

1 Teoria da interpolação

1.1 Interpolação polinomial de Lagrange

Seja f uma função definida num intervalo $[a, b]$. Sejam x_0, x_1, \dots, x_n $n + 1$ pontos distintos de $[a, b]$ e defina-se $f_j := f(x_j)$, $j = 0, 1, \dots, n$. O problema que estudaremos será o de determinar uma função g tal que

$$g(x_j) = f_j, \quad j = 0, 1, \dots, n. \quad (1)$$

Os pontos x_j são chamados *pontos de interpolação* ou *nós* e os valores f_j os *valores interpolados*. Uma função g que verifique (1) é chamada *função interpoladora*. Em particular, vão ser do nosso interesse funções interpoladoras polinomiais.

Teorema 1.1.1 *Sejam x_0, x_1, \dots, x_n $n + 1$ pontos distintos do intervalo $[a, b]$ e sejam f_0, f_1, \dots, f_n os correspondentes valores de uma função f nesses pontos. Existe um único polinómio p_n de grau $\leq n$ que interpola f nos pontos $x_j, j = 0, 1, \dots, n$, ou seja, tal que:*

$$p_n(x_j) = f_j, \quad j = 0, 1, \dots, n. \quad (2)$$

Demonstração 1

Seja

$$p_n(x) := a_0 + a_1x + a_2x^2 + \dots + a_nx^n, \quad (3)$$

onde os coeficientes a_j são constantes a determinar. As condições (2) traduzem-se nas $n + 1$ equações

$$p_n(x_j) = a_0 + a_1x_j + a_2x_j^2 + \dots + a_nx_j^n, \quad j = 0, 1, \dots, n.$$

Estas constituem um sistema da forma $\mathbf{V}\mathbf{a} = \mathbf{f}$, onde \mathbf{V} é a matriz definida por

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix}$$

e onde $\mathbf{a} = [a_0 a_1 \cdots a_n]^T$ e $\mathbf{f} = [f_0 f_1 \cdots f_n]^T$. A matriz \mathbf{V} é uma matriz de Vandermonde e pode verificar-se que o seu determinante é dado por

$$\det \mathbf{V} = \prod_{0 \leq j < i \leq n} (x_i - x_j)$$

Sendo os x_j distintos, será $\det \mathbf{V} \neq 0$. Então \mathbf{V} é não singular e o sistema $\mathbf{V}\mathbf{a} = \mathbf{f}$ tem uma solução única \mathbf{a} . Fica provada a existência dum único polinómio interpolador de grau $\leq n$. Para se obter os coeficientes a_0, a_1, \dots, a_n que devem ser inseridos em (3), é necessário resolver o sistema $\mathbf{V}\mathbf{a} = \mathbf{f}$. Este processo de determinar uma expressão para p_n é pouco conveniente do ponto de vista computacional. Outros modos de construir p_n serão tratados neste capítulo.

Demonstração 2

A seguinte demonstração alternativa fornece uma expressão explícita para o polinómio interpolador.

Existência

Para cada $0 \leq j \leq n$, seja l_j o polinómio de grau n definido por

$$l_j(x) := \frac{(x - x_0)(x - x_1) \cdots (x - x_{j-1})(x - x_{j+1}) \cdots (x - x_n)}{(x_j - x_0)(x_j - x_1) \cdots (x_j - x_{j-1})(x_j - x_{j+1}) \cdots (x_j - x_n)}.$$

Os l_j são chamados *polinómios de Lagrange*.

Facilmente se verifica que

$$l_j(x_i) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \quad (4)$$

Vamos provar que p_n pode ser representado na seguinte forma

$$p_n(x) = \sum_{j=0}^n f_j l_j(x).$$

Comecemos por notar que a soma de polinómios de grau n é um polinómio de grau $\leq n$. Além disso, atendendo a (4) tem-se, para $0 \leq i \leq n$,

$$\begin{aligned} p_n(x_i) &= f_0 l_0(x_i) + \cdots + f_n l_n(x_i) \\ &= f_i l_i(x_i) = f_i, \end{aligned}$$

o que prova a existência dum polinómio interpolador.

Unicidade

Para provar a unicidade, suponhamos que existe um outro polinómio $q_n \in \mathcal{P}_n$, tal que $q_n(x_j) = p_n(x_j) = f_j$, $0 \leq j \leq n$. Então $r := q_n - p_n$ está ainda em \mathcal{P}_n e r tem $n + 1$ zeros. Ter-se-á necessariamente $r(x) \equiv 0$, donde $q_n \equiv p_n$, o que completa a demonstração 2.

A fórmula

$$p_n(x) = \sum_{j=0}^n f_j l_j(x), \quad (5)$$

com

$$l_j(x) := \prod_{\substack{0 \leq i \leq n \\ i \neq j}} \frac{(x - x_i)}{(x_j - x_i)} \quad (6)$$

é chamada *fórmula de Lagrange* para o polinómio interpolador.

Exemplo 1.1.1 Determinar o polinómio interpolador de grau ≤ 2 duma certa função f nos seguintes pontos tabelados

$$\begin{array}{c|ccc} x_i & 1 & -1 & 2 \\ \hline f(x_i) & 0 & -3 & 4 \end{array}$$

Tem-se

$$\begin{aligned} l_0(x) &= \frac{(x+1)(x-2)}{(1+1)(1-2)} = \frac{x^2 - x - 2}{-2} \\ l_1(x) &= \frac{(x-1)(x-2)}{(-1-1)(-1-2)} = \frac{x^2 - 3x + 2}{6} \\ l_2(x) &= \frac{(x-1)(x+1)}{(2-1)(2+1)} = \frac{x^2 - 1}{3} \end{aligned}$$

A fórmula (5) dá-nos então

$$p_2(x) = 0 - 3 \frac{x^2 - 3x + 2}{6} + 4 \frac{x^2 - 1}{3}.$$

Se quisermos representar p_2 em potências de x , basta-nos efectuar os cálculos na expressão obtida. Tem-se, neste caso, $p_2(x) = \frac{1}{6}(5x^2 + 9x - 14)$, que está na forma (3).

É importante distinguir entre a *função* p_n e as várias *representações* de p_n . Pelo teorema anterior, sabemos que p_n é único, para cada conjunto de $n + 1$ valores. Contudo, poderá haver muitas maneiras de representar explicitamente p_n . Cada uma dessas *fórmulas* sugere um dado algoritmo para calcular p_n . Na secção 1.3, estudaremos outra representação para p_n .

1.2 Erro de interpolação polinomial

Como pretendemos usar o polinómio interpolador para aproximar a função f em pontos distintos dos *pontos de interpolação* x_j , interessa-nos uma estimativa da diferença $f(x) - p_n(x)$, para $x \in [a, b]$.

Teorema 1.2.1 *Seja $p_n \in \mathcal{P}_n$ o polinómio interpolador de f em $n + 1$ pontos distintos x_0, x_1, \dots, x_n de $[a, b]$. Se $f \in C^{n+1}[a, b]$, então para cada $x \in [a, b]$ existe um ponto $\xi = \xi(x)$ pertencente ao intervalo $I_p(x_0, x_1, \dots, x_n, x) := (\min\{x_0, x_1, \dots, x_n, x\}, \max\{x_0, x_1, \dots, x_n, x\})$ tal que*

$$e_n(x) := f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} W(x), \quad (7)$$

onde

$$W(x) := (x - x_0)(x - x_1) \cdots (x - x_n).$$

Demonstração

Se $x = x_j$, para algum j , o resultado do teorema é trivialmente válido. Seja x um valor diferente dos x_j e considere-se a função auxiliar F definida por

$$F(t) := f(t) - p_n(t) - CW(t), \quad (8)$$

onde

$$C = \frac{f(x) - p_n(x)}{W(x)}.$$

Tem-se

$$F(x_j) = f(x_j) - p_n(x_j) - CW(x_j) = 0, \quad j = 0, 1, \dots, n,$$

e também

$$F(x) = f(x) - p_n(x) - CW(x) = 0,$$

atendendo à definição de C . A função F tem pelo menos $n + 2$ zeros distintos em $[a, b]$. Pelo teorema de Rolle, a derivada F' tem pelo menos $n + 1$ zeros no intervalo $I_p(x_0, x_1, \dots, x_n, x)$. A segunda derivada terá pelo menos n zeros e, por um processo indutivo, concluímos que a derivada de ordem $(n + 1)$ tem pelo menos um zero. Seja ξ esse zero. Derivando $n + 1$ vezes a equação (8) e fazendo $t = \xi$, obtemos

$$0 = F^{(n+1)}(\xi) = f^{(n+1)}(\xi) - C(n + 1)!,$$

o que, depois de substituir C pela sua expressão, conduz ao resultado pretendido.

É evidente que a equação (7) não pode ser usada para calcular o valor exacto do erro $e_n(x)$, visto que $\xi = \xi(x)$ é em geral uma função desconhecida. (Uma excepção é o caso em que a derivada de ordem $(n + 1)$ de f é uma constante).

Porém, sendo $f^{(n+1)}$ contínua, existirá uma constante M_{n+1} tal que $|f^{(n+1)}(\xi)| \leq M_{n+1}$ para qualquer $\xi \in [a, b]$. Isto, juntamente com a equação (7), resulta na fórmula

$$|e_n(x)| = |f(x) - p_n(x)| \leq \frac{M_{n+1}}{(n + 1)!} |W(x)| \quad (9)$$

Enunciamos ainda um corolário do teorema anterior.

Corolário 1.1.1 Nas condições do Teorema 1.2.1, a função $f^{(n+1)}(\xi)$ pode ser prolongada como uma função contínua de x , para $x \in [a, b]$.

1.3 Fórmula interpoladora de Newton com diferenças divididas

Uma outra forma de representar o polinómio interpolador é usando diferenças divididas, que a seguir definiremos. Este método, devido a Isaac Newton (1642-1727), permite passar dum polinómio interpolador dum certo grau $n - 1$ para o polinómio de grau superior n , a partir do polinómio de grau inferior.

Mais precisamente, sejam $p_{n-1} \in \mathcal{P}_{n-1}$ e $p_n \in \mathcal{P}_n$ os polinómios interpoladores de f nos pontos x_i , para $i = 0, 1, \dots, n - 1$ e $i = 0, 1, \dots, n$, respectivamente. Vamos provar que é possível escrever

$$p_n(x) = p_{n-1}(x) + C_n(x), \quad (10)$$

onde C_n é uma função que a seguir caracterizamos.

Como $C_n \in \mathcal{P}_n$ e $C_n(x_i) = 0, i = 0, 1, \dots, n - 1$, então C_n é da forma

$$C_n(x) = a_n(x - x_0)(x - x_1) \cdots (x - x_{n-1})$$

com a_n constante real. Substituindo em (10), vem

$$p_n(x) = p_{n-1}(x) + a_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}). \quad (11)$$

Começando com $p_0(x) = f_0$, obtém-se desta relação

$$p_1(x) = f_0 + a_1(x - x_0) \quad (12)$$

$$\begin{aligned} p_2(x) &= p_1(x) + a_2(x - x_0)(x - x_1) \\ &= f_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) \end{aligned} \quad (13)$$

$$\begin{aligned} p_3(x) &= p_2(x) + a_3(x - x_0)(x - x_1)(x - x_2) \\ &= f_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + a_3(x - x_0)(x - x_1)(x - x_2) \end{aligned} \quad (14)$$

\vdots

$$\begin{aligned} p_n(x) &= f_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + a_3(x - x_0)(x - x_1)(x - x_2) + \cdots \\ &+ a_n(x - x_0) \cdots (x - x_{n-1}) \end{aligned} \quad (15)$$

Substituindo x por x_1, x_2, \dots, x_n , respectivamente nas equações (12), ..., (15), obtemos um sistema de equações (de matriz triangular) nas incógnitas a_1, \dots, a_n . É fácil de verificar que este sistema tem uma solução única.

Assim, de $p_1(x_1) = f_1$, sai que

$$a_1 = \frac{f_1 - f_0}{x_1 - x_0}, \quad (16)$$

que, substituído na equação $p_2(x_2) = f_2$, permite obter a_2 . E assim sucessivamente, até obter a_n .

Podemos concluir de (15) que a_n é o coeficiente do termo em x^n de p_n . Para obtermos uma expressão para a_n , voltemos à fórmula de Lagrange

$$p_n(x) = \sum_{j=0}^n \frac{(x-x_0)(x-x_1)\cdots(x-x_{j-1})(x-x_{j+1})\cdots(x-x_n)}{(x_j-x_0)(x_j-x_1)\cdots(x_j-x_{j-1})(x_j-x_{j+1})\cdots(x_j-x_n)} f_j, \quad (17)$$

que se pode escrever na forma

$$p_n(x) = \sum_{j=0}^n \frac{f_j}{\prod_{\substack{0 \leq i \leq n \\ i \neq j}} (x_j - x_i)} [x^n + p_{n-1,j}^*].$$

Então

$$a_n = \sum_{j=0}^n \frac{f_j}{\prod_{\substack{0 \leq i \leq n \\ i \neq j}} (x_j - x_i)} \quad (18)$$

Para realçar a dependência de a_n em relação aos valores $x_i, f_i, 0 \leq i \leq n$, usa-se a seguinte notação para designar o valor a_n ,

$$f[x_0, x_1, \dots, x_n] := a_n \quad (19)$$

a que se chama *diferença dividida de ordem n da função f , associada aos pontos x_0, x_1, \dots, x_n* .

Com a notação das diferenças divididas, podemos escrever o polinómio interpolador na forma

$$p_n(x) = f[x_0] + f[x_0, x_1](x-x_0) + f[x_0, x_1, x_2](x-x_0)(x-x_1) + \cdots + f[x_0, x_1, \dots, x_n](x-x_0)\cdots(x-x_{n-1}) \quad (20)$$

a que se chama *fórmula de Newton com diferenças divididas*. Por uma questão de uniformidade, no lugar de $f(x_i)$ usa-se $f[x_i]$, a que se chama diferença de ordem zero.

Exemplos

Resulta de (18) e (19) que

$$f[x_0, x_1] = \frac{f_0}{x_0 - x_1} + \frac{f_1}{x_1 - x_0}$$

o que confirma a expressão já obtida em (16). De modo geral, a diferença dividida de ordem 1, nos pontos x_i, x_{i+1} , é dada por

$$f[x_i, x_{i+1}] = \frac{f_{i+1} - f_i}{x_{i+1} - x_i}.$$

Como exemplo de diferença de ordem 2, tem-se

$$f[x_0, x_1, x_2] = \frac{f_0}{(x_0 - x_1)(x_0 - x_2)} + \frac{f_1}{(x_1 - x_0)(x_1 - x_2)} + \frac{f_2}{(x_2 - x_0)(x_2 - x_1)}$$

A seguinte propriedade permite relacionar diferenças de uma certa ordem k com diferenças de ordem $k - 1$.

$$f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i} \quad (21)$$

Resulta também de (18) que $f[x_0, \dots, x_n]$ é uma função simétrica de x_0, \dots, x_n , isto é, qualquer permutação dos x_i conduz ao mesmo valor da diferença dividida.

A fórmula (21) permite calcular as diferenças divididas usando um esquema recursivo que é em geral apresentado sob a forma duma tabela. Na 1ª coluna escrevem-se os pontos $x_i, 0 \leq i \leq n$ e na 2ª coluna os valores $f(x_i), 0 \leq i \leq n$. Depois de obter a terceira coluna da tabela, ou seja, as diferenças de 1ª ordem, por aplicação de (21) podem obter-se as diferenças de 2ª ordem na coluna seguinte, e assim sucessivamente. Por exemplo, depois de se calcular

$$\begin{aligned} f[x_0, x_1] &= \frac{f[x_1] - f[x_0]}{x_1 - x_0}, & f[x_1, x_2] &= \frac{f[x_2] - f[x_1]}{x_2 - x_1} \\ f[x_2, x_3] &= \frac{f[x_3] - f[x_2]}{x_3 - x_2}, & f[x_3, x_4] &= \frac{f[x_4] - f[x_3]}{x_4 - x_3} \end{aligned}$$

pode-se obter

$$\begin{aligned} f[x_0, x_1, x_2] &= \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} \\ f[x_1, x_2, x_3] &= \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1} \end{aligned}$$

1.4 Relação entre as diferenças divididas e as derivadas de f

Seja $f \in C^{n+1}[a, b]$ e sejam x_0, x_1, \dots, x_n pontos distintos de $[a, b]$. Usando a fórmula de Newton com diferenças divididas, podemos escrever a igualdade

$$f(x) = f[x_0] + f[x_0, x_1](x - x_0) + \dots + f[x_0, x_1, \dots, x_n](x - x_0) \dots (x - x_{n-1}) + \frac{f^{(n+1)}(\xi)}{(n+1)!} W(x), \quad (22)$$

Table 1: Tabela de diferenças divididas

x_i	$f[x_i]$	$f[x_i, x_{i+1}]$	$f[, ,]$	$f[, , ,]$	$f[, , , ,]$
x_0	$f[x_0]$				
		$f[x_0, x_1]$			
x_1	$\cdots f[x_1]$		$f[x_0, x_1, x_2]$		
		\ddots		$f[x_0, x_1, x_2, x_3]$	
		$f[x_1, x_2]$	\ddots		
x_2	$f[x_2]$		$f[x_1, x_2, x_3]$		$f[x_0, x_1, x_2, x_3, x_4]$
		\ddots		$f[x_1, x_2, x_3, x_4]$	
		$f[x_2, x_3]$	\ddots		
x_3	$\cdots f[x_3]$		$f[x_2, x_3, x_4]$		
		\ddots			
		$f[x_3, x_4]$			
x_4	$f[x_4]$				

onde $W(x) = (x - x_0)(x - x_1) \cdots (x - x_n)$ e $\xi \in I_p(x_0, \dots, x_n, x)$. O último termo do lado direito da equação representa o erro de interpolação (ver eq. (7)).

Comparando (22) com a fórmula de Taylor

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots + f^{(n)}(x_0) \frac{(x - x_0)^n}{n!} + f^{(n+1)}(\gamma) \frac{(x - x_0)^{n+1}}{(n+1)!}, \quad \gamma \text{ entre } x_0 \text{ e } x,$$

constatamos uma certa semelhança entre as duas fórmulas, sugerindo a *existência duma relação entre as diferenças divididas e as derivadas de f* . Com efeito, basta notar por exemplo que sendo

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

resulta do Teorema do valor médio de Lagrange que $f[x_0, x_1]$ coincidirá com o valor da derivada f' nalgum ponto situado entre x_0 e x_1 .

Vamos agora generalizar este resultado às diferenças divididas de ordem superior e, simultaneamente, obter uma fórmula para o erro de interpolação expresso em diferenças divididas.

Juntemos aos $n+1$ pontos x_0, \dots, x_n um outro ponto qualquer do intervalo $[a, b]$ (distinto dos x_i), que designaremos por \hat{x} . A estes $n+2$ pontos está associado um único polinómio interpolador de f de grau $\leq n+1$. Designe-se por p_{n+1} .

Usando a fórmula de recorrência (11), com $n + 1$ no lugar de n , podemos escrever

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

e, como se tem $p_{n+1}(\hat{x}) = f(\hat{x})$ pela definição de polinómio interpolador, obtém-se que

$$f(\hat{x}) = p_n(\hat{x}) + f[x_0, \dots, x_n, \hat{x}](\hat{x} - x_0) \cdots (\hat{x} - x_n)$$

donde resulta a seguinte fórmula para o erro de interpolação

$$f(\hat{x}) - p_n(\hat{x}) = f[x_0, \dots, x_n, \hat{x}](\hat{x} - x_0) \cdots (\hat{x} - x_n) \quad (23)$$

Por outro lado, sabemos que

$$f(\hat{x}) - p_n(\hat{x}) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(\hat{x} - x_0) \cdots (\hat{x} - x_n), \quad \xi \in I_p(x_0, \dots, x_n, \hat{x}). \quad (24)$$

Comparando as duas equações anteriores, segue que

$$f[x_0, x_1, \dots, x_n, \hat{x}] = \frac{f^{(n+1)}(\xi)}{(n+1)!}, \quad \xi \in I_p(x_0, \dots, x_n, \hat{x}). \quad (25)$$

Esta fórmula é válida quaisquer que sejam os $n + 2$ pontos distintos considerados. Por facilidade de notação, podemos designar \hat{x} por x_{n+1} e (25) pode-se escrever como

$$f[x_0, x_1, \dots, x_{n+1}] = \frac{f^{(n+1)}(\xi)}{(n+1)!}, \quad \xi \in I_p(x_0, \dots, x_n, x_{n+1}). \quad (26)$$

Nota

Conclui-se de (26) que se $f \in \mathcal{P}_n$, as diferenças divididas de f de ordem superior a n são zero.

1.5 Interpolação em pontos igualmente espaçados. Fórmula de Newton com diferenças progressivas

Até aqui considerámos o problema de interpolar uma dada função f , conhecendo-se o seu valor em $n + 1$ pontos distintos quaisquer. Vamos agora supor que esses pontos estão ordenados e são igualmente espaçados, com espaçamento (ou passo) h . Mais precisamente, sejam

$$x_0 < x_1 < x_2 < \cdots < x_n$$

pontos do intervalo $[a, b]$, verificando $x_j = x_0 + jh$, $0 \leq j \leq n$. Veremos que a tabela de diferenças divididas pode ser simplificada, bem como a fórmula de Newton baseada nessas diferenças.

Definição 1.5.1 Dada uma função f e um passo h fixo, definimos o operador de diferenças progressivas (ou descendentes) Δ do seguinte modo

$$\Delta f(x) := f(x+h) - f(x).$$

Por exemplo, se $f(x) = \cos x$ e $h = 0.1$, então

$$\Delta f(x) = \cos(x+0.1) - \cos x.$$

Usaremos a notação $\Delta f_i := \Delta f(x_i)$. Tem-se assim

$$\begin{aligned} \Delta f_0 &= f(x_0+h) - f(x_0) = f_1 - f_0 \\ \Delta f_1 &= f_2 - f_1 \\ \text{e, em geral } \Delta f_i &= f_{i+1} - f_i. \end{aligned}$$

Designaremos por $\Delta^2 f(x)$ as diferenças progressivas de 2.^a ordem, definidas por

$$\Delta^2 f(x) := \Delta(\Delta f(x)),$$

ou seja,

$$\Delta^2 f(x) = \Delta(f(x+h) - f(x)) = f(x+2h) - 2f(x+h) + f(x).$$

Assim, por exemplo,

$$\begin{aligned} \Delta^2 f_0 &= f_2 - 2f_1 + f_0 \\ \Delta^2 f_1 &= f_3 - 2f_2 + f_1. \end{aligned}$$

Analogamente se definem as diferenças progressivas de 3.^a ordem

$$\Delta^3 f(x) := \Delta(\Delta^2 f(x)) = f(x+3h) - 3f(x+2h) + 3f(x+h) - f(x).$$

Em particular, $\Delta^3 f_0 = f_3 - 3f_2 + 3f_1 - f_0$. Por um processo recursivo, definimos de modo geral

$$\Delta^k f(x) := \Delta(\Delta^{k-1} f(x)), \quad k = 1, 2, \dots \quad (27)$$

$$\text{com } \Delta^0 f(x) := f(x). \quad (28)$$

Prova-se que

$$\Delta^k f(x) = \sum_{i=0}^k \binom{k}{i} (-1)^i f(x + (k-i)h), \quad (29)$$

onde

$$\binom{k}{i} := \frac{k!}{i!(k-i)!}, \quad \binom{0}{0} = \binom{1}{1} = 1.$$

Tem-se a seguinte relação entre diferenças divididas e diferenças progressivas

$$f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{\Delta^k f_i}{k!h^k}. \quad (30)$$

Table 2: Diferenças divididas expressas em termos de diferenças progressivas

x_0	f_0				
			$\frac{1}{h}\Delta f_0$		
x_1	$\cdots f_1$			$\frac{1}{2!h^2}\Delta^2 f_0$	
		\ddots	$\frac{1}{h}\Delta f_1$		$\frac{1}{3!h^3}\Delta^3 f_0$
x_2	f_2			$\frac{1}{2!h^2}\Delta^2 f_1$	$\frac{1}{4!h^4}\Delta^4 f_0$
		\ddots	$\frac{1}{h}\Delta f_2$		$\frac{1}{3!h^3}\Delta^3 f_1$
x_3	$\cdots f_3$			$\frac{1}{2!h^2}\Delta^2 f_2$	
		\ddots	$\frac{1}{h}\Delta f_3$		
x_4	f_4				

Com a notação das diferenças progressivas, a tabela de diferenças divididas toma a forma da tabela 2.

Por outro lado, podemos reescrever a fórmula (20) como

$$\begin{aligned}
 p_n(x) &= \Delta^0 f_0 + \frac{\Delta f_0}{h}(x - x_0) + \frac{\Delta^2 f_0}{2!h^2}(x - x_0)(x - x_1) + \frac{\Delta^3 f_0}{3!h^3}(x - x_0)(x - x_1)(x - x_2) + \dots \\
 &+ \frac{\Delta^n f_0}{n!h^n}(x - x_0)(x - x_1) \dots (x - x_{n-1}). \tag{31}
 \end{aligned}$$

Efectuando a mudança de variável $x = x_0 + sh$ e denotando

$$\binom{s}{j} := \frac{s(s-1)\dots(s-j+1)}{j!}$$

obtém-se de (31)

$$p_n(x_0 + sh) = f_0 + \binom{s}{1} \Delta f_0 + \binom{s}{2} \Delta^2 f_0 + \dots + \binom{s}{n} \Delta^n f_0, \tag{32}$$

a que se chama *fórmula de Newton com diferenças progressivas*. Este resultado foi devido a James Gregory (1638-1675), um contemporâneo escocês de Isaac Newton. Notemos que as diferenças progressivas podem ser calculadas a partir da seguinte tabela.

Table 3: Tabela de diferenças progressivas

x_i	f_i	Δf_i	$\Delta^2 f_i$	$\Delta^3 f_i$	$\Delta^4 f_i$
x_0	f_0				
		Δf_0			
x_1	$\cdots f_1$		$\Delta^2 f_0$		
		\ddots		$\Delta^3 f_0$	
		Δf_1			
x_2	f_2		$\Delta^2 f_1$		$\Delta^4 f_0$
		\ddots		$\Delta^3 f_1$	
		Δf_2			
x_3	$\cdots f_3$		$\Delta^2 f_2$		
		\ddots			
		Δf_3			
x_4	f_4				

1.6 Considerações sobre o erro de interpolação de Lagrange. Nós de Chebyshev

Consideremos de novo a fórmula de majoração do erro de interpolação (equação (9))

$$|e_n(x)| = |f(x) - p_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |W(x)|,$$

onde

$$W(x) = (x - x_0)(x - x_1) \cdots (x - x_n).$$

e

$$|f^{(n+1)}(x)| \leq M_{n+1}, \forall x \in [a, b].$$

Para cada valor de x , a fórmula anterior dá-nos um majorante para o erro cometido ao aproximar $f(x)$ por $p_n(x)$. Contudo se pretendermos um majorante para $|e_n(x)|$ válido para todos os valores de $x \in [a, b]$, devemos calcular

$$\|W\|_\infty := \max_{x \in [a, b]} |W(x)|,$$

o que pode ser feito numericamente, através da determinação dos zeros de $|W'(x)|$, e obtém-se a desigualdade

$$|e(x)| = |f(x) - p_n(x)| \leq \frac{M_{n+1}}{(n+1)!} \|W\|_\infty, \quad \forall x \in [a, b]. \quad (33)$$

Ao aplicarmos (9), é de notar-se a importância da contribuição de $|W(x)|$. É claro que, para x perto de cada ponto de interpolação x_j , $|W(x)|$ será pequeno. Vamos começar por analisar o caso especial em que os pontos de interpolação (ou nós) são igualmente espaçados.

1.6.1 Nós igualmente espaçados

Consideremos os pontos $x_0 < x_1 < \dots < x_n$ no intervalo $[a, b]$ tal que

$$x_i = x_0 + ih, \quad i = 0, 1, \dots, n. \quad (34)$$

Com a mudança de variável $x = x_0 + th$, $0 \leq t \leq n$, a fórmula (33) pode-se escrever como

$$|e(x)| \leq \frac{M_{n+1}}{(n+1)!} \max_{t \in [0, n]} |W(x_0 + th)|, \quad (35)$$

com

$$W(x_0 + th) = h^{n+1} \psi_n(t) \quad (36)$$

e onde

$$\psi_n(t) := \begin{cases} t & n = 0 \\ t(t-1) \dots (t-n) & n = 1, 2, \dots \end{cases}$$

A função $\psi_n(t)$ é um polinómio de grau $n+1$ na variável t e anula-se para os $n+1$ valores reais $0, 1, \dots, n$. São válidas as seguintes propriedades de simetria em relação ao ponto $t = n/2$, que é o ponto médio do intervalo $[0, n]$.

Teorema 1.6.1 *Para n ímpar,*

$$\psi_n\left(\frac{n}{2} - \tau\right) = \psi_n\left(\frac{n}{2} + \tau\right), \quad (37)$$

ou seja, $\psi_n(t)$ é simétrica em relação a $t = n/2$; para n par

$$\psi_n\left(\frac{n}{2} - \tau\right) = -\psi_n\left(\frac{n}{2} + \tau\right), \quad (38)$$

ou seja, $\psi_n(t)$ é anti-simétrica em relação a $t = n/2$.

O seguinte resultado dá uma comparação dos valores de $\psi_n(t)$ com $\psi_n(t+1)$.

Teorema 1.6.2

a) *Se $t+1$ é um valor não inteiro, com $1 < t+1 < n/2$,*

$$|\psi_n(t+1)| < |\psi_n(t)| \quad (39)$$

b) *Se t é um valor não inteiro, com $n/2 \leq t < n-1$,*

$$|\psi_n(t)| < |\psi_n(t+1)|. \quad (40)$$

As propriedades enunciadas nos teoremas 1.6.1 e 1.6.2 estão ilustradas nas figuras 1.6.1 e 1.6.2, onde se apresentam os gráficos de $\psi_n(t)$, para vários valores de n . Concluimos, em particular, que os valores dos máximos relativos de $|\psi_n(t)|$ decrescem à medida que t se aproxima de $n/2$. Então o máximo de $|\psi_n(t)|$, $0 \leq t \leq n$, ocorre no intervalo $(0, 1)$ (e, por simetria, também no intervalo $(n - 1, n)$). Daqui podemos obter a seguinte majoração

$$\max_{t \in [0, n]} |\psi_n(t)| < n! \tag{41}$$

Atendendo à fórmula (35) e ao que se disse sobre a função $\psi_n(t)$, na interpolação de $f(x)$ por $p_n(x)$ os pontos de interpolação (34) devem ser escolhidos por forma a que x esteja perto do ponto médio do intervalo $[x_0, x_n]$. Observamos ainda que $|\psi_n(t)|$ cresce rapidamente para $t < 0$ e $t > n$. Assim, se pretendermos aproximar o valor $f(x)$ por *extrapolação* (isto é, se x está fora do intervalo $[x_0, x_n]$), é de esperar que o erro cometido cresça muito rapidamente à medida que x se afasta quer de x_0 ou de x_n .

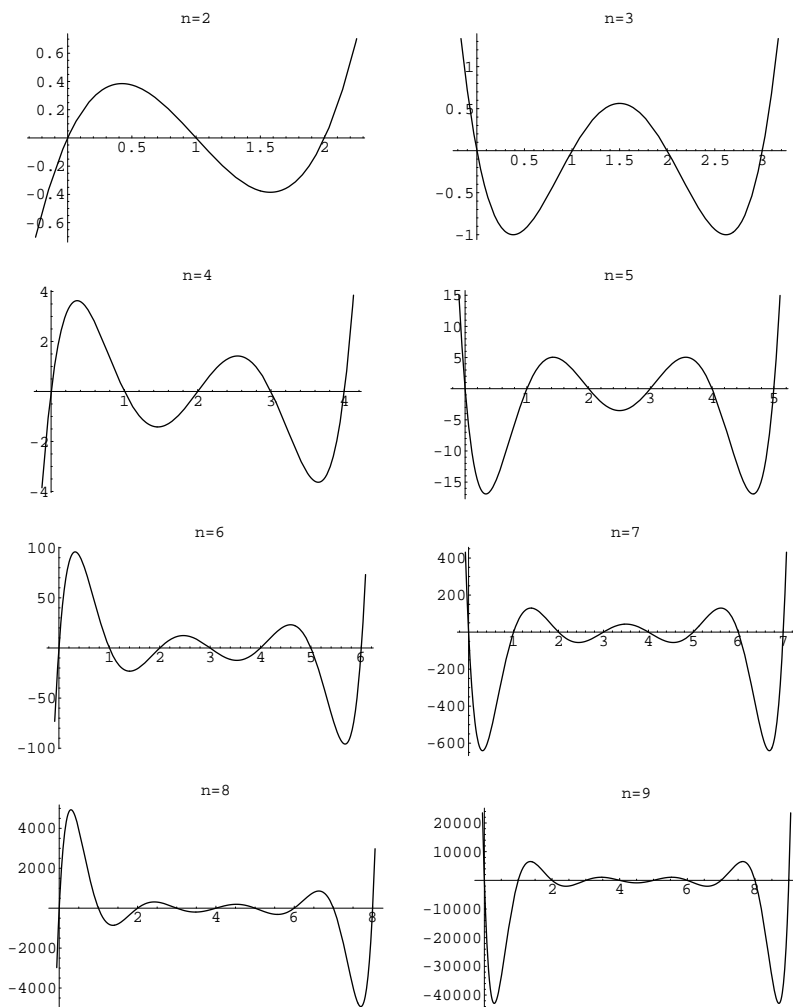


Fig 1.6.1 $\psi_n(n \text{ par})$;

Fig 1.6.2 $\psi_n(n \text{ ímpar})$

1.6.2 Nós de Chebyshev

Retomando a fórmula (33), vamos agora considerar o problema de escolher os pontos de interpolação $\{x_j\}_{j=0}^n$ por forma a minimizar $\|W\|_\infty$. Este problema foi tratado por P. L. Chebyshev (1821-1894) para o caso de $[a, b] = [-1, 1]$, e a solução envolve os chamados polinómios de Chebyshev, que definiremos a seguir.

À função

$$T_n(x) := \cos(n \arccos x), \quad n = 0, 1, 2, \dots, x \in [-1, 1] \quad (42)$$

chama-se polinómio de Chebyshev de grau n . Vejamos que se trata de facto de um polinómio. Resulta de (42) que

$$T_0(x) = 1, \quad T_1(x) = x.$$

Fazendo a substituição $x = \cos \theta, \theta \in [0, \pi]$, tem-se

$$T_n(x) = T_n(\cos \theta) = \cos(n\theta) =: t_n(\theta).$$

Por outro lado, fazendo uso de desigualdades trigonométricas elementares, obtém-se

$$T_{n+1}(x) = t_{n+1}(\theta) = \cos(n\theta + \theta) = \cos(n\theta) \cos(\theta) - \sin(n\theta) \sin(\theta), \quad n \geq 0,$$

$$T_{n-1}(x) = t_{n-1}(\theta) = \cos(n\theta - \theta) = \cos(n\theta) \cos(\theta) + \sin(n\theta) \sin(\theta), \quad n \geq 1.$$

Somando as duas equações anteriores, resulta

$$T_{n+1}(x) + T_{n-1}(x) = t_{n+1}(\theta) + t_n(\theta) = 2 \cos(n\theta) \cos(\theta),$$

ou seja, é válida a seguinte relação de recorrência

$$T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x), \quad n \geq 1. \quad (43)$$

Então, sendo T_0 e T_1 polinómios de graus zero e um, respectivamente, de (43) concluímos que T_2 é um polinómio de grau 2. Em geral, se T_{n-1} e T_n forem polinómios de graus $n-1$ e n , respectivamente, então da fórmula anterior vem que T_{n+1} é um polinómio de grau $n+1$. Dando valores a n , obtém-se

$$\begin{aligned} T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \\ &\vdots \end{aligned}$$

Provemos agora algumas propriedades dos polinómios de Chebyshev.

Propriedade 1 *Se $n \geq 1$, $T_n(x)$ tem n zeros simples em $[-1, 1]$ nos pontos*

$$x_k = \cos\left(\frac{2k+1}{2n} \pi\right), \quad k = 0, 1, \dots, n-1. \quad (44)$$

e é válida a seguinte factorização

$$T_n(x) = 2^{n-1} \prod_{k=0}^{n-1} \left[x - \cos\left(\frac{2k+1}{2n} \pi\right) \right]. \quad (45)$$

Demonstração Os zeros de T_n são os pontos x_k tais que

$$n \arccos x_k = \left(\frac{2k+1}{2}\right)\pi, \quad k = 0, 1, \dots, n-1,$$

donde

$$x_k = \cos\left(\frac{2k+1}{2n} \pi\right), \quad k = 0, 1, \dots, n-1.$$

Provemos agora, por indução, que o coeficiente do termo em x^n de $T_n(x)$ é 2^{n-1} , para $n \geq 1$. Se $n = 1$, $T_1(x) = x = 2^{1-1}x$. Seja agora a hipótese de indução que o coeficiente do termo em x^n de $T_n(x)$ é 2^{n-1} . De (43), segue imediatamente que o coeficiente de x^{n+1} é 2^n . Daqui, juntamente com (44), sai (45).

Propriedade 2 Os extremos de $T_n(x)$ para $x \in [-1, 1]$, são 1 e -1 , tomados alternadamente nos $n+1$ pontos

$$t_k = \cos \frac{k\pi}{n}, \quad k = 0, 1, \dots, n. \quad (46)$$

Demonstração Os extremos da função $\cos x$ ocorrem para $x = k\pi$ e têm o valor $(-1)^k$, $k \in \mathbf{Z}$. Então, atendendo a (42), os extremos de $T_n(x)$ para $x \in [-1, 1]$, ocorrem nos pontos t_k tais que

$$n \arccos t_k = k\pi, \quad k = 0, 1, \dots, n,$$

donde os t_k são dados por (46). Tendo-se $T_n(t_k) = \cos(k\pi) = (-1)^k$, $k = 0, 1, \dots, n$, conclui-se que t_k é ponto de *máximo* se k é par e ponto de *mínimo* se k é ímpar.

Teorema 1.6.3 *Definam-se*

$$\bar{T}_0(x) := 1 \quad (47)$$

$$\bar{T}_n(x) := \frac{1}{2^{n-1}} T_n(x), \quad n = 1, 2, \dots \quad (48)$$

e seja $\bar{\pi}_n$ o conjunto dos polinómios mónicos de grau n . Tem-se

$$\max_{x \in [-1, 1]} |\bar{T}_n(x)| \leq \max_{x \in [-1, 1]} |q_n(x)|, \quad \forall q_n \in \bar{\pi}_n. \quad (49)$$

Demonstração

Com efeito, \bar{T}_n anula-se nos pontos x_k definidos por (44) e os seus extremos são os valores

$$\bar{T}_n(t_k) = \frac{(-1)^k}{2^{n-1}}, \quad k = 0, 1, \dots, n,$$

donde

$$\|\overline{T}_n\|_\infty = \max_{x \in [-1,1]} |\overline{T}_n(x)| = \frac{1}{2^{n-1}}. \quad (50)$$

Para provar (49), suponhamos que existe $q_n \in \overline{\pi}_n$ tal que

$$\max_{x \in [-1,1]} |q_n(x)| < \frac{1}{2^{n-1}} \quad (51)$$

e seja $r_n(x) := \overline{T}_n(x) - q_n(x)$, $x \in [-1, 1]$. Nos pontos t_k dados por (46), tem-se

$$r_n(t_k) = \frac{(-1)^k}{2^{n-1}} - q_n(t_k) = (-1)^k \left[\frac{1}{2^{n-1}} - (-1)^k q_n(t_k) \right], \quad k = 0, 1, \dots, n.$$

Atendendo a (51), a quantidade entre parêntesis é sempre positiva. Então $r_n(t)$ tem sinal alternado nos $n + 1$ pontos t_0, t_1, \dots, t_n e, sendo r_n uma função contínua, terá pelo menos n zeros distintos. Por outro lado, r_n é um polinómio de grau $\leq n - 1$, visto tratar-se da diferença entre dois polinómios mónicos de grau n . Então só poderá ter-se $r_n \equiv 0$, donde $q_n(x) = \overline{T}_n(x)$, $x \in [-1, 1]$. Consequentemente,

$$\max_{x \in [-1,1]} |q_n(x)| = \frac{1}{2^{n-1}},$$

o que contradiz a hipótese (51). Fica assim provada a relação (49).

Estamos agora em condições de resolver o seguinte problema.

Seja $[a, b] = [-1, 1]$. É possível escolher pontos distintos x_0, x_1, \dots, x_n em $[-1, 1]$ de modo a que o factor $\|W\|_\infty$ da fórmula (33),

$$\|W\|_\infty = \max_{x \in [-1,1]} |(x - x_0)(x - x_1) \cdots (x - x_n)|,$$

seja o menor possível? A resposta a esta questão é dada por

Teorema 1.6.4 *De entre todas as escolhas dos pontos distintos x_0, x_1, \dots, x_n em $[-1, 1]$, $\|W\|_\infty$ é minimizado se*

$$x_k = \cos\left(\frac{2k+1}{2n+2}\pi\right), \quad k = 0, 1, \dots, n, \quad (52)$$

ou seja, se os x_k forem os zeros do polinómio de Chebyshev T_{n+1} . Se $f \in C^{(n+1)}[-1, 1]$, a fórmula (33) de majoração do erro de interpolação reescreve-se como:

$$|e(x)| = |f(x) - p_n(x)| \leq \frac{M_{n+1}}{(n+1)! 2^n} \quad (53)$$

Demonstração

Notamos que $W \in \overline{\pi}_{n+1}$, ou seja, W é um polinómio mónico de grau $n + 1$. Pelo teorema 1.6.3, sabemos que

$$\|\overline{T}_{n+1}\|_\infty \leq \|q_{n+1}\|_\infty, \quad \forall q_{n+1} \in \overline{\pi}_{n+1}.$$

Então a escolha dos x_k por forma que $W(x) = \overline{T}_{n+1}$ resolve o nosso problema de minimização. Mas isso corresponde a tomar os x_k dados por (52). Para obter (53), basta usar (50).

1.7 Sucessões de polinómios interpoladores

Contrariamente ao que se possa pensar, não é em geral verdade que o uso de polinómios interpoladores de graus mais elevados conduza a aproximações com erros menores.

Consideremos uma sucessão de polinómios interpoladores

$$p_1(x), p_2(x), p_3(x), \dots, p_n(x), \dots$$

de uma dada função f , tal que:

$$\begin{aligned} p_1(x) & \text{ interpola } f \text{ nos pontos } x_0^{(1)}, x_1^{(1)}, \\ p_2(x) & \text{ interpola } f \text{ nos pontos } x_0^{(2)}, x_1^{(2)}, x_2^{(2)}, \\ p_3(x) & \text{ interpola } f \text{ nos pontos } x_0^{(3)}, x_1^{(3)}, x_2^{(3)}, x_3^{(3)}, \\ & \dots \\ p_n(x) & \text{ interpola } f \text{ nos pontos } x_0^{(n)}, x_1^{(n)}, \dots, x_n^{(n)}, \\ & \dots \end{aligned}$$

Admita-se também que todos os pontos $x_0^{(n)}, \dots, x_n^{(n)}$, $n = 1, 2, \dots$ pertencem a um intervalo finito $[a, b]$. Consideremos a seguinte questão:

Em que condições se tem

$$\lim_{m \rightarrow \infty} p_n(x) = f(x), \quad (54)$$

para todo o $x \in [a, b]$?

Exemplo 1.7.1 Seja a função $f(x) := |x|$ e $[a, b] = [-1, 1]$. Suponhamos que os polinómios interpoladores são construídos usando pontos igualmente espaçados

$$x_k^{(n)} = -1 + k \frac{2}{n}, \quad k = 0, \dots, n,$$

ou seja,

$$\begin{aligned} x_0^{(1)} &= -1, x_1^{(1)} = 1, \\ x_0^{(2)} &= -1, x_1^{(2)} = 0, x_2^{(2)} = 1, \\ x_0^{(3)} &= -1, x_1^{(3)} = -\frac{1}{3}, x_2^{(3)} = \frac{1}{3}, x_3^{(3)} = 1, \\ x_0^{(4)} &= -1, x_1^{(4)} = -\frac{1}{2}, x_2^{(4)} = 0, x_3^{(4)} = \frac{1}{2}, x_4^{(4)} = 1 \\ & \dots \end{aligned} \quad (55)$$

Bernstein (1912) provou que para valores de x diferentes de $-1, 0, 1$, então $p_n(x)$ não converge para $f(x)$. As figuras 1.7.1-1.7.3 mostram os gráficos dos polinómios p_4 , p_{10} e p_{20} .

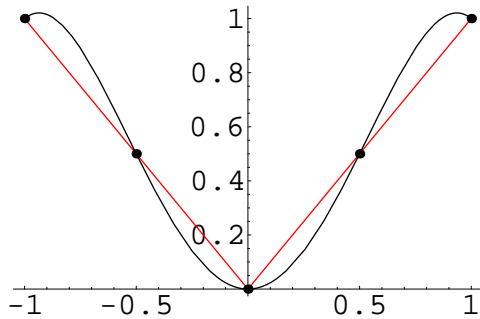


Fig. 1.7.1 p_4

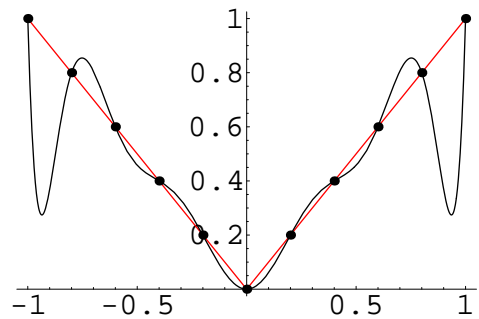


Fig. 1.7.2 p_{10}

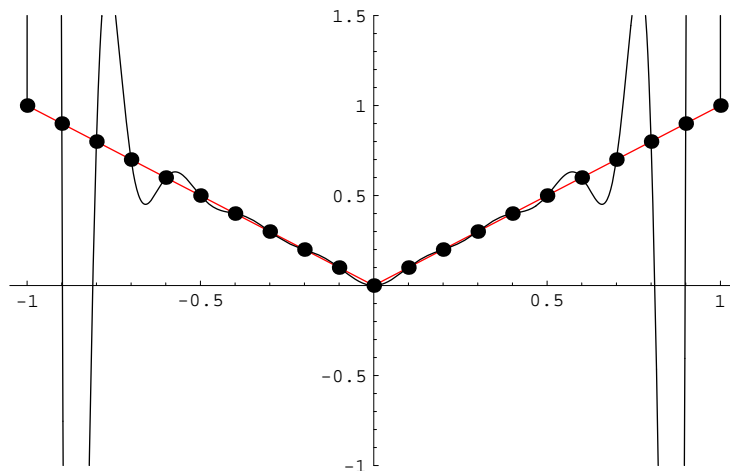


Fig. 1.7.3 p_{20}

Poder-se ia pensar que uma razão para a divergência desta sucessão de polinómios reside no facto de f não ser diferenciável na origem.

No entanto, C. Runge (em 1901) provou que, usando pontos igualmente espaçados, existem funções indefinidamente diferenciáveis para as quais (54) não se dá. Um exemplo é o caso da função

$$f(x) = \frac{1}{1 + 25x^2}, \quad x \in [-1, 1].$$

Runge considerou os polinómios $p_n(x)$ de grau $\leq n$ que interpolam f nos pontos equidistantes (55).

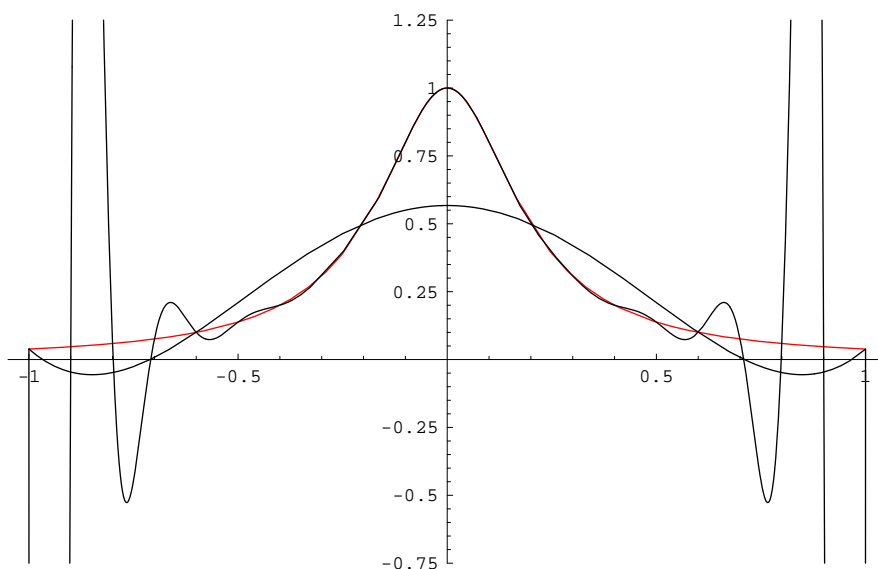


Fig 1.7.4 Gráficos de $f(x) = 1/(1 + 25x^2)$ e dos polinômios interpoladores em pontos igualmente espaçados p_5 e p_{20}

À medida que n tende para infinito, pode mostrar-se que

$$|e_n(x)| = |f(x) - p_n(x)| \rightarrow \infty, \quad \text{se } 0.726 \dots \leq |x| < 1.$$

Este facto é ilustrado graficamente na figura 1.7.4. Observamos que a aproximação melhora na parte central do intervalo, mas observam-se grandes oscilações perto dos extremos do intervalo $[-1, 1]$. Isto está relacionado com o comportamento da componente $\psi_n(t)$ do erro de interpolação (ver secção 1.6.1). Contudo, se examinarmos a fórmula do erro de interpolação para o caso de pontos igualmente espaçados, poderemos deduzir condições a impor à função f de modo a "compensar o comportamento oscilatório" de $\psi_n(t)$ e tal que a igualdade (54) seja válida. Suponhamos então que $f \in C^{n+1}[a, b]$. Usando (41), tem-se, para $t \in [0, n]$,

$$|f(x_0 + th) - p_n(x_0 + th)| \leq \left(\frac{b-a}{n}\right)^{n+1} \frac{1}{n+1} \max_{\xi \in [a,b]} |f^{(n+1)}(\xi)| \quad (56)$$

Se

$$\lim_{n \rightarrow \infty} \frac{(b-a)^{n+1}}{n^{n+1}} \max_{\xi \in [a,b]} \frac{|f^{(n+1)}(\xi)|}{n+1} = 0,$$

então (54) verifica-se. Isto acontece se por exemplo

$$\max_{\xi \in [a,b]} |f^{(n)}(\xi)| \leq M^n, \quad (M \in \mathcal{R}) \quad n = 1, 2, \dots$$

Incluimos ainda o seguinte resultado.

Teorema 1.7.1 *Se $f \in C^2[-1, 1]$, então a sucessão de polinómios interpoladores de grau $\leq n + 1$ que interpolam f nos zeros do polinómio de Chebyshev de grau $n + 1$, $n \geq 1$, ou seja, nos pontos*

$$x_k^{(n)} = \cos\left(\frac{2k + 1}{2n + 2}\pi\right), \quad k = 0, 1, \dots, n, \quad n \geq 1,$$

converge uniformemente para f no intervalo $[-1, 1]$, quando $n \rightarrow \infty$.

1.8 Interpolação de Hermite

Nas secções anteriores estudámos o chamado "polinómio interpolador de Lagrange de grau $\leq n$ " de uma dada função f . A existência desse polinómio p_n baseia-se no facto de o sistema de $n + 1$ equações lineares, $p_n(x_j) = f_j$, $0 \leq j \leq n$, nas $n + 1$ incógnitas a_0, a_1, \dots, a_n (os coeficientes de p_n) possuir uma única solução (ver teorema 1.1.1). É natural pensarmos que poderemos obter uma melhor aproximação polinomial para $f(x)$, impondo que algumas derivadas de f sejam também interpoladas em determinados pontos.

1.8.1 Interpolação de Hermite simples

Neste tipo de interpolação, procura-se um polinómio H que interpole f nos pontos x_j e de modo que também H' interpole f' nesses pontos. Temos o seguinte teorema de existência e unicidade.

Teorema 1.8.1 *Sejam x_0, x_1, \dots, x_n pontos distintos do intervalo $[a, b]$ e designem-se por f_j e f'_j , $0 \leq j \leq n$, os valores, respectivamente, de f e f' naqueles pontos. Existe um único polinómio H_{2n+1} de grau $\leq 2n + 1$ que verifica*

$$H_{2n+1}(x_j) = f_j, \quad H'_{2n+1}(x_j) = f'_j, \quad j = 0, 1, \dots, n. \quad (57)$$

Chamaremos a H_{2n+1} polinómio de Hermite de grau $2n + 1$ interpolador de f e f' nos pontos $\{x_i\}$.

Demonstração

Existência

Sejam $l_j(x)$ os polinómios de grau n dados por (6), ou seja,

$$l_j(x) = \prod_{\substack{0 \leq i \leq n \\ i \neq j}} \frac{(x - x_i)}{(x_j - x_i)},$$

e definam-se as funções

$$\phi_j(x) := (1 - 2l'_j(x_j)(x - x_j))(l_j(x))^2 \quad (58)$$

$$\bar{\phi}_j(x) := (x - x_j)(l_j(x))^2. \quad (59)$$

Provemos que

$$H_{2n+1}(x) = \sum_{j=0}^n f_j \phi_j(x) + \sum_{j=0}^n f'_j \bar{\phi}_j(x). \quad (60)$$

Basta mostrar que

i)

$$\phi_j(x_i) = \begin{cases} 0 & \text{se } j \neq i \\ 1 & \text{se } j = i, \end{cases}$$

$$\text{e } \bar{\phi}_j(x_i) = 0, \quad \forall i.$$

ii)

$$\bar{\phi}_j'(x_i) = \begin{cases} 0 & \text{se } j \neq i \\ 1 & \text{se } j = i, \end{cases}$$

$$\text{e } \phi_j'(x_i) = 0, \quad \forall i.$$

Deixamos a prova ao cuidado do leitor.

Unicidade

Suponhamos que além de H_{2n+1} , existe outro polinómio, Q_{2n+1} , de grau $\leq 2n + 1$, que resolve o problema de interpolação (57). Então $r(x) = H_{2n+1}(x) - Q_{2n+1}(x)$ é um polinómio de grau $\leq 2n + 1$ que verifica

$$r(x_j) = 0, \quad r'(x_j) = 0, \quad 0 \leq j \leq n,$$

donde sai que $r(x)$ tem $2n + 2$ zeros. Necessariamente $r \equiv 0$, ou seja, $H_{2n+1} \equiv Q_{2n+1}$, o que completa a demonstração do teorema.

Exemplo 1.8.1 Determinemos uma expressão para $H_3(x)$, o polinómio de Hermite de grau ≤ 3 interpolador de f e f' nos pontos $x_0 < x_1$ de $[a, b]$.

Tem-se

$$l_0(x) = \frac{x - x_1}{x_0 - x_1}, \quad l_1(x) = \frac{x - x_0}{x_1 - x_0}.$$

Aplicando (60), vem

$$\begin{aligned} H_3(x) &= f(x_0)\left(1 - 2\frac{x - x_0}{x_0 - x_1}\right)\left(\frac{x - x_1}{x_0 - x_1}\right)^2 + f(x_1)\left(1 - 2\frac{x - x_1}{x_1 - x_0}\right)\left(\frac{x - x_0}{x_1 - x_0}\right)^2 \\ &+ f'(x_0)(x - x_0)\left(\frac{x - x_1}{x_0 - x_1}\right)^2 + f'(x_1)(x - x_1)\left(\frac{x - x_0}{x_1 - x_0}\right)^2. \end{aligned} \quad (61)$$

O teorema seguinte dá-nos uma expressão para o erro na interpolação polinomial de Hermite.

Teorema 1.8.2 *Seja H_{2n+1} o polinómio de Hermite interpolador de f e f' nos pontos distintos x_0, x_1, \dots, x_n . Se $f \in C^{2n+2}[a, b]$ então, para cada $x \in [a, b]$, existe $\xi = \xi(x) \in I_p(x_0, x_1, \dots, x_n, x)$ tal que*

$$e_{2n+1}(x) := f(x) - H_{2n+1}(x) = \frac{f^{(2n+2)}(\xi)}{(2n+2)!} \prod_{i=0}^n (x - x_i)^2. \quad (62)$$

A demonstração é análoga à do teorema 1.2.1, mas neste caso tomamos para função auxiliar a função

$$F(t) = f(t) - H_{2n+1}(t) - \frac{f(x) - H_{2n+1}(x)}{[W(x)]^2} [W(t)]^2, \quad (x \neq x_i),$$

com $W(t) := \prod_{i=0}^n (t - x_i)$.

1.8.2 Interpolação de Hermite geral

O teorema 1.8.1 pode generalizar-se do seguinte modo.

Teorema 1.8.3 *Sejam x_0, x_1, \dots, x_n $n + 1$ pontos distintos de $[a, b]$ e m_0, m_1, \dots, m_n números inteiros positivos. Seja $N := m_0 + m_1 + \dots + m_n + n$. Então existe um único polinómio H_N de grau $\leq N$ tal que*

$$\begin{aligned} H^{(j)}(x_0) &= f^{(j)}(x_0), & j = 0, 1, \dots, m_0 \\ H^{(j)}(x_1) &= f^{(j)}(x_1), & j = 0, 1, \dots, m_1 \\ &\dots & \\ H^{(j)}(x_n) &= f^{(j)}(x_n), & j = 0, 1, \dots, m_n. \end{aligned} \quad (63)$$

1.9 Interpolação por polinómios de definição seccional. Splines polinomiais

Com o fim de sugerir uma motivação para o uso de interpolação polinomial de definição seccional, começamos por apresentar um exemplo.

Exemplo 1.9.1 A figura 1.9.1 mostra o gráfico dos valores de uma função $f(x)$ nos pontos 1, 3, 4, 6, 8, 10.

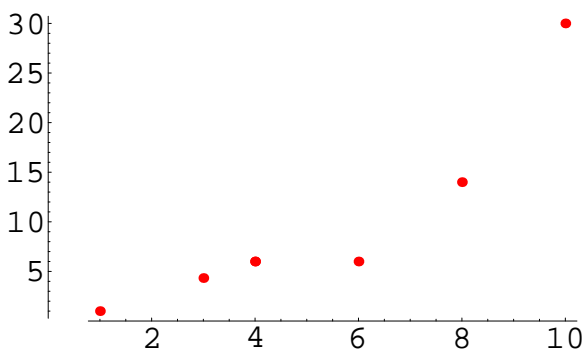


Fig 1.9.1 Conjunto $\{(x_i, f_i), 0 \leq i \leq 5\}$,

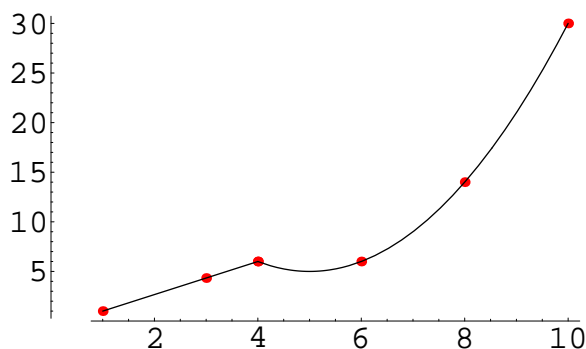


Fig 1.9.2 Função interpoladora g

Uma maneira de aproximar $f(x), x \in [1, 10]$ será através do seu polinómio interpolador de Lagrange, p_5 , de grau ≤ 5 . Contudo, observamos que f parece ter um comportamento diferente em duas regiões. O gráfico sugere que aproximemos f por um polinómio de grau ≤ 1 no subintervalo $[1, 4]$ e por um polinómio de grau ≤ 2 em $[4, 10]$ (ver figura 1.9.2). Isto é, considera-se a seguinte função seccionalmente polinomial

$$g(x) := \begin{cases} g_1(x) = a_0 + a_1x & \text{se } x \in [1, 4], \\ g_2(x) = b_0 + b_1x + b_2x^2 & \text{se } x \in [4, 10] \end{cases}$$

para aproximar f .

Em geral, quando se pretende uma aproximação $g(x)$ tal que $g(x_i) = f_i, i = 0, 1, \dots, n$, a interpolação seccionalmente polinomial pode ter vantagens sobre o uso de um polinómio interpolador $p_n(x)$ de grau $\leq n$. Basta lembrar que, à medida que n aumenta, p_n tende a apresentar oscilações mais acentuadas. Na interpolação seccional, vários polinómios de grau baixo são juntos de maneira contínua, de modo a que a função $g(x)$ resultante seja interpoladora de f . Um caso extremo é o da linha quebrada passando pelos pontos (x_i, f_i) .

A função $q(x)$ da figura 1.9.3 foi obtida unindo os pontos (x_i, f_i) e (x_{i+1}, f_{i+1}) por um segmento de recta, para $i = 0, 1, \dots, 5$. Ou seja, $q(x_i) = f_i$ e $q(x)$ é um polinómio de grau ≤ 1 em cada subintervalo $[x_i, x_{i+1}]$. A função $q(x)$ tem a desvantagem de, nos pontos de junção, não ser diferenciável.

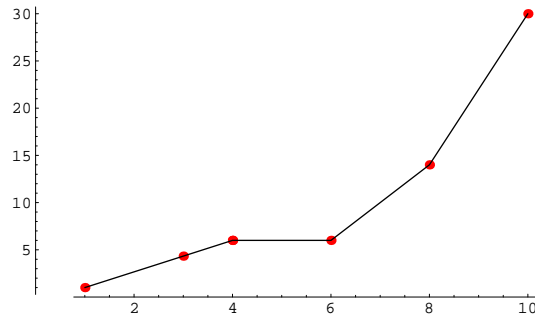


Fig 1.9.3 Interpolação linear de definição seccional

Consideramos em seguida uma classe de funções seccionalmente polinomiais, das quais são caso particular os chamados *splines cúbicos interpoladores*, muito usados nas aplicações.

Seja Δ a seguinte partição do intervalo $[a, b]$

$$\Delta : a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b. \quad (64)$$

Definição 1.9.1

$S_r(\Delta)$ é o conjunto das funções reais s definidas em $[a, b]$, tal que

- i) $s \in C^{r-1}[a, b]$,
- ii) s é um polinómio de grau $\leq r$ em cada subintervalo $[x_i, x_{i+1}]$, $i = 0, 1, \dots, N - 1$.

Os elementos s de $S_r(\Delta)$ são chamados *splines polinomiais de grau r* . São funções seccionalmente polinomiais de grau $\leq r$, que têm derivadas contínuas até à ordem $r - 1$ nos pontos de junção x_i , $i = 1, 2, \dots, N$. Tem-se

$$s(x) := \begin{cases} s_0(x) & x \in [x_0, x_1] \\ s_1(x) & x \in [x_1, x_2] \\ \vdots & \\ s_{N-1}(x) & x \in [x_{N-1}, x_N] \end{cases}$$

e, atendendo à condição ii), verificam-se as igualdades, para $j = 0, 1, \dots, r - 1$,

$$\begin{aligned} s_0^{(j)}(x_1) &= s_1^{(j)}(x_1) \\ s_1^{(j)}(x_2) &= s_2^{(j)}(x_2) \\ &\vdots \\ s_{N-1}^{(j)}(x_{N-1}) &= s_N^{(j)}(x_{N-1}). \end{aligned}$$

Teorema 1.9.1 Com as operações de soma de duas funções e produto de uma função por um escalar, $S_r(\Delta)$ é um espaço vectorial de dimensão $N + r$. As funções

$$\{1, (x - x_0), (x - x_0)^2, \dots, (x - x_0)^r, (x - x_1)_+^r, (x - x_2)_+^r, \dots, (x - x_{N-1})_+^r\}, \quad (65)$$

onde, para $j = 1, 2, \dots, N - 1$,

$$(x - x_j)_+^r := \begin{cases} (x - x_j)^r & \text{se } x > x_j \\ 0 & \text{se } x \leq x_j \end{cases} \quad (66)$$

constituem uma base para $S_r(\Delta)$.

Teorema 1.9.2 Se $s \in S_r(\Delta)$, então s pode ser representado de forma única

$$s(x) = \sum_{i=0}^r a_i (x - x_0)^i + \sum_{j=1}^{N-1} c_j (x - x_j)_+^r, \quad (67)$$

onde os a_i, c_j são constantes reais.

Do ponto de vista computacional, esta base não é conveniente. Se pretendermos calcular o valor de $s(x)$ quando x é um ponto perto de x_N , teremos de usar quase todas as funções base. É no entanto possível construir uma outra base, formada por funções chamadas B-splines, que são funções com suporte compacto.

Em seguida damos alguns exemplos de espaços $S_r(\Delta)$.

I O espaço $S_1(\Delta)$

Os elementos deste espaço são funções contínuas, seccionalmente polinomiais de grau ≤ 1 . Ou seja, cada elemento $s \in S_1(\Delta)$ é a junção de polinómios de grau ≤ 1 .

Se, no exemplo 1.9.1, definirmos $\Delta := \{1 < 3 < 4 < 6 < 8 < 10\}$ como uma partição do intervalo $[1, 10]$, então a função aproximadora $q(x)$ é um elemento de $S_1(\Delta)$.

O teorema seguinte diz-nos que um elemento $s \in S_1(\Delta)$ fica bem determinado se conhecermos os valores $s(x_i), i = 0, \dots, N$.

Teorema 1.9.3 Sejam dados valores $f_j, 0 \leq j \leq N$, de uma função f nos pontos x_j da partição Δ . Então existe um único $s \in S_1(\Delta)$ tal que

$$s(x_j) = f_j, \quad j = 0, 1, \dots, N. \quad (68)$$

A função s é chamada *spline interpolador de grau 1*.

Demonstração

Seja $s_j(x)$ a restrição de $s(x)$ a $[x_j, x_{j+1}]$. Tem-se

$$s_j(x) = f_j + \frac{f_{j+1} - f_j}{x_{j+1} - x_j} (x - x_j),$$

que é o polinómio de Lagrange de grau ≤ 1 que interpola f nos pontos x_j, x_{j+1} . A existência e unicidade de s_j estão garantidos pelo teorema 1.1.1.

Estes splines têm aplicação na resolução numérica de equações integrais e equações diferenciais. Contudo, como já foi dito anteriormente, possuem o inconveniente de, em geral, não possuírem derivada nos pontos de junção. Os splines diferenciáveis mais usados são os splines cúbicos, que estudaremos em seguida.

II O espaço $S_3(\Delta)$

Os elementos de $S_3(\Delta)$ são funções s que verificam

- i) $s \in C^2[a, b]$
 - ii) s é um polinómio de grau ≤ 3 em cada subintervalo $[x_i, x_{i+1}]$, $i = 0, 1, \dots, N - 1$.
- (69)

Teorema 1.9.4 *Sejam dados os valores $f_j, 0 \leq j \leq N$, de uma função f nos pontos x_j da partição Δ , bem como os valores $f'_0 := f'(x_0), f'_N := f'(x_N)$. Existe um único $s \in S_3(\Delta)$ tal que*

$$s(x_j) = f_j, \quad j = 0, 1, \dots, N, \quad (70)$$

$$s'(x_j) = f'_j, \quad j = 0, N. \quad (71)$$

A função s é chamada *spline cúbico interpolador*.

Demonstração

Designando por $s_j(x)$ a restrição de $s(x)$ a $[x_j, x_{j+1}]$, $j = 0, 1, \dots, N - 1$, devemos provar que as condições (70) e (71), permitem construir cada s_j de maneira única. Começamos por notar que da definição do espaço $S_3(\Delta)$ e da condição (70) resulta

$$s_j(x_j) = f_j, \quad s_j(x_{j+1}) = f_{j+1}, \quad j = 0, 1, \dots, N - 1. \quad (72)$$

Além disso, impondo a continuidade da primeira derivada, deve ter-se

$$s'_j(x_j) = s'_{j-1}(x_j), \quad s'_j(x_{j+1}) = s'_{j+1}(x_{j+1}), \quad j = 1, 2, \dots, N - 1. \quad (73)$$

Designemos por k_j o valor de $s'(x_j)$, $j = 1, \dots, N - 1$. Então, considerando o intervalo genérico $[x_j, x_{j+1}]$, resulta de (72) e (73) que a componente s_j de s verifica

$$s_j(x_j) = f_j, \quad s_j(x_{j+1}) = f_{j+1} \quad (74)$$

$$s'_j(x_j) = k_j, \quad s'_j(x_{j+1}) = k_{j+1}, \quad (75)$$

para $j = 0, 1, \dots, N - 1$. Pelo teorema de interpolação de Hermite, sabemos que existe um único polinómio de grau ≤ 3 que satisfaz estas condições. Denotando $h_j := x_{j+1} - x_j$, uma expressão para s_j pode ser obtida de (62), com as necessárias adaptações. Assim

$$\begin{aligned}
s_j(x) &= f_j \frac{(x - x_{j+1})^2}{h_j^2} [1 + 2(x - x_j)/h_j] + f_{j+1} \frac{(x - x_j)^2}{h_j^2} [1 - 2(x - x_{j+1})/h_j] \\
&+ k_j \frac{(x - x_j)}{(h_j)^2} (x - x_{j+1})^2 + k_{j+1} \frac{(x - x_j)^2}{(h_j)^2} (x - x_{j+1}).
\end{aligned} \tag{76}$$

Note-se que os valores k_1, k_2, \dots, k_{N-1} são desconhecidos à partida. Vamos provar que esses valores podem ser obtidos impondo-se a continuidade das segundas derivadas de s . Derivando (76) duas vezes, obtém-se

$$s_j''(x_j) = -\frac{6}{h_j^2} f_j + \frac{6}{h_j^2} f_{j+1} - \frac{4}{h_j} k_j - \frac{2}{h_j} k_{j+1} \tag{77}$$

$$s_j''(x_{j+1}) = \frac{6}{h_j^2} f_j - \frac{6}{h_j^2} f_{j+1} + \frac{2}{h_j} k_j + \frac{4}{h_j} k_{j+1}. \tag{78}$$

Impondo a igualdade $s_j''(x_j) = s_{j-1}''(x_j)$, $j = 1, 2, \dots, N - 1$, vem

$$\frac{k_{j-1}}{h_{j-1}} + 2\left(\frac{1}{h_{j-1}} + \frac{1}{h_j}\right)k_j + \frac{1}{h_j}k_{j+1} = 3\left(\frac{\Delta f_{j-1}}{(h_{j-1})^2} + \frac{\Delta f_j}{(h_j)^2}\right),$$

onde $\Delta f_j := f_{j+1} - f_j$.

Definindo ainda $\tau_j := 1/h_j$, obtém-se, fazendo $j = 1, 2, \dots, N - 1$, o seguinte sistema de $N - 1$ equações nas $N - 1$ incógnitas k_1, k_2, \dots, k_{N-1} :

$$\begin{aligned}
\tau_0 k_0 + 2(\tau_0 + \tau_1)k_1 + \tau_1 k_2 &= b_1 \\
\tau_1 k_1 + 2(\tau_1 + \tau_2)k_2 + \tau_2 k_3 &= b_2 \\
\tau_2 k_2 + 2(\tau_2 + \tau_3)k_3 + \tau_3 k_4 &= b_3 \\
&\vdots \\
\tau_{N-3} k_{N-3} + 2(\tau_{N-3} + \tau_{N-2})k_{N-2} + \tau_{N-2} k_{N-1} &= b_{N-2} \\
\tau_{N-2} k_{N-2} + 2(\tau_{N-2} + \tau_{N-1})k_{N-1} + \tau_{N-1} k_N &= b_{N-1},
\end{aligned} \tag{79}$$

onde k_0, k_N são conhecidos. A matriz do sistema (79) é a seguinte

$$\begin{bmatrix}
2(\tau_0 + \tau_1) & \tau_1 & 0 & 0 & \cdots & 0 \\
\tau_1 & 2(\tau_1 + \tau_2) & \tau_2 & 0 & \cdots & 0 \\
0 & \tau_2 & 2(\tau_2 + \tau_3) & \tau_3 & \cdots & 0 \\
\vdots & & \ddots & & & \vdots \\
\vdots & & & \ddots & & \vdots \\
0 & & \cdots & \tau_{N-3} & 2(\tau_{N-3} + \tau_{N-2}) & \tau_{N-2} \\
& & & & \tau_{N-2} & 2(\tau_{N-2} + \tau_{N-1})
\end{bmatrix}$$

Trata-se de uma matriz tridiagonal de diagonal estritamente dominante (é também simétrica e definida positiva). Consequentemente, é uma matriz não singular, pelo que o sistema (79)

tem uma solução única bem determinada k_1, k_2, \dots, k_{N-1} . Estes valores são depois substituídos em (76), ficando deste modo cada s_j bem determinado. Isto completa a demonstração.

O teorema seguinte é uma propriedade dos splines cúbicos interpoladores.

Teorema 1.9.5 *Dada uma função $f \in C^3[a, b]$, seja s o spline cúbico interpolador de f associado aos pontos $a = x_0 < x_1 < x_2 < \dots < x_N = b$. Seja ainda $g \in C^2[a, b]$ uma função qualquer tal que g interpola f nos pontos x_i e verifica $g'(a) = f'(a)$, $g'(b) = f'(b)$. Então*

$$\int_a^b [g''(x)]^2 dx \geq \int_a^b [s''(x)]^2 dx$$

e a igualdade é atingida se e só se $g = s$. Isto significa que de entre as funções $g \in C^2[a, b]$ que interpolam f em x_0, \dots, x_n e verificam $g'(a) = f'(a)$, $g'(b) = f'(b)$, a função com curvatura mínima é $g = s$.

Incluimos ainda dois resultados sobre o erro de interpolação associado aos spline cúbicos interpoladores.

Teorema 1.9.6 *Seja $f \in C^2[a, b]$ e seja s o spline cúbico interpolador de f associado aos pontos $a = x_0 < x_1 < x_2 < \dots < x_N = b$. Designemos por $h := \max_{0 \leq j \leq N-1} x_{j+1} - x_j$ e seja x um ponto qualquer de $[a, b]$. Então tem-se*

$$|f(x) - s(x)| \leq h^{3/2} \left(\int_a^b (f''(t))^2 dt \right)^{1/2} \quad (80)$$

Deste teorema podemos concluir que, fazendo tender para infinito o número de pontos da partição Δ , de modo que $h \rightarrow 0$, então $s(x)$ converge uniformemente para $f(x)$.

Teorema 1.9.7 *Se $f \in C^4[a, b]$ e s é o spline cúbico interpolador de f , com h definido como no teorema anterior, então*

$$\|f - s\|_\infty \leq \frac{5}{384} h^4 \|f^{(4)}\|_\infty. \quad (81)$$

2 Teoria da aproximação

2.1 Introdução

Continua a ser do nosso interesse o problema de aproximar uma função "complicada", $f(x)$, por um função "mais simples". Já anteriormente considerámos exemplos de funções que podem ser usadas como funções aproximadoras:

- (I) Os polinómios interpoladores
- (II) Os splines interpoladores de graus 1 e 3
- (III) Os polinómios de Taylor $q_n(x)$, definidos por

$$q_n(x) := \sum_{j=0}^n \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j.$$

Relembremos que q_n é o único polinómio de grau menor ou igual a n que verifica as condições:

$$q_n^{(j)}(x_0) = f^{(j)}(x_0), \quad 0 \leq j \leq n.$$

Então as somas parciais das séries de Taylor dão resposta a um problema que pode ser considerado como um caso especial de interpolação de Hermite. Se f for de classe C^{n+1} , o erro da aproximação de $f(x)$ por $q_n(x)$ é dado por

$$\frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1},$$

o que nos leva a concluir que a aproximação será boa se x estiver suficientemente próximo de x_0 , tornando-se em geral pior à medida que x se afasta de x_0 .

Neste capítulo, consideraremos outros tipos de funções aproximadoras para $f(x)$, sem a exigência de que elas interpoem f . Com efeito, em muitas situações não é conveniente que a função aproximadora seja uma função interpoladora. Por exemplo, suponhamos que a função f traduz uma relação entre duas grandezas físicas, x e $f(x)$. Suponhamos ainda que os valores $f(x_i)$ $i = 0, 1, \dots, N$ foram obtidos experimentalmente; ou seja, apenas dispomos dos valores $f(x_i) + \epsilon_i$, $i = 0, 1, \dots, N$, onde os erros ϵ_i são desconhecidos. Neste caso, não seria razoável aproximar f por um polinómio passando por $(x_i, f(x_i) + \epsilon_i)$, $i = 0, 1, \dots, N$, a não ser que os erros ϵ_i fossem pequenos. Uma solução mais adequada seria utilizar um polinómio obtido pelo chamado método dos mínimos quadrados (ver secção 2.5), onde se reduz o efeito dos erros dos dados. A situação já é diferente quando se conhece uma expressão analítica da função f que se pretende aproximar. Nesse caso poderemos por exemplo estar interessados em determinar o polinómio com o menor grau possível que aproxima f num dado intervalo, com um erro máximo inferior a um certo δ . Idealmente, uma função aproximadora deve traduzir, o mais fielmente possível, o comportamento da função que ela vai representar. Por isso, em muitos casos, poderá ser mais natural utilizar funções trigonométricas, exponenciais ou racionais.

O que se entende por uma "boa" aproximação? Uma certa função ϕ^* será uma boa aproximação para f se a função $f - \phi^*$ for pequena nalgum sentido.

O teorema seguinte motiva o uso de uma certa classe de funções, os polinómios, para aproximar funções contínuas.

Teorema 2.1.1 Teorema de Weierstrass Seja $[a, b] \subset \mathcal{R}$ e $f \in C[a, b]$. Então dado $\epsilon > 0$, existe um polinómio P_n de grau $n = n(\epsilon)$, tal que

$$|f(x) - P_n(x)| < \epsilon, \quad \text{para qualquer } x \in [a, b]$$

Uma demonstração deste teorema consiste na construção dos chamados *polinómios de Bernstein*. Seja $\{B_n(f, x)\}$, com $x \in [0, 1]$, o polinómio de Bernstein

$$B_n(f, x) := \sum_{k=0}^n \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}, \quad n \geq 1.$$

Tem-se o seguinte resultado de convergência uniforme:

$$\lim_{n \rightarrow \infty} \|f^{(k)}(x) - B_n(f, x)\|_{\infty} = 0$$

2.2 Melhor aproximação num espaço métrico

Definição Seja E um conjunto não vazio. Uma função $d : E \times E \rightarrow \mathcal{R}$, que verifica

$$\text{(M1)} \quad d(f, g) \geq 0 \text{ ed } d(f, g) = 0 \text{ se e só se } f = g, \quad \forall f, g \in E \quad (1)$$

$$\text{(M2)} \quad d(f, g) = d(g, f), \quad \forall f, g \in E, \quad (2)$$

$$\text{(M3)} \quad d(f, g) \leq d(f, h) + d(h, g), \quad \forall f, g, h \in E, \quad (3)$$

é uma *métrica* ou *distância*. O conjunto E onde está definida a métrica diz-se um *espaço métrico*.

Nos problemas que trataremos, o elemento f que se pretende aproximar pertencerá a um certo espaço métrico E e o conjunto F onde se vai procurar o elemento aproximador será um subespaço de E . Já que uma das propriedades básicas de um espaço métrico é possuir uma função distância (ou métrica), ele fornece um modo natural de medir a qualidade de uma aproximação. Será, então, natural usar o seguinte critério: o elemento f_1 "aproxima melhor" f do que o elemento f_2 se $d(f, f_1) < d(f, f_2)$. Somos conduzidos ao seguinte *problema geral de aproximação*:

Seja d^* a quantidade

$$d^* := \inf_{\phi \in F} d(f, \phi), \quad (4)$$

Pretende-se saber se existe algum elemento $\phi^* \in F$ para o qual este ínfimo é atingido, e, em caso afirmativo, saber se ele é único.

E tem-se a seguinte definição:

Definição 2.2.1 Um elemento $\phi^* \in F$ para o qual se tem

$$\begin{aligned} d(f, \phi^*) &= \inf_{\phi \in F} d(f, \phi) = d^*, \\ \text{ou seja} \\ d(f, \phi^*) &\leq d(f, \phi), \quad \forall \phi \in F \end{aligned}$$

é chamado uma *melhor aproximação* (m. a.) do elemento $f \in E$ em F , relativamente à métrica ou distância d .

O seguinte Teorema dá uma resposta ao problema de existência de melhor aproximação num espaço métrico.

Teorema 2.2.1 *Seja F um subconjunto compacto de um espaço métrico E . Então, dado um elemento f de E existe pelo menos um elemento ϕ^* de F tal que $d(f, \phi^*) = d(f, F) = \inf_{\phi \in F} d(f, \phi)$.*

Demonstração A função $\phi \in F \rightarrow d(f, \phi)$ é contínua e, sendo F compacto, ela atinge um mínimo nalgum elemento ϕ^* ; ou seja, existe pelo menos um elemento $\phi^* \in F$ tal que $d(f, \phi^*) = \inf_{\phi \in F} d(f, \phi) = \min_{\phi \in F} d(f, \phi)$. \diamond

2.3 Melhor aproximação em espaços normados e seminormados

Seja E um espaço vectorial. Uma aplicação $\|\cdot\| : E \rightarrow \mathcal{R}$, que verifica

$$(N1) \quad \|f\| = 0 \Leftrightarrow f = 0, \quad \forall f \in E \tag{5}$$

$$(N2) \quad \|\lambda f\| = |\lambda| \|f\|, \quad \forall f \in E, \lambda \in \mathcal{R} \tag{6}$$

$$(N3) \quad \|f + g\| \leq \|f\| + \|g\|, \quad \forall f, g \in E, \tag{7}$$

é uma *norma*. O espaço E onde está definida a norma diz-se um *espaço normado*.

Uma norma induz uma métrica. Então E será um espaço métrico com métrica

$$d(f, g) = \|f - g\|, \quad \forall f, g \in E. \tag{8}$$

Da Definição 2.2.1 resulta a seguinte

Definição 2.3.1 *Seja E um espaço normado e seja $F \subset E$ um subespaço de E . Dado um elemento f de E , se existe algum elemento ϕ^* de F tal que*

$$\|f - \phi^*\| = \inf_{\phi \in F} \|f - \phi\|, \tag{9}$$

então ϕ^* é chamado uma melhor aproximação de f em F , em relação à norma $\|\cdot\|$.

Seminormas

Uma aplicação $\|\cdot\| : E \rightarrow \mathcal{R}$, que verifica as condições **N2**, **N3** e, em vez de **N1**, apenas

$$(\mathbf{N1}') : \|f\| = 0 \text{ se } f = 0,$$

é chamada uma *seminorma*. O espaço onde está definida uma seminorma dz-se um *espaço seminormado*.

Exemplos

Seja $C([a, b])$ o espaço das funções definidas e contínuas num intervalo fechado $[a, b]$. Dados $N+1$ pontos distintos $\pi = \{x_0, x_1, x_2, \dots, x_N\}$ de $[a, b]$, a aplicação $\|\cdot\| : C([a, b]) \rightarrow \mathcal{R}$ definida por

$$\|f\|_{2,\pi} = \left(\sum_{i=0}^N \omega(x_i) (f(x_i))^2 \right)^{1/2}, \quad (10)$$

onde $\omega \in C([a, b])$ e $w(x_i) > 0$, $i = 0, 1, \dots, N$, é uma seminorma, a que chamaremos seminorma Euclideana.

Teorema 2.3.1 *Seja E um espaço vectorial seminormado e seja $F \subset E$ um subespaço de dimensão finita. Então para cada elemento f de E existe pelo menos um elemento ϕ^* que é melhor aproximação de f em F , em relação à seminorma definida em F .*

2.4 Melhor aproximação uniforme

Dependendo da escolha da norma ou seminorma obtêm-se tipos diferentes de aproximação. Esta secção é dedicada à *melhor aproximação uniforme*, também chamada *aproximação de Chebyshev* ou ainda *aproximação minimax*. Consideremos o espaço real $E = C([a, b])$ e seja $F \subset C([a, b])$ um subespaço com dimensão finita n . Para cada função $f \in E$, o Teorema 2.3.1 garante a existência de uma melhor aproximação f em F . O problema da unicidade requer uma análise especial e, para isso, apresentamos o seguinte conceito devido a A. Haar (1918).

Definição 2.4.1 *Seja $F \subset C([a, b])$ um subespaço com dimensão finita n . Diz-se que F satisfaz a *condição de Haar* se todo o elemento f não nulo de F tem quando muito $n - 1$ zeros em $[a, b]$.*

Exemplo *Seja $F = \mathbf{P}_{n-1}([a, b])$ o subespaço de $C([a, b])$ constituído pelo polinómio nulo e ainda os polinómios de grau menor ou igual a $n - 1$. Facilmente se verifica que F satisfaz a condição de Haar.*

Teorema 2.4.1 *A condição de Haar é equivalente à seguinte: para qualquer base $\{\psi_1, \psi_2, \dots, \psi_n\}$ de F e qualquer que seja o conjunto de n pontos distintos $\{t_1, t_2, \dots, t_n\}$ de $[a, b]$, tem-se*

$$D = \begin{vmatrix} \psi_1(t_1) & \psi_1(t_2) & \cdots & \psi_1(t_n) \\ \psi_2(t_1) & \psi_2(t_2) & \cdots & \psi_2(t_n) \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \psi_n(t_1) & \psi_n(t_2) & \cdots & \psi_n(t_n) \end{vmatrix} \neq 0 \quad (11)$$

Também se diz que o conjunto $\{\psi_1, \psi_2, \dots, \psi_n\}$ é um *sistema de Chebyshev*.

Teorema 2.4.2 (Teor. de Haar) *Seja $F \subset C([a, b])$ um subespaço com dimensão finita n . A m.a. uniforme de f em F é única, qualquer que seja $f \in C([a, b])$, se e só se F satisfaz a condição de Haar.*

Corolário A m. a. uniforme de $f \in C([a, b])$ em $\mathbf{P}_n([a, b])$ é única.

Teorema 2.4.3 (Teor. de La Vallée-Poussin) *Seja $F \subset C([a, b])$ um subespaço com dimensão finita n que satisfaz a condição de Haar. Consideremos uma função $f \in C([a, b])$. Suponhamos que existe $\phi \in F$ tal que, em $n + 1$ pontos de $[a, b]$ $x_0 < x_1 < \dots < x_n$, a diferença $f - \phi$ verifica*

$$f(x_i) - \phi(x_i) = \theta (-1)^i \lambda_i, \quad i = 0, 1, 2, \dots, n, \quad (12)$$

com $\theta = 1$ ou $\theta = -1$ e $\lambda_i > 0$.

Então a m.a. uniforme ϕ^* de f em F é tal que

$$\|f - \phi^*\|_\infty = \max_{x \in [a, b]} |f(x) - \phi^*(x)| \geq \min_{i=0,1,\dots,n} \lambda_i \quad (13)$$

Demonstração Tem-se

$$\begin{aligned} f(x_0) - \phi(x_0) &= \theta \lambda_0 \\ f(x_1) - \phi(x_1) &= \theta \lambda_1 \\ f(x_2) - \phi(x_2) &= \theta \lambda_2 \\ &\vdots \\ f(x_n) - \phi(x_n) &= \theta \lambda_n \end{aligned}$$

Pretende-se provar que $\|f - \phi^*\|_\infty \geq \min \lambda_i$. Neguemos a tese, ou seja, suponhamos que

$$\|f - \phi^*\|_\infty < \min \lambda_i. \quad (14)$$

Resulta então

$$|f(x) - \phi^*(x)| \leq \|f - \phi^*\|_\infty < \min \lambda_i \leq \lambda_i \quad (15)$$

e tem-se a desigualdade

$$|f(x) - \phi^*(x)| < \lambda_i, \quad i = 0, 1, \dots, n. \quad (16)$$

Por outro lado,

$$\phi(x_i) - \phi^*(x_i) = \underbrace{(f(x_i) - \phi^*(x_i))}_{<\lambda_i} - \underbrace{(f(x_i) - \phi(x_i))}_{\pm\lambda_i}, \quad i = 0, 1, \dots, n, \quad (17)$$

donde concluimos que $\phi(x_i) - \phi^*(x_i)$ tem o sinal de $-(f(x_i) - \phi(x_i)) = \theta (-1)^{i+1} \lambda_i$. Ou seja, $\phi - \phi^*$ alterna nos $n + 1$ pontos x_i e, portanto, tem pelo menos n zeros em $[a, b]$. Mas por outro lado $\phi - \phi^*$ é por sua vez um elemento de F , subespaço dimensão n verificando a condição de Haar. Atendendo à Definição 2.4.1, terá de ter-se $\phi - \phi^* \equiv 0$. Usando esta igualdade acima em (17), vem $f(x_i) - \phi^*(x_i) = f(x_i) - \phi(x_i)$, o que contraria (14). \diamond

Caracterização da m.a. uniforme

Teorema 2.4.4 (Teor. de Chebyshev ou da oscilação uniforme) *Seja $F \subset C([a, b])$ um subespaço com dimensão finita n que satisfaz a condição de Haar. Dada uma função $f \in C([a, b])$, é condição necessária e suficiente para que ϕ seja a m.a. uniforme de f em F , que existam em $[a, b]$ $n + 1$ pontos distintos $x_0 < x_1 < \dots < x_n$, tal que*

$$(i) \quad f(x_i) - \phi(x_i) = -(f(x_{i+1}) - \phi(x_{i+1})), \quad i = 0, 1, \dots, n-1. \quad (18)$$

$$(ii) \quad |f(x_i) - \phi(x_i)| = \|f - \phi\|_\infty = \delta, \quad i = 0, 1, \dots, n. \quad (19)$$

Demonstração Consideramos aqui apenas a condição suficiente. Resulta de (18)-(19) que ϕ verifica uma relação do tipo (13), com $\lambda_i = \delta$, $i = 0, 1, \dots, n$. Então, se for ϕ^* a m.a. de f em F , resulta de (13)

$$\|f - \phi^*\|_\infty \geq \delta \quad (20)$$

Por outro lado, da definição de m.a., vem $\|f - \phi^*\|_\infty \leq \|f - \phi\|_\infty$, o que, juntamente com (20), implica $\|f - \phi^*\|_\infty = \|f - \phi\|_\infty$. Daqui sai $\phi = \phi^*$, ou seja, ϕ é a m.a. de f em F . \diamond

2.5 Melhor aproximação em espaços pré-Hilberteanos

Dada uma função $f \in E$, o problema de determinar a melhor aproximação ϕ^* de f em $F \subset E$ é em geral difícil. Para a melhor aproximação uniforme é possível obter aproximações usando certos métodos iterativos, mas não se consegue, na maioria dos casos, uma fórmula explícita para ϕ^* . Vamos agora considerar o importante caso em que E é um espaço pré-Hilberteano, isto é, um espaço vectorial munido dum produto interno. Veremos que a determinação da melhor aproximação em relação à norma induzida pelo produto interno pode ser reduzida à resolução de um sistema de equações lineares. Começemos por relembrar alguns conceitos.

Seja E um espaço vectorial. Uma aplicação $\langle \cdot, \cdot \rangle: E \times E \rightarrow \mathcal{R}$, que verifica

$$(P1) \quad \langle f, f \rangle \geq 0 \text{ e } \langle f, f \rangle = 0 \Leftrightarrow f = 0, \quad \forall f \in E \quad (21)$$

$$(P2) \quad \langle f, g \rangle = \langle g, f \rangle \quad \forall f, g \in E, \quad (22)$$

$$(P3) \quad \langle f, \alpha g + \mu h \rangle = \alpha \langle f, g \rangle + \mu \langle f, h \rangle, \quad \forall f, g, h \in E, \alpha, \mu \in \mathcal{R} \quad (23)$$

é um *produto interno*. O espaço E onde está definido um produto interno diz-se um *espaço pré-Hilbertiano*.

Um produto interno induz a norma

$$\|f\| = \sqrt{\langle f, f \rangle} \quad (24)$$

e E passa a ser um espaço normado.

Um exemplo importante de espaço pré-Hilbertiano é o espaço $C([a,b])$, com

$$\langle f, g \rangle = \int_a^b \omega(x) f(x)g(x)dx, \quad (25)$$

onde $\omega(x) \geq 0$ é uma função contínua que se anula quando muito num número finito de pontos.

Pode-se então definir

$$\|f\|_2 = \sqrt{\int_a^b \omega(x) (f(x))^2 dx} \quad (26)$$

que se chama norma Euclideana.

São de especial interesse os seguintes problemas, em particular quando F é um espaço de polinómios.

PROBLEMA A: dada $f \in C([a, b])$ e o subespaço $F \subset C([a, b])$ de dimensão finita, determinar o elemento $\phi^* \in F$ tal que

$$\|f - \phi^*\|_2 = \min_{g \in F} \|f - g\|_2$$

Ou seja, ϕ^* é o elemento de F que que minimiza a norma Euclideana do erro $f - \phi^*$, isto é, a expressão

$$\left(\int_a^b \omega(x) (f(x) - g(x))^2 dx \right)^{1/2}, \quad (27)$$

de entre todas as funções $g \in F$. Chamamos a ϕ^* a *melhor aproximação mínimos quadrados* de f em F .

O correspondente problema discreto é o seguinte:

PROBLEMA B: Sejam dados os valores $f(x_i), i = 0, 1, \dots, N$ de uma função $f \in C([a, b])$ em $N + 1$ pontos distintos $\pi = \{x_0, \dots, x_N\}$ de $[a, b]$. Sejam ainda $w_i = \omega(x_i), i = 0, 1, \dots, N$, valores positivos. Considere o problema de determinar o elemento ϕ^* dum subespaço $F \subset C([a, b])$ de dimensão finita que minimiza a seminorma Euclideana do erro (ver (10))

$$\|f - \phi\|_{2,\pi} = \left(\sum_{i=0}^N w_i (f(x_i) - \phi(x_i))^2 \right)^{1/2}, \quad \phi \in F \quad (28)$$

Chamamos a ϕ^* a *melhor aproximação mínimos quadrados discreta* de f em F .

Note que a existência de pelo menos uma solução para cada um dos PROBLEMAS A,B fica assegurada pelo Teorema 2.3.1.

Caracterização da m. a. num espaço pré -Hilberteano

Teorema 2.5.1 *Seja E um espaço pré-Hilberteano, F um subespaço linear de E e f um elemento de E . O elemento $\phi^* \in F$ é uma m. a. de f em F se e somente se*

$$\langle f - \phi^*, \phi \rangle = 0, \quad \forall \phi \in F. \quad (29)$$

Demonstração (i) $\phi^* \in F$ é uma m. a. de $f \Rightarrow \langle f - \phi^*, \phi \rangle = 0, \quad \forall \phi \in F$.

Suponhamos que existe $\phi_1 : \langle f - \phi^*, \phi_1 \rangle = \alpha \neq 0$

Seja $\phi_2 = \phi^* + \beta \phi_1$, com $\beta = \alpha / (\|\phi_1\|^2)$.

Tem-se

$$\begin{aligned} \|f - \phi_2\|^2 &= \langle f - \phi_2, f - \phi_2 \rangle = \langle f - \phi^* - \beta\phi_1, f - \phi^* - \beta\phi_1 \rangle \\ &= \langle f - \phi^*, f - \phi^* \rangle - \beta \langle f - \phi^*, \phi_1 \rangle + \beta^2 \langle \phi_1, \phi_1 \rangle - \beta \langle f - \phi^*, \phi_1 \rangle \\ &= \|f - \phi^*\|^2 - 2 \frac{\alpha}{\|\phi_1\|^2} \langle f - \phi^*, \phi_1 \rangle + \frac{\alpha^2}{\|\phi_1\|^4} \|\phi_1\|^2 \\ &= \|f - \phi^*\|^2 - \frac{\alpha^2}{\|\phi_1\|^2} \end{aligned}$$

Sai que $\|f - \phi^*\|^2 = \|f - \phi_2\|^2 + \alpha^2 / (\|\phi_1\|^2)$, logo $\|f - \phi^*\|^2 > \|f - \phi_2\|^2$, o que contraria a hipótese de ϕ^* ser uma m.a. de f em F .

(ii) $\langle f - \phi^*, \phi \rangle = 0 \quad \forall \phi \in F \Rightarrow \phi^*$ é um m.a. de f em F .

Tem-se, com $\phi \in F$ arbitrária,

$$\begin{aligned} \|f - \phi\|^2 &= \langle f - \phi, f - \phi \rangle = \langle (f - \phi^*) - (\phi - \phi^*), (f - \phi^*) - (\phi - \phi^*) \rangle \\ &= \langle f - \phi^*, f - \phi^* \rangle - 2 \langle f - \phi^*, \phi - \phi^* \rangle + \langle \phi - \phi^*, \phi - \phi^* \rangle \\ &= \|f - \phi^*\|^2 + \|\phi - \phi^*\|^2 \end{aligned}$$

onde se usou o facto de $\langle f - \phi^*, \phi - \phi^* \rangle = 0$, visto $\phi - \phi^*$ ser um elemento de F . Finalmente, da igualdade

$$\|f - \phi^*\|^2 = \|f - \phi\|^2 - \|\phi - \phi^*\|^2, \quad \phi \in F \quad (30)$$

sai que $\|f - \phi^*\|^2 \leq \|f - \phi\|^2$ e, portanto, $\|f - \phi^*\| \leq \|f - \phi\|$. Sendo $\phi \in F$ qualquer, concluímos que ϕ^* é um m.a. de f . \diamond

Vejamos ainda o seguinte resultado de unicidade.

Teorema 2.5.2 *Seja E um espaço pré-Hilberteano, F um subespaço linear de E e f um elemento de E . A m. a. de f em F é única.*

Demonstração Suponha-se a existência de outra m.a. de f (diferente de ϕ^*) e designe-se por ϕ_1^* . Em particular, a igualdade (30) é válida para esse elemento ϕ_1^* . Substituindo então em (30), ϕ por ϕ_1^* , resulta

$$\|f - \phi^*\|^2 = \|f - \phi_1^*\|^2 - \|\phi_1^* - \phi^*\|^2.$$

Daqui obtém-se a relação $\|f - \phi_1^*\|^2 > \|f - \phi^*\|^2$, o que contraria a hipótese de ϕ_1^* ser m.a..

Cálculo da m. a. num espaço p.H.

Seja F um subespaço com dimensão finita. Pelo Teorema 2.3.1 fica garantida a existência de m.a. em relação à norma induzida pelo produto interno. O teorema 2.5.2 garante a unicidade.

Seja $\{\varphi_0, \dots, \varphi_n\}$ uma base para F . Então é válida a representação, para qualquer $\phi \in F$,

$$\phi = \sum_{k=0}^n a_k \varphi_k, \quad a_0, a_1, \dots, a_n \text{ reais}$$

e, em particular,

$$\phi^* = \sum_{k=0}^n a_k^* \varphi_k \quad (31)$$

Substituindo na relação (29) que caracteriza a m.a., veremos que se obtém um sistema linear de equações nas incógnitas φ_k .

Tem-se

$$\begin{aligned} \langle f - \phi^*, \phi \rangle = 0, \quad \forall \phi \in F &\Leftrightarrow \langle f - \phi^*, \varphi_j \rangle = 0, \quad j = 0, 1, \dots, n \\ &\Leftrightarrow \langle f - \sum_{k=0}^n a_k^* \varphi_k, \varphi_j \rangle = 0, \quad j = 0, \dots, n \\ &\Leftrightarrow \langle \sum_{k=0}^n a_k^* \varphi_k, \varphi_j \rangle = \langle f, \varphi_j \rangle, \quad j = 0, \dots, n \\ &\Leftrightarrow \sum_{k=0}^n a_k^* \langle \varphi_k, \varphi_j \rangle = \langle f, \varphi_j \rangle, \quad j = 0, \dots, n \end{aligned} \quad (32)$$

O sistema de equações normais (32) pode-se escrever na forma matricial

$$\begin{bmatrix} \langle \varphi_0, \varphi_0 \rangle & \langle \varphi_0, \varphi_1 \rangle & \langle \varphi_0, \varphi_2 \rangle & \cdots & \langle \varphi_0, \varphi_n \rangle \\ \langle \varphi_1, \varphi_0 \rangle & \langle \varphi_1, \varphi_1 \rangle & \langle \varphi_1, \varphi_2 \rangle & \cdots & \langle \varphi_1, \varphi_n \rangle \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \langle \varphi_n, \varphi_0 \rangle & \langle \varphi_n, \varphi_1 \rangle & \langle \varphi_n, \varphi_2 \rangle & \cdots & \langle \varphi_n, \varphi_n \rangle \end{bmatrix} \begin{bmatrix} a_0^* \\ a_1^* \\ \vdots \\ a_n^* \end{bmatrix} = \begin{bmatrix} \langle f, \varphi_0 \rangle \\ \langle f, \varphi_1 \rangle \\ \vdots \\ \langle f, \varphi_n \rangle \end{bmatrix} \quad (33)$$

Trata-se de um sistema não homogéneo de $n + 1$ equações lineares nas $n + 1$ incógnitas $a_k^*, k = 0, 1, \dots, n$. Note-se que a matriz do sistema é simétrica. Como o problema de m.a. considerado tem solução única e obter a sua solução é equivalente a resolver (33), então a matriz tem de ser invertível.

Exemplo Determine a recta $p_1^*(x) = a_0^* + a_1^*x$ que é a m. a. mínimos quadrados da função $f(x) = \exp(x)$ no intervalo $[0, 1]$ (com função peso $\omega(x) \equiv 1$).

Pretende-se o elemento $p_1^*(x) \in \mathbf{P}_1([0, 1])$ que minimiza a norma do erro (comparar (27)):

$$\min_{p_1 \in \mathbf{P}_1} \left(\int_0^1 (\exp(x) - p_1(x))^2 dx \right)^{1/2}$$

Esse elemento existe e é único, pelos Teoremas 2.3.1 e 2.5.2. Os coeficientes a_0^* e a_1^* são obtidos resolvendo o sistema (33). Seja $\varphi_0(x) = 1$ e $\varphi_1(x) = x$ a base escolhida para $\mathbf{P}_1([0, 1])$. Tem-se

$$\begin{aligned} \langle \varphi_0, \varphi_0 \rangle &= \int_0^1 1 dx = 1, \\ \langle \varphi_0, \varphi_1 \rangle &= \langle \varphi_1, \varphi_0 \rangle = \int_0^1 x dx = 1/2, \\ \langle \varphi_1, \varphi_1 \rangle &= \int_0^1 x^2 dx = 1/3, \\ \langle f, \varphi_0 \rangle &= \int_0^1 e^x dx = e - 1, \\ \langle f, \varphi_1 \rangle &= \int_0^1 x e^x dx = 1 \end{aligned}$$

e obtém-se o sistema normal

$$\begin{aligned} a_0^* + (1/2)a_1^* &= e - 1 \\ (1/2)a_0^* + (1/3)a_1^* &= 1 \end{aligned}$$

donde $a_0^* = 4e - 10 \simeq 0.8731$, $a_1^* = 6(3 - e) \simeq 1.6903$.

Finalmente, $p_1^*(x) = 0.8731 + 1.6903x$ é a m.a.pretendida.

2.6 Sistemas ortogonais

Na prática, o sistema de equações normais (33) pode ser mal condicionado e isso causa dificuldades numéricas na sua resolução. Este problema pode ser ultrapassado escolhendo uma base ortogonal para o espaço F .

Como exemplo, suponhamos que se pretende determinar a m.a. mínimos quadrados polinomial, p_n^* , de uma função f , no intervalo $[0, 1]$. Expressando p_n^* como uma combinação linear de funções $\varphi_j(x) = x^j$, $j = 0, 1, \dots, n$, obtém-se

$$\langle \varphi_j, \varphi_k \rangle = \int_0^1 x^{j+k} dx = \frac{1}{j+k+1}$$

e a matriz do sistema normal é matriz de Hilbert

$$\begin{bmatrix} 1 & 1/2 & 1/3 & \cdots & 1/(n+1) \\ 1/2 & 1/3 & 1/4 & \cdots & 1/(n+2) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1/(n+1) & 1/(n+2) & 1/(n+3) & \cdots & 1/(2n+1) \end{bmatrix}$$

Trata-se de uma matriz mal condicionada. Pode-se evitar este tipo de dificuldade escolhendo uma base ortogonal para o espaço dos polinómios de grau $\leq n$.

Definição 2.6.1 Um conjunto S de elementos de um espaço pré-Hilbertiano E diz-se ortogonal se $\langle u, v \rangle = 0$ quaisquer que sejam $u, v \in S$, com $u \neq v$ (e usa-se a notação $u \perp v$). Se, além disso, $\langle u, u \rangle = 1$ o sistema diz-se ortonormado.

Teorema 2.6.1 (Teor. de Pitágoras): se u_1, u_2, \dots, u_n são ortogonais então

$$\|u_1 + u_2 + \dots + u_n\|^2 = \|u_1\|^2 + \|u_2\|^2 + \dots + \|u_n\|^2$$

Teorema 2.6.2 Um conjunto finito de elementos ortogonais não nulos, $\{u_0, \dots, u_n\}$, é um conjunto de elementos linearmente independentes.

Demonstração Considere-se a combinação linear nula $c_0 u_0 + c_1 u_1 + \dots + c_n u_n = 0$. Então, para cada $k = 0, 1, \dots, n$, vem

$$0 = \langle \sum_{i=0}^n c_i u_i, u_k \rangle = c_k \langle u_k, u_k \rangle \Rightarrow c_k = 0, \quad k = 0, 1, \dots, n. \quad \diamond$$

O seguinte teorema diz-nos que é possível ortogonalizar um conjunto finito de elementos linearmente independentes.

Teorema 2.6.3 (Processo de ortogonalização de Gram Schmidt) Seja $\{e_0, e_1, \dots, e_n\}$ um conjunto de elementos linearmente independentes do espaço p.H. E . Sejam $\{\varphi_0, \varphi_1, \dots, \varphi_n\}$ os elementos definidos por

$$\varphi_j = \frac{v_j}{\|v_j\|}, \quad j = 0, 1, 2, \dots, n \quad \text{onde} \quad v_0 = e_0, \quad v_j = e_j - \sum_{k=0}^{j-1} \langle e_j, \varphi_k \rangle \varphi_k, \quad j \geq 1, \quad (34)$$

Então o conjunto $\{v_0, v_1, \dots, v_n\}$ é uma base ortogonal para E o conjunto $\{\varphi_0, \varphi_1, \dots, \varphi_n\}$ é uma base ortonormada.

Expressão da m.a. num espaço p.H. usando uma base ortogonal

Suponhamos então que a base considerada para F , $\{\varphi_0, \dots, \varphi_n\}$, é um sistema ortogonal. Nesse caso, tem-se $\langle \varphi_j, \varphi_k \rangle = 0$ se $i \neq j$ e a matriz do sistema de equações normais (32) reduz-se a uma matriz diagonal.

$$a_k^* \langle \varphi_k, \varphi_k \rangle = \langle f, \varphi_k \rangle, \quad k = 0, 1, \dots, n. \quad (35)$$

donde

$$a_k^* = \frac{\langle f, \varphi_k \rangle}{\langle \varphi_k, \varphi_k \rangle}, \quad k = 0, 1, \dots, n. \quad (36)$$

Note-se que se a base for ortonormada, então vem simplesmente $a_k^* = \langle f, \varphi_k \rangle$, $k = 0, 1, \dots, n$.

Substituindo (36) em (31) obtém-se a seguinte expressão para a melhor aproximação de f em F

$$\phi^* = \sum_{k=0}^n \frac{\langle f, \varphi_k \rangle}{\langle \varphi_k, \varphi_k \rangle} \varphi_k \quad (37)$$

Polinómios ortogonais

Consideremos o espaço $C([a, b])$ com o produto interno definido por (25). O processo de ortogonalização de Gram Schmidt aplicado ao conjunto $\{1, x, x^2, \dots, x^n\}$ permite obter diferentes bases ortogonais $\{\varphi_0, \varphi_1, \dots, \varphi_n\}$ para P_n , correspondentes a diferentes funções peso $\omega(x)$. Uma maneira alternativa de construir polinómios ortogonais é usar o teorema seguinte.

Teorema 2.6.4 (Relação de recorrência tripla) *O conjunto de polinómios $\{\varphi_0, \varphi_1, \dots, \varphi_n\}$, definidos da maneira a seguir indicada, é ortogonal em $[a, b]$ em relação produto interno definido por (25), isto é,*

$$\langle \varphi_j, \varphi_k \rangle = \int_a^b \omega(x) \varphi_j(x) \varphi_k(x) dx = 0, \quad j \neq k.$$

Os polinómios são definidos recursivamente através de

$$\varphi_0(x) = 1, \quad \varphi_1(x) = x - B_1, \quad \text{onde } B_1 = \frac{\langle x\varphi_0, \varphi_0 \rangle}{\langle \varphi_0, \varphi_0 \rangle} \quad (38)$$

e

$$\varphi_k(x) = (x - B_k)\varphi_{k-1}(x) - C_k\varphi_{k-2}(x), \quad k \geq 2 \quad (39)$$

com

$$B_k = \frac{\langle x\varphi_{k-1}, \varphi_{k-1} \rangle}{\langle \varphi_{k-1}, \varphi_{k-1} \rangle} \quad C_k = \frac{\langle x\varphi_{k-1}, \varphi_{k-2} \rangle}{\langle \varphi_{k-2}, \varphi_{k-2} \rangle} \quad (40)$$

Note-se que a relação (39) produz polinómios ortogonais com coeficiente unitário da potência mais elevada.

Exemplo Os polinómios de Legendre são ortogonais no intervalo $[-1, 1]$ com função peso $\omega(x) \equiv 1$. Na sua definição clássica, eles verificam $\varphi_n(1) = 1, n \geq 1$. Utilizando a relação (39), vamos obter o que chamaremos de polinómios de Legendre não normalizados. Aplicando o Teorema 2.6.4, vem $\varphi_0(x) = 1$,

$$B_1 = \frac{\int_{-1}^1 x dx}{\int_{-1}^1 dx} = 0, \quad \varphi_1(x) = (x - B_1)P_0(x) = x.$$

Ainda

$$B_2 = \frac{\int_{-1}^1 x^3 dx}{\int_{-1}^1 x^2 dx} = 0, \quad C_2 = \frac{\int_{-1}^1 x^2 dx}{\int_{-1}^1 1 dx} = \frac{1}{3}$$

logo

$$\varphi_2(x) = (x - B_2)\varphi_1(x) - C_2\varphi_0(x) = (x - 0)x - \frac{1}{3} \cdot 1 = x^2 - \frac{1}{3}.$$

Analogamente se poderiam obter polinómios de grau mais elevado: $\varphi_3(x) = x^3 - (3/5)x$, $\varphi_4(x) = x^4 - (6/7)x^2 + 3/35$

A seguir damos exemplos de polinómios ortogonais obtidos pelo processo de Gram Schmidt, os quais podem também ser gerados pelas relações de recorrência indicadas para cada caso. Eles diferem por uma constante multiplicativa dos que se obtêm usando (39). Do ponto de vista da aproximação é irrelevante se os coeficientes dos polinómios da base diferem por uma constante multiplicativa.

Exemplos de polinómios ortogonais

(1) Polinómios de Legendre, $P_n(x)$

$$\omega(x) \equiv 1, \quad x \in [a, b] = [-1, 1].$$

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

com $P_0(x) = 1$ Tem-se que P_n é um polinómio de grau n , $P_n(1) = 1, n \geq 1$ Além disso

$$\langle P_n, P_m \rangle = \int_{-1}^1 P_n(x)P_m(x)dx = \begin{cases} 0, & n \neq m \\ \frac{2}{2n+1}, & n = m. \end{cases}$$

Fórmula de recorrência:

$$P_1(x) = x$$

$$P_{n+1}(x) = \frac{2n+1}{n+1}xP_n(x) - \frac{n}{n+1}P_{n-1}(x), \quad n \geq 1$$

(2) Polinômios de Chebyshev, $T_n(x)$

Já foram definidos anteriormente. Eles são dados por $T_n(x) = \cos(n \arccos x)$, $n \geq 0$ e são ortogonais com respeito à função peso seguinte

$$\omega(x) = \frac{1}{\sqrt{1-x^2}}, \quad x \in [a, b] = [-1, 1].$$

Tem-se que T_n é de grau n , $T_n(1) = 1, \forall n$ e ainda

$$\langle T_n, T_m \rangle = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) dx = \begin{cases} 0, & n \neq m \\ \pi, & m = n = 0 \\ \frac{\pi}{2}, & n = m > 0. \end{cases}$$

Fórmula de recorrência:

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

(3) Polinômios de Laguerre, $L_n(x)$

$$\omega(x) = e^{-x}, \quad x \in [a, b] = [0, \infty).$$

$$L_n(x) = \frac{1}{n!e^{-x}} \frac{d^n}{dx^n} (x^n e^{-x}), \quad n \geq 0$$

L_n é de grau n , $L_0(x) = 1, L_1(x) = 1 - x$

$$\langle L_n, L_m \rangle = \int_0^\infty e^{-x} L_n(x) L_m(x) dx = \begin{cases} 0, & n \neq m \\ 1, & m = n \end{cases}$$

Fórmula de recorrência:

$$L_{n+1}(x) = \frac{1}{n+1}(2n+1-x)L_n(x) - \frac{n}{n+1}L_{n-1}(x), \quad n \geq 1$$

Propriedades dos polinómios ortogonais

Teorema 2.6.5 *Seja $\{\varphi_j(x), j \geq 0\}$ um sistema ortogonal em $[a, b]$ com função peso $\omega(x)$. (Admite-se implicitamente que o grau de φ_j é $j, j \geq 0$). Se $p(x)$ é um polinómio de grau $\leq m$ então*

$$p(x) = \sum_{j=0}^m \frac{\langle p, \varphi_j \rangle}{\langle \varphi_j, \varphi_j \rangle} \varphi_j(x)$$

Demonstração Comece-se por escrever $p(x)$ na forma

$$p(x) = \sum_{k=0}^m \beta_k \varphi_k(x)$$

Então, para cada $j = 0, 1, \dots, m$,

$$\langle p, \varphi_j \rangle = \sum_{k=0}^m \beta_k \langle \varphi_k, \varphi_j \rangle = \beta_j \langle \varphi_j, \varphi_j \rangle$$

donde o resultado pretendido.

Corolário *Se $p(x)$ é um polinómio de grau $\leq m - 1$ então $\langle p, \varphi_m \rangle = 0$.*

Teorema 2.6.6 *Seja $\{\varphi_j(x), j \geq 0\}$ um sistema ortogonal em $[a, b]$ com função peso $\omega(x)$. Então o polinómio $\varphi_n(x)$ tem exactamente n raízes reais distintas no intervalo aberto (a, b) .*

Exercícios

1. Considere a seguinte tabela de valores de uma função f

x_i	-1	0	1	2
$f(x_i)$	1	1	1	2

- (a) Usando a fórmula de Newton com diferenças divididas, construa o polinómio interpolador de f de grau menor ou igual a 3.
- (b) Sabendo que $f'''(x) = 4x - 1$, utilize a alínea anterior para determinar a expressão exacta de f . (EXAME 04.03.94)
2. Seja f uma função real definida em $[a, b]$ e suponhamos conhecidos os valores de $f(x_i)$ em $n+1$ pontos distintos $x_0, \dots, x_n \in [a, b]$. Prove que

$$f[x_0, \dots, x_n] = \sum_{j=0}^n \frac{f_j}{\prod_{\substack{0 \leq i \leq n \\ i \neq j}} (x_j - x_i)}$$

e obtenha a seguinte fórmula de recorrência

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

3. Considere a seguinte tabela de valores de uma função f

x_i	-2	0	2	4
$f(x_i)$	-17	5	-5	c

que se sabe ser um polinómio da forma $f(x) = x^3 + a_2x^2 + a_1x + 5$, a_1, a_2 reais.

Que relação existe entre o polinómio interpolador de f nos 3 primeiros pontos e a função f ? Determine o valor de $f(4)$. (TESTE 15.12.95)

4. Sejam x_1, x_2, \dots, x_M valores reais distintos e f_1, f_2, \dots, f_M os valores correspondentes de uma função f nesses pontos. Prove que existe uma única função F_M da forma

$$F_M(x) = \sum_{j=1}^M c_j e^{jx},$$

para a qual se tem $F_M(x_i) = f_i, i = 1, 2, \dots, M$. (EXAME 17.01.94)

5. Seja $f : [0, 1] \rightarrow \mathcal{R}$ uma função de classe C^3 .

(a) Mostre que existe um e um só polinómio p de grau ≤ 2 tal que

$$p(0) = f(0); \quad p(1) = f(1); \quad \int_0^1 p(x) dx = \int_0^1 f(x) dx.$$

(b) Mostre que existe $\xi \in]0, 1[$ tal que $p(\xi) = f(\xi)$.

(c) Utilize o resultado da alínea anterior para mostrar que existe $M \in \mathcal{R}_0^+$ tal que

$$\max_{x \in [0,1]} |f(x) - p(x)| \leq M/6.$$

6. Seja $f \in C^1[a, b]$ e suponha que $f(a)f(b) < 0$ e $f'(x) \neq 0, \forall x \in [a, b]$; logo existe um único $z \in (a, b)$ tal que $f(z) = 0$.

(a) Sejam $x_0, x_1, \dots, x_n, n + 1$ pontos distintos em $[a, b]$ e sejam y_0, y_1, \dots, y_n os valores de f nesses pontos, i.e. $y_j = f(x_j), j = 0, \dots, n$. Dê uma expressão para o polinómio $q_n \in \mathcal{P}_n$ que interpola $f^{-1}(x)$ nos nós y_0, y_1, \dots, y_n .

(b) Notando que $z = f^{-1}(0)$, utilize o polinómio q_3 e os seguintes valores tabelados para aproximar a raiz da equação $f(x) = e^{-x} - x = 0$.

x	0.3	0.4	0.5	0.6
e^{-x}	0.740818	0.670320	0.606531	0.548812

7. Seja $p_n(x)$ o polinómio interpolador de f em $n + 1$ pontos x_0, x_1, \dots, x_n , igualmente espaçados no intervalo $[a, b]$ e tais que $x_0 = a$ e $x_n = b$.

(a) Mostre que, no caso de existir uma constante $M > 0$ tal que $\|f^{(n+1)}\|_\infty \leq M^n$, tem-se $\lim_{n \rightarrow \infty} \|f - p_n\|_\infty = 0$. Conclua que para a função $f(x) = e^{3x}$ se verifica

$$\lim_{n \rightarrow \infty} p_n(x) = f(x)$$

(b) Prove que $\lim_{n \rightarrow \infty} \|f - p_n\|_\infty = 0$ também no caso em que $\|f^{(n+1)}\|_\infty \leq M^n(n+1)!$, com $M \in (0, e/(b-a))$.

8. Seja f uma função indefinidamente diferenciável para $x \geq 0$ e tal que $|f^m(x)| \leq M$, para todo $x \geq 0$, e para todo $m = 0, 1, 2, \dots$. Seja p_n o polinómio interpolador de f nos pontos $0, \frac{1}{2}, \frac{2}{2}, \frac{3}{2}, \dots, \frac{n}{2}$. Mostre que

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

para cada $x \geq 0$. (EXAME 31.01.96)

9. Seja p_4 o polinómio de grau ≤ 4 que resolve o seguinte problema de interpolação de Hermite

$$p_4^{(j)}(a) = f^{(j)}(a), j = 0, 1 \quad p_4^{(k)}(b) = f^{(k)}(b), k = 0, 1, 2.$$

Prove que, para cada $x \in [a, b]$, existe $\xi = \xi(x) \in (a, b)$, tal que

$$f(x) - p_4(x) = \frac{(x-a)^2(x-b)^3}{5!} f^{(5)}(\xi).$$

Sugestão: faça $F(t) = (f(t) - p_4(t)) - (f(x) - p_4(x)) \frac{\Omega(t)}{\Omega(x)}$, $\Omega(t) = (t-a)^2(t-b)^3$. (EXAME 31.01.97)

10. Demonstre a condição suficiente do Teorema de Chebyshev (utilize o Teorema de La Vallée Poussin).
11. Considere a função $f(x) = |x|$, $-1 \leq x \leq 1$. Por aplicação do Teorema de Chebyshev, prove que a melhor aproximação uniforme de f em $\mathcal{P}_3[-1, 1]$ é dada por

$$p_3^*(x) = x^2 + \frac{1}{8}, \quad -1 \leq x \leq 1.$$

12. Determine, utilizando o Teorema de Chebyshev e observações geométricas, a melhor aproximação uniforme da função $f(x) = e^x$, $x \in [0, 1]$: no espaço $\mathcal{P}_0[0, 1]$; no espaço $\mathcal{P}_1[0, 1]$.
13. Seja $f(x) = (x^3 + x^2)/100$, $x \in [-1, 1]$. Determine o polinómio $p_2 \in \mathcal{P}_2[-1, 1]$ tal que

$$\max_{x \in [-1, 1]} |f(x) - p_2(x)| \quad \text{é mínimo.}$$

Indique o valor do mínimo.

14. (a) Prove que $p_{n-1}^*(x) = x^n - 2^{1-n}T_n(x)$ é a melhor aproximação uniforme da função $f(x) = x^n$, $x \in [-1, 1]$, em $\mathcal{P}_{n-1}[-1, 1]$.
- (b) Calcule $\|f - p_{n-1}^*\|_\infty$
15. Seja $T_n(x)$, $x \in [-1, 1]$ o polinómio de Chebyshev de grau n . Prove que no conjunto $\bar{\pi}_n$ dos polinómios mónicos de grau n , definidos em $[-1, 1]$, o polinómio

$$\bar{T}_n = \frac{T_n}{2^{n-1}}$$

é a melhor aproximação uniforme da função nula.

16. Considere o espaço linear $C^1[a, b]$ e o operador definido por

$$L(f) = \left(\int_a^b (f'(x))^2 dx \right)^{1/2}, \quad f \in C^1[a, b].$$

- (a) Sabendo que $L(f + g) \leq L(f) + L(g)$, $\forall f, g \in C^1([a, b])$, prove que L define uma seminorma (que não é norma) em $C^1[a, b]$.
- (b) Recorrendo à teoria da melhor aproximação e usando L , mostre que existem constantes reais $\bar{\alpha}, \bar{\beta}$ tais que

$$\int_a^b (\cos x + \bar{\alpha}x + \bar{\beta})^2 dx \leq \int_a^b (\cos x + \alpha x + \beta)^2 dx, \quad \forall \alpha, \beta \in \mathcal{R}.$$

17. Utilizando o exercício 14, prove que a melhor aproximação uniforme da função nula, relativamente ao intervalo $[a, b]$, no conjunto $\bar{\pi}_n$ dos polinómios mónicos de grau n , é o polinómio p^* definido por

$$p^*(x) = 2 \left(\frac{b-a}{4} \right)^n T_n \left(\frac{2x}{b-a} - \frac{a+b}{b-a} \right).$$

18. Determine a melhor aproximação da função nula relativamente a $[-1, 3]$, da forma

$$\frac{x^3}{3} + bx^2 + cx + d.$$

19. Obtenha os 3 primeiros polinômios ortogonais de Legendre, utilizando a relação de recorrência tripla para polinômios ortogonais. Determine em seguida a melhor aproximação dos mínimos quadrados da função $f(x) = x^4$ no espaço $\mathcal{P}_2[-1, 1]$.

20. Demonstre a seguinte propriedade dos polinômios de Chebyshev $T_n(x)$, $n = 0, 1, \dots$

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

21. Seja E um espaço com produto interno e seja $\{\psi_0, \psi_1, \dots, \psi_n, \dots\}$ um sistema de polinômios ortogonais em $[a, b]$ (por ex. obtidos pelo p.o. G.Schmidt).

(a) Demonstre as seguintes propriedades:

(i) Se q_n é um polinômio de grau $\leq n$, então

$$q_n(x) = \sum_{k=0}^n \frac{\langle q_n, \psi_k \rangle}{\langle \psi_k, \psi_k \rangle} \psi_k(x).$$

(ii) Se q_k é um polinômio de grau $\leq k$, então

$$\langle q_k, \psi_n \rangle = 0, \quad \text{para } k \leq n - 1.$$

(b) Seja $f \in E$ e considere a norma induzida pelo produto interno. Mostre que

(i)

$$\|f - \sum_{k=0}^n a_k \psi_k\| \text{ é mínima quando } a_k = \frac{\langle f, \psi_k \rangle}{\|\psi_k\|^2}.$$

(ii)

$$\sum_{k=0}^{+\infty} \frac{(\langle f, \psi_k \rangle)^2}{\|\psi_k\|^2} \leq \|f\|^2. \quad (\text{Desigualdade de Bessel})$$

22. Considere os seguintes problemas de minimização

a)

$$\min_{\alpha \in \mathbb{R}} \max_{x \in [-1, 1]} |x - \alpha x^2|$$

Há alguma contradição em relação ao que se poderia concluir pelo teorema de Haar ?

b)

$$\min_{\alpha \in \mathbb{R}} \max_{x \in [-1, 1]} |x^2 - 1 - \alpha x^2|$$

Bibliografia

- Atkinson, K., *An Introduction to Numerical Analysis*, Wiley, 1978.
- Blum, E. K., *Numerical Analysis and Computation: Theory and Practice*, Addison-Wesley, 1972.
- Davis, P. J., *Interpolation and Approximation*, Dover, 1975.
- Isaacson, E. & Keller, H. B., *Analysis of Numerical Methods*, John Wiley & Sons, 1966.
- Johnson, L. W. & Riess, R. D., *Numerical Analysis*, Addison-Wesley, 1977.
- Kreysig, E., *Introductory Functional Analysis with Applications*, John Wiley & Sons 1978.
- Schwarz, H. R., *Numerical Analysis, A Comprehensive Introduction*, John Wiley & Sons, 1989.
- Stoer, J. & Burlich, R., *Introduction to Numerical Analysis*, Springer-Verlag 1980.
- Stroud, A. H., *Numerical Quadrature and Solution of Ordinary Differential Equations*, Applied Mathematical Sciences, vol. 10, Springer-Verlag 1974.
- Valença, M. R., *Métodos Numéricos*, Instituto Nacional de Investigação Científica 1990.