, necesitamos Aplicar una transformación de Liouville

$$y = \alpha(x) z,$$

que nos permite escribir

$$\begin{aligned} \mathcal{L}(y) &= p \frac{d^2(\alpha z)}{dx^2} + q \frac{d(\alpha z)}{dx} + r \alpha z \\ &= p \frac{d}{dx} \left(\frac{d\alpha}{dx} z + \alpha \frac{dz}{dx} \right) + q \left(\frac{d\alpha}{dx} z + \alpha \frac{dz}{dx} \right) + r \alpha z \\ &= p \left(\frac{d^2\alpha}{dx^2} z + \alpha \frac{d^2z}{dx^2} + 2 \frac{d\alpha}{dx} \frac{dz}{dx} \right) + q \left(\frac{d\alpha}{dx} z + \alpha \frac{dz}{dx} \right) + r \alpha z \\ &= \alpha p \frac{d^2z}{dx^2} + \left(q \alpha + 2p \frac{d\alpha}{dx} \right) \frac{dz}{dx} + z \left(p \frac{d^2\alpha}{dx^2} + q \frac{d\alpha}{dx} + r \alpha \right), \end{aligned}$$

y si ahora podemos elegir α tal que se haga cero el término de primer orden en la derivada

$$q \alpha + 2p \frac{d\alpha}{dx} = 0, \quad \frac{d\alpha}{\alpha} = -\frac{q}{2p} dx$$

$$\ln \alpha = -\frac{1}{2} \int_0^x \frac{q}{p} dx,$$

entonces

$$\mathcal{L}(y) = \mathcal{L}(\alpha z) = \alpha p \frac{d^2z}{dx^2} + z \left(p \frac{d^2\alpha}{dx^2} + q \frac{d\alpha}{dx} + r \alpha \right).$$

La ecuación diferencial de segundo orden no homogénea,

$$\mathcal{L}(y) = p \frac{d^2 y}{dx^2} + q \frac{dy}{dx} + r y = p \left(\frac{d^2 y}{dx^2} + \frac{q}{p} \frac{dy}{dx} + \frac{r}{p} y \right) = f(x)$$

si $p(x) \neq 0$, se puede escribir

$$\frac{d^2 y}{dx^2} + \frac{q}{p} \frac{dy}{dx} + \frac{r}{p} y = \frac{f}{p},$$

y por tanto, tiene forma de Sturm-Liouville

$$\frac{d}{dx} \left(\alpha \frac{dy}{dx} \right) + \beta y = \gamma,$$

es decir,

$$\alpha \frac{d^2 y}{dx^2} + \frac{d\alpha}{dx} \frac{dy}{dx} + \beta y = \gamma,$$

$$\frac{d^2 y}{dx^2} + \frac{1}{\alpha} \frac{d\alpha}{dx} \frac{dy}{dx} + \frac{\beta}{\alpha} y = \frac{\gamma}{\alpha},$$

si igualamos términos

$$\frac{1}{\alpha} \frac{d\alpha}{dx} = \frac{q}{p}, \quad \ln \alpha = \int_0^x \frac{q}{p} dx,$$

$$\frac{\beta}{\alpha} = \frac{r}{p}, \quad \beta = \frac{r\alpha}{p} = \frac{r}{p} \exp \left(\int_0^x \frac{q}{p} dx \right),$$

$$\frac{\gamma}{\alpha} = \frac{f}{p}, \quad \gamma = \frac{f\alpha}{p} = \frac{f}{p} \exp \left(\int_0^x \frac{q}{p} dx \right).$$

Es decir, toda ecuación de segundo orden se puede reducir a una ecuación con la forma de Sturm-Liouville.

La ecuación de Sturm-Liouville tiene la ventaja de que es auto-adjunta. Escribamos

$$\frac{d}{dx} \left(p \frac{dy}{dx} \right) + q y = f,$$

y calculemos su operador autoadjunto

$$\begin{aligned} \int_a^b z \mathcal{L}(y) dx &= \int_a^b \left(z \frac{d}{dx} \left(p \frac{dy}{dx} \right) + q y z \right) dx \\ &= \left[z p \frac{dy}{dx} \right]_a^b - \int_a^b p \frac{dz}{dx} \frac{dy}{dx} dx + \int_a^b q y z dx \\ &= \left[z p \frac{dy}{dx} - p \frac{dz}{dx} y \right]_a^b + \int_a^b y \underbrace{\left(\frac{d}{dx} \left(p \frac{dz}{dx} \right) + q z \right)}_{\mathcal{L}^*(z)} dx, \end{aligned}$$

luego $\mathcal{L}(y) = \mathcal{L}^*(y)$ y el operador de Sturm-Liouville es autoadjunto. Sin embargo, el problema diferencial será autoadjunto

$$\mathcal{L}(y) = f = \mathcal{L}^*(y),$$

solamente si las condiciones de contorno garantizan que

$$\left[z p \frac{dy}{dx} - p \frac{dz}{dx} y \right]_a^b = 0.$$

Ahora vamos a estudiar el problema de valores propios para el operador autoadjunto de Sturm-Liouville con condiciones de contorno de Robin,

$$\frac{d}{dx} \left(p \frac{dy}{dx} \right) + (q + \lambda r) y = 0,$$

$$\alpha y(a) + \beta \frac{dy(a)}{dx} = 0, \quad \gamma y(b) + \delta \frac{dy(b)}{dx} = 0.$$

Este problema es autoadjunto, ya que las condiciones de contorno de Robin cumplen

$$\begin{aligned} & \left[z p \frac{dy}{dx} - p \frac{dz}{dx} y \right]_a^b \\ &= z(b) p(b) \frac{dy(b)}{dx} - p(b) \frac{dz(b)}{dx} y(b) - z(a) p(a) \frac{dy(a)}{dx} + p(a) \frac{dz(a)}{dx} y(a) \\ &= p(b) \left(z(b) \frac{dy(b)}{dx} - \frac{dz(b)}{dx} y(b) \right) - p(a) \left(z(a) \frac{dy(a)}{dx} - \frac{dz(a)}{dx} y(a) \right) \\ &= p(b) y(b) \left(-\frac{\gamma}{\delta} z(b) - \frac{dz(b)}{dx} \right) - p(a) y(a) \left(-\frac{\alpha}{\beta} z(a) - \frac{dz(a)}{dx} \right) = 0, \end{aligned}$$

si el operador adjunto está sometido también a la mismas condiciones de contorno de Robin.

Una solución trivial de este problema autoadjunto (como de cualquier problema de valores propios) es la trivial $y \equiv 0$ que no es de interés. ¿Para qué valores de λ se obtendrán soluciones no nulas $y \neq 0$? Estudiaremos el caso $r > 0$.

Primero vamos a demostrar que las autofunciones para autovalores distintos son ortogonales entre sí respecto al producto interior asociado a la función r . Consideremos dos autofunciones y_i e $f(x_j)$,

$$\frac{d}{dx} \left(p \frac{dy_i}{dx} \right) + (q + \lambda_i r) y_i = 0,$$

$$\frac{d}{dx} \left(p \frac{df(x_j)}{dx} \right) + (q + \lambda_j r) f(x_j) = 0,$$

y restando estas ecuaciones para que desaparezca el término en q ,

$$f(x_j) \frac{d}{dx} \left(p \frac{dy_i}{dx} \right) - y_i \frac{d}{dx} \left(p \frac{df(x_j)}{dx} \right) + (\lambda_i - \lambda_j) r y_i f(x_j) = 0,$$

$$\frac{d}{dx} \left(p \left(f(x_j) \frac{dy_i}{dx} - y_i \frac{df(x_j)}{dx} \right) \right) + (\lambda_i - \lambda_j) r y_i f(x_j) = 0,$$

e integrando

$$\int_a^b (\lambda_i - \lambda_j) r y_i f(x_j) dx + \left[p \left(f(x_j) \frac{dy_i}{dx} - y_i \frac{df(x_j)}{dx} \right) \right]_a^b = 0,$$

obtenemos

$$(\lambda_i - \lambda_j) \int_a^b r y_i f(x_j) dx = 0$$

que si $\lambda_i \neq \lambda_j$ es la condición de r -ortogonalidad de las autofunciones

$$\int_a^b r y_i f(x_j) dx = \langle y_i, f(x_j) \rangle_r = 0.$$

Por tanto, las autofunciones se pueden ortonormalizar fácilmente

$$\hat{y}_i = \frac{y_i}{\sqrt{\int_a^b r y_i^2 dx}}, \quad \int_a^b r \hat{y}_i^2 dx = 1,$$

es decir, $\langle y_i, f(x_j) \rangle_r = \delta_{ij}$.

Además, si $p, q, r, \alpha, \beta, \gamma$ y δ son reales, los autovalores también son reales. Veámoslo,

$$\begin{aligned} \frac{d}{dx} \left(p \frac{dy_i}{dx} \right) + (q + \lambda_i r) y_i &= 0, \\ \frac{d}{dx} \left(p \frac{d\bar{y}_i}{dx} \right) + (q + \bar{\lambda}_i r) \bar{y}_i &= 0, \end{aligned}$$

con lo que

$$\begin{aligned} \bar{y}_i \frac{d}{dx} \left(p \frac{dy_i}{dx} \right) - y_i \frac{d}{dx} \left(p \frac{d\bar{y}_i}{dx} \right) + (\lambda_i - \bar{\lambda}_i) r y_i \bar{y}_i &= 0, \\ \frac{d}{dx} \left(p \left(\bar{y}_i \frac{dy_i}{dx} - y_i \frac{d\bar{y}_i}{dx} \right) \right) + (\lambda_i - \bar{\lambda}_i) r |y_i|^2 &= 0, \end{aligned}$$

con lo que

$$(\lambda_i - \bar{\lambda}_i) \int_a^b r |y_i|^2 dx = 0,$$

lo que implica que $\lambda_i = \bar{\lambda}_i$, es decir, los autovalores son reales.

Se puede probar también que hay un número infinito numerable de autovalores λ_i y que las autofunciones y_i forman un espacio de funciones completo, es decir, que dada una función $f(x)$ en $[a, b]$ que cumpla con las condiciones de contorno de Robin puede desarrollarse en una serie uniformemente convergente

$$f(x) = \sum_{i=1}^{\infty} \alpha_i y_i(x),$$

donde los coeficientes α_i se calculan

$$\langle f(x_j), f \rangle_r = \int_a^b r f(x_j) f dx = \int_a^b \sum_{i=1}^{\infty} \alpha_i r y_i f(x_j) dx,$$

que por la convergencia uniforme

$$\sum_{i=1}^{\infty} \alpha_i \int_a^b r y_i f(x_j) dx = \alpha_j \int_a^b r f(x_j)^2 dx,$$

luego

$$\alpha_j = \frac{\int_a^b r f(x_j) f dx}{\int_a^b r f(x_j)^2 dx} = \frac{\langle f(x_j), f \rangle_r}{\langle f(x_j), f(x_j) \rangle_r}.$$

6.6.3 Funciones trigonométricas de Legendre, Bessel, Hermite y Chebyshev

Consideremos el problema de Sturm-Liouville más sencillo, es decir, la siguiente ecuación diferencial lineal con condiciones de contorno de Dirichlet homogéneas

$$\frac{d^2y}{dx^2} + \lambda y = 0, \quad y(a) = y(b) = 0.$$

Aplicaremos una traslación en $x \rightarrow x - a$ y escribamos $L = b - a$, por lo que tenemos el problema

$$\frac{d^2y}{dx^2} + \lambda y = 0, \quad y(0) = y(L) = 0.$$

Las soluciones de este problema diferencial son bien conocidas y dependen del signo de λ . Sea $\lambda > 0$, entonces

$$y = A \cos \sqrt{\lambda} x + B \sin \sqrt{\lambda} x,$$

y aplicando las condiciones de contorno

$$y(0) = 0 = A, \quad y(L) = B \sin \sqrt{\lambda} L = 0,$$

por lo que, o bien $B = 0$, en cuyo caso obtenemos la solución trivial $y = 0$ que no tiene interés, o bien

$$\sin \sqrt{\lambda} L = 0, \quad \sqrt{\lambda_n} L = n \pi,$$

y hemos obtenidos los valores propios y las autofunciones

$$\lambda_n = \frac{n^2 \pi^2}{L^2}, \quad y_n(x) = B_n \sin \frac{n \pi x}{L}.$$

Nótese que hemos sustituido B por B_n ya que como la ecuación es lineal, dada una solución $z(x)$ entonces $\alpha z(x)$ es también otra solución con α una constante.

Para $\lambda = 0$,

$$\frac{d^2y}{dx^2} = 0, \quad y = Ax + B,$$

y por tanto

$$y(0) = 0 = B, \quad y(L) = AL = 0 = A, \quad y = 0,$$

hemos obtenido la solución trivial, que no tiene interés.

Para $\lambda < 0$, tomamos $\lambda = -\mu$, con $\mu > 0$, y la ecuación

$$\frac{d^2y}{dx^2} - \mu y = 0, \quad y(a) = y(b) = 0,$$

tiene una solución de la forma

$$y = A \cosh \sqrt{\lambda} x + B \sinh \sqrt{\lambda} x,$$

y aplicando las condiciones de contorno

$$y(0) = 0 = A, \quad y(L) = B \sinh \sqrt{\lambda} L = 0 = 0,$$

y volvemos a obtener la solución trivial. En resumen, hemos obtenido los autovalores reales y las autofunciones

$$\lambda_n = \frac{n^2 \pi^2}{L^2}, \quad y_n(x) = B_n \sin \frac{n \pi x}{L}, \quad n = 1, 2, \dots,$$

que forman un conjunto completo para todas las funciones definidas en el intervalo $[0, L]$ y que son nulas en sus extremos. Estas autofunciones están en la base de las series de Fourier.

Seguidamente vamos a presentar otros ejemplos de ecuaciones de Sturm-Liouville con sus correspondientes autovalores y autofunciones (que son llamadas funciones especiales).

Polinomios de Legendre. En el intervalo $[-1, 1]$, las soluciones de la ecuación

$$(1 - x^2) \frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + n(n + 1)y = 0,$$

que puede escribirse

$$\frac{d}{dx} \left((1 - x^2) \frac{dy}{dx} \right) + \lambda_n y = 0,$$

cuyos autovalores son $\lambda_n = n(n + 1)$ y cuyas autofunciones son los polinomios de Legendre $P_n(x)$, que pueden determinarse por la siguiente fórmula de Rodrigues

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

Estos polinomios pueden determinarse por la fórmula de recurrencia

$$(n + 1) P_{n+1}(x) - (2n + 1)x P_n(x) + n P_{n-1}(x) = 0,$$

a partir de $P_0(x) = 1$ y $P_1(x) = x$; por ejemplo,

$$P_2(x) = \frac{1}{2} (3x^2 - 1), \quad P_3(x) = \frac{1}{2} (5x^3 - 3x), \quad \dots$$

Toda función en $[-1, 1]$ puede aproximarse en desarrollo de Fourier con polinomios de Legendre de la forma

$$f(x) = \sum_{i=0}^{\infty} \alpha_n P_n(x), \quad \alpha_n = \frac{2n + 1}{2} \int_{-1}^1 f(x) P_n(x) dx.$$

Funciones de Bessel. Consideremos la ecuación de Bessel

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - n^2) y = 0,$$

en el intervalo $[0, \infty]$. Esta ecuación puede escribirse como un problema de Sturm-Liouville,

$$\frac{d}{dx} \left(p \frac{dy}{dx} \right) + (q + \lambda r) y = 0,$$

donde $p = x$, $q = x$ y

$$r = \frac{1}{x},$$

es decir,

$$\frac{d}{dx} \left(x \frac{dy}{dx} \right) + \left(x + \frac{\lambda}{x} \right) y = 0,$$

cuyos autovalores son $\lambda_n = -n^2$ y cuyas autofunciones son $J_n(x)$ y $J_{-n}(x)$. Hay varios posibles desarrollos en serie de Fourier para las funciones de Bessel, por ejemplo,

$$f(x) = \sum_{i=1}^{\infty} \alpha_i J_n(\kappa_i x), \quad \alpha_i = \frac{2}{J_{n+1}^2(\kappa_i)} \int_0^1 x f(x) J_n(\kappa_i x) dx,$$

donde κ_i son las raíces positivas de $J_n(x) = 0$.

Polinomios de Hermite. Consideremos la ecuación de Hermite en el intervalo $(-\infty, \infty)$,

$$\frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + 2n y = 0,$$

que podemos escribir en forma de Sturm-Liouville

$$\frac{d^2 y}{dx^2} + \frac{1}{p} \frac{dp}{dx} \frac{dy}{dx} + \left(\frac{q}{p} + \lambda \frac{r}{p} \right) y = 0,$$

donde

$$\frac{1}{p} \frac{dp}{dx} = -2x, \quad p = e^{-x^2},$$

$q = 0$, $r = p$ y $\lambda_n = 2n$. Los polinomios de Hermite se pueden calcular por la fórmula de Rodrigues

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}),$$

o por la fórmula de recurrencia

$$H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x),$$

donde $H_0(x) = 1$, $H_1(x) = 2x$; por ejemplo,

$$H_2(x) = 4x^2 - 2, \quad H_3(x) = 8x^3 - 12x, \quad \dots,$$

Polinomios de Chebyshev. Consideremos la ecuación de Chebyshev en el intervalo $(-1, 1)$,

$$(1 - x^2) \frac{d^2 y}{dx^2} - x \frac{dy}{dx} + n^2 y = 0,$$

que podemos escribir en forma de Sturm-Liouville donde

$$\frac{1}{p} \frac{dp}{dx} = -\frac{x}{1-x^2}, \quad p = \sqrt{1-x^2},$$

$$\frac{r}{p} = \frac{1}{1-x^2}, \quad r = \frac{1}{\sqrt{1-x^2}},$$

$q = 0$ y $\lambda_n = n^2$. Los polinomios de Chebyshev se pueden calcular por la fórmula

$$T_n(x) = \cos(n \arccos x), \quad n \geq 0.$$

Los primeros polinomios de Chebyshev son

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x, \quad \dots$$

El desarrollo de Fourier para los polinomios de Chebyshev toma la forma

$$f(x) = \frac{\alpha_0}{2} T_0(x) + \sum_{n=1}^{\infty} \alpha_n T_n(x), \quad \alpha_n = \frac{2}{\pi} \int_{-1}^1 \frac{f(x) T_n(x)}{\sqrt{1-x^2}} dx.$$

Funciones de Jacobi. La ecuación de Jacobi

$$x(1-x) \frac{d^2y}{dx^2} + (a - (1+b)x) \frac{dy}{dx} + n(b+n)y = 0,$$

que es un caso particular de la ecuación diferencial hipergeométrica, también se puede escribir como una ecuación de Sturm-Liouville. Omitiremos los detalles.

Polinomios de Laguerre. La ecuación de Laguerre

$$x \frac{d^2y}{dx^2} + (1-x) \frac{dy}{dx} + ny = 0,$$

en el intervalo $[0, \infty)$ se puede escribir en la forma de Sturm-Liouville de la forma

$$\frac{1}{p} \frac{dp}{dx} = \frac{1-x}{x}, \quad p = x e^{-x},$$

$$\frac{r}{p} = \frac{1}{x}, \quad r = e^{-x},$$

$q = 0$ y $\lambda_n = n$. Los polinomios de Laguerre se pueden calcular mediante la fórmula de Rodrigues

$$L_n(x) = e^x \frac{d^n}{dx^n} (x^n e^{-x}), \quad n \geq 0,$$

o mediante la fórmula de recurrencia

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

donde $L_0(x) = 1$, $L_1(x) = 1-x$; por ejemplo,

$$L_2(x) = x^2 - 4x + 2, \quad L_3(x) = -x^3 + 9x^2 - 18x + 6, \quad \dots$$

y tienen como desarrollo de Fourier,

$$f(x) = \sum_{i=0}^{\infty} \alpha_i L_i(x), \quad \alpha_i = \frac{(i!)^2}{\pi} \int_0^{\infty} e^{-x} f(x) L_i(x) dx.$$

6.6.4 Aproximaciones continuas mediante mínimos cuadrados

La aproximación mínimo-cuadrática de una función $f(x)$ mediante otra función $p(x)$ se basa en elegir esta última de forma que se minimice la norma- r de la diferencia, es decir,

$$\langle f - p, f - p \rangle_r = \int_a^b (f - p)^2 r \, dx.$$

La función p se obtiene truncando una serie de Fourier para una ecuación de Sturm-Liouville,

$$p(x) = \sum_{i=0}^n \alpha_i P_i(x).$$

Las funciones P_i son normalmente polinomios ortogonales. De esta forma, el error mínimo cuadrático se escribe

$$E(\alpha_0, \alpha_1, \dots, \alpha_n) = \int_a^b r \left(f(x) - \sum_{i=0}^n \alpha_i P_i(x) \right)^2 dx,$$

y será mínimo si

$$\begin{aligned} 0 &= \frac{\partial E}{\partial \alpha_j} = -2 \int_a^b r \left(f - \sum_{i=0}^n \alpha_i P_i \right) P_j \, dx \\ &= -2 \int_a^b r f P_j \, dx + 2 \sum_{i=0}^n \alpha_i \int_a^b r P_i P_j \, dx \\ &= -2 \int_a^b r f P_j \, dx + 2 \alpha_j \int_a^b r P_j^2 \, dx, \quad j = 0, \dots, n, \end{aligned}$$

por lo que los coeficientes

$$\alpha_j = \frac{\int_a^b r f P_j \, dx}{\int_a^b r P_j^2 \, dx} = \frac{\langle f, P_j \rangle_r}{\langle P_j, P_j \rangle_r}.$$

Como vemos, la teoría de aproximación mínimo-cuadrática está conectada con la teoría de la aproximación de Sturm-Liouville y consiste en truncar a un número finito de términos el desarrollo de Fourier en las funciones (polinomios) ortogonales para la función considerada. Es decir, la mejor aproximación en el sentido mínimo-cuadrático es la proyección ortogonal de la función en un subespacio de dimensión finita. Estos espacios y subespacios se construyen mediante la teoría de Sturm-Liouville.

En la sección anterior hemos considerados aproximaciones continuas mediante mínimos cuadrados, sin embargo, reemplazando las integrales por sumas también podemos obtener aproximaciones discretas. Estas aproximaciones se basan en minimizar el error

$$\langle f - p, f - p \rangle_r = \sum_{n=1}^N (f(x_n) - p(x_n))^2 r(x_n), \quad r(x_n) > 0.$$

Omitiremos detalles adicionales.

6.6.5 *Aproximación polinómica uniforme ¡¡¡NO!!!

Denotemos Π_n el espacio de los polinomios de grado n a lo sumo. Un polinomio $p_n^* \in \Pi_n$ es la mejor aproximación a una función f en norma uniforme si

$$\|f - p_n^*\|_\infty = \min_{p \in \Pi_n} \|f - p\|_\infty = \text{dist}_\infty(f, \Pi_n),$$

donde, como ya sabemos,

$$\|f - p\|_\infty = \max_{a \leq x \leq b} |f(x) - p(x)|,$$

y por ello a este tipo de aproximación se la denomina minimax.

Podemos obtener una cota inferior para la distancia uniforme mínima. Recordamos de secciones anteriores que las diferencias divididas de Newton se pueden escribir

$$f[x_0, x_1, \dots, x_{n+1}] = \sum_{i=0}^{n+1} \frac{f(x_i)}{W'(x_i)},$$

donde

$$W(x) = (x - x_0) \cdots (x - x_{n+1}),$$

cumpliendo que

$$f[x_0, x_1, \dots, x_{n+1}] = f[x_0, x_1, \dots, x_{n+1}] - p[x_0, x_1, \dots, x_{n+1}] = \sum_{i=0}^{n+1} \frac{f(x_i) - p(x_i)}{W'(x_i)},$$

ya que para $p \in \Pi_n$,

$$p[x_0, x_1, \dots, x_{n+1}] = \sum_{i=0}^{n+1} \frac{p(x_i)}{W'(x_i)} = 0.$$

De esta forma podemos acotar la diferencia dividida por

$$|f[x_0, x_1, \dots, x_{n+1}]| = \left| \sum_{i=0}^{n+1} \frac{f(x_i) - p(x_i)}{W'(x_i)} \right| \leq \|f - p\|_\infty W(x_0, x_1, \dots, x_{n+1}),$$

donde

$$W(x_0, x_1, \dots, x_{n+1}) = \sum_{i=0}^{n+1} \frac{1}{|W'(x_i)|}.$$

La desigualdad anterior se cumple incluso para el polinomio p^* que es el que mejor aproxima en norma uniforme a la función f y entonces tenemos la cota

$$\|f - p^*\|_\infty = \text{dist}_\infty(f, \Pi_n) \geq \frac{|f[x_0, x_1, \dots, x_{n+1}]|}{W(x_0, x_1, \dots, x_{n+1})}.$$

Esta cota inferior tiene el problema de que requiere calcular W , lo que, en general es bastante costoso.

Para evitar el cálculo de $W(x_0, \dots, x_{n+1})$ podemos utilizar el siguiente teorema, aunque requiere conocer una aproximación $p \in \Pi_n$ de signo alternante.

Teorema de Vallée-Poussin. Si $f(x) - p(x)$ con $p \in \Pi_n$ cambia de signo en los puntos $a \leq x_0 < x_1 < \dots < x_{n+1} \leq b$, es decir,

$$(-1)^i (f(x_i) - p(x_i)) \operatorname{sign}(f(x_0) - p(x_0)) \geq 0, \quad i = 0, 1, \dots, n+1,$$

entonces

$$\operatorname{dist}_\infty(f, \Pi_n) \geq \min_i |f(x_i) - p(x_i)|.$$

Este teorema implica que $(-1)^{n+1-i} W'(x_i) > 0$, por lo que

$$(f - p)[x_0, \dots, x_{n+1}] = \sum_{i=0}^{n+1} \frac{f(x_i) - p(x_i)}{W'(x_i)}$$

BIBLIOGRAFÍA

- [1] David Kincaid, Ward Cheney, “Análisis Numérico. Las matemáticas del cálculo científico,” Addison-Wesley Iberoamericana, 1994.
- [2] Francis B. Hildebrand, “Introduction to Numerical Analysis,” (2nd edition), McGraw-Hill, Inc., 1974 (republished by Dover Publications, Inc., 1987).