

ANÁLISE NUMÉRICA II

Lic Matemática

Maria Joana Soares

2013/2104

Interpolação Polinomial

Problema de Interpolação Polinomial

Dados

- ▶ x_1, \dots, x_n distintos
- ▶ y_1, \dots, y_n

Construir um polinómio $P \in \mathcal{P}_{n-1}$ que satisfaça

$$P(x_i) = y_i, \quad i = 1, 2, \dots, n.$$

- ▶ $\mathcal{P}_{n-1} \equiv$ espaço dos polinómios (de coef.reais) de grau não superior a $n-1$
- ▶ Geralmente $y_i = y(x_i)$; nesse caso, dizemos que P interpola y nos nós x_i .

INTERPOLAÇÃO POLINOMIAL

Interpolação Polinomial

Existência e Unicidade

Teorema

Sejam $x_1, \dots, x_n \in \mathbb{R}$, distintos, e sejam y_1, \dots, y_n números reais dados. Então, **existe um e um só** polinómio $P \in \mathcal{P}_{n-1}$ que satisfaz

$$P(x_i) = y_i; \quad i = 1, 2, \dots, n. \quad (1)$$

Dem:

Existência

Se encontrarmos n polinómios de grau $n-1$, $L_1(x), L_2(x), \dots, L_n(x)$, tais que $L_i(x_j) = 1$ se $i = j$ e $L_i(x_j) = 0$ se $i \neq j$, então o polinómio P dado por

$$P(x) = L_1(x)y_1 + L_2(x)y_2 + \dots + L_n(x)y_n$$

estará em \mathcal{P}_{n-1} e satisfará (1).

Como encontrar os polinómios $L_i(x)$?

1

$$L_i \text{ de grau } n-1 \text{ e } L_i(x_j) = 0; j \neq i \\ \iff L_i(x) = A_i(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)$$

2

$$L_i(x_i) = 1 \iff A_i(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n) = 1 \\ \iff A_i = \frac{1}{(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)}$$

Logo:

$$L_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^n \left(\frac{x-x_j}{x_i-x_j} \right); i = 1, \dots, n. \quad (2)$$

Definição (Polinómios de Lagrange)

Os polinómios dados por (2) são chamados polinómios de Lagrange relativos aos pontos x_i .

Conclusão:

O polinómio dado por

$$P(x) = \sum_{i=1}^n \left\{ \prod_{\substack{j=1 \\ j \neq i}}^n \left(\frac{x-x_j}{x_i-x_j} \right) \right\} y_i \quad (3)$$

é o único polinómio de grau não superior a $n-1$ que satisfaz as condições de interpolação

$$P(x_i) = y_i; i = 1, \dots, n.$$

Definição

A forma (3) do polinómio interpolador é dita **forma de Lagrange** ou **forma cardinal** desse polinómio.

Nota: Em geral, designaremos o polinómio anterior por P_{n-1} , para explicitar qual o seu grau.

Dem (cont):

Unicidade

Sejam P e Q polinómios em \mathcal{P}_{n-1} tais que

$$P(x_i) = y_i \text{ e } Q(x_i) = y_i; i = 1, \dots, n.$$

Consideremos o polinómio D que é a diferença de P e Q , i.e. $D := P - Q$. Então:

1 D é um polinómio em \mathcal{P}_{n-1} ;

2

$$D(x_i) = P(x_i) - Q(x_i) = y_i - y_i = 0; i = 1, \dots, n$$

$D \in \mathcal{P}_{n-1}$ e tem n zeros $\Rightarrow D = P - Q$ é o polinómio identicamente nulo $\Rightarrow P$ e Q são o mesmo polinómio.

Exemplo (Interpolação Linear, $n = 2$)

$$x_1, x_2, y_1, y_2 \rightarrow P = P_1 \in \mathcal{P}_1 \\ L_1(x) = \frac{x-x_2}{x_1-x_2}, \quad L_2(x) = \frac{x-x_1}{x_2-x_1}.$$

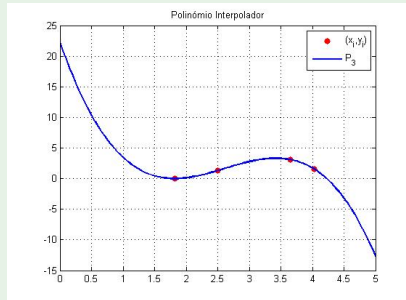
$$P_1(x) = \frac{x-x_2}{x_1-x_2}y_1 + \frac{x-x_1}{x_2-x_1}y_2 \\ = \left(\frac{x-x_1+x_1-x_2}{x_1-x_2} \right) y_1 + \frac{x-x_1}{x_2-x_1}y_2 \\ = y_1 + \left(\frac{y_2-y_1}{x_2-x_1} \right) (x-x_1)$$

↓

Equação da reta que passa pelos pontos (x_1, y_1) e (x_2, y_2) .

Exemplo

x_i	1.82	2.50	3.65	4.03
y_i	0.00	1.30	3.10	2.52



Erro em Interpolação Polinomial

$$y_i = y(x_i) \rightarrow y(x) \approx P_{n-1}(x) \rightarrow y(x) - P_{n-1}(x) ?$$

Teorema

Seja P_{n-1} o polinômio de grau $\leq n-1$ interpolador de uma dada função y em n pontos distintos x_1, x_2, \dots, x_n . Se $y \in C^n[a, b]$, onde $[a, b]$ é um intervalo que contém os pontos x_i , então, para todo o ponto $x \in [a, b]$, existe $\xi_x = \xi_x(x) \in (a, b)$ tal que

$$E_{n-1}(x) := y(x) - P_{n-1}(x) = \frac{y^{(n)}(\xi_x)}{n!} \prod_{i=1}^n (x - x_i)$$

Nota: Se $x \in [\min\{x_1, \dots, x_n\}, \max\{x_1, \dots, x_n\}] \rightarrow$ *interpolação*.
Se $x \notin [\min\{x_1, \dots, x_n\}, \max\{x_1, \dots, x_n\}] \rightarrow$ *extrapolação*.

Demonstração fórmula do erro

- ▶ Se $x = x_i$ para algum i , o resultado é trivialmente verdadeiro.
- ▶ Se x não é nenhum dos pontos de interpolação, definamos a seguinte função, na variável t , considerando o valor de x fixo:

$$Y(t) := y(t) - P_{n-1}(t) - \left\{ \frac{y(x) - P_{n-1}(x)}{\Pi_n(x)} \right\} \Pi_n(t),$$

onde $\Pi_n(t) := \prod_{i=1}^n (t - x_i)$. Então:

- 1 $Y \in C^n[a, b]$.
- 2 $Y(x_i) = y(x_i) - P_{n-1}(x_i) - E_{n-1}(x) \frac{\Pi_n(x_i)}{\Pi_n(x)} = 0$; $i = 1, \dots, n$.
- 3 $Y(x) = y(x) - P_{n-1}(x) - E_{n-1}(x) \frac{\Pi_n(x)}{\Pi_n(x)} = 0$.

⇓ (T. Rolle generalizado)

$$\exists \xi_x \in (\min\{x_1, \dots, x_n, x\}, \max\{x_1, \dots, x_n, x\}) : Y^{(n)}(\xi_x) = 0.$$

Demonstração (cont.)

Mas,

$$Y^{(n)}(t) = y^{(n)}(t) - \frac{E_{n-1}(x)}{\Pi_n(x)} n!,$$

donde, temos

$$\begin{aligned} Y^{(n)}(\xi_x) = 0 &\iff y^{(n)}(\xi_x) - \frac{E_{n-1}(x)}{\Pi_n(x)} n! = 0 \\ &\iff E_{n-1}(x) = \frac{y^{(n)}(\xi_x)}{n!} \Pi_n(x), \\ &\iff E_{n-1}(x) = \frac{y^{(n)}(\xi_x)}{n!} \prod_{i=1}^n (x - x_i), \end{aligned}$$

como pretendíamos provar.

Erro em Interpolação Polinomial

Em geral, não conhecemos $y^{(n)}$ nem $\xi_x = \xi(x)$. Se conhecermos M_n tal que $\max_{x \in [a,b]} |f^{(n)}(x)| \leq M_n$ teremos o seguinte **majorante** para o erro (em valor absoluto)

$$|E_{n-1}(x)| = |y(x) - P_{n-1}(x)| \leq \frac{M_n}{n!} \prod_{i=1}^n |x - x_i|.$$

Para um valor de x concreto podemos usar a fórmula anterior. Interessa-nos também obter um majorante que seja válido para qualquer ponto x no intervalo definido pelos pontos x_i , isto é, para todo o x em $[\min\{x_i\}, \max\{x_i\}]$. Para isso, bastará encontrar o máximo valor de

$$\prod_{i=1}^n |x - x_i|.$$

Forma de Newton do polinómio interpolador

P_{n-2} – pol. interpolador dos valores y_1, \dots, y_{n-1} nos pontos x_1, \dots, x_{n-1} .

P_{n-1} – pol. interpolador dos valores y_1, \dots, y_{n-1}, y_n nos pontos x_1, \dots, x_{n-1}, x_n .

Problema

Obter uma representação do polinómio interpolador que permita obter P_{n-1} , a partir de P_{n-2} , por junção de mais um termo; isto é, determinar $Q_{n-1}(x)$ de tal modo que

$$\begin{aligned} P_{n-1}(x) &= P_{n-2}(x) + Q_{n-1}(x); \quad n = 2, 3, \dots, \\ P_0(x) &= y_1 \end{aligned}$$

- 1 Q_{n-1} deve ser um polinómio de grau $\leq n - 1$.
- 2 $P_{n-1}(x_i) = P_{n-2}(x_i) = y_i; i = 1, \dots, n - 1$
 $\implies Q_{n-1}(x_i) = 0; i = 1, \dots, n - 1$
 $\implies Q_{n-1}(x) = a_{n-1}(x - x_1) \dots (x - x_{n-1})$.

Exemplo (Interpolação Linear $n = 2$)

Temos

$$|y(x) - P_1(x)| \leq \frac{M_2}{2} |(x - x_1)(x - x_2)|.$$

Se $h := x_2 - x_1$, facilmente se verifica que

$$\max_{x \in [x_1, x_2]} |(x - x_1)(x - x_2)| = \frac{h}{2} \cdot \frac{h}{2} = \frac{h^2}{4}.$$

Assim, tem-se

$$|E_1(x)| \leq \frac{h^2}{8} M_2, \quad \forall x \in [x_1, x_2].$$

Forma de Newton do polinómio interpolador

$$P_{n-1}(x) = P_{n-2}(x) + a_{n-1} \prod_{i=1}^{n-1} (x - x_i)$$

Problema

Determinação da constante a_{n-1} .

Temos

$$P_{n-1}(x) = \sum_{i=1}^n \left\{ \prod_{\substack{j=1 \\ j \neq i}}^n \left(\frac{x - x_j}{x_i - x_j} \right) \right\} y_i$$

Igualando o coeficiente de x^{n-1} nas duas expressões acima, tem-se

$$a_{n-1} = \sum_{i=1}^n \left\{ \frac{y_i}{\prod_{\substack{j=1 \\ j \neq i}}^n (x_i - x_j)} \right\}.$$

Diferença dividida

Definição (Diferença dividida)

Chamamos *diferença dividida* de ordem $n - 1$ dos valores y_1, \dots, y_n (relativa aos pontos x_1, \dots, x_n) e denotamos por $[y_1, \dots, y_n]$ o valor dado pela expressão anterior, i.e.

$$[y_1, \dots, y_n] = \sum_{i=1}^n \left\{ \frac{y_i}{\prod_{\substack{j=1 \\ j \neq i}}^n (x_i - x_j)} \right\}$$

Se $y_i = y(x_i)$, usamos também a notação $y[x_1, \dots, x_n]$ e falamos da diferença dividida da função y relativa aos pontos x_1, \dots, x_n .

Nota

A diferença dividida $[y_1, \dots, y_n]$ (relativa aos pontos x_1, \dots, x_n) é o coeficiente de x^{n-1} do polinómio P_{n-1} interpolador de y_1, \dots, y_n em x_1, \dots, x_n .

Polinómio interpolador

Forma de Newton com diferenças divididas

Assim, temos a seguinte expressão para o polinómio P_{n-1} interpolador de y_1, \dots, y_n nos pontos x_1, \dots, x_n , a que chamamos *forma de Newton com diferenças divididas*:

$$\begin{aligned} P_{n-1}(x) &= y_1 + [y_1, y_2](x - x_1) + [y_1, y_2, y_3](x - x_1)(x - x_2) \\ &\quad + \dots + [y_1, y_2, \dots, y_n](x - x_1)(x - x_2) \dots (x - x_{n-1}) \\ &= y_1 + \sum_{i=2}^n [y_1, \dots, y_i](x - x_1) \dots (x - x_{i-1}) \end{aligned}$$

Exemplo (Polinómio linear)

Temos x_1, x_2 (distintos) e y_1, y_2 .

A diferença dividida $[y_1, y_2]$ é dada por

$$\begin{aligned} [y_1, y_2] &= \frac{y_1}{(x_1 - x_2)} + \frac{y_2}{(x_2 - x_1)} \\ &= \frac{y_2 - y_1}{x_2 - x_1} \end{aligned}$$

Temos, então

$$\begin{aligned} P_1(x) &= y_1 + [y_1, y_2](x - x_1) \\ &= y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) \end{aligned}$$

→ Forma mais habitual da equação da reta que passa pelos pontos (x_1, y_1) e (x_2, y_2) .

Relembremos

$$[y_1, y_2, \dots, y_n] = \sum_{i=1}^n \left\{ \frac{y_i}{\prod_{\substack{j=1 \\ j \neq i}}^n (x_i - x_j)} \right\}$$

- 1 Porquê o nome *diferenças divididas*?
- 2 Haverá uma forma mais simples de as calcular?

Começemos por notar que a fórmula acima mostra que o valor de uma diferença dividida $[y_1, \dots, y_n]$ é independente da ordem dos seus argumentos, isto é, se p for uma permutação do conjunto $\{1, \dots, n\}$, tem-se

$$[y_1, y_2, \dots, y_n] = [y_{p(1)}, y_{p(2)}, \dots, y_{p(n)}].$$

Seja $P_{n-1}(x)$ o polinómio interpolador de y_1, \dots, y_n nos pontos x_1, \dots, x_n .

Consideremos a sua construção na forma de Newton com diferenças divididas, tomando a seguinte ordem dos pontos de interpolação:

$x_2, \dots, x_{n-1}, x_1, x_n$.

$$P_{n-1}(x) = y_2 + [y_2, y_3](x - x_2) + \dots + [y_2, \dots, y_{n-1}, y_1](x - x_2) \dots (x - x_{n-1}) + [y_2, \dots, y_{n-1}, y_1, y_n](x - x_2) \dots (x - x_{n-1})(x - x_1)$$

Considerando agora os pontos de interpolação pela ordem $x_2, \dots, x_{n-1}, x_n, x_1$, tem-se

$$P_{n-1}(x) = y_2 + [y_2, y_3](x - x_2) + \dots + [y_2, \dots, y_{n-1}, y_n](x - x_2) \dots (x - x_{n-1}) + [y_2, \dots, y_{n-1}, y_n, y_1](x - x_2) \dots (x - x_{n-1})(x - x_n)$$

Cancelamento o termo $(x - x_2) \dots (x - x_{n-1})$, comum a todas as parcelas de ambos os lados da igualdade anterior, vem

$$[y_1, y_2, \dots, y_{n-1}] + [y_1, y_2, \dots, y_n](x - x_1) = [y_2, \dots, y_{n-1}, y_n] + [y_1, y_2, \dots, y_n](x - x_n)$$

de onde se obtém

$$[y_1, y_2, \dots, y_n](x - x_1 - x + x_n) = [y_2, \dots, y_{n-1}, y_n] - [y_1, y_2, \dots, y_{n-1}],$$

ou, finalmete

$$[y_1, y_2, \dots, y_n] = \frac{[y_2, \dots, y_{n-1}, y_n] - [y_1, y_2, \dots, y_{n-1}]}{x_n - x_1}$$

Igualando estas duas formas do mesmo polinómio e cancelando os termos iguais, vem

$$[y_2, \dots, y_{n-1}, y_1](x - x_2) \dots (x - x_{n-1}) + [y_2, \dots, y_{n-1}, y_1, y_n](x - x_2) \dots (x - x_{n-1})(x - x_1) = [y_2, \dots, y_{n-1}, y_n](x - x_2) \dots (x - x_{n-1}) + [y_2, \dots, y_{n-1}, y_n, y_1](x - x_2) \dots (x - x_{n-1})(x - x_n)$$

Tendo em conta o que referimos sobre as diferenças divididas serem independentes da ordem dos argumentos, vem

$$[y_1, y_2, \dots, y_{n-1}](x - x_2) \dots (x - x_{n-1}) + [y_1, y_2, \dots, y_n](x - x_1)(x - x_2) \dots (x - x_{n-1}) = [y_2, \dots, y_{n-1}, y_n](x - x_2) \dots (x - x_{n-1}) + [y_1, y_2, \dots, y_n](x - x_2) \dots (x - x_{n-1})(x - x_n)$$

Cálculo recursivo das diferenças divididas

Concluimos assim que as diferenças divididas podem calcular-se pela seguinte fórmula recursiva:

$$[y_1, \dots, y_n] = \frac{[y_2, \dots, y_n] - [y_1, \dots, y_{n-1}]}{x_n - x_1}$$

onde as diferenças de ordem zero $[y_i]$ devem ser entendidas como y_i .

Exemplo ($n = 4$)

x	y	$[\cdot, \cdot]$	$[\cdot, \cdot, \cdot]$	$[\cdot, \cdot, \cdot, \cdot]$
x_1	y_1	$[y_1, y_2]$		
x_2	y_2	$[y_2, y_3]$	$[y_1, y_2, y_3]$	
x_3	y_3	$[y_3, y_4]$	$[y_2, y_3, y_4]$	$[y_1, y_2, y_3, y_4]$
x_4	y_4			

Fórmula do erro com diferenças divididas

Seja P_{n-1} o pol. interp. de y em x_1, \dots, x_n .

Seja $x \neq x_i$ e seja P_n o pol. interp. de y em x_1, \dots, x_n e x . Então

$$P_n(t) = P_{n-1}(t) + y[x_1, \dots, x_n, x](t - x_1) \dots (t - x_n),$$

pelo que

$$P_n(x) = P_{n-1}(x) + y[x_1, \dots, x_n, x](x - x_1) \dots (x - x_n).$$

Mas, $P_n(x) = y(x)$, donde, temos

$$y(x) = P_{n-1}(x) + y[x_1, \dots, x_n, x](x - x_1) \dots (x - x_n)$$

ou seja, temos

$$E_{n-1}(x) = y(x) - P_{n-1}(x) = y[x_1, \dots, x_n, x](x - x_1) \dots (x - x_n).$$

Para calcularmos $y[x_1, \dots, x_n, x]$ precisaríamos de saber o valor de $y(x)$!

Diferenças divididas e derivadas

Teorema

Se $y \in C^{n-1}[a, b]$ e $x_1, \dots, x_n \in [a, b]$ (distintos), então,

$$y[x_1, x_2, \dots, x_n] = \frac{y^{(n-1)}(\xi)}{(n-1)!}$$

para um certo $\xi \in (a, b)$.

Dem: Seja P_{n-2} o pol. interp. de y em x_1, \dots, x_{n-1} . Então, temos

$$y(x_n) - P_{n-2}(x_n) = \frac{y^{(n-1)}(\xi)}{(n-1)!} (x_n - x_1) \dots (x_n - x_{n-1}).$$

Por outro lado, temos também

$$y(x_n) - P_{n-2}(x_n) = y[x_1, x_2, \dots, x_n](x_n - x_1) \dots (x_n - x_{n-1})$$

e o resultado segue-se de imediato.

Nota: Caso $n = 2$ é o Teorema do Valor Médio.

$$E_{n-1}(x) = y(x) - P_{n-1}(x) = y[x_1, \dots, x_n, x](x - x_1) \dots (x - x_n).$$

Se as diferenças divididas de ordem n não variarem muito, podemos tomar $y[x_1, \dots, x_n, x_{n+1}]$ como aproximação para $y[x_1, \dots, x_n, x]$, obtendo, assim, a seguinte **estimativa** para o erro:

$$E_{n-1}(x) \approx y[x_1, \dots, x_n, x_{n+1}](x - x_1) \dots (x - x_n)$$

Estimativa do próximo termo

A fórmula anterior mostra que o erro é dado, aproximadamente, pelo valor do "próximo termo".

Interpolação em pontos equidistantes

Se os pontos de interpolação forem igualmente espaçados, é possível escrever fórmulas mais simples para o polinómio interpolador, com as chamadas **diferenças finitas**.

No que se segue, assumimos que são dados valores y_i correspondentes a uma sequência de pontos igualmente espaçados

$$x_i = x_1 + (i - 1)h; i = 1, 2, \dots$$

Podem definir-se diferenças finitas descendentes, ascendentes ou centradas. Neste curso, vamos apenas falar de diferenças adscendentes e referir formas do polinómio interpolador com esse tipo de diferenças.

Diferenças descendentes

Definição (Diferenças descendentes)

Chama-se **diferença descendente** de ordem k de y_i e denota-se por $\Delta^k y_i$ (ou, por vezes, apenas por Δ_i^k) à quantidade definida por:

- ▶ $\Delta^1 y_i = y_{i+1} - y_i$
- ▶ Para $k = 2, 3, \dots$, $\Delta^k y_i = \Delta(\Delta^{k-1} y_i)$

Nota: Escreve-se simplesmente Δy_i para denotar $\Delta^1 y_i$.

Exemplo

y	Δ	Δ^2	Δ^3
y_1	Δy_1		
y_2	Δy_2	$\Delta^2 y_1$	
y_3	Δy_3	$\Delta^2 y_2$	$\Delta^3 y_1$
y_4			

Diferenças ascendentes

Definição (Diferenças ascendentes)

Chama-se **diferença ascendente** de ordem k de y_i e denota-se por $\nabla^k y_i$ (ou ∇_i^k) à quantidade definida por:

- ▶ $\nabla^1 y_i = y_i - y_{i-1}$
- ▶ Para $k = 2, 3, \dots$, $\nabla^k y_i = \nabla(\nabla^{k-1} y_i)$

Exemplo

y	∇	∇^2	∇^3
y_1			
y_2	∇y_2		
y_3	∇y_3	$\nabla^2 y_3$	
y_4	∇y_4	$\nabla^2 y_4$	$\nabla^3 y_4$

Propriedades das diferenças finitas

- ① $\Delta^k y_i = \nabla^k y_{i+k}$
- ② $\Delta^k y_i = \sum_{j=0}^k (-1)^{k+j} \binom{k}{j} y_{i+j}$
- ③ $\nabla^k y_i = \sum_{j=0}^k (-1)^j \binom{k}{j} y_{i-k+j}$
- ④ $[y_1, y_2, \dots, y_k] = \frac{\Delta^{k-1} y_1}{h^{k-1}(k-1)!}$
- ⑤ $[y_n, \dots, y_{n-k+1}] = \frac{\nabla^{k-1} y_n}{h^{k-1}(k-1)!}$

Dem: A demonstração de todas estas propriedades faz-se por indução sobre k .

Forma de Newton com diferenças descendentes

Relembre:

①

$$P_{n-1}(x) = y_1 + \sum_{k=2}^n [y_1, \dots, y_k] (x - x_1) \dots (x - x_{k-1})$$

②

$$x_i = x_1 + (i-1)h \Rightarrow [y_1, y_2, \dots, y_k] = \frac{\Delta^{k-1} y_1}{h^{k-1}(k-1)!}$$

Então, no caso de nós igualmente espaçados, temos:

$$\begin{aligned} P_{n-1}(x) &= y_1 + \sum_{k=2}^n \frac{\Delta^{k-1} y_1}{h^{k-1}(k-1)!} (x - x_1) \dots (x - x_{k-1}) \\ &= y_1 + \sum_{k=1}^{n-1} \frac{\Delta^k y_1}{h^k k!} (x - x_1) \dots (x - x_k) \end{aligned}$$

Forma de Newton com diferenças descendentes

Efetuem uma mudança de variável definida por

$$s = \frac{x - x_1}{h} \iff x = x_1 + sh$$

Temos

$$x = x_1 + sh \iff x - x_i = x - (x_1 + (i - 1)h) = (s - i + 1)h.$$

Vem, então

$$\begin{aligned} P_{n-1}(x_1 + sh) &= y_1 + \sum_{k=1}^{n-1} \frac{\Delta^k y_1}{h^k k!} s h (s - 1) h \dots (s - k + 1) h \\ &= y_1 + \sum_{k=1}^{n-1} \frac{\Delta^k y_1}{k!} s (s - 1) \dots (s - k + 1) \end{aligned}$$

Polinómio interpolador

Forma de Newton com diferenças ascendentes

De modo totalmente análogo se deduz a forma de Newton do polinómio interpolador com diferenças ascendentes (escrita na forma encaixada), para o caso de nós igualmente espaçados:

$$\begin{aligned} P_{n-1}(x_n + th) &= y_n + t \left\{ \nabla y_n + \frac{(t+1)}{2} \left\{ \nabla^2 y_n + \frac{(t+2)}{3} \left\{ \nabla^3 y_n \right. \right. \right. \\ &\quad \left. \left. \left. + \dots \left\{ \nabla^{n-2} y_n + \frac{(s+n-2)}{n-1} \nabla^{n-1} y_n \right\} \dots \right\} \right\}. \end{aligned}$$

Forma de Newton com diferenças descendentes (forma encaixada)

A fórmula anterior pode escrever-se também na seguinte forma, conhecida por **forma encaixada** do polinómio interpolador com diferenças descendentes

$$\begin{aligned} P_{n-1}(x_1 + sh) &= y_1 + s \left\{ \Delta y_1 + \frac{(s-1)}{2} \left\{ \Delta^2 y_1 + \frac{(s-2)}{3} \left\{ \Delta^3 y_1 \right. \right. \right. \\ &\quad \left. \left. \left. + \dots \left\{ \Delta^{n-2} y_1 + \frac{(s-n+2)}{n-1} \Delta^{n-1} y_1 \right\} \dots \right\} \right\}, \end{aligned}$$

a qual, em relação à fórmula anterior, tem a vantagem de diminuir o número cálculos a efetuar.

Diferenças de valores de um polinómio

Seja $y(x) = p_n(x)$ um polinómio de grau n e sejam dados os valores $y_i = p_n(x_i)$ num conjunto de pontos igualmente espaçados $x_i = x_1 + (i - 1)h, i = 1, 2, \dots$. Atendendo à relação entre diferenças divididas e derivadas e às fórmulas que relacionam diferenças finitas (ascendentes ou descendentes) com diferenças divididas, podemos concluir que diferenças finitas de ordem $n + 1$ desse polinómio serão nulas.

Exemplo

$$y(x) = p_3(x) = x^3 - x + 1$$

x	p_3	Δ	Δ^2	Δ^3	Δ^4
1	1				
2	7	6			
3	25	18	12		
4	61	36	18	6	0
5	121	60	24	6	0
6	211	90	30	6	0
7	337	126	36	6	

Diferenças de valores arredondados de um polinómio

Consideremos agora o caso das diferenças de um polinómio, mas em que os seus valores não são dados exatamente, mas sim **arredondados** para um certo número de casas decimais.

Exemplo

$y(x) = p_2(x) = \frac{(x+1)^2}{15} \rightarrow$ valores arredondados para **4 casas decimais**.

x	\tilde{p}_2	Δ	Δ^2	Δ^3
1	0.2667			
2	0.6000	0.3333	0.1334	
3	1.0667	0.4667	0.1333	-0.0001
4	1.6667	0.6000	0.1333	0.0000
5	2.4000	0.7333	0.1334	0.0001
6	3.2667	0.8667	0.1333	-0.0001
7	4.2667	1.0000		

As diferenças de ordem 3 já não são nulas, mas “oscilam” em torno de zero!

- Se os valores de um polinómio p de grau n são arredondados para d casas decimais (i.e. $|E_{p(x)}| \leq 0.5 \times 10^{-d}$), as diferenças de 1ª ordem terão um erro que, em valor absoluto, não excede

$$2 \times 0.5 \times 10^{-d} = 10^{-d} = 2^0 \times 10^{-d}$$

- as diferenças de 2ª ordem terão um erro não superior $2 \times 10^{-d} = 2^1 \times 10^{-d}$
- as diferenças de 3ª ordem terão um erro não superior $4 \times 10^{-d} = 2^2 \times 10^{-d}$
- ...
- As diferenças de ordem $n + 1$ (que deveriam ser nulas, se os valores do polinómio não tivessem sido arredondados) terão um erro que, em valor absoluto, não excede $2^n \times 10^{-d}$, ou seja: as diferenças de ordem $n + 1$ serão, em módulo, não superiores a $2^n \times 10^{-d}$.

Escolha do grau do polinómio interpolador com base no comportamento de uma tabela de diferenças

Sejam dados valores aproximados de uma função y num conjunto de pontos igualmente espaçados

$$x_i = x_1 + (i - 1)h; i = 1, 2, \dots, N + 1,$$

sendo as aproximações dadas com uma precisão de d casas decimais.

Se, ao formarmos a tabela das diferenças relativas a esses valores, verificarmos que, para um certo inteiro n ($0 \leq n \leq N - 1$) as diferenças de ordem $n + 1$ são, em módulo, não superiores a $2^n \times 10^{-d}$, isto é, se essas diferenças se comportarem como as de um polinómio de grau n com valores arredondado para d casas decimais, fará sentido interpolar y por um polinómio de grau n .

Exemplo

$y(x) = \log_{10}(x) \rightarrow$ arredondamento para 5 c.d.

x	$\tilde{y}(x)$	Δ	Δ^2	Δ^3	Δ^4	Δ^5	Δ^6
2.1	0.32222						
2.2	0.34242	0.02020					
2.3	0.36173	0.01931	-0.00089				
2.4	0.38021	0.01848	-0.00083	0.00006			
2.5	0.39794	0.01773	-0.00075	0.00008	0.00002		
2.6	0.41497	0.01703	-0.00070	0.00005	-0.00003	-0.00005	
2.7	0.43136	0.01639	-0.00064	0.00006	0.00001	0.00004	0.00009

- $\max |E_{\tilde{y}}| \leq 0.5 \times 10^{-5}$
- $\max |\Delta| = 0.02020 > 2^0 \times 10^{-5} = 0.00001$
- $\max |\Delta^2| = 0.00089 > 2^1 \times 10^{-5} = 0.00002$
- $\max |\Delta^3| = 0.00008 > 2^2 \times 10^{-5} = 0.00004$
- $\max |\Delta^4| = 0.00003 < 2^3 \times 10^{-5} = 0.00008$

As diferenças comportam-se como as de um polinómio de grau 3 arredondado para 5 c.d. \Rightarrow Será razoável fazer interpolação por um polinómio cúbico.

Considerações sobre o erro de interpolação

Fórmula do erro em interpolação polinomial

$$y(x) - P_{n-1}(x) = \frac{y^{(n)}(\xi_x)}{n!} (x - x_1) \dots (x - x_n)$$

Se $\max_{x \in [a,b]} |y^{(n)}(x)| \leq M_n$, então

$$|y(x) - P_n(x)| \leq \frac{M_n}{n!} |(x - x_1) \dots (x - x_n)|.$$

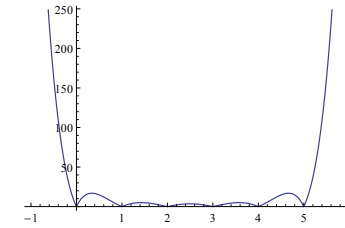


Gráfico de $|x(x-1)(x-2)\dots(x-5)|$

- ▶ O produto $|(x - x_1) \dots (x - x_n)|$ tende a crescer quando x se aproxima dos extremos do intervalo de interpolação e cresce muito rapidamente à medida que x se afasta de $[a, b]$ (Extrapolação). O produto é relativamente pequeno para x situado no “meio” dos nós. Por outro lado, o erro é nulo quando x é algum dos pontos de interpolação, devendo, por continuidade, ser “pequeno” se x estiver próximo de um nó. \Rightarrow Devemos interpolar em nós situados à volta do ponto onde pretendemos o valor de P_{n-1} .

- ▶ Suponhamos que os pontos de interpolação são igualmente espaçados no intervalo $[a, b]$,

$$x_i = a + (i - 1)h; i = 1, \dots, n$$

onde $h = (b - a)/(n - 1)$. É fácil de mostrar que

$$\max |(x - x_1) \dots (x - x_n)| \leq \frac{h^n (n - 1)!}{4}$$

e, portanto, tem-se

$$\max_{x \in [a,b]} |y(x) - P_{n-1}(x)| \leq \frac{h^n M_n}{4n}.$$

Assim, se a função y for infinitamente derivável e tiver derivadas **uniformemente limitadas** em $[a, b]$, ou seja, se existir uma constante K , **independente de n** , tal que, para todo o $x \in [a, b]$, se tem $|y^{(n)}(x)| \leq K$, poderemos concluir que a sequência de polinómios interpoladores (P_n) (construídos aumentando sucessivamente o número de pontos de interpolação) convergirá uniformemente para a função y em $[a, b]$, ou seja, teremos

$$\lim_{n \rightarrow \infty} \max_{x \in [a,b]} |y(x) - P_n(x)| = 0.$$

Este será o caso, por exemplo, da função $y(x) = \sin(x)$ no intervalo $[0, \pi]$.

Fenómeno de Runge

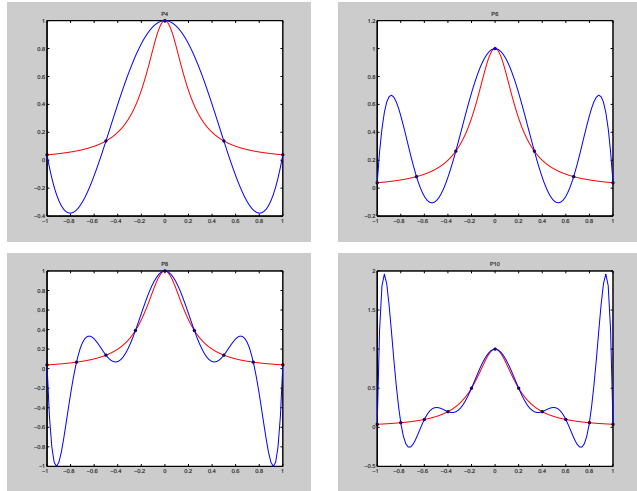
- ▶ No entanto, existem funções para as quais a interpolação num número crescente de pontos não produz uma sequência de polinómios convergindo uniformemente para y . Um exemplo clássico é a chamada **função de Runge**, definida, em $[-1, 1]$, por

$$y(x) = \frac{1}{1 + 25x^2}.$$

Esta função é infinitamente derivável. No entanto, sendo P_n o polinómio de grau não superior a n interpolador de y em $n + 1$ pontos igualmente espaçados no intervalo $[-1, 1]$, pode provar-se que a sequência de polinómios (P_n) não converge uniformemente para y no intervalo $[-1, 1]$, tendo-se

$$\lim_{n \rightarrow \infty} \max_{x \in [-1,1]} |y(x) - P_n(x)| = +\infty.$$

Fenómeno de Runge



Interpolação da função de Runge por P_4, P_6, P_8 e P_{10} (nós igualmente espaçados)

Algumas notações

- ▶ Uma sequência de pontos $\Omega_n = \{x_i\}_{i=1}^n$, onde

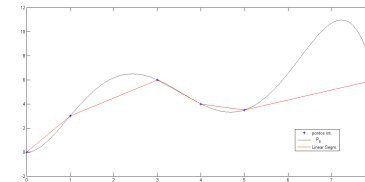
$$a = x_1 < x_2 < \dots < x_{n-1} < x_n = b,$$

é chamada uma **sequência de nós** no intervalo $[a, b]$.

- ▶ Os nós x_2, \dots, x_{n-1} são chamados **nós interiores**.
- ▶ Os nós $x_1 = a$ e $x_n = b$ são chamados **nós fronteiros**.

Interpolação segmentada

- ↓ Vimos que a interpolação num número crescente de nós não conduz necessariamente a melhores resultados (em todo o intervalo de interpolação).
- ↓ Polinómios de grau muito elevado *oscilam* muito...
- ↑ “Partir” o intervalo de interpolação em subintervalos e usar polinómios diferentes nos subintervalos (*polinómios segmentados*). Por exemplo, podemos tomar, em cada subintervalo $[x_{i-1}, x_i]$, o polinómio linear interpolador dos pontos (x_i, y_i) e (x_{i+1}, y_{i+1}) → interpolação linear segmentada .



Interpolação polinomial de grau 5 e interpolação linear segmentada de 6 pontos

Funções *spline*

Definição

Seja dada uma sequência de nós $\Omega_n : a = x_1 < x_2 < \dots < x_n = b$ de um intervalo $[a, b]$ e seja $k \in \mathbb{N}$. Uma função $s : [a, b] \rightarrow \mathbb{R}$ diz-se uma função **spline de grau k com nós Ω_n** , se:

- ▶ Em cada intervalo $[x_i, x_{i+1}]$, s é um polinómio de grau não superior a k ;
- ▶ $s \in C^{(k-1)}[a, b]$.

Nota: Como só definimos *splines* de grau $k \geq 1$, teremos sempre funções **contínuas**.

Por esse motivo, na definição anterior podemos usar os intervalos $[x_i, x_{i+1}]$ fechados em ambos os extremos, não havendo qualquer incompatibilidade na definição de s nos nós interiores.

De entre as funções *spline*, são especialmente importantes as funções *spline* cúbicas e será destas que iremos tratar de seguida com mais pormenor.

Splines cúbicas interpoladoras

Seja dada uma sequência de nós $\Omega_n : a = x_1 < x_2 < \dots < x_n = b$ de um intervalo $[a, b]$ e n valores y_i (em geral, valores de uma dada função y nos nós x_i .)

Problema

Determinar uma função *spline* cúbica s , com nós Ω_n , e *interpoladora* dos valores y_i nos nós x_i . Por outras palavras, pretende-se encontrar uma função s , que satisfaça:

- ▶ Em cada um dos subintervalos $[x_i, x_{i+1}]$; $i = 1, \dots, n - 1$, s é um polinómio de grau não superior a três;
- ▶ $s \in C^2[a, b]$;
- ▶ $s(x_i) = y_i$; $i = 1, 2, \dots, n$.

Graus de liberdade: $4 \times (n - 1) = 4n - 4$

Condições a impor: Continuidade de s, s' e s'' nos nós interiores + condições de interpolação $\rightarrow 3 \times (n - 2) + n = 4n - 6 \implies$ o problema não está totalmente definido, sendo necessário especificar duas condições adicionais.

Integrando duas vezes a expressão anterior, vem

$$s(x) = \frac{(x_{i+1} - x)^3 M_i + (x - x_i)^3 M_{i+1}}{6h} + Ax + B,$$

com A, B constantes. Por uma questão de conveniência, reescrevamos a expressão acima como

$$s(x) = \frac{(x_{i+1} - x)^3 M_i + (x - x_i)^3 M_{i+1}}{6h} + \tilde{A}(x_{i+1} - x) + \tilde{B}(x - x_i). \quad (4)$$

Da condição de interpolação $s(x_i) = y_i$, vem

$$\frac{h^3}{6h} M_i + \tilde{A}h = y_i$$

de onde se obtém a seguinte expressão para \tilde{A} :

$$\tilde{A} = \frac{y_i}{h} - \frac{h}{6} M_i. \quad (5)$$

De modo análogo, usando a condição $s(x_{i+1}) = y_{i+1}$ obtém-se a seguinte expressão para \tilde{B} :

$$\tilde{B} = \frac{y_{i+1}}{h} - \frac{h}{6} M_{i+1}. \quad (6)$$

Caso de nós igualmente espaçados

Por uma questão de simplicidade, consideramos apenas o caso em que os nós são igualmente espaçados, com espaçamento h , i.e., em que temos

$$x_i = a + (i - 1)h \quad i = 1, 2, \dots, n, \quad h = \frac{b - a}{n - 1}.$$

Vamos introduzir a seguinte notação para os valores que a segunda derivada de s assume em cada um dos nós

$$M_i := s''(x_i)$$

Se os valores M_i forem conhecidos, facilmente se obtém a expressão do polinómio cúbico que “forma” a função *spline* s num dado intervalo $[x_i, x_{i+1}]$.

De facto, em $[x_i, x_{i+1}]$, a segunda derivada de s será um polinómio linear, cuja expressão será dada por

$$s''(x) = \frac{(x_{i+1} - x)M_i + (x - x_i)M_{i+1}}{h}.$$

(Porquê?)

Expressão da *spline* em $[x_i, x_{i+1}]$

Substituindo as expressões (5) e (6) em (4), obtemos a seguinte expressão para o polinómio cúbico que define s no intervalo $[x_i, x_{i+1}]$:

$$s(x) = \frac{(x_{i+1} - x)^3 M_i + (x - x_i)^3 M_{i+1}}{6h} + \left(\frac{y_i}{h} - \frac{h}{6} M_i \right) (x_{i+1} - x) + \left(\frac{y_{i+1}}{h} - \frac{h}{6} M_{i+1} \right) (x - x_i) \quad (7)$$

Problema

Como determinar os valores M_i a usar na fórmula acima?

A função s cuja expressão, em cada um dos intervalos $[x_i, x_{i+1}]$, é dada por (7) é tal que $s(x_i^-) = s(x_i^+) = y_i$ e $s''(x_i^-) = s''(x_i^+) = M_i$ em cada nó interior. Para que s seja uma função spline cúbica, deveremos também ter

$$s'(x_i^+) = s'(x_i^-); i = 2, \dots, n-1.$$

Mas:

- ▶ Em $[x_i, x_{i+1}]; i = 1, 2, \dots, n-1$, tem-se

$$s'(x) = \frac{-(x_{i+1} - x)^2 M_i + (x - x_i)^2 M_{i+1}}{2h} + \frac{y_{i+1} - y_i}{h} - \frac{h}{6}(M_{i+1} - M_i).$$

- ▶ Em $[x_{i-1}, x_i]; i = 2, 3, \dots, n-1, n$, tem-se

$$s'(x) = \frac{-(x_i - x)^2 M_{i-1} + (x - x_{i-1})^2 M_i}{2h} + \frac{y_i - y_{i-1}}{h} - \frac{h}{6}(M_i - M_{i-1}).$$

Assim, tem-se

$$s'(x_i^+) = -\frac{h}{2}M_i + \frac{y_{i+1} - y_i}{h} - \frac{h}{6}(M_{i+1} - M_i); i = 1, 2, \dots, n-1$$

e

$$s'(x_i^-) = \frac{h}{2}M_i + \frac{y_i - y_{i-1}}{h} - \frac{h}{6}(M_i - M_{i-1}), i = 2, 3, \dots, n-1, n$$

Condições finais

Spline natural e spline completa

As escolhas mais usuais para as condições finais são as seguintes, tomando as respectivas funções *spline* os nomes indicados:

- ▶ $s''(x_1) = s''(x_n) = 0 \rightarrow$ *spline natural*.
- ▶ $s'(x_1) = d_1$ e $s'(x_n) = d_n$, com d_1 e d_n dados \rightarrow *spline completa*.

Geralmente construímos este tipo de função spline quando estamos a interpolar uma função y nos nós x_i , isto é, quando $y_i = y(x_i)$ e, além disso, conhecemos os valores da primeira derivada de y nos nós fronteiros x_1 e x_n ; nesse caso, devemos tomar $d_1 = y'(x_1)$ e $d_n = y'(x_n)$.

Relações de consistência para os M_i

Igualando as expressões anteriores (para $i = 2, \dots, n-1$) e efetuando alguma manipulação simples, vem que teremos de ter

$$M_{i-1} + 4M_i + M_{i+1} = \frac{6}{h^2}(y_{i-1} - 2y_i + y_{i+1}); i = 2, \dots, n-1.$$

As $n-2$ equações anteriores são conhecidas como *relações de consistência* para os valores $M_i = s''(x_i)$.

Elas não são suficientes para determinar as n incógnitas $M_i; i = 1, \dots, n$ necessárias para a construção de s .

Uma vez mais confirmamos que o problema da determinação de uma função *spline* cúbica interpoladora de certos valores nos nós, i.e, satisfazendo as condições de interpolação $s(x_i) = y_i; i = 1, \dots, n$, não está completamente definido, sendo necessário especificar **duas** condições adicionais. Essas condições são, geralmente, impostas nos nós fronteiros (ou em nós próximos destes) e tomam o nome de *condições finais*.

Condições finais (cont.)

Spline sem-nó

- ▶ $s^{(3)}(x_2^+) = s^{(3)}(x_2^-)$ e $s^{(3)}(x_{n-1}^+) = s^{(3)}(x_{n-1}^-) \rightarrow$ *spline sem-nó*.

A condição de continuidade imposta à terceira derivada de s no nó x_2 (juntamente com a continuidade de s, s' e s'' nesse ponto) implica que as cúbicas que definem a *spline* nos intervalos $[x_1, x_2]$ e $[x_2, x_3]$, sejam **a mesma**, isto é, que não haja um verdadeiro nó em x_2 (o mesmo se passando, naturalmente, com o nó x_{n-1}). Isto justifica a designação escolhida para esta função *spline*.

Naturalmente, teremos de verificar que, para quaisquer das condições finais indicadas, o problema da construção da respetiva função *spline* interpoladora tem uma e uma só solução.

Spline cúbica natural interpoladora

Existência e unicidade

Analisemos, então, o caso da *spline* natural.

As relações de consistência fornecem-nos $n - 2$ equações para a determinação dos valores M_i ; como, neste caso, $M_1 = M_n = 0$, na realidade há apenas $n - 2$ incógnitas a determinar.

O sistema para a determinação dessas incógnitas é dado por

$$\begin{cases} 4M_2 + M_3 = b_2 \\ M_{i-1} + 4M_i + M_{i+1} = b_i; i = 3, \dots, n-2, \\ M_{n-2} + 4M_{n-1} = b_{n-1} \end{cases}$$

onde

$$b_i = \frac{6}{h^2}(y_{i-1} - 2y_i + y_{i+1}).$$

De modo análogo, embora um pouco mais trabalhoso, se podem analisar os casos das funções *spline* sem-nó ou *spline* completa. Haverá, nestes casos, que expressar as condições finais respetivas em termos dos M_i ; por exemplo, as condições finais $s'(x_1) = d_1$ e $s'(x_n) = d_n$ da *spline* completa podem escrever-se como

$$2M_1 + M_2 = \frac{6}{h^2}(y_2 - y_1) - \frac{6}{h}d_1$$

e

$$M_{n-1} + 2M_n = \frac{6}{h}d_n - \frac{6}{h^2}(y_n - y_{n-1}),$$

respetivamente. (Verifique!) Adicionando estas duas equações às $n - 2$ relações de consistência, obtém-se novamente um sistema de matriz de diagonal estritamente dominante e tridiagonal.

A matriz do sistema anterior é

$$\begin{bmatrix} 4 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 4 & 1 & 0 & \cdots & 0 \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & 1 & 4 & 1 \\ 0 & 0 & \cdots & & 1 & 4 \end{bmatrix}.$$

Trata-se de uma matriz de diagonal estritamente dominante, portanto invertível, pelo que o sistema correspondente tem solução única. Podemos, assim, concluir que o problema da determinação de uma função *spline* cúbica natural interpoladora de n valores y_i em n nós x_i tem uma e uma só solução.

Nota: Além disso, a matriz do sistema é tridiagonal, o que torna o sistema especialmente simples de resolver.

Construção de *splines* cúbicas interpoladoras

Passos na construção de uma *spline* cúbica interpoladora:

- 1 Uso das relações de consistência + condições finais (expressas em termos dos M_i) para obter um sistema para determinação dos M_i .
- 2 Resolução desse sistema
- 3 Uso da fórmula (7) para obter a expressão de s em cada intervalo $[x_i, x_{i+1}]$.

Erro em interpolação por *splines* completas

Teorema

Seja s a função spline cúbica completa interpoladora de uma dada função y nos nós $x_i = a + (i - 1)h$, $i = 1 \dots, n$; $h = \frac{b-a}{n-1}$. Suponhamos que $y \in C^4[a, b]$ e seja M_4 tal que $\max_{x \in [a, b]} |y^{(4)}(x)| \leq M_4$. Então, tem-se o seguinte resultado:^a

$$\max_{x \in [a, b]} |s(x) - y(x)| \leq \frac{5}{384} h^4 M_4. \quad (8)$$

^aEstamos aqui a admitir que os valores d_1 e d_n usados nas condições finais são os valores da primeira derivada de y nos extremos, isto é, que $d_1 = y'(x_1)$ e $d_n = y'(x_n)$.

Omitiremos a demonstração, por ser demasiado técnica; ver, e.g. Hall, C.A., Meyer, W., *Optimal Error Bounds for Spline Interpolation*, J. Approx. Theory, **16**, 105-112 (1976).

Na figura seguinte apresenta-se novamente a função de Runge e a função spline cúbica sem-nó interpoladora dessa função em 11 nós igualmente espaçados no intervalo $[-1, 1]$.

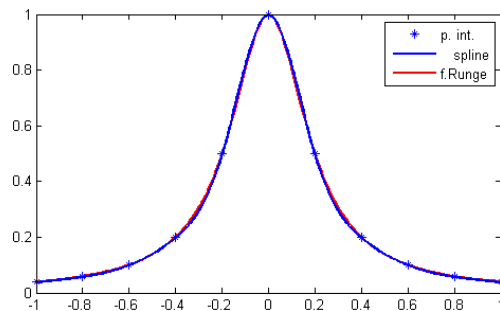


Figura : Função de Runge e função spline cúbica sem-nó interpoladora em 11 pontos igualmente espaçados

O resultado anterior costuma escrever-se, usando o símbolo de Landau \mathcal{O} , da seguinte forma:

$$\|s - y\|_\infty := \max_{x \in [a, b]} |s(x) - y(x)| = \mathcal{O}(h^4),$$

quando $h \rightarrow 0$.

Podemos assim concluir que, se (s_n) for uma sequência de *splines* completas interpoladoras de uma determinada função $y \in C^4[a, b]$ em n nós igualmente espaçados nesse intervalo $[a, b]$, então (s_n) converge uniformemente para y quando $n \rightarrow \infty$ (ou seja, quando o espaçamento h entre os nós tende para zero).

Podemos provar-se também que a função spline sem-nó satisfaz

$$\|s - y\|_\infty = \mathcal{O}(h^4),$$

ou seja que a ordem de convergência deste tipo de *splines* é idêntica à da *spline* completa.

Quanto à função *spline* natural, tem-se o seguinte resultado

$$\|s - y\|_\infty = \mathcal{O}(h^2).$$

Isto significa que, contrariamente ao que o nome **natural** possa sugerir, a interpolação por este tipo de *splines* é, do ponto de vista da aproximação obtida, pouco recomendável.

Nota: No entanto, pode mostrar-se que a influência negativa das condições finais da função *spline* natural diminui à medida que se consideram valores de x mais "interiores" no intervalo de interpolação.

Algumas notas

- ▶ Pode provar-se que, de todas as funções f duas vezes continuamente diferenciáveis num certo intervalo $[a, b]$ e cujo gráfico passa por um certo conjunto de pontos (x_i, y_i) (com x_i abcissas distintas em $[a, b]$), a função *spline* cúbica natural interpoladora s é a (única) que minimiza

$$J(f) = \int_a^b [f''(x)]^2 dx.$$

Esta propriedade é conhecida por **propriedade minimal** da *spline* cúbica natural.

- ▶ Como $f''(x)$ é um indicador da curvatura da curva plana definida por $y = f(x)$, o valor $\int_a^b [f''(x)]^2 dx$ pode ser visto como uma espécie de “curvatura média” da curva no intervalo $[a, b]$.

Podemos, assim, interpretar a propriedade minimal da *spline* natural, do seguinte modo: **De entre todas funções duas vezes continuamente diferenciáveis cujo gráfico passa por determinados pontos, a *spline* interpoladora natural é aquela a que tem menor “curvatura média”.**

- ▶ Um *spline* é um instrumento mecânico usado para traçar curvas suaves passando por determinados pontos (por exemplo, pontos de um mapa, estabelecendo a rota de um navio). Entende-se, assim, a escolha deste nome para designar as funções que temos vindo a estudar.



Figura : Um *spline* mecânico

QUADRATURA

Quadratura Numérica

O problema da **integração numérica** ou **quadratura numérica**¹ é o de estimar o valor de um certo integral $\int_a^b f(x)dx$, o qual supomos existir.²

Existem várias razões que justificam a necessidade do recurso a métodos numéricos para aproximar integrais.

- ▶ Há funções cuja primitiva não é uma função que se possa expressar em termos de funções elementares. Exemplo: $\phi(x) = e^{-x^2}$.
- ▶ Por vezes, embora a primitiva da função a integrar seja conhecida, a sua expressão é de tal modo complicada que não é eficiente o seu uso.
- ▶ Poderemos ter necessidade de integrar uma função da qual conhecemos apenas um conjunto de valores, obtidos, por exemplo, experimentalmente.

¹Quadratura é um termo histórico associado ao cálculo de áreas.

²Para já, consideramos a e b finitos.

Regras básicas

Grande parte dos processos de integração numérica para estimar o valor de um integral

$$I[f] = \int_a^b f(x)dx \quad (9)$$

consistem na utilização de fórmulas simples, chamadas **fórmulas** ou **regras de quadratura**, as quais são, geralmente, do tipo

$$Q_n[f] = \sum_{i=1}^n w_i f(x_i) \quad (10)$$

Assim, o integral é simplesmente substituído por uma combinação linear de valores da função integranda em certos pontos.

- ▶ Os pontos x_i são chamados **abscissas** (ou **pontos de quadratura**) da regra de quadratura.
- ▶ Os coeficientes w_i são chamados **pesos** dessa regra.

Regras de Newton-Cotes (fechadas)

As chamadas regras de quadratura de Newton-Cotes (fechadas) são uma família Q_n ($n \in \mathbb{N}, n \geq 2$) de regras do tipo (10). A regra Q_n tem:

- 1 como abscissas, os n pontos igualmente espaçados no intervalo $[a, b]$,

$$x_i = a + (i - 1)h; \quad i = 1, \dots, n; \quad h = \frac{b - a}{n - 1}. \quad (12)$$

- 2 como pesos, os valores obtidos substituindo a função integranda f pelo polinómio P_{n-1} de grau não superior a $n - 1$ que interpola f nas n abscissas.

Nota: Estas regras são conhecidas como regras de Newton-Cotes **fechadas** porque incluem os extremos do intervalo de integração como abscissas; existem também regras de quadratura de Newton-Cotes **abertas**, em que a e b não são tomados como pontos de quadratura. Se nada dissermos em contrário, quando nos referirmos a regras de quadratura de Newton-Cotes, queremos significar as regras fechadas.

Erro de quadratura

Nota: Em geral, os pesos e as abscissas não dependem da função integranda f e, além disso, os pesos satisfazem

$$\sum_{i=1}^n w_i = b - a.$$

Ao erro resultante da substituição do integral pelo valor dado pela regra, chamamos **erro de quadratura** da regra em questão. Designando por $E_n[f]$ o erro de quadratura associado à regra $Q_n[f]$, tem-se, então

$$\int_a^b f(x)dx = Q_n[f] + E_n[f] \quad (11)$$

Pesos da regra de Newton-Cotes com n abscissas

Relembrando a fórmula de Lagrange do polinómio interpolador P_{n-1} :

$$P_{n-1}(x) = \sum_{i=1}^n L_i(x)f(x_i),$$

onde $L_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^n \left(\frac{x - x_j}{x_i - x_j} \right)$, temos, então, que

$$\begin{aligned} Q_n[f] &= \int_a^b P_{n-1}(x)dx = \int_a^b \left\{ \sum_{i=1}^n L_i(x)f(x_i) \right\} dx \\ &= \sum_{i=1}^n \underbrace{\left\{ \int_a^b L_i(x)dx \right\}}_{w_i} f(x_i). \end{aligned}$$

⇒ Os pesos w_i da regra de Newton-Cotes com n abscissas são dados por

$$w_i = \int_a^b L_i(x)dx; \quad i = 1, \dots, n,$$

onde $L_i(x)$ são os polinómios de Lagrange relativos às abscissas x_1, \dots, x_n .

Erro nas fórmulas de Newton-Cotes

Quanto ao erro de quadratura da regra Q_n , temos

$$E_n[f] = \int_a^b f(x)dx - \int_a^b P_{n-1}(x)dx = \int_{a=x_1}^{b=x_n} (f(x) - P_{n-1}(x)) dx$$

Supondo que $f \in C^n[a, b]$, e tendo em conta a expressão deduzida anteriormente para o erro em interpolação polinomial, vem, então:

$$E_n[f] = \int_{x_1}^{x_n} \frac{f^{(n)}(\xi_x)}{n!} \prod_{i=1}^n (x - x_i) dx. \quad (13)$$

A fórmula anterior mostra que a regra de quadratura de Newton-Cotes Q_n correspondente ao uso de n abcissas é exata (ou seja, tem erro nulo) quando for aplicada a um polinómio de grau não superior a $n - 1$.

Caso $n = 2$ - Regra do trapézio

Consideremos o caso em que $n = 2$, ou seja, o caso em que temos apenas dois pontos de quadratura, $x_1 = a$ e $x_2 = b$, sendo $h = b - a$. Neste caso,

$$L_1(x) = -\frac{x - x_2}{h} \quad \text{e} \quad L_2(x) = \frac{x - x_1}{h}.$$

Assim, os pesos desta regra são dados por

$$w_1 = \int_{x_1}^{x_2} L_1(x)dx = -\frac{1}{h} \int_{x_1}^{x_2} (x - x_2)dx = -\frac{1}{h} \left[\frac{(x - x_2)^2}{2} \right]_{x_1}^{x_2} = \frac{h}{2}$$

e

$$w_2 = \int_{x_1}^{x_2} L_2(x)dx = \frac{1}{h} \int_{x_1}^{x_2} (x - x_1)dx = \frac{1}{h} \left[\frac{(x - x_1)^2}{2} \right]_{x_1}^{x_2} = \frac{h}{2}.$$

Logo, a regra de quadratura de Newton-Cotes para $n = 2$ é dada por

$$Q_2[f] = \frac{h}{2} [f(x_1) + f(x_2)].$$

Grau de exatidão de uma regra de quadratura

Definição

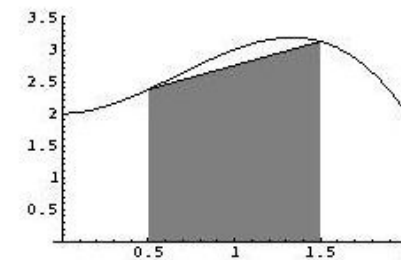
Diz-se que uma regra de quadratura tem **precisão m** ou é de grau (de exatidão) m se for exata para todos os polinómios de grau não superior a m e existir, pelo menos, um polinómio de grau $m + 1$ que não é integrado exatamente por essa fórmula.

Nota: Vemos assim que a precisão da regra de Newton-Cotes com n abcissas é, no mínimo, $n - 1$.

Vamos deduzir agora a expressão das regras de quadratura correspondentes aos primeiros valores de n .

Regra do trapézio

Esta regra é conhecida pelo nome de **regra do trapézio**, por razões que a seguinte figura ilustra claramente.



Regra do trapézio

Erro da regra do trapézio

O erro de quadratura da regra do trapézio é dado por

$$E_2[f] = \frac{1}{2} \int_{x_1}^{x_2} f^{(2)}(\xi_x)(x - x_1)(x - x_2)dx.$$

Tendo em conta que o produto $(x - x_1)(x - x_2)$ não muda de sinal no intervalo (x_1, x_2) , vem, aplicando o *Teorema do valor médio para integrais*,

$$\begin{aligned} E_2[f] &= \frac{1}{2} f^{(2)}(\eta) \int_{x_1}^{x_2} (x - x_1)(x - x_2)dx \\ &= -\frac{h^3}{12} f^{(2)}(\eta), \end{aligned}$$

para um certo $\eta \in (x_1, x_2)$.

Podemos, então, resumir os resultados anteriores no seguinte teorema.

Teorema (Regra do Trapézio e Erro)

Seja f uma função duas vezes continuamente diferenciável num certo intervalo $[x_1, x_2]$ e seja $h = x_2 - x_1$. Então,

$$\int_{x_1}^{x_2} f(x)dx = \frac{h}{2} [f(x_1) + f(x_2)] - \frac{h^3}{12} f^{(2)}(\eta),$$

para um certo $\eta \in (x_1, x_2)$.

Caso $n = 3$ - Regra de Simpson

Neste caso, a função integranda f é aproximada pelo polinómio de grau ≤ 2 que a interpola nos três pontos $a = x_1$, $x_2 = x_1 + h$, $x_3 = x_1 + 2h = b$, onde $h = (b - a)/2$. Temos, então, para pesos desta regra

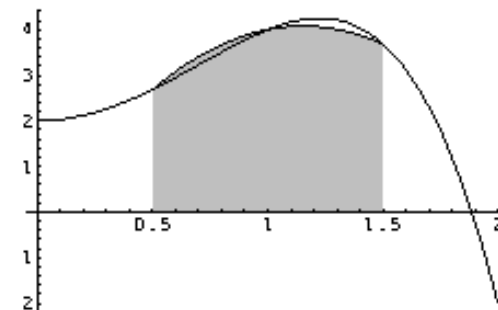
$$\begin{aligned} w_1 &= \int_{x_1}^{x_3} L_1(x)dx = \frac{1}{2h^2} \int_{x_1}^{x_3} (x - x_2)(x - x_3)dx = \frac{h}{3} \\ w_2 &= \int_{x_1}^{x_3} L_2(x)dx = -\frac{1}{h^2} \int_{x_1}^{x_3} (x - x_1)(x - x_3)dx = \frac{4h}{3} \\ w_3 &= \int_{x_1}^{x_3} L_3(x)dx = \frac{1}{2h^2} \int_{x_1}^{x_3} (x - x_1)(x - x_2)dx = \frac{h}{3}. \end{aligned}$$

Assim, tem-se

$$\int_{x_1}^{x_3} f(x)dx \approx Q_3[f] = \frac{h}{3} [f(x_1) + 4f(x_2) + f(x_3)], \quad (14)$$

onde $x_2 = x_1 + h$, $h = (x_3 - x_1)/2$. Esta regra de quadratura é conhecida por **regra de Simpson**.

Regra de Simson



Regra de Simpson

Quanto ao erro de quadratura, tem-se, supondo $f \in C^3[a, b]$,

$$E_3[f] = \frac{1}{3!} \int_{x_1}^{x_3} f^{(3)}(\xi_x)(x - x_1)(x - x_2)(x - x_3)dx.$$

Neste caso, $\Pi_3(x) = (x - x_1)(x - x_2)(x - x_3)$ muda de sinal em (x_1, x_3) , pelo que não podemos aplicar o Teorema do valor médio. Pode, no entanto, provar-se o seguinte resultado (veja, e.g. Valença, M.R., *Análise Numérica*, Univ. Aberta (1996), pp. 211-212): Se $f \in C^4[a, b]$, então

$$E_3(f) = -\frac{h^5}{90}f^{(4)}(\eta).$$

Assim tem-se o seguinte teorema.

Teorema (Regra de Simpson e Erro)

Seja $f \in C^4[a, b]$ e sejam $x_i = a + (i - 1)h, i = 1, 2, 3; h = (b - a)/2$. Então,

$$\int_{a=x_1}^{b=x_3} f(x)dx = \frac{h}{3} [f(x_1) + 4f(x_2) + f(x_3)] - \frac{h^5}{90}f^{(4)}(\eta),$$

para um certo $\eta \in (x_1, x_3)$.

O resultado anterior mostra que a regra de Simpson (caso $n = 3$ das regras de Newton-Cotes fechadas) tem ordem de precisão 3, ou seja tem

Regras de Newton-Cotes fechadas

n	C_n	w_1	w_2	w_3	w_4	w_5	$E_n(f)$
2	$\frac{1}{2}$	1					$-\frac{h^3}{12}f^{(2)}(\eta)$
3	$\frac{1}{3}$	1	4				$-\frac{h^5}{90}f^{(4)}(\eta)$
4	$\frac{3}{8}$	1	3				$-\frac{3h^5}{180}f^{(4)}(\eta)$
5	$\frac{2}{45}$	7	32	12			$-\frac{8h^7}{945}f^{(6)}(\eta)$
6	$\frac{5}{288}$	19	75	50			$-\frac{275h^7}{12096}f^{(6)}(\eta)$
7	$\frac{1}{140}$	41	216	27	272		$-\frac{9h^9}{180}f^{(8)}(\eta)$
8	$\frac{7}{17280}$	751	3577	1323	2989		$-\frac{8183h^9}{518400}f^{(8)}(\eta)$
9	$\frac{4}{14175}$	989	5888	-928	10496	-4540	$-\frac{2368h^{11}}{467775}f^{(10)}(\eta)$

De um modo análogo ao que fizemos para deduzir as regras do trapézio e de Simpson, podem deduzir-se regras de Newton-Cotes fechadas para outros valores de n .

Apresentamos, de seguida, uma tabela com os pesos e a expressão do erro das fórmulas de Newton-Cotes (fechadas) para $n = 2, \dots, 9$.

Para simplificar a tabela, consideramos as fórmulas escritas na forma

$$Q_n[f] = C_n h [w_1 f_1 + \dots + w_n f_n]$$

e indicamos, em cada caso, o valor de C_n e dos coeficientes

$w_i; i = 1, \dots, \lfloor (n + 1)/2 \rfloor$; os restantes coeficientes w_i podem ser obtidos por simetria, isto é, $w_{n+1-i} = w_i$.

- ▶ As regras de Newton-Cotes para $n \geq 9$ têm pesos grandes em valor absoluto, alguns positivos e outros negativos. O uso destas regras é, assim, mais sujeito ao aparecimento de cancelamento subtrativo, pelo que elas são raramente utilizadas.
- ▶ A observação da tabela mostra que as regras com n ímpar têm grau de exatidão n (e não apenas de grau $n - 1$, como seria de esperar). Não se trata de uma coincidência para as regras aí apresentadas, mas de um resultado válido em geral, como indica o teorema seguinte (cuja demonstração omitiremos).

Teorema

Designemos por $E_n(f)$ o erro da fórmula de Newton-Cotes correspondente ao uso de n abcissas $x_i = a + (i - 1)h; i = 1, \dots, n; h = \frac{b-a}{n-1}$. Então:

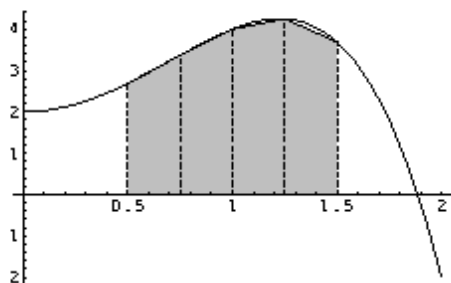
► Se n é par e $f \in C^n[a, b], \exists \eta \in (a, b)$ tal que

$$E_n(f) = \frac{K_n}{n!} f^{(n)}(\eta), \quad \text{com } K_n = \int_a^b \prod_{i=1}^n (x - x_i) dx.$$

► Se n é ímpar e $f \in C^{n+1}[a, b], \exists \eta \in (a, b)$ tal que

$$E_n(f) = \frac{K_n}{(n+1)!} f^{(n+1)}(\eta), \quad \text{com } K_n = \int_a^b x \prod_{i=1}^n (x - x_i) dx.$$

Nota: Por uma questão de simplicidade, usamos a notação $f_i := f(x_i)$. A regra anterior é conhecida por **regra do trapézio composta** (com N intervalos) e designada por $T_N[f]$.



Regra do trapézio composta

Regras compostas

A ideia das regras **compostas** é subdividir o intervalo de integração $[a, b]$ num certo número de subintervalos e aplicar regras simples em subintervalos.

Regra do trapézio composta

Seja $N \in \mathbb{N}$ e sejam $x_i = a + (i - 1)h; i = 1, \dots, N + 1$, com $h = (b - a)/N$. Temos, então

$$\int_{a=x_1}^{b=x_{N+1}} f(x) dx = \sum_{i=1}^N \int_{x_i}^{x_{i+1}} f(x) dx.$$

Usando a regra do trapézio em cada subintervalo $[x_i, x_{i+1}]$, vem

$$\begin{aligned} \int_a^b f(x) dx &\approx \sum_{i=1}^N \left[\frac{h}{2} (f_i + f_{i+1}) \right] \\ &= \frac{h}{2} [(f_1 + f_2) + (f_2 + f_3) + \dots + (f_{N-1} + f_N) + (f_N + f_{N+1})] \\ &= \frac{h}{2} [f_1 + 2(f_2 + \dots + f_N) + f_{N+1}] \end{aligned}$$

Erro da regra do trapézio composta

Quanto ao erro desta regra, tem-se, supondo que $f \in C^2[a, b]$, e recordando a expressão do erro da regra do trapézio,

$$\begin{aligned} E_{T_N}[f] &:= \int_a^b f(x) dx - T_N[f] \\ &= \sum_{i=1}^N \left(-\frac{h^3}{12} f^{(2)}(\eta_i) \right), \end{aligned}$$

onde $\eta_i \in (x_i, x_{i+1})$. Aplicando o Teorema do valor médio para somas, vem

$$\begin{aligned} E_{T_N}[f] &= -\frac{h^3}{12} f^{(2)}(\eta) \sum_{i=1}^N 1 = -\frac{h^3}{12} N f^{(2)}(\eta) \\ &= -\frac{(b-a)h^2}{12} f^{(2)}(\eta), \quad \eta \in (a, b), \end{aligned}$$

atendendo a que $N = \frac{(b-a)}{h}$.

Regra do trapézio composta e erro

Temos, assim, o seguinte teorema.

Teorema (Regra do trapézio composta e erro)

Seja $f \in C^2[a, b]$ e sejam $x_i = a + (i - 1)h, i = 1, \dots, N + 1$, com $h = \frac{b-a}{N}, N \in \mathbb{N}$. Então, tem-se

$$\int_a^b f(x)dx = \frac{h}{2} [f_1 + 2(f_2 + \dots + f_N) + f_{N+1}] - \frac{(b-a)h^2}{12} f^{(2)}(\eta),$$

para um certo $\eta \in (a, b)$.

Nota

O caso $N = 1$ corresponde à regra do trapézio (simples).

A regra anterior é conhecida por regra de Simpson composta com $N = 2m$ subintervalos e denotada por S_{2m} .

Nota

O caso $m = 1$ (i.e. $N = 2$) corresponde à regra de Simpson (simples).

Quanto ao erro desta regra, vem

$$\begin{aligned} E_{S_{2m}}[f] &:= \int_a^b f(x)dx - S_{2m}[f] \\ &= -\frac{h^5}{90} \sum_{i=1}^m f^{(4)}(\eta_i) \\ &= -\frac{h^5}{90} m f^{(4)}(\eta) \\ &= -\frac{(b-a)h^4}{180} f^{(4)}(\eta), \quad \eta \in (a, b), \end{aligned}$$

onde se usou, uma vez mais, o Teorema do valor médio para somas e se teve em conta que $m = \frac{b-a}{2h}$.

Regra de Simpson composta

Consideremos agora o intervalo $[a, b]$ partido num número par $N = 2m$ de subintervalos, isto é, sejam $x_i = a + (i - 1)h; i = 1, \dots, 2m + 1$, com $h = \frac{b-a}{2m}$. Temos, então

$$\begin{aligned} \int_{a=x_1}^{b=x_{2m+1}} f(x)dx &= \int_{x_1}^{x_3} f(x)dx + \int_{x_3}^{x_5} f(x)dx + \dots + \int_{x_{2m-1}}^{x_{2m+1}} f(x)dx \\ &= \sum_{i=1}^m \int_{x_{2i-1}}^{x_{2i+1}} f(x)dx. \end{aligned}$$

Usando a regra de Simpson em cada duplo subintervalo $[x_{2i-1}, x_{2i+1}]$, vem

$$\begin{aligned} \int_a^b f(x)dx &\approx \sum_{i=1}^m \left[\frac{h}{3} (f_{2i-1} + 4f_{2i} + f_{2i+1}) \right] \\ &= \frac{h}{3} [f_1 + 4(f_2 + \dots + f_{2m}) + 2(f_3 + \dots + f_{2m-1}) + f_{2m+1}] \end{aligned}$$

Regra de Simpson composta e erro

Temos, assim, o seguinte teorema.

Teorema (Regra de Simpson composta e erro)

Seja $f \in C^4[a, b]$ e sejam $x_i = a + (i - 1)h; i = 1, \dots, 2m + 1$, com $h = (b - a)/2m$. Então,

$$\begin{aligned} \int_a^b f(x)dx &= \frac{h}{3} [f_1 + 4(f_2 + \dots + f_{2m}) + \\ &\quad + 2(f_3 + \dots + f_{2m-1}) + f_{2m+1}] \\ &\quad - \frac{(b-a)h^4}{180} f^{(4)}(\eta) \end{aligned}$$

para um certo $\eta \in (a, b)$.

Quadratura Gaussiana

As regras de quadratura de Newton-Cotes são da forma

$$I[f] = \int_a^b f(x)dx \approx \sum_{i=1}^n w_i f(x_i) \quad (15)$$

onde as abscissas x_i são igualmente espaçadas em $[a, b]$ (e onde os pesos w_i são dados por $\int_a^b L_i(x)dx$). Tais fórmulas, como vimos, têm grau de exatidão $n - 1$, se n é ímpar e grau de exatidão n , se n é par.

É, no entanto, possível obter regras do tipo (15) que sejam exatas para todos os polinómios de grau (não superior a) $2n - 1$, escolhendo as abscissas e os pesos **apropriadamente**. Esta é a ideia básica das regras de quadratura Gaussiana.

Nota: É muito simples de provar que uma regra do tipo (15) não pode ter grau de precisão superior a $2n - 1$, pois existe sempre um polinómio de grau $2n$ que não pode ser integrado exatamente por uma tal regra. (Qual?)

Exemplo

No que se segue, vamos considerar que $[a, b] = [-1, 1]$.

Note-se que temos sempre

$$I = \int_a^b f(x)dx = \frac{(b-a)}{2} \int_{-1}^1 f\left(\frac{a+b}{2} + \frac{b-a}{2}u\right) du,$$

pelo que não há perda de generalidade na hipótese acima considerada.

Vejamos, então, se será possível encontrar uma regra de quadratura do tipo

$$I[f] = \int_{-1}^1 f(x)dx \approx w_1 f(x_1) + w_2 f(x_2) \quad (16)$$

que seja exata para todos os polinómios de grau não superior a 3.

Como a integração é um processo linear, a fórmula (16) será exata para polinómios de grau ≤ 3 se e só se integrar exactamente os polinómios $p_1(x) = 1, p_2(x) = x, p_3(x) = x^2$ e $p_3(x) = x^3$.

Exemplo (cont.)

Temos, então, que a regra será exata para todos os polinómios de grau não superior a 3 se e só se tivermos

$$\begin{cases} \int_{-1}^1 1dx = w_1 \times 1 + w_2 \times 1 \\ \int_{-1}^1 xdx = w_1 x_1 + w_2 x_2 \\ \int_{-1}^1 x^2 dx = w_1 x_1^2 + w_2 x_2^2 \\ \int_{-1}^1 x^3 dx = w_1 x_1^3 + w_2 x_2^3 \end{cases} \iff \begin{cases} w_1 + w_2 = 2 \\ w_1 x_1 + w_2 x_2 = 0 \\ w_1 x_1^2 + w_2 x_2^2 = \frac{2}{3} \\ w_1 x_1^3 + w_2 x_2^3 = 0 \end{cases}$$

Exemplo (cont.)

Facilmente se verifica que $w_1 = w_2 = 1$ e $x_1 = -\frac{\sqrt{3}}{3}, x_2 = \frac{\sqrt{3}}{3}$ é solução do sistema não linear anterior.

Temos, então, que a seguinte regra de quadratura

$$\int_{-1}^1 f(x)dx \approx f\left(-\frac{\sqrt{3}}{3}\right) + f\left(\frac{\sqrt{3}}{3}\right) \quad (17)$$

é exata para polinómios de grau não superior a 3; como ela não pode ter grau de precisão superior a 3, concluímos que tem grau precisão 3.

O processo descrito para obter a regra anterior pode generalizar-se: para obter um regra de quadratura da forma

$$I = \int_{-1}^1 f(x)dx \approx w_1 f(x_1) + \dots + w_n f(x_n)$$

que seja exata para polinómios de grau $\leq 2n - 1$, deveremos resolver as seguintes $2n$ equações não lineares

$$\sum_{j=1}^n w_j x_j^k = \alpha_k; k = 0, \dots, 2n - 1,$$

onde

$$\alpha_k = \int_{-1}^1 x^k dx = \begin{cases} \frac{2}{k+1}, & \text{se } k \text{ par} \\ 0, & \text{se } k \text{ ímpar} \end{cases}$$

Pode provar-se que tais equações têm solução (real), pelo que, se encontrarmos a sua solução, obteremos os pesos w_i e as abcissas x_i que procuramos.

Intermezzo....

Polinómios Ortogonais Reais

Referência: P.J. Davis, *Interpolation and Approximation*, Dover Publ., New York, 1975.

Espaços com produto interno

Seja E um espaço vectorial real e considere-se uma aplicação $\langle \cdot | \cdot \rangle : E \times E \rightarrow \mathbb{R}$ que satisfaça as seguintes propriedades:

$$\text{PI1 } \forall x \in E, \quad \langle x | x \rangle \geq 0 \quad \text{e} \quad \langle x | x \rangle = 0 \iff x = 0.$$

$$\text{PI2 } \forall x, y \in E, \quad \langle x | y \rangle = \langle y | x \rangle.$$

$$\text{PI3 } \forall x, y, z \in E, \forall \alpha, \beta \in \mathbb{R}, \langle \alpha x + \beta y | z \rangle = \alpha \langle x | z \rangle + \beta \langle y | z \rangle.$$

Então, $\langle \cdot | \cdot \rangle$ diz-se um produto interno em E (e E é dito um espaço com produto interno).

Nota: Pode dar-se uma definição análoga para e.v. complexos; nesse caso, a condição PI2 deve ser substituída por $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

Esta abordagem **algébrica** de obter estas regras de quadratura não é, no entanto, muito adequada.

Na realidade, é possível deduzi-las por um processo diferente (analítico), o qual tem ainda a vantagem de determinar uma fórmula para o erro de quadratura associado a cada uma das regras.

Para fazer essa dedução, necessitaríamos, no entanto, de resultados sobre interpolação (nomeadamente a chamada interpolação de Hermite) que não fazem parte do programa deste curso, pelo que esta não será apresentada aqui.

Ainda assim, será conveniente, a este propósito, efetuar um estudo breve sobre os chamados **polinómios ortogonais clássicos**, o qual apresentamos de seguida.

Exemplos

E1. $E = \mathbb{R}^n$ e

$$\langle x | y \rangle = \sum_{i=1}^n \omega_i x_i y_i, \quad \omega_i > 0.$$

E2. $E = C[a, b]$ (a, b finitos) e

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

onde $\omega(x)$ é uma função positiva e integrável em (a, b) .

E3. $E = \mathcal{P}$, onde \mathcal{P} designa o espaço de todos os polinómios de coeficientes reais e

$$\langle p | q \rangle = \int_a^b \omega(x) p(x) q(x) dx \tag{18}$$

(a e b não necessariamente finitos), onde $\omega(x)$ designa uma função positiva definida em (a, b) e tal que $\int_a^b \omega(x) |x|^n dx$ existe e é finito para todo o $n \geq 0$.

Ortogonalidade

Como casos particulares do produto interno considerado no Exemplo **E3** salientam-se, pela sua importância, os seguintes:

E3(i) $a = -1, b = 1, w(x) = 1.$

E3(ii) $a = -1, b = 1, w(x) = \frac{1}{\sqrt{1-x^2}}.$

E3(iii) $a = 0, b = +\infty, w(x) = e^{-x}.$

E3(iv) $a = -\infty, b = +\infty, w(x) = e^{-x^2}.$

Definição

Seja E um espaço com produto interno $\langle \cdot | \cdot \rangle$.

- ▶ Dois elementos $x, y \in E$ dizem-se **ortogonais** se $\langle x | y \rangle = 0$.
- ▶ Um subconjunto $S \subseteq E$ diz-se **ortogonal** se $\langle x | y \rangle = 0$, para quaisquer $x, y \in S, x \neq y$.

Facilmente se verifica que a aplicação $\| \cdot \|$ definida em E por

$$\|x\| = \langle x | x \rangle^{1/2}$$

é uma norma, a qual se diz induzida pelo produto interno.

Se $S \subseteq E$ é ortogonal e, além disso, se tiver $\|x\| = 1$, para todo o $x \in S$, dizemos que S é **ortonormado** (o.n.).

Facilmente se prova o seguinte resultado.

Teorema

Se $S \subseteq E$ é um conjunto ortogonal de elementos não nulos, então S é formado por vectores linearmente independentes (l.i.).

O teorema anterior admite o seguinte “recíproco parcial”.

Teorema (Processo de ortogonalização de Gram-Schmidt)

Seja x_1, x_2, x_3, \dots uma sequência de vectores linearmente independentes de um espaço com produto interno $\langle \cdot | \cdot \rangle$. Então, a seguinte sequência de vectores

$$x_1^* = x_1, x_k^* = x_k - \sum_{j=1}^{k-1} \frac{\langle x_k | x_j^* \rangle}{\langle x_j^* | x_j^* \rangle} x_j^*; k = 2, 3, \dots$$

tem as seguintes propriedades:

- ▶ é formada por vectores ortogonais (dois a dois);
- ▶ para cada n , o espaço gerado pelos vectores x_1^*, \dots, x_n^* coincide com o espaço gerado por x_1, \dots, x_n .

Sequência de polinómios ortogonais

A partir de agora, vamos considerar o caso em que $E = \mathcal{P}$ (espaço vectorial real de todos os polinómios de coeficientes reais) e onde se supõe definido um produto interno da forma descrita no Exemplo **E3**.

Definição

Uma sequência $(p_k)_{k=0}^{\infty}$ de polinómios diz-se uma **sequência de polinómios ortogonais** (relativa a um produto interno $\langle \cdot | \cdot \rangle$), se:

- Cada polinómio p_k tem grau exactamente igual a k ;
- Polinómios distintos são ortogonais, i.e., $\langle p_k | p_j \rangle = 0, j \neq k$.

Nota: Se (p_k) for uma sequência de polinómios ortogonais relativa a um certo produto interno, então é imediato reconhecer que qualquer sequência da forma $(\alpha_k p_k)$, com $\alpha_k \neq 0$, é também uma sequência de polinómios ortogonais relativa ao mesmo produto interno (verifique). É, por vezes, conveniente trabalhar com sequências de polinómios ortogonais “normalizados” de uma certa forma; por exemplo, poderá interessar que os polinómios p_k sejam mónicos (i.e. que o coeficiente do termo de maior grau de cada polinómio seja igual a 1) ou que a soma dos coeficientes de cada polinómio seja igual a 1 (i.e. que os polinómios

Questões

- 1 Fixado em \mathcal{P} um produto interno, existirá sempre uma sequência de polinômios ortogonais relativa a esse produto interno?
- 2 Será que, a haver uma sequência, essa sequência é única (a menos, naturalmente, do produto de cada polinômio por uma constante arbitrária não nula)?

A resposta à primeira questão é afirmativa: basta notar que a sequência de polinômios $(p_k)_{k=0}^{\infty}$ onde $p_k(x) = x^k, k = 0, 1, 2, \dots$ é formada por polinômios linearmente independentes e que, se lhe aplicarmos o processo de Gram-Schmidt, obtemos uma sequência (p_k^*) de polinômios, ortogonais dois a dois, e onde cada polinômio p_k^* tem grau exatamente igual a k , já que o coeficiente de x^k em p_k^* é igual a 1 (ou seja esta sequência é formada por polinômios mônicos).

A resposta à segunda questão é também afirmativa, como demonstraremos nas aulas práticas; mais precisamente, demonstraremos que não existem duas sequências distintas de polinômios ortogonais mônicos relativas a um dado produto interno.

A demonstração de algumas destas propriedades será feita nas aulas práticas.

Os casos especiais de polinômios ortogonais relativos a cada um dos produtos internos **E3(i) – E3(iv)** referidos anteriormente foram estudados intensivamente, constituindo os chamados **polinômios ortogonais clássicos**. Damos, de seguida, um resumo dos resultados mais importantes relativos a esses polinômios.

Propriedades dos polinômios ortogonais

Seja $(p_k)_{k=0}^{\infty}$ uma sequência de polinômios ortogonais, relativa a um produto interno do tipo $\langle p, q \rangle = \int_a^b \omega(x)p(x)q(x)dx$. Então, são válidas as seguintes propriedades:

PO1 Os primeiros $n + 1$ polinômios, p_0, p_1, \dots, p_n , formam uma base do espaço \mathcal{P}_n dos polinômios de grau não superior a n .

PO2 Dado $n \in \mathbb{N}$, o polinômio p_n é ortogonal a qualquer polinômio de grau inferior a n , isto é,

$$\langle p_n | q \rangle = 0, \forall q \in \mathcal{P}_{n-1}.$$

PO3 Os polinômios ortogonais p_k satisfazem uma relação de recorrência a três termos do tipo:

$$p_{k+1}(x) = A_k(x - B_k)p_k(x) - C_k p_{k-1}(x); \quad k = 2, 3, \dots,$$

com p_0 e p_1 dados.

PO4 Os zeros de p_k ($k \geq 1$) são reais, simples e estão no intervalo (a, b) .

Polinômios de Legendre

$$a = -1, \quad b = 1, \quad \omega(x) = 1$$

Símbolo: P_n

Normalização: $P_n(1) = 1$

Norma: $\|P_n\|^2 = \int_{-1}^1 P_n^2(x)dx = \frac{2}{2n+1}$

Relação de recorrência: $(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$

Expressão explícita: $P_n(x) = \frac{1}{2^n} \sum_{m=0}^{\lfloor n/2 \rfloor} (-1)^m \binom{n}{m} \binom{2n-2m}{n} x^{n-2m}$

Fórmula de Rodrigues: $P_n(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} \{(1-x^2)^n\}$

Desigualdade: $|P_n(x)| \leq 1, \quad x \in [-1, 1]$.

Primeiros polinômios de Legendre

n	P_n
0	1
1	x
2	$\frac{1}{2}(3x^2 - 1)$
3	$\frac{1}{2}(5x^3 - 3x)$
4	$\frac{1}{8}(35x^4 - 30x^2 + 3)$
5	$\frac{1}{8}(63x^5 - 70x^3 + 15x)$
6	$\frac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5)$

Primeiros polinômios de Chebyshev

n	T_n
0	1
1	x
2	$2x^2 - 1$
3	$4x^3 - 3x$
4	$8x^4 - 8x^2 + 1$
5	$16x^5 - 20x^3 + 5x$
6	$32x^6 - 48x^4 + 18x^2 - 1$

Polinômios de Chebyshev

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

Símbolo: T_n

Normalização: $T_n(1) = 1$

Norma:

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

Relação de recorrência: $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$

Expressão explícita: $T_n(x) = \cos(n \arccos x)$

Fórmula de Rodrigues: $T_n(x) = \frac{(-1)^n (1-x^2)^{1/2} \sqrt{\pi}}{2^{n+1} \Gamma(n+\frac{1}{2})} \frac{d^n}{dx^n} \{(1-x^2)^{n-\frac{1}{2}}\}$

Desigualdade: $|T_n(x)| \leq 1, \quad x \in [-1, 1].$

Polinômios de Laguerre

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

Símbolo: L_n

Normalização: $L_n(x) = \frac{(-1)^n}{n!} x^n + \dots$

Norma: $\|L_n\|^2 = \int_0^\infty e^{-x} L_n^2(x) dx = 1$

Relação de recorrência: $(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x)$

Expressão explícita: $L_n(x) = \sum_{m=0}^n (-1)^m \binom{n}{n-m} \frac{1}{m!} x^m$

Fórmula de Rodrigues: $L_n(x) = \frac{1}{n! x e^{-x}} \frac{d^n}{dx^n} \{x^n e^{-x}\}$

Primeiros polinómios de Laguerre

n	L_n
0	1
1	$-(x - 1)$
2	$\frac{1}{2}(x^2 - 4x + 2)$
3	$-\frac{1}{6}(x^3 - 9x^2 + 18x - 6)$
4	$\frac{1}{24}(x^4 - 16x^3 + 72x^2 - 86x + 24)$
5	$-\frac{1}{120}(x^5 - 25x^4 + 200x^3 - 600x^2 + 600x - 120)$
6	$\frac{1}{720}(x^6 - 36x^5 + 450x^4 - 2400x^3 + 5400x^2 - 4320x + 720)$

Polinómios de Hermite

$$a = -\infty, \quad b = \infty, \quad \omega(x) = e^{-x^2}$$

Símbolo: H_n

Normalização: $H_n(x) = 2^n x^n + \dots$

Norma: $\|H_n\|^2 = \int_{-\infty}^{\infty} e^{-x^2} H_n^2(x) dx = \sqrt{\pi} 2^n n!$

Relação de recorrência: $H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$

Expressão explícita: $H_n(x) = n! \sum_{m=0}^{\lfloor n/2 \rfloor} (-1)^m \frac{(2x)^{n-2m}}{m!(n-2m)!}$

Fórmula de Rodrigues: $H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} \{e^{-x^2}\}$

Primeiros polinómios de Hermite

n	H_n
0	1
1	$2x$
2	$4x^2 - 2$
3	$8x^3 - 12x$
4	$16x^4 - 48x^2 + 12$
5	$32x^5 - 160x^3 + 120$
6	$64x^6 - 480x^4 + 720x^2 - 120$

De volta às regras de quadratura...

Se consideramos o polinómio de Legendre de grau 2

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

vemos que os seus dois zeros são

$$x_1 = -\frac{\sqrt{3}}{3} \quad \text{e} \quad x_2 = \frac{\sqrt{3}}{3},$$

ou seja, são precisamente as abscissas da regra de quadratura (17) anteriormente deduzida (a que, com apenas dois pontos de quadratura, é exacta para polinómios de grau ≤ 3).

Não se trata de uma mera coincidência, mas de um resultado particular do seguinte teorema.

Regras de Gauss-Legendre

Teorema

Uma regra de quadratura da forma

$$\int_{-1}^1 f(x) dx \approx \sum_{i=1}^n w_i f(x_i) \quad (19)$$

é exata para todos os polinómios em \mathcal{P}_{2n-1} , se e só se:

- 1 as abcissas x_i forem os n zeros do polinómio ortogonal de Legendre de grau n ,^a
- 2 os pesos forem dados por $w_i = \int_{-1}^1 L_i(x) dx$ onde L_i são os polinómios de Lagrange (relativos às abcissas consideradas).^b

^aNote-se que estes zeros são, de acordo com a propriedade PO5 anteriormente referida dos polinómios ortogonais, números reais, distintos e pertencentes ao intervalo $(-1, 1)$.

^bPode provar-se que $w_i = \int_{-1}^1 L_i(x) dx = \int_{-1}^1 L_i^2(x) dx$, o que garante que os pesos são positivos.

de Gauss-Legendre (com n pontos) e denotada por Q_n^{GL} .

Demonstração (cont.)

\Leftarrow) Suponhamos que (1) e (2) se verificam e vejamos que a regra (19) com essas abcissas e pesos é exata para todo o polinómio em \mathcal{P}_{2n-1} . Seja então p um polinómio qualquer em \mathcal{P}_{2n-1} . Esse polinómio pode escrever-se como

$$p(x) = q(x)P_n(x) + r(x),$$

onde P_n é o polinómio de Legendre de grau n e $q, r \in \mathcal{P}_{n-1}$. Como $P_n(x_i) = 0; i = 1, \dots, n$, segue-se que $p(x_i) = r(x_i); i = 1, \dots, n$. Usando a fórmula de Lagrange do polinómio interpolador (e tendo em conta que r é um polinómio em \mathcal{P}_{n-1} , pelo que coincide com o polinómio que o interpola em n abcissas), temos

$$r(x) = \sum_{i=1}^n r(x_i)L_i(x) = \sum_{i=1}^n p(x_i)L_i(x).$$

Temos, então

$$\begin{aligned} \int_{-1}^1 p(x) dx &= \int_{-1}^1 q(x)P_n(x) dx + \int_{-1}^1 r(x) dx = 0 + \int_{-1}^1 \sum_{i=1}^n L_i(x)p(x_i) dx \\ &= p(x_i) \sum_{i=1}^n \int_{-1}^1 L_i(x) dx = \sum_{i=1}^n w_i p(x_i). \end{aligned}$$

Demonstração

\Rightarrow) Suponhamos que a regra é exata para todos os polinómios em \mathcal{P}_{2n-1} e mostremos que x_1, \dots, x_n são os zeros do polinómio de Legendre de grau n . Seja A_n o coeficiente de x^n do polinómio de Legendre de grau n , P_n . Consideremos o polinómio $q_n(x) = A_n(x - x_1) \dots (x - x_n)$; o que pretendemos mostrar é que q_n coincide com P_n . Aplicando a fórmula (19) aos polinómios $q_n(x)x^k; k = 0, 1, \dots, n - 1$, tem-se, atendendo a que esses polinómios estão em \mathcal{P}_{2n-1} , pelo que a regra será exata para esses polinómios:

$$\int_{-1}^1 q_n(x)x^k dx = \sum_{i=1}^n w_i q_n(x_i)x_i^k = 0.$$

Isto mostra que q_n é ortogonal a cada um dos polinómios $1, x, \dots, x^{n-1}$, ou seja, que é ortogonal a qualquer polinómio em \mathcal{P}_{n-1} . Como q_n e P_n são ambos polinómios de grau n com o mesmo coeficiente de x^n , tem-se que $q_n - P_n \in \mathcal{P}_{n-1}$. Assim sendo, temos

$$\langle q_n - P_n, q_n - P_n \rangle = \langle q_n - P_n, q_n \rangle - \langle q_n - P_n, P_n \rangle = 0.$$

Isto significa que $q_n - P_n \equiv 0$ ou seja, que $q_n = P_n$.

Para demonstrar 2 basta aplicar a fórmula cada um dos polinómios $L_i(x)$ e ter em atenção que $L_i(x_k) = \delta_{ik}$.

Erro da fórmula de Gauss-Legendre

Pode mostrar-se que, se $f \in C^{2n}[-1, 1]$, é válida a seguinte fórmula para o erro de quadratura relativo à regra de Gauss-Legendre com n pontos (isto é, à regra referida no teorema anterior):

$$E_n^{GL}[f] = \int_{-1}^1 f(x) dx - Q_n^{GL}[f] = C_n \frac{f^{(2n)}(\eta)}{(2n)!}, \eta \in (-1, 1),$$

onde

$$C_n = \frac{2^{2n+1}(n!)^4}{(2n+1)((2n)!)^2}.$$

Os pesos e abcissas das fórmulas de Gauss-Legendre para n com interesse prático encontram-se tabelados; veja, e.g. Stroud, A.H. e Secrest, D., *Gaussian Quadrature Formulae*, Prentice-Hall (1966). Segue-se uma tabela para alguns valores de n .

Pesos e abscissas das fórmulas de Gauss-Legendre

n	x_i	w_i
2	± 0.5773502691896258	1
3	0 ± 0.7745966692414834	0.8888888888888889 0.5555555555555556
4	± 0.3399810435848563 ± 0.8611363115940526	0.6521451548625461 0.3478548451374538
5	0 ± 0.5384693101056831 ± 0.9061798459386640	0.5688888888888889 0.4786286704993665 0.2386191860831969
6	± 0.2386191860831969 ± 0.6612093864662645 ± 0.9324695142031521	0.4679139345726910 0.3607615730481386 0.1713244923791704

Outras fórmulas de quadratura Gaussiana

As fórmulas de Gauss-Legendre são, como vimos, do tipo

$$\int_{-1}^1 f(x)dx \approx \sum_{i=1}^n w_i f(x_i)$$

sendo portanto aplicáveis ao cálculo de integrais que estejam na forma $\int_{-1}^1 f(x)dx$. É também possível obter fórmulas de quadratura do tipo

$$\int_a^b \omega(x)f(x)dx \approx \sum_{i=1}^n w_i f(x_i),$$

para outros valores de a e b e para outras funções peso $\omega(x)$, exatas para polinómios de grau $\leq 2n - 1$, escolhendo para abscissas os zeros do polinómio ortogonal de grau n correspondente ao produto interno associado aos extremos a e b e à função peso ω .

Por exemplo, uma fórmula de quadratura do tipo

$$\int_0^{\infty} e^{-x} f(x)dx \approx \sum_{i=1}^n w_i f(x_i)$$

com x_i os n zeros do polinómio de Laguerre de grau n e cujos pesos são dados por

$$w_i = \int_0^{\infty} e^{-x} L_i(x)dx$$

é exata para todos os polinómios de grau $\leq 2n - 1$.³ Esta fórmula é conhecida por [fórmula de Gauss-Laguerre com \$n\$ pontos](#).

De modo análogo, têm-se regras de quadratura de [Gauss-Chebyshev](#) para estimar integrais do tipo

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} f(x)dx$$

e [regras de Gauss-Hermite](#) para integrais da forma

$$\int_{-\infty}^{\infty} e^{-x^2} f(x)dx.$$

³E esta é a única escolha de abscissas e pesos que garante esse grau de exatidão.

APROXIMAÇÃO

O problema de aproximação

O conceito de aproximação desempenha um papel essencial em matemática, em especial, em matemática aplicada. De um modo muito genérico, um problema de aproximação pressupõe a existência de:

- ▶ um determinado objecto u (uma função, um conjunto de dados, etc.) que pretendemos aproximar;
- ▶ um conjunto \mathcal{A} (de “aproximações”);
- ▶ um processo de seleccionar uma aproximação para u de entre os elementos de \mathcal{A} .

É muito frequente escolher como aproximações elementos do espaço \mathcal{P}_n (dos polinómios de coeficientes reais, de grau não superior a n , com $n \in \mathbb{N}$, fixado), sendo esse o caso a que nos dedicaremos neste curso.

Distância induzida por uma norma

Estaremos especialmente interessados em distâncias que sejam induzidas por normas. Mais precisamente, se $(V, \|\cdot\|)$ é um espaço vetorial normado, então é fácil de verificar que a função $d: V \times V \rightarrow \mathbb{R}$ definida por

$$d(u, v) = \|u - v\| \quad (20)$$

é uma métrica em V . Por outras palavras, todo o espaço vetorial normado é um espaço métrico e é neste tipo de espaços que iremos trabalhar.

Conceito de distância

É, naturalmente, importante dispor de uma maneira de medir a qualidade da aproximação, isto é, de medir a “distância” entre o objecto u a aproximar e o elemento aproximador. A este propósito, tem-se a seguinte definição.

Definição (Distância)

Seja F um conjunto qualquer. Uma aplicação $d: F \times F \rightarrow \mathbb{R}$ diz-se uma **métrica** ou **distância** definida em F , se verificar:

$$D1 \quad \forall f, g \in F, d(f, g) \geq 0.$$

$$D2 \quad \forall f, g \in F, d(f, g) = 0 \iff f = g.$$

$$D3 \quad \forall f, g \in F, d(f, g) = d(g, f).$$

$$D4 \quad \forall f, g, h \in F, d(f, g) \leq d(f, h) + d(h, g).$$

Um conjunto no qual esteja definida uma métrica diz-se um espaço métrico.

Melhor aproximação (num espaço normado)

Começamos por introduzir a seguinte definição.

Definição

Seja $(V, \|\cdot\|)$ um espaço vetorial normado e seja \mathcal{A} um subconjunto não vazio de V . Dado um elemento $v \in V$, dizemos que $v^* \in \mathcal{A}$ é uma **melhor aproximação** para v por elementos de \mathcal{A} ou uma **melhor \mathcal{A} -aproximação** para v , se v^* satisfizer

$$\|v - v^*\| \leq \|v - u\|, \forall u \in \mathcal{A}. \quad (21)$$

Por outras palavras, v^* é uma melhor \mathcal{A} -aproximação para v se não existir em \mathcal{A} nenhum elemento cuja distância a v seja inferior à de v^* .

Nota: Naturalmente, se $v \in \mathcal{A}$, então v é a única melhor aproximação para v por elementos de \mathcal{A} .

Distância de um elemento a um subconjunto

Definição

Dado $v \in V$ e $\emptyset \neq \mathcal{A} \subset V$, ao valor $d(v, \mathcal{A})$ definido por

$$d(v, \mathcal{A}) := \inf_{u \in \mathcal{A}} \|v - u\| \quad (22)$$

chamamos **distância de v ao conjunto \mathcal{A}** .

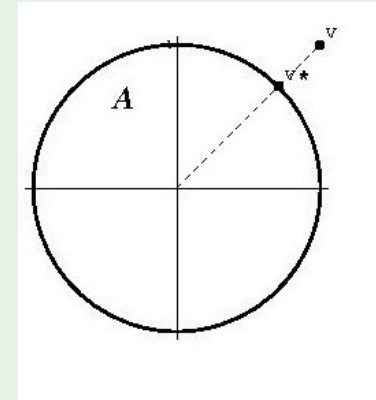
Nota: Se $v^* \in \mathcal{A}$ é uma melhor \mathcal{A} -aproximação para v , então

$$\|v - v^*\| = d(v, \mathcal{A}) = \min_{u \in \mathcal{A}} \|v - u\|. \quad (23)$$

Os dois exemplos seguintes mostram que uma melhor aproximação pode, ou não, existir.

Exemplo

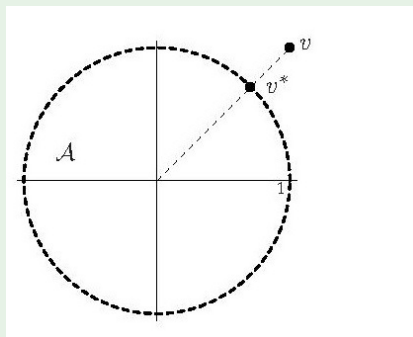
Seja $V = \mathbb{R}^2$, com a norma euclidiana e seja $\mathcal{A} = \{u \in V : \|u\| \leq 1\}$. Facilmente se verifica que, para qualquer elemento $v \in \mathbb{R}^2$, existe uma melhor aproximação $v^* \in \mathcal{A}$.



Exemplo

Seja, agora, $\mathcal{A} = \{u \in V : \|u\| < 1\}$. Então, dado um qualquer elemento $v \in \mathbb{R}^2$, não existe $v^* \in \mathcal{A}$ tal que

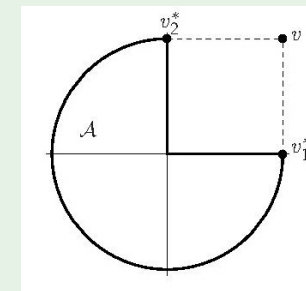
$$\|v - v^*\| \leq \|v - u\|, \quad \forall u \in \mathcal{A}.$$



Para além da questão da existência ou não de uma melhor aproximação, convém também referir que, quando ela existe, poderá não ser única.

Exemplo

Seja, agora $\mathcal{A} = \{u = (x, y) \in V : \|u\| \leq 1 \text{ e } (x \leq 0 \text{ ou } y \leq 0)\}$.



Neste caso, ambos os pontos $v_1^* = (1, 0)$ e $v_2^* = (0, 1)$ são \mathcal{A} -melhores aproximações para $v = (1, 1)$.

Nas aplicações, é especialmente importante o caso em que $\mathcal{A} = U$, onde U é um **subespaço vetorial de dimensão finita** do espaço vetorial normado V . Neste caso, pode estabelecer-se a existência de melhor aproximação, ou seja, tem-se o seguinte teorema (veja, e.g. G. Hämerlin e K.-H. Hoffmann, *Numerical Mathematics*, p. 139.)

Teorema

Se U é um **subespaço vetorial de dimensão finita** de um espaço vetorial normado V , então todo o elemento $v \in V$ tem (pelo menos) uma melhor aproximação por elementos de U .

Definição

Seja $(V, \|\cdot\|)$ um espaço vetorial normado. A norma $\|\cdot\|$ diz-se uma norma **estrita** (e V diz-se um espaço **estritamente normado**) se for válida a seguinte propriedade:

$$\forall u, v \in V \quad \|u + v\| = \|u\| + \|v\| \implies u \text{ e } v \text{ são linearmente dependentes.}$$

Pode mostrar-se (demonstração nas aulas práticas) que, se U é um subespaço vetorial de um espaço estritamente normado, então todo o elemento de $v \in V$ tem, **no máximo**, uma melhor aproximação por elementos de U .

Num espaço com produto interno é válida a seguinte desigualdade, conhecida por **desigualdade de Cauchy-Schwarz**:

$$|\langle u|v \rangle| \leq \|u\| \|v\|,$$

tendo-se, além disso:

$$|\langle u|v \rangle| = \|u\| \|v\| \iff u \text{ e } v \text{ são linearmente dependentes.}$$

Pode mostrar-se (demonstração nas aulas práticas) que a norma induzida por um produto interno é uma norma estrita.

Assim, temos:

Melhor aproximação num espaço com produto interno

Se U for um **subespaço vetorial de dimensão finita** de um espaço com produto interno V , então todo o elemento $v \in V$ admite **uma e uma só** melhor aproximação por elementos de U .

Em resumo, tem-se o seguinte resultado:

Melhor aproximação num espaço estritamente normado

Num espaço vetorial **estritamente normado**, qualquer elemento admite **uma e uma só** melhor aproximação por elementos de um seu **subespaço de dimensão finita**.

Aproximação em espaços com produto interno

Se V for um espaço onde esteja definido um produto interno $\langle \cdot | \cdot \rangle$, então é possível definir em V uma norma, do seguinte modo:

$$\|v\| = \sqrt{\langle v|v \rangle}.$$

Esta norma diz-se **induzida** pelo produto interno considerado.

Por outras palavras, todo o espaço com produto interno é um espaço normado; quando, num espaço com produto interno, falarmos em norma (e correspondente distância) será sempre a esta norma induzida que nos queremos referir.

O teorema seguinte caracteriza essa melhor aproximação.

Nota: No que se segue, por uma questão de simplicidade, assumimos que V é um espaço vetorial **real**; as adaptações para o caso complexo são muito simples.

Teorema

Seja V um espaço vetorial com produto interno $\langle \cdot | \cdot \rangle$ e seja U um seu subespaço vetorial de dimensão finita. Dado $v \in V$, tem-se que $v^* \in U$ é **a melhor U -aproximação** para v se e só se $v - v^*$ for ortogonal a todos os elementos de U , isto é, se e só se

$$\langle v - v^* | u \rangle = 0, \quad \forall u \in U. \quad (24)$$

Nota: O subconjunto de V formado pelos vetores que são ortogonais a todos os elementos de U é chamado complemento ortogonal de U (em V) e denotado por U^\perp , i.e.

$$U^\perp = \{w \in V : \langle w|u \rangle = 0, \forall u \in U\}.$$

Demonstração

Condição Suficiente

Suponhamos que $v^* \in U$ satisfaz a condição (24). Dado um elemento arbitrário $u \in U$, escrevamo-lo como

$$u = v^* + \underbrace{(u - v^*)}_{u'}$$

onde $u' = u - v^* \in U$ (porque U é um subespaço vetorial de V). Então, tem-se

$$\begin{aligned}\|v - u\|^2 &= \|v - v^* - u'\|^2 = \|v - v^*\|^2 + \|u'\|^2 - 2\underbrace{\langle v - v^* | u' \rangle}_{=0} \\ &= \|v - v^*\|^2 + \|u'\|^2\end{aligned}$$

(Na segunda igualdade usámos o resultado do Exercício 1. a) das folha de exercícios sobre produto interno). Como $\|u'\|^2 \geq 0$, segue-se que $\|v - v^*\|^2 \leq \|v - u\|^2$ ou seja, que

$$\|v - v^*\| \leq \|v - u\|.$$

Relembrando que u é um elemento arbitrário de U , a desigualdade anterior mostra que v^* é (um)a melhor aproximação para v por elementos de U .

Corolário

Nas condições do teorema anterior, a distância $v - v^*$ é dada por

$$\|v - v^*\| = \sqrt{\|v\|^2 - \|v^*\|^2} \quad (25)$$

Dem: Temos

$$\begin{aligned}\|v\|^2 &= \|v - v^* + v^*\|^2 \\ &= \|v - v^*\|^2 + \|v^*\|^2 + 2\langle v - v^* | v^* \rangle \\ &= \|v - v^*\|^2 + \|v^*\|^2\end{aligned}$$

já que $v^* \in U$ e, portanto, $\langle v - v^* | v^* \rangle = 0$. Tem-se, então

$$\|v - v^*\|^2 = \|v\|^2 - \|v^*\|^2$$

e o resultado segue-se de imediato.

Demonstração (cont.)

Condição necessária

Seja v^* a melhor U -aproximação para v e suponhamos que, para um certo elemento $u \in U$, se tem $\langle v - v^* | u \rangle = c \neq 0$. Vejamos que isso leva a uma contradição. Note-se que, neste caso, $u \neq 0$ e considere-se, então, o vetor

$$w = v^* + c \frac{u}{\|u\|^2}.$$

Atendendo a que U é um subespaço vetorial de V e v^* e u são elementos de U , podemos concluir que $w \in U$. Além disso, tem-se

$$\begin{aligned}\|v - w\|^2 &= \langle v - v^* - c \frac{u}{\|u\|^2} | v - v^* - c \frac{u}{\|u\|^2} \rangle \\ &= \|v - v^*\|^2 - 2 \frac{c}{\|u\|^2} \langle v - v^* | u \rangle + \frac{c^2}{\|u\|^2} \\ &= \|v - v^*\|^2 - \frac{c^2}{\|u\|^2} < \|v - v^*\|^2,\end{aligned}$$

que contradiz o facto de v^* ser a melhor aproximação para v por elementos de U .

O teorema seguinte fornece um processo de determinação da melhor aproximação para um dado elemento de um espaço com produto interno por elementos de um seu subespaço de dimensão finita.

Teorema

Seja V um espaço vetorial com um produto interno $\langle \cdot | \cdot \rangle$, seja U um seu subespaço de dimensão finita e seja (u_1, \dots, u_m) uma base de U . Dado $v \in V$, tem-se que $v^* \in U$ é a melhor U -aproximação para v se e só se v^* for dado por

$$v^* = \sum_{j=1}^m c_j u_j \quad (26)$$

com os coeficientes c_j satisfazendo

$$\sum_{j=1}^m \langle u_i | u_j \rangle c_j = \langle v | u_i \rangle; \quad i = 1, 2, \dots, m. \quad (27)$$

Demonstração

Como (u_1, \dots, u_m) é uma base de U e $v^* \in U$, é evidente que v^* tem uma e uma só expansão na forma (26). Resta-nos mostrar que os coeficientes c_j satisfazem as equações (27). De acordo com o teorema anterior, sabemos que v^* é a melhor U -aproximação para v se e só se tivermos

$$\langle v - v^* | u \rangle = \langle v - \sum_{j=1}^m c_j u_j | u \rangle = 0, \quad \forall u \in U.$$

Atendendo às propriedades do produto interno e a ao facto de $(u_i)_{i=1}^m$ ser uma base de U , é imediato reconhecer que a condição anterior é equivalente às seguintes equações

$$\langle v - \sum_{j=1}^m c_j u_j | u_i \rangle = 0, \quad i = 1, \dots, m$$

ou seja, a termos

$$\sum_{j=1}^m \langle u_j | u_i \rangle c_j = \langle v | u_i \rangle, \quad i = 1, \dots, m,$$

de onde se obtém de imediato o resultado, atendendo a que $\langle u_j | u_i \rangle = \langle u_i | u_j \rangle$ (recorde que estamos a assumir que o espaço vetorial é real).

Calculada a U -aproximação v^* para v , a distância $d(v, U) = \|v - v^*\|$ pode ser calculada facilmente. Tem-se (ver corolário da pg. 16)

$$\|v - v^*\|^2 = \|v\|^2 - \|v^*\|^2.$$

Mas,

$$\begin{aligned} \|v^*\|^2 &= \langle v^* | v^* \rangle = \langle v^* - v + v | v^* \rangle \\ &= \langle v^* - v | v^* \rangle + \langle v | v^* \rangle \\ &= \langle v | v^* \rangle = \sum_{j=1}^m c_j \langle v | u_j \rangle. \end{aligned}$$

Assim, temos que

$$\|v - v^*\| = \sqrt{\|v\|^2 - \sum_{j=1}^m c_j \langle v | u_j \rangle}. \quad (28)$$

Equações normais

As equações (27) são chamadas **equações normais** para a determinação dos coeficientes da melhor aproximação (na base escolhida). Estas equações podem escrever-se na forma matricial como

$$G\vec{c} = \vec{b}$$

onde:

- ▶ $G = [g_{ij}]$, com $g_{ij} = \langle u_i | u_j \rangle$; $i, j = 1, \dots, m$
- ▶ $\vec{c} = [c_1, \dots, c_m]^T$
- ▶ $\vec{b} = [b_1, \dots, b_m]^T$, com $b_i = \langle v | u_i \rangle$.

A matriz $G = [\langle u_i | u_j \rangle]$ é dita **matriz de Gram**. Pode mostrar-se que esta matriz é simétrica e definida positiva, o que confirma que o sistema tem solução única (como teria de ser, uma vez que sabemos que a melhor aproximação existe é única, neste caso).

Aproximação Polinomial dos Mínimos Quadrados

Aproximação de uma função (contínua)

Consideremos agora o caso particular em que:

- ▶ $V = C[a, b]$;
- ▶ em V se considera um produto interno da forma

$$\langle f | g \rangle = \int_a^b w(x) f(x) g(x) dx \quad (29)$$

onde $w(x)$ é uma função definida em (a, b) , positiva e integrável nesse intervalo;

- ▶ se toma para subespaço U de aproximações o espaço vetorial real \mathcal{P}_n dos polinómios de coeficientes reais de grau não superior a n , $n \in \mathbb{N}$ (fixado). (Naturalmente, consideramos os polinómios como funções restringidas a $[a, b]$).

Note-se que \mathcal{P}_n é um subespaço vetorial de $C[a, b]$ de dimensão $n + 1$. Assim, dada uma função $f \in C[a, b]$, existe um e um só um polinómio $p_n^* \in \mathcal{P}_n$ que, de entre todos os polinómios de \mathcal{P}_n , melhor aproxima f relativamente à distância correspondente à norma induzida pelo produto interno, i.e., é, que satisfaz

$$\|f - p_n^*\|^2 = \min_{p \in \mathcal{P}_n} \|f - p\|^2 = \min_{p \in \mathcal{P}_n} \int_a^b w(x) (f(x) - p(x))^2 dx. \quad (30)$$

Tal polinómio p_n^* é chamado **aproximação polinomial dos mínimos quadrados (ponderados, se $w(x) \neq 1$) de grau n para f ou, mais simplesmente, polinómio (de grau n) dos mínimos quadrados (ponderados) para f .**

Exemplo

Seja $V = C[0, 1]$ com produto interno $\langle f|g \rangle = \int_0^1 f(x)g(x)dx$ e suponhamos que escolhíamos para base de \mathcal{P}_n a base canónica $1, x, \dots, x^n$. Então, a matriz G das equações normais seria $G = [g_{ij}]$ onde

$$g_{ij} = \int_0^1 x^{i-1}x^{j-1}dx = \frac{1}{i+j-1}; i, j = 1, \dots, n+1.$$

Assim,

$$G = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n+1} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+2} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n+1} \end{bmatrix}$$

seria a matriz de Hilbert de ordem $n + 1$, a qual, como sabemos, é uma matriz mal condicionada (mesmo para valores de n razoavelmente pequenos). O sistema das equações normais seria, assim, extremamente sensível a pequenas alterações nos seus coeficientes ou no vetor dos termos independentes.

De acordo com o teorema anterior, escolhida uma base $(\phi_1, \dots, \phi_{n+1})$ para \mathcal{P}_n , o polinómio p_n^* admite uma e uma única expansão nessa base

$$p_n^* = \sum_{j=1}^{n+1} c_j \phi_j, \quad (31)$$

sendo os coeficientes c_j obtidos como solução do sistema

$$\sum_{j=1}^{n+1} \langle \phi_i | \phi_j \rangle c_j = \langle f | \phi_i \rangle; i = 1, \dots, n+1$$

O exemplo anterior mostra como é importante a escolha da base de \mathcal{P}_n na qual se representa a aproximação dos mínimos quadrados. A solução do sistema de equações normais será particularmente simples se escolhermos a base de \mathcal{P}_n formada pelos primeiros $n + 1$ **polinómios ortogonais** relativamente ao produto interno considerado. De facto, nesse caso, tem-se

$$g_{ij} = \langle \phi_i | \phi_j \rangle = 0, \text{ se } i \neq j,$$

isto é, a matriz G é uma **matriz diagonal**, vindo os coeficientes c_j dados simplesmente por

$$c_i = \frac{\langle f, \phi_i \rangle}{\langle \phi_i, \phi_i \rangle}; i = 1, \dots, n+1, \quad (32)$$

não havendo problemas de instabilidade no seu cálculo.

Exemplo

Determinar a aproximação polinomial dos mínimos quadrados de grau 3 para a função $f(x) = e^x, x \in [-1, 1]$, relativamente ao produto interno $\langle f|g \rangle = \int_{-1}^1 f(x)g(x)dx$.

Esta aproximação é dada por

$$p_3^*(x) = c_0P_0(x) + c_1P_1(x) + c_2P_2(x) + c_3P_3(x),$$

onde P_0, P_1, P_2, P_3 são os primeiros quatro polinómios de Legendre, $P_0(x) = 1, P_1(x) = x, P_2(x) = \frac{1}{2}(3x^2 - 1), P_3(x) = \frac{1}{2}(5x^3 - 3x)$, sendo os coeficientes c_i dados simplesmente por

$$c_i = \frac{\langle f|P_i \rangle}{\langle P_i|P_i \rangle}; i = 0, 1, 2, 3.$$

Mas (veja propriedades dos pol. Legendre), tem-se

$$\langle P_i|P_i \rangle = \|P_i\|^2 = \frac{2}{2i+1}.$$

Exemplo (cont.)

Resta-nos, portanto, calcular

$$I_0 = \langle f|P_0 \rangle = \int_{-1}^1 e^x dx, \quad I_1 = \langle f|P_1 \rangle = \int_{-1}^1 x e^x dx$$

$$I_2 = \langle f|P_2 \rangle = \frac{1}{2} \int_{-1}^1 (3x^2 - 1)e^x dx \quad \text{e} \quad I_3 = \langle f|P_3 \rangle = \frac{1}{2} \int_{-1}^1 (5x^3 - 3x)e^x dx.$$

Tem-se (com 8 dígitos de precisão)

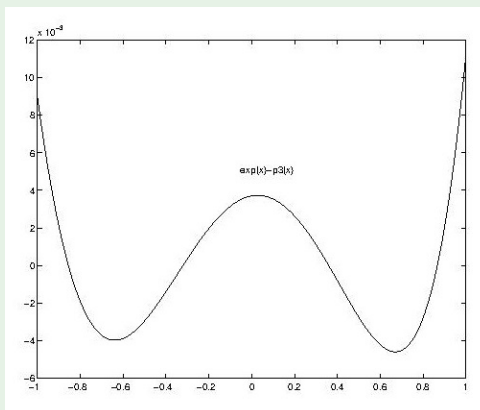
$$I_0 = 2.3504024, \quad I_1 = 0.73575888, \quad I_2 = 0.14312574, \quad I_3 = 0.020130181,$$

pelo que

$$p_3^* = \frac{2.3504024}{2} + \frac{3}{2}0.73575888x + \frac{5}{4}0.14312574(3x^2 - 1) + \frac{7}{4}0.020130181(5x^3 - 3x).$$

Exemplo (cont.)

Na figura seguinte apresenta-se o gráfico da função $E(x) = e^x - p_3^*(x)$, no intervalo $[-1, 1]$.



Aproximação Polinomial dos Mínimos Quadrados

Caso de um conjunto de pontos

Sejam dados m pontos $(x_1, y_1), \dots, (x_m, y_m)$ (x_i distintos), onde cada y_i foi, por exemplo, obtido experimentalmente e aproxima o valor de uma certa função y em x_i , e suponhamos que se pretende “reconstruir” y , isto é, determinar uma aproximação para a função y , usando a informação dada.

Seja \mathcal{P}_n , com $n < m$, o espaço de onde se pretende extrair essa aproximação.

Consideremos o espaço vetorial $V = \mathbb{R}^m$, com um produto interno

$$\langle \vec{u}, \vec{v} \rangle = \sum_{i=1}^m w_i u_i v_i, \quad w_i > 0, \quad (33)$$

(nesta secção, por uma questão de clareza, os vetores do espaço \mathbb{R}^m serão denotados por \vec{u}, \vec{v} etc.) e correspondente norma

$$\|\vec{u}\| = \sqrt{\langle \vec{u}|\vec{u} \rangle} = \sqrt{\sum_{i=1}^m w_i u_i^2}. \quad (34)$$

Considere-se a aplicação

$$\begin{aligned} \pi : \mathcal{P}_n &\longrightarrow \mathbb{R}^m \\ p &\mapsto \vec{p} = [p(x_1), \dots, p(x_m)]^T \end{aligned}$$

e seja U o seu contradomínio, isto é:

$$U := \pi(\mathcal{P}_n) = \{[p(x_1), \dots, p(x_m)]^T : p \in \mathcal{P}_n\}.$$

É fácil de verificar (verifique!) que:

- ▶ A aplicação π é injetiva (tenha em atenção que $n < m$ e relembre os resultados sobre unicidade do polinómio interpolador);
- ▶ o contradomínio de π é um subespaço vetorial (de dimensão finita, naturalmente) de \mathbb{R}^m .

Escolhida uma base $(\phi_1, \dots, \phi_{n+1})$ de \mathcal{P}_n , é imediato verificar que $(\vec{\phi}_1, \vec{\phi}_2, \dots, \vec{\phi}_n)$, onde

$$\vec{\phi}_k = \pi(\phi_k) = [\phi_k(x_1), \dots, \phi_k(x_m)]^T$$

constitui uma base do espaço $U = \pi(\mathcal{P}_n)$. Então, um decalque do estudo que fizemos para o caso contínuo permite-nos concluir que o polinómio

$$p_n^* = \sum_{j=1}^{n+1} c_j \phi_j$$

será a aproximação polinomial dos mínimos quadrados de grau n para o conjunto de pontos $(x_1, y_1), \dots, (x_m, y_m)$ se e só se os coeficientes c_j satisfizerem as equações normais

$$\sum_{j=1}^{n+1} \langle \vec{\phi}_i | \vec{\phi}_j \rangle c_j = \langle \vec{y} | \vec{\phi}_i \rangle; \quad i = 1, \dots, n+1, \quad (35)$$

onde o produto interno em causa é o produto que considerámos em \mathbb{R}^m , isto é

$$\langle \vec{u} | \vec{v} \rangle = \sum_{i=1}^m w_i u_i v_i.$$

Como $\vec{y} = [y_1, \dots, y_m]^T \in \mathbb{R}^m$, sabemos que existe um e um só vetor \vec{y}^* no subespaço U , que melhor aproxima \vec{y} relativamente à norma considerada, isto é, tal que

$$\|\vec{y} - \vec{y}^*\| \leq \|\vec{y} - \vec{u}\|, \quad \forall \vec{u} \in U.$$

Mas, atendendo à definição de U , $U = \pi(\mathcal{P}_n)$, sabemos que esse vetor \vec{y}^* é necessariamente da forma

$$\vec{y}^* = \pi(p_n^*) = [p_n^*(x_1), \dots, p_n^*(x_m)]^T,$$

para um certo polinómio $p_n^* \in \mathcal{P}_n$. Como π é injetiva, esse polinómio p_n^* é univocamente determinado. Assim, podemos concluir que **existe um e um só polinómio** $p_n^* \in \mathcal{P}_n$ que satisfaz

$$\sum_{i=1}^m w_i (y_i - p_n^*(x_i))^2 \leq \sum_{i=1}^m w_i (y_i - p(x_i))^2, \quad \forall p \in \mathcal{P}_n$$

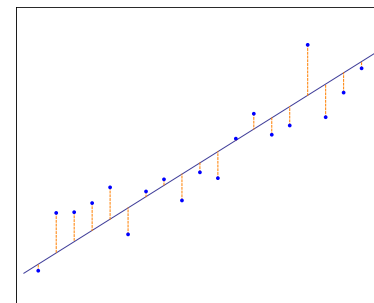
A esse polinómio p_n^* chamamos polinómio dos mínimos quadrados (ponderados, se os w_i não forem todos iguais a 1) de grau n para os pontos $(x_1, y_1), \dots, (x_m, y_m)$.

Exemplo: Retas de regressão

Considere-se o caso em que são dados m pontos

$$(x_i, y_i); \quad i = 1, \dots, m,$$

e se pretende obter o polinómio dos mínimos quadrados **de grau ≤ 1** para esse conjunto de pontos (considerando os pesos $w_i \equiv 1$). Este polinómio é vulgarmente designado por **reta de regressão** relativa a esse conjunto de pontos.



Sejam $\phi_1(x) = 1$ e $\phi_2(x) = x$, os polinómios da base canónica de \mathcal{P}_1 e determinemos, então, as correspondentes equações normais.

Ter-se-á:

►
$$\vec{\phi}_1 = \pi(\phi_1) = [1, 1, \dots, 1]^T$$

►
$$\vec{\phi}_2 = \pi(\phi_2) = [x_1, x_2, \dots, x_m]^T.$$

Então, vem:

$$\langle \vec{\phi}_1 | \vec{\phi}_1 \rangle = \sum_{i=1}^m (1 \times 1) = m$$

$$\langle \vec{\phi}_1 | \vec{\phi}_2 \rangle = \langle \vec{\phi}_2 | \vec{\phi}_1 \rangle = \sum_{i=1}^m (1 \times x_i) = \sum_{i=1}^m x_i$$

$$\langle \vec{\phi}_2 | \vec{\phi}_2 \rangle = \sum_{i=1}^m (x_i \times x_i) = \sum_{i=1}^m x_i^2.$$

$$\langle \vec{y} | \vec{\phi}_1 \rangle = \sum_{i=1}^m (y_i \times 1) = \sum_{i=1}^m y_i$$

$$\langle \vec{y} | \vec{\phi}_2 \rangle = \sum_{i=1}^m (y_i \times x_i)$$

Assim, temos que

$$r(x) = a + bx \tag{36}$$

é a **reta de regressão** para o conjunto de pontos $\{(x_i, y_i) : i = 1, \dots, m\}$ se e só se os coeficientes a, b forem solução do sistema

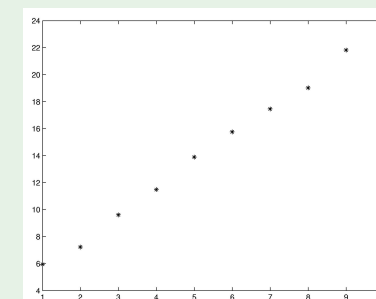
$$\begin{bmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \end{bmatrix} \tag{37}$$

Exemplo

Considerem-se os dados constantes da seguinte tabela

x_i	y_i	x_i	y_i
1	5.6644	6	15.5560
2	7.6427	7	17.5343
3	9.6210	8	19.5127
4	11.5994	9	21.4910
5	13.5777	10	23.4693

e representados graficamente na figura abaixo:



Exemplo (cont.)

É bem evidente que os dados “quase” se dispõem sobre uma reta, ou seja que fará sentido ajustar uma função do tipo

$$y(x) = a + bx$$

aos dados em causa, isto é, procurar uma função aproximadora do espaço \mathcal{P}_1 . Neste caso, obtém-se o seguinte sistema de equações normais

$$\begin{bmatrix} 10 & 55 \\ 55 & 385 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 145.6686 \\ 964.3889 \end{bmatrix},$$

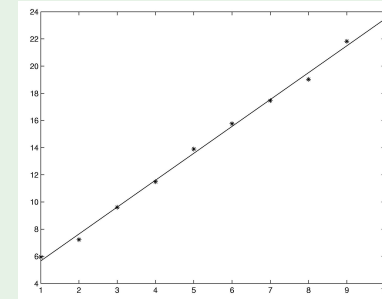
cuja solução é $a = 3.6861$ e $b = 1.9783$.

Assim, a reta dos mínimos quadrados para os dados considerados é

$$y = 3.6861 + 1.9783x.$$

Exemplo (cont.)

Na figura seguinte, apresentam-se os dados tabelados e a respectiva reta dos mínimos quadrados:



Exemplo

Considerem-se agora os dados da seguinte tabela

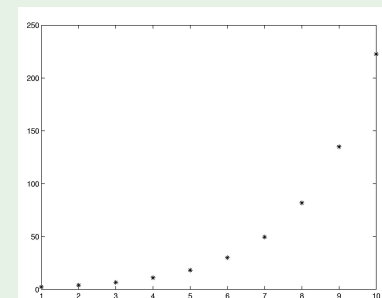
x_i	y_i	x_i	y_i
1	2.4739	6	30.1293
2	4.0784	7	49.6735
3	6.7231	8	81.8975
4	11.0845	9	135.0266
5	18.2739	10	222.6205

Exemplo (cont.)

A representação gráfica desses dados, apresentada na figura seguinte, sugere um crescimento exponencial de y com x , pelo que será mais razoável ajustar uma curva da forma

$$y = Ae^{Bx} \quad (38)$$

a esses dados. Como poderemos determinar A e B ?



Exemplo (cont.)

Aplicando logaritmos a ambos os lados da expressão (38), obtém-se

$$\ln y = \ln A + Bx \quad (39)$$

pelo que poderemos resolver o problema ajustando uma reta aos valores

$$(x_i, \ln y_i).$$

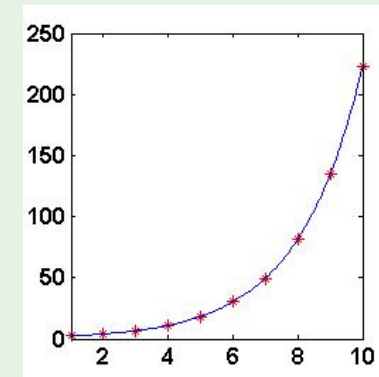
Encontrada a solução $y = a + bx$ que se ajusta a esses dados, virá, de (39),

$$a = \ln A \Rightarrow A = \exp(a) \quad \text{e} \quad B = b.$$

Para os dados da tabela, obtém-se $a = 0.4056$, $b = 0.7$, donde se segue, $A = 1.5002$ e $B = 0.7$.

Exemplo (cont.)

A figura seguinte apresenta a curva $y = 1.5002e^{0.7x}$ e os dados tabelados.



Linearização

Na tabela seguinte indicam-se transformações adequadas para linearizar os dados, para diversas relações funcionais entre y e x .

$y = Ae^{Bx}$	$\ln y = \ln A + Bx$
$y = AB^x$	$\ln y = \ln A + (\ln B)x$
$y = Ax^B$	$\ln y = \ln A + B \ln x$
$y = \frac{1}{1+Ae^{Bx}}$	$\ln\left(\frac{1}{y} - 1\right) = \ln A + Bx$

Equações Diferenciais Ordinárias

As equações diferenciais são usadas para descrever um grande número de fenómenos em ciências e tecnologia. Para grande parte das equações diferenciais que aparecem nos problemas não é possível encontrar uma solução por processos analíticos, sendo, assim, indispensável o recurso a métodos numéricos para a obtenção de uma solução (aproximada). De facto, a necessidade de resolver numericamente equações diferenciais foi uma das grandes motivações para o desenvolvimento dos computadores. Ainda hoje, grande parte do tempo de “computação científica” é gasto na resolução de problemas de equações diferenciais.

Nesta UC, daremos uma breve introdução ao problema da resolução numérica de equações diferenciais **ordinárias**.

Naturalmente, muitos problemas envolvem equações diferenciais de ordem superior à primeira; note-se, todavia, que tais equações podem ser reformuladas como um **sistema** de equações diferenciais de primeira ordem. De facto, dada uma equação diferencial de ordem k da forma

$$y^{(k)} = f(x, y, y', \dots, y^{(k-1)}),$$

se definirmos as funções auxiliares $Y_1(x) := y(x)$, $Y_2(x) := y'(x)$, ..., $Y_k(x) := y^{(k-1)}(x)$, poderemos escrever a equação dada, de forma equivalente, como o seguinte sistema de primeira ordem:

Mais precisamente, vamos considerar apenas equações diferenciais ordinárias **de primeira ordem** que possam ser escritas na forma

$$y'(x) = f(x, y(x)) \tag{40}$$

(geralmente, abreviada para $y' = f(x, y)$), onde f é uma função dada de duas variáveis (definida e contínua num certo aberto $D \subseteq \mathbb{R}^2$). Começamos por recordar o que se entende por uma solução da equação diferencial (40).

Definição

Seja $f : D \rightarrow \mathbb{R}$ uma função definida e contínua num certo aberto $D \subseteq \mathbb{R}^2$ e seja I um intervalo de \mathbb{R} com interior não vazio. Uma função $u : I \rightarrow \mathbb{R}$ diz-se uma solução, em I , da equação diferencial (40) se e só se $u \in C^1(I)$ e u satisfaz essa equação em I , i.e. $u'(x) = f(x, u(x))$ para todo o $x \in I$; se x for um extremo do intervalo I , a derivada $u'(x)$ deve ser encarada como a derivada lateral adequada.

$$\begin{cases} Y_1'(x) = Y_2(x) \\ Y_2'(x) = Y_3(x) \\ \vdots \\ Y_{k-1}'(x) = Y_k(x) \\ Y_k'(x) = f(x, Y_1(x), \dots, Y_k(x)) \end{cases}$$

o qual se costuma escrever, em notação vetorizada, como

$$\mathbf{Y}' = \mathbf{f}(x, \mathbf{Y}).$$

Os métodos numéricos que descreveremos para uma equação de primeira ordem generalizam-se de maneira simples para sistemas de tais equações, pelo que o caso aqui abordado não é tão restritivo quanto parece.

Em geral, como sabemos, uma equação do tipo (40) admite uma infinidade de soluções.

Exemplo

Consideremos a equação diferencial $y'(x) = y(x)$, isto é, seja $f(x, y) = y$. Facilmente se verifica que a equação dada é satisfeita por qualquer função do tipo $y(x) = Ce^x$, $C \in \mathbb{R}$.

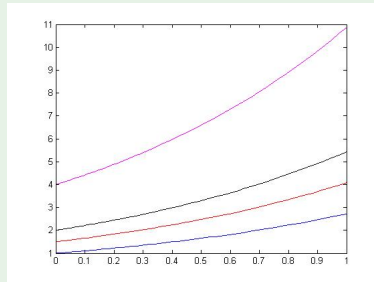


Figura : Várias soluções da equação $y' = y$.

Problema de valor inicial (PVI)

Normalmente, pretende-se apenas encontrar uma solução particular da equação (40); o caso mais frequente consiste em procurar (um)a solução que satisfaça uma condição adicional do tipo

$$y(a) = \alpha, \quad (41)$$

para $(a, \alpha) \in D$. A condição anterior é referida como **condição inicial** e o problema de resolver

$$\begin{cases} y'(x) = f(x, y), \\ y(a) = \alpha \end{cases} \quad (42)$$

é dito um **problema de valor inicial (PVI)**.

Por exemplo, o PVI

$$y'(x) = y(x), \quad y(0) = 1$$

tem como (única) solução a função $y(x) = e^x$.

Métodos de variável discreta

Não entraremos aqui na discussão de problemas de existência e unicidade de solução de um PVI⁴ e admitiremos tacitamente que os problemas que vamos considerar admitem uma única solução, definida num certo intervalo $I = [a, b] \subset \mathbb{R}$, a qual pretendemos aproximar numericamente.

Os métodos numéricos que vamos descrever são baseados na seguinte ideia: não conseguindo determinar $y(x)$ para todo o $x \in [a, b]$, vamos apenas considerar um **conjunto discreto de pontos** x_k em $[a, b]$ e calcular aproximações y_k para os valores exatos $y(x_k)$, i.e. para os valores que y assume em x_k .

Os métodos que determinam aproximações para a solução do problema num conjunto discreto de pontos da variável independente são chamados **métodos de variável discreta**.

Para simplificar, começaremos por supor que esses pontos são igualmente espaçados em $[a, b]$, com espaçamento (ou **passo**) h , isto é, consideraremos

$$x_k = a + (k - 1)h; \quad k = 1, \dots, N + 1, \quad h = \frac{b - a}{N}, \quad (43)$$

para um determinado inteiro N .

O valor inicial dá-nos

$$y_1 = y(x_1) = y(a) = \alpha.$$

A ideia é tentar, a partir deste valor, determinar y_2 como aproximação para $y(x_2)$; depois, obter y_3 como aproximação para $y(x_3)$, etc.

⁴Objeto de estudo de outra UC do seu curso...

Métodos de passo único

Se a fórmula usada para obter y_{k+1} depender apenas de y_k , mas não (diretamente) de y_{k-1}, y_{k-2}, \dots , diremos que temos um **método de passo único**. Mais precisamente, temos a seguinte definição:

Definição

Um método de passo único (explícito)^a para aproximar a solução do PVI (42), correspondente ao uso dos pontos (43), é um método da forma

$$y_{k+1} = y_k + h\phi(x_k, y_k, h); \quad k = 1, \dots, N; \quad y_1 = \alpha, \quad (44)$$

para uma certa função iterativa $\phi : [a, b] \times \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}$.

^aO que definimos aqui é um método de passo único *explícito*; métodos implícitos serão abordados mais à frente.

Método de Euler

Esta é a base do **método de Euler** para aproximar a solução do PVI (42), o qual será, portanto, definido do seguinte modo.

Método de Euler

$$y_{k+1} = y_k + hf(x_k, y_k); \quad k = 1, \dots, N,$$
$$y_1 = \alpha.$$

Note-se que o método de Euler é um método de passo único (44), com função iterativa ϕ dada por

$$\phi(x_k, y_k, h) = f(x_k, y_k).$$

O mais simples deste tipo de métodos é o chamado **método de Euler**, que passamos a descrever.

Nota

Em tudo quanto se segue, admitimos que y admite derivadas contínuas em $[a, b]$, até à ordem indicada.

Se expandirmos $y(x_{k+1}) = y(x_k + h)$ em série de Taylor em torno do ponto x_k , até à 2ª ordem, vem

$$y(x_{k+1}) = y(x_k + h) = y(x_k) + hy'(x_k) + \frac{h^2}{2}y''(\xi_{k+1})$$
$$= y(x_k) + hf(x_k, y(x_k)) + \frac{h^2}{2}y''(\xi_k), \quad \xi_{k+1} \in]x_k, x_{k+1}[. \quad (45)$$

Sendo y'' limitada em $[a, b]$, se o passo h for “pequeno”, será natural ignorar o último termo em (45), tendo-se

$$y(x_{k+1}) \approx y(x_k) + hf(x_k, y(x_k)).$$

Exemplo

Consideremos o PVI

$$y'(x) = xy(x), \quad y(0) = 1,$$

cuja solução exata é $y(x) = e^{x^2/2}$. Suponhamos que se pretende determinar uma aproximação no ponto $x = 0.4$, isto é, que se pretende um valor aproximado para $y(0.4)$. Começemos por considerar $h = 0.2$. Tem-se, $x_1 = 0, x_2 = 0.2$ e $x_3 = 0.4$, sendo y_3 o valor pretendido. Neste caso, $f(x, y) = xy$, pelo que, aplicando o método de Euler, vem

$$y_2 = y_1 + hf(x_1, y_1) = y_1 + hx_1y_1 = 1 + 0.2 \times 0 \times 1 = 1$$

$$y_3 = y_2 + hf(x_2, y_2) = y_2 + hx_2y_2 = 1 + 0.2 \times 0.2 \times 1 = 1.04.$$

Note-se que o valor de $y(0.4)$ (com 3 c.d.) é 1.083, pelo que o erro em $x = 0.4$ é dado por

$$|y(0.4) - y_3| \approx |1.083 - 1.04| = 0.043.$$

Exemplo (cont.)

Se usarmos $h = 0.1$ e aplicarmos novamente o método de Euler, obtemos os valores (com 3 c. d.) constantes da seguinte tabela :

k	x_k	y_k	$y(x_k)$	erro
2	0.1	1.000	1.005	0.005
3	0.2	1.010	1.020	0.010
4	0.3	1.030	1.046	0.016
5	0.4	1.062	1.083	0.022

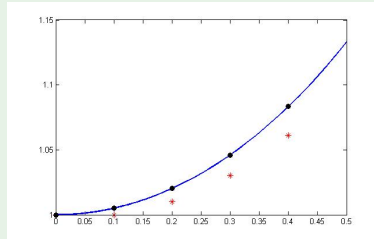


Figura : Solução exata e solução aproximada pelo método de Euler

Erros de discretização local e global

Vamos distinguir dois tipos de erros de discretização num dado ponto x_k : o **erro local** e o **erro global**, que passamos a definir.

Definição (Erro de discretização local num ponto)

Dado um método de passo único (44) (correspondente ao conjunto de pontos (43)) para a solução do PVI $y' = f(x, y)$, $y(a) = \alpha$, chama-se **erro de discretização local** no ponto x_k , $k = 2, \dots, N + 1$, e denota-se por $\mathcal{L}_h(x_k)$, ao erro produzido apenas pela aplicação do **passo $k - 1$** do método, isto é, ao erro cometido ao passar de y_{k-1} para y_k , admitindo que y_{k-1} não tem erro, i.e., **supondo que $y_{k-1} = y(x_{k-1})$** . Por outras palavras, tem-se

$$\mathcal{L}_h(x_k) = y(x_k) - [y(x_{k-1}) + h\phi(x_{k-1}, y(x_{k-1}), h)]. \quad (46)$$

Nota : Naturalmente, no ponto x_1 não há erro, já que $y(x_1) = y_1 = \alpha$; no entanto, se for conveniente, falaremos em $\mathcal{L}_h(x_1)$ (o qual valerá zero).

Como a tabela e a figura anterior indicam, a solução calculada é afetada de erro. Uma questão importante na utilização de um método numérico é, naturalmente, saber qual o tipo de precisão das aproximações y_k para $y(x_k)$. O erro nas aproximações vem de duas fontes:

- ▶ erros de arredondamento, devido ao facto de a aritmética de vírgula flutuante ter precisão finita;
- ▶ erro de truncatura (ou discretização), inerente ao próprio método (e que se manteria, mesmo que a aritmética fosse exata).

O erro de discretização é, em geral, o fator dominante que determina a qualidade da aproximação.

Nota importante

No que se segue, vamos concentrar-nos apenas no estudo do erro de discretização. Assim, nas definições e resultados seguintes, os valores indicados referem-se a valores calculados **ignorando os erros de arredondamento**.

Na prática, ao calcularmos o valor y_k , usaremos y_k , que é uma aproximação para $y(x_{k-1})$, valor que já foi calculado usando y_{k-2} , etc. Há, assim, que ter em conta o efeito acumulado dos erros de discretização locais até se chegar ao valor y_k .

Definição (Erro de discretização global)

Dado um método de passo único (44) (associado ao conjunto de pontos (43)) para a solução do PVI $y' = f(x, y)$, $y(a) = \alpha$, chama-se **erro de discretização global no ponto x_k** , $k = 2, \dots, N + 1$, denota-se por $\epsilon_h(x_k)$, ao erro acumulado na aplicação dos diversos passos até se chegar a y_k , isto é

$$\epsilon_h(x_k) = y(x_k) - y_k. \quad (47)$$

Ao valor dado por

$$E(h) := \max_{k=2, \dots, N+1} |\epsilon_h(x_k)| = \max_{k=2, \dots, N+1} |y(x_k) - y_k| \quad (48)$$

chamamos **erro de discretização global do método**.

Nota : Uma vez mais, poderá ser conveniente referimo-nos a $\epsilon_h(x_1)$ (o qual, naturalmente, terá o valor zero).

No caso do método de Euler, temos $\phi(x_k, y_k, h) = f(x_k, y_k)$ e portanto vem, para o erro de discretização local:

$$\begin{aligned}\mathcal{L}_h(x_k) &= y(x_k) - [y(x_{k-1}) + hf(x_{k-1}, y(x_{k-1}))] \\ &= y(x_k) - [y(x_{k-1}) + hy'(x_{k-1})] \\ &= y(x_{k-1}) + hy'(x_{k-1}) + \frac{h^2}{2}y''(\xi_k) - [y(x_{k-1}) + hy'(x_{k-1})] \\ &= \frac{h^2}{2}y''(\xi_k), \quad \xi_k \in]x_{k-1}, x_k[.\end{aligned}$$

Se $y \in C^2[a, b]$ e se for $M = \max_{a \leq x \leq b} |y''(x)|$, ter-se-á

$$|\mathcal{L}_h(x_k)| \leq \frac{h^2}{2} M.$$

Erro de discretização local do método de Euler

Nas condições anteriores, podemos concluir que o erro de discretização local do método de Euler é $\mathcal{O}(h^2)$.

Assim, temos

$$\begin{aligned}|\epsilon_h(x_k)| &\leq (1 + hL)|\epsilon_h(x_{k-1})| + \frac{h^2}{2}M \\ &\leq (1 + hL) \left[(1 + hL)|\epsilon_h(x_{k-2})| + \frac{h^2}{2}M \right] + \frac{h^2}{2}M \\ &= (1 + hL)^2 |\epsilon_h(x_{k-2})| + (1 + hL) \frac{h^2}{2}M + \frac{h^2}{2}M \\ &\leq (1 + hL)^3 |\epsilon_h(x_{k-3})| + (1 + hL)^2 \frac{h^2}{2}M + (1 + hL) \frac{h^2}{2}M + \frac{h^2}{2}M \\ &\vdots \\ &\leq (1 + hL)^{k-1} \underbrace{|\epsilon_h(x_1)|}_{=0} + (1 + hL)^{k-2} \frac{h^2}{2}M + \dots + (1 + hL) \frac{h^2}{2}M + \frac{h^2}{2}M \\ &= \frac{h^2}{2}M [1 + (1 + hL) + \dots + (1 + hL)^{k-2}] \\ &= \frac{h^2}{2}M \left[\frac{1 - (1 + hL)^{k-1}}{1 - (1 + hL)} \right] \\ &= \frac{hM}{2L} [(1 + hL)^{k-1} - 1]\end{aligned}$$

Passamos agora a deduzir um majorante para erro de discretização global do método de Euler.

Suponhamos que $y \in C^2[a, b]$, sendo $M = \max_{a \leq x \leq b} |y''(x)|$ e que f é diferenciável em relação à segunda variável, com derivada parcial $\frac{\partial f}{\partial y}$ limitada, i.e., existe $L > 0$ tal que $\left| \frac{\partial f(x, y)}{\partial y} \right| \leq L, \forall x \in [a, b], (x, y) \in D$. Temos, então

$$\begin{aligned}|\epsilon_h(x_k)| &= |y(x_k) - y_k| \\ &= \left| y(x_{k-1}) + hf(x_{k-1}, y(x_{k-1})) + \frac{h^2}{2}y''(\xi_k) - y_{k-1} - hf(x_{k-1}, y_{k-1}) \right| \\ &\leq |y(x_{k-1}) - y_{k-1}| + h|f(x_{k-1}, y(x_{k-1})) - f(x_{k-1}, y_{k-1})| + \frac{h^2}{2}M \\ &= |\epsilon_h(x_{k-1})| + h \left| \frac{\partial f}{\partial y}(x_{k-1}, \eta_k) (y(x_{k-1}) - y_{k-1}) \right| + \frac{h^2}{2}M \\ &\leq |\epsilon_h(x_{k-1})| + hL |y(x_{k-1}) - y_{k-1}| + \frac{h^2}{2}M \\ &= |\epsilon_h(x_{k-1})| + hL |\epsilon_h(x_{k-1})| + \frac{h^2}{2}M \\ &= (1 + hL) |\epsilon_h(x_{k-1})| + \frac{h^2}{2}M\end{aligned}$$

Como $hL > 0$, tem-se $1 + hL < e^{hL}$, ou seja, tem-se $(1 + hL)^{k-1} < (e^{hL})^{k-1} = e^{h(k-1)L}$. Logo, da desigualdade anterior, vem

$$|\epsilon_h(x_k)| \leq \frac{hM}{2L} (e^{h(k-1)L} - 1) = \frac{hM}{2L} (e^{(x_k - x_1)L} - 1). \quad (49)$$

Temos, então,

$$E(h) = \max_{2 \leq k \leq N+1} |\epsilon_h(x_k)| \leq \frac{hM}{2L} (e^{(b-a)L} - 1). \quad (50)$$

Erro de discretização global do método de Euler

Nas condições anteriores, podemos concluir que o erro de discretização global do método de Euler é $\mathcal{O}(h)$. Costumamos dizer que o método de Euler é um método de primeira ordem.

Geralmente, sob condições razoáveis, se o erro de discretização local de um método é $\mathcal{O}(h^{p+1})$, então o erro de discretização global é $\mathcal{O}(h^p)$.

Estimativa simples para o erro no método de Euler

Se considerarmos o erro num ponto x_{2k+1} correspondente ao uso de um número par (igual a $2k$) de passadas, tem-se

$$|\epsilon_h(x_{2k+1})| = |y(x_{2k+1}) - y_{2k+1}| \leq Ch,$$

onde C é uma constante independente de h . Se denotarmos por Y_{k+1} a aproximação no mesmo ponto, mas obtida usando apenas metade das passadas, isto é, usando k passadas, o que corresponde ao uso de um passo igual a $2h$, ter-se-á

$$|y(x_{2k+1}) - Y_{k+1}| \leq C(2h).$$

Temos, então que

$$\begin{aligned} |y(x_{2k+1}) - y_{2k+1}| &\approx \frac{1}{2} |y(x_{2k+1}) - Y_{k+1}| = \frac{1}{2} |y(x_{2k+1}) - y_{2k+1} + y_{2k+1} - Y_{k+1}| \\ &\leq \frac{1}{2} |y(x_{2k+1}) - y_{2k+1}| + \frac{1}{2} |y_{2k+1} - Y_{k+1}|, \end{aligned}$$

de onde se obtém a seguinte **estimativa** para o erro de discretização no ponto x_{2k+1} (correspondente ao uso do passo h)

$$|\epsilon_h(x_{2k+1})| = |y(x_{2k+1}) - y_{2k+1}| \approx |y_{2k+1} - Y_{k+1}|.$$

Métodos de Runge-Kutta

Uma vez que $y' = f(x, y)$, ao calcular as derivadas de y ordem superior à primeira, será necessário calcular as derivadas parciais da função f (como vimos no exemplo anterior), o que pode tornar estes métodos bastante trabalhosos. Os chamados métodos de Runge-Kutta foram desenvolvidos com o objectivo de produzir resultados com a mesma ordem de precisão dos métodos obtidos pela expansão em série de Taylor, mas evitando o cálculo das diversas derivadas da função f .

Limitar-nos-emos, aqui, a deduzir a fórmula do **método de Runge-Kutta de 2ª ordem**.

A fórmula do método de 4ª ordem (um dos mais populares) será apresentada sem a sua dedução, a qual, no entanto, poderia fazer-se seguindo o raciocínio usado na dedução da fórmula que apresentamos.

Métodos baseados na série de Taylor

Recordemos que o método de Euler foi deduzido truncando a expansão em série de Taylor de $y(x_{k+1})$ em torno de x_k antes do termo $\mathcal{O}(h^2)$. Métodos de ordem superior poderiam ser obtidos de modo análogo, retendo mais termos da série de Taylor. Por exemplo, tem-se

$$\begin{aligned} y(x_{k+1}) &= y(x_k) + hy'(x_k) + \frac{h^2}{2} y''(x_k) + \mathcal{O}(h^3) \\ &= y(x_k) + hf(x_k, y(x_k)) + \frac{h^2}{2} \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y' \right) (x_k, y(x_k)) + \mathcal{O}(h^3) \\ &= y(x_k) + hf(x_k, y(x_k)) + \frac{h^2}{2} \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} f \right) (x_k, y(x_k)) + \mathcal{O}(h^3) \end{aligned} \quad (51)$$

a que corresponderia o método

$$y_{k+1} = y_k + hf(x_k, y_k) + \frac{h^2}{2} \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} f \right) (x_k, y_k),$$

o qual teria ordem de convergência local $\mathcal{O}(h^3)$ e ordem de convergência global $\mathcal{O}(h^2)$.

Pretende-se determinar constantes A, B, α e β de tal modo que a fórmula

$$y(x_{k+1}) = y(x_k) + h \left[Af(x_k, y(x_k)) + Bf(x_k + \alpha h, y(x_k)) + \beta hf(x_k, y(x_k)) \right] \quad (52)$$

coincida, quando expandida até termos de $\mathcal{O}(h^3)$, com a respetiva parte da expansão em série de Taylor (51).

Usando a fórmula de Taylor para funções de duas variáveis, tem-se

$$\begin{aligned} f(x_k + \alpha h, y(x_k)) + \beta hf(x_k, y(x_k)) &= f(x_k, y(x_k)) \\ &+ \left(\alpha h \frac{\partial f}{\partial x} + \beta h f \frac{\partial f}{\partial y} \right) (x_k, y(x_k)) + \mathcal{O}(h^2). \end{aligned} \quad (53)$$

Substituindo (53) em (52) e colecionando os termos das diversas potências de h , obtém-se

$$y(x_{k+1}) = y(x_k) + (A + B)hf(x_k, y(x_k)) + Bh^2 \left(\alpha \frac{\partial f}{\partial x} + \beta \frac{\partial f}{\partial y} f \right) (x_k, y(x_k)) + \mathcal{O}(h^3) \quad (54)$$

Comparando a expansão anterior com a fórmula (51), vemos elas coincidam até aos termos $\mathcal{O}(h^3)$ se e só se as constantes A, B, α e β satisfizerem

$$\begin{cases} A + B = 1 \\ B\alpha = \frac{1}{2} \\ B\beta = \frac{1}{2} \end{cases}$$

O método correspondente será, então, definido pela fórmula

$$y_{k+1} = y_k + h \left[\frac{1}{2} f(x_k, y_k) + \frac{1}{2} f(x_k + h, y_k + hf(x_k, y_k)) \right]$$

sendo mais frequente apresentá-lo de seguinte modo:

Método de Runge-Kutta de 2ª ordem

$$y_{k+1} = y_k + \frac{1}{2}(k_1 + k_2); \quad k = 1, \dots, N,$$

onde

$$\begin{aligned} k_1 &= hf(x_k, y_k), \\ k_2 &= hf(x_k + h, y_k + k_1), \\ y_1 &= \alpha. \end{aligned}$$

Trata-se de um método cuja erro de discretização local é $\mathcal{O}(h^3)$ e cujo erro global é $\mathcal{O}(h^2)$.

Um solução particularmente simples do sistema será dada por

$$\begin{aligned} A &= B = \frac{1}{2} \\ \alpha &= \beta = 1. \end{aligned}$$

Nota: O sistema anterior é um sistema de 3 equações em 4 incógnitas, pelo que temos ainda um grau de liberdade na sua solução. Poderíamos usar esse grau de liberdade para tentar obter concordância no coeficiente de h^3 . No entanto, pode provar-se que tal não é possível (para todas as funções $f(x, y)$).

De um modo análogo, se deduz a fórmula do método de Runge-Kutta de 4ª ordem.

Método de Runge-Kutta de 4ª ordem

$$y_{k+1} = y_k + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4); \quad k = 1, \dots, N,$$

onde

$$\begin{aligned} k_1 &= hf(x_k, y_k), \\ k_2 &= hf\left(x_k + \frac{h}{2}, y_k + \frac{k_1}{2}\right), \\ k_3 &= hf\left(x_k + \frac{h}{2}, y_k + \frac{k_2}{2}\right), \\ k_4 &= hf(x_k + h, y_k + k_3), \\ y_1 &= \alpha. \end{aligned}$$

Estimativas para o erro nos métodos de Runge-Kutta

Sabendo que a ordem de convergência dos métodos de Runge-Kutta é $\mathcal{O}(h^2)$ e $\mathcal{O}(h^4)$ para os métodos de 2ª e de 4ª ordem, respetivamente, deixamos ao cuidado dos alunos a dedução, de forma totalmente análoga ao que fizemos para o método de Euler, das seguintes **estimativas** para o erro destes métodos (num ponto obtido com um número par de passadas, e onde se usa a notação anterior):

- ▶ Para o método de 2ª ordem,

$$|\epsilon_h(x_{2k+1})| \approx \frac{1}{3} |y_{2k+1} - Y_{k+1}|.$$

- ▶ Para o método de 4ª ordem,

$$|\epsilon_h(x_{2k+1})| \approx \frac{1}{15} |y_{2k+1} - Y_{k+1}|.$$

Se integramos a equação diferencial $y'(x) = f(x, y(x))$ entre o ponto x_k e o ponto x_{k+1} , vem

$$\int_{x_k}^{x_{k+1}} y'(x) dx = \int_{x_k}^{x_{k+1}} f(x, y(x)) dx,$$

ou seja, vem

$$y(x_{k+1}) - y(x_k) = \int_{x_k}^{x_{k+1}} f(x, y(x)) dx.$$

Os chamados métodos de Adams são obtidos substituindo a função integranda $f(x, y(x))$ por um polinómio p escolhido adequadamente, sendo portanto da forma

$$y_{k+1} = y_k + \int_{x_k}^{x_{k+1}} p(x) dx. \quad (55)$$

Dependendo da escolha do polinómio p , obtemos diferentes métodos.

Métodos de Passo Múltiplo

O método de Euler e os métodos de Runge-Kutta são, como dissemos, exemplos de métodos de passo único, em que a fórmula usada para obter y_{k+1} depende apenas (diretamente) do valor aproximado y_k , e não de aproximações de y noutros pontos.

Nos chamados **métodos de passo múltiplo**, a aproximação da solução num certo ponto é calculada usando informação acerca do valor aproximado de y em vários pontos.

Uma classe importante desses métodos é baseada no princípio de integração numérica.

Métodos de Adams-Bashforth

No caso dos chamados **métodos de Adams-Bashforth** (AB), tomamos para p o polinómio p_m de grau não superior a m ($m \in \mathbb{N}$), interpolador dos valores

$$f_i := f(x_i, y_i); \quad i = k, k-1, \dots, k-m,$$

onde $y_k, y_{k-1}, \dots, y_{k-m}$ são aproximações para y nos pontos $x_k, x_{k-1}, \dots, x_{k-m}$ ($k \geq m+1$), que admitimos conhecer.

Podemos provar-se que o método de Adams-Bashforth correspondente ao uso do polinómio de grau m tem erro global $\mathcal{O}(h^{m+1})$, isto é, é um método de ordem $m+1$.

Método de Adams-Bashforth de ordem 2 ($m = 1$)

Neste caso, tomamos para p o polinómio linear p_1 que satisfaz $p_1(x_{k-1}) = f_{k-1}$ e $p_1(x_k) = f_k$, ou seja, o polinómio dado por

$$p_1(x) = -\frac{x - x_k}{h} f_{k-1} + \frac{x - x_{k-1}}{h} f_k,$$

pelo que o método será definido por

$$y_{k+1} = y_k + \frac{1}{h} \int_{x_k}^{x_{k+1}} [-(x - x_k)f_{k-1} + (x - x_{k-1})f_k] dx, \quad k \geq 2$$

Calculando o integral do lado direito da equação anterior, vem, então:

$$y_{k+1} = y_k + \frac{h}{2}(3f_k - f_{k-1}), \quad k \geq 2.$$

Inicialização dos métodos de passo múltiplo

Os métodos de passo múltiplo têm um problema que não encontramos para os de passo único. O método de ordem $m + 1$ só pode ser usado para calcular y_{k+1} para $k \geq m + 1$; veja, por exemplo, fórmula do método de Adams-Bashforth de 4ª ordem, a qual só pode ser usada para $k \geq 4$, uma vez que envolve o valor de $f_{k-3} = f(x_{k-3}, y_{k-3})$. Isto mostra que os métodos de passo múltiplo precisam de *ajuda* para se iniciar a sua utilização. Como resolver esta questão?

- ▶ A solução usual consiste em usar um método de passo único até que se tenham calculado os valores necessários para a utilização do método de passo múltiplo.
- ▶ Em alternativa, poderá usar-se um método de passo único no primeiro passo, um método de passo duplo no segundo passo, etc.
- ▶ É importante que os valores iniciais obtidos desta maneira sejam da mesma ordem de precisão daqueles que irão ser obtidos pelo método final. Se os métodos iniciais forem de ordem mais baixa do que o método de passo múltiplo que queremos usar, será necessário utilizá-los com um passo mais pequeno, gerando resultados intermédios.

De modo análogo se deduz o **método de Adams-Bashforth de 3ª ordem** (caso $m = 2$), dado por:

$$y_{k+1} = y_k + \frac{h}{12}(23f_k - 16f_{k-1} + 5f_{k-2}); \quad k \geq 3,$$

e o **método de Adams-Bashforth de 4ª ordem** (caso $m = 3$), dado por:

$$y_{k+1} = y_k + \frac{h}{24}(55f_k - 59f_{k-1} + 37f_{k-2} - 9f_{k-3}); \quad k \geq 4.$$

Nota: O caso $m = 0$, ou seja, o caso em que $p = p_0$ é o polinómio constante $p_0(x) = f_k$, conduziria ao método de Euler, o qual, no entanto, não é um método de passo múltiplo.

Métodos de Adams-Moulton

Os métodos de Adams-Bashforth foram obtidos usando informação já calculada no ponto x_k e em pontos anteriores x_{k-1}, \dots, x_{k-m} . Em princípio, poder-se-á formar o polinómio interpolador usando informação em x_{k+1}, x_{k+2} etc. No caso mais simples, usar-se-ão os pontos x_{k-m}, \dots, x_k ($m \in \mathbb{N}_0$) e x_{k+1} e determinar-se-á o polinómio de grau não superior a $m + 1$ satisfazendo

$$p_{m+1}(x_i) = f_i; \quad i = k - m, \dots, k, k + 1.$$

Este processo gera uma família de métodos conhecidos por **métodos de Adams-Moulton**.

O método de Adams-Moulton para um certo valor de m tem ordem $m + 2$.

Método de Adams-Moulton de ordem 2 ($m = 0$)

Neste caso, $p = p_1$ é o polinómio linear que satisfaz $p_1(x_k) = f_k$ e $p_1(x_{k+1}) = f_{k+1}$, isto é, é o polinómio dado por

$$p_1(x) = -\frac{x - x_{k+1}}{h} f_k + \frac{x - x_k}{h} f_{k+1}.$$

Então, o método virá definido por

$$\begin{aligned} y_{k+1} &= y_k + \frac{1}{h} \int_{x_k}^{x_{k+1}} [-(x - x_{k+1})f_k + (x - x_k)f_{k+1}] dx \\ &= y_k + \frac{h}{2}(f_k + f_{k+1}); \quad k \geq 1. \end{aligned}$$

Note-se que, na fórmula anterior, se tem $f_{k+1} = f(x_{k+1}, y_{k+1})$, isto é, o método é definido por

$$y_{k+1} = y_k + \frac{h}{2} \left(f(x_k, y_k) + f(x_{k+1}, y_{k+1}) \right)$$

o que mostra que o valor de y_{k+1} está definido implicitamente. Ele é, por isso, chamado um **método implícito**.

Métodos Preditores-Corretores

Os métodos de Adams-Moulton têm a grande desvantagem de, para calcular o valor de y_{k+1} , ser necessário resolver uma equação, a qual, muitas vezes (dependendo da expressão de f) é não linear. Nesse caso, em cada passo, será preciso usar um método iterativo (tipo secante, Newton, etc) para resolver uma equação, o que torna o método muito “caro”. No entanto, a qualidade da aproximação obtida com estes métodos é, em geral, muito boa.

Na prática, geralmente, os métodos implícitos são usados apenas para corrigir aproximações obtidas por fórmulas explícitas. Deste modo, obtêm-se métodos conhecidos por **métodos preditores-corretores**.

Os outros métodos de Adams-Moulton são obtidos de modo análogo, para diferentes valores de m . Por exemplo, tem-se o **método de Adams-Moulton de 3ª ordem**, correspondente ao caso $m = 1$:

$$y_{k+1} = y_k + \frac{h}{12}(5f_{k+1} + 8f_k - f_{k-1}); \quad k \geq 2$$

e o **método Adams-Moulton de 4ª ordem**, correspondente ao caso $m = 2$:

$$y_{k+1} = y_k + \frac{h}{24}(9f_{k+1} + 19f_k - 5f_{k-1} + f_{k-2}); \quad k \geq 3.$$

Note-se que todos os métodos de Adams-Moulton são métodos implícitos, contrariamente aos métodos de Adams-Bashforth, que são métodos explícitos.

Método AB4-AM4

Um método preditor-corretor frequentemente utilizado combina os métodos de Adams-Bashforth e de Adams-Moulton de 4ª ordem, do seguinte modo:

$$y_{k+1}^{(p)} = y_k + \frac{h}{24}(55f_k - 59f_{k-1} + 37f_{k-2} - 9f_{k-3})$$

$$f_{k+1}^{(p)} = f(x_{k+1}, y_{k+1}^{(p)})$$

$$y_{k+1} = y_k + \frac{h}{24}(9f_{k+1}^{(p)} + 19f_k - 5f_{k-1} + f_{k-2})$$

Note-se que este método é totalmente explícito.

- ▶ Primeiro utiliza-se o método explícito de Adams-Bashforth para *predizer* uma primeira aproximação $y_{k+1}^{(p)}$;
- ▶ esta aproximação é utilizada para obter um valor aproximado de f_{k+1} , $f_{k+1}^{(p)}$;
- ▶ esse valor é, então, usado na fórmula do método de Adams-Moulton para *corrigir* o valor de $y_{k+1}^{(p)}$ e obter y_{k+1} .