

# Cálculo Numérico e Computacional – CNC

Luiza Amalia Pinto Cantão

`luiza@sorocaba.unesp.br`



---

# Sumário

---

<b>1</b>	<b>Introdução à Teoria de Erros e Estabilidade</b>	<b>3</b>
1.1	Representação de Números . . . . .	3
1.1.1	Representação de um Número Inteiro . . . . .	4
1.1.2	Representação de um Número Real . . . . .	4
1.2	Erros . . . . .	7
1.2.1	Representação de Números no Sistema $F(\beta, t, m, M)$ . . . . .	8
1.3	Operações Aritméticas em Ponto Flutuante . . . . .	9
1.4	Exercícios . . . . .	10
<b>2</b>	<b>Sistemas de Equações Lineares</b>	<b>13</b>
2.1	Métodos Diretos . . . . .	14
2.1.1	Sistemas Triangulares . . . . .	14
2.1.2	Método de Eliminação de Gauss . . . . .	15
2.1.3	Fatoração <b>LU</b> . . . . .	18
2.1.4	Cálculo da Matriz Inversa . . . . .	21
2.2	Métodos Iterativos . . . . .	22
2.2.1	Método de Jacobi . . . . .	23
2.2.2	Método de Gauss-Seidel . . . . .	23
2.2.3	Convergência dos Métodos de Jacobi e Gauss-Seidel . . . . .	24
2.3	Exercícios . . . . .	25
<b>3</b>	<b>Equações Não-Lineares</b>	<b>28</b>
3.1	<b>Fase I:</b> Isolamento das Raízes . . . . .	29
3.2	<b>Fase II:</b> Refinamento . . . . .	31
3.2.1	Critério de Parada . . . . .	31
3.2.2	Método da Bissecção . . . . .	31
3.2.3	Método de Newton . . . . .	32
3.2.4	Método da Secante . . . . .	35
3.3	Sistemas Não-Lineares . . . . .	36
3.3.1	Método de Newton para Sistemas Não-Lineares . . . . .	37
3.4	Exercícios . . . . .	40
<b>4</b>	<b>Ajuste de Curvas</b>	<b>42</b>
4.1	Interpolação . . . . .	42
4.1.1	Polinômio Interpolador de Lagrange . . . . .	43
4.1.2	Erro na Interpolação . . . . .	46

4.1.3	Polinômio Interpolador de Newton . . . . .	47
4.2	Quadrados Mínimos Lineares . . . . .	50
4.3	Interpolação com Splines . . . . .	53
4.3.1	Interpolação por Spline Linear . . . . .	53
4.3.2	Interpolação por Spline Cúbico . . . . .	54
4.4	Exercícios . . . . .	59
<b>5</b>	<b>Integração Numérica</b>	<b>63</b>
5.1	Fórmula de Newton-Cotes . . . . .	63
5.1.1	Fórmula dos Trapézios: $n = 1$ . . . . .	64
5.1.2	Fórmula de Simpson: $n = 2$ . . . . .	64
5.1.3	Fórmulas de Newton-Cotes para $n = 3$ e $n = 4$ . . . . .	65
5.2	Fórmulas Repetidas . . . . .	66
5.3	Erro nas Fórmulas de Newton-Cotes . . . . .	66
5.3.1	Erro na Fórmula do Trapézio . . . . .	67
5.3.2	Erro na Fórmula de Simpson . . . . .	68
5.4	Integração de Romberg . . . . .	69
5.4.1	Extrapolação de Richardson . . . . .	69
5.4.2	Integração de Romberg . . . . .	70
5.5	Quadratura de Gauss-Legendre . . . . .	71
5.5.1	Fórmula para dois pontos . . . . .	71
5.6	Exercícios . . . . .	73
<b>6</b>	<b>Aproximações para Equações Diferenciais Ordinárias</b>	<b>76</b>
6.1	Diferenças Finitas . . . . .	77
6.2	Métodos de Runge-Kutta . . . . .	80
6.2.1	Método de Runge-Kutta de ordem 2 . . . . .	81
6.2.2	Método de Runge-Kutta de ordem 3 . . . . .	83
6.2.3	Método de Runge-Kutta de ordem 4 . . . . .	84
6.3	Exercícios . . . . .	85
	<b>Referências Bibliográficas</b>	<b>87</b>

---

## Introdução à Teoria de Erros e Estabilidade

---

### Introdução

Segundo [4], o conjunto de números representáveis em qualquer máquina é finito, portanto discreto, ou seja, não é possível representar em uma máquina todos os números de um dado intervalo  $[a, b]$ . A implicação imediata desse fato é que o resultado de uma simples operação aritmética ou o cálculo de uma função, realizadas neste conjunto de números, podem conter erros. A menos que medidas apropriadas sejam tomadas, essas imprecisões causadas, por exemplo, por

1. simplificações no modelo matemático, necessárias para se obter um modelo matemático solúvel;
2. erros de truncamento (troca de uma série infinita por uma finita);
3. erros de arredondamento, devido à própria estrutura da máquina, ou ainda
4. erro nos dados, dados imprecisos obtidos experimentalmente, ou arredondados na entrada

podem diminuir, e algumas vezes destruir, a precisão dos resultados.

Assim, o objetivo deste capítulo é o estudo dos erros inerentes à estrutura computacional, dar subsídios para evitá-los e interpretar os resultados obtidos.

### 1.1 Representação de Números

**Exemplo 1.** Calcule a área de uma circunferência de raio igual a 100 m.

*Resultados Obtidos:*

1.  $A = 31400 \text{ m}^2$ ;
2.  $A = 31416 \text{ m}^2$ ;
3.  $A = 31415.92654 \text{ m}^2$ .

Como justificar as diferenças entre os resultados apresentados no exemplo 1? É possível obter exatamente esta área?

Os erros ocorridos dependem da representação do número (neste caso, do número  $\pi$ ) na máquina utilizada<sup>1</sup> e do número máximo de dígitos usados na sua representação.

O número  $\pi$ , por exemplo, não pode ser representado através de um número finito de dígitos decimais. No exemplo 1, o número  $\pi$  foi escrito como 3.14, 3.1416 e 3.141592654 respectivamente. Para cada representação foi obtido um resultado diferente, e o erro neste caso depende exclusivamente da aproximação escolhida para  $\pi$ . Qualquer que seja a circunferência, a sua área nunca será obtida exatamente de forma numérica!

Logo, qualquer cálculo que envolva números que não podem ser representados através de um número finito de dígitos não fornecerá como resultado um valor exato.

### 1.1.1 Representação de um Número Inteiro

A representação de um número inteiro não apresenta dificuldade. Qualquer computador trabalha internamente com uma base fixa  $\beta$ , onde  $\beta$  é um inteiro  $\geq 2$  e é escolhido como uma potência de 2.

Assim, dado um número inteiro  $n \neq 0$ , ele possui uma única representação:

$$n = \pm (n_{-k} n_{-k+1} \dots n_{-1} n_0) = \pm (n_0 \beta^0 + n_{-1} \beta^1 + \dots + n_{-k} \beta^k)$$

onde  $n_i$ ,  $i = 0, -1, \dots, -k$  são inteiros satisfazendo  $0 \leq n_i \leq \beta$  e  $n_{-k} \neq 0$ .

**Exemplo 2.** Na base  $\beta = 10$ , o número 1997 é representado por:

$$1997 = 7 \times 10^0 + 9 \times 10^1 + 9 \times 10^2 + 1 \times 10^3$$

e é armazenado como  $n_{-3} n_{-2} n_{-1} n_0$ .

### 1.1.2 Representação de um Número Real

A representação de um número real no computador pode ser feita por representação em *Ponto Fixo* ou *Ponto Flutuante*, como apresentado em [4].

**(i) Representação em Ponto Fixo:** Dado um número real  $x \neq 0$ , ele será representado em ponto fixo por:

$$x = \pm \sum_{i=k}^n x_i \beta^{-i},$$

onde  $k$  e  $n$  são inteiros satisfazendo  $k < n$  e, usualmente,  $k \leq 0$  e  $n > 0$  e os  $x_i$  são inteiros satisfazendo  $0 \leq x_i < \beta$ .

**Exemplo 3.** Na base 10, o número 1997.16 é representado por:

$$1997.16 = \sum_{i=-3}^2 x_i \beta^{-i} = 1 \times 10^3 + 9 \times 10^2 + 9 \times 10^1 + 7 \times 10^0 + 1 \times 10^{-1} + 6 \times 10^{-2}$$

e é armazenado como  $x_{-3} x_{-2} x_{-1} x_0 \cdot x_1 x_2$ .

**(ii) Representação em Ponto Flutuante** Esta representação é universalmente utilizada atualmente. Dado um número real  $x \neq 0$ , ele pode ser representado por:

$$x = \pm d \times \beta^e$$

<sup>1</sup>Calculadora ou computador.

onde:

- $d_i$ , para  $i = 1, 2, 3, \dots, t$ , são os dígitos da parte fracionária, tal que  $0 \leq d_i \leq 9$ . Se  $d_1 \neq 0$  diz-se que o número está normalizado.
- $t$  é o número de dígitos na mantissa.
- $e$  é um expoente inteiro.

A mantissa é um número em ponto fixo, isto é:

$$d = \sum_{i=k}^n d_i \beta^{-i}$$

onde  $k = 1$ , tal que se  $x \neq 0$ , então  $d_1 \neq 0$ ;  $0 \leq d_i < \beta$ ,  $i = 1, 2, \dots, t$ , com  $t$  a quantidade de dígitos significativos ou precisão do sistema,  $\beta^{-1} \leq d < 1$  e  $-m \leq e \leq M$ .

**Exemplo 4.** Escreva os números:

$$x_1 = 0.35 \quad x_2 = -5.172 \quad x_3 = 0.0123 \quad x_4 = 5391.3 \quad x_5 = 0.0003,$$

onde todos estão na base  $\beta = 10$ , em ponto flutuante na forma normalizada.

Temos:

$$\begin{aligned} 0.35 &= (3 \times 10^{-1} + 5 \times 10^{-2}) \times 10^0 = 0.35 \times 10^0 \\ -5.172 &= -(5 \times 10^{-1} + 1 \times 10^{-2} + 7 \times 10^{-3} + 2 \times 10^{-4}) \times 10^1 = -0.5172 \times 10^1 \\ 0.0123 &= (1 \times 10^{-1} + 2 \times 10^{-2} + 3 \times 10^{-3}) \times 10^{-1} = 0.123 \times 10^{-1} \\ 5391.3 &= (5 \times 10^{-1} + 3 \times 10^{-2} + 9 \times 10^{-3} + 1 \times 10^{-4} + 3 \times 10^{-5}) \times 10^4 = 0.53913 \times 10^4 \\ 0.0003 &= (3 \times 10^{-1}) \times 10^{-3} = 0.3 \times 10^{-3} \end{aligned}$$

**Notação:** para representar um sistema numérico em ponto flutuante normalizado, na base  $\beta$ , com  $t$  dígitos significativos e com limites dos expoentes  $m$  e  $M$ , usamos  $F(\beta, t, m, M)$ .

Assim, um número em  $F(\beta, t, m, M)$  será representado por:

$$\pm 0.d_1 d_2 \dots, d_t \times \beta^e$$

onde  $d_1 \neq 0$  e  $-m \leq e \leq M$ .

**Exemplo 5.** Considere o sistema  $F(10, 3, 2, 2)$ . Represente neste sistema os números do Exemplo 4.

Neste sistema, um número será representado por:

$$0.d_1 d_2 d_3 \times 10^e,$$

onde  $-2 \leq e \leq 2$ . Assim:

$$\begin{aligned} 0.35 &= 0.350 \times 10^0, \\ -5.172 &= -0.517 \times 10^1, \\ 0.0123 &= 0.123 \times 10^{-1} \end{aligned}$$

Note que os números 5391.3 e 0.0003 não podem ser representados no sistema, pois  $5391.3 = 0.539 \times 10^4$  e, portanto, o expoente é maior que 2, causando **overflow**. Por outro lado,  $0.0003 = 0.300 \times 10^{-3}$  e, assim, o expoente é menor do que  $-2$ , causando **underflow**.

**(iii) Mudança de Base** A maioria dos computadores trabalha na base  $\beta$  (inteiro  $\geq 2$ ), normalmente escolhido como uma potência de 2. Assim um mesmo número pode ser representado em mais de uma base. Além disso, sabemos que, através de uma mudança de base, é sempre possível determinar a representação em uma nova base.

Em particular, um número representado na base  $\beta = 2$  é conhecido como um **número binário**.

**Exemplo 6.** Mudar a representação dos números:

a. 1101 da base 2 para a base 10

Neste caso, o procedimento é multiplicar cada algarismo do número na base 2 por potências crescentes de 2, da direita para a esquerda e somar todas as parcelas:

$$1101 = 1 \times 2^0 + 0 \times 2^1 + 1 \times 2^2 + 1 \times 2^3 = 1 + 0 + 4 + 8 = 13$$

b. 0.110 da base 2 para a base 10

Multiplicar cada algarismo do número na base 2, após o ponto, por potências decrescentes de 2, da esquerda para a direita e somar as parcelas.

$$0.110 = 1 \times 2^{-1} + 1 \times 2^{-2} + 0 \times 2^{-3} = \frac{1}{2} + \frac{1}{4} + 0 = 0.75$$

c. 13 da base 10 para a base 2

O procedimento é dividir o número por 2. A seguir, continuar dividindo o quociente por 2, até que o último quociente seja igual a 1. O número na base 2 será então obtido tomando-se o último quociente e todos os restos das divisões anteriores.

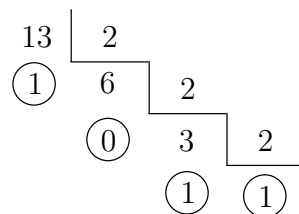


Figura 1.1:  $(13)_{10} = (1101)_2$ .

d. 0.75 da base 10 para a base 2;

Multiplique a parcela decimal por 2. Continue multiplicando a parte decimal do resultado obtido por 2. O número na base 2 será então obtido tomando-se a parte inteira do resultado de cada multiplicação. Assim,

$$0.75 \times 2 = \mathbf{1.50}$$

$$0.50 \times 2 = \mathbf{1.00}$$

$$0.00 \times 2 = \mathbf{0.00}$$

Logo:  $(0.75)_{10} = (0.110)_2$ .

e. 3.8 da base 10 para a base 2.

O procedimento é transformar a parte inteira seguindo o item “c”, o que nos fornece  $(3)_{10} = (11)_2$ , e parte decimal seguindo o item “d”.

$$\begin{aligned}
 0.8 \times 2 &= \mathbf{1.6} \\
 0.6 \times 2 &= \mathbf{1.2} \\
 0.2 \times 2 &= \mathbf{0.4} \\
 0.4 \times 2 &= \mathbf{0.8} \\
 0.8 \times 2 &= \dots
 \end{aligned}$$

Logo:  $(3.8)_{10} = (11.11001100\dots)_2$ . Portanto, o número  $(3.8)_{10}$  não tem representação exata na base 2. Este exemplo ilustra também o caso de erro de arredondamento nos dados, que será abordado na seção 1.2.

No exemplo 6, mudamos a representação de números na base 10 para a base 2 e vice-versa. O mesmo procedimento pode ser utilizado para mudar da base 10 para outra base qualquer e vice-versa. A pergunta que surge é: qual o procedimento para representar um número que está numa dada base  $\beta_1$  em uma outra base  $\beta_2$ , onde  $\beta_1 \neq \beta_2 \neq 10$ ? Neste caso, devemos representar o número que está na base  $\beta_1$ , na base 10 e, a seguir, o número obtido na base 10, na base  $\beta_2$ .

**Exemplo 7.** Dado o número  $12.20$  que está na base 4, representá-lo na base 3.

1. Representar o  $(12.20)_4$  na base 10:

$$\begin{aligned}
 12 &= 2 \times 4^0 + 1 \times 4^1 = 6, \\
 0.20 &= 2 \times 4^{-1} + 0 \times 4^{-2} = \frac{2}{4} = 0.5
 \end{aligned}$$

Portanto:  $(12.20)_4 = (6.5)_{10}$ .

2. Representar  $(6.5)_{10}$  na base 3:

$$\begin{array}{r|l}
 6 & 3 \\
 \hline
 \textcircled{0} & \textcircled{2}
 \end{array}$$

Figura 1.2:  $(6)_{10} = (20)_3$ .

$$\begin{aligned}
 0.5 \times 3 &= \mathbf{1.5} \\
 0.5 \times 3 &= \mathbf{1.5} \\
 &\vdots
 \end{aligned}$$

Assim:  $(12.20)_4 = (20.111\dots)_3$ . Observe que o número dado na base 4 tem representação exata na base 4, mas não na base 3.

## 1.2 Erros

O formato de um número em aritmética de ponto flutuante limita a mantissa em  $k$  dígitos decimais. Existem duas maneiras de obter essa limitação. Um método, chamado de truncamento, consiste em simplesmente cortar os dígitos  $d_{k+1}d_{k+2}\dots$ .

O outro método, chamado de arredondamento trunca a mantissa em  $k$  dígitos (como no caso acima), porém duas situações podem ocorrer:

1. Se  $d_{k+1} \geq 5$ ,  $d_k = d_k + 1$ ;



2. Se  $d_{k+1} < 5$ ,  $d_k = d_k$ .

**Exemplo 8.** Podemos escrever o número  $\pi$  na forma de aritmética de ponto flutuante com 5 dígitos usando:

1. O método de Truncamento:  $\pi = 0.31415 \times 10^1$ ;
2. O método de Arredondamento:  $\pi = 0.31416 \times 10^1$ .

Estes dois processos geram erros nos cálculos numéricos e são conhecidos como erros de truncamento e erros de arredondamento, respectivamente.

## Erros Absolutos e Relativos

O erro absoluto é a diferença entre o valor exato de um número  $x$  e seu valor aproximado  $\bar{x}$ :

$$EA_x = |x - \bar{x}|. \quad (1.1)$$

Em geral, apenas o valor  $\bar{x}$  é conhecido, e neste caso, é impossível obter o valor exato do erro absoluto. O que se faz é obter um limitante superior ou uma estimativa para o módulo do erro absoluto.

O erro relativo é definido como o erro absoluto dividido pelo valor aproximado:

$$ER_x = \frac{EA_x}{|\bar{x}|} = \frac{|x - \bar{x}|}{|\bar{x}|}, \quad \bar{x} \neq 0. \quad (1.2)$$

### 1.2.1 Representação de Números no Sistema $F(\beta, t, m, M)$

Os números reais podem ser representados por uma reta contínua. Entretanto, em ponto flutuante, podemos representar apenas pontos discretos na reta real.

**Exemplo 9.** Considere o número  $F(2, 3, 1, 2)$ . Quantos e quais números podem ser representados neste sistema?

Temos que  $\beta = 2$ , então os dígitos podem ser 0 ou 1;  $m = 1$  e  $M = 2$ , então,  $-1 \leq e \leq 2$  e  $t = 3$ . Assim, os números são da forma:

$$\pm 0.d_1 d_2 d_3 \times \beta^e$$

Logo, temos: duas possibilidades para o sinal, uma possibilidade para  $d_1$ , duas possibilidades para  $d_2$  e  $d_3$  e quatro possibilidades para  $\beta^e$ . Fazendo o produto  $2 \times 1 \times 2 \times 2 \times 4 = 32$ . Assim, neste sistema podemos representar 33 números, visto que o zero faz parte de qualquer sistema.

Para responder quais são os números, notemos que as formas da mantissa são: 0.100, 0.101, 0.110 e 0.111, e as formas de  $\beta^e$  são:  $2^{-1}$ ,  $2^0$ ,  $2^1$  e  $2^2$ . Assim, obtemos os seguintes números:

$$0.100 \times \begin{cases} 2^{-1} = (0.25)_{10} \\ 2^0 = (0.5)_{10} \\ 2^1 = (1.0)_{10} \\ 2^2 = (2.0)_{10} \end{cases}$$

desde que  $(0.100)_2 = (0.5)_{10}$ ;

$$0.101 \times \begin{cases} 2^{-1} = (0.3125)_{10} \\ 2^0 = (0.625)_{10} \\ 2^1 = (1.25)_{10} \\ 2^2 = (2.5)_{10} \end{cases}$$

desde que  $(0.101)_2 = (0.625)_{10}$ ;

$$0.110 \times \begin{cases} 2^{-1} = (0.375)_{10} \\ 2^0 = (0.75)_{10} \\ 2^1 = (1.5)_{10} \\ 2^2 = (3.0)_{10} \end{cases}$$

desde que  $(0.110)_2 = (0.75)_{10}$ ;

$$0.111 \times \begin{cases} 2^{-1} = (0.4375)_{10} \\ 2^0 = (0.875)_{10} \\ 2^1 = (1.75)_{10} \\ 2^2 = (3.5)_{10} \end{cases}$$

desde que  $(0.111)_2 = (0.875)_{10}$ .

**Exemplo 10.** Considerando o mesmo sistema do Exercício 9, represente os números:  $x_1 = 0.38$ ,  $x_2 = 5.3$  e  $x_3 = 0.15$  dados na base 10.

Fazendo os cálculos, obtemos:  $(0.38)_{10} = 0.110 \times 2^{-1}$ ,  $(5.3)_{10} = 0.101 \times 2^3$  e  $(0.15)_{10} = 0.100 \times 2^{-2}$ . Assim, apenas o primeiro número pode ser representado no sistema, pois para o segundo teremos **overflow** e, para o terceiro, **underflow**.

Todas as operações num computador são arredondadas. Considere o exemplo a seguir.

**Exemplo 11.** Calcular o quociente entre 15 e 7.

Temos três representações alternativas:

$$x_1 = \frac{15}{7}, \quad x_2 = 2\frac{1}{7}, \quad x_3 = 2.142857.$$

As representações dadas por  $x_1$  e  $x_2$  são exatas e  $x_3$  é uma aproximação do quociente. Suponha agora que só dispomos de quatro dígitos para representar o quociente entre 15 e 7. Daí,  $\frac{15}{7} = 2.142$ . Mas não seria melhor aproximarmos  $\frac{15}{7} = 2.143$ ? Sim e isto significa que o número foi arredondado.

## Arredondamento em Ponto Flutuante

**Definição 1.** [4] Arredondar um número  $x$ , por outro com um número menor de dígitos significativos, consiste em encontrar um número  $\bar{x}$ , pertencente ao sistema de numeração, tal que  $|x - \bar{x}|$  seja o menor possível.

## 1.3 Operações Aritméticas em Ponto Flutuante

Considere uma máquina qualquer e uma série de operações aritméticas. Pelo fato do arredondamento ser feito após cada operação, temos, ao contrário do que é válido para números reais, que as operações aritméticas (adição, subtração, divisão e multiplicação) não são nem associativas e nem distributivas.

Para os exemplos abaixo, considere o sistema com base  $\beta = 10$  e três dígitos significativos.

**Exemplo 12.** Efetue as operações indicadas:

a)  $(11.4 + 3.18) + 5.05$  e  $11.4 + (3.18 + 5.05)$

$$\begin{aligned}(11.4 + 3.18) + 5.05 &= 14.6 + 5.05 = 19.7 \\ 11.4 + (3.18 + 5.05) &= 11.4 + 8.23 = 19.6\end{aligned}$$

b)  $\frac{3.18 \times 11.4}{5.05}$  e  $\left(\frac{3.18}{5.05}\right) \times 11.4$

$$\begin{aligned}\frac{3.18 \times 11.4}{5.05} &= \frac{36.3}{11.4} = 7.18 \\ \left(\frac{3.18}{5.05}\right) \times 11.4 &= 3.18 \times 16.5 = 7.19\end{aligned}$$

c)  $3.18 \times (5.05 + 11.4)$  e  $3.18 \times 5.05 + 3.18 \times 11.4$

$$\begin{aligned}3.18 \times (5.05 + 11.4) &= 3.18 \times 16.5 = 52.3 \\ 3.18 \times 5.05 + 3.18 \times 11.4 &= 16.1 + 36.3 = 52.4\end{aligned}$$

## 1.4 Exercícios

1. Considere o sistema  $F(10, 4, 4, 4)$ . Represente neste sistema os números:

(a)  $x_1 = 4321.24$

(c)  $x_3 = 125.64$

(e)  $x_5 = 0.00034$

(b)  $x_2 = -0.0013523$

(d)  $x_4 = 57481.23$

2. Represente no sistema  $F(10, 3, 1, 3)$  os números do exercício acima.

3. Considere os seguintes números:  $x_1 = 34$ ,  $x_2 = 0.125$  e  $x_3 = 33.023$  que estão na base 10. Escreva-os na base 2.

4. Considere os seguintes números:  $x_1 = 110111$ ,  $x_2 = 0.01011$  e  $x_3 = 11.0101$  que estão na base 2. Escreva-os na base 10.

5. Considere os seguintes números:  $x_1 = 33$ ,  $x_2 = 0.132$  e  $x_3 = 32.013$  que estão na base 4. Escreva-os na base 5.

6. Considere o sistema  $F(3, 3, 2, 1)$ .

(a) Quantos e quais números podemos representar neste sistema?

(b) Represente no sistema os números:  $x_1 = (0.40)_{10}$  e  $x_2 = (2.8)_{10}$ .

7. Considere o sistema  $F(2, 5, 3, 1)$ . Quantos números podemos representar neste sistema?

8. Calcule o erro absoluto e o erro relativo nas aproximações de  $p$  e  $\bar{p}$ :

(a)  $p = \pi$ ,  $\bar{p} = 22/7$ ; ( $EA = 0.001264$  e  $ER = 4.025 \times 10^{-4}$ )

(b)  $p = \pi$ ,  $\bar{p} = 3.1416$ ; ( $EA = 7.346 \times 10^{-6}$  e  $ER = 2.338 \times 10^{-6}$ )

(c)  $p = e$ ,  $\bar{p} = 2.718$ ; ( $EA = 2.818 \times 10^{-4}$  e  $ER = 1.037 \times 10^{-4}$ )

(d)  $p = e^{10}$ ,  $\bar{p} = 22000$ ; ( $EA = 1.454 \times 10$  e  $ER = 1.05 \times 10^{-2}$ )

- (e)  $p = 8!$ ,  $\bar{p} = 39900$ ; ( $EA = 420$  e  $ER = 1.042 \times 10^{-2}$ )  
 (f)  $p = \sqrt{2}$ ,  $\bar{p} = 1.414$ ; ( $EA = 2.136 \times 10^{-4}$  e  $1.51 \times 10^{-4}$ )  
 (g)  $p = 9!$ ,  $\bar{p} = \sqrt{18\pi}(9/e)^9$ . ( $EA = 3.343 \times 10^3$  e  $9.213 \times 10^{-3}$ )
9. O número  $e$  pode ser definido por  $e = \sum_{n=0}^{\infty} (1/n!)$ , onde  $n! = n \cdot (n-1) \cdots 2 \cdot 1$  para  $n \neq 0$  e  $0! = 1$ . Calcule o erro absoluto e o erro relativo nas seguintes aproximações de  $e$ :

(a)  $\sum_{n=0}^5 \frac{1}{n!}$ ;

(b)  $\sum_{n=0}^{10} \frac{1}{n!}$ .

10. Seja  $f(x) = \frac{x \cos x - \sin x}{x - \sin x}$ .

- (a) Encontre o  $\lim_{x \rightarrow 0} f(x)$ . ( $-2$ )  
 (b) Utilize aritmética com arredondamento para valores de quatro dígitos para calcular  $f(0.1)$ . ( $-1.941$ )  
 (c) O valor real é  $f(0.1) = -1.99899998$ . Encontre o erro absoluto e relativo para o valor encontrado no item anterior.
11. Efetue os somatórios seguintes em uma calculadora e em um computador:

$$S = \sum_{i=1}^{30000} x_i$$

- (a) para  $x_i = 0.5$ ;  
 (b) para  $x_i = 0.11$ ;  
 (c) para  $x_i = 3.8$ , para  $i = 1 : 10000$ .
12. Os primeiros três termos diferentes de zeros da série de MacLaurin para a função arcotangente são  $x - (1/3)x^3 + (1/5)x^5$ . Calcule o erro absoluto e o erro relativo para as seguintes aproximações de  $\pi$  utilizando o polinômio em lugar da função arcotangente:
- (a)  $4 \left[ \arctan\left(\frac{1}{2}\right) + \arctan\left(\frac{1}{3}\right) \right]$  (Aprox.  $3.14557613$ ,  $EA = 3.983 \times 10^{-3}$  e  $ER = 1.268 \times 10^{-3}$ )  
 (b)  $16 \arctan\left(\frac{1}{5}\right) - 4 \arctan\left(\frac{1}{239}\right)$  (Aprox.  $3.14162103$ ,  $EA = 2.838 \times 10^{-5}$  e  $ER = 9.032 \times 10^{-6}$ )
13. Use a aritmética com números de três dígitos para executar os cálculos a seguir. Calcule os erros absolutos e relativos comparando-os com o valor exato determinado com pelo menos cinco dígitos.

- (a)  $133 + 0.921$  (Aprox.  $134$ ,  $EA = 0.079$  e  $ER = 5.9 \times 10^{-4}$ )  
 (b)  $133 - 0.499$  (Aprox.  $133$ ,  $EA = 0.499$  e  $ER = 3.77 \times 10^{-3}$ )  
 (c)  $(121 - 0.327) - 119$  (Aprox.  $2.00$ ,  $EA = 0.327$  e  $ER = 0.195$ )  
 (d)  $(121 - 119) - 0.327$  (Aprox.  $1.67$ ,  $EA = 0.003$  e  $ER = 1.79 \times 10^{-3}$ )  
 (e)  $\frac{\frac{13}{14} - \frac{6}{7}}{2e - 5.4}$  (Aprox.  $1.954$ ,  $EA = 0.00046$  e  $ER = 0.0002353$ )  
 (f)  $-10\pi + 6e - \frac{3}{62}$  (Aprox.  $-15.1$ ,  $EA = 0.0546$  e  $ER = 3.6 \times 10^{-3}$ )  
 (g)  $\left(\frac{2}{9}\right) \cdot \left(\frac{9}{7}\right)$  (Aprox.  $0.286$ ,  $EA = 2.86 \times 10^{-4}$  e  $ER = 10^{-3}$ )

(h)  $\frac{\pi - \frac{22}{7}}{\frac{1}{17}}$  (Aprox. 0.00,  $EA = 0.0215$  e  $ER = 1.00$ )

14. A fórmula quadrática estabelece que as raízes da equação  $ax^2 + bx + c = 0$ , quando  $a \neq 0$ , são:

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{e} \quad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

Considere a equação  $x^2 + 62.1x + 1 = 0$ , cujas raízes são aproximadamente  $x_1 = -0.01610723$  e  $x_2 = -62.0839$ .

Cálculo a equação acima utilizando arredondamento para quatro dígitos e posteriormente avalie o erro absoluto e relativo para cada raiz.

15. Utilize a aritmética com arredondamento para quatro dígitos e as fórmulas do exercício acima para encontrar os valores aproximados mais precisos para as raízes das equações quadráticas a seguir. Calcule os erros absolutos e relativos.

(a)  $\frac{1}{3}x^2 - \frac{123}{4}x + \frac{1}{6} = 0$

(b)  $\frac{1}{3}x^2 + \frac{123}{4}x - \frac{1}{6} = 0$

(c)  $1.002x^2 - 11.01x + 0.01265 = 0$

(d)  $1.002x^2 + 11.01x + 0.01265 = 0$

Questão	$x_1$	$EA$	$ER$	$x_2$	$EA$	$ER$
7	-0.02		$2.4 \times 10^{-1}$	-62.1		$3.2 \times 10^{-4}$
8 (a)	92.26	0.1542	$1.672 \times 10^{-4}$	0.005419	$6.273 \times 10^{-7}$	$1.157 \times 10^{-4}$
(b)	0.005421	$1.264 \times 10^{-6}$	$2.333 \times 10^{-4}$	$4.58 \times 10^{-3}$	$4.58 \times 10^{-3}$	$4.965 \times 10^{-5}$
(c)	10.98	$6.875 \times 10^{-3}$	$6.257 \times 10^{-4}$	$7.566 \times 10^{-8}$	$7.566 \times 10^{-8}$	$6.584 \times 10^{-5}$
(d)	-0.001149	$7.566 \times 10^{-8}$	$6.584 \times 10^{-5}$	$6.875 \times 10^{-3}$	$6.875 \times 10^{-3}$	$6.257 \times 10^{-4}$

16. Considere o sistema  $F(10, 3, 5, 5)$ . Efetue as operações indicadas:

(a)  $(1.386 - 0.987) + 7.6485$  e  $1.386 - (0.987 + 7.6485)$ ;

(b)  $\frac{1.338 - 2.038}{4.577}$  e  $\left(\frac{1.338}{4.577}\right) - \left(\frac{2.038}{4.577}\right)$ .

17. Seja  $x = \frac{17.678}{3.471} + \frac{(9.617)^2}{3.716 \times 1.85}$

(a) Calcule  $x$  com todos os algarismos da sua calculadora, sem efetuar arredondamentos.

(b) Calcule  $x$  considerando o sistema  $F(10, 3, 4, 3)$ . Faça arredondamento a cada operação efetuada.

18. Seja  $P(x) = 2.3x^3 - 0.5x^2 + 1.8x - 2.2$ . Deseja-se obter o valor de  $P(x)$  para  $x = 1.61$ .

(a) Calcule  $P(1.61)$  com todos os algarismos da sua calculadora, sem efetuar arredondamentos.

(b) Calcule  $P(1.61)$  considerando o sistema  $F(10, 3, 4, 3)$ . Faça arredondamento a cada operação efetuada.

# Sistemas de Equações Lineares

# Introdução

A solução de um sistema de equações lineares é provavelmente o processo numérico mais utilizado para simular situações do mundo real. É uma etapa fundamental na resolução de vários problemas que envolvam, por exemplo, equações diferenciais, otimização, regressão e sistemas não-lineares. Portanto, é de extrema importância que se tenha uma implementação eficiente do método para solução do sistema linear, pois geralmente esta é a fase que demanda a maior parte do tempo de processamento para resolver o problema.

Veremos aqui técnicas diretas e iterativas para resolver o sistema linear:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \quad \quad \quad \ddots \quad \quad \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases} \quad (2.1)$$

para  $x_1, x_2, \dots, x_n$ , dadas as constantes  $a_{ij}$  para cada  $i, j = 1, 2, \dots, n$  e  $b_i$  para cada  $i = 1, 2, \dots, n$ .

As técnicas diretas são métodos que dão uma resposta em um número finito de passos, sujeitos apenas aos erros de arredondamento. As técnicas iterativas geram, a partir de uma solução inicial, uma sequência de soluções que deve convergir para a solução do sistema.

Uma outra maneira de escrever o sistema (2.1) é usando a forma matricial, denotada por  $\mathbf{Ax} = \mathbf{b}$  e genericamente apresentada como:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad (2.2)$$

Note que  $\mathbf{A}_{n \times n}$  denota a matriz de coeficientes,  $\mathbf{x}$  o vetor das incógnitas e  $\mathbf{b}$  o vetor com os valores do lado direito do sistema (2.1).

Se admitirmos que  $\mathbf{A}$  é uma matriz inversível, ou seja,  $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ , onde  $\mathbf{I}$  é a matriz Identidade, então o sistema (2.1) ou (2.2) tem solução única  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ . Porém, calcular explicitamente  $\mathbf{A}^{-1}$  e em seguida  $\mathbf{A}^{-1}\mathbf{b}$  é desaconselhável, uma vez que o número de operações envolvidas é grande, o que torna este processo não competitivo com os métodos que estudaremos aqui.

## 2.1 Métodos Diretos

Apresentaremos inicialmente a resolução de Sistemas Triangulares Superior e Inferior, usados nos Métodos de Eliminação de Gauss e Fatoração **LU**.

### 2.1.1 Sistemas Triangulares

Considere os seguintes sistemas lineares:

$$\begin{cases} 2x_1 - x_2 + x_3 = 2 \\ \phantom{2}x_2 + 2x_3 = 3 \\ \phantom{2}x_3 = 1 \end{cases} \quad (2.3a)$$

$$\begin{cases} x_1 = 2 \\ 2x_1 + x_2 = 3 \\ 2x_1 - x_2 + 5x_3 = 2 \end{cases} \quad (2.3b)$$

cujas respectivas soluções podem ser obtidas diretamente, ou seja, para os problemas (2.3a) e (2.3b), temos:

$$\begin{cases} x_3 = 1 \\ x_2 = \frac{3-2}{1} = 1 \\ x_1 = \frac{2+1-1}{2} = 1 \end{cases} \quad \begin{cases} x_1 = 2 \\ x_2 = \frac{3-4}{1} = -1 \\ x_3 = \frac{2-4-1}{5} = -\frac{3}{5} \end{cases}$$

Os problemas (2.3a) e (2.3b) podem ser generalizados como seguem:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1,n-1}x_{n-1} + a_{1n}x_n = b_1 \\ \phantom{a_{11}x_1} a_{22}x_2 + \cdots + a_{2,n-1}x_{n-1} + a_{2n}x_n = b_2 \\ \phantom{a_{11}x_1} \vdots \\ \phantom{a_{11}x_1} a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1} \\ \phantom{a_{11}x_1} a_{nn}x_n = b_n \end{cases} \quad (2.4a)$$

$$\begin{cases} a_{11}x_1 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \\ \phantom{a_{21}x_1} \vdots \\ a_{n-1,1}x_1 + a_{n-1,2}x_2 + \cdots + a_{n-1,n-1}x_{n-1} = b_{n-1} \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{n,n-1}x_{n-1} + a_{nn}x_n = b_n \end{cases} \quad (2.4b)$$

Os Algoritmos (1) e (2) apresentam os procedimentos de resolução para sistemas nas formas (2.4a) e (2.4b), que são generalizações para os problemas (2.3a) e (2.3b), respectivamente.

**Dado**  $n$ ,  $\mathbf{A}_{n \times n}$  (matriz triangular superior),  $\mathbf{b}_{n \times 1}$  e  $\mathbf{x}_{n \times 1}$

- 1: **Faça**  $x_n = \frac{b_n}{a_{nn}}$
- 2: **Para**  $k = n - 1$  **até**  $k = 1$  **faça**
- 3:      $soma = b_k$
- 4:     **Para**  $j = k + 1$  **até**  $j = n$  **faça**
- 5:          $soma = soma - a_{kj}x_j$
- 6:     **Fim do laço**
- 7:      $x_k = \frac{soma}{a_{kk}}$
- 8: **Fim do laço**

Algoritmo 1: Solução de Sistemas Triangulares Superiores

**Dado**  $n$ ,  $\mathbf{A}_{n \times n}$  (matriz triangular inferior),  $\mathbf{b}_{n \times 1}$  e  $\mathbf{x}_{n \times 1}$ .

- 1: **Faça**  $x_1 = \frac{b_1}{a_{11}}$
- 2: **Para**  $k = 2$  **até**  $k = n$  **faça**
- 3:      $soma = b_k$
- 4:     **Para**  $j = 1$  **até**  $j = n - 1$  **faça**
- 5:          $soma = soma - a_{kj}x_j$
- 6:     **Fim do laço**
- 7:      $x_k = \frac{soma}{a_{kk}}$
- 8: **Fim do laço**

Algoritmo 2: Solução de Sistemas Triangulares Inferior

### 2.1.2 Método de Eliminação de Gauss

Os métodos diretos mais comuns têm como base as seguintes propriedades elementares de sistemas de equações lineares.

**Propriedade 1.** A solução do sistema  $\mathbf{Ax} = \mathbf{b}$  não se altera se o submetermos a uma seqüência de operações do tipo:

1. Multiplicação de uma equação por uma constante não-nula;
2. Soma do múltiplo de uma equação à outra;
3. Troca da ordem das equações.

Estas operações geram um sistema  $\tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{b}}$  equivalente ao sistema original  $\mathbf{Ax} = \mathbf{b}$

O método de Eliminação de Gauss usa esta propriedade para transformar a matriz  $\mathbf{A}$  numa matriz triangular superior equivalente. Suponha aqui, que  $\det(\mathbf{A}) \neq 0$ .

Reescrevemos a matriz  $\mathbf{A}$  e o vetor  $\mathbf{b}$  na forma de uma matriz expandida:



$$\left[ \begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} & b_i \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} & b_n \end{array} \right] \quad \begin{array}{c} l_1^{(1)} \\ l_2^{(1)} \\ \vdots \\ l_i^{(1)} \\ \vdots \\ l_n^{(1)} \end{array}$$

os elementos  $l_i^{(1)}$ , para  $i = 1, 2, \dots, n$ , representam as equações do sistema linear (2.1) a ser triangularizado.

### Eliminação da Primeira Coluna

Suponha que  $a_{11} \neq 0$ . Para eliminar a incógnita  $x_1$  das  $n - 1$  equações, subtraímos a primeira linha multiplicada pelo fator

$$m_{i1} = \frac{a_{i1}}{a_{11}}$$

de todas as outras linhas  $l_i$ ,  $i = 2, 3, \dots, n$

Dessa maneira,  $l_i^{(2)} = l_i^{(1)} - m_{i1}l_1^{(1)}$ , para  $i = 2, 3, \dots, n$ , ou ainda,

$$\begin{aligned} \text{Para } i = 2 : n \quad a_{ij}^{(2)} &= a_{ij} - m_{i1}a_{1j}, \quad j = 2 : n \\ b_i^{(2)} &= b_i - m_{i1}b_1. \end{aligned}$$

O índice superior  $^{(2)}$  indica que usaremos um segundo valor para  $a_{ij}$  e  $b_i$ .

No final deste estágio, os coeficientes da matriz aumentada foram modificados de modo que a matriz assume a seguinte configuração:

$$\left[ \begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{i2}^{(2)} & a_{i3}^{(2)} & \cdots & a_{in}^{(2)} & b_i^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array} \right] \quad \begin{array}{c} l_1^{(1)} \\ l_2^{(2)} \\ \vdots \\ l_i^{(2)} \\ \vdots \\ l_n^{(2)} \end{array}$$

### Eliminação da Segunda Coluna

Para eliminar a incógnita  $x_2$  das  $n - 2$  últimas equações repetimos o procedimento anterior tomando agora a segunda linha como auxiliar no processo de eliminação, isto é:

$$l_i^{(3)} = l_i^{(2)} - m_{i2}l_2^{(2)}, \quad i = 3 : n, \quad \text{onde} \quad m_{i2} = \frac{a_{i2}}{a_{22}}, \quad i = 3 : n,$$

supondo que  $a_{22} \neq 0$ .

Os coeficientes serão modificados segundo as relações:

$$\begin{aligned} \text{Para } i = 3 : n \quad a_{ij}^{(3)} &= a_{ij}^{(2)} - m_{i2}a_{2j}^{(2)}, \quad j = 3 : n \\ b_i^{(3)} &= b_i^{(2)} - m_{i2}b_2^{(2)}. \end{aligned}$$

Seguindo raciocínio análogo, procede-se até  $i = n$  e a matriz resultante será:

$$\left[ \begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} & b_n^{(n)} \end{array} \right] \Longleftrightarrow \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{22}^{(2)}x_2 + \cdots + a_{2n}^{(2)}x_n = b_2^{(2)} \\ \vdots \\ a_{nn}^{(n)}x_n = b_n^{(n)} \end{cases}$$

No processo de eliminação, os elementos  $a_{11}, a_{22}^{(2)}, a_{33}^{(3)}, \dots, a_{nn}^{(n)}$  que aparecem na diagonal principal da matriz  $\mathbf{A}$  são chamados pivôs e os elementos  $m_{ij}$ , para  $i = 1, 2, \dots, n$  e  $k = i + 1, \dots, n$ , os multiplicadores.

Se no processo de eliminação um dos pivôs se anular, devemos trocar linhas (sempre escolhendo aquelas abaixo da diagonal para não perder a eliminação anterior), de modo que escolhamos elementos não nulos para pivôs.

Dado  $n$ ,  $\mathbf{A}_{n \times n}$ ,  $\mathbf{b}_{n \times 1}$  e  $\mathbf{x}_{n \times 1}$ .

- 1: **Para**  $k = 1$  **até**  $k = n - 1$  **faça**
- 2:     **Selecione**  $i \geq k$  tal que  $a_{ik} \neq 0$
- 3:     **Se**  $a_{ii} = 0$  para todo  $i \geq k$  **então**
- 4:          $\mathbf{A}$  não é inversível. **PARE**
- 5:     **Caso contrário**
- 6:         **Se**  $i \neq k$  **então**
- 7:             **Troque** a linha  $k$  com a linha  $i$
- 8:         **Fim do condicional**
- 9:     **Fim do condicional**
- 10:    **Para**  $i = k + 1$  **até**  $i = n$  **faça**
- 11:          $m = m_{ik} = \frac{a_{ik}}{a_{kk}}$
- 12:          $b_i = b_i - mb_k$
- 13:         **Para**  $j = k + 1$  **até**  $j = n$  **faça**
- 14:              $a_{ij} = a_{ij} - ma_{kj}$
- 15:         **Fim do laço**
- 16:     **Fim do laço**
- 17: **Fim do laço**
- 18: Execute o algoritmo (1).

Algoritmo 3: Solução de (2.2) via Eliminação de Gauss

## Estratégia de Pivoteamento

**Exemplo 13.** Resolva o sistema abaixo usando o algoritmo (3):

$$\begin{aligned} 0.004x_1 + 15.73x_2 &= 15.77 \\ 0.423x_1 - 24.72x_2 &= -20.49 \end{aligned}$$

Use quatro dígitos na representação em ponto flutuante e arredondamento ao desprezar o quinto dígito.  
Troque a ordem das equações lineares e resolva novamente o problema usando o mesmo algoritmo.

**Solução 1.** Para eliminarmos  $x_1$ , obtemos  $m = 105.8$  e:

$$\begin{array}{rcl} 0.004x_1 & + & 15.73x_2 = 15.77 \\ & - & 1689x_2 = -1688 \end{array}$$

Cuja solução obtida é  $x_1 = 15$  e  $0.999$ . A solução correta do sistema é  $x_1 = 10$  e  $x_2 = 1.0$ . Calcule o erro relativo da solução obtida.

Resolvendo o mesmo problema com a ordem das equações trocadas temos  $m = 0.9456 \times 10^{-2}$  e:

$$\begin{array}{rcl} 0.423x_1 & - & 24.72x_2 = -20.49 \\ & & 15.96x_2 = 15.96 \end{array}$$

cuja solução é  $x_1 = 10$  e  $x_2 = 1$ , que é a solução exata do sistema.

O procedimento usado no exemplo 13 para obter a solução correta do sistema é chamado de estratégia de pivoteamento, que consiste na troca sistemática das linhas, de modo que o pivô seja o maior elemento, em valor absoluto, da coluna que estamos eliminando. Assim,

1. no  $k$ -ésimo passo procuramos o elemento pivô de maior valor absoluto entre os coeficientes:

$$|a_{rk}| = \max_{i \leq k \leq n} |a_{ik}|;$$

2. trocamos as linhas  $k$  e  $r$  se for necessário.

O algoritmo (4) ilustra esta estratégia.

Existem dois casos nos quais o método de eliminação pode ser aplicado sem pivoteamento:

1. Uma matriz é diagonalmente dominante, ou seja, seus elementos satisfazem a  $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ ,  $i = 1, 2, \dots, n$ , para todo  $i$ . Neste caso, os cálculos serão estáveis com respeito ao crescimento dos erros de arredondamento.
2. Uma matriz é simétrica  $\mathbf{A}^T = \mathbf{A}$  e positiva definida  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ , para todo vetor  $\mathbf{x} \neq \mathbf{0}$ .

### 2.1.3 Fatoração LU

Os métodos de Eliminação de Gauss e Eliminação de Gauss com Pivoteamento podem ser usados economicamente quando precisamos resolver vários sistemas com a mesma matriz dos coeficientes  $\mathbf{A}$  e diversos termos independentes  $\mathbf{b}$ .

Uma opção seria guardar os coeficientes  $m_{ij}$  calculados no processo de eliminação e usá-los na atualização dos termos independentes  $\mathbf{b}$ . Computacionalmente, esta alternativa é conhecida como Fatoração **LU** da matriz  $\mathbf{A}$ .

Suponha que seja possível fatorar a matriz  $\mathbf{A}$  num produto de uma matriz triangular inferior (com elementos da diagonal principal iguais a 1)  $\mathbf{L}$ , e uma matriz triangular superior  $\mathbf{U}$ , isto é:

$$\mathbf{A} = \mathbf{LU} \implies \mathbf{Ax} = \mathbf{b} \iff \mathbf{LUx} = \mathbf{b}.$$

O sistema  $\mathbf{LUx} = \mathbf{b}$  permite o desmembramento em dois sistemas triangulares:

$$\mathbf{Ly} = \mathbf{b} \quad \text{e} \quad \mathbf{Ux} = \mathbf{y}.$$

**Dado**  $n$ ,  $\mathbf{A}_{n \times n}$ ,  $\mathbf{b}_{n \times 1}$  e  $\mathbf{x}_{n \times 1}$ .

- 1: **Para**  $k = 1$  **até**  $k = n - 1$  **faça**
- 2:      $w = |a_{kk}|$  e  $r = k$
- 3:     **Para**  $j = k$  **até**  $j = n$  **faça**
- 4:         **Se**  $|a_{kj}| > w$  **então**
- 5:              $w = |a_{kj}|$  e  $r = j$
- 6:         **Fim do condicional**
- 7:     **Fim do laço**
- 8:     **Se**  $w = 0$  **então**
- 9:          $\mathbf{A}$  não é inversível. **PARE**
- 10:    **Caso contrário**
- 11:        **Troque** a linha  $k$  com a linha  $r$
- 12:    **Fim do condicional**
- 13:    **Para**  $i = k + 1$  **até**  $i = n$  **faça**
- 14:         $m = m_{ik} = \frac{a_{ik}}{a_{kk}}$
- 15:         $b_i = b_i - mb_k$
- 16:        **Para**  $j = k + 1$  **até**  $j = n$  **faça**
- 17:             $a_{ij} = a_{ij} - ma_{kj}$
- 18:        **Fim do laço**
- 19:    **Fim do laço**
- 20: **Fim do laço**
- 21: Execute o algoritmo (1).

Algoritmo 4: Solução de (2.2) via Eliminação de Gauss com Pivoteamento

Resolvendo o primeiro sistema, calculamos  $\mathbf{y}$  que, usado no segundo sistema, fornecerá o vetor procurado  $\mathbf{x}$ .

**Teorema 1.** Dada uma matriz  $\mathbf{A}_{n \times n}$ , seja  $\mathbf{A}_k$  a matriz constituída das primeiras  $k$  linhas e colunas de  $\mathbf{A}$ . Suponha que  $\det(\mathbf{A}_k) \neq 0$  para  $k = 1, 2, \dots, (n - 1)$ . Então, existe uma única matriz triangular inferior  $\mathbf{L} = (m_{ij})$ , com  $m_{ij} = 1$ ,  $1 \leq i \leq n$  e uma matriz triangular superior  $\mathbf{U} = (u_{ij})$  tais que  $\mathbf{LU} = \mathbf{A}$ . Ainda mais,  $\det(\mathbf{A}) = u_{11}u_{22} \cdots u_{nn}$ .

**Exemplo 14.** Resolva o sistema linear a seguir usando a fatoração  $\mathbf{LU}$ :

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 = 1 \\ x_1 + x_2 + 2x_3 = 2 \\ 4x_1 + 3x_2 + 2x_3 = 3 \end{cases}$$

Dado  $n$ ,  $\mathbf{A}_{n \times n}$ ,  $\mathbf{b}_{n \times 1}$  e  $\mathbf{x}_{n \times 1}$ .

- 1: Para  $i = 1$  até  $i = n$  faça
- 2:     Para  $j = i$  até  $j = n$  faça
- 3:          $u_{ij} = a_{ij} - \sum_{k=1}^{i-1} m_{ik} u_{kj}$
- 4:         Para  $j = i + 1$  até  $j = n$  faça
- 5:              $m_{ji} = \left( a_{ji} - \sum_{k=1}^{i-1} m_{jk} u_{ki} \right) / u_{ii}$
- 6:         Fim do laço
- 7:     Fim do laço
- 8: Fim do laço
- 9: Execute o algoritmo (2) para resolver  $\mathbf{L}\mathbf{y} = \mathbf{b}$  e o algoritmo (1) para resolver  $\mathbf{U}\mathbf{x} = \mathbf{y}$

Algoritmo 5: Solução de (2.2) via Fatoração  $\mathbf{LU}$

**Solução 2.** Considere a matriz abaixo onde a primeira parte refere-se a matriz  $\mathbf{L}$  e a segunda parte, a matriz  $\mathbf{U}$ :

$$\left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 2 & 4 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 4 & 3 & 2 \end{array} \right] \xrightarrow{1^a \text{ Etapa}} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 2 & 4 \\ 1/3 & 1 & 0 & 0 & 1/3 & 2/3 \\ 4/3 & 0 & 1 & 0 & 1/3 & -10/3 \end{array} \right] \xrightarrow{2^a \text{ Etapa}} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 2 & 4 \\ 1/3 & 1 & 0 & 0 & 1/3 & 2/3 \\ 4/3 & 1 & 1 & 0 & 0 & -4 \end{array} \right]$$

$1^a \text{ Etapa}$

Pivô:  $a_{11} = u_{11} = 3$  e multiplicadores:  $m_{21} = \frac{a_{21}}{u_{11}} = \frac{1}{3}$  e  $m_{31} = \frac{a_{31}}{u_{11}} = \frac{4}{3}$ .

$$\begin{aligned} u_{2j} &= a_{2j} - m_{21} a_{1j} & j &= 2, 3 \\ u_{3j} &= a_{3j} - m_{31} a_{1j} & j &= 2, 3 \end{aligned}$$

$2^a \text{ Etapa}$

Pivô:  $u_{22} = \frac{1}{3}$  e multiplicadores:  $m_{32} = \frac{u_{32}}{u_{22}} = \frac{1/3}{1/3} = 1$ .

$$u_{33} = u_{33} - m_{32} u_{23}$$

Resolvendo  $\mathbf{L}(\mathbf{U}\mathbf{x}) = \mathbf{b}$

(i)  $\mathbf{L}\mathbf{y} = \mathbf{b}$

$$\begin{cases} y_1 & = 1 \\ 1/3 y_1 + y_2 & = 2 \\ 4/3 y_1 + y_2 + y_3 & = 3 \end{cases} \implies \mathbf{y} = \left( 1, \frac{5}{3}, 0 \right)^T$$

(ii)  $\mathbf{U}\mathbf{x} = \mathbf{y}$

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 & = 1 \\ 1/3 x_2 + 2/3 x_3 & = 5/3 \\ -4x_3 & = 0 \end{cases} \implies \mathbf{x} = (-3, 5, 0)^T$$

### 2.1.4 Cálculo da Matriz Inversa

Sejam  $\mathbf{A}$  uma matriz quadrada não singular ( $\det(\mathbf{A}) \neq 0$ ) e  $\mathbf{A}^{-1} = [\mathbf{b}_1 | \mathbf{b}_2 | \dots | \mathbf{b}_n]$  a matriz inversa de  $\mathbf{A}$ , onde  $\mathbf{b}_j$  é a  $j$ -ésima coluna da matriz  $\mathbf{A}^{-1}$  e  $\mathbf{e}_j$  é a  $j$ -ésima coluna da matriz identidade  $\mathbf{I}$ . De  $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}$ , isto é, de:

$$\mathbf{A} \cdot [\mathbf{b}_1 | \mathbf{b}_2 | \dots | \mathbf{b}_n] = [\mathbf{e}_1 | \mathbf{e}_2 | \dots | \mathbf{e}_n]$$

resulta  $\mathbf{A} \cdot \mathbf{b}_j = \mathbf{e}_j$ ,  $j = 1 : n$ .

Assim, podemos calcular as colunas  $j$ ,  $j = 1 : n$ , da matriz  $\mathbf{A}^{-1}$  resolvendo os sistemas lineares anteriores e, portanto podemos inverter uma matriz utilizando qualquer método de resolução de sistemas lineares:

- a. Usando **Fatoração LU**, obtemos as colunas de  $\mathbf{A}^{-1}$  fazendo:

$$\mathbf{LU} \cdot \mathbf{b}_j = \mathbf{e}_j, \quad j = 1 : n$$

isto é, resolvendo os sistemas lineares  $\mathbf{Ly}_j = \mathbf{e}_j$  e  $\mathbf{Ub}_j = \mathbf{y}_j$ , para  $j = 1 : n$ .

- b. Usando o método de **Eliminação de Gauss**, obtemos as colunas de  $\mathbf{A}^{-1}$  resolvendo os sistemas lineares:

$$\mathbf{A} \cdot \mathbf{b}_j = \mathbf{e}_j, \quad j = 1 : n$$

**Exemplo 15.** Considere a matriz:

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 3 \\ 2 & -2 & 1 \\ 1 & 2 & 0 \end{bmatrix}$$

- a. Calcule  $\mathbf{A}^{-1}$  usando a Fatoração **LU**;  
b. Calcule  $\mathbf{A}^{-1}$  utilizando o método de Eliminação de Gauss.

**Solução 3.**

- a. Inicialmente procedemos com a fatoração da matriz  $\mathbf{A}$ :

$$\left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 0 & 3 \\ 0 & 1 & 0 & 2 & -2 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \end{array} \right] \xrightarrow{1^a \text{ Etapa}} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 0 & 3 \\ 2/3 & 1 & 0 & 0 & -2 & -1 \\ 1/3 & 0 & 1 & 0 & 2 & -1 \end{array} \right] \xrightarrow{2^a \text{ Etapa}} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 0 & 3 \\ 2/3 & 1 & 0 & 0 & -2 & -1 \\ 1/3 & -1 & 1 & 0 & 0 & -2 \end{array} \right]$$

Agora resolvemos  $\mathbf{LU}_j = \mathbf{e}_j$ , para  $j = 1 : 3$ .

- $j = 1$ :

$$\mathbf{Ly}_1 = \mathbf{b}_1 \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \longrightarrow \mathbf{y}_1 = \begin{bmatrix} 1 \\ -\frac{2}{3} \\ -1 \end{bmatrix}$$

$$\mathbf{Ub}_1 = \mathbf{y}_1 \longrightarrow \begin{bmatrix} 3 & 0 & 3 \\ 0 & -2 & -1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{2}{3} \\ -1 \end{bmatrix} \longrightarrow \mathbf{b}_1 = \begin{bmatrix} -\frac{1}{6} \\ \frac{1}{12} \\ \frac{1}{2} \end{bmatrix}$$

O vetor  $\mathbf{b}_1$  representa a primeira coluna da matriz  $\mathbf{A}^{-1}$ .

- $j = 2$ :

$$\mathbf{L}\mathbf{y}_2 = \mathbf{b}_2 \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \longrightarrow \mathbf{y}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{U}\mathbf{b}_2 = \mathbf{y}_2 \longrightarrow \begin{bmatrix} 3 & 0 & 3 \\ 0 & -2 & -1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \longrightarrow \mathbf{b}_2 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{4} \\ -\frac{1}{2} \end{bmatrix}$$

O vetor  $\mathbf{b}_2$  representa a segunda coluna da matriz  $\mathbf{A}^{-1}$ .

- $j = 3$ :

$$\mathbf{L}\mathbf{y}_3 = \mathbf{b}_3 \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \longrightarrow \mathbf{y}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{U}\mathbf{b}_3 = \mathbf{y}_3 \longrightarrow \begin{bmatrix} 3 & 0 & 3 \\ 0 & -2 & -1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} b_{13} \\ b_{23} \\ b_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \longrightarrow \mathbf{b}_3 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ -\frac{1}{2} \end{bmatrix}$$

O vetor  $\mathbf{b}_3$  representa a terceira coluna da matriz  $\mathbf{A}^{-1}$ .

$$\text{Portanto, } \mathbf{A}^{-1} = [\mathbf{b}_1 | \mathbf{b}_2 | \mathbf{b}_3] = \begin{bmatrix} -\frac{1}{6} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

b. usando o método de Eliminação de Gauss, podemos escrever a matriz  $\mathbf{A}$  e os vetores  $\mathbf{e}_1$ ,  $\mathbf{e}_2$  e  $\mathbf{e}_3$ , na forma expandida:

$$\left[ \begin{array}{ccc|ccc} 3 & 0 & 3 & 1 & 0 & 0 \\ 2 & -2 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array} \right] \xrightarrow{1^{\text{a}} \text{ Coluna}} \left[ \begin{array}{ccc|ccc} 3 & 0 & 3 & 1 & 0 & 0 \\ 0 & -2 & 1 & -\frac{2}{3} & 1 & 0 \\ 0 & 2 & -1 & -\frac{1}{3} & 0 & 1 \end{array} \right] \xrightarrow{2^{\text{a}} \text{ Coluna}} \left[ \begin{array}{ccc|ccc} 3 & 0 & 3 & 1 & 0 & 0 \\ 0 & -2 & 1 & -\frac{2}{3} & 1 & 0 \\ 0 & 0 & -2 & -1 & 1 & 1 \end{array} \right]$$

Note que, a quarta, quinta e sexta coluna na matriz expandida representam, respectivamente, os vetores  $\mathbf{y}_1$ ,  $\mathbf{y}_2$  e  $\mathbf{y}_3$ . Resolvendo o sistema triangular superior (primeira, segunda e terceira coluna da matriz expandida) para  $\mathbf{y}_1$ ,  $\mathbf{y}_2$  e  $\mathbf{y}_3$  ( $\mathbf{A}\mathbf{e}_j = \mathbf{y}_j$ ,  $j = 1 : 3$ ) obtemos:

$$\mathbf{A}^{-1} = \begin{bmatrix} -\frac{1}{6} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

## 2.2 Métodos Iterativos

Se no sistema  $\mathbf{Ax} = \mathbf{b}$  os elementos da diagonal são diferentes de zero,  $a_{ii} \neq 0$ ,  $i = 1, 2, \dots, n$ , então podemos explicitar  $x_1$  usando a primeira equação,  $x_2$  usando a segunda equação e assim sucessivamente. Ou seja,

reescrevemos o sistema (2.1) numa forma que é conveniente para os métodos iterativos:

$$\begin{aligned} x_1 &= \left( b_1 - \sum_{j=2}^n a_{1j}x_j \right) / a_{11} \\ &\vdots \\ x_i &= \left( b_i - \sum_{j \neq i}^n a_{ij}x_j \right) / a_{ii} \\ &\vdots \\ x_n &= \left( b_n - \sum_{j=1}^{n-1} a_{nj}x_j \right) / a_{nn} \end{aligned} \quad (2.5)$$

### 2.2.1 Método de Jacobi

Neste método, as equações (2.5) são usadas para calcular uma sequência de vetores aproximações  $\mathbf{x}^{(1)}$ ,  $\mathbf{x}^{(2)}$ , ...,  $\mathbf{x}^{(k)}$ . Dada uma aproximação inicial  $\mathbf{x}^{(0)}$ , usamos o primeiro termo à direita da  $i$ -ésima equação para definir uma nova aproximação para  $x_i$ :

$$x_i^{(1)} = \left( b_i - \sum_{j \neq i}^n a_{ij}x_j^{(0)} \right) / a_{ii}, \quad i = 1, 2, \dots, n. \quad (2.6)$$

Usamos agora o vetor  $\mathbf{x}^{(1)}$  nas equações (2.5) para calcular o novo vetor das aproximações,  $\mathbf{x}^{(2)} = (x_1^{(2)}, x_2^{(2)}, \dots, x_n^{(2)})^T$ . Em resumo, o método de Jacobi consiste em calcularmos as componentes dos vetores  $\mathbf{x}^{(1)}$ ,  $\mathbf{x}^{(2)}$ , ... usando (2.5).

Verificamos se este método converge fazendo:

$$EA = \max_{1 \leq i \leq n} |x_i^{(k+1)} - x_i^{(k)}| < \epsilon \quad \text{ou} \quad ER = \max_{1 \leq i \leq n} \frac{|x_i^{(k+1)} - x_i^{(k)}|}{|x_i^{(k+1)}|} < \epsilon$$

onde  $\epsilon$  é uma tolerância suficientemente pequena. Note que, as medidas de erro acima são os erros calculados em (1.1) e (1.2), respectivamente, sendo  $x_i^{(k+1)}$  a solução procurada, isto é,  $x_i^{(k+1)} = \bar{x}$ . Em testes computacionais usamos também como teste de parada um número máximo de iterações.

Para iniciar o processo iterativo é necessário fornecer uma aproximação inicial  $\mathbf{x}^{(0)}$ . **Na falta de informação sobre a solução, tomamos  $\mathbf{x}^{(0)} = \mathbf{0}$ .** O Algoritmo 6 apresenta os passos do método de Jacobi.

### 2.2.2 Método de Gauss-Seidel

Este método consiste em uma modificação do método de Jacobi. Nele, as iterações serão calculadas usando as equações (2.5), mas aproveitando os cálculos já atualizados de outras componentes, para atualizar a componente que está sendo calculada. Assim, o valor recém calculado para  $x_1^{(k+1)}$  será usado no cálculo de  $x_2^{(k+1)}$ , como mostra o Algoritmo 7.



**Dado**  $n$ ,  $\mathbf{A}_{n \times n}$ ,  $\mathbf{b}_{n \times 1}$  e  $\mathbf{x}_{n \times 1}^{(0)}$ ,  $\max$ ,  $\epsilon$

- 1: **Para**  $k = 0$  **até**  $k = \max$  **faça**
- 2:     **Para**  $i = 1$  **até**  $i = n$  **faça**
- 3:         
$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)} \right)$$
- 4:         **Se**  $\max_{1 \leq i \leq n} |x_i^{(k+1)} - x_i^{(k)}| < \epsilon$  **ou**  $\frac{|x_i^{(k+1)} - x_i^{(k)}|}{|x_i^{(k+1)}|} < \epsilon$  **então**
- 5:              $\mathbf{x} = \mathbf{x}^{(k+1)}$
- 6:         **Caso contrário**
- 7:             **Se**  $k = \max$  **então**
- 8:                 **PARE:** não houve convergência.
- 9:             **Fim do condicional**
- 10:         **Fim do condicional**
- 11:     **Fim do laço**
- 12: **Fim do laço**

Algoritmo 6: Método de Jacobi

**Dado**  $n$ ,  $\mathbf{A}_{n \times n}$ ,  $\mathbf{b}_{n \times 1}$  e  $\mathbf{x}_{n \times 1}^{(0)}$ ,  $\max$ ,  $\epsilon$

- 1: **Para**  $k = 0$  **até**  $k = \max$  **faça**
- 2:     **Para**  $i = 1$  **até**  $i = n$  **faça**
- 3:         
$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right)$$
- 4:         **Se**  $\max_{1 \leq i \leq n} |x_i^{(k+1)} - x_i^{(k)}| < \epsilon$  **ou**  $\frac{|x_i^{(k+1)} - x_i^{(k)}|}{|x_i^{(k+1)}|} < \epsilon$  **então**
- 5:              $\mathbf{x} = \mathbf{x}^{(k+1)}$
- 6:         **Caso contrário**
- 7:             **Se**  $k = \max$  **então**
- 8:                 **PARE:** não houve convergência.
- 9:             **Fim do condicional**
- 10:         **Fim do condicional**
- 11:     **Fim do laço**
- 12: **Fim do laço**

Algoritmo 7: Método de Gauss-Seidel

**Exemplo 16.** Resolva o sistema abaixo usando os métodos de Jacobi e Gauss-Seidel:

$$\begin{cases} 4.00x_1 + 0.24x_2 - 0.08x_3 = 8.00 \\ 0.09x_1 + 3.00x_2 - 0.15x_3 = 9.00 \\ 0.04x_1 - 0.08x_2 + 4.00x_3 = 20.00 \end{cases}$$

### 2.2.3 Convergência dos Métodos de Jacobi e Gauss-Seidel

As condições de convergência de verificação simples para os métodos de Jacobi e Gauss-Seidel são:

- Os métodos iterativos de Jacobi e Gauss-Seidel convergem se a matriz  $\mathbf{A}$  é diagonalmente dominante, ou seja:  $|a_{ii}| > \sum_{j \neq i}^n |a_{ij}|$ ,  $i = 1 : n$ .
- Os métodos iterativos de Jacobi e Gauss-Seidel convergem se  $\mathbf{A}$  é uma matriz positiva definida:  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$  para todo  $\mathbf{x} \neq \mathbf{0}$ .

## 2.3 Exercícios

- Para cada um dos sistemas lineares seguintes, obtenha uma solução por meio de métodos gráficos, se possível. Explique os resultados do ponto de vista geométrico.

$$(a) \begin{cases} x_1 + 2x_2 = 3 \\ x_1 - x_2 = 0 \end{cases}$$

$$(b) \begin{cases} x_1 + x_2 = 0 \\ x_1 - x_2 = 0 \end{cases}$$

$$(c) \begin{cases} x_1 + 2x_2 = 3 \\ 2x_1 + 4x_2 = 6 \end{cases}$$

$$(d) \begin{cases} x_1 + x_2 = 3 \\ -2x_1 - 4x_2 = 6 \end{cases}$$

$$(e) \begin{cases} x_1 + 2x_2 = 0 \\ 2x_1 + 4x_2 = 0 \end{cases}$$

$$(f) \begin{cases} 2x_1 + x_2 = -1 \\ x_1 + x_2 = 2 \\ x_1 - 3x_2 = 5 \end{cases}$$

$$(g) \begin{cases} 2x_1 + x_2 = -1 \\ 4x_1 + 2x_2 = -2 \\ x_1 - 3x_2 = 5 \end{cases}$$

$$(h) \begin{cases} 2x_1 + x_2 + x_3 = 1 \\ 2x_1 + 4x_2 - x_3 = -1 \end{cases}$$

- Utilize o método de Gauss usando operações com arredondamento para dois dígitos para resolver os sistemas lineares a seguir. **Não** reordene as equações. (A solução exata para cada sistema é  $x_1 = 1$ ,  $x_2 = -1$  e  $x_3 = 3$ .)

(a)

$$\begin{aligned} 4x_1 - x_2 + x_3 &= 8 \\ 2x_1 + 5x_2 + 2x_3 &= 3 \\ x_1 + 2x_2 + 4x_3 &= 11 \end{aligned}$$

(b)

$$\begin{aligned} 4x_1 + x_2 + 2x_3 &= 9 \\ 2x_1 + 4x_2 - x_3 &= -5 \\ x_1 + x_2 - 3x_3 &= -9 \end{aligned}$$

- Considere os seguintes sistemas:

- (i) Solução real:  $(0, 10, 1/7)^T$

$$\begin{cases} 3.03x_1 - 12.1x_2 + 14x_3 = -119 \\ -3.03x_1 + 12.1x_2 - 7x_3 = 120 \\ 6.11x_1 - 14.2x_2 + 21x_3 = -139 \end{cases}$$

- (ii) Solução real:  $(10, 1)^T$

$$\begin{cases} 0.03x_1 + 58.9x_2 = 59.2 \\ 5.31x_1 - 6.10x_2 = 47.0 \end{cases}$$

- (iii) Solução real:  $(0.17682530, 0.01269269, -0.02065405, -1.18260870)^T$

$$\begin{cases} 1.19x_1 + 2.11x_2 - 100x_3 + x_4 = 1.12 \\ 14.2x_1 - 0.122x_2 + 12.2x_3 - x_4 = 3.44 \\ 100x_2 - 99.9x_3 + x_4 = 2.15 \\ 15.3x_1 + 0.110x_2 - 13.1x_3 - x_4 = 4.16 \end{cases}$$

- (iv) Solução real:  $(0.78839378, -3.12541367, 10.16759660, 4.55700252)^T$

$$\begin{cases} \pi x_1 - e x_2 + \sqrt{2}x_3 - \sqrt{3}x_4 = \sqrt{11} \\ \pi^2 x_1 + e x_2 - e^2 x_3 + \frac{3}{7}x_4 = 0 \\ \sqrt{5}x_1 - \sqrt{6}x_2 + x_3 - \sqrt{2}x_4 = \pi \\ \pi^3 x_1 + e^2 x_2 - \sqrt{7}x_3 + \frac{1}{9}x_4 = \sqrt{2} \end{cases}$$

- (a) Resolva os sistemas acima usando o método de Eliminação de Gauss e operações aritméticas com aproximação de três dígitos por truncamento.
- (b) Repita o item acima usando o método de Eliminação de Gauss com Pivoteamento.
- (c) Resolva os sistemas acima usando Fatoração **LU** e operações aritméticas com aproximação de três dígitos por truncamento.
- (d) Calcule o erro relativo dos itens acima.
- (e) Verifique se os sistemas acima são convergente se aplicarmos os Métodos Iterativos de Jacobi e Gauss-Seidel. Justifique sua resposta.

4. Sejam os sistemas lineares:

$$(a) \begin{cases} x_1 - 0.5x_2 + 0.5x_3 = 3 \\ x_1 + x_2 + x_3 = 12 \\ -0.5x_1 - 0.5x_2 + x_3 = 3 \end{cases} \quad (b) \begin{bmatrix} 10 & 2 & -3 & 5 \\ 1 & 8 & -1 & 2 \\ 2 & -1 & -5 & 1 \\ -1 & 2 & 3 & 20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 48 \\ 4 \\ -11 \\ 150 \end{bmatrix}$$

Resolva-os usando os métodos de Jacobi e Gauss-Seidel e verifique a convergência dos métodos. Justifique os resultados. [ (a)  $\bar{x} = (2, 4, 6)^T$ , (b)  $\bar{x} = (3, -1, 5, 7)^T$  ]

5. Resolva os sistemas lineares:

$$(a) \begin{cases} x_1 - 3x_2 + 5x_3 + 6x_4 = 17 \\ -8x_1 + 4x_2 - x_3 = 29 \\ 3x_1 + 2x_2 - 2x_3 + 7x_4 = -11 \\ x_1 + 2x_2 + 5x_3 - 4x_4 = 7 \end{cases} \quad (b) \begin{bmatrix} 1 & 2 & 4 \\ -3 & -1 & 4 \\ 2 & 14 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 13 \\ 8 \\ 50 \end{bmatrix}$$

usando o método de Eliminação de Gauss com e sem pivoteamento. Compare e verifique os resultados. [ (a)  $\bar{x} = (-4.6077, 0.3757, 3.1215, 1.1878)^T$  ]

6. Efetue os cálculos, utilizando apenas 4 casas decimais:

$$(a) \begin{bmatrix} 2 & 6 & -3 \\ 1 & 3.001 & 2 \\ 4 & -1 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 9 \\ 29 \end{bmatrix} \quad (b) \begin{cases} x_1 + 2x_2 + 3x_3 = 17 \\ -5x_1 - x_2 + 4x_3 = -2 \\ 2x_1 + 4x_2 + x_3 = 25 \end{cases}$$

usando Fatoração **LU**. Verifique a unicidade e a exatidão das soluções.

7. Calcule o determinante da matriz  $A$  usando Fatoração  $LU$ .

$$A = \begin{bmatrix} 0 & 2 & 1 & 4 & -1 & 3 \\ 1 & 2 & -1 & 3 & 4 & 0 \\ 0 & 1 & 1 & -1 & 2 & -1 \\ 2 & 3 & -4 & 2 & 0 & 5 \\ 1 & 1 & 1 & 3 & 0 & 2 \\ -1 & -1 & 2 & -1 & 2 & 0 \end{bmatrix}$$

8. Resolva os sistemas abaixo usando Fatoração  $LU$ :

(a)

$$\begin{aligned} 1.012x_1 - 2.132x_2 + 3.104x_3 &= 1.984 \\ -2.132x_1 + 4.096x_2 - 7.013x_3 &= -5.049 \\ 3.104x_1 - 7.013x_2 + 0.014x_3 &= -3.895 \end{aligned}$$

(b)

$$\begin{aligned} 2.1756x_1 + 4.0231x_2 - 2.1732x_3 + 5.1967x_4 &= 17.102 \\ -4.0231x_1 + 6.0000x_2 + 1.1973x_4 &= -6.1593 \\ -1.0000x_1 - 5.2107x_2 + 1.1111x_3 &= 3.0004 \\ 6.0235x_1 + 7.0000x_2 - 4.1561x_4 &= 0.0000 \end{aligned}$$

9. Considere as matrizes abaixo:

$$(a) \mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

$$(c) \mathbf{C} = \begin{bmatrix} 2 & 4 & 6 \\ 1 & -3 & -1 \\ 2 & 1 & 1 \end{bmatrix}$$

$$(b) \mathbf{B} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 10 & 2 \\ -1 & 2 & 4 \end{bmatrix}$$

$$(d) \mathbf{D} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 2 \\ 4 & -1 & 3 \end{bmatrix}$$

Calcule as suas matrizes inversas usando:

- Fatoração **LU**;
- Método de Eliminação de Gauss.

10. Seja o sistema linear  $\mathbf{Ax} = \mathbf{b}$ , dado por:

$$\begin{pmatrix} 10 & 7 & 8 \\ 7 & 5 & 6 \\ 8 & 6 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3 \\ -1 \\ 7 \end{pmatrix}$$

- (a) Determine a inversa da matriz dos coeficientes pelo método de Eliminação de Gauss.  
 (b) Resolva o sistema linear dado utilizando a matriz inversa obtida no item (a).

## Equações Não-Lineares

### Introdução

A necessidade de encontrar valores de  $x = \bar{x}$  que satisfaçam a equação  $f(\bar{x}) = 0$  aparece frequentemente em uma ampla variedade de problemas provenientes das Ciências e das Engenharias. Estes valores especiais são chamados de *raízes da equação*  $f(x) = 0$ , ou *zeros da função*  $f$ , os quais podem ser vistos na figura 3.1, que mostra quando a função  $f(x) = \cos x = 0$ .

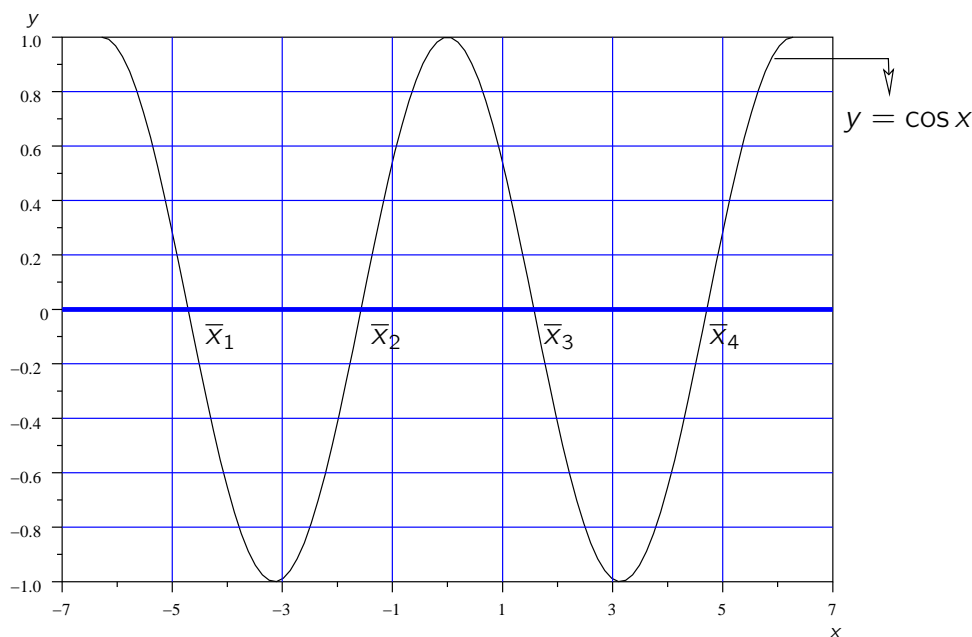


Figura 3.1: Raízes da função  $f(x) = \cos x$ .

Para equações algébricas de grau até quatro, as raízes podem ser calculadas analiticamente por meio de uma expressão tal como  $x = (-b \pm \sqrt{b^2 - 4ac})/2$  para determinar as duas raízes de  $f(x) = ax^2 + bx + c = 0$ . No entanto, para equações algébricas de grau superior a quatro e para a grande maioria das equações transcendentais, as raízes não podem ser calculadas analiticamente. Nestes casos tem de usar métodos que encontrem uma solução **aproximada** para estas raízes.

O problema de calcular uma raiz pode ser dividido em duas fases:

**Fase I:** Isolamento da raiz, isto é, encontrar um intervalo  $[a, b]$  que contenha uma, e somente uma, raiz de  $f(x) = 0$ .

**Fase II:** Refinamento da raiz, que consiste em, escolhidas aproximações iniciais no intervalo encontrado na Fase I, melhorá-las sucessivamente até obter uma aproximação para a raiz dentro de uma precisão  $\epsilon$  pré-fixada.

### 3.1 Fase I: Isolamento das Raízes

Nesta fase é feita uma análise teórica e gráfica da função  $f(x)$ . Na análise teórica usamos frequentemente o teorema:

**Teorema 2.** *Seja  $f(x)$  uma função **contínua** num intervalo  $[a, b]$ . Se  $f(a) \cdot f(b) < 0$  então existe pelo menos um ponto  $\bar{x} \in [a, b]$  tal que  $f(\bar{x}) = 0$ . Além disso, se  $f'(x)$  não muda de sinal em  $[a, b]$  então  $\bar{x}$  é a única raiz de  $f(x)$  neste intervalo.*

A Figura 3.2 exemplifica o Teorema 2 e a Figura 3.3, o caso onde há raiz mas **não** satisfaz as condições do Teorema.

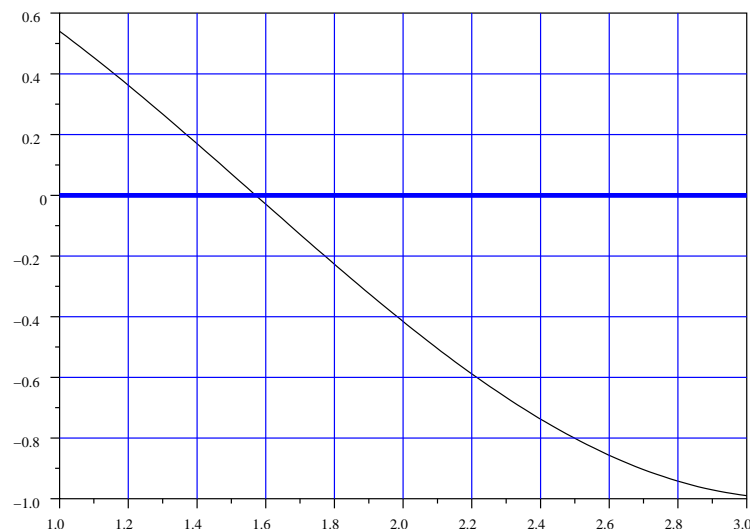


Figura 3.2: Ampliação da função  $f(x) = \cos x$  no intervalo  $x \in [1, 3]$ . Note que  $f'(x) = -\sin x$  e  $f'(x) < 0$  neste intervalo.

A análise gráfica da função  $f(x)$  ou da equação  $f(x) = 0$  é fundamental para obter boas aproximações para a raiz. Para tanto, podemos executar os seguintes procedimentos:

1. Esboçar o gráfico da função  $f(x)$  e localizar as abscissas dos pontos onde a curva intercepta o eixo  $\vec{OX}$ ;
2. A partir da equação  $f(x) = 0$ , obter a equação equivalente  $g(x) = h(x)$ , esboçar os gráficos das funções  $g(x)$  e  $h(x)$  no mesmo eixo cartesiano e localizar os pontos  $x$  onde as duas curvas se interceptam, pois neste caso  $f(\bar{x}) = 0 \iff g(\bar{x}) = h(\bar{x})$ ;
3. Usar os programas que traçam gráficos de funções, disponíveis em algumas calculadoras ou softwares matemáticos.

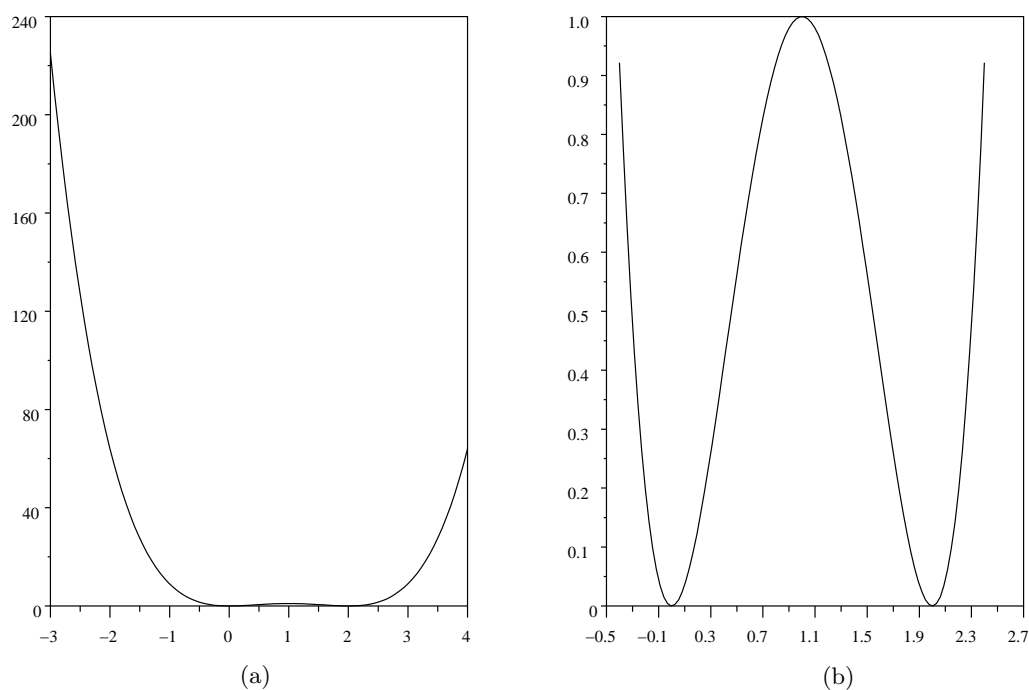


Figura 3.3:  $f(x) = 4x^2 - 4x^3 + x^4$ .

### Exemplo 17.

1.  $f(x) = x^3 - 9x + 3$ : Neste caso, somente o passo (1) é necessário para localizar as possíveis raízes. Porém a Figura 3.4 mostra os três passos.

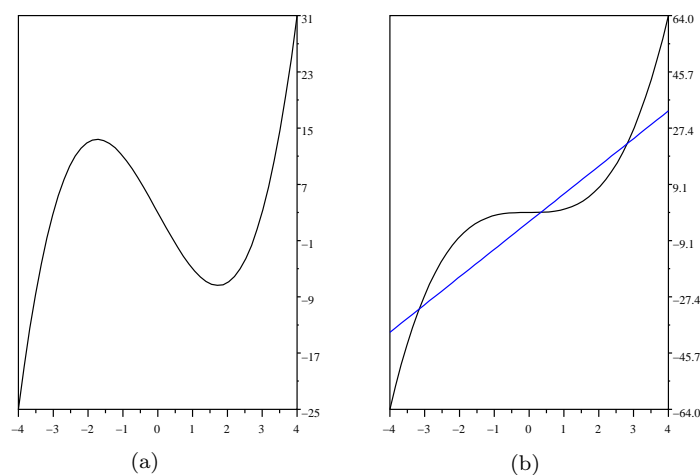


Figura 3.4: (a)  $f(x) = x^3 - 9x + 3$ , (b)  $g(x) = x^3$  e  $h(x) = 9x - 3$ .

2.  $f(x) = \sqrt{x} - 5e^{-x}$ : Neste caso, é mais conveniente usar os passos (2) e (3). Assim,  $\sqrt{x} - 5e^{-x} = 0 \iff \sqrt{x} = 5e^{-x} \implies g(x) = \sqrt{x}$  e  $h(x) = 5e^{-x}$ . A Figura 3.5 ilustra as funções  $f(x)$ ,  $g(x)$  e  $h(x)$ .

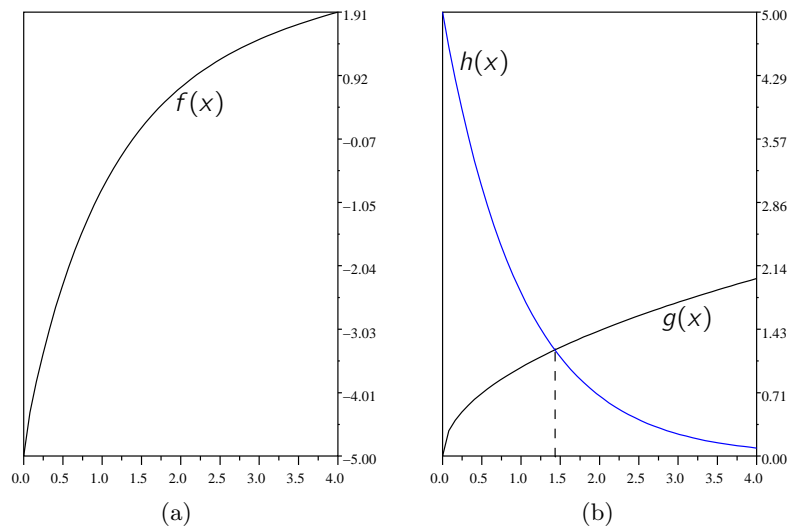


Figura 3.5: (a)  $f(x) = \sqrt{x} - 5e^{-x}$ , (b)  $g(x) = \sqrt{x}$  e  $h(x) = 5e^{-x}$ .

## 3.2 Fase II: Refinamento

Serão apresentados aqui três métodos numéricos de refinamento da raiz: o método da Bissecção, de Newton e das Secantes. A forma como se efetua o refinamento é o que diferencia os métodos. Porém, antes de descrever estes métodos, veremos os critérios de parada adotados.

### 3.2.1 Critério de Parada

O *critério de parada* interrompe a sequência gerada pelos métodos. Este deve avaliar quando  $x_k$ , na  $k$ -ésima iteração, está suficientemente próximo da raiz exata. Contudo, o valor exato da raiz é desconhecido na maioria dos casos, logo, o processo é interrompido quando pelo menos um dos critérios a seguir é satisfeito:

1. Avaliação do ponto na função:

$$|f(x_k)| \leq \epsilon;$$

2. Avaliação do tamanho do intervalo:

$$|x_k - x_{k-1}| \leq \epsilon \quad \text{ou} \quad \left| \frac{x_k - x_{k-1}}{x_k} \right| \leq \epsilon;$$

para  $\epsilon$  suficientemente pequeno (precisão desejada).

### 3.2.2 Método da Bissecção

Seja  $f(x)$  uma função contínua no intervalo  $[a, b]$  e tal que  $f(a) \cdot f(b) < 0$ , ou seja,  $f$  satisfaz as condições do Teorema 2.

Suponha que o intervalo  $(a, b)$  contenha uma única raiz da equação  $f(x) = 0$ . O objetivo deste método é reduzir a amplitude do intervalo que contém a raiz até se atingir a precisão requerida, neste caso  $(b - a) < \epsilon$ , usando para isto a sucessiva divisão de  $[a, b]$  ao meio. A Figura 3.6 mostra este processo.

Os passos para a execução deste método são dados pelo Algoritmo 8.



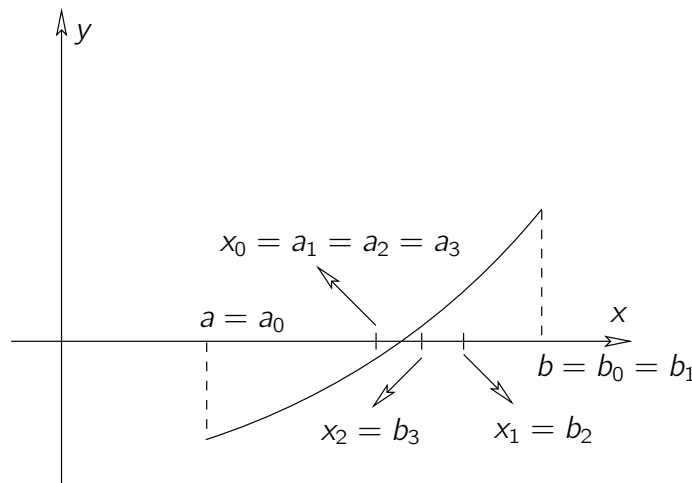


Figura 3.6: O método da Bisseção graficamente.

**Dado**  $f(x)$ ,  $a$  e  $b$  tais que  $f(a) \cdot f(b) < 0$ ,  $\max, \epsilon$

- 1: **Para**  $k = 0$  **até**  $k = \max$  **faça**
- 2:      $x_k = \frac{a + b}{2}$
- 3:     **Se**  $f(a) \cdot f(x_k) < 0$  **então**
- 4:          $b = x_k$
- 5:     **Caso contrário**
- 6:          $a = x_k$
- 7:     **Fim do condicional**
- 8:     **Se**  $|b - a| \leq \epsilon$  **então**
- 9:         **PARE**,  $\bar{x} = \frac{a + b}{2}$  é raiz aproximada de  $f(x)$
- 10:     **Fim do condicional**
- 11: **Fim do laço**

Algoritmo 8: Método da Bisseção

**Exemplo 18.** Aplique o método da Bisseção para encontrar a raiz de  $f(x) = x^3 - 9x + 3$  no intervalo  $I = [0, 1]$ , para  $\epsilon = 10^{-3}$  (a Tabela 3.1 apresenta os passos deste exemplo).

### 3.2.3 Método de Newton

Isaac Newton (1642–1727) publicou seu método para encontrar raízes de equações não-lineares em 1687. Este método também é conhecido como *Newton-Raphson*, devido à sistematização apresentada por Joseph Raphson em 1690.

O método de Newton combina duas idéias comuns nas aproximações numéricas: **linearização** e **iteração**. A linearização substitui a curva  $y = f(x)$  por sua reta tangente.

Seja  $x_0$  uma aproximação inicial da raiz, como ilustra a Figura 3.7. Aproximando a curva  $y = f(x)$  por sua reta tangente traçada no ponto  $(x_0, f(x_0))$  obtemos a aproximação linear. Encontrando o ponto de intersecção desta reta com o eixo  $x$ , obteremos uma nova aproximação para a raiz, o ponto  $x_1$  da figura.

$k$	$a$	$b$	$x_k = (a + b)/2$	$f(x_k)$	$f(a) \cdot f(x_k)$	$ b - a $
0	0	1	0.5	-1.375	(-)	0.5
1	0	0.5	0.25	0.765625	(+)	0.25
2	0.25	0.5	0.375	-0.3222656	(-)	0.125
3	0.25	0.375	0.3125	0.2180176	(+)	0.06
4	0.3125	0.375	0.34375	-0.0531311	(-)	0.03125
5	0.3125	0.34375	0.328125	0.0713582	(+)	0.014375
6	0.328125	0.34375	0.3359375	0.0144744	(+)	0.0078125
7	0.3359375	0.34375	0.3398438	-0.0193443	(-)	0.0039063
8	0.3359375	0.3398438	0.78906	-0.0024384	(-)	0.0019531
9	0.3359375	0.3378906	0.336914	0.0060175	(+)	0.0009766
10	0.336914	0.3378906	0.3374023	0.0017893		

Tabela 3.1: Resultado do Exemplo 18.

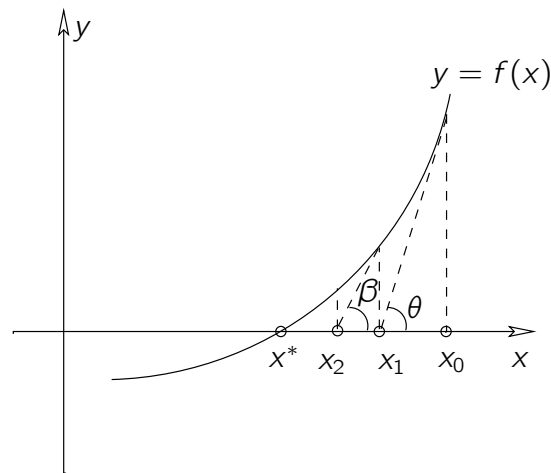


Figura 3.7: O método de Newton-Raphson.

O processo iterativo será obtido pela repetição do procedimento. A nova aproximação, representada na Figura 3.7 pelo ponto  $x_2$ , foi calculada a partir da reta tangente à curva  $y = f(x)$  no ponto  $(x_1, f(x_1))$ .

Para estabelecer expressões analíticas que permitam o cálculo de  $x_1, x_2, \dots$  observamos que a tangente do ângulo  $\theta$  pode ser obtida tanto da definição da função trigonométrica tangente quanto pela derivada de  $f(x)$  no ponto  $x_0$  (inclinação da reta tangente). Assim, da Figura 3.7, temos:

$$\begin{aligned} \operatorname{tg}(\theta) &= \frac{f(x_0)}{x_0 - x_1} = f'(x_0) \quad \rightarrow \quad x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \\ \operatorname{tg}(\beta) &= \frac{f(x_1)}{x_1 - x_2} = f'(x_1) \quad \rightarrow \quad x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} \end{aligned}$$

Genericamente, o processo consiste em evoluir da aproximação  $x_k$  para a aproximação  $x_{k+1}$  usando a fórmula:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad (3.1)$$

O método de Newton-Raphson está esquematizado no Algoritmo 9.

**Exemplo 19.** Aplique o método de Newton para encontrar a raiz de  $f(x) = x^3 - 9x + 3$  tomando  $x_0 = 0.5$ , para  $\epsilon = 10^{-4}$  (a Tabela 3.2 apresenta os passos deste exemplo).

**Dado**  $x_0, f(x), f'(x)$ ,  $\max$  e  $\epsilon$

- 1: **Para**  $k = 0$  **até**  $k = \max$  **faça**
- 2:      $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$
- 3:     **Se**  $|f(x_{k+1})| \leq \epsilon$  **ou**  $|x_{k+1} - x_k| \leq \epsilon$  **então**
- 4:         **PARE**,  $\bar{x} = x_{k+1}$
- 5:     **Fim do condicional**
- 6:     **Se**  $k = \max$  **então**
- 7:         **PARE**, o método **não** converge para a solução.
- 8:     **Fim do condicional**
- 9: **Fim do laço**

Algoritmo 9: Método de Newton-Raphson

$k$	$x_k$	$f(x_k)$	$f'(x_k)$	$ x_{k+1} - x_k $
0	0.5	-1.375	-8.25	1.6667
1	0.3333	0.0370	-8.6667	0.0042735
2	0.3376068	0.0000183	-8.6581	0.0000021
3	0.337609	$4.545 \times 10^{-12}$		

Tabela 3.2: Resultado do Exemplo 19.

### Ordem de Convergência — Método de Newton

**Definição 2.** Sejam  $\{x_k\}$  o resultado da aplicação de um método numérico na iteração  $k$  e  $e_k = |x_k - \bar{x}|$  o seu erro. Se existirem um número  $p$  e uma constante  $C$  tais que:

$$\lim_{k \rightarrow \infty} \frac{|e_{k+1}|}{|e_k|^p} = C$$

então a ordem de convergência é  $p$  desse método.

**Teorema 3.** [4] Se  $f$ ,  $f'$  e  $f''$  são contínuas em um intervalo  $I$  cujo centro  $\bar{x}$  é solução de  $f(x) = 0$  e se  $f'(x) \neq 0$  então a ordem de convergência do Método de Newton é quadrática, ou seja,  $p = 2$ .

**Prova 1.** Seja  $\Psi(x) = x - \frac{f(x)}{f'(x)}$ . Subtraindo a equação  $\bar{x} = \Psi(\bar{x})$  de (3.1), obtemos:

$$x_{k+1} - \bar{x} = \Psi(x_k) - \Psi(\bar{x}).$$

Desenvolvendo  $\Psi(x_k)$  em série de Taylor em torno do ponto  $\bar{x}$ , obtemos:

$$x_{k+1} - \bar{x} = \Psi(\bar{x}) + (x_k - \bar{x})\Psi'(\bar{x}) + \frac{(x_k - \bar{x})^2}{2!}\Psi''(\xi_k) - \Psi(\bar{x})$$

Derivando  $\Psi(\bar{x})$ , temos:

$$\Psi'(\bar{x}) = 1 - \frac{f'(\bar{x}) \cdot f'(\bar{x}) - f(\bar{x}) \cdot f''(\bar{x})}{(f'(\bar{x}))^2} = 1 - \frac{(f'(\bar{x}))^2}{(f'(\bar{x}))^2} = 1 - 1 = 0$$

lembrando que  $f(\bar{x}) = 0$ . Então:

$$x_{k+1} - \bar{x} = \psi(\bar{x}) + \frac{(x_k - \bar{x})^2}{2!} \cdot \psi''(\bar{x})(\xi_k) - \psi(\bar{x})$$

Escrevendo:

$$\frac{|x_{k+1} - \bar{x}|}{|x_k - \bar{x}|^2} = \frac{\psi''(\xi_k)}{2!} \leq C.$$

Logo, a ordem de convergência é  $p = 2$ .

Assim, o método de Newton tem convergência quadrática. Isso significa que a quantidade de dígitos significativos corretos duplica à medida que os valores da sequência se aproxima de  $\bar{x}$ . A desvantagem do método de Newton está no fato de termos de calcular a derivada da função e, em cada iteração, calcular o seu valor numérico, pode ser muito caro computacionalmente. Além disso, a função pode ser não diferenciável em alguns pontos do domínio.

### 3.2.4 Método da Secante

Uma grande desvantagem do método de Newton é a necessidade de se obter  $f'(x)$  e calcular seu valor numérico a cada iteração. Uma alternativa é usar retas secantes como aproximações lineares locais da função, em vez de tangentes. Neste caso, são necessárias duas aproximações para inicializarmos o processo,  $x_0$  e  $x_1$ .

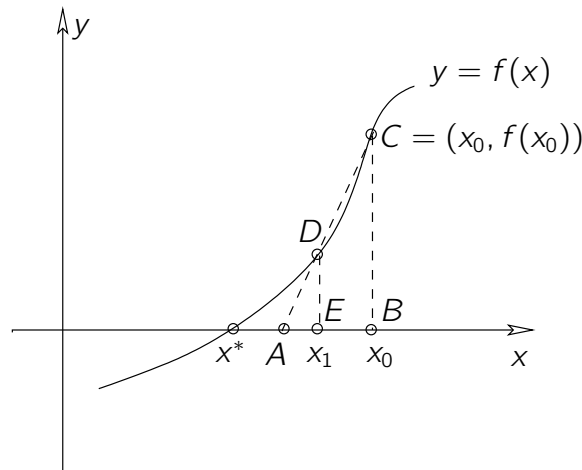


Figura 3.8: O método das Secantes.

No método da Secante, tomamos a reta que passa pelos pontos  $(x_0, f(x_0))$  e  $(x_1, f(x_1))$  como uma aproximação linear da curva  $y = f(x)$ , como indica a Figura 3.8.

Para estabelecermos a relação de recorrência do *Método das Secantes*, usamos a semelhança de triângulos  $ABC$  e  $AED$ :

$$\frac{f(x_0)}{x_0 - x_2} = \frac{f(x_1)}{x_1 - x_2}$$

onde  $x_2$  é o ponto denotado por A na Figura 3.8. Explicitando o valor da incógnita  $x_2$  teremos:

$$x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)}.$$

Generalizando, no método das secantes usamos duas aproximações  $x_{k-1}$  e  $x_k$ , para calcular uma nova apro-

ximação  $x_{k+1}$ , através da fórmula:

$$x_{k+1} = \frac{x_{k-1}f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}.$$

O Algoritmo 10 sistematiza o procedimento para o método das Secantes.

**Dado**  $x_0, x_1, f(x), \max$  e  $\epsilon$

- 1: **Para**  $k = 1$  **até**  $k = \max$  **faça**
- 2:  $x_{k+1} = \frac{x_{k-1}f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}$
- 3: **Se**  $|f(x_{k+1})| \leq \epsilon$  **ou**  $|x_{k+1} - x_k| \leq \epsilon$  **então**
- 4: **PARE**,  $\bar{x} = x_{k+1}$
- 5: **Fim do condicional**
- 6: **Se**  $k = \max$  **então**
- 7: **PARE**, o método **não** converge para a solução.
- 8: **Fim do condicional**
- 9: **Fim do laço**

Algoritmo 10: Algoritmo para o Método das Secantes

**Exemplo 20.** Aplique o método das Secantes para encontrar a raiz de  $f(x) = x^3 - 9x + 3$  tomando  $x_0 = 0$  e  $x_1 = 1$ , para  $\epsilon = 10^{-4}$  (a Tabela 3.3 apresenta os passos deste exemplo).

$k$	$x_{k-1}$	$x_k$	$f(x_{k-1})$	$f(x_k)$	$ x_{k+1} - x_k $
1	0	1	3	-5	0.625
2	1	0.375	-5	-0.3222656	0.0430585
3	0.375	0.3319415	-0.3222656	0.0491011	0.0056931
4	0.3319415	0.3376346	0.0491011	0.0002222	0.0000256
5	0.3376346	0.337609	0.0002222	$-1.464 \times 10^{-7}$	

Tabela 3.3: Resultado do Exemplo 20.

### Ordem de Convergência — Método da Secante

**Teorema 4.** [4] A ordem de convergência do método da Secante é  $p = \frac{1 + \sqrt{5}}{2} \simeq 1.618$ .

Note que, apesar da ordem de convergência do método da Secante ser inferior à do método de Newton, ele fornece uma alternativa viável, desde que requer somente um cálculo da função  $f$  por passo, enquanto dois cálculos ( $f(x_k)$  e  $f'(x_k)$ ) são necessários para o método de Newton.

## 3.3 Sistemas Não-Lineares

Resolver um problema de equações não-lineares é uma tarefa complexa, eventualmente contornada aproximando-se o problema original por um sistema de equações lineares. Quando essa solução é insatisfatória, o problema deve ser atacado diretamente.

A principal ferramenta para resolver uma equação não-linear é o *Método de Newton*. Essa técnica será modificada para resolver um sistema de equações não-lineares.

### 3.3.1 Método de Newton para Sistemas Não-Lineares

Considere  $n$  funções :  $f_1, f_2, \dots, f_n$ , onde cada uma delas é uma função de  $n$  variáveis  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$  para  $i = 1, 2, \dots, n$ , ou seja,  $x_1, x_2, \dots, x_n$ . Queremos encontrar os valores de  $x_1^*, x_2^*, \dots, x_n^*$  tais que:

$$\begin{cases} f_1(x_1^*, x_2^*, \dots, x_n^*) = 0 \\ f_2(x_1^*, x_2^*, \dots, x_n^*) = 0 \\ \vdots \\ f_n(x_1^*, x_2^*, \dots, x_n^*) = 0 \end{cases} \quad (3.2)$$

Considere o seguinte exemplo:

$$\begin{cases} -x_1(x_1 + 1) + 2x_2 = 18 \\ (x_1 - 1)^2 + (x_2 - 6)^2 = 25 \end{cases}$$

Aqui, procuramos  $x_1$  e  $x_2$ , tais que:

$$\begin{aligned} f_1(x_1^*, x_2^*) &= -x_1^*(x_1^* + 1) + 2x_2^* - 18 = 0 \\ f_2(x_1^*, x_2^*) &= (x_1^* - 1)^2 + (x_2^* - 6)^2 - 25 = 0 \end{aligned}$$

Usando a notação vetorial, temos:

$$\begin{aligned} \mathbf{x} &= (x_1, x_2, \dots, x_n), \\ \mathbf{F}(\mathbf{x}) &= (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x})) \end{aligned}$$

Assim, o sistema (3.2) pode ser representado por uma única equação vetorial,  $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ .

Como foi feito no caso de uma equação não-linear, a idéia da linearização será usada no caso de sistemas. Tomando a *série de Taylor* como uma aproximação de funções, em torno de  $x_k$ , para uma função escalar, temos:

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{f''(x_k)}{2!}(x - x_k)^2 + \dots + \frac{f^{(n)}(x_k)}{n!}(x - x_k)^n + \frac{f^{(n+1)}(\alpha)}{(n+1)!}(x - x_k)^{n+1}$$

onde  $x \leq \alpha \leq x_k$ . O último termo desta expressão representa o erro da aproximação de  $f(x)$  pelo polinômio de grau  $n$ , o polinômio de Taylor, formado pelos  $n + 1$  primeiros termos da expansão em série de Taylor.

Observe que, se tomarmos a aproximação linear (usando os dois primeiros termos da série de Taylor), teremos:

$$f(x) \approx f(x_k) + f'(x_k)(x - x_k).$$

Se usarmos esta aproximação linear para encontrar a raiz de  $f(x) = 0$ ,

$$f(\bar{x}) \approx 0 \quad \text{se} \quad f(x_k) + f'(x_k)(\bar{x} - x_k) = 0,$$

ou ainda, explicitando  $\bar{x}$  temos,

$$\bar{x} = x_k - \frac{f(x_k)}{f'(x_k)}. \quad (3.3)$$

Assim, a série de Taylor é outra maneira de se obter a função de iteração do método de Newton.

Aplicando esta idéia para resolver o sistema (3.2), temos a aproximação de Taylor como:

$$\begin{aligned}\mathbf{F}(\mathbf{x}) &= \mathbf{F}(\mathbf{x}^k) + \mathbf{F}'(\mathbf{x}^k)(\mathbf{x} - \mathbf{x}^k) + \mathbf{E} \\ &= \text{aproximação linear} + \text{erro}\end{aligned}\quad (3.4)$$

onde  $\mathbf{x}^k$  = vetor aproximação na  $k$ -ésima iteração e  $\mathbf{E}$  é um vetor que representa o erro da aproximação linear. Analogamente à série de Taylor de funções de uma variável,  $\|\mathbf{E}\| \leq C\|\mathbf{x} - \mathbf{x}^k\|^2$  para alguma constante  $C$ , quando  $\mathbf{x}^k \rightarrow \mathbf{x}$ .

Note que,  $\mathbf{x}^k$  denota o vetor obtido na  $k$ -ésima iteração:  $\mathbf{x}^k = (x_1^k, x_2^k, \dots, x_n^k)$ .

Em (3.4),  $\mathbf{F}'(\mathbf{x})$  é a derivada de uma função vetorial com variáveis vetoriais. Esta derivada é uma matriz que contém todas as derivadas parciais de todos os componentes da função  $\mathbf{F}(\mathbf{x})$ , a matriz *Jacobiana* de  $\mathbf{F}(\mathbf{x})$ ,

$$\mathbf{J}(\mathbf{x}) = \mathbf{F}'(\mathbf{x}) = [J_{ij}] = \left[ \frac{\partial f_i(\mathbf{x})}{\partial x_j} \right].$$

Para estabelecer o método iterativo, a aproximação na iteração  $k+1$  será definida pelo vetor que anula a parte linear da equação (3.4), isto é, o vetor  $\mathbf{x}^{k+1}$  é tal que:

$$\begin{aligned}\mathbf{F}(\mathbf{x}^k) + \mathbf{J}(\mathbf{x}^k)(\mathbf{x}^{k+1} - \mathbf{x}^k) &= \mathbf{0} \\ \mathbf{J}(\mathbf{x}^k)(\mathbf{x}^{k+1} - \mathbf{x}^k) &= -\mathbf{F}(\mathbf{x}^k)\end{aligned}\quad (3.5)$$

Para explicitar  $\mathbf{x}^{k+1}$ , multiplicamos a equação acima pela inversa da Jacobiana e teremos:

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{J}^{-1}(\mathbf{x}^k) \mathbf{F}(\mathbf{x}^k),$$

que é uma generalização de (3.3) para o caso de sistemas não-lineares.

Como a inversão de matrizes é um processo trabalhoso e computacionalmente caro, evitamos este cálculo fazendo:

$$\begin{aligned}\mathbf{v} &= -\mathbf{J}^{-1}(\mathbf{x}^k) \mathbf{F}(\mathbf{x}^k) \\ \mathbf{J}(\mathbf{x}^k) \mathbf{v} &= -\mathbf{F}(\mathbf{x}^k)\end{aligned}$$

Ou seja, o cálculo das equações acima envolve um sistema de equações lineares para a aproximação da iteração  $k+1$ . Assim, a nova aproximação será:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{v}$$

O método de Newton para Sistemas não-Lineares é dado pelo algoritmo (11)

**Exemplo 21.** Resolva o sistema que segue, usando o método de Newton para equações não-lineares.

$$\begin{aligned}f_1(x_1, x_2) &= 2x_1^3 - x_2^2 - 1 = 0 \\ f_2(x_1, x_2) &= x_1x_2^3 - x_2 - 4 = 0\end{aligned}$$

As derivadas parciais desta função são:

$$\begin{aligned}\frac{\partial f_1}{\partial x_1} &= 6x_1^2 & \frac{\partial f_1}{\partial x_2} &= -2x_2 \\ \frac{\partial f_2}{\partial x_1} &= x_2^3 & \frac{\partial f_2}{\partial x_2} &= 3x_1x_2^2 - 1\end{aligned}$$

**Dado**  $\mathbf{x}^0$ ,  $f_i$  para  $i = 1, 2, \dots, n$ ,  $\frac{\partial f_i}{\partial x_j}$  para  $j = 1, 2, \dots, n$ ,  $\epsilon_1$  e  $\epsilon_2$

```

1: Para  $k = 1$  até  $k = \max$  faça
2:   Para  $i = 1$  até  $i = n$  faça
3:      $F_i = f_i(\mathbf{x}^{k-1})$ 
4:     Para  $j = 1$  até  $j = n$  faça
5:        $J_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}^{k-1})$ 
6:     Fim do laço
7:   Fim do laço
8:   Resolva o sistema  $\mathbf{J}\mathbf{v} = -\mathbf{F}$ 
9:    $\mathbf{x}^k = \mathbf{x}^{k-1} + \mathbf{v}$ 
10:  Se  $\max_{1 \leq i \leq n} |f_i(\mathbf{x}^k)| < \epsilon_1$  então
11:    Então  $\mathbf{x}^* = \mathbf{x}^k$ 
12:  Fim do condicional
13:  ... ou ...
14:  Se  $\max_{1 \leq i \leq n} |x_i^k - x_i^{k-1}| < \epsilon_2$  então
15:    Então  $\mathbf{x}^* = \mathbf{x}^k$ 
16:  Fim do condicional
17: Fim do laço

```

Algoritmo 11: Método de Newton para Sistemas Não-Lineares

e, portanto, a matriz Jacobiana é:

$$\mathbf{J}(x_1, x_2) = \begin{bmatrix} 6x_1^2 & -2x_2 \\ x_2^3 & 3x_1x_2^2 - 1 \end{bmatrix}$$

Tomando uma aproximação inicial  $\mathbf{x}^0 = (1.2, 1.7)$  temos:

$$\mathbf{J}(\mathbf{x}^0) = \begin{bmatrix} 8.64 & -3.4 \\ 4.91 & 9.4 \end{bmatrix} \quad \text{e} \quad \mathbf{F}(\mathbf{x}^0) = \begin{bmatrix} -0.4340 \\ 0.1956 \end{bmatrix}$$

Resolvendo o sistema  $\mathbf{J}(\mathbf{x}^k)\mathbf{v} = -\mathbf{F}(\mathbf{x}^k)$  encontramos:

$$\mathbf{v} = (0.03488, -0.03902)^T$$

Assim,

$$\mathbf{x}^1 = (1.2349, 1.6610)^T.$$

Para o cálculo de uma nova aproximação recalculamos  $\mathbf{F}(\mathbf{x}^1)$  e a matriz Jacobiana em  $\mathbf{x}^1$ . Resolvendo o sistema linear para obter  $\mathbf{v}$ , temos que:

$$\mathbf{x}^2 = (1.2343, 1.6615)^T$$

Podemos testar as aproximações obtidas calculando os respectivos  $f_1$  e  $f_2$ :

$$\begin{array}{ll} f_1(\mathbf{x}^0) = -0.4340 & f_2(\mathbf{x}^0) = 0.1956 \\ f_1(\mathbf{x}^1) = 7.4607 \times 10^{-3} & f_2(\mathbf{x}^1) = -1.9876 \times 10^{-3} \\ f_1(\mathbf{x}^2) = -1.2779 \times 10^{-4} & f_2(\mathbf{x}^2) = 3.202 \times 10^{-4} \end{array}$$



Observe que a solução do sistema e o cálculo do Jacobiano, a cada iteração, são dispendiosos. Para economizar os cálculos, podemos manter uma matriz Jacobiana fixa em algumas iterações. Neste caso, a decomposição  $LU$  é adequada na solução do sistema, pois a mesma é usada em algumas iterações, atualizando-se apenas o termo independente do sistema.

### 3.4 Exercícios

1. Localize **graficamente** as raízes das equações a seguir:

$$\begin{array}{lll} \text{(a)} \quad 4 \cos x - e^{2x} = 0; & \text{(c)} \quad 1 - x \ln x = 0; & \text{(e)} \quad x^3 + x - 1000 = 0. \\ \text{(b)} \quad \frac{x}{2} - \operatorname{tg} x = 0 & \text{(d)} \quad 2^x - 3x = 0; & \end{array}$$

2. Calcular pelo menos uma raiz de cada equação abaixo com  $\epsilon \leq 0.001$  pelo método da Bissecção.

$$\begin{array}{ll} \text{(a)} \quad f(x) = e^{2x} - 2x^3 - 5 = 0; & \text{(b)} \quad f(x) = 2x^3 - 5x^2 - x + 3 = 0; \\ \text{(c)} \quad f(x) = 5x^2 + \log(x+1) - 2 = 0. & \end{array}$$

3. Seja  $f(x) = (x+2)(x+1)x(x-1)^3(x-2)$ . Para que zero de  $f$  o método da Bissecção vai convergir quando for aplicado nos seguintes intervalos ?

$$\begin{array}{llll} \text{(a)} \quad [-3, 2.5] & \text{(b)} \quad [-2.5, 3] & \text{(c)} \quad [-1.75, 1.5] & \text{(d)} \quad [-1.5, 1.75] \\ (\bar{x} = 2) & (\bar{x} = -2) & (\bar{x} = -1) & (\bar{x} = 1) \end{array}$$

4. Encontre um valor aproximado para  $\sqrt{3}$  com precisão de  $10^{-4}$  utilizando:

$$\text{(a)} \quad \text{Método da Bissecção}; \quad \text{(b)} \quad \text{Método da Secante}; \quad \text{(c)} \quad \text{Método de Newton.}$$

5. Determinar pelo menos uma raiz de cada equação abaixo com  $\epsilon \leq 10^{-4}$  usando o método das Secantes.

$$\begin{array}{ll} \text{(a)} \quad f(x) = 2x^3 - 5x^2 - 10x + 20 = 0; & \text{(b)} \quad f(x) = 5 \log x + 3x^4 - 7 = 0; \\ \text{(c)} \quad f(x) = 2^x + (\cos x)x^2 = 0 & \end{array}$$

6. Achar pelo menos uma raiz positiva de cada equação com  $\epsilon \leq 10^{-5}$  pelo método de Newton.

$$\begin{array}{ll} \text{(a)} \quad 4x^3 + x + \cos x - 10 = 0; & \text{(b)} \quad x^4 - 2x^3 + 2x - 1 = 0; \\ \text{(c)} \quad f(x) = (x-2)(e^{x-2} - 1) = 0. & \end{array}$$

7. A função  $f(x) = (4x-7)/(x-2)$  se anula em  $x = 7/4$ . Calcule as iterações do método de Newton partindo das aproximações iniciais:

$$\begin{array}{lll} \text{(a)} \quad x_0 = 1.625; & \text{(c)} \quad x_0 = 1.5; & \text{(e)} \quad x_0 = 3; \\ \text{(b)} \quad x_0 = 1.875; & \text{(d)} \quad x_0 = 1.95; & \text{(f)} \quad x_0 = 7. \end{array}$$

Explique graficamente seus resultados.

8. Seja  $f(x) = (4x - 7)/(x - 2)$ . Verifique que  $f(1.8)f(3) < 0$ . É possível usar o método da Bissecção para localizar raízes neste intervalo? Explique.

9. Demonstrar que a raiz  $\sqrt[p]{a}$ ,  $a \geq 0$  pode ser calculada pela fórmula de recorrência:

$$x_{k+1} = \frac{1}{p} \left( (p-1)x_k + \frac{a}{x_k^{p-1}} \right), \quad x_0 > 0.$$

10. A função  $f(x) = \operatorname{tg}(\pi x) - 6$  tem um zero em  $\bar{x} = (1/\pi) \arctan 6 \approx 0.447431543$ . Considerando  $x_0 = 0$  e  $x_1 = 0.48$ , use 10 iterações para cada um dos seguintes métodos para calcular o valor aproximado dessa raiz. Que método é mais bem-sucedido? Por quê?

(a) Método da Bissecção;

(b) Método da Secante.

11. Use o Método de Newton com  $\mathbf{x}^0 = \mathbf{0}$  para calcular  $\mathbf{x}^2$  para cada um dos seguintes sistemas não lineares:

$$(a) \begin{cases} 4x_1^2 - 20x_1 + \frac{1}{4}x_2^2 + 8 = 0 \\ \frac{1}{2}x_1x_2^2 + 2x_1 - 5x_2 + 8 = 0 \end{cases}$$

$$\text{Resp.: } \mathbf{x}^2 = (0.4958936, 1.983423)^T$$

$$(c) \begin{cases} 3x_1 - \cos(x_2x_3) - \frac{1}{2} = 0 \\ 4x_1^2 - 625x_2^2 + 2x_2 - 1 = 0 \\ e^{-x_1x_2} + 20x_3 + \frac{10\pi-3}{3} = 0 \end{cases}$$

$$\text{Resp.: } \mathbf{x}^2 = (0.5001667, 0.2508036, -0.5173874)^T$$

$$(b) \begin{cases} \sin(4\pi x_1x_2) - 2x_2 - x_1 = 0 \\ \left( \frac{4\pi-1}{4\pi} \right) (e^{2x_1} - e) + 4ex_2^2 - 2ex_1 = 0 \end{cases}$$

$$\text{Resp.: } \mathbf{x}^2 = (-0.5131616, -0.01837622)^T$$

$$(d) \begin{cases} x_1^2 + x_2 - 37 = 0 \\ x_1 - x_2^2 - 5 = 0 \\ x_1 + x_2 + x_3 - 3 = 0 \end{cases}$$

$$\text{Resp.: } \mathbf{x}^2 = (4.350877, 18.49123, -19.84211)^T$$

12. Use o método de Newton para encontrar uma solução para os seguintes sistemas não-lineares nos domínios dados. Faça iterações até  $\|\mathbf{x}^k - \mathbf{x}^{k-1}\| < 10^{-6}$ .

(a) Use  $\mathbf{x}^0 = (1, 1)^T$

$$\begin{cases} 3x_1^2 - x_2^2 = 0 \\ 3x_1x_2^2 - x_1^3 - 1 = 0 \end{cases}$$

$$\text{Resp.: } \mathbf{x}^5 = (0.5000, 0.8660254)^T$$

(c) Use  $\mathbf{x}^0 = (-1, -2, 1)^T$

$$\begin{cases} x_1^3 + x_1^2x_2 - x_1x_3 + 6 = 0 \\ -e^{x_1} + e^{x_2} - x_3 = 0 \\ x_2^2 - 2x_1x_3 = 4 \end{cases}$$

$$\text{Resp.: } \mathbf{x}^5 = (-1.456043, -1.66423, 0.4224934)^T$$

(b) Use  $\mathbf{x}^0 = (2, 2)^T$

$$\begin{cases} \ln(x_1^2 + x_2^2) - \sin(x_1x_2) = \ln 2 + \ln \pi \\ e^{x_1-x_2} + \cos(x_1x_2) = 0 \end{cases}$$

$$\text{Resp.: } \mathbf{x}^6 = (1.772454, 1.772454)^T$$

(d) Use  $\mathbf{x}^0 = (0, 0, 0)^T$

$$\begin{cases} 6x_1 - 2\cos(x_2x_3) - 1 = 0 \\ 9x_2 + \sqrt{x_1^2 + \sin x_3 + 1.06} + 0.9 = 0 \\ 60x_3 + 3e^{-x_1x_2} + 10\pi - 3 = 0 \end{cases}$$

$$\text{Resp.: } \mathbf{x}^4 = (0.4981447, -0.1996059, -0.528826)^T$$

## Ajuste de Curvas

### Introdução

Apresentaremos aqui os aspectos básicos e teóricos para *ajuste de curvas* usando interpolação e aproximação de funções de uma variável real.

Os problemas de interpolação e aproximação estudados surgem ao se aproximar uma função  $f$  por outra função  $g$  mais apropriada. Ou ainda, quando há a necessidade de aproximar dados tabelados através de uma função. Para isso existem duas classes de métodos e a distinção entre elas está em considerarmos, ou não, a existência de erros nos dados.

1. **Interpolação:** os dados são precisos e portanto pode-se exigir que a curva de ajuste passe pelos pontos dados. Em geral, as funções escolhidas para o ajuste são polinômios.

Estabelece um método que permite encontrar um polinômio que passe por  $n+1$  pontos conhecidos, denotados por  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$  onde  $x_0 \neq x_1 \neq x_2 \neq \dots \neq x_n$ .

2. **Método dos Quadrados Mínimos:** considera-se possíveis erros introduzidos na obtenção dos dados.

### 4.1 Interpolação

Seja a Tabela 4.1.

$x_i$	0.1	0.6	0.8
$f(x_i)$	1.221	3.320	4.953

Tabela 4.1: Dados para interpolação.

O problema consiste em encontrar o valor correspondente de  $y = f(x)$  para um dado  $x$  **não** pertencente à tabela. Um modo de resolver este problema é obter uma função que relaciona as variáveis  $x$  e  $y$ . Considerando que os polinômios são as funções mais simples e estudadas, então eles são os mais utilizados para determinar esta *relação*. Um polinômio construído com o intuito de aproximar uma função é denominado *polinômio interpolador*. Assim, para resolver o problema basta avaliar o polinômio no ponto desejado.

Existem vários métodos para construir um polinômio interpolador a partir de um conjunto de pares de dados. Aqui, estudaremos o Polinômio Interpolador de Lagrange e o de Newton.

### 4.1.1 Polinômio Interpolador de Lagrange

Dados  $y_0 = f(x_0)$ ,  $y_1 = f(x_1)$ , ...,  $y_n = f(x_n)$ , desejamos encontrar um polinômio  $p(x)$  tal que  $p(x_0) = f(x_0)$ ,  $p(x_1) = f(x_1)$ , ...,  $p(x_n) = f(x_n)$ . A Figura 4.1 ilustra este caso.

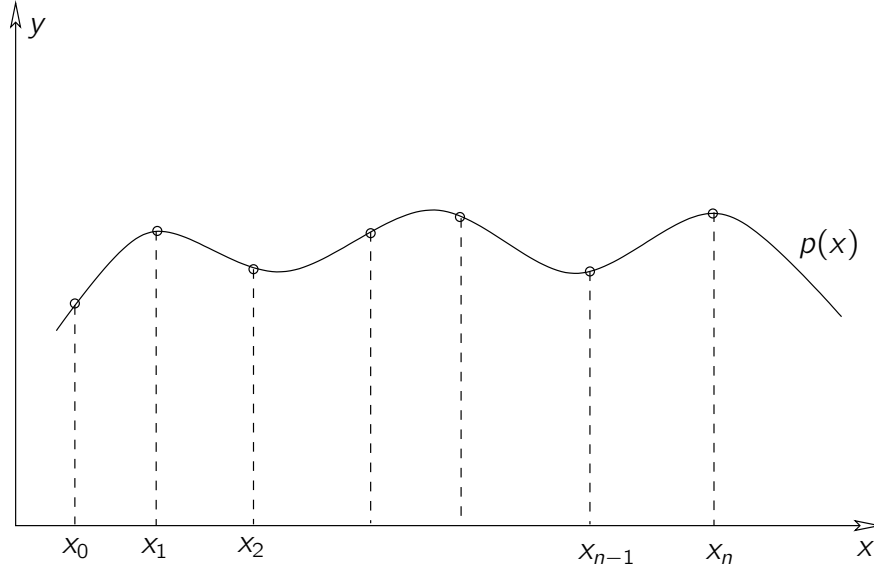


Figura 4.1: Exemplo de Interpolação.

**Teorema 5.** *Seja  $f(x)$  uma função conhecida nos  $n + 1$  pontos distintos  $x_0, x_1, x_2, \dots, x_n$ . Existe um único polinômio  $p(x)$ , de grau menor ou igual a  $n$ , tal que:*

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

**Prova 2.** *Seja:*

$$\begin{aligned} p(x) &= f(x_0)l_0^n(x) + f(x_1)l_1^n(x) + f(x_2)l_2^n(x) + \dots + f(x_n)l_n^n(x) \\ &= \sum_{k=0}^n f(x_k)l_k^n(x) \end{aligned} \quad (4.1)$$

onde  $l_k^n$  são os polinômios de Lagrange definidos por:

$$\begin{aligned} l_k^n &= \frac{(x - x_0)(x - x_1) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n)}{(x_k - x_0)(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n)} \\ &= \prod_{i=0, i \neq k}^n \frac{(x - x_i)}{(x_k - x_i)}. \end{aligned} \quad (4.2)$$

Sabemos que:

$$l_k^n(x_k) = 1 \quad \text{e} \quad l_k^n(x_i) = 0 \quad \text{para } i \neq k.$$

Com isso, verificamos que  $p(x)$  é o polinômio interpolador.

A unicidade é consequência de ser  $p(x) \equiv 0$  o único polinômio de grau  $n$  que se anula em  $n + 1$  pontos distintos. De fato, supondo que existam dois polinômios interpoladores  $p_1(x)$  e  $p_2(x)$ , usamos (4.1) com  $p(x) = p_1(x) - p_2(x)$  para mostrar que  $p(x) \equiv 0$  e portanto,  $p_1(x) \equiv p_2(x)$ .

O polinômio interpolador (4.1) é conhecido como *Fórmula de Lagrange*.

**Exemplo 22.** Encontre o polinômio que interpola  $f(x) = \frac{1}{x^2}$  nos pontos  $x_0 = 2$ ,  $x_1 = 2.5$  e  $x_2 = 4$ .

Como serão usados três pontos, o polinômio interpolador tem grau 2. Pela substituição dos pontos de interpolação na expressão 4.2, calculamos os polinômios de Lagrange:

$$\begin{aligned}l_0^2(x) &= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 2.5)(x - 4)}{(2 - 2.5)(2 - 4)} = x^2 - 6.5x + 10 \\l_1^2(x) &= \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 2)(x - 4)}{(2.5 - 2)(2.5 - 4)} = \frac{-4x^2 + 24x + 32}{3} \\l_2^2(x) &= \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 2)(x - 2.5)}{(4 - 2)(4 - 2.5)} = \frac{x^2 - 4.5x + 5}{3}\end{aligned}$$

Lembrando que  $f(x) = \frac{1}{x^2}$ ,  $f(x_0) = f(2) = 0.25$ ,  $f(x_1) = f(2.5) = 0.16$  e  $f(x_2) = f(4) = 0.0625$ . Usando a expressão (4.1) teremos o polinômio interpolador  $p_2(x)$ :

$$\begin{aligned}p_2(x) &= 0.25(x^2 - 6.5x + 10) + \frac{0.16}{3}(-4x^2 + 24x + 32) + \frac{0.0625}{3}(x^2 - 4.5x + 5) \\&= 0.0575x^2 - 0.4388x + 0.8975\end{aligned}$$

Apenas para comparação, tomando  $x = 3$  teremos  $p_2(x) = 0.0986$ , que é uma aproximação de  $f(3) = 0.1111$ , com erro absoluto de 0.0125.

A Figura 4.2 ilustra este exemplo.

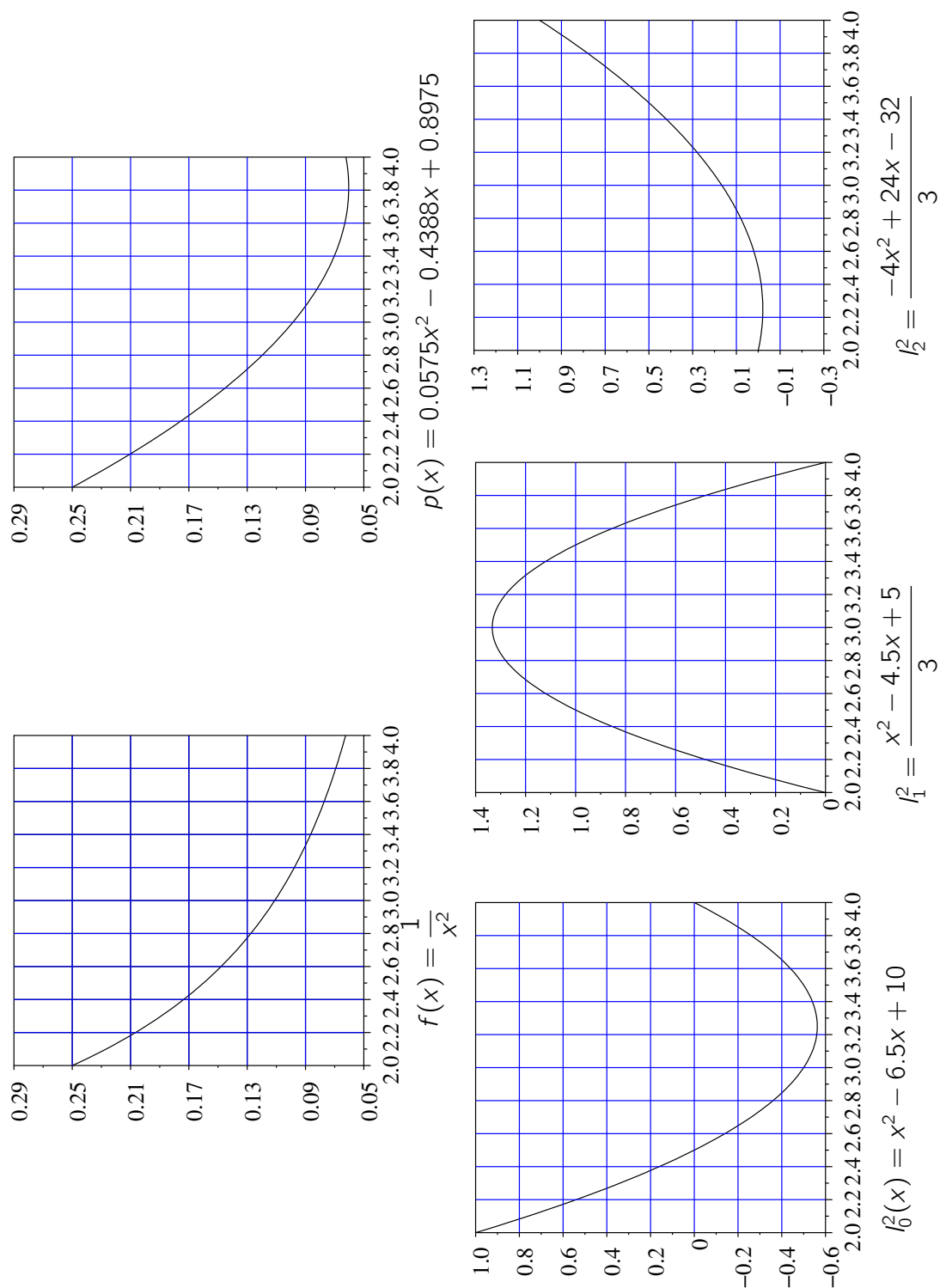


Figura 4.2: Exemplo sobre Interpolação de Lagrange.

### 4.1.2 Erro na Interpolação

O polinômio interpolador  $P_n(x)$  para uma função  $y = f(x)$  sobre um conjunto de pontos distintos  $x_0, x_1, \dots, x_n$  tem a propriedade:

$$P_n(x_k) = f(x_k), \quad k = 0, 1, \dots, n.$$

Nos pontos  $\bar{x} \neq x_k$  nem sempre é verdade que  $P_n(\bar{x}) = f(\bar{x})$ . Entretanto, para avaliar  $f(x)$  nos pontos  $\bar{x} \neq x_k$ ,  $k = 0, 1, \dots, n$ , consideramos  $P_n(x)$  como uma aproximação para a função  $y = f(x)$  em um intervalo que contenha os pontos  $x_0, x_1, \dots, x_n$  e calculamos  $f(\bar{x})$  através de  $P_n(\bar{x})$ . Algumas perguntas que surgem são:

- i) Podemos ter idéia do erro que cometemos quando substituímos  $f(x)$  por  $P_n(x)$ ?
- ii) O polinômio da interpolação é uma boa aproximação para  $f(x)$ ?

**Teorema 6.** (Teorema de Rolle) *Seja  $f(x)$  contínua em  $[a, b]$  e diferenciável em cada ponto de  $(a, b)$ . Se  $f(a) = f(b)$ , então existe um ponto  $x = \xi$ ,  $a < \xi < b$ , tal que  $f'(\xi) = 0$ .*

**Prova 3.** Vide [8] em “Aplicações da Derivada”.

**Teorema 7.** (Teorema de Rolle generalizado) *Seja  $n \geq 2$ . Suponhamos que  $f(x)$  seja contínua em  $[a, b]$  e que  $f^{(n-1)}(x)$  exista em cada ponto de  $(a, b)$ . Suponhamos que  $f(x_1) = f(x_2) = \dots = 0$  para  $a \leq x_1 < x_2 < \dots < x_n \leq b$ . Então existe um ponto  $\xi$ ,  $x_1 < \xi < x_n$ , tal que  $f^{(n-1)}(\xi) = 0$ .*

**Prova 4.** Aplicação sucessiva do Teorema de Rolle.

Vejamos agora um teorema que nos fornece uma expressão do termo do erro.

**Teorema 8.** *Seja  $f(x)$  contínua em  $[a, b]$  e suponhamos que  $f^{(n+1)}(x)$  exista em cada ponto  $(a, b)$ . Se  $a \leq x_0 < x_1 < \dots < x_n \leq b$ , então:*

$$R_n(f; x) = f(x) - P_n(f; x) = \frac{(x - x_0)(x - x_1) \dots (x - x_n)}{(n+1)!} f^{(n+1)}(\xi) \quad (4.3)$$

onde  $\min\{x, x_0, x_1, \dots, x_n\} < \xi < \max\{x, x_0, x_1, \dots, x_n\}$ . O ponto  $\xi$  depende de  $x$ .

**Prova 5.** Vide [4].

O termo  $R_n(f; x)$  na expressão  $f(x) = P_n(f; x) + R_n(f; x)$  é chamado *termo do erro* ou *erro de truncamento*: é o erro que se comete no ponto  $x$  quando se substitui a função por seu polinômio de interpolação calculado em  $x$ .

A importância do teorema é mais teórica do que prática, visto que não conseguimos determinar o ponto  $\xi$  de modo que seja válida a igualdade (4.3). Na prática, para estimar o erro cometido ao aproximar o valor da função num ponto por seu polinômio de interpolação, utilizamos o corolário a seguir.

**Corolário 4.1.1.** *Seja  $R_n(f; x) = f(x) - P_n(f; x)$ . Se  $f(x)$  e suas derivadas até a ordem  $n+1$  são contínuas em  $[a, b]$  então:*

$$|R_n(f; x)| \leq \frac{|x - x_0| |x - x_1| \dots |x - x_n|}{(n+1)!} \max_{a \leq t \leq b} |f^{(n+1)}(t)|. \quad (4.4)$$

**Exemplo 23.** Dada a tabela:

$x$	0	0.1	0.2	0.3	0.4	0.5
$e^{3x}$	1	1.3499	1.8221	2.4596	3.3201	4.4817

calcule um limitante superior para o erro de truncamento quando avaliamos  $f(0.25)$ , onde  $f(x) = x e^{3x}$  usando o polinômio de interpolação do 2º grau.

**Solução:** Tomemos a equação (4.4):

$$|R_n(f; x)| \leq \frac{|x - x_0| |x - x_1| |x - x_2|}{3!} \max_{x_0 \leq t \leq x_2} |f'''(t)|$$

Como  $f(t) = t e^{3t}$  segue que:

$$\begin{aligned} f'(t) &= e^{3t} + 3t e^{3t} &= e^{3t}(1 + 3t) \\ f''(t) &= 3e^{3t} + 3e^{3t} + 9t e^{3t} &= 6e^{3t} + 9t e^{3t} \\ f'''(t) &= 18e^{3t} + 27t e^{3t} + 9e^{3t} &= 27e^{3t} + 27t e^{3t} = 27e^{3t}(1 + t) \end{aligned}$$

Como queremos estimar o valor de  $f(0.25)$  usando o polinômio do 2º grau, devemos tomar três pontos consecutivos na vizinhança de 0.25. Tomando então  $x_0 = 0.2$ ,  $x_1 = 0.3$  e  $x_2 = 0.4$ , obtemos:

$$\max_{0.2 \leq t \leq 0.4} |f'''(0.4)| = 27 e^{3(0.3)}(1 + 0.4) = 125.50042$$

Estamos, portanto, em condições de calcular um limitante superior para o erro de truncamento. Assim,

$$|R_2 f(f; x)| \leq \frac{|0.25 - 0.2| \cdot |0.25 - 0.3| \cdot |0.25 - 0.4|}{3!} \cdot (125.50042) \approx 0.0078 \approx 8 \times 10^{-3}.$$

Pelo resultado obtido, vemos que, se tomarmos um polinômio do 2º para avaliar  $f(0.25)$ , obteremos o resultado com duas casas decimais corretas.

Note que o número de zeros depois do ponto decimal, no resultado do erro, fornece o número de dígitos significativos corretos que teremos na aproximação.

### 4.1.3 Polinômio Interpolador de Newton

Nem sempre temos conhecimento, *a priori*, do grau adequado para o polinômio interpolador. Um teste razoável consiste em aumentar o número de pontos de interpolação, crescendo portanto o grau do polinômio interpolador, e testar se houve melhoria nos cálculos. Assim, é importante que o trabalho efetuado anteriormente possa ser reutilizado para estabelecer a expressão do novo polinômio. Note que, ao acrescentarmos um novo ponto ao polinômio de Lagrange, teremos de *refazer os cálculos*. Por outro lado, o *Polinômio de Newton* acrescenta um novo termo aos cálculos efetuados anteriormente.

Considere a forma do polinômio interpolador de Newton:

$$p_n(x) = d_0 + d_1(x - x_0) + d_2(x - x_0)(x - x_1) + \dots + d_n(x - x_0)(x - x_1) \dots (x - x_{n-1}) \quad (4.5)$$

onde cada  $d_i$ ,  $i = 0, 1, 2, \dots, n$  é conhecido como *operador diferenças divididas*.



### Operador Diferenças Divididas

**Definição 3.** [4] Sejam  $x_0, x_1, \dots, x_n$ ,  $n + 1$  pontos distintos no intervalo  $[a, b]$ , e sejam  $y_0, y_1, \dots, y_n$ ,  $n + 1$  valores de uma função  $y = f(x)$  sobre  $x = x_k$ ,  $k = 0, 1, \dots, n$ . Define-se:

$$\begin{aligned} f[x_k] &= f(x_k) & k = 0, 1, \dots, n \\ f[x_0, x_1, \dots, x_n] &= \frac{f[x_1, x_2, \dots, x_n] - f[x_0, x_1, \dots, x_{n-1}]}{x_n - x_0}, \end{aligned}$$

onde  $f[x_0, x_1, \dots, x_n]$  é a **diferença dividida** de ordem  $n$  da função  $f(x)$  sobre os pontos  $x_0, x_1, \dots, x_n$ .

Assim, o operador de diferenças divididas, denotado por  $\Delta$ , é definido como:

$$\begin{aligned} \text{Ordem 0: } \Delta^0 y_i &= y_i & &= f[x_i] \\ \text{Ordem 1: } \Delta y_i &= \frac{\Delta^0 y_{i+1} - \Delta^0 y_i}{x_{i+1} - x_i} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} & &= f[x_i, x_{i+1}] \\ \text{Ordem 2: } \Delta^2 y_i &= \frac{\Delta y_{i+1} - \Delta y_i}{x_{i+2} - x_i} & &= f[x_i, x_{i+1}, x_{i+2}] \\ &\dots & & \\ \text{Ordem } n: \Delta^n y_i &= \frac{\Delta^{n-1} y_{i+1} - \Delta^{n-1} y_i}{x_{i+n} - x_i} & &= f[x_i, x_{i+1}, x_{i+2}, \dots, x_{i+n}] \end{aligned}$$

### Cálculo Sistemático das Diferenças Divididas

Para calcular as diferenças divididas de uma função  $y = f(x)$  sobre os pontos  $x_0, x_1, \dots, x_n$ , construímos a tabela de diferenças divididas (Tabela 4.2).

$i$	$x_i$	$y_i = f[x_i]$	$\Delta y_i = f[x_i, x_{i+1}]$	$\Delta^2 y_i = f[x_i, x_{i+1}, x_{i+2}]$	$\dots$
0	$x_0$	$y_0 = f[x_0]$	$f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0}$	$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$	$\dots$
1	$x_1$	$y_1 = f[x_1]$	$f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1}$	$f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1}$	$\dots$
2	$x_2$	$y_2 = f[x_2]$	$f[x_2, x_3] = \frac{f[x_3] - f[x_2]}{x_3 - x_2}$	$f[x_2, x_3, x_4] = \frac{f[x_3, x_4] - f[x_2, x_3]}{x_4 - x_2}$	$\dots$
3	$x_3$	$y_3 = f[x_3]$	$f[x_3, x_4] = \frac{f[x_4] - f[x_3]}{x_4 - x_3}$	$\vdots$	$\dots$
4	$x_4$	$y_4 = f[x_4]$	$\vdots$	$\dots$	$\dots$

Tabela 4.2: Cálculo Sistemático

A Tabela 4.2 é construída da seguinte maneira:

1. A primeira coluna atribuímos os índices  $i$  das variáveis,  $i = 0, 1, \dots, n$ .
2. A segunda coluna é constituída dos pontos  $x_i$ ,  $i = 0, 1, \dots, n$ ;
3. A terceira coluna contém os valores de  $y_i = f(x_i)$ ,  $i = 0, 1, \dots, n$ ;
4. nas colunas 4, 5, 6,  $\dots$  estão as diferenças divididas de ordem 1, 2, 3,  $\dots$ . Cada uma destas diferenças é uma fração cujo numerador é sempre a diferença entre duas diferenças divididas consecutivas e de ordem imediatamente inferior, e cujo denominador é a diferença entre os dois extremos dos pontos envolvidos.

$x_i$	-2	-1	0	1	2
$f(x_i)$	-2	29	30	31	62

**Exemplo 24.** Para a seguinte função tabelada:  
construir a tabela de diferenças divididas.

**Solução:** Usando o esquema da Tabela 4.2, obtemos:

$i$	$x_i$	$y_i = f[x_i]$	$\Delta y_i$	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$
0	-2	-2	$\frac{29 - (-2)}{-1 - (-2)} = 31$	$\frac{1 - (-31)}{0 - (-2)} = -15$	$\frac{0 - (-15)}{1 - (-2)} = 5$	$\frac{5 - 5}{2 - (-2)} = 0$
1	-1	29	$\frac{30 - 29}{0 - (-1)} = 1$	$\frac{1 - 1}{1 - (-1)} = 0$	$\frac{15 - 0}{2 - (-1)} = 5$	
2	0	30	$\frac{31 - 30}{1 - 0} = 1$	$\frac{31 - 1}{2 - 0} = 15$		
3	1	31	$\frac{62 - 31}{2 - 1} = 31$			
4	2	62				

Note que o elemento 0 corresponde à diferença dividida:

$$\Delta^2 y_i = f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1} = \frac{1 - 1}{1 - (-1)} = 0.$$

Os resultados a serem utilizados na construção do polinômio de interpolação na fórmula de Newton são os primeiros valores em cada coluna de diferenças, embora tenhamos de construir toda a tabela, pois os valores são dependentes uns dos outros.

### Forma de Newton para o Polinômio Interpolador

Considere  $f(x)$  contínua e derivável no intervalo  $[a, b]$ , que os pontos  $x_0, x_1, \dots, x_n$  sejam distintos neste intervalo e que  $P(x_i) = y_i$  ( $P(x)$  um polinômio de grau  $n$ ). Pela definição de diferenças divididas, tem-se:

$$f[x_0, x] = \frac{f[x] - f[x_0]}{x - x_0}, \text{ definida em } [a, b] \text{ para } x \neq x_0 \text{ e } n = 1,$$

ou

$$f(x) = f[x_0] + f[x_0, x](x - x_0). \quad (4.6)$$

Para  $n = 2$ , procedemos

$$\begin{aligned} f[x_0, x_1, x] &= \frac{f[x_0, x] - f[x_0, x_1]}{x - x_1}, \text{ definida em } [a, b], \text{ para } x \neq x_0 \text{ e } x \neq x_1, \\ f[x_0, x_1, x](x - x_1) &= f[x_0, x] - f[x_0, x_1] \end{aligned}$$

Substituindo esta equação em (4.6), tem-se:

$$\begin{aligned} f[x_0, x_1, x](x - x_1) &= \frac{f[x] - f[x_0]}{x - x_0} - f[x_0, x_1] \\ f(x) &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x](x - x_0)(x - x_1). \end{aligned} \quad (4.7)$$

De maneira análoga, para  $n + 1$  pontos, temos:

$$f[x_0, x_1, \dots, x_n, x] = \frac{f[x_0, x_1, \dots, x_{n-1}, x] - f[x_0, x_1, \dots, x_n]}{x - x_n}, \text{ definida em } [a, b] \text{ para } x \neq n_k, k = 0, 1, \dots, n,$$

ou

$$\begin{aligned} f(x) &= \{f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ &+ f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2) + \dots \\ &+ f[x_0, x_1, \dots, x_n](x - x_0)(x - x_1) \dots (x - x_{n-1})\}_1 \\ &+ \{f[x_0, x_1, \dots, x_n, x](x - x_0)(x - x_1) \dots (x - x_{n-1})\}_2 \end{aligned} \quad (4.8)$$

onde, da fórmula (4.8), o termo  $\{\cdot\}_1$  é a fórmula de recorrência para  $f(x)$  e  $\{\cdot\}_2$  é o **termo do erro** ou **erro de truncamento**, que é o mesmo da forma de Lagrange (para maiores detalhes veja [4]).

Assim, o polinômio:

$$P_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \dots + f[x_0, x_1, \dots, x_n](x - x_0)(x - x_1) \dots (x - x_{n-1}) \quad (4.9)$$

é o polinômio de interpolação da função  $y = f(x)$  sobre os pontos  $x_0, x_1, \dots, x_n$ , conhecido como **Fórmula de Newton do Polinômio Interpolador**.

**Exemplo 25.** Use o polinômio de Newton para interpolar os pontos (1, 0), (2, 2), (4, 12) e (5, 20).

Organizando os cálculos através de uma tabela, temos:

$i$	$x_i$	$y_i$	$\Delta y_i$	$\Delta^2 y_i$	$\Delta^3 y_i$
0	1	0	2	1	0
1	2	2	5	1	
2	4	12	8		
3	5	20			

Tabela 4.3: Tabela de Diferenças Divididas.

Escrevendo o polinômio de Newton, temos:

$$\begin{aligned} P_3(x) &= y_0 + \Delta y_0(x - x_0) + \Delta^2 y_0(x - x_0)(x - x_1) + \Delta^3 y_0(x - x_0)(x - x_1)(x - x_2) \\ &= 0 + 2(x - 1) + 1(x - 1)(x - 2) + 0(x - 1)(x - 2)(x - 4) \\ &= x(x - 1) \end{aligned}$$

Note que, neste exemplo, apesar de usarmos quatro pontos na interpolação (portanto esperávamos obter um polinômio de grau 3), o resultado foi um polinômio de grau 2. Isto se deve aos quatro pontos estarem sobre uma parábola.

## 4.2 Quadrados Mínimos Lineares

O método dos Quadrados Mínimos é a técnica de aproximação mais usada na análise numérica e em problemas práticos, pois fornece uma aproximação para dados que são medidos com um certo grau de incerteza, buscando uma aproximação que minimiza os resíduos.

Seja  $f(x)$  uma função que será aproximada por  $g(x)$ . No caso dos *quadrados mínimos lineares* partimos da hipótese de que temos alguma informação sobre o comportamento de  $g(x)$ .

Suponha que os dados são aproximados por uma função do tipo:

$$f(x) \simeq g(x) = c_1\phi_1(x) + c_2\phi_2(x) + \dots c_n\phi_n(x) \quad (4.10)$$

onde as funções  $\phi_1(x)$ ,  $\phi_2(x)$ , ...,  $\phi_n(x)$  são preestabelecidas.

Para cada conjunto de coeficientes  $c_i$ ,  $i = 1 : n$ , o resíduo de (4.10), em  $x_k$ , será:

$$\begin{aligned} r(x_k) &= f(x_k) - g(x_k) \\ &= f(x_k) - [c_1\phi_1(x_k) + c_2\phi_2(x_k) + \dots c_n\phi_n(x_k)] \\ &= r(x_k; c_1, c_2, \dots, c_n). \end{aligned} \quad (4.11)$$

Assim, precisamos determinar  $c_1$ ,  $c_2$ , ...,  $c_n$  de maneira a *melhorar a aproximação*. A idéia do método dos quadrados mínimos é a de minimizar a soma dos quadrados dos resíduos, ou seja, minimizar:

$$\sum_{i=1}^m r^2(x_i) = \sum_{i=1}^m (f(x_i) - g(x_i))^2,$$

onde  $f(x_i)$ ,  $i = 1 : m$ , são os dados que serão aproximados.

Quando temos o caso de aproximar  $f(x_k)$ , uma função conhecida para todo  $x$ , por uma expressão mais simples  $g(x)$ , o método dos quadrados mínimos minimiza:

$$\int_a^b r^2(x) dx = \int_a^b (f(x) - g(x))^2 dx.$$

**Definição 4.** O produto escalar entre duas funções  $f(x)$  e  $g(x)$  é definido por:

$$\langle f, g \rangle = \begin{cases} \sum_{i=1}^m f(x_i)g(x_i), & \text{no caso discreto, isto é, se } f(x) \text{ e } g(x) \\ & \text{são conhecidos em } x_i, i = 1 : m. \\ \int_a^b f(x)g(x) dx, & \text{no caso contínuo, isto é, se } f(x) \text{ e } g(x) \\ & \text{são conhecidas e integráveis em } (a, b). \end{cases} \quad (4.12)$$

Note que, (4.12) possui as seguintes propriedades:

1.  $\langle \alpha_1 g_1 + \alpha_2 g_2, h \rangle = \alpha_1 \langle g_1, h \rangle + \alpha_2 \langle g_2, h \rangle$  – Linearidade.
2.  $\langle f, g \rangle = \langle g, f \rangle$  – Comutatividade.
3.  $\langle g, g \rangle \geq 0$  e  $\langle g, g \rangle = 0 \Leftrightarrow g \equiv 0$  – Positividade.

Usando (4.12), minimizando  $\langle r, r \rangle$ , que é uma função de  $c_1, c_2, \dots, c_n$ .

**Exemplo 26.** Regressão Linear (quadrados mínimos com uma reta)

Dada uma tabela com  $m$  valores  $(x_i, f(x_i))$ ,  $i = 1 : m$ , queremos encontrar a reta ( $f(x) \simeq g(x) = c_1\phi_1(x) + c_2\phi_2(x)$ ) que melhor ajusta esta tabela, no sentido dos quadrados mínimos. Tomando  $\phi_1(x) = 1$  e  $\phi_2(x) = x$ , temos:

$$f(x) \simeq g(x) = c_1 + c_2x.$$

O resíduo para cada par  $(c_1, c_2)$  e para cada  $x$ , será  $r(c_1, c_2; x) = f(x) - c_1 - c_2x$ . Assim, pelo método dos quadrados mínimos devemos procurar  $c_1$  e  $c_2$  que minimizem a função:

$$\langle r, r \rangle (c_1, c_2) = \langle f(x) - c_1 - c_2x, f(x) - c_1 - c_2x \rangle = \sum_{i=1}^m [f(x_i) - c_1 - c_2x_i]^2.$$

Do Cálculo Diferencial sabemos que a condição necessária do ponto crítico é que as derivadas nele se anulem, isto é,

$$\frac{\partial}{\partial c_1} \langle r, r \rangle = \frac{\partial}{\partial c_2} \langle r, r \rangle = 0,$$

ou ainda, procedidas as respectivas derivações com respeito a  $c_1$  e  $c_2$  em  $\langle r, r \rangle$  temos:

$$-\sum_{i=1}^m 2(f(x_i) - c_1 - c_2x_i) = 0 \text{ e } -\sum_{i=1}^m 2x_i(f(x_i) - c_1 - c_2x_i) = 0$$

Estas duas equações formam um sistema linear nas incógnitas  $c_1$  e  $c_2$ , que pode ser reescrito na forma:

$$\begin{cases} mc_1 + c_2 \sum_{i=1}^m x_i = \sum_{i=1}^m f(x_i) \\ c_1 \sum_{i=1}^m x_i + c_2 \sum_{i=1}^m x_i^2 = \sum_{i=1}^m x_i f(x_i) \end{cases}$$

Para ilustrar este procedimento, considere o ajuste da tabela abaixo por uma reta:

$x_i$	0	0.25	0.5	0.75	1.0
$f(x_i)$	1.000	1.2840	1.6487	2.1170	2.7183

Tabela 4.4: Dados para ajuste de curva.

Usando os valores da tabela temos:

$$\begin{aligned} \sum_{i=1}^5 x_i &= 2.5, & \sum_{i=1}^5 x_i^2 &= 1.875 \\ \sum_{i=1}^5 f(x_i) &= 8.768 & \sum_{i=1}^5 x_i f(x_i) &= 5.4514 \end{aligned}$$

Assim, os valores de  $c_1$  e  $c_2$  da melhor reta (no sentido dos quadrados mínimos) são obtidos pelo sistema:

$$\begin{aligned} 5c_1 + 2.5c_2 &= 8.768 \\ 2.5c_1 + 1.875c_2 &= 5.4514 \end{aligned}$$

Resolvendo o sistema, obtemos  $c_1 = 0.89968$  e  $c_2 = 1.70784$ , ou seja  $g(x) = 0.89968 + 1.70784x$ . Os valores de  $g(x_i)$  e os respectivos resíduos estão na tabela abaixo.

$x_i$	0	0.25	0.5	0.75	1.0
$f(x_i)$	1.000	1.2840	1.6487	2.1170	2.7183
$g(x_i)$	0.89968	1.32664	1.7536	2.18056	2.60752
$r(x_i)$	0.10032	-0.04264	-0.1049	-0.06356	0.11078

Neste exemplo, a soma dos quadrados do resíduo é:

$$\sum_{i=1}^5 r^2(x_i) = 0.039198$$

## 4.3 Interpolação com Splines

Se a função  $f(x)$  está tabelada em  $(n+1)$  pontos e a aproximarmos por um polinômio de grau  $n$  que a interpole sobre os pontos tabelados, o resultado dessa aproximação pode ser ruim quando  $n$  cresce.

Uma alternativa é interpolar  $f(x)$  em grupos de poucos pontos, obtendo-se polinômio de grau menor e impor condições para que a função de aproximação seja contínua e tenha derivadas contínuas até uma certa ordem. Para que isto aconteça, devemos respeitar as condições da Definição 5.

**Definição 5.** Considere a função  $f(x)$  tabelada nos pontos  $x_0 < x_1 < \dots < x_n$ . Uma função  $S_p(x)$  é denominada spline interpolante de grau  $p$  com nós nos pontos  $x_j$ ,  $j = 0, 1, \dots, n$ , se satisfaz as seguintes condições:

1. em cada subintervalo  $[x_j, x_{j+1}]$ ,  $j = 0, 1, \dots, (n-1)$ ,  $S_p(x)$  é um polinômio de grau  $p$ ;
2.  $S_p(x)$  é contínua e tem derivada contínua até ordem  $(p-1)$  em  $[a, b]$ .
3.  $S_p(x_j) = f(x_j)$ ,  $j = 0, 1, \dots, n$ .

A origem do nome spline vem de uma régua elástica, usada em desenhos de engenharia, que pode ser curvada de forma a passar por um dado conjunto de pontos  $(x_j, y_j)$ , que tem o nome de spline. Sob certas hipóteses (de acordo com a teoria da elasticidade) a curva definida pela régua pode ser descrita aproximadamente como sendo uma função por partes, cada qual um polinômio cúbico, de tal forma que ela e suas duas primeiras derivadas são contínuas sempre. A terceira derivada, entretanto, pode ter descontinuidades nos pontos  $x_j$ . Esta função é uma spline cúbica interpolante com nós nos pontos  $x_j$ , segundo a Definição 5.

Estudaremos aqui dois tipos de Interpolação por Splines: *Interpolação por Spline Linear* e *Interpolação por Spline Cúbica*.

### 4.3.1 Interpolação por Spline Linear

A interpolação por spline linear de  $f(x)$ ,  $S(x)$ , nos pontos  $x_1, x_2, \dots, x_n$  pode ser escrito na forma de subintervalos  $[x_{j-1}, x_j]$ ,  $j = 1, \dots, n$  como:

$$S_j(x) = f(x_{j-1}) \frac{x_j - x}{x_j - x_{j-1}} + f(x_j) \frac{x - x_{j-1}}{x_j - x_{j-1}}, \quad \forall x \in [x_{j-1}, x_j].$$

Assim:

1.  $S(x)$  é um polinômio de grau 1 em cada subintervalo  $[x_{j-1}, x_j]$ ;
2.  $S(x)$  é contínua em  $(x_{j-1}, x_j)$  e nos pontos  $x_j$ , uma vez que  $S(x)$  está bem definida, pois:  
 $S_j(x_j) = S_{j+1}(x_j) = f(x_j) \implies S(x)$  é contínua em  $[a, b]$  e, portanto,  $S(x)$  é uma spline linear;
3.  $S(x_j) = S_j(x_j) = f(x_j) \implies S(x)$  é a interpolação por spline linear de  $f(x)$  nos pontos  $x_0, x_1, \dots, x_n$ .

**Exemplo 27.** Achar a função de interpolação por spline linear de:

De acordo com a definição:

$x_i$	1.0	2.0	5.0	7.0
$f(x_i)$	1.0	2.0	3.0	2.5

$$\begin{aligned} S_1(x) &= f(x_0) \frac{x_1 - x}{x_1 - x_0} + f(x_1) \frac{x - x_0}{x_1 - x_0} \\ &= 1 \frac{2 - x}{2 - 1} + 2 \frac{x - 1}{2 - 1} = 2 - x + 2x - 2 = x \quad x \in [1, 2] \end{aligned}$$

$$\begin{aligned} S_2(x) &= f(x_1) \frac{x_2 - x}{x_2 - x_1} + f(x_2) \frac{x - x_1}{x_2 - x_1} \\ &= 2 \frac{5 - x}{5 - 2} + 3 \frac{x - 2}{5 - 2} = \frac{2}{3}(5 - x) + x - 2 = \frac{1}{3}(x + 4) \quad x \in [2, 5] \end{aligned}$$

$$\begin{aligned} S_3(x) &= f(x_2) \frac{x_3 - x}{x_3 - x_2} + f(x_3) \frac{x - x_2}{x_3 - x_2} \\ &= 3 \frac{7 - x}{7 - 5} + 2.5 \frac{x - 5}{7 - 5} = \frac{1}{2}(-0.5x + 8.5), \quad x \in [5, 7] \end{aligned}$$

ou ainda,

$$S(x) = \begin{cases} x & \text{se } x \in [1, 2] \\ \frac{1}{3}(x + 4) & \text{se } x \in [2, 5] \\ \frac{1}{2}(-0.5x + 8.5) & \text{se } x \in [5, 7] \end{cases}$$

### 4.3.2 Interpolação por Spline Cúbico

A spline linear apresenta a desvantagem de ter derivada primeira descontínua nos pontos  $x_j$ ,  $j = 1, 2, \dots, n$ . Se usarmos splines quadráticas, teremos que  $S_2(x)$  tem derivadas contínuas até a ordem 1 apenas e, portanto, a curvatura de  $S_2(x)$  pode trocar nos pontos  $x_j$ . Por esta razão, as splines cúbicas são mais usadas.

**Definição 6.** Dada uma função  $f(x)$  definida em  $[a, b]$  e um conjunto de nós  $a = x_0 < x_1 < \dots < x_n = b$ , um spline cúbico interpolador  $S(x)$  para  $f$  é uma função que satisfaz as seguintes condições:

1.  $S(x)$  é um polinômio cúbico, indicado por  $S_j(x)$ , no intervalo  $[x_j, x_{j+1}]$  para cada  $j = 0, 1, \dots, (n-1)$ ;
2.  $S(x_j) = f(x_j)$  para cada  $j = 0, 1, \dots, n$ ;
3.  $S_{j+1}(x_{j+1}) = S_j(x_{j+1})$  para cada  $j = 0, 1, \dots, (n-2)$ ;
4.  $S'_{j+1}(x_{j+1}) = S'_j(x_{j+1})$  para cada  $j = 0, 1, \dots, (n-2)$ ;
5.  $S''_{j+1}(x_{j+1}) = S''_j(x_{j+1})$  para cada  $j = 0, 1, \dots, (n-2)$ ;
6. Um dos seguintes conjuntos de condições de contorno é satisfeito:
  - (a)  $S''(x_0) = S''(x_n) = 0$  (condição de contorno livre ou natural);
  - (b)  $S'(x_0) = f'(x_0)$  e  $S'(x_n) = f'(x_n)$  (contorno restrito).

Para se construir o spline cúbico interpolador para uma função dada, as condições indicadas na definição são aplicadas aos polinômios cúbicos

$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3$$

para cada  $j = 0, 1, \dots, (n-1)$  (condição (1)).

Da condição (2), temos:

$$\begin{aligned} S_j(x_j) &= a_j + b_j(x_j - x_j) + c_j(x_j - x_j)^2 + d_j(x_j - x_j)^3 \\ &= a_j \\ S_j(x_j) &= a_j = f(x_j), \end{aligned}$$

aplicando a condição (3),

$$a_{j+1} = S_{j+1}(x_{j+1}) = S_j(x_{j+1}) = a_j + b_j(x_{j+1} - x_j) + c_j(x_{j+1} - x_j)^2 + d_j(x_{j+1} - x_j)^3$$

para cada  $j = 0, 1, \dots, (n-2)$ .

Denotando  $h_j = x_{j+1} - x_j$ , simplificamos a notação para cada  $j = 0, 1, \dots, (n-1)$ . Se definirmos  $a_n = f(x_n)$ , então a equação:

$$a_{j+1} = a_j + b_j h_j + c_j h_j^2 + d_j h_j^3 \quad (4.13)$$

é válida para cada  $j = 0, 1, 2, \dots, (n-1)$ .

Tomando  $S'_j(x)$ , temos:

$$S'_j(x) = b_j + 2c_j(x - x_j) + 3d_j(x - x_j)^2$$

implica que:

$$S'_j(x_j) = b_j + 2c_j(x_j - x_j) + 3d_j(x_j - x_j)^2 \implies S'_j(x_j) = b_j$$

para cada  $j = 0, 1, 2, \dots, (n-1)$ . Aplicando a condição (4), temos:

$$b_{j+1} = b_j + 2c_j h_j + 3d_j h_j^2 \quad (4.14)$$

para cada  $j = 0, 1, 2, \dots, (n-1)$ .

Sendo  $S''_j(x) = 2c_j + 6d_j(x - x_j)$ ,

$$\begin{aligned} S''_j(x_j) &= 2c_j + 6d_j(x_j - x_j) \\ &= 2c_j \\ \frac{S''_j(x_j)}{2} &= c_j \end{aligned}$$

Aplicando na condição (5), para cada  $j = 0, 1, \dots, (n-1)$ ,

$$\begin{aligned} 2c_{j+1} &= 2c_j + 6d_j h_j \\ c_{j+1} &= c_j + 3d_j h_j \end{aligned}$$

Isolando  $d_j$ :

$$d_j = \frac{c_{j+1} - c_j}{3h_j} \quad (4.15)$$

Substituindo a equação (4.15) nas equações (4.13) e (4.14), respectivamente, temos:

$$\begin{aligned} a_{j+1} &= a_j + b_j h_j + c_j h_j^2 + \left( \frac{c_{j+1} - c_j}{3h_j} \right) h_j^3 \\ &= a_j + b_j h_j + c_j h_j^2 + \left( \frac{c_{j+1} - c_j}{3} \right) h_j^2 \end{aligned} \quad (4.16)$$



$$\begin{aligned}
b_{j+1} &= b_j + 2c_j h_j + 3 \left( \frac{c_{j+1} - c_j}{3h_j} \right) h_j^2 \\
&= b_j + 2c_j h_j + (c_{j+1} - c_j) h_j \\
&= b_j + (c_{j+1} + c_j) h_j
\end{aligned} \tag{4.17}$$

Resolvendo a equação (4.16) para  $b_j$ , temos:

$$\begin{aligned}
a_{j+1} &= a_j + b_j h_j + c_j h_j^2 + \left( \frac{c_{j+1} - c_j}{3h_j} \right) h_j^3 \\
b_j h_j &= (a_{j+1} - a_j) - \frac{h_j^2}{3} (2c_j + c_{j+1}) \\
b_j &= \frac{1}{h_j} (a_{j+1} - a_j) - \frac{h_j}{3} (2c_j + c_{j+1})
\end{aligned} \tag{4.18}$$

Reduzindo os índices para  $b_{j-1}$ :

$$b_{j-1} = \frac{1}{h_{j-1}} (a_j - a_{j-1}) - \frac{h_{j-1}}{3} (2c_{j-1} + c_j)$$

Substituindo esses valores na equação (4.17), com os índices reduzidos, temos o sistema linear de equações:

$$\begin{aligned}
b_j &= b_{j-1} + (c_j + c_{j-1}) h_{j-1} \\
\frac{1}{h_j} (a_{j+1} - a_j) - \frac{h_j}{3} (2c_j + c_{j+1}) &= \frac{1}{h_{j-1}} (a_j - a_{j-1}) - \frac{h_{j-1}}{3} (2c_{j-1} + c_j) + (c_j + c_{j-1}) h_{j-1} \\
h_{j-1} c_{j-1} + 2c_j (h_{j-1} + h_j) + h_j c_{j+1} &= \frac{3}{h_j} (a_{j+1} - a_j) - \frac{3}{h_{j-1}} (a_j - a_{j-1})
\end{aligned} \tag{4.19}$$

para cada  $j = 1, 2, \dots, (n-1)$ . Esse sistema envolve apenas  $\{c_j\}_{j=0}^n$  como incógnitas, na medida em que os valores de  $\{h_j\}_{j=0}^{n-1}$  e  $\{a_j\}_{j=0}^n$  são dados, respectivamente, pelo espaçamento dos nós  $\{x_j\}_{j=0}^n$  e os valores de  $f$  nos nós.

Note que, uma vez que os valores de  $\{c_j\}_{j=0}^n$  são determinados, essa é uma maneira simples de encontrar as constantes  $\{b_j\}_{j=0}^{n-1}$  remanescentes na equação (4.18) e  $\{d_j\}_{j=0}^n$  da equação (4.15), e de se construir os polinômios cúbicos  $\{S_j(x)\}_{j=0}^{n-1}$ .

A questão maior que fica em relação a essa construção é se os valores de  $\{c_j\}_{j=0}^n$  podem ser encontrados utilizando-se o sistema de equações (4.19) e, se a resposta for positiva, se esses valores são únicos. Os teoremas a seguir indicam esse é o caso quando qualquer das condições de contorno dadas na parte (6) da Definição 6 é imposta.

**Teorema 9.** *Se  $f$  é definida em  $a = x_0 < x_1 < \dots < x_n = b$ , então  $f$  tem um único spline interpolador natural (ou livre)  $S$  nos nós  $x_0, x_1, \dots, x_n$ ; isto é, um spline interpolador que satisfaz as condições de contorno  $S''(a) = 0$  e  $S''(b) = 0$ .*

**Prova 6.** *As condições de contorno nesse caso implicam que  $c_n = S''(x_n)/2 = 0$  e que:*

$$0 = S''(x_0) = 2c_0 + 6d_0(x_0 - x_0);$$

e portanto  $c_0 = 0$ .

As duas condições  $c_0 = 0$  e  $c_n = 0$ , juntamente com as equações (4.19), produzem um sistema linear descrito

pela equação vetorial  $\mathbf{Ax} = \mathbf{b}$ , onde  $\mathbf{A}$  é uma matriz  $(n+1) \times (n+1)$ :

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ h_0 & 2(h_0 + h_1) & h_1 & \ddots & \ddots & \vdots \\ 0 & h_1 & 2(h_1 + h_2) & h_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & h_{n-2} & 2(h_{n-2} + h_{n-1}) & h_{n-1} \\ 0 & \cdots & \cdots & 0 & 0 & 1 \end{bmatrix}$$

e  $\mathbf{b}$  e  $\mathbf{x}$  são os vetores:

$$\mathbf{b} = \begin{bmatrix} 0 \\ \frac{3}{h_1}(a_2 - a_1) - \frac{3}{h_0}(a_1 - a_0) \\ \vdots \\ \frac{3}{h_{n-1}}(a_n - a_{n-1}) - \frac{3}{h_{n-2}}(a_{n-1} - a_{n-2}) \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{x} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

A matriz  $\mathbf{A}$  é diagonalmente dominante, e portanto, tem solução única para  $c_0, c_1, \dots, c_n$ .

**Teorema 10.** Se  $f$  é definida em  $a = x_0 < x_1 < \dots < x_n = b$ , e diferenciável em  $a$  e  $b$ , então  $f$  tem um spline interpolador restrito  $S$  nos nós  $x_0, x_1, \dots, x_n$ ; isto é, um spline interpolador que satisfaz as condições de contorno  $S'(a) = f'(a)$  e  $S'(b) = f'(b)$ .

**Prova 7.** Como  $f'(a) = S'(a) = S'(x_0) = b_0$ , aplicando-se a equação (4.18), com  $j = 0$ , temos:

$$\begin{aligned} f'(a) &= \frac{1}{h_0}(a_1 - a_0) - \frac{h_0}{3}(2c_0 + c_1) \\ 3f'(a) &= \frac{3}{h_0}(a_1 - a_0) - h_0(2c_0 + c_1) \\ 2h_0c_0 + h_0c_1 &= \frac{3}{h_0}(a_1 - a_0) - 3f'(a) \end{aligned}$$

Do mesmo modo (equação (4.17)):

$$f'(b) = b_n = b_{n-1} + h_{n-1}(c_{n-1} + c_n)$$

de maneira que a equação (4.18) nos permite dizer que  $j = n - 1$  implica que:

$$\begin{aligned} f'(b) &= \frac{a_n - a_{n-1}}{h_{n-1}} - \frac{h_{n-1}}{3}(2c_{n-1} + c_n) + h_{n-1}(c_{n-1} + c_n) \\ f'(b) &= \frac{a_n - a_{n-1}}{h_{n-1}} + \frac{h_{n-1}}{3}(c_{n-1} + 2c_n) \\ 3f'(b) &= \frac{3}{h_{n-1}}(a_n - a_{n-1}) + h_{n-1}(c_{n-1} + 2c_n) \\ h_{n-1}c_{n-1} + 2h_{n-1}c_n &= 3f'(b) - \frac{3}{h_{n-1}}(a_n - a_{n-1}) \end{aligned}$$

A equação (4.19), juntamente com as equações:

$$\begin{aligned} 2h_0c_0 + h_0c_1 &= \frac{3}{h_0}(a_1 - a_0) - 3f'(a) \\ h_{n-1}c_{n-1} + 2h_{n-1}c_n &= 3f'(b) - \frac{3}{h_{n-1}}(a_n - a_{n-1}) \end{aligned}$$

determinam o sistema linear  $\mathbf{Ax} = \mathbf{b}$ , onde

$$\mathbf{A} = \begin{bmatrix} 2h_0 & h_0 & 0 & \cdots & \cdots & 0 \\ h_0 & 2(h_0 + h_1) & h_1 & \ddots & \ddots & \vdots \\ 0 & h_1 & 2(h_1 + h_2) & h_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & h_{n-2} & 2(h_{n-2} + h_{n-1}) & h_{n-1} \\ 0 & \cdots & \cdots & 0 & h_{n-1} & 2h_{n-1} \end{bmatrix}$$

e  $\mathbf{b}$  e  $\mathbf{x}$  são os vetores:

$$\mathbf{b} = \begin{bmatrix} \frac{3}{h_0}(a_1 - a_0) - 3f'(a) \\ \frac{3}{h_1}(a_2 - a_1) - \frac{3}{h_0}(a_1 - a_0) \\ \vdots \\ \frac{3}{h_{n-1}}(a_n - a_{n-1}) - \frac{3}{h_{n-2}}(a_{n-1} - a_{n-2}) \\ 3f'(b) - \frac{3}{h_{n-1}}(a_n - a_{n-1}) \end{bmatrix} \quad e \quad \mathbf{x} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

A matriz  $\mathbf{A}$  é diagonalmente dominante, e portanto, tem solução única para  $c_0, c_1, \dots, c_n$ .

**Exemplo 28.** Construa os splines cúbicos naturais, com  $S''(x_0) = S''(x_n) = 0$ , para:

$x_i$	0.1	0.2	0.3	0.4
$y_i$	-0.62049958	-0.28398668	0.00660095	0.2484244

Como os intervalos estão igualmente espaçados,  $h_j = 0.1$ , para  $j = 0, 1, 2, 3$ . Assim, usando o Teorema 9, temos o seguinte sistema tridiagonal linear:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.1 & 0.4 & 0.1 & 0 \\ 0 & 0.1 & 0.4 & 0.1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1.377581 \\ -1.4629254 \\ 0 \end{bmatrix}$$

Resolvendo o sistema acima, temos  $\mathbf{c}^T = [0, -2.698738, -2.982629, 0]^T$ .

Como  $a_j = f(x_j)$ , para  $j = 0, 1, 2, 3$ , podemos determinar  $b_j$  e  $d_j$ , para  $j = 0, 1, 2$  usando as equações (4.18) e (4.15), respectivamente, como mostra a Tabela 28.

A partir da Tabela 28, podemos montar a função spline:

$$S(x) = \begin{cases} -0.62049958 + 3.4550869(x - 0.1) - 8.9957933(x - 0.1)^3 & x \in [0.1, 0.2) \\ -0.28398668 + 3.1852131(x - 0.2) - 2.698738(x - 0.2)^2 - 0.9463033(x - 0.2)^3 & x \in [0.2, 0.3) \\ 0.00660095 + 2.6170764(x - 0.3) - 2.982629(x - 0.3)^2 + 9.9420967(x - 0.3)^3 & x \in [0.3, 0.4] \end{cases}$$

$a_j$	$b_j$	$c_j$	$d_j$
-0.62049958	3.4550869	0	-8.9957933
-0.28398668	3.1852131	-2.698738	-0.9463033
0.00660095	2.6170764	-2.982629	9.9420967
0.24842440	—	0	—

Tabela 4.5: Coeficientes para os splines naturais.

**Exemplo 29.** Construa os splines cúbicos restritos, com  $f'(0.1) = 3.58502082$  e  $f'(0.4) = 2.16529366$ , para os dados do exemplo anterior.

Como os intervalos estão igualmente espaçados,  $h_j = 0.1$ , para  $j = 0, 1, 2, 3$ . Assim, usando o Teorema 10, temos o seguinte sistema tridiagonal linear:

$$\begin{bmatrix} 0.2 & 0.1 & 0 & 0 \\ 0.1 & 0.4 & 0.1 & 0 \\ 0 & 0.1 & 0.4 & 0.1 \\ 0 & 0 & 0.1 & 0.2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -0.6596755 \\ -1.377581 \\ -1.4629254 \\ -0.7588225 \end{bmatrix}$$

Resolvendo o sistema acima, temos  $\mathbf{c}^T = [-2.1498408, -2.29703, -2.4394481, -2.5743885]^T$ .

Como  $a_j = f(x_j)$ , para  $j = 0, 1, 2, 3$ , podemos determinar  $b_j$  e  $d_j$ , para  $j = 0, 1, 2$  usando as equações (4.18) e (4.15), respectivamente, como mostra a Tabela 29.

$a_j$	$b_j$	$c_j$	$d_j$
-0.62049958	3.5850208	-2.1498408	-0.4907741
-0.28398668	3.1403294	-2.297073	-0.4745836
0.00660095	2.6666773	-2.4394481	-0.4498015
0.24842440	—	-2.5743885	—

Tabela 4.6: Coeficientes para os splines restrito.

A partir da Tabela 29, podemos montar a função spline:

$$S(x) = \begin{cases} -0.62049958 + 3.5850208(x - 0.1) - 2.1498408(x - 0.1)^2 - 0.4907741(x - 0.1)^3 & x \in [0.1, 0.2) \\ -0.28398668 + 3.1403294(x - 0.2) - 2.297073(x - 0.2)^2 - 0.4745836(x - 0.2)^3 & x \in [0.2, 0.3) \\ 0.00660095 + 2.6666773(x - 0.3) - 2.4394481(x - 0.3)^2 - 0.4498015(x - 0.3)^3 & x \in [0.3, 0.4] \end{cases}$$

## 4.4 Exercícios

1. Obtenha o polinômio interpolador de Lagrange e Newton para as seguintes funções :

- (a)  $f(x) = e^{2x} \cos 3x$ ,  $x_0 = 0$ ,  $x_1 = 0.3$  e  $x_2 = 0.6$ ,  $n = 2$ ;
- (b)  $f(x) = \sin(\ln x)$ ,  $x_0 = 2.0$ ,  $x_1 = 2.4$  e  $x_2 = 2.6$ ,  $n = 2$ ;
- (c)  $f(x) = \ln x$ ,  $x_0 = 1$ ,  $x_1 = 1.1$ ,  $x_2 = 1.3$  e  $x_3 = 1.4$ ,  $n = 3$ ;
- (d)  $f(x) = \cos x + \sin x$ ,  $x_0 = 0$ ,  $x_1 = 0.25$ ,  $x_2 = 0.5$  e  $x_3 = 1.0$ ,  $n = 3$ .

2. Utilize o polinômio interpolador de Lagrange e de Newton de grau 1, 2 e 3 para aproximar cada um dos seguintes itens:

- (a)  $f(8.4)$  se  $f(8.1) = 16.94410$ ,  $f(8.3) = 17.56492$ ,  $f(8.6) = 18.50515$  e  $f(8.7) = 18.82091$ ;
- (b)  $f(-\frac{1}{3})$  se  $f(-0.75) = -0.07181250$ ,  $f(-0.5) = -0.02475000$ ,  $f(-0.25) = 0.33493750$  e  $f(0) = 1.10100000$ ;

- (c)  $f(0.25)$  se  $f(0.1) = 0.62049958$ ,  $f(0.2) = -0.28398668$ ,  $f(0.3) = 0.00660095$  e  $f(0.4) = 0.24842440$ ;  
 (d)  $f(0.9)$  se  $f(0.6) = -0.17694460$ ,  $f(0.7) = 0.01375227$ ,  $f(0.8) = 0.22363362$  e  $f(1.0) = 0.65809197$ .
3. Seja  $p_3(x)$  o polinômio interpolador para os dados  $(0, 0)$ ,  $(0.5, y)$ ,  $(1, 3)$  e  $(2, 2)$ . Encontre  $y$  para o caso em que o coeficiente de  $x^3$  em  $P_3(x)$  é 6 (use o polinômio interpolador de Lagrange).
4. Utilize os seguintes valores e aritmética com arredondamento de quatro dígitos para obter uma aproximação para  $f(1.09)$ , utilizando um polinômio de Lagrange e um de Newton de 3º grau. A função que será aproximada é  $f(x) = \log_{10}(\operatorname{tg} x)$ . Utilize esse conhecimento para encontrar um limite para o erro na aproximação:

$$f(1.00) = 0.1924 \quad f(1.05) = 0.2414 \quad f(1.10) = 0.2933 \quad f(1.15) = 0.3492$$

5. Seja  $f(x) = 7x^5 - 3x^2 - 1$ .
- (a) Calcular  $f(x)$  nos pontos  $x = 0$ ,  $x = \pm 1$ ,  $x = \pm 2$  e  $x = \pm 3$ . Construir a seguir a tabela segundo os valores crescente de  $x$ .
- (b) Construir o polinômio de interpolação para esta função sobre os pontos  $-2$ ,  $-1$ ,  $0$  e  $1$ .
- (c) Determinar, pela fórmula (4.4), um limitante para o erro de truncamento em  $x = -0.5$  e  $x = 0.5$ .

Resposta:

(b)  $P_3(x) = \frac{104}{3}x^3 - 3x^2 - \frac{83}{3}x - 1$ .

(c)  $|R_3(-0.5)| \leq 39.375$  e  $|R_3(0.5)| \leq 65.625$ .

6. Conhecendo-se a tabela: calcular um limitante superior para o erro de truncamento quando calculamos

$x$	0.8	0.9	1.0	1.1	1.3	1.5
$\cos x$	0.6967	0.6216	0.5403	0.4536	0.2675	0.0707

$\cos 1.05$  usando polinômio de interpolação sobre quatro pontos.

Resposta: Considerando os pontos  $x_0 = 1.0$ ,  $x_1 = 1.1$ ,  $x_2 = 1.3$  e  $x_3 = 1.5$ , temos  $|R_3(x)| \approx 1.453 \times 10^{-6}$ .

7. Apresentando os dados

$x_i$	4.0	4.2	4.5	4.7	5.1	5.5	5.9	6.3	6.8	7.1
$y_i$	102.56	113.18	130.11	142.05	167.53	195.14	224.87	256.73	299.50	326.72

- (a) Construa o polinômio de quadrados mínimos de grau 1 e calcule o erro.  
 (b) Construa o polinômio de quadrados mínimos de grau 2 e calcule o erro.

Resposta: (a)  $y = 72.0845x - 194.138$ , (b)  $y = 6.61821x^2 - 1.14352x + 1.23556$ .

8. Repita o exercício anterior para os seguintes dados:

$x_i$	0.2	0.3	0.6	0.9	1.1	1.3	1.4	1.6
$y_i$	0.050446	0.098426	0.33277	0.72660	1.0972	1.5697	1.8487	2.5015

9. Encontre os polinômios de quadrados mínimos de graus 1, 2 e 3 para os dados apresentados na tabela abaixo. Calcule o resíduo para cada caso (erro).

Resposta:  $y_1 = 0.620895 + 1.219621x$ ,  $y_2 = 0.5965807 + 1.253293x - 0.01085343x^2$  e  $y_3 = 0.6290193 + 1.18501x + 0.03533252x^2 - 0.01004723x^3$ .

10. Transforme os modelos abaixo em relações lineares:

$x_i$	1.0	1.1	1.3	1.5	1.9	2.1
$y_i$	1.84	1.96	2.21	2.45	2.94	3.18

(a)  $y = \frac{a}{b + cx}$

(b)  $y = ab^x$

(c)  $y = \frac{a}{b + x}$

(d)  $y = \frac{1}{1 + e^{bx}}$

11. Suponha que num laboratório obtivemos experimentalmente os seguintes valores para  $f(x)$  sobre os pontos  $x_i$ ,  $i = 1 : 8$ :

$x_i$	-1.0	-0.7	-0.4	-0.1	0.2	0.5	0.8	1.0
$f(x_i)$	36.547	17.264	8.155	3.852	1.820	0.860	0.406	0.246

- (a) Faça o diagrama da dispersão dos dados;  
 (b) Ajuste os dados, usando o Método dos Quadrados Mínimos, para a função  $y = c_1 e^{-c_2 x}$ ;  
 (c) Calcule o resíduo da sua aproximação.

Resposta:  $y = 3.001 e^{-2.5x}$ .

12. Ajuste os dados:

$x$	-8	-6	-4	-2	0	2	4
$y$	30	10	9	6	5	4	4

- (a) usando a aproximação  $y = \frac{1}{(c_0 + c_1 x)}$ . Faça o gráfico para  $1/y$  e verifique que esta aproximação é viável;  
 (b) idem para  $y = ab^x$ ;  
 (c) compare os resultados (a) e (b).

Resposta: (a)  $y = \frac{1}{0.195 + 0.0185x}$ , (b)  $y = 5.5199(0.8597)^x$ .

13. O número de bactérias, por unidade de volume, existente em uma cultura após  $x$  horas é apresentado na tabela:

nº de horas ( $x$ )	0	1	2	3	4	5	6
nº de bactérias por vol. unitário ( $y$ )	32	47	65	92	132	190	275

- (a) Verifique que uma curva para se ajustar ao diagrama de dispersão é do tipo exponencial;  
 (b) Ajuste aos dados as curvas  $y = ab^x$  e  $y = ax^b$ ; compare os valores obtidos por meio destas equações com os dados experimentais;

Resposta:  $y = 32.14685(1.42696)^x$  e  $y = 38.83871x^{0.963}$ .

- (c) Avalie da melhor forma o valor de  $y(x)$  para  $x = 7$ .

14. Determine o spline cúbico livre  $S$  que interpole os dados  $f(0) = 0$ ,  $f(1) = 1$  e  $f(2) = 2$ .

Resposta:  $S(x) = x$ , em  $[0, 2]$ .

15. Determine o spline cúbico restrito  $S$  que interpole os dados  $f(0) = 0$ ,  $f(1) = 1$  e  $f(2) = 2$  e satisfaz  $s'(0) = s'(2) = 1$ .

16. Construa os splines cúbicos livres para os dados que se seguem.

- (a)  $f(8.3) = 17.56492$  e  $f(8.6) = 18.50515$ ;
- (b)  $f(0.8) = 0.22363362$  e  $f(1.0) = 0.65809197$ ;
- (c)  $f(-0.5) = -0.02475$ ,  $f(0.25) = 0.3349375$  e  $f(0) = 1.101$
17. Os dados no Exercício acima (14) foram gerados utilizando-se as funções a seguir. Use os splines cúbicos construídos para os valores dados de  $x$ , para aproximar  $f(x)$  e  $f'(x)$ , e calcule o erro absoluto.
- (a)  $f(x) = x \ln x$ ; aproxime  $f(8.4)$  e  $f'(8.4)$ ;
- (b)  $f(x) = \sin(e^x - 2)$ ; aproxime  $f(0.9)$  e  $f'(0.9)$ ;
- (c)  $f(x) = x^3 + 4.001x^2 + 4.002x + 1.101$ ; aproxime  $f(-1/3)$  e  $f'(1/3)$ ;
18. Construa os splines cúbicos restritos utilizando os dados do Exercício (14) e o fato de que:
- (a)  $f'(8.3) = 1.116256$  e  $f'(8.6) = 1.151762$ ;
- (b)  $f'(0.8) = 2.1691753$  e  $f'(1.0) = 2.0466965$ ;
- (c)  $f'(-0.5) = 0.751$  e  $f'(0) = 4.002$ ;
19. Repita o Exercício (15) utilizando os splines gerados no Exercício (16).
20. Um spline cúbico natural  $S$  em  $[0, 2]$  é definido por:

$$S(x) = \begin{cases} S_0(x) &= 1 + 2x - x^3 & \text{se } 0 \leq x < 1 \\ S_1(x) &= 2 + b(x-1) + c(x-1)^2 + d(x-1)^3 & \text{se } 1 \leq x \leq 2 \end{cases}$$

Encontre  $b$ ,  $c$  e  $d$ .

Resposta:  $b = -1$ ,  $c = -3$  e  $d = 1$ .

21. Um spline cúbico restrito  $S$  para uma função  $f$  é definido em  $[1, 3]$  por:

$$S(x) = \begin{cases} S_0(x) &= 3(x-1) + 2(x-1)^2 - (x-1)^3 & \text{se } 1 \leq x < 2 \\ S_1(x) &= a + b(x-2) + c(x-2)^2 + d(x-2)^3 & \text{se } 2 \leq x \leq 3 \end{cases}$$

Dados  $f'(1) = f'(3)$ , encontre  $a$ ,  $b$ ,  $c$  e  $d$ .

22. Construa um spline cúbico livre para aproximar  $f(x) = \cos(\pi x)$  utilizando os valores dados por  $f(x)$  em  $x = 0, 0.25, 0.5, 0.75, 1.0$ . Integre o spline em  $[0, 1]$  e compare o resultado com  $\int_0^1 \cos(\pi x) dx = 0$ . Use as derivadas do spline para aproximar  $f'(0.5)$  e  $f''(0.5)$ . Compare essas aproximações com os resultados reais.
23. Repita o exercício acima, construindo desta vez um spline cúbico restrito com  $f'(0) = f'(1) = 0$ .
24. Dada a partição  $x_0 = 0$ ,  $x_1 = 0.05$  e  $x_2 = 0.1$  de  $[0, 0.1]$ , encontre a função interpoladora linear para  $f(x) = e^{2x}$ . Aproxime  $\int_0^{0.1} e^{2x} dx$  com  $\int_0^{0.1} S(x) dx$ , e compare os resultados com os valores reais.

---

# Integração Numérica

---

## Introdução

*Por que Integração Numérica ?* Isto é: por que não restringir o cálculo de integrais ao uso das técnicas de integração estudadas no Cálculo Diferencial e Integral ? A resposta para essa questão tem por base dois fatos:

1. Geralmente em problemas envolvendo o cálculo de integrais não se conhece a expressão analítica da função integrando, somente os valores dessa função, o que inviabiliza o uso das técnicas de integração do Cálculo, mas que são os dados necessários para a integração numérica;
2. Mesmo quando se conhece a expressão analítica da função integrando, o cálculo da função primitiva pode ser trabalhoso e nem sempre simples. Por exemplo, a integral

$$\int e^{-x^2} dx$$

resulta em uma função que não pode ser expressa em termos de combinações finitas de outras funções algébricas, logarítmicas ou exponenciais.

A idéia básica da integração numérica reside na aproximação da função integrando por um polinômio. As fórmulas de integração são somatórios cujas parcelas são valores da função  $f(x)$  calculados em pontos e multiplicados por pesos convenientemente escolhidos. Assim, vamos procurar desenvolver fórmulas de integração do tipo:

$$\int_a^b f(x) dx \cong \sum_{i=0}^n w_i f(x_i), \quad (5.1)$$

onde  $a \leq x_0 < x_1 < \dots < x_n \leq b$  são chamados *pontos de integração* e  $w_i$  são os pesos da fórmula de integração.

## 5.1 Fórmula de Newton-Cotes

Neste caso, os pontos de integração são igualmente espaçados em  $(a, b)$ , tal que  $h = \frac{b-a}{n}$ , onde  $n$  é um número inteiro. Os pontos de integração (para este caso) são:

$$x_j = a + jh \quad j = 0 : n.$$



Considere agora o polinômio de Lagrange de grau  $n$  que interpola os  $(n + 1)$  pontos  $(x_i, f(x_i))$ ,  $i = 0 : n$

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

Integrando esta última expressão no intervalo  $(a, b)$ , temos:

$$\int_a^b f(x) l(x) dx = \sum_{i=0}^n f(x_i) \int_a^b l_i(x) dx.$$

Assim, de (5.1), o cálculo de  $w_i$  é obtido pela integração de  $l_i(x)$ , isto é:

$$w_i = \int_a^b l_i(x) dx = \int_a^b \frac{(x - x_0) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)} dx. \quad (5.2)$$

Através da equação (5.2) podemos obter fórmulas do tipo *Newton-Cotes* para polinômios de qualquer grau.

### 5.1.1 Fórmula dos Trapézios: $n = 1$

A fórmula dos trapézios corresponde à interpolação da função a ser integrada por um polinômio de grau  $n = 1$ . Como a interpolação linear pede 2 pontos, tomaremos os extremos do intervalo de integração, isto é,  $a = x_0$  e  $b = x_1$ .

A expressão (5.2) nos permite encontrar os pesos da regra dos trapézios:

$$\begin{aligned} w_0 &= \int_{x_0}^{x_1} \frac{(x - x_1)}{(x_0 - x_1)} dx = \frac{1}{-h} \int_{x_0}^{x_1} (x - x_1) dx = - \frac{(x - x_1)^2}{2h} \Big|_{x=x_0}^{x=x_1} = \frac{h}{2} \\ w_1 &= \int_{x_0}^{x_1} \frac{(x - x_0)}{(x_1 - x_0)} dx = \frac{1}{h} \int_{x_0}^{x_1} (x - x_0) dx = \frac{(x - x_0)^2}{2h} \Big|_{x=x_0}^{x=x_1} = \frac{h}{2} \end{aligned}$$

Com isso, podemos estabelecer a *fórmula dos trapézios* para a integração no intervalo  $(x_0, x_1)$ :

$$\int_{x_0}^{x_1} f(x) dx = \frac{h}{2} [f(x_0) + f(x_1)]. \quad (5.3)$$

### 5.1.2 Fórmula de Simpson: $n = 2$

Para estabelecer a *fórmula de Simpson*, interpolamos  $f(x)$  usando um polinômio de grau 2 que coincide com essa função nos pontos  $x_0, x_1$  e  $x_2$ . Assim, tomamos  $n = 2$ ,  $x_0 = a$ ,  $x_1 = \frac{(a+b)}{2}$  e  $x_2 = b$  em (5.2). Integrando os polinômios de grau 2, estabelecemos os pesos da *fórmula de Simpson*:

$$\begin{aligned} w_0 &= \int_{x_0}^{x_2} \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} dx = \frac{h}{3} \\ w_1 &= \int_{x_0}^{x_2} \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} dx = \frac{4h}{3} \\ w_2 &= \int_{x_0}^{x_2} \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} dx = \frac{h}{3} \end{aligned}$$

O cálculo das integrais acima podem ser simplificados lembrando que:

$$x_1 - x_0 = h, \quad x_2 - x_1 = h, \quad x_2 - x_0 = 2h$$

e usando a substituição de variáveis  $t = x - x_1$  e portanto  $t + h = x - x_1 + h = x - x_0$ ,  $t - h = x - x_1 - h = x - x_2$ . Por exemplo, fazendo estas substituições no cálculo do peso  $w_0$  temos:

$$w_0 = \int_{x_0}^{x_1} \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} dx = \int_{-h}^h \frac{t(t - h)}{2h^2} dt = \frac{h}{3}$$

Dessa maneira, usando polinômios interpoladores de grau 2, estabelecemos a fórmula de Simpson:

$$\int_{x_0}^{x_2} f(x) dx = \left(\frac{h}{3}\right) [f(x_0) + 4f(x_1) + f(x_2)]. \quad (5.4)$$

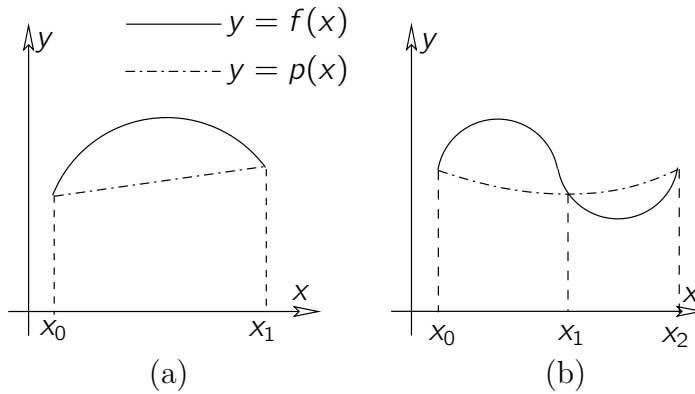


Figura 5.1: Aproximação de integral por trapézio e Simpson.

A Figura 5.1(a) mostra a área sob a curva aproximada pela área do trapézio e a Figura 5.1(b), a área sob a parábola que passa pelos pontos  $(x_0, f(x_0))$ ,  $(x_1, f(x_1))$  e  $(x_2, f(x_2))$ .

### 5.1.3 Fórmulas de Newton-Cotes para $n = 3$ e $n = 4$

- $n = 3$

$$\int_{x_0}^{x_3} f(x) l(x) dx = \left(\frac{3h}{8}\right) [f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)] \quad (5.5)$$

- $n = 4$

$$\int_{x_0}^{x_4} f(x) l(x) dx = \left(\frac{2h}{45}\right) [7f(x_0) + 32f(x_1) + 12f(x_2) + 32f(x_3) + 7f(x_4)] \quad (5.6)$$

**Exemplo 30.** Sabemos que

$$\ln 2 = \int_1^2 \frac{1}{x} dx \cong 0.69314718.$$

Use as fórmulas de Newton-Cotes apresentadas (fórmulas (5.3), (5.4), (5.5) e (5.6)) para obter aproximações para  $\ln 2$ . Calcule o erro absoluto de cada aproximação.

## 5.2 Fórmulas Repetidas

Divida o intervalo de integração  $[a, b]$  em  $n$  subintervalos de igual comprimento  $h = \frac{(b-a)}{n}$ . Sejam  $x_0 = a$ ,  $x_i = x_{i-1} + h$  e  $x_n = b$ . Podemos aplicar a regra dos trapézios para cada um dos subintervalos. Assim, lembrando que  $x_i - x_{i-1} = h$ , e as propriedades de integrais, temos:

$$\begin{aligned} \int_a^b f(x) dx &= \int_{x_0}^{x_1} f(x) dx + \int_{x_1}^{x_2} f(x) dx + \int_{x_2}^{x_3} f(x) dx \dots + \int_{x_{n-1}}^{x_n} f(x) dx \\ &\cong \left(\frac{h}{2}\right) [f(x_0) + 2f(x_1) + 2f(x_2) \dots + 2f(x_{n-1}) + f(x_n)]. \end{aligned} \quad (5.7)$$

Se optarmos por aplicar a fórmula de Simpson repetida, devemos repartir o intervalo num número par de subintervalos, uma vez que cada parábola requer três pontos de interpolação. Assim, se  $n$  é um número par:

$$\begin{aligned} \int_a^b f(x) dx &= \int_{x_0}^{x_2} f(x) dx + \int_{x_2}^{x_4} f(x) dx + \dots + \int_{x_{n-2}}^{x_n} f(x) dx \\ &= \left(\frac{h}{3}\right) [f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \dots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)] \\ &= \frac{h}{3} \{f(x_0) + 4[f(x_1) + f(x_3) + \dots + f(x_{n-1})] + \\ &\quad 2[f(x_2) + f(x_4) + \dots + f(x_{n-2})] + f(x_n)\} \end{aligned} \quad (5.8)$$

**Exemplo 31.** Ainda calculando aproximações para  $\ln 2$ , aplique as fórmulas (5.7) e (5.8) no intervalo  $[1, 2]$  e  $h = 0.25$ . Calcule o erro absoluto para cada aproximação.

## 5.3 Erro nas Fórmulas de Newton-Cotes

**Teorema 11. (Ímpar)** Se os pontos  $x_j = x_0 + jh$ ,  $j = 0 : n$ , dividem  $[a, b]$  ( $x_0 = a$  e  $x_n = b$ ) em um número ímpar de intervalos iguais e  $f(x)$  tem derivada de ordem  $(n+1)$  contínua em  $[a, b]$  então a expressão do erro para as fórmulas de Newton-Cotes do tipo fechado, com  $n$  ímpar, é dada por:

$$R(f) = \frac{h^{n+2} f^{(n+1)}(\xi)}{(n+1)!} \int_0^n u(u-1) \dots (u-n) du$$

para algum ponto  $\xi \in [a, b]$ .

**Teorema 12. (Par)** Se os pontos  $x_j = x_0 + jh$ ,  $j = 0 : n$ , dividem  $[a, b]$  ( $x_0 = a$  e  $x_n = b$ ) em um número par de intervalos iguais e  $f(x)$  tem derivada de ordem  $(n+2)$  contínua em  $[a, b]$ , então a expressão do erro para as fórmulas de Newton-Cotes do tipo fechado, com  $n$  par, é dada por:

$$R(f) = \frac{h^{n+3} f^{(n+2)}(\xi)}{(n+2)!} \int_0^n \left(u - \frac{n}{2}\right) u(u-1) \dots (u-n) du$$

para algum ponto  $\xi \in [a, b]$ .

### 5.3.1 Erro na Fórmula do Trapézio

Considere o intervalo  $[x_0, x_1]$ , ou seja,  $n = 1$ . Usando o Teorema 11, temos:

$$\begin{aligned} R(f) &= \frac{h^3 f''(\xi)}{2!} \int_0^1 [u(u-1)] du = \frac{h^3 f''(\xi)}{2!} \int_0^1 (u^2 - u) du \\ &= \frac{h^3 f''(\xi)}{2!} \left[ \frac{u^3}{3} - \frac{u^2}{2} \right]_0^1 = \frac{h^3 f''(\xi)}{2!} \left( -\frac{1}{6} \right) \\ &= -\frac{h^3 f''(\xi)}{12} \end{aligned}$$

Assim,

$$\int_{x_0}^{x_1} f(x) dx = \frac{2}{h} [f(x_0) + f(x_n)] - \frac{h^3}{12} f''(\xi), \quad x_0 < \xi < x_1$$

O erro na fórmula do trapézio repetido é obtido adicionando-se  $N$  erros na fórmula  $R(f)$  acima, onde  $N = \frac{b-a}{h}$ . Logo,

$$\begin{aligned} \int_{x_0}^{x_n} f(x) dx &= \left( \frac{h}{2} \right) [f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)] - \frac{Nh^3}{12} f''(\xi), \\ &= \left( \frac{h}{2} \right) [f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)] - \frac{(b-a)h^2}{12} f''(\xi), \quad x_0 < \xi < x_n \end{aligned}$$

Note que o termo do erro **não** será, na prática, subtraído do resultado aproximado; assim nunca conseguiremos o resultado exato, pois o ponto  $\xi$  que fornece a igualdade é único, mas não há como determiná-lo. A aplicação da fórmula do termo do resto é útil quando queremos o resultado com precisão pré-fixada.

**Exemplo 32.** Determinar o número de intervalos em que podemos dividir  $[0, 1.2]$  para obter

$$\int_0^{1.2} e^x \cos x dx$$

pela regra do trapézio com três casas decimais.

**Solução:**  $R(f) = -\frac{Nh^3}{12} \max_{0 \leq t \leq 1.2} |f''(t)|$ ,  $x_0 < \xi < x_n$ .

Calculando:

$$f'(t) = e^t \cos t - e^t \sin t = e^t (\cos t - \sin t)$$

$$f''(t) = e^t (\cos t - \sin t) + e^t (-\sin t - \cos t) = -2e^t \sin t$$

$$\therefore \max_{0 \leq t \leq 1.2} |f''(t)| = |f''(1.2)| = 2(3.320)(0.932) = 6.188$$

$$\text{Na regra do trapézio: } h = \frac{b-a}{N} = \frac{1.2-0}{N} = \frac{1.2}{N}.$$

Impondo erro  $\leq 0.5 \times 10^{-3}$ , temos:

$$R(f) \leq \frac{1.2h^2}{12} (6.188) \leq 0.0005 \implies h^2 \leq 0.0000808 \implies h \leq 0.02842$$

Observe que, devemos escolher o menor  $h$  que seja menor ou igual a 0.02842, mas que divida exatamente o intervalo  $[0, 1.2]$ . Assim, tomamos  $h = 0.025$ :

$$N = \frac{1.2}{0.025} \implies N = 48$$

### 5.3.2 Erro na Fórmula de Simpson

Para obtermos o erro na fórmula de Simpson, sobre o intervalo  $[x_0, x_2]$  fazemos  $n = 2$  no Teorema 12. Assim:

$$\begin{aligned} R(f) &= \frac{h^{2+3} f^{(2+2)}(\xi)}{(2+2)!} \int_0^2 \left(u - \frac{2}{2}\right) u(u-1)(u-2) du \\ &= \frac{h^5 f^{(IV)}(\xi)}{4!} \int_0^2 (u^4 - 4u^3 + 5u^2 - 2u) du \\ &= \frac{h^5 f^{(IV)}(\xi)}{4!} \left[ \frac{u^5}{5} - u^4 + \frac{5u^3}{3} - u^2 \right]_0^2 \\ &= \frac{h^5 f^{(IV)}(\xi)}{4!} \left( -\frac{4}{15} \right) = \frac{-h^5 f^{(IV)}(\xi)}{90} \end{aligned}$$

Então, podemos escrever:

$$\int_{x_0}^{x_2} f(x) dx = \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] - \frac{h^5}{90} f^{(IV)}(\xi), \quad x_0 < \xi < x_2.$$

O erro na fórmula de Simpson Repetido é obtido adicionando-se  $N$  erros da fórmula acima, onde  $N = \frac{b-a}{2h}$ . Assim:

$$\begin{aligned} \int_{x_0}^{x_n} f(x) dx &= \frac{h}{3} [f(x_0) + 4f(x_1) + 2f(x_2) + \cdots + 2f(x_{2N-2}) + 4f(x_{2N-1}) + f(x_{2N})] - \frac{Nh^5}{90} f^{(IV)}(\xi), \\ \int_{x_0}^{x_n} f(x) dx &= \frac{h}{3} [f(x_0) + 4f(x_1) + 2f(x_2) + \cdots + 2f(x_{2N-2}) + 4f(x_{2N-1}) + f(x_{2N})] - \frac{(b-a)h^4}{180} f^{(IV)}(\xi), \end{aligned}$$

para  $x_0 < \xi < x_n$ .

Note que:

- Comparando as expressões de erro, vemos que a fórmula de Simpson repetido é da ordem de  $h^4$  (em símbolos,  $O(h^4)$ ), enquanto que a regra do trapézio repetido é  $O(h^2)$ . Assim, a regra de Simpson possui uma ordem de convergência maior, resultando em erros de aproximação menores para um mesmo  $h$ .
- Para obter o resultado da integral com uma determinada precisão, podemos utilizar a fórmula do erro impondo, em módulo, seja inferior a  $0.5 \times 10^{-k}$ , onde  $k$  é o número de casas decimais corretas que desejamos no resultado, e assim obter o número de intervalos necessários, ou ir aumentando o número de pontos e comparando dois resultados consecutivos até obter a precisão desejada. Na prática, é mais comum usarmos esta segunda possibilidade.

**Exemplo 33.** Usando a Regra de Simpson Repetido, obter a integral do Exemplo 32, com duas casas decimais corretas.

**Solução:** Inicialmente, calculamos a integral usando 3 (três) pontos:

$$I_3 = \int_0^{1.2} e^x \cos x dx = \frac{1}{3} h [f(0) + 4f(0.6) + f(1.2)] = \frac{0.6}{3} [1 + 4(1.503) + 1.202] = 1.6428$$

Agora calculamos a integral com 5 (cinco) pontos:

$$I_5 = \int_0^{1.2} e^x \cos x dx = \frac{0.3}{3} [f(0) + 4f(0.3) + 2f(0.6) + 4f(0.9) + f(1.2)] = 1.6464$$

Calculando o erro relativo:

$$ER = \frac{|I_5 - I_e|}{|I_5|} = \frac{|1.6464 - 1.6428|}{|1.6464|} = 0.0022 < 10^{-2}$$

Portanto, o valor da integral com duas casas decimal de precisão é 1.6464.

O Exemplo 33 ilustra o procedimento computacional a ser implementado. Logo, devemos tomar  $h \rightarrow 0$ , ou seja, devemos fazer  $h = 0.1, 0.01, \dots$  e ir comparando os resultados obtidos através do erro relativo, isto é, se

$$ER = \frac{|I_r - I_s|}{|I_r|} < \epsilon$$

onde  $I_r$  e  $I_s$  são dois resultados consecutivos, e  $\epsilon$  uma precisão pré-fixada, então paramos o processo.

## 5.4 Integração de Romberg

As fórmulas de Newton-Cotes podem ser combinadas de modo a melhorar a aproximação da integração numérica. Uma técnica eficiente para estabelecer esta combinação é através da *integração de Romberg*, que é baseada na *extrapolação de Richardson*.

### 5.4.1 Extrapolação de Richardson

Inicialmente, seja  $g$  uma função com derivadas de todas as ordens em algum intervalo contendo  $x_0$  como um ponto interior. Então a expansão da *série de Taylor* gerada por  $g$  em  $x = x_0$  é:

$$\begin{aligned} g(x) &= \sum_{k=0}^{\infty} \frac{g^{(k)}(x_0)}{k!} (x - x_0)^k \\ &= g(x_0) + g'(x_0)(x - x_0) + \frac{g''(x_0)}{2!} (x - x_0)^2 + \dots + \frac{g^n(x_0)}{n!} (x - x_0)^n + \dots \end{aligned} \quad (5.9)$$

Seja  $g(0)$  um valor a ser calculado, e  $g(h)$  sua aproximação, calculada em função do parâmetro  $h$  (isto é,  $h = x - x_0$ ). Usando a equação (5.9) em torno de zero, de  $g(h)$  e  $g(h/2)$ , teremos, respectivamente:

$$\begin{aligned} g(h) &= g(0) + g'(0)h + g''(0)\frac{h^2}{2} + g'''(0)\frac{h^3}{3!} + \dots \\ g\left(\frac{h}{2}\right) &= g(0) + g'(0)\frac{h}{2} + g''(0)\frac{h^2}{8} + g'''(0)\frac{h^3}{48} + \dots \end{aligned}$$

Tomando os dois primeiros termos do lado direito destas duas expressões, vemos que  $g(h)$  e  $g(h/2)$  são aproximações para  $g(0)$ . Combinando convenientemente estes dois somatórios definimos uma aproximação mais precisa:

$$\begin{aligned} g_1(h) &= 2g\left(\frac{h}{2}\right) - g(h) = \frac{1}{2^1 - 1} \left[ 2^1 g\left(\frac{h}{2}\right) - g(h) \right] \\ &= \left[ 2g(0) + 2g'(0)\frac{h}{2} + 2g''(0)\frac{h^2}{8} + 2g'''(0)\frac{h^3}{48} + \dots \right] - \\ &\quad \left[ g(0) + g'(0)h + g''(0)\frac{h^2}{2} + g'''(0)\frac{h^3}{6} \right] = \\ &= g(0) - g''(0)\frac{h^2}{4} - 3g'''(0)\frac{h^3}{24} + \dots \end{aligned}$$

Se  $g_1(h)$  for usada para calcular  $g(0)$ , então este será mais preciso. Repetindo a idéia, tomamos agora  $g_1(h)$  e  $g_1(h/2)$  para definir outra aproximação  $g_2(h)$ . Sejam:

$$\begin{aligned} g_1(h) &= g(0) - g''(0)\frac{h^2}{4} - 3g'''(0)\frac{h^3}{24} + \dots \\ g_1\left(\frac{h}{2}\right) &= g(0) - g''(0)\frac{h^2}{16} - 3g'''(0)\frac{h^3}{192} + \dots \end{aligned}$$

Subtraindo a primeira equação da segunda, multiplicada por 4, teremos:

$$4g_1\left(\frac{h}{2}\right) - g_1(h) = 3g(0) + 3g'''(0)\frac{h^3}{48} + \dots$$

Dessa maneira:

$$\begin{aligned} g_2(h) &= \frac{1}{3} \left[ 4g_1\left(\frac{h}{2}\right) - g_1(h) \right] = \frac{1}{2^2 - 1} \left[ 2^2 g_1\left(\frac{h}{2}\right) - g_1(h) \right] \\ &= g(0) + g'''(0)\frac{h^3}{48} + \dots \end{aligned} \quad (5.10)$$

Em resumo, a idéia é aplicar sucessivamente combinações de aproximações para melhorar a aproximação final. Pode-se verificar que a expressão genérica do  $n$ -ésimo estágio deste procedimento é:

$$g_n(h) = \frac{2^n g_{n-1}(h/2) - g_{n-1}(h)}{2^n - 1} = \frac{1}{2^n - 1} \left[ 2^n g_{n-1}\left(\frac{h}{2}\right) - g_{n-1}(h) \right] \quad (5.11)$$

## 5.4.2 Integração de Romberg

Chamemos  $I_1(h)$  e  $I_1(h/2)$  os resultados obtidos pela regra dos trapézios tomando  $2h$  e  $h$ , respectivamente, em  $(x_0, x_2)$ , como ilustra a Figura 5.2.

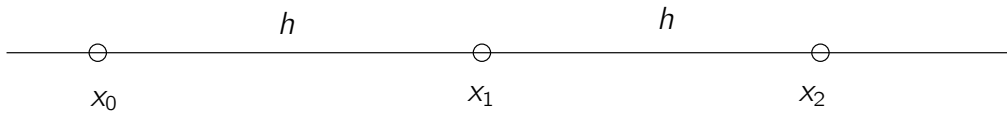


Figura 5.2: Intervalo  $(x_0, x_2)$ .

Aplicando a fórmula dos trapézios simples e repetida temos:

$$I_1(h) = \frac{2h}{2} [f(x_0) + f(x_2)] \quad \text{e} \quad I_1\left(\frac{h}{2}\right) = \frac{h}{2} [f(x_0) + 2f(x_1) + f(x_2)].$$

Usando a fórmula de extrapolação (5.10), temos:

$$\begin{aligned} I_2(h) &= \frac{1}{3} \left[ 4I_1\left(\frac{h}{2}\right) - I_1(h) \right] \\ &= \frac{1}{3} \left[ \frac{4h}{2} [f(x_0) + 2f(x_1) + f(x_2)] - h [f(x_0) + f(x_2)] \right] \\ &= \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] \end{aligned}$$

Note que, a expressão de  $I_2(h)$  corresponde a fórmula de Simpson aplicada no intervalo  $(x_0, x_2)$ , que é uma

aproximação melhor do que a fórmula do Trapézio. Se quisermos uma aproximação melhor, podemos usar a fórmula geral (5.11).

**Exemplo 34.** Calcule a integral abaixo usando a Extrapolação de Richardson com  $n = 4$ .

$$\int_0^1 e^{-x^2} dx$$

## 5.5 Quadratura de Gauss-Legendre

Segundo [2], o fato de escolher pontos igualmente espaçados nas fórmulas de Newton-Cotes certamente simplifica os cálculos. Contudo, se as abscissas não tiverem esta imposição de espaçamento constante, então podem ser obtidas fórmulas que forneçam uma maior exatidão.

### 5.5.1 Fórmula para dois pontos

A integração de uma função  $f(x)$  pela regra do Trapézio é baseada em um polinômio de grau 1 que passa pelos pontos  $a$  e  $b$ , do intervalo de integração  $[a, b]$ . Porém, pode-se escolher dois pontos  $c$  e  $d$ , de maneira que a área do trapézio seja a mais próxima da área sob a curva, sendo  $c \in [a, b]$  e  $d \in [a, b]$ .

Com este objetivo é feita, inicialmente, uma mudança de variável de  $x$  para  $t$ , definida num intervalo  $[-1, 1]$ , por intermédio da expressão

$$x = x(t) = \frac{b-a}{2}t + \frac{a+b}{2} \quad (5.12)$$

de modo que

$$dx = \frac{b-a}{2} dt$$

Definindo

$$F(t) = \frac{b-a}{2} f(x(t)) \quad (5.13)$$

então

$$\int_a^b f(x) dx = \int_{-1}^1 \frac{b-a}{2} F(t) \frac{b-a}{2} dt \rightarrow \int_a^b f(x) dx = \int_{-1}^1 F(t) dt.$$

Considere o ponto  $c$ , com coordenadas  $c[t_1, F(t_1)]$  e o ponto  $d$  com coordenadas  $d[t_2, F(t_2)]$ ; então deseja-se que,

$$\int_a^b f(x) dx = \int_{-1}^1 F(t) dt \approx I_2 = A_1 F(t_1) + A_2 F(t_2) \quad (5.14)$$

expressão que é análoga à regra do trapézio

$$\int_a^b f(x) dx \approx \frac{h}{2} [f(a) + f(b)].$$

O problema consiste em encontrar valores  $t_1$ ,  $t_2$ ,  $A_1$  e  $A_2$  que tornem a exatidão a maior possível. Para isto, o método é construído de modo a ser exato para polinômios de grau até 3. Fazendo

$$F(t) = t^k, \quad k = 0, 1, 2, 3,$$

e impondo (5.14) ser igual à integral analítica de  $F(t)$ , então:

para  $k = 0$ :

$$F(t) = 1 \rightarrow \int_{-1}^1 1 dt = 1 - (-1) = 2 = A_1 + A_2$$



para  $k = 1$ :

$$F(t) = t \rightarrow \int_{-1}^1 t dt = \frac{t^2}{2} \Big|_{-1}^1 = \frac{1}{2} - \frac{1}{2} = 0 = A_1 t_1 + A_2 t_2$$

para  $k = 2$ :

$$F(t) = t^2 \rightarrow \int_{-1}^1 t^2 dt = \frac{t^3}{3} \Big|_{-1}^1 = \frac{1}{3} - \left(-\frac{1}{3}\right) = \frac{2}{3} = A_1 t_1^2 + A_2 t_2^2$$

para  $k = 3$ :

$$F(t) = t^3 \rightarrow \int_{-1}^1 t^3 dt = \frac{t^4}{4} \Big|_{-1}^1 = \frac{1}{4} - \frac{1}{4} = 0 = A_1 t_1^3 + A_2 t_2^3$$

As expressões acima constituem um sistema de equações não lineares de ordem 4:

$$\begin{cases} A_1 + A_2 = 2 \\ A_1 t_1 + A_2 t_2 = 0 \\ A_1 t_1^2 + A_2 t_2^2 = \frac{2}{3} \\ A_1 t_1^3 + A_2 t_2^3 = 0 \end{cases}$$

cuja solução é

$$t_1 = \frac{-1}{\sqrt{3}} \approx -0.5774, \quad t_2 = \frac{1}{\sqrt{3}} \approx 0.5774, \quad A_1 = 1 \quad \text{e} \quad A_2 = 1.$$

**Exemplo 35.** Calcular  $\int_1^5 (2x^3 + 3x^2 + 6x + 1)dx$  usando a Quadratura de Gauss (fórmula (5.14)).

**Solução:** Inicialmente, realizamos uma mudança de variável, de  $x$  para  $t$ , através da equação (5.12):

$$x_i = \frac{b-a}{2}t_i + \frac{a+b}{2} = \frac{5-1}{2}t_i + \frac{1+5}{2} \rightarrow x_i = 2t_i + 3$$

Por (5.13), a alteração da função em  $x$  para uma função em  $t$ :

$$F(t_i) = \frac{b-a}{2}f(x_i) = \frac{5-1}{2}f(x_i) \rightarrow F(t_i) = 2f(2t_i + 3),$$

ou seja,

$$F(t_i) = 2[2(2t_i + 3)^3 + 3(2t_i + 3)^2 + 6(2t_i + 3) + 1]$$

Os cálculos podem ser sistematizados em uma tabela composta por 4 (quatro) colunas, como mostra a Tabela (5.1).

$i$	$t_i$	$F(t_i)$	$A_i$
1	$-\frac{1}{\sqrt{3}}$	69.7083	1
2	$\frac{1}{\sqrt{3}}$	442.2917	1

Tabela 5.1: Aproximação para  $n = 2$ .

$$I_2 = A_1 F(t_1) + A_2 F(t_2) = 1 \cdot 69.7083 + 1 \cdot 442.2917 \rightarrow I_2 = 512.000$$

Este resultado é exato porque  $\int_1^5 (2x^3 + 3x^2 + 6x + 1)dx = \left(\frac{x^4}{2} + x^3 + 3x^2 + x\right) \Big|_1^5 = 517.5 - 5.5 = 512$ .

**Exemplo 36.** Calcular  $\int_0^\pi (e^x + \sin x + 2) dx$ , usando Quadratura Gaussiana, para  $n = 2$ .

Por (5.12),

$$x_i = \frac{b-a}{2}t_i + \frac{a+b}{2} = \frac{\pi-0}{2}t_i + \frac{0+\pi}{2} \rightarrow x_i = \frac{\pi}{2}(t_i + 1)$$

Usando (5.13),

$$F(t_i) = \frac{b-a}{2}f(x_i) = \frac{\pi-0}{2}f(x_i) \rightarrow F(t_i) = \frac{\pi}{2}f\left(\frac{\pi}{2}(t_i + 1)\right)$$

Temos a Tabela (5.2).

$i$	$t_i$	$F(t_i)$	$A_i$
1	$-\frac{1}{\sqrt{3}}$	7.1605	1
2	$\frac{1}{\sqrt{3}}$	22.8236	1

Tabela 5.2: Aproximação para  $n = 2$ .

Assim,

$$I_2 = A_1F(t_1) + A_2F(t_2) = 1 \cdot 7.1605 + 1 \cdot 22.8236 = 29.9841$$

O valor exato desta integral é  $\int_0^\pi (e^x + \sin x + 2) dx = (e^x - \cos x + 2x)\Big|_0^\pi = 30.4239$ .

Portanto, o erro cometido pela Quadratura Gaussiana com 2 pontos foi de  $|30.4239 - 29.9841| = 0.4398$ , que é mais exato que o obtido pela regra do trapézio repetido com 6 subintervalos, equivalente a 7 pontos.

## 5.6 Exercícios

- Use a regra do Trapézio (Tr) ( $n = 1$ ) e a regra de Simpson (Sp) ( $n = 2$ ) para aproximar as seguintes integrais:

(a)  $\int_{0.5}^1 x^4 dx;$

(Resposta: (Tr) 0.265625, (Sp) 0.1940104)

(b)  $\int_0^{0.5} \frac{2}{x-4} dx;$

(Resposta: (Tr) -0.2678571, (Sp) -0.2670635)

(c)  $\int_1^{1.5} x^2 \ln x dx;$

(Resposta: (Tr) -0.17776434, (Sp) 0.1922453)

(d)  $\int_0^1 x^2 e^{-x} dx;$

(Resposta: (Tr) 0.1839397, (Sp) 0.16240168)

(e)  $\int_1^{1.6} \frac{2x}{x^2-4} dx;$

(Resposta: (Tr) -0.8666667, (Sp) -0.7391053)

(f)  $\int_0^{0.35} \frac{2}{x^2-4} dx;$

(Resposta: (Tr) -0.1777643, (Sp) -0.1768216)

(g)  $\int_0^{\pi/4} x \sin x dx;$

(Resposta: (Tr) 0.2180895, (Sp) 0.1513826)

(h)  $\int_0^{\pi/4} e^{3x} \sin 2x dx;$

(Resposta: (Tr) 4.1432597, (Sp) 2.5836964)

- Utilize a regra Trapezoidal Composta (Tc) e Simpson Composta (Sc) com os valores indicados de  $n$  para aproximar as seguintes integrais:

(a)  $\int_1^2 x \ln x dx, n = 4;$

(Resposta: (Tc) 0.639900)

(b)  $\int_{-2}^2 x^3 e^x dx, n = 4;$

(Resposta: (Tc) 31.3653)

$$(c) \int_0^2 \frac{2}{x^2 + 4} dx, n = 6;$$

(Resposta: (Tc) 0.784241)

$$(d) \int_0^\pi x^2 \cos x dx, n = 6;$$

(Resposta: (Tc) -6.42872)

$$(e) \int_0^2 e^{2x} \sin 3x dx, n = 8;$$

(Resposta: (Tc) -13.5760)

$$(f) \int_0^2 \frac{1}{x^2 + 4} dx, n = 8;$$

(Resposta: (Tc) 0.476977)

$$(g) \int_3^5 \frac{x}{\sqrt{x^2 - 4}} dx, n = 8;$$

(Resposta: (Tc) 0.605498)

$$(h) \int_0^{3\pi/8} \tan x dx, n = 8;$$

(Resposta: (Tc) 0.970926)

3. Suponha que  $f(0) = 1$ ,  $f(0.5) = 2.5$ ,  $f(1) = 2$  e  $f(0.25) = f(0.75) = \alpha$ . Determine  $\alpha$  se a regra Trapezoidal Composta com  $n = 4$  dá 1.75 para  $\int_0^1 f(x) dx$ .

4. Utilize a integração de Romberg para calcular as seguintes integrais, com  $n = 3$ :

$$(a) \int_1^{1.5} x^2 \ln x dx; \text{ (Resposta: 0.1922593)}$$

$$(e) \int_0^{\pi/4} e^{3x} \sin 2x dx; \text{ (Resposta: 2.5879685)}$$

$$(b) \int_0^1 x^2 e^{-x} dx; \text{ (Resposta: 0.1606105)}$$

$$(f) \int_1^{1.6} \frac{2x}{x^2 - 4} dx; \text{ (Resposta: -0.7341567)}$$

$$(c) \int_0^{0.35} \frac{2}{x^2 - 4} dx; \text{ (Resposta: -0.1768200)}$$

$$(g) \int_3^{3.5} \frac{x}{\sqrt{x^2 - 4}} dx; \text{ (Resposta: 0.6362135)}$$

$$(d) \int_0^{\pi/4} x^2 \sin x dx; \text{ (Resposta: 0.08875677)}$$

$$(h) \int_0^{\pi/4} (\cos x)^2 dx; \text{ (Resposta: 0.6426970)}$$

5. Utilize os seguintes dados para aproximar  $\int_1^2 f(x) dx$  usando a integração de Romberg.

$x_i$	1.0	1.25	1.5	1.75	2.0
$f(x_i)$	1.0	0.8	0.6666	0.5714	0.5

6. Use a regra de Simpson com  $n = 8$  para aproximar o valor médio de  $f$  no intervalo dado.

$$(a) f(x) = \frac{1}{x^4 + 1}, [0, 4]. \text{ (Resposta: 0.28)}$$

$$(b) f(x) = \sqrt{\cos x}, [-1, 1]$$

7. Para controlar a poluição térmica de um rio, um biólogo registra a temperatura (em °F) a cada hora, de 9h da manhã às 5h da tarde. Os dados constam da tabela abaixo:

Hora	9	10	11	12	1	2	3	4	5
Temperatura	75.3	77.0	83.1	84.8	86.5	86.4	81.1	78.6	75.1

Use a regra de Simpson e o Teorema do Valor Médio (Cálculo) para estimar a temperatura média da água entre 9h da manhã e 5h da tarde.

8. Calcule  $\int_0^{10} \sqrt{x^4} dx$  usando:

(a) a regra do Trapézio com  $n = 5$ ; (Resposta: 341.36)

(b) a regra de Simpson com  $n = 8$ ; (Resposta: 334.42)

(c) a regra de Romberg com  $n = 4$ .

**Atenção:** Use aproximações com quatro casas decimais para  $f(x_k)$  e arredonde as respostas para duas casas decimais.

9. Aproxime a integral imprópria fazendo a substituição  $u = 1/x$  e aplicando então a regra de Simpson com  $n = 4$ .

(a)  $\int_2^{\infty} \frac{1}{\sqrt{x^4 + x}} dx$  (Resposta: 0.49)

(c)  $\int_{-\infty}^{-10} \frac{\sqrt{|x|}}{x^3 + 1} dx$

(b)  $\int_1^{-\infty} e^{-x^2} dx$  (Resposta: 0.14)

(d)  $\int_1^{\infty} e^{-x} \sin \sqrt{x} dx$

10. Obtenha uma aproximação de  $\int_0^1 \frac{\cos x}{\sqrt{x}} dx$  fazendo a substituição  $u = \sqrt{x}$  e aplicando a regra do trapézio com  $n = 4$ . (Resposta: 1.79)

11. Determine  $h$  de modo que a Fórmula do Trapézio forneça o valor de:

$$I = \int_0^1 e^{-x^2} dx$$

com erro inferior a  $0.5 \times 10^{-6}$ . (Resposta:  $h < 0.00245 \implies h = 0.002$ .)

12. Achar o número mínimo de intervalos que se pode usar para, utilizando a Fórmula de Simpson, obter:

$$\int_0^{\frac{\pi}{2}} e^{-x} \cos x dx,$$

com quatro casas decimais corretas. (Resposta:  $N = 10$  intervalos.)

13. Calcular:

(a)  $\int_{-1}^1 (z^3 + z^2 + z + 1) dz$

(b)  $\int_{-2}^0 (x^2 - 1) dx$

por quadratura gaussiana e diretamente. Calcule o erro absoluto.

(Resposta: (a)  $\approx 2.6667$  e  $\frac{8}{3}$ ; (b)  $\approx 0.6667$  e  $\frac{2}{3}$ .)

14. Calcular  $\int_0^1 e^x dx$ , usando a regra do trapézio com  $n = 1$  e Quadratura Gaussiana com  $n = 2$ , compare os resultados obtidos com o valor analítico.

15. Considere a integral  $I = \int_0^1 e^{-x^2} dx$ .

(a) Estime  $I$  pela regra de Simpson usando  $h = 0.25$ ; (Resposta:  $I = 0.746855$ .)

(b) Estime  $I$  por Quadratura Gaussiana com 2 pontos; (Resposta:  $I = 0.746594$ .)

(c) Sabendo que o valor exato de  $I$  (com 5 casas decimais) é 0.74682, pede-se:

i. compare as estimativas (a) e (b);

ii. quantos pontos seriam necessários para que a regra dos Trapézios obtivesse a mesma precisão que a estimativa para  $I$  em (b) ? (Resposta:  $m = 27$ .)

---

# Aproximações para Equações Diferenciais Ordinárias

---

## Introdução

**O que é uma Equação Diferencial Ordinária (EDO) ?** É uma relação que envolve uma “função incógnita” e suas derivadas ou diferenciais. Por exemplo:

1.  $y'(x) = f(x)$ , onde  $y'(x)$  denota  $\frac{dy}{dx}$ ;
2.  $y''(x) + y(x) = 0$ ;
3.  $y^{(3)}(x) + (\sin x)y''(x) + 5xy(x) = 0$ .

São equações cujas incógnitas são funções de uma variável e suas derivadas (caso unidimensional).

**Qual é a ordem de uma EDO ?** É a ordem da mais alta derivada da função incógnita.

**Qual é a solução de uma EDO ?** É uma *função* definida num intervalo que, juntamente com suas derivadas, satisfaz a equação diferencial dada.

**Tipos de problemas com EDO** Para definir uma única função  $y(x)$ , temos de fornecer dados adicionais à equação diferencial; estes dados definem dois tipos de problemas:

1. *Problema de valor inicial* (PVI): procuramos  $y(x)$  que, além de satisfazer a equação diferencial para  $x > a$ , atende às condições preestabelecidas no início do intervalo onde vamos resolvê-la:

$$y(a) = y_0 \quad \text{e, às vezes,} \quad y'(a) = v_0,$$

onde  $y_0$  e  $v_0$  são valores conhecidos.

2. *Problema de valor de contorno* (PVC): a equação diferencial deverá ser satisfeita no intervalo  $(a, b)$  e são preestabelecidos valores para  $y(x)$  nos extremos deste intervalo:

$$y(a) = y_0 \quad \text{e} \quad y(b) = y_n,$$

com  $y_0$  e  $y_n$  conhecidos.

**Objetivos** Apresentar métodos numéricos que nos conduzam a aproximações da função de uma única variável  $y(x)$ , solução das EDOs de tipos:

$$y'(x) = f(x, y(x)) \quad \text{e} \quad y''(x) = f(x, y(x), y'(x)), \quad (6.1)$$

onde  $f(x, y)$  e  $f(x, y, y')$  são funções conhecidas.

**Um exemplo** (Desintegração radioativa) Quando observamos a desintegração (variação) de uma substância radioativa, podemos constatar que “o número de desintegrações por unidade de tempo é proporcional à quantidade de substância presente em cada instante”. Assim, se  $x = x(t)$  representa a quantidade de uma substância radioativa presente em cada instante  $t$ , o modelo matemático que representa o fenômeno da desintegração é dado por:

$$\frac{dx(t)}{dt} = -\alpha x(t)$$

onde  $\frac{dx}{dt}$  é a variação instantânea (desintegração) sofrida pela substância e o parâmetro  $\alpha > 0$  representa o coeficiente de proporcionalidade, que é constante para cada substância específica. Usamos o sinal negativo porque o número de átomos diminui com o passar do tempo e, portanto,  $\frac{dx}{dt} < 0$ .

## 6.1 Diferenças Finitas

A essência dos métodos numéricos está na discretização do contínuo. É esta discretização que torna “finito” o problema e, portanto, viabiliza sua solução computacional.

A incógnita de uma equação diferencial ordinária é uma função  $y(x)$  definida em todos os pontos do intervalo no qual a equação está sendo resolvida. O primeiro passo de qualquer método destinado à solução numérica de equações diferenciais é discretizar a região onde procuramos a solução. Neste passo, definimos uma *malha*, que é um conjunto finito de pontos, chamados de nós da malha.

Seja  $x_0$  um ponto de referência e  $h$  um número positivo. A malha de passo  $h$  associada a  $x_0$  é o conjunto de pontos:

$$x_i = x_0 \pm ih, \quad i = 1, 2, \dots$$

As aproximações de  $y(x)$  serão calculadas nos pontos desta malha.

No *método de Diferenças Finitas*, o segundo passo consiste na discretização das derivadas que estão na equação diferencial. Neste passo, as derivadas são aproximadas por diferenças entre valores da solução discretizada.

A ferramenta matemática básica de aproximação para as derivadas é a série de Taylor, pois esta nos dá informações sobre a função, no ponto  $x$ , e sua avaliação numa vizinhança de  $x$ , em  $x + h$ . Se assumirmos que  $y(x)$  tem derivadas até a ordem  $n + 1$  em  $x$ , sua expansão em série de Taylor é:

$$y(x + h) = y(x) + hy'(x) + \frac{h^2}{2!}y''(x) + \dots + \frac{h^n}{n!}y^{(n)}(x) + \frac{h^{n+1}}{(n+1)!}y^{(n+1)}(\xi), \quad (6.2)$$

onde  $\xi$  está entre  $x$  e  $x + h$ .

Se tomarmos  $n = 1$  em (6.2), teremos:

$$y(x + h) = y(x) + hy'(x) + \frac{h^2}{2!}y''(\xi).$$

Isolando  $y'(x)$  na equação acima, a *fórmula avançada* para a discretização da derivada e seu erro é:

$$y'(x) = \frac{y(x+h) - y(x)}{h} - \frac{h}{2}y''(\xi). \quad (6.3)$$

De modo semelhante, tomando  $-h$  em (6.2), ainda com  $n = 1$ , temos:

$$y(x-h) = y(x) - hy'(x) + \frac{h^2}{2!}y''(\xi).$$

Isolando  $y'(x)$  na equação acima, a *fórmula atrasada* para a discretização da derivada e seu erro é:

$$y'(x) = \frac{y(x) - y(x-h)}{h} + \frac{h}{2}y''(\xi). \quad (6.4)$$

Tomemos agora  $n = 2$  em (6.2), reescrevemos as fórmulas avançada e atrasada, temos:

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2}y''(x) + \frac{h^3}{3!}y'''(\xi)$$

$$y(x-h) = y(x) - hy'(x) + \frac{h^2}{2}y''(x) - \frac{h^3}{3!}y'''(\xi)$$

Subtraindo a penúltima expressão da última, temos:

$$y(x-h) - y(x+h) = -2hy'(x) - 2\frac{h^3}{3!}y'''(\xi).$$

Rearranjando a expressão acima, obtemos a *fórmula centrada* para discretização da derivada e o seu erro:

$$y'(x) = \frac{y(x+h) - y(x-h)}{2h} - \frac{h^2}{3!}y'''(\xi) \quad (6.5)$$

Seguindo as mesmas idéias, podemos estabelecer uma expressão para o cálculo da aproximação para a segunda derivada. Tomando  $n = 3$  em (6.2), temos:

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2}y''(x) + \frac{h^3}{3!}y'''(x) + \frac{h^4}{4!}y^{(iv)}(\xi)$$

$$y(x-h) = y(x) - hy'(x) + \frac{h^2}{2}y''(x) - \frac{h^3}{3!}y'''(x) + \frac{h^4}{4!}y^{(iv)}(\xi)$$

Somando estas duas últimas fórmulas e explicitando  $y''(x)$ , obtemos uma fórmula para discretizar a derivada de segunda ordem, bem como o erro associado a esta discretização:

$$y''(x) = \frac{y(x+h) - 2y(x) + y(x-h)}{h^2} - \frac{h^4}{12}y^{(iv)}(\xi) \quad (6.6)$$

para algum  $\xi \in (x-h, x+h)$ .

No *método das Diferenças Finitas*, as derivadas presentes na EDO são substituídas por aproximações descritas acima.

Resumindo, para cada ponto da malha no interior do intervalo onde a EDO está definida, denotamos por  $y_i$  a aproximação de  $y(x_i)$  (ou seja,  $y_i \cong y(x_i)$ ), usamos:

$$y'(x_i) \cong \frac{y_{i+1} - y_i}{h}, \quad y'(x_i) \cong \frac{y_i - y_{i-1}}{h}, \quad \text{ou} \quad y'(x_i) \cong \frac{y_{i+1} - y_{i-1}}{2h}$$

e

$$y''(x) = \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2},$$

como discretização das derivadas de primeira e segunda ordens que aparecem na equação diferencial.

O método das Diferenças Finitas pode ser generalizado em um algoritmo para EDO de segunda ordem.

Sejam  $p(x)$ ,  $q(x)$  e  $r(x)$  funções contínuas definidas em  $[a, b]$ . O problema de encontrar  $y(x)$  tal que:

$$y'' + p(x)y' + q(x)y = r(x), \quad a < x < b,$$

com valores de contorno

$$y(a) = y_0, \quad y(b) = y_n,$$

é conhecido como problema de *Sturm-Liouville*. A hipótese

$$q(x) \leq 0, \quad a \leq x \leq b,$$

e a escolha de  $h$  tal que:

$$h \cdot \left( \max_{[a,b]} |p(x)| \right) < 2$$

tornam o sistema diagonalmente dominante, com todas as desigualdades estritas, o que garante a existência de uma única solução do sistema linear.

Dividindo o intervalo  $[a, b]$  em  $n$  partes iguais de comprimento  $h$ , introduzimos a malha  $a = x_0 < x_1 < \dots < x_n = b$ . Se em cada ponto interior  $x_i$  usamos as aproximações :

$$y'(x_i) \cong \frac{y_{i+1} - y_{i-1}}{2h} \quad \text{e} \quad y''(x_i) \cong \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2},$$

obtendo a discretização da EDO:

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + p(x_i) \frac{y_{i+1} - y_{i-1}}{2h} + q(x_i)y_i = r(x_i), \quad i = 1 : n - 1$$

ou ainda, denotando  $p_i = p(x_i)$ ,  $q_i = q(x_i)$  e  $r_i = r(x_i)$ ,

$$\left(1 - \frac{h}{2}p_i\right)y_{i-1} + (-2 + h^2q_i)y_i + \left(1 + \frac{h}{2}p_i\right)y_{i+1} = h^2r_i,$$

para  $i = 1 : n - 1$ . Note que a primeira e a última equações devem ser modificadas pela utilização da condição de contorno.

**Exemplo 37.** Aproxime numericamente a solução da EDO abaixo usando o método de Diferenças Finitas:

$$y'' - y' + xy = e^x(x^2 + 1), \quad x \in (0, 1)$$

com condições de contorno:

$$y(0) = 0, \quad y(1) = e$$

usando  $h = 0.1$  e a fórmula centrada para aproximar  $y'(x_i)$ .



**Dado**  $a, b, p(x), q(x), r(x), y_0, y_n, n$

- 1:  $h = (b - a)/n$
- 2: **Para**  $j = 1$  **até**  $j = n - 1$  **faça**
- 3:      $x = a + jh$
- 4:      $di(j) = 1 - \frac{h}{2}p(x)$
- 5:      $d(j) = -2 + h^2q(x)$
- 6:      $ds(j) = 1 + \frac{h}{2}p(x)$
- 7:      $ti(j) = h^2r(x)$
- 8: **Fim do laço**
- 9:  $ti(1) = ti(1) - y_0(1 - (h/2)p(a + h))$
- 10:  $ti(n - 1) = ti(n - 1) - y_n(1 + (h/2)p(b - h))$
- 11: Resolver o sistema.

Algoritmo 12: Método de Diferenças Finitas de Sturm-Liouville

## 6.2 Métodos de Runge-Kutta

Tomemos a EDO de primeira ordem, com dado inicial:

$$\begin{aligned} y'(x) &= f(x, y) \\ y(x_0) &= y_0. \end{aligned} \tag{6.7}$$

Considerando a malha definida pelo passo  $h$ , podemos usar a fórmula de diferenças finitas avançada para discretizar a derivada de  $y(x)$  no ponto  $x_k$ . Assim, obtemos uma versão discretizada de (6.7):

$$\frac{y_{i+1} - y_i}{h} = f(x_i, y_i). \tag{6.8}$$

Esta equação permite que calculemos  $y_{i+1}$  a partir de  $y_i$ , e define o *Método de Euler*. Sua aplicação é muito simples:

$$\begin{aligned} y_0 &= y(x_0) \\ y_{i+1} &= y_i + hf(x_i, y_i), \quad \text{para } i = 0, 1, \dots \end{aligned} \tag{6.9}$$

O Algoritmo 13 ilustra este método.

**Dado**  $a, b, n, y_0$  e  $f(x, y)$

- 1:  $h = (b - a)/n$
- 2: **Para**  $i = 0$  **até**  $i = n - 1$  **faça**
- 3:      $x_i = a + ih$
- 4:      $y_{i+1} = y_i + hf(x_i, y_i)$
- 5: **Fim do laço**

Algoritmo 13: Método de Euler

A idéia de Euler (6.9) tem versões de maior precisão nos trabalhos de Runge (1895), para o caso de uma

equação, e Kutta (1901), para o caso de sistemas de EDO de primeira ordem.

**Definição 7.** [4] (*Passo Simples*) Um método é de passo simples quando a aproximação  $y_{i+1}$  for calculada a partir do valor  $y_i$  do passo anterior. Sendo  $\phi$  a função incremento, um método de passo simples é definido na forma:

$$y_{i+1} = y_i + h\phi(x_i, y_i; h).$$

**Definição 8.** [4] (*Ordem*) Um método de passo simples tem ordem  $q$  se a função incremento  $\phi$  for tal que

$$y_{i+1} = y_i + h\phi(x_i, y_i; h) + O(h^{q+1})$$

onde  $O(\cdot)$  é a ordem convergência.

**Definição 9.** O Método de Runge-Kutta de ordem  $R$  é definida por:

$$y_{i+1} - y_i = h \cdot \phi(x_i, y_i; h)$$

onde

$$\begin{aligned} \phi(x, y; h) &= \sum_{r=0}^{R-1} c_r k_r \\ m_0 &= f(x, y) \\ m_r &= f\left(x + a_r h, y + h \sum_{s=1}^r b_{rs} k_s\right) \quad r = 1, 2, \dots, R-1 \\ a_r &= \sum_{s=1}^r b_{rs} \quad r = 1, 2, \dots, R-1 \end{aligned} \tag{6.10}$$

Para obter o método de Runge-Kutta para uma ordem específica, devemos determinar as constantes  $c_r$ ,  $a_r$  e  $b_{rs}$  da Definição 9. Determinamos essas constantes comparando a expressão da função  $\phi(x, y; h)$ , definida pelo sistema (6.10), em potências de  $h$ , com a função:

$$\phi_T(x, y; h) = f(x, y) + \frac{h}{2!} f'(x, y) + \dots + \frac{h^{q-1}}{q!} f^{(q-1)}(x, y) \tag{6.11}$$

conhecida como função do Método de Taylor.

### 6.2.1 Método de Runge-Kutta de ordem 2

Considere a Definição 9, com  $R = 2$ :

$$\begin{aligned} \phi(x, y; h) &= c_0 m_0 + c_1 m_1 \\ m_0 &= f(x, y) \\ m_1 &= f(x + a_1 h, y + h b_{11} m_0) \\ a_1 &= b_{11} \end{aligned}$$

Substituindo  $b_{11}$  e  $m_0$  em  $m_1$ , temos:

$$m_1 = f(x + a_1 h; y + h a_1 f(x, y))$$

Desenvolvendo  $k_2$  em série de Taylor em torno do ponto  $(x, y)$ , obtemos:

$$\begin{aligned} m_1 = & f(x, y) + (a_1 h) f_x(x, y) + (a_1 h f(x, y)) f_y(x, y) + \frac{(a_1 h)^2}{2!} f_{xx}(x, y) \\ & + (a_1 h)(a_1 h f(x, y)) f_{xy}(x, y) + \frac{(a_1 h f(x, y))^2}{2!} f_{yy}(x, y) + O(h^3) \end{aligned}$$

Substituindo o valor de  $m_0$  por  $f(x, y)$  e  $m_1$  pela expressão anterior, em  $\phi(x, y; h)$ , segue que:

$$\begin{aligned} \phi(x, y; h) &= c_0 f(x, y) + c_1 \left[ f(x, y) + (a_1 h) f_x(x, y) + (a_1 h f(x, y)) f_y(x, y) + \frac{(a_1 h)^2}{2!} f_{xx}(x, y) \right. \\ &\quad \left. + (a_1 h)^2 f(x, y) f_{xy}(x, y) + \frac{(a_1 h f(x, y))^2}{2!} f_{yy}(x, y) + O(h^3) \right] \\ &= (c_0 + c_1) f(x, y) + c_1 a_1 h (f_x(x, y) + f_y(x, y) f(x, y)) + \frac{(a_1 h^2)^2}{2!} c_1 \\ &\quad \cdot [f_{xx}(x, y) + 2f(x, y) f_{xy}(x, y) + f_{yy}(x, y) f^2(x, y)] + O(h^3) \end{aligned}$$

onde agrupamos os termos de mesma potência de  $h$ . Observe que, na expressão anterior,  $f$  e suas derivadas estão calculadas em  $(x, y)$ . Denotando por

$$F = f_x(x, y) + f_y(x, y) f(x, y) \quad \text{e} \quad G = f_{xx}(x, y) + 2f(x, y) f_{xy}(x, y) + f_{yy}(x, y) f^2(x, y),$$

obtemos:

$$\phi(x, y; h) = (c_0 + c_1) f(x, y) + c_1 a_1 h F + \frac{(a_1 h)^2}{2!} c_1 G + O(h^3) \quad (6.12)$$

Note que, podemos escrever a função  $\phi_T(x, y; h)$ , equação (6.11), com  $q = 3$ , como:

$$\begin{aligned} \phi_T(x, y; h) &= f(x, y) + \frac{h}{2!} f'(x, y) + \frac{h^2}{3!} f''(x, y) + O(h^3) \\ &= f + \frac{h}{2!} (f_x + f_y f) + \frac{h^2}{3!} (f_{xx} + 2f_{xy} f + f_{yy} f^2 + f_x f_y + f_y^2 f) + O(h^3) \\ &= f + \frac{h}{2!} (f_x + f_y f) + \frac{h^2}{3!} (f_{xx} + 2f_{xy} f + f_{yy} f^2 + f_y (f_x + f_y f)) + O(h^3) \end{aligned}$$

Usando  $F$  e  $G$ , obtemos

$$\phi_T(x, y; h) = f + \frac{h}{2} F + \frac{h^2}{3!} [G + f_y F] + O(h^3) \quad (6.13)$$

Para determinarmos o método de Runge-Kutta de ordem 2, comparamos as equações (6.12) e (6.13), obtendo:

$$\begin{cases} c_0 + c_1 = 1 \\ c_1 \cdot a_1 = \frac{1}{2} \end{cases}$$

Resolvendo o sistema, teremos o Método de Runge-Kutta de ordem 2, porém o sistema possui 2 (duas) equações e 3 (três) incógnitas, possuindo infinitas soluções. Tomando  $c_0 = \frac{1}{2}$ , obtemos  $c_1 = \frac{1}{2}$  e  $a_1 = 1$ . Portanto:

$$\begin{aligned} y_{i+1} &= y_i + \frac{h}{2} (m_0 + m_1) \\ m_0 &= f(x_i, y_i) \\ m_1 &= f(x_i + h, y_i + h m_0) \end{aligned}$$

que é conhecido como *Método de Euler Melhorado*. O Algoritmo 14, ilustra o método abordado.

**Dado**  $a, b, n, y_0$  e  $f(x, y)$

- 1:  $h = (b - a)/n$
- 2: **Para**  $k = 0$  **até**  $k = n - 1$  **faça**
- 3:      $x_k = a + kh$
- 4:      $m_0 = hf(x_k, y_k)$
- 5:      $m_1 = hf(x_k + h, y_k + m_0)$
- 6:      $y_{k+1} = y_k + \frac{1}{2}(m_0 + m_1)$
- 7: **Fim do laço**

Algoritmo 14: Método de Runge-Kutta de segunda ordem

### 6.2.2 Método de Runge-Kutta de ordem 3

Neste caso, temos:

$$y_{i+1} = y_i + h(c_0 m_0 + c_1 m_1 + c_2 m_2)$$

onde  $m_0$  e  $m_1$  possuem as mesmas condições do método de ordem 2 e

$$\begin{aligned} m_2 &= f(x + ha_2, y + hb_{21}m_0 + b_{22}m_1) \\ &= f(x + ha_2, y + h(a_2 - b_{22})m_0 + b_{22}m_1) \end{aligned}$$

desde que  $a_2 = b_{21} + b_{22}$ . Devemos então agrupar os termos semelhantes e compará-los com  $\phi_T(x, y; h)$ , obtendo o sistema:

$$\begin{cases} c_0 + c_1 + c_2 = 1 \\ a_1 c_1 + a_2 c_2 = \frac{1}{2} \\ c_2 b_{22} a_1 = \frac{1}{6} \\ c_1 a_1^2 + c_2 a_2^2 = \frac{1}{3} \end{cases}$$

que é um sistema de 4 equações e 6 incógnitas, sendo possível obter vários métodos de ordem 3 com diferentes conjuntos de constantes.

Um desses métodos é chamado de *Runge-Kutta de ordem 3*. Os valores das oito incógnitas são:

$$\begin{aligned} c_0 &= \frac{1}{6} & b_{11} &= \frac{1}{2} \\ c_1 &= \frac{4}{6} & a_1 &= \frac{1}{2} & b_{21} &= -1 \\ c_2 &= \frac{1}{6} & a_2 &= 1 & b_{22} &= 2 \end{aligned}$$

Com estas constantes, as equações formam o método de Runge-Kutta de ordem 3, representado no Algo-

ritmo 15.

$$\begin{aligned}y_{i+1} &= y_i + \left(\frac{1}{6}\right)(m_0 + 4m_1 + m_2) \\m_0 &= f(x_i, y_i) \\m_1 &= f\left(x_i + \frac{h}{2}, y_i + \frac{m_0}{2}h\right) \\m_2 &= f(x_i + h, y_i - m_0h + 2m_1h)\end{aligned}$$

**Dado**  $a, b, n, y_0$  e  $f(x, y)$

- 1:  $h = (b - a)/n$
- 2: **Para**  $k = 0$  **até**  $k = n - 1$  **faça**
- 3:      $x_k = a + kh$
- 4:      $m_0 = hf(x_k, y_k)$
- 5:      $m_1 = hf\left(x_k + \frac{h}{2}, y_k + \frac{m_0}{2}h\right)$
- 6:      $m_2 = hf(x_{k+1}, y_k - m_0h + 2m_1h)$
- 7:      $y_{k+1} = y_k + \frac{1}{6}(m_0 + 4m_1 + m_2)h$
- 8: **Fim do laço**

Algoritmo 15: Método de Runge-Kutta de terceira ordem

### 6.2.3 Método de Runge-Kutta de ordem 4

Neste caso, a comparação de  $\phi$  com  $\phi_T$ , para se obter métodos de Runge-Kutta de ordem 4, fornece um sistema de 11 (onze) equações e 13 (treze) incógnitas. Cada solução deste sistema define um método de Runge-Kutta com ordem 4. Um dos mais utilizados é dado pelo Algoritmo 16.

**Dado**  $a, b, n, y_0$  e  $f(x, y)$

- 1:  $h = (b - a)/n$
- 2: **Para**  $k = 0$  **até**  $k = n - 1$  **faça**
- 3:      $x_k = a + kh$
- 4:      $m_0 = hf(x_k, y_k)$
- 5:      $m_1 = hf\left(x_k + \frac{h}{2}, y_k + \frac{m_0}{2}h\right)$
- 6:      $m_2 = hf\left(x_k + \frac{h}{2}, y_k + \frac{m_1}{2}h\right)$
- 7:      $m_3 = hf(x_{k+1}, y_k + m_2h)$
- 8:      $y_{k+1} = y_k + \frac{1}{6}(m_0 + 2m_1 + 2m_2 + m_3)h$
- 9: **Fim do laço**

Algoritmo 16: Método de Runge-Kutta de quarta ordem

Os métodos de *Runge-Kutta* (como são conhecidos) para EDO's de primeira ordem.

**Exemplo 38.** Seja a EDO:

$$\begin{aligned}y'(x) &= \frac{1}{1+x^2} - 2y^2 \\y(0) &= 0\end{aligned}$$

Resolva-a usando os métodos de Euler, Runge-Kutta de segunda, terceira e quarta ordens para  $x \in (0, 1)$  e  $h = 0.1$ . Compare os resultados obtidos com a sua solução analítica (calcule o erro relativo):

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

## 6.3 Exercícios

1. Aplique o método de Euler, Runge-Kutta de segunda, terceira e quarta ordens para aproximar as soluções dos seguintes PVI e compare os resultados com os valores reais:

(a)  $y' = y/t - (y/t)^2$ ,  $1 \leq t \leq 2$ ,  $y(1) = 1$ , com  $h = 0.1$  e  $x \in [1, 2]$ ; solução real:  $y(t) = t/(1 + \ln t)$ .

(b)  $y' = 1 + y/t + (y/t)^2$ ,  $1 \leq t \leq 3$ ,  $y(1) = 0$  com  $h = 0.2$  e  $x \in [1, 2]$ ; solução real:  $y(t) = t \tan(\ln t)$ .

(c)  $y' = -\sin(x)y$ ,  $x \in [0, \pi]$ ,  $y(0) = -1$ ,  $m = 6$ ; solução real:  $y(x) = -e^{\cos(x)-1}$ .

(d)  $y' = (\sqrt{x} + 1)y$ ,  $x \in [1, 3]$ ,  $y(1) = 1$ ,  $m = 5$ ; solução real:  $y(x) = e^{2x^{1.5}/3 + x - 5/3}$

Atenção:  $m$  é o número de subintervalos.

2. Dado o PVI

$$y' = \frac{2}{t}y + t^2 e^t, \quad 1 \leq t \leq 2 \quad y(1) = 0,$$

com as soluções exatas  $y(t) = t^2(e^t - e)$ . Use o método de Euler com  $h = 0.1$  para aproximar a solução e compare-a com os valores reais de  $y$ .

3. Resolver os problemas de valor inicial abaixo, utilizando os métodos de Euler, Runge-Kutta de segunda, terceira e quarta ordens com o número de subintervalos  $n$  indicado:

(a)  $y' = \sqrt{x}$ ,  $y(0) = 0$ ,  $x \in [0, 2]$  e  $n = 5$ . (solução usando o método de Euler:  $y_5 = 1.55490$ );

(b)  $y' = x^2 + y^2$ ,  $y(1) = 0$ ,  $x \in [1, 2]$  e  $n = 8$  (solução usando o método de Euler:  $y_8 = 3.39195$ );

(c)  $y' = xy$ ,  $y(0) = 1$ ,  $x \in [0, 1]$  e  $n = 10$  (solução usando o método de Euler:  $y_{10} = 1.54711$ ).

4. Use o Método das Diferenças Finitas para aproximar as soluções dos problemas abaixo e compare com as soluções reais.

(a)  $y'' + y = 0$ ,  $0 \leq x \leq (\pi/4)$ ,  $y(0) = 1$ ,  $y(\pi/4) = 1$ ; use  $h = (\pi/20)$ ;

solução real:  $y(x) = \cos x + (\sqrt{2} - 1) \sin x$ .

(b)  $y'' + 4y = \cos x$ ,  $0 \leq x \leq (\pi/4)$ ,  $y(0) = 0$ ,  $y(\pi/4) = 0$ ; use  $h = (\pi/20)$ ;

solução real:  $y(x) = -\frac{1}{3} \cos 2x - \frac{\sqrt{2}}{6} \sin 2x + \frac{1}{3} \cos x$ .

(c)  $y'' = -\frac{4}{x}y' - \frac{2}{x^2}y + \frac{2 \ln x}{x^2}$ ,  $1 \leq x \leq 2$ ,  $y(1) = \frac{1}{2}$ ,  $y(2) = \ln 2$ ; use  $h = 0.05$ ;

solução real  $y(x) = \frac{4}{x} - \frac{2}{x^2} + \ln x - \frac{3}{2}$ .

(d)  $y'' = 2y' - y + x e^x - x$ ,  $0 \leq x \leq 2$ ,  $y(0) = 0$ ,  $y(2) = -4$ ; use  $h = 0.2$ ;

solução real:  $y(x) = \frac{1}{6}x^3 e^x - \frac{5}{3}x e^x + 2e^x - x - 2$ .

5. Use o Método das Diferenças Finitas para aproximar as soluções dos seguintes PVC:

(a)  $y'' = -\frac{4}{x}y' + \frac{2}{x^2}y - \frac{2}{x^2} \ln x$ ,  $1 \leq x \leq 2$ ,  $y(1) = -\frac{1}{2}$ ,  $y(2) = \ln 2$ ; use  $h = 0.05$ ;

(b)  $y'' = \frac{y'}{x} + \frac{3}{x^2}y + \frac{\ln x}{x} - 1$ ,  $1 \leq x \leq 2$ ,  $y(1) = y(2) = 0$ ; use  $h = 0.1$

6. Considere o PVI:  $y' = yx^2 - y$ , com  $y(0) = 1$ .

- (a) Encontre a solução aproximada usando o método de Euler com  $h = 0.5$  e  $h = 0.25$ , considerando  $x \in [0, 2]$ ;
- (b) idem, usando Runge-Kutta de 4ª ;
- (c) Sabendo que a solução analítica do problema é  $y = e^{-x+x^3/3}$ , coloque num mesmo gráfico a solução analítica e as soluções numéricas encontradas nos itens anteriores. Compare seus resultados.

7. Dado o PVI  $y' = \frac{x}{y}$ ,  $y(0) = 20$ , deseja-se encontrar aproximações para  $y(16)$ . Resolva por:

- (a) Runge-Kutta de 2ª ordem,  $h = 2$ . (Resposta:  $y(16) = 12.00999$ )
- (b) Runge-Kutta de 4ª ordem,  $h = 4$ . (Resposta:  $y(16) = 11.998$ )

8. Calcule  $y(1)$  para  $y' = y - x$ ;  $y(0) = 2$ , utilizando Euler e Runge-Kutta de 4ª ordem com  $h = 0.2$ . Compare seus resultados com os valores exatos de  $y(x)$  nos pontos  $x_i$ , sabendo que  $y(x) = e^x + x + 1$ .

Resposta: Euler:  $y(1) = 4.488320$ , Runge-Kutta de 4ª ordem:  $y(1) = 4.78251$ , solução exata:  $y(1) = 4.718282$ .

9. Resolva pelo método de diferenças finitas, o PVC:

$$\begin{cases} y'' + 2y' + y = x \\ y(0) = 2 \\ y(1) = 0 \end{cases}$$

usando  $h = 0.25$ . (Resposta:  $y(0.25) = 1.107487$ ,  $y(0.5) = 0.529106$ ,  $y(0.75) = 0.180622$ )

10. O PVC:

$$y'' = 4(y - x), \quad 0 \leq x \leq 1, \quad y(0) = 0, \quad y(1) = 2$$

tem a solução  $y(x) = e^2(e^4 - 1)^{-1}(e^{2x} - e^{-2x}) + x$ . Use o Método das Diferenças Finitas para aproximar a solução e compare os resultados com a solução real.

- (a) Com  $h = \frac{1}{2}$ .
- (b) Com  $h = \frac{1}{4}$ .
- (c) Qual o valor de  $y(1/2)$  nos itens anteriores.

11. O PVC:

$$y'' = y' + 2y + \cos x, \quad 0 \leq x \leq \frac{\pi}{2}, \quad y(0) = -0.3, \quad y\left(\frac{\pi}{2}\right) = -0.1$$

tem a solução  $y(x) = -\frac{1}{10}(\sin x + 3 \cos x)$ . Use o Método das Diferenças Finitas para aproximar a solução e compare os resultados com a solução real.

- (a) Com  $h = \frac{\pi}{4}$ .
- (b) Com  $h = \frac{\pi}{8}$ .
- (c) Qual o valor de  $y(\pi/4)$  nos itens anteriores.

12. Use o Método das Diferenças Finitas para aproximar as soluções de  $y = e^{-10x}$  para o PVC:

$$y'' = 100y, \quad 0 \leq x \leq 1, \quad y(0) = 1, \quad y(1) = e^{-10}.$$

Use  $h = 0.1$  e  $h = 0.25$ .

---

## Referências Bibliográficas

---

- [1] R. L. Burden, J. D. Faires. *Análise Numérica*. Pioneira Thomson Learning, 2003.
- [2] F. F. Campos Fo. *Algoritmos Numéricos*. LTC Editora, 2001.
- [3] M. C. C. Cunha. *Métodos Numéricos*. 2ª Edição, Editora da Unicamp, 2000.
- [4] N. B. Franco. *Cálculo Numérico*. Pearson Prentice Hall, 2006.
- [5] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*, SIAM – Society for Industrial and Applied Mathematics, 1996.
- [6] M. A. G. Ruggiero, V. L. R. Lopes *Cálculo Numérico: Aspectos Teóricos e Computacionais*. 2ª Edição, Makron Books, 1996.
- [7] D. Sperandio, J. T. Mendes, L. H. M. e Silva. *Cálculo Numérico: Características Matemáticas e Computacionais dos Métodos Numéricos*. Prentice Hall, 2003.
- [8] J. Stewart. *Cálculo*. Volume I. 5ª Edição, Pioneira Thomson Learning, 2006.
- [9] E. W. Sowokowski. *Cálculo com Geometria Analítica*. Volume 1. 2ª Edição, Makron Books, 1994.
- [10] [www.nrbook.com/b/bookcpdf.php](http://www.nrbook.com/b/bookcpdf.php)
- [11] [www.mathworks.com/moler/chapters.html](http://www.mathworks.com/moler/chapters.html)