

A thick yellow line starts at the bottom left, moves up and right in a jagged, step-like fashion, then continues up and right, ending at the top right. It separates a blue area on the left from a white area on the right.

# Cálculo Numérico

## Volume 1





# Cálculo Numérico

## Volume 1

Germán Ramón Canahualpa Suazo  
Márcia Rosales Ribeiro Simch  
André Luis Andrejew Ferreira

Pelotas 2011

Ministério da Educação  
Universidade Federal de Pelotas  
Curso de Licenciatura em Matemática a Distância



## Obra publicada pela Universidade Federal de Pelotas

*Reitor: Prof. Dr. Antonio Cesar Gonçalves Borges*  
*Vice-Reitor: Prof. Dr. Manoel Luiz Brenner de Moraes*  
*Pró-Reitor de Extensão e Cultura: Prof. Dr. Luiz Ernani Gonçalves Ávila*  
*Pró-Reitora de Graduação: Prof. Dra. Eliana Póvoas Brito*  
*Pró-Reitor de Pesquisa e Pós-Graduação: Prof. Dr. Manoel de Souza Maia*  
*Pró-Reitor Administrativo: Eng. Francisco Carlos Gomes Luzzardi*  
*Pró-Reitor de Planejamento e Desenvolvimento: Prof. Ms. Élio Paulo Zonta*  
*Pró-Reitor de Recursos Humanos: Admin. Roberta Trierweiler*  
*Pró-Reitor de Infra-Estrutura: Mario Renato Cardoso Amaral*  
*Pró-Reitora de Assistência Estudantil: Assistente Social Carmen de Fátima de Mattos do Nascimento*

### CONSELHO EDITORIAL

*Prof. Dr. Antonio Jorge Amaral Bezerra*  
*Prof. Dra. Isabel Porto Nogueira*  
*Profa. Lígia Antunes Leivas*  
*Prof. Dr. Renato Luiz Mello Varoto*  
*Prof. Dr. Volmar Geraldo da Silva Nunes*  
*Prof. Dr. Elomar Antonio Callegaro Tambara*  
*Prof. Dr. José Justino Faleiros*  
*Profa. Dra. Neusa Mariza Leite Rodrigues Felix*  
*Prof. Ms. Valter Eliogabalos Azambuja*  
*Prof. Dr. Wilson Marcelino Miranda*

### CURSO DE LICENCIATURA EM MATEMÁTICA A DISTÂNCIA

*Coordenador do Colegiado: Dr. Luiz Alberto Brettas*

### PROJETO GRÁFICO E DIAGRAMAÇÃO

*Celina Bastos Lemos*  
*Eduardo Harry Luersen*  
*Guilherme Camargo*  
*Mateus Dias Vilela*  
*Rodrigo Pizarro dos Santos*

### FOTOGRAFIA DA CAPA

*Grande Hotel*  
*por Marcio Kinzeski*



**Editora e Gráfica Universitária**  
R Lobo da Costa, 447 – Pelotas, RS – CEP 96010-150  
Fone/fax: (53) 3227 8411  
e-mail: editora@ufpel.edu.br

*Diretor da Editora e Gráfica Universitária: Prof. Dr. Volmar Geraldo da Silva Nunes*  
*Gerência Operacional: Carlos Gilberto Costa da Silva*

*Impresso no Brasil*  
*Edição: 2009*

*ISBN : 978-85-7192-484-0*  
*Tiragem: xxx exemplares*

Dados de Catalogação na Fonte Internacional:  
(Bibliotecária Daiane Schramm – CRB-10/1881 )



Material produzido pelo Laboratório de Ensino de Matemática a Distância - LEMAD  
Direitos autorais do Curso de Licenciatura em Matemática a Distância da UFPel

## Apresentação

A literatura didática para o ensino de matemática recebe, através dessa publicação, um novo fôlego em material preparado para o ensino. Trata-se de um novo caminho que começamos a trilhar, aprendendo através da pesquisa, da discussão e, acima de tudo, da dedicação de nossos professores. Esta obra nasce a partir do esforço conjunto de professores da Universidade Federal de Santa Catarina, da Universidade Federal do Rio Grande do Sul, da Universidade Estadual de Maringá, da Universidade Federal de Pelotas, de alunos e técnicos do Laboratório de Ensino de Matemática a Distância. Estes profissionais, apoiados em recursos do Ministério da Educação, aceitaram o desafio de preparar os materiais necessários para a execução do Curso de Licenciatura em Matemática a Distância da UFPel. Esperamos que este livro, que ora apresentamos, seja um instrumento útil nas mãos de alunos e professores e sirva de apoio para a melhoria da educação em nosso país.

**Dr. Luiz Alberto Brettas**

Coordenador do Colegiado do Curso de Licenciatura  
em Matemática a Distância

# Sumário

## Capítulo 1 – Introdução: Grandezas Físicas, Representação Vetorial e Sistemas de Unidades

1.1 Erros.....	11
1.1.1 Erro Inerente.....	11
1.1.2 Erros de Discretização.....	12
1.1.3 Erros de Arredondamento.....	12
1.1.4 Erros na Fase da Resolução.....	12
1.2 Conversão de Bases.....	13
1.2.1 Conversão de Inteiros da Base Decimal para a Base Binária.....	13
1.2.2. Conversão de Inteiros da Base Binária para a Base Decimal.....	14
1.2.2.1. O Esquema de Horner - Que Consiste em Calcular a Seqüência.....	14
1.2.2.2. A Divisão de Ruffini - Que é equivalente ao primeiro método e só difere na disposição dos coeficientes $a_j$ e $b_j$ .....	15
1.2.3 Conversão de Números Fracionários da Base Decimal para a Base Binária.....	16
1.2.4 Conversão de Números Fracionários da Base Binária para a Base Decimal.....	16
1.2.4.1 O Esquema de Horner.....	17
1.2.4.2 A Divisão de Ruffini.....	17
1.3 Aritmética de Ponto Flutuante.....	18

## Capítulo 2 – Equações Algébricas e Transcendentais

2.1 Introdução.....	22
2.2 Métodos Diretos e Indiretos.....	24
2.2.1 Métodos Diretos.....	24
2.2.2 Métodos Indiretos.....	24
2.2.3 Aproximações Iniciais.....	25
2.2.4 Geração de Aproximantes e Critérios de Parada.....	26
2.3 Equações Algébricas Polinomiais.....	29
2.3.1 Teorema de Lagrange.....	29
2.3.2 Regra de Sinais de Descartes.....	30
2.3.3 Seqüência de Sturm.....	31
2.4 Equações Transcendentais.....	34
2.4.1 Método da Bissecção.....	34
2.4.2 Método da Falsa Posição.....	37
2.4.3 Método das Cordas.....	39
2.4.4 Método de Newton-Raphson.....	41

## Capítulo 3 – Sistemas de Equações Lineares Algébricas

3.1 Introdução.....	46
3.2 Métodos Diretos.....	47
3.2.1 Regra de Cramer.....	47
3.2.2 Eliminação Gaussiana.....	48
3.2.3 Eliminação de Decomposição LU.....	50
3.3 Métodos Iterativos.....	51
3.3.1 Método de Jacobi.....	52
3.3.2 Método de Gauss-Seidel.....	54
3.3.3 Convergência dos Métodos Iterativos.....	57

## Capítulo 4 – Interpolação

4.1 Introdução.....	60
4.2 Interpolação de Van der Monde.....	60
4.3 Interpolação de Lagrange.....	61

4.4 Interpolação de Newton.....	69
---------------------------------	----

## Capítulo 5 – Ajustes por Quadros Mínimos

5.1 Ajuste Discreto Linear por Quadrados Mínimos.....	76
5.1.1 Ajuste de uma Reta por Quadrados Mínimos.....	78
5.2 Ajuste Discreto Não Linear por Quadrados Mínimos.....	83
5.3 Ajuste Contínuo Linear por Quadrados Mínimos.....	87

## Capítulo 6 – Integração Numérica

6.1 Regra do Trapézio.....	92
6.2 Regra de Simpson.....	94
6.3 Técnicas Compostas de Integração Numérica.....	98
6.3.1 Regra Composta do Trapézio.....	99
6.4 Quadratura Gaussiana.....	105



# Introdução a Análise de Erros

Capítulo

1

O Cálculo Numérico corresponde a um conjunto de ferramentas ou métodos usados para se obter a solução de problemas matemáticos de forma aproximada. Esses métodos se aplicam principalmente a problemas que não apresentam uma solução exata, portanto precisam ser resolvidos numericamente.

Na resolução de um problema matemático o objetivo geral é obter a solução. Esse objetivo pode ser alcançado através de várias etapas que podem ser resumidas pela figura a seguir.

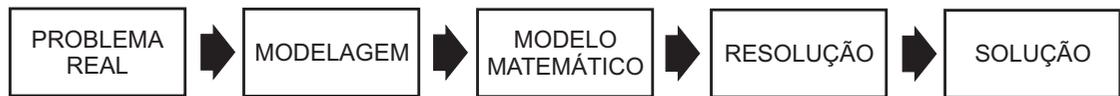


FIGURA 1: Fases da resolução de um problema

Nos problemas reais, os dados são medidas e, como tais, não são exatos. Uma medida física não é um número, é um intervalo, pela própria imprecisão das medidas. Daí trabalha-se sempre com a figura do erro, inerente à própria medição. Logo solução obtida para um problema real pode estar sujeita a algum tipo de erro na fase da modelagem e da resolução. Essas fases são as maiores fontes de erros na resolução numérica de um problema real.

O erro pode ser entendido como a diferença entre o valor real e o valor obtido como solução para um determinado problema.

**Exemplo 1:** Calcular a área de uma circunferência de raio 100 cm.

Resultados Obtidos:

- a) Área = 31400 cm<sup>2</sup>
- b) Área = 31416 cm<sup>2</sup>
- c) Área = 31415,92654 cm<sup>2</sup>

Como justificar as diferenças entre os resultados? É possível obter “exatamente” esta área?

**Exemplo 2:** Efetuar os somatórios seguintes em uma calculadora e em um computador:

$$S = \sum_{i=1}^{30000} x_i \text{ para } x_i = 0,4 \text{ e para } x_i = 0,11$$

Resultados Obtidos:

- a) Para  $x_i = 0,4$ :
  - Na calculadora:  $S = 12000$
  - No computador:  $S = 12000$
- b) Para  $x_i = 0,11$ :
  - Na calculadora:  $S = 3300$
  - No computador:  $S = 3299,99691$

Como justificar a diferença entre os resultados obtidos pela calculadora e pelo computador para  $x_i = 0,11$ ?

Os erros observados nos dois problemas dependem da representação dos números na máquina utilizada.

A representação de um número depende da base escolhida ou disponível na máquina em uso e do número de dígitos usados na sua representação. Qualquer cálculo que envolva números, que não podem ser representados através de um número finito de dígitos não fornecerá como resultado um valor exato. Quanto maior o número de dígitos utilizados, melhor a precisão obtida. Por isso, o valor obtido no cálculo da área da circunferência é aquela obtida no caso (c).

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** no primeiro caso, **3,1416** no segundo e **3,141592654**. Em cada caso foi obtido um resultado distinto, e o erro depende exclusivamente da aproximação escolhida para  $\pi$ . Qualquer que seja a circunferência, a sua área nunca será obtida exatamente, pois o número  $\pi$  é um número irracional.

No exemplo 2, observe o que acontece na interação entre o usuário e o computador: os dados de entrada são enviados ao computador pelo usuário no sistema decimal (0, 1, 2, ..., 7, 8 e 9); toda essa informação é convertida para o sistema binário (0 e 1), e as operações todas serão efetuadas neste sistema. Os resultados finais serão convertidos para o sistema decimal e, finalmente, serão apresentados ao usuário. Todo esse processo de conversão é uma fonte de erros que afetam o resultado final dos cálculos.

## 1.1 Erros

Existem vários tipos de erros e suas possíveis causas. A seguir serão abordados três tipos de erros: os inerentes, os de discretização e os de arredondamento.

### 1.1.1 Erro Inerente

Os erros inerentes aparecem na criação ou simplificação de um modelo matemático de determinado sistema, ou ainda nas coletas das medidas, em geral. Os valores de medidas como tempo, temperatura, distância, intensidade luminosa, e outras representações de grandezas físicas são obtidos de instrumentos que têm precisão limitada. Sobre isso o analista numérico não tem meios de evitá-los ou mesmo minimizar seu efeito aplicados ao modelo em questão.

**Exemplo 3:** Considere a equação de queda livre de um corpo dada pela expressão  $h = h_0 - v_0 t - \frac{1}{2}gt^2$ , sendo:

- h: altura do corpo com relação ao solo no instante t (m);
- $h_0$ : altura inicial (t = 0) do corpo com relação ao solo (m);
- $v_0$ : velocidade inicial (t = 0) do corpo (m/s);
- g: aceleração da gravidade (m/s<sup>2</sup>);
- t: tempo (s).

A equação acima tem simplificações tais como a não inclusão da resistência do ar, a velocidade e direção do vento, etc. Suponha que se quer determinar a altura de um prédio mediante um cronômetro e uma bolinha de metal, utilizando a fórmula da queda livre acima. Sobe-se no topo do edifício e larga-se a bolinha, medindo o tempo que a bolinha gasta para tocar o solo. Se o cronômetro marcou 3 segundos, então, como no solo a altura  $h = 0$ , resulta que  $h = 44,1$ m. Mas este resultado pode não ser confiável, pois se considerarmos um tempo de 3,1 segundos a altura resultaria em  $h = 47,1$ m. A diferença entre as medições pode ser devido à percepção do observador bem como a imprecisões próprias do instrumento utilizado.

### 1.1.2 Erros de Discretização

Os erros de discretização, ou de aproximação, ou truncamento são erros cometidos quando se substitui qualquer processo infinito por um processo finito ou discreto.

**Exemplo 4:** Determinação do cálculo da constante  $e$ , que é dada por.

Se quisermos obter o valor de  $e$ , por esta série teremos de efetuar o cálculo de várias parcelas e depois parar, ou seja, truncando a série, nesse caso cometemos um erro causado pelo abandono das parcelas que não foram somadas.

### 1.1.3 Erros de Arredondamento

Os erros de arredondamento surgem quando trabalhamos com máquinas digitais para representar os números reais. Em geral, trabalhamos com arredondamento para o número de ponto flutuante mais próximo ou com o arredondamento por falta. A diferença entre o valor arredondado e o valor exato pode ser medida pelo erro absoluto ou relativo.

**Definição 1:** Seja  $x$  o valor exato de um número e  $\bar{x}$  o seu valor aproximado. O erro absoluto ( $EA_x$ ) é definido como a diferença entre o valor exato e o valor aproximado do número, isto é,  $EA_x = x - \bar{x}$ .

**Definição 2:** Seja  $x$  o valor exato de um número e  $\bar{x}$  o seu valor aproximado. O erro relativo ( $ER_x$ ) é definido como o erro absoluto dividido por  $\bar{x}$ , isto é,  $ER_x = \frac{EA_x}{\bar{x}} = \frac{x - \bar{x}}{\bar{x}}$ .

**Exemplo 5:** Seja  $x = 0.00006$  e  $\bar{x} = 0.00005$ . Utilizando as expressões para determinar os erros, temos que:

$$EA_x = 0.00001 \text{ e } ER_x = \frac{0.00001}{0.0005} = 0.2$$

Neste caso podemos até dizer que  $EA_x$  é pequeno; no entanto, não há dúvida a respeito do  $ER_x$  que é da ordem de 20%.

### 1.1.4 Erros na Fase da Resolução

O objetivo da fase da resolução é obter a solução do modelo matemático junto com a informação dos dados, mediante a aplicação de métodos numéricos.

Nesta fase, tem-se as seguintes fontes de erros:

- Conversão numérica;
- Aritmética de ponto flutuante.

1.1.4.1. Erros de conversão numérica: originam-se quando a informação dos dados numéricos na base decimal é transferida para a calculadora ou computador, ou quando o resultado numérico obtido no computador é transferido para o usuário. Estes erros podem ser de arredondamento ou de truncamento.

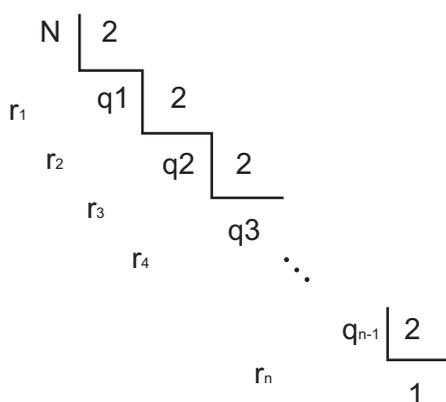
1.1.4.2. Erros de aritmética de ponto flutuante: que se originam quando os dados são processados no computador mediante operações aritméticas. Cabe ressaltar que acontece um fenômeno chamado propagação de erros, que é um processo em que os erros cometidos em um cálculo influenciam nos cálculos subsequentes.

Nas seções seguintes, estudam-se estes processos.

## 1.2 Conversão de Bases

### 1.2.1 Conversão de Inteiros da Base Decimal para a Base Binária

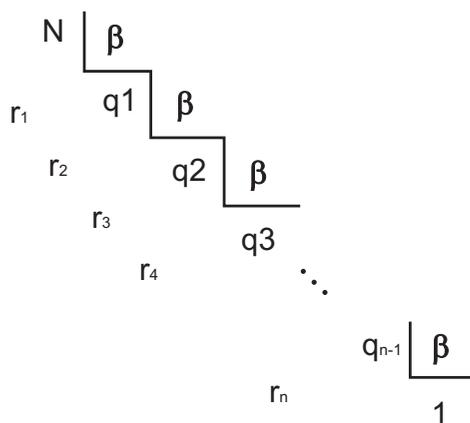
Aplica-se o método das divisões sucessivas, que consiste em dividir o número por 2 e os sucessivos quocientes por 2 até que o último quociente seja 1. A representação esquemática:



A representação binária do número  $N$  será dada por  $N = 1.(r_n r_{n-1} \dots r_3 r_2 r_1)_2$ .

Este método pode ser utilizado para converter qualquer número inteiro da base decimal para uma base  $\beta$  qualquer, para isso basta dividir o número pela base  $\beta$  desejada e os sucessivos quocientes por  $\beta$  até que o último quociente seja um inteiro  $1 \leq q_n < \beta$ .

Generalizando para uma base qualquer  $\beta$  temos:

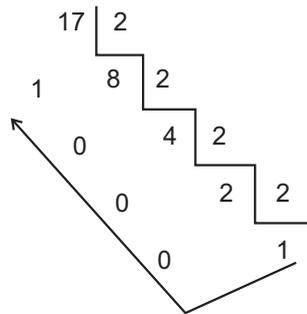


A representação do número será dada por  $N = q_n.(r_n r_{n-1} \dots r_3 r_2 r_1)_\beta$ , onde:

- $N$  = número;
- $q$  = quociente;
- $r$  = resto;
- $\beta$  = base.

**Exemplo 6:** Faça a conversão dos números inteiros na base decimal para a base binária.

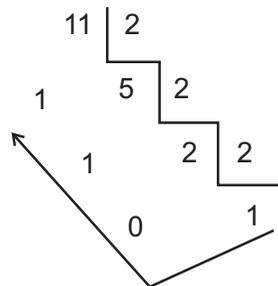
a)  $(17)_{10} = (?)_2$



Então:  $(17)_{10} = (10001)_2$ , logo,  $17 = 1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 16 + 0 + 0 + 0 + 1$ .

A construção do número acompanha o sentido da última figura.

b)  $(11)_{10} = (?)_2$



Então:  $(11)_{10} = (1011)_2$ , logo,  $11 = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 8 + 0 + 2 + 1$ .

### 1.2.2. Conversão de Inteiros da Base Binária para a Base Decimal

O número representado por  $(a_m \dots a_2 a_1 a_0)_2$  tem o valor decimal igual a  $a_m 2^m + \dots + a_2 2^2 + a_1 2^1 + a_0 2^0$ . Portanto, para conhecer o valor decimal do número  $(a_m \dots a_2 a_1 a_0)_2$  basta calcular a soma anterior.

Esta soma pode ser calculada diretamente ou utilizando dois métodos equivalentes:

#### 1.2.2.1. O Esquema de Horner - Que Consiste em Calcular a Seqüência

$$\begin{aligned} b_m &= a_m, \\ b_{m-1} &= a_{m-1} + 2b_m, \\ b_{m-2} &= a_{m-2} + 2b_{m-1}, \\ &\dots \\ b_1 &= a_1 + 2b_2, \\ b_0 &= a_0 + 2b_1. \end{aligned}$$

Onde o valor decimal do número  $(a_m \dots a_2 a_1 a_0)_2$  é então, simplesmente,  $b_0$ .

1.2.2.2. A Divisão de Ruffini - Que é equivalente ao primeiro método e só difere na disposição dos coeficientes  $a_j$  e  $b_j$

	$a_m$	$a_{m-1}$	...	$a_2$	$a_1$	$a_0$
2		$2b_m$	...	$2b_3$	$2b_2$	$2b_1$
	$b_m$	$b_{m-1}$	...	$b_2$	$b_1$	$b_0$

**Exemplo 7:** O número  $(11101)_2$  forma a seqüência de Horner, então tem-se:

$$\begin{aligned}
 b_4 &= a_4 = 1, \\
 b_3 &= a_3 + 2b_4 = 1 + 2 \cdot 1 = 3; \\
 b_2 &= a_2 + 2b_3 = 1 + 2 \cdot 3 = 7; \\
 b_1 &= a_1 + 2b_2 = 0 + 2 \cdot 7 = 14; \\
 b_0 &= a_0 + 2b_1 = 1 + 2 \cdot 14 = 29.
 \end{aligned}$$

Logo,  $(11101)_2 = (29)_{10}$ .

**Exemplo 8:** O número  $(101100)_2$  forma o esquema de Ruffini, então tem-se:

	1	0	1	1	0	0
2		2	4	10	22	44
	1	2	5	11	22	44

Portanto,  $(101100)_2 = (44)_{10}$ .

Um outro método prático é realizar a operação de multiplicação sobre o conjunto de dígitos binários, obedecendo a ordem decrescente dos expoentes, isto é, do expoente mais significativo até o menor.

Exemplo 8: Faça a conversão dos números inteiros na base binária para a base decimal.

a)  $(10101)_2 = (?)_{10}$

Então:  $(10101)_2 = 1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 16 + 0 + 4 + 0 + 1 = 21$ .

Logo  $(10101)_2 = (21)_{10}$

b)  $(01010)_2 = (?)_{10}$

Então:  $(01010)_2 = 0 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 0 + 8 + 0 + 2 + 0 = 10$ .

Logo  $(01010)_2 = (10)_{10}$

### 1.2.3 Conversão de Números Fracionários da Base Decimal para a Base Binária

Aplica-se o método das multiplicações sucessivas, que consiste em multiplicar o número por 2 e extrair a parte inteira (podendo ser 0). O processo se repete, onde o resto fracionário é multiplicado novamente por 2 e a parte inteira sempre é extraída. Quando o resultado da multiplicação for maior que um, então o dígito binário é 1, caso contrário o resultado será o dígito binário 0. Este processo deve-se repetir até ter um resto fracionário igual a 0 ou até observar um padrão repetitivo, em cujo caso tratar-se-á de um número fracionário periódico (puro ou misto).

**Exemplo 9:** Faça a conversão dos números fracionários na base decimal para a base binária.

a)  $(0,3125)_{10} = (?)_2$

$0,3125 \times 2 = 0,6250 \therefore \text{resultado} < 1 \therefore \text{dígito binário } 0$

$0,6250 \times 2 = 1,250 \therefore \text{resultado} > 1 \therefore \text{dígito binário } 1 \therefore \text{extrai-se a parte inteira}$

$0,250 \times 2 = 0,500 \therefore \text{resultado} < 1 \therefore \text{dígito binário } 0$

$0,500 \times 2 = 1,000 \therefore \text{resultado igual a } 1 \therefore \text{dígito binário } 1 \therefore \text{para o processo.}$

Então:  $(0,3125)_{10} = (0,0101)_2$

b)  $(0,8)_{10} = (?)_2$

$0,8 \times 2 = 1,6 \therefore \text{resultado} > 1 \therefore \text{dígito binário } 1 \therefore \text{extrai-se a parte inteira}$

$0,6 \times 2 = 1,2 \therefore \text{resultado} > 1 \therefore \text{dígito binário } 1 \therefore \text{extrai-se a parte inteira}$

$0,2 \times 2 = 0,4 \therefore \text{resultado} < 1 \therefore \text{dígito binário } 0$

$0,4 \times 2 = 0,8 \therefore \text{resultado} < 1 \therefore \text{dígito binário } 0$

$0,8 \times 2 = 1,6 \therefore \text{resultado} > 1 \therefore \text{dígito binário } 1 \therefore \text{extrai-se a parte inteira}$

Observe que o último resultado é igual ao primeiro, nesse caso começa o processo de repetição.

Então:  $(0,8)_{10} = (0,11001100\dots)_2 = (0,\overline{1100})_2$

Considere o número composto por parte inteira e fracionária, por exemplo,  $(12,125)_{10} = (12)_{10} + (0,125)_{10}$ . Para transformá-lo para a base binária faz-se uso dos dois últimos processos, isto é, aplica-se sobre a parte inteira o método das divisões sucessivas, e para a parte fracionária, o método das multiplicações sucessivas.

### 1.2.4 Conversão de Números Fracionários da Base Binária para a Base Decimal

Considere o número fracionário na base binária  $(0, a_1 a_2 \dots a_n)_2$ .

Este número tem representação finita e o valor decimal é dado por:

$$a_1 2^{-1} + a_2 2^{-2} + \dots + a_n 2^{-n}.$$

Esta soma pode ser calculada diretamente ou utilizando qualquer um dos dois métodos já enunciados na subseção anterior com algumas modificações.

### 1.2.4.1 O Esquema de Horner

$$\begin{aligned}
 b_n &= a_n, \\
 b_{n-1} &= a_{n-2} + \frac{1}{2} \cdot b_n, \\
 b_{n-2} &= a_{n-3} + \frac{1}{2} b_{n-1}, \\
 &\dots \\
 b_2 &= a_2 + \frac{1}{2} b_3, \\
 b_1 &= a_1 + \frac{1}{2} b_2, \\
 b_0 &= b_1.
 \end{aligned}
 \tag{3.10}$$

O valor decimal do número  $(0, a_1 a_2 \dots a_n)_2$  é então, simplesmente,  $b_0$ .

### 1.2.4.2 A Divisão de Ruffini

	$a_n$	$a_{n-1}$	$\dots$	$a_2$	$a_1$	$a_0$
$\frac{1}{2}$		$\frac{1}{2}b_n$	$\dots$	$\frac{1}{2}b_3$	$\frac{1}{2}b_2$	$\frac{1}{2}b_1$
	$b_n$	$b_{n-1}$	$\dots$	$b_2$	$b_1$	$b_0$

**Exemplo 10:** O número  $(0,10111)_2$  origina a seqüência de Horner. Temos então:

$$\begin{aligned}
 b_5 &= a_5 = 1, \\
 b_4 &= a_4 + \frac{1}{2} \cdot b_5 = 1 + \frac{1}{2} \cdot 1 = 3/2, \\
 b_3 &= a_3 + \frac{1}{2} \cdot b_4 = 1 + \frac{1}{2} \cdot 3/2 = 7/4, \\
 b_2 &= a_2 + \frac{1}{2} \cdot b_3 = 0 + \frac{1}{2} \cdot 7/4 = 7/8, \\
 b_1 &= a_1 + \frac{1}{2} \cdot b_2 = 1 + \frac{1}{2} \cdot 7/8 = 23/16, \\
 b_0 &= \frac{1}{2} \cdot b_1 = \frac{1}{2} \cdot 23/16 = 23/32 = 0,71875.
 \end{aligned}$$

Portanto,  $(0,10111)_2 = (0,71875)_{10}$ .

**Exemplo 11:** O número  $(0,00111)_2$  origina a divisão de Ruffini. Temos então:

	1	1	1	0	0	0
$\frac{1}{2}$		$\frac{1}{2}$	$3/4$	$7/8$	$7/16$	$7/32$
	1	$3/2$	$7/4$	$7/8$	$7/16$	$7/32$

Portanto,  $(0,00111)_2 = 7/32 = (0,21875)_{10}$ .

Se o número fracionário tem representação binária infinita, a sua representação é:  $(0, a_1 a_2 \dots a_n \overline{b_1 b_2 \dots b_m})_2$

então o valor decimal é dado por

$$a_1 2^{-1} + a_2 2^{-2} + \dots + a_n 2^{-n} + b_1 2^{-n-1} + b_2 2^{-n-2} + \dots + b_m 2^{-n-m} \\ + b_1 2^{-n-m-1} + b_2 2^{-n-m-2} + \dots + b_m 2^{-n-2m} \\ + b_1 2^{-n-2m-1} + b_2 2^{-n-2m-2} + \dots + b_m 2^{-n-3m} + \dots$$

que pode ser escrito como

$a_1 2^{-1} + a_2 2^{-2} + \dots + a_n 2^{-n} + b_1 2^{-n-1} + b_2 2^{-n-2} + \dots + b_m 2^{-n-m} \cdot (2^{m-n}/2^m - 1)$ , onde as duas expressões entre parênteses tem a mesma forma que a soma na equação e podem ser calculadas diretamente ou utilizando qualquer dos métodos descritos anteriormente.

Exemplo 12: Dado o número fracionário  $(0,11010)_2$ , determinando o seu respectivo valor decimal, temos:  $(1.2^{-1} + 1.2^{-2}) + (0.2^{-1} + 1.2^{-2} + 0.2^{-3}) \cdot 1/(2^3-1) = (1/2 + 1/4) + 1/4 \cdot 1/7 = 11/14$ .

Portanto,  $(0,11010)_2 = 11/14 = (0,7857142)_{10}$

Em geral, se o número fracionário tem representação infinita na base  $\beta$ ,  $(0, a_1 a_2 \dots a_n \overline{b_1 b_2 \dots b_m})_\beta$ , então o valor decimal é:

$$(a_1 \beta^{-1} + a_2 \beta^{-2} + \dots + a_n \beta^{-n}) + (b_1 \beta^{-1} + b_2 \beta^{-2} + \dots + b_m \beta^{-m}) \cdot \frac{\beta^{m-n}}{\beta^m - 1},$$

onde as duas expressões entre parênteses podem ser calculadas diretamente ou utilizando qualquer um dos métodos descritos anteriormente (Divisão de Ruffini ou Método de Horner).

### 1.3 Aritmética de Ponto Flutuante

As calculadoras e computadores possuem um número finito de dígitos para representar os números. Formaliza-se esta afirmação na seguinte definição:

Definição 3: Um sistema normalizado de aritmética de ponto flutuante consiste de zeros e todos os números que podem ser representados na forma  $\pm(0, d_1 d_2 \dots d_t)_\beta \cdot \beta^{\text{exp}}$ , onde  $\pm$  é um inteiro maior ou igual que 2 denominado base do sistema,  $t$  é um inteiro maior ou igual que 1 denominado número de dígitos do sistema,  $d_j$  são inteiros tais que  $0 \leq d_j \leq \beta - 1, \forall j = 2, \dots, t$ , e  $0 < d_1 < \beta - 1$ , e  $\text{exp}$  é um inteiro tal que satisfaz  $m < \text{exp} < M$ , sendo  $m < M$  dois inteiros,  $\text{exp}$  é denominado expoente do número. Tal sistema é representado por  $F(\beta, t, m, M)$ .

Em um sistema normalizado de aritmética de ponto flutuante, a parte fracionária de um número na base,  $(d_1 d_2 \dots d_t)$  é denominada mantissa.

Os parâmetros  $t, m$  e  $M$  dependem exclusivamente do computador.

Máquina		$t$	$m$	$M$
Burroughs 5500	8	13	-51	77
Burroughs 6700	8	13	-63	63
HP 45	10	10	-98	100
Texas SR-5X	10	12	-98	100
PDP-11	2	24	-128	127
IBM/360	16	6	-64	63
IBM/370	16	14	-64	63
Quartzil QI 800	2	24	-127	127

No sistema  $F(\beta, t, m, M)$  podem ser observadas as seguintes características:

1. O menor número positivo possível de ser representado neste sistema é

$$X_{\text{MIN}} = (0,10\dots0)_{\beta} \times \beta^m = \beta^{m-1}$$

2. A região de “underflow” é definida como o intervalo  $(-X_{\text{MIN}}, X_{\text{MIN}})$ , isto é, qualquer número real cujo valor verdadeiro está nesse intervalo, será considerado zero neste sistema.

3. O maior número positivo possível de ser representado neste sistema é

$$x_{\text{MAX}} = (0,(\beta-1)(\beta-1)\dots(\beta-1))_{\beta} \cdot \beta^M = \frac{\beta^t - 1}{\beta^t} \beta^M$$

4. A região do chamado “overflow” é definida como sendo a união de dois intervalos  $(-\infty, -X_{\text{MAX}}) \cup (X_{\text{MAX}}, \infty)$ , isto é, qualquer número real cujo valor verdadeiro está nessa união de intervalos, será considerado  $-1$  ou  $+1$ , dependendo de se o número está no primeiro ou no segundo intervalo, respectivamente.

5. O número total de elementos do sistema é  $2(\beta-1)(M-m+1)\beta^{t-1} + 1$ .

**Exemplo 13:** Considere o sistema  $F(2, 3, -1, 2)$ . Temos que:

O menor elemento positivo deste sistema é  $X_{\text{MIN}} = (0,100)_2 \times 2^{-1} = 2^{-1-1} = 2^{-2} = 1/4$ , portanto a região de *underflow* é  $(-1/4; 1/4)$ .

O maior elemento positivo deste sistema é  $X_{\text{MAX}} = (0,111)_2 \cdot 2^2 = \frac{2^3 - 1}{2^3} \cdot 2^2 = \frac{7}{2}$ , portanto, a região de “overflow” deste sistema é  $(-\infty, -7/2) \cup (7/2, +\infty)$ .

Mais ainda, como os valores dos parâmetros do sistema são pequenos, os valores na base decimal dos elementos positivos podem ser enumerados em um arranjo. Segue a tabela que demonstra isso:

Mantissa ↓	exp=-1	exp=0	exp=1	exp=2
100	1/4	1/2	1	2
101	5/16	5/8	5/4	5/2
110	3/8	3/4	3/2	3
111	7/16	7/8	7/4	7/2

O número de elementos do sistema  $F(2, 3, -1, 2)$  é

$$2(\beta-1)(M-m+1)\beta^{t-1} + 1 = 2 \cdot (2-1)(2+1+1) \cdot 2^{3-1} + 1 = 33$$

# ANOTAÇÕES

---

Lined area for notes, consisting of multiple horizontal blue lines.

Equações Algébricas e  
Transcendentais

2  
*Capítulo*

As **equações algébricas** são igualdades que dependem de uma ou mais variáveis relacionadas mediante um número finito de adições, subtrações, multiplicações, divisões e potenciação e radiciação inteiras. As operações mencionadas anteriormente são denominadas **operações algébricas**. Exemplos de equações algébricas são

$$\sqrt{x^2 - x + 1} + \frac{x}{\sqrt[3]{x^2 + 1}} - 2 = 0, \quad x^{3/5} - y^2 + \frac{1}{z} + 5 = 0.$$

Neste capítulo, estudar-se-ão as **equações polinomiais** de uma variável, ou seja, equações da forma  $a_n x^n + \dots + a_1 x + a_0 = 0$ , sendo  $a_n, \dots, a_1, a_0$  números reais. As equações polinomiais são casos particulares de equações algébricas.

Uma **equação transcendental** de uma ou mais variáveis é uma igualdade que não pode ser reduzida a uma equação algébrica. Por exemplo,  $x^{3/8} - \cos(x) - 1 = 0$  ou  $e^{x+\sqrt{y}} - \ln(2xy) = 0$ . Em geral, as equações transcendentais envolvem as assim chamadas **funções transcendentais**: funções exponenciais, logarítmicas, trigonométricas e hiperbólicas, entre outras. Os métodos numéricos desenvolvidos neste capítulo servirão apenas para funções transcendentais de uma variável.

## 2.1 Introdução

Uma equação (algébrica ou transcendental) de uma variável pode ser escrita sempre como  $f(x) = 0$ , sendo  $f(x)$  uma função na variável escalar  $x$ .

### Definição 2.1: (Solução, raiz ou zero de uma equação)

Um número  $\xi$  é uma **solução (raiz ou zero)** de uma equação  $f(x) = 0$  se  $f(\xi) = 0$ , ou seja, se a igualdade  $f(x) = 0$  é válida quando  $x = \xi$ .

### Exemplo 2.1:

Considere a equação  $x^5 - 4x^4 + 2x + 1 = 0$  na variável  $x$ . Então observe que o número  $\xi = 1$  é uma raiz de tal equação, pois quando substitui-se  $x$  por 1 na equação dada, tem-se que  $1^5 - 4 \cdot 1^4 + 2 \cdot 1 + 1 = 1 - 4 + 2 + 1 = 0$ .

### Definição 2.2: (Multiplicidade algébrica de uma raiz)

Uma raiz  $\xi$  da equação  $f(x) = 0$  é de **multiplicidade  $m$**  se as derivadas até de ordem  $m$  de  $f(x)$  em  $x = \xi$  existem  $f(\xi) = f'(\xi) = \dots = f^{(m-1)}(\xi) = 0$  e  $f^{(m)}(\xi) \neq 0$ .

### Exemplo 2.2:

Sabe-se, pelo exemplo anterior, que  $\xi = 1$  é uma raiz da equação polinomial  $x^5 - 4x^4 + 2x + 1 = 0$ . Considerando  $f(x) = x^5 - 4x^4 + 2x + 1$  tem-se que  $f(\xi) = 0$  e  $t_1(x) = 4x^3 - 15x^2 + 18x - 7$ . Assim,  $\xi = 1$  é uma raiz de multiplicidade  $m = 1$ . Neste caso, diz-se que  $\xi = 1$  é uma **raiz simples**.

### Exemplo 2.3:

Pode-se verificar que  $\xi = 1$  é uma raiz da equação polinomial  $x^5 - 2x^4 + x^3 - x^2 + 2x - 1 = 0$ . Seja  $f(x) = x^5 - 2x^4 + x^3 - x^2 + 2x - 1$ . Assim,

$k$	$f^{(k)}(x)$	$f^{(k)}(\xi)$
0	$x^5 - 2x^4 + x^3 - x^2 + 2x - 1$	0
1	$5x^4 - 8x^3 + 3x^2 - 2x + 2$	0
2	$ax^3 + bx^2 + cx + d = 0$	0
3	$60x^2 - 24x + 6$	42

Portanto,  $\xi = 1$  é uma raiz de multiplicidade  $m = 3$ .

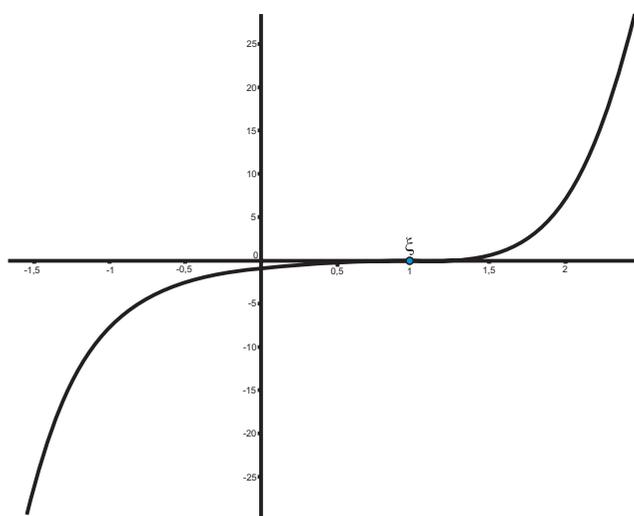
**Exemplo 2.4:**

Pode-se verificar que  $\xi = 0$  é uma raiz da equação transcendental  $x \operatorname{sen}(x) = 0$ . Seja  $f(x) = x \operatorname{sen}(x)$ . Assim,

$k$	$f^{(k)}(x)$	$f^{(k)}(\xi)$
0	$x \operatorname{sen}(x)$	0
1	$\operatorname{sen}(x) + x \cos(x)$	0
2	$2 \cos(x) - x \operatorname{sen}(x)$	2

Portanto,  $\xi = 0$  é uma raiz de multiplicidade  $m = 2$ .

**Observação 2.1:** No plano cartesiano, a raiz  $\xi$  da equação  $f(x) = 0$  corresponde ao intercepto do gráfico de  $f(x)$  com o eixo  $x$ . Por exemplo, considere a equação  $x^5 - 2x^4 + x^3 - x^2 + 2x - 1 = 0$ . Utilizando um recurso computacional, obtém-se

**ATIVIDADE 1**

1) Classifique as seguintes equações em algébricas (polinomial ou não) ou transcendentais, Justifique sua resposta.

$$\sqrt{x} + \sqrt[3]{x} = 1,$$

$$\sqrt{x} + \sqrt[3]{x \cos(x)} = 1,$$

$$e^2 x + x^3 + x^5 = -3,$$

$$e^x x + x^3 + x^5 = -3,$$

$$1 + x + x^2 + x^3 + \dots + x^n = \sqrt{2},$$

$$1 + x + x^2 + x^3 + \dots + x^n + \dots = \sqrt{2}.$$

2) Verifique se o valor de  $x$  dado é solução da equação respectiva.

$$x = 0 \text{ em } x \cdot \ln(1+x) = 0,$$

$$x = 1 \text{ em } (x-1)^2 \cdot \ln(1+x) = 0,$$

$$x = 1 \text{ em } x^9 - 8x^3 - 7 = 0,$$

$$x = -1 \text{ em } x^9 - 8x^3 - 7 = 0,$$

$$x = 2 \text{ em } x^6 - 7x^5 + 15x^4 - 40x^2 + 48x - 6 = 0.$$

3) Determine a multiplicidade algébrica das seguintes raízes em relação à equação respectiva.

$$x = 1 \text{ em } (x-1)^2 \cdot \ln(1+x) = 0,$$

$$x = 0 \text{ em } x \cdot \ln(1+x) = 0,$$

$$x = 0 \text{ em } x \cdot \cos(x) = 0,$$

$$x = 1 \text{ em } x^6 - 2x^5 - 3x^4 + 12x^3 - 13x^2 + 6x - 1 = 0.$$

## 2.2 Métodos Diretos e Indiretos

Os métodos para resolver equações (ou sistemas de equações) algébricas e transcendentais de uma (ou várias) podem ser classificados em diretos e indiretos.

### 2.2.1 Métodos Diretos

Os métodos diretos são de natureza analítica, ou seja, sob certas condições, eles fornecem todas as soluções de uma equação mediante fórmulas algébricas fechadas.

1) Dada uma *equação linear* ou *de primeiro grau* em  $x$ ,  $ax + b = 0$ , com  $a \neq 0$ , a solução de tal equação é

$$x = -\frac{b}{a}.$$

2) Dada uma *equação quadrática* ou *de segundo grau* em  $x$ ,  $ax^2 + bx + c = 0$  com  $a \neq 0$ , as soluções

(em geral complexas) estão dadas pela fórmula  $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ .

3) Dada uma *equação cúbica* ou *de terceiro grau* em  $x$ ,  $ax^3 + bx^2 + cx + d = 0$  com  $a \neq 0$ , as soluções podem ser obtidas mediante procedimentos algébricos baseados na fórmula de Tartália para equações do tipo  $y^3 + py + q = 0$ .

4) Dada uma *equação quártica* ou *de quarto grau* em  $x$ ,  $ax^4 + bx^3 + cx^2 + dx + e = 0$  com  $a \neq 0$ , as soluções podem ser obtidas mediante procedimentos algébricos baseados no método de Ferrari para equações do tipo  $y^3 + py + q = 0$ .

5) O matemático Abel provou que não existe uma fórmula geral envolvendo operações algébricas para encontrar as soluções de uma equação polinomial de grau maior que 4.

### 2.2.2 Métodos Indiretos

Os métodos indiretos são de natureza iterativa, ou seja, geram, em cada execução, mediante equações em diferença, uma seqüência de números,  $\{x_k\}$ , denominados *aproximantes* ou *iteradas*, tais que  $\lim_{k \rightarrow +\infty} x_k = \xi$ , sendo  $\xi$  uma raiz da equação que se pretende resolver.

### Definição 2.3: (Método Iterativo de $m$ pontos)

Um método iterativo de  $m$  pontos está definido mediante uma equação recursiva e certas condições iniciais como segue:

$$\begin{cases} x_0, x_1, \dots, x_{m-1} \text{ números conhecidos, e} \\ x_{k+1} = \phi(x_k, x_{k-1}, \dots, x_{k-m+1}) \text{ para } k = m-1, m, m+1, \dots \end{cases}$$

onde  $\phi$  é uma função real de  $m$  variáveis reais denominada a *função iterativa* do método e os valores  $x_0, x_1, \dots, x_{m-1}$  formam o conjunto de *condições iniciais* do método.

O processo de encontrar iterativamente uma raiz  $\xi$  da equação  $f(x) = 0$  pode ser dividido em duas etapas:

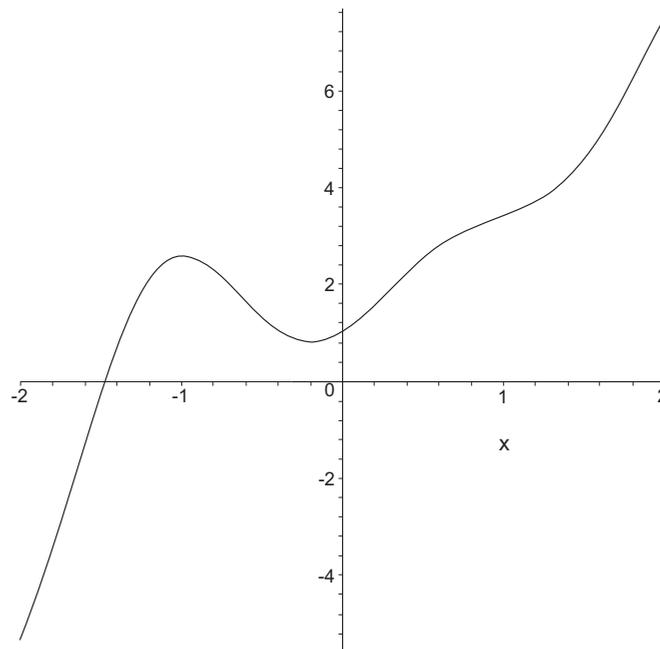
1. determinação de um conjunto de  $m$  aproximantes da raiz  $\xi$ , e
2. escolha de uma certa função iterativa de  $m$  pontos que gere a seqüência finita de aproximantes  $\{x_k\}$  de  $\xi$ .

### 2.2.3 Aproximações Iniciais

Com freqüência, as aproximações iniciais de uma raiz  $\xi$  são tomadas a partir de considerações particulares do problema. Mas também existem recursos tais como a localização das raízes mediante critérios analíticos ou gráficos. Por exemplo, na equação  $f(x) = 0$ , se é tomado um certo valor  $x_0$  tal que  $f(x_0) \approx 0$ , então o valor  $x_0$  pode fornecer uma aproximação inicial adequada. Se estão disponíveis recursos gráficos, outra abordagem interessante é tomar como aproximação inicial algum valor  $x_0$  tal que o gráfico da função  $f(x)$  corte o eixo  $x$  bem perto de  $x_0$ .

**Exemplo 2.5:** Seja  $f(x) = \cos(3x) + x\sqrt{x^2 + 1}$ . Observe que  $f(-0,4) = -0,0684554299 \approx 0$ . Assim,  $x_0 = -0,4$  pode ser uma boa aproximação inicial da raiz da equação  $\cos(3x) + x\sqrt{x^2 + 1} = 0$ .

**Exemplo 2.6:** Se o gráfico de certa função é apresentado na figura a seguir



pode-se dizer que uma boa aproximação inicial da solução da equação  $f(x) = 0$  está dada por  $x_0 = -1,5$ .

Outro método muito comum utilizado para obter uma aproximação inicial de uma raiz está baseado no Teorema do Valor Intermediário:

**Teorema 2.1:** Se  $f(x)$  é uma função contínua no intervalo fechado  $[a, b]$  e  $f(a)f(b) < 0$ , então a equação  $f(x) = 0$  tem pelo menos uma raiz no intervalo aberto  $(a, b)$ .

Uma abordagem interessante para determinar uma aproximação inicial é construir uma tabela de valores da função  $f(x)$  e observa os intervalos onde acontece a mudança de sinal da função.

**Exemplo 2.7:** Se  $f(x) = e^{-x} - \text{sen}(x/2)$ , então tem-se a tabela de valores

$x$	$f(x)$
-4	55,507
-3	21,083
-2	8,231
-1	3,198
0	1
1	-0,112
2	-0,706
3	-0,948
4	-0,891

Observa-se que a função  $f(x)$  é contínua e existe uma mudança de sinal entre  $x = 0$  e  $x = 1$ , logo uma aproximação inicial adequada da raiz da equação  $e^{-x} - \text{sen}(x/2) = 0$  seria  $x_0 = 1$  que dá o valor mais próximo de zero.

#### 2.2.4 Geração de Aproximantes e Critérios de Parada

A partir de uma aproximação inicial  $x_0$ , um método iterativo gera uma seqüência de aproximantes  $x_1, x_2, \dots, x_k, \dots$  de uma raiz da equação  $f(x) = 0$ . Devido à limitações físicas, tal seqüência não pode ser infinita.

Então existe a necessidade de estabelecer critérios de parada para construir uma seqüência finita de aproximantes  $x_1, x_2, \dots, x_k$ . Entre os critérios mais utilizados tem-se os seguintes

- 1) Número Fixo de Iterações: ou seja, o método iterativo deve gerar um número prescrito,  $N$ , de aproximantes.
- 2) Erro Absoluto prescrito do Aproximante: dada uma tolerância de erro,  $\varepsilon$ , deve-se parar se  $|x_k - x_{k-1}| < \varepsilon$ .
- 3) Erro Relativo prescrito do Aproximante: dada uma tolerância de erro,  $\varepsilon$ , deve-se parar se  $\left| \frac{x_k - x_{k-1}}{x_k} \right| < \varepsilon$ .
- 4) Número de Dígitos Significativos Exatos do Aproximante: prescrito um número de dígitos significativos

exatos,  $M$ , deve-se parar quando  $\left| \frac{x_k - x_{k-1}}{x_k} \right| < 0,5 \times 10^{-M+1}$ .

## ATIVIDADE 2

1) Faça uma tabela de valores para as funções  $f(x)$ , para  $x$  inteiro de  $-5$  a  $5$  para descobrir um valor inicial adequado  $x_0$  que pode ser utilizado para começar um método numérico que aproxime uma raiz da equação  $f(x) = 0$ .

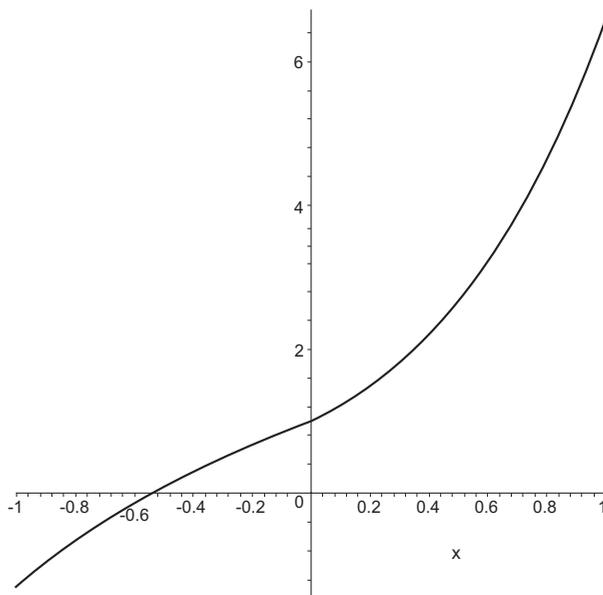
a)  $f(x) = x^6 - x^5 - 2$

b)  $f(x) = x^6 + x^5 - 2e^x$

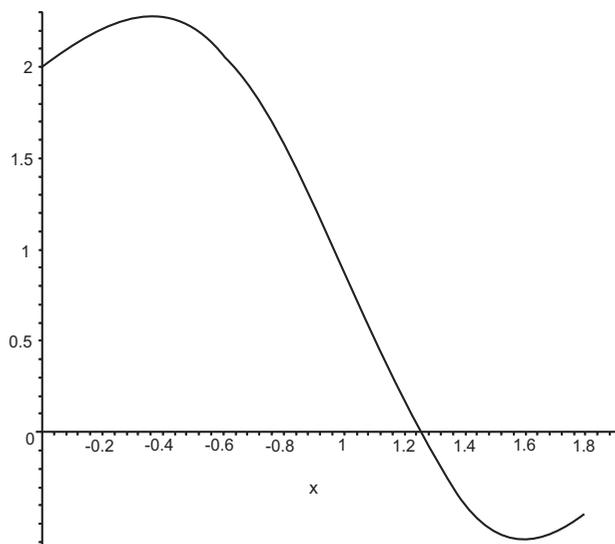
c)  $f(x) = 3e^x - x \cos(2x)$

2) Dados os seguintes gráficos da função  $f(x)$ , após observar o intercepto do gráfico com o eixo  $x$  diga qual pode ser um valor inicial adequado

a)

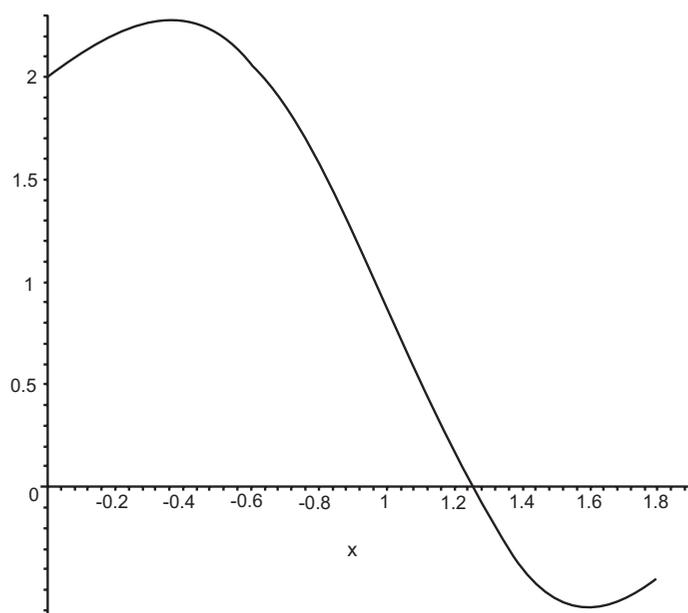


b)

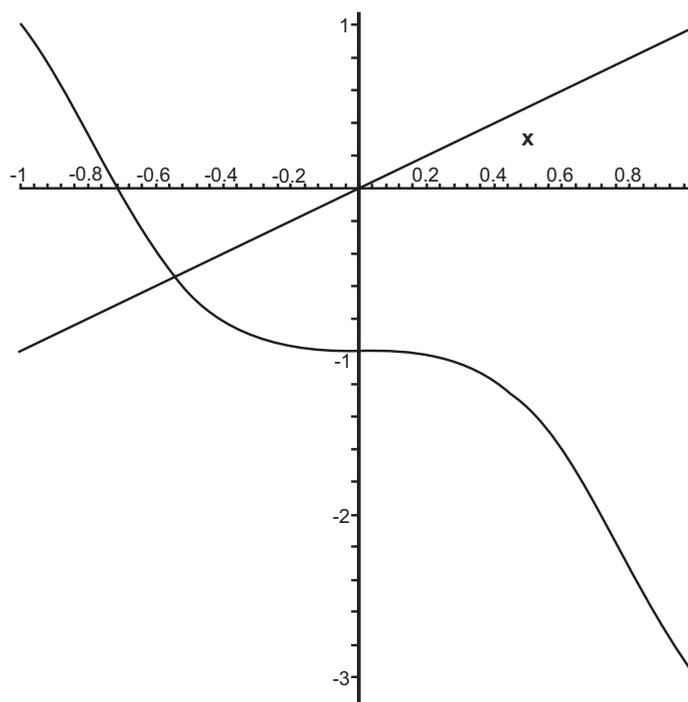


Dados os gráficos de  $x$  e uma função  $g(x)$  diga qual seria um valor inicial adequado para resolver a equação  $g(x) = x$  :

a)



b)



4) Se um aproximante de uma raiz da equação  $x^5 - x - 1 = 0$  é  $1,167303542$  com um erro relativo não maior que  $0,0002321$ , quantos dígitos significativos exatos possui tal aproximante?

5) Faça uma pesquisa bibliográfica ou na internet, para detalhar as fórmulas de Tartália e resolva algebricamente a equação  $x^3 - 4x^2 + 5x - 1 = 0$ .

## 2.3 Equações Algébricas Polinomiais

Nesta seção, o interesse está em obter as raízes de uma equação polinomial

$$p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$$

onde  $a_n \neq 0$ ,  $a_{n-1}$ ,  $a_1$ ,  $a_0$  são números reais.

Existem critérios que permitem conhecer o número de raízes reais (positivas e negativas) e complexas, ou seja, classificar as raízes. Também, existem critérios para saber a localização das raízes. Este será objeto de nosso estudo nas seguintes subseções.

### 2.3.1 Teorema de Lagrange

É possível determinar intervalos que contem as raízes reais de uma equação polinomial mediante o seguinte teorema:

**Teorema 2.2: (Teorema de Lagrange)** Seja a equação polinomial  $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$  onde  $a_n > 0$ . Então uma cota superior para as raízes positivas da equação polinomial está dada por

$$L = 1 + \sqrt[n-k]{\frac{B}{a_n}}, \text{ onde } B \text{ é igual ao maior valor absoluto dos coeficientes negativos existentes no}$$

polinômio ou zero em caso contrário (se não existem coeficientes negativos) e  $k$  é a maior potência de  $x$  correspondente a um coeficiente negativo. O número  $L$  é denominado a **cota de Lagrange** da equação  $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$ .

**Exemplo 2.8:** Considere a equação polinomial  $x^6 - 3x^4 + 8x^3 - 5x^2 + x - 4 = 0$ . Neste caso, de acordo com as condições do teorema,  $B = 5$  e  $k = 4$ . Observe que  $n = 6$ . Assim, uma cota superior para as raízes

$$\text{positivas é } L = 1 + \sqrt[6-4]{\frac{5}{1}} = 1 + \sqrt{5}.$$

**Corolário 2.1:** Seja a equação polinomial  $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$ , onde  $a_n > 0$ .

1) Se  $L_1$  a cota de Lagrange da equação polinomial  $x^n p_n(1/x) = 0$ , então uma cota inferior positiva das raízes positivas da equação  $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$  é  $\frac{1}{L_1}$ .

2) Se  $L_2$  a cota de Lagrange da equação polinomial  $p_n(-x) = 0$ , então uma cota inferior negativa das raízes negativas da equação  $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$  é  $-L_2$ .

3) Se  $L_3$  a cota de Lagrange da equação polinomial  $x^n p_n(-1/x) = 0$ , então uma cota superior negativa das raízes negativas da equação  $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$  é  $-\frac{1}{L_3}$ .

Assim, todas as raízes positivas da equação polinomial  $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$  encontram-se no intervalo  $\left[\frac{1}{L_1}, L\right]$ , e todas as raízes negativas, encontram-se no intervalo  $\left[-L_2, -\frac{1}{L_3}\right]$ .

**Exemplo 2.9:** Seja a equação polinomial  $p_n(x) = x^4 - 7x^3 - 9x^2 + 30x + 31 = 0$ . Então

$$x^4 p_n(1/x) = 31x^4 + 30x^3 - 9x^2 - 7x + 1,$$

$$p_n(-x) = x^4 + 7x^3 - 9x^2 - 30x + 31 \text{ e}$$

$$x^4 p_n(-1/x) = 31x^4 - 30x^3 - 9x^2 + 7x + 1$$

Então

$$L = 1 + 4\sqrt[3]{\frac{9}{1}} = 10, \quad L_1 = 1 + 4\sqrt[3]{\frac{9}{31}} \approx 1,5388 \Rightarrow \frac{1}{L_1} = 0,6498,$$

$$L_2 = 1 + 4\sqrt[3]{\frac{30}{1}} \approx 6,4772 \Rightarrow -L_2 = -6,4772 \text{ e } L_3 = 1 + 4\sqrt[3]{\frac{30}{31}} \approx 1,9677 \Rightarrow -\frac{1}{L_3} \approx -0,5081.$$

Logo, as raízes positivas da equação polinomial  $p_n(x) = x^4 - 7x^3 - 9x^2 + 30x + 31 = 0$ , se existirem, encontram-se no intervalo  $[0,6498;10]$ , e as raízes negativas da mesma equação, se existirem, encontram-se no intervalo  $[-6,4772;-0,5081]$ .

**Exemplo 2.10:** Seja a equação polinomial  $p_n(x) = 3x^7 + x^5 - 3x^3 + 8x^2 + 10x + 1 = 0$ . Então

$$x^7 p_n(1/x) = x^7 + 10x^6 + 8x^5 - 3x^4 + x^2 + 3,$$

$$p_n(x) = 3x^7 + x^5 - 3x^3 + 8x^2 + 10x + 1 = 0 \text{ e}$$

$$x^7 p_n(-1/x) = x^7 - 10x^6 + 8x^5 + 3x^4 - x^2 - 3$$

Observe que para satisfazer as condições do teorema de Lagrange, o sinal de  $p_n(-x)$  foi trocado. Então

$$L = 1 + 7\sqrt[3]{\frac{3}{3}} = 2, \quad L_1 = 1 + 7\sqrt[3]{\frac{3}{1}} \approx 2,4422 \Rightarrow \frac{1}{L_1} = 0,4095,$$

$$L_2 = 1 + 7\sqrt[3]{\frac{8}{3}} \approx 2,2779 \Rightarrow -L_2 = -2,2779 \text{ e } L_3 = 1 + 7\sqrt[3]{\frac{10}{1}} \approx 11 \Rightarrow -\frac{1}{L_3} \approx -0,0909$$

Logo, as raízes positivas da equação polinomial  $p_n(x) = 3x^7 + x^5 - 3x^3 + 8x^2 + 10x + 1 = 0$ , se existirem, encontram-se no intervalo  $[0,4095;2]$ , e as raízes negativas da mesma equação, se existirem, encontram-se no intervalo  $[-2,2779;-0,0909]$ .

### 2.3.2 Regra de Sinais de Descartes

Quanto ao número de raízes (reais ou complexas), sabe-se, pelo Teorema Fundamental da Álgebra, que uma equação polinomial  $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0$ ,  $a_n \neq 0$ , possui  $n$  raízes reais ou complexas, incluindo a multiplicidade algébrica delas.

Mas nesta subseção, são destacados alguns critérios mais específicos quanto ao número de raízes reais, positivas e negativas de uma equação polinomial.

**Definição 2.4:** Considere um polinômio  $p_n(x)$  com os termos escritos em ordem decrescente em relação à potência de  $x$ , sem termos com coeficientes nulos. Então se os coeficientes de dois termos  $a_r x^r$  e  $a_s x^s$  consecutivos, sendo  $r > s$ , tem sinais iguais, diz-se que aconteceu uma **permanência de sinal** entre tais termos, mas se estes coeficientes tem sinais diferentes, diz-se que ocorreu uma **mudança de sinal** entre esses termos.

**Exemplo 2.11:** O polinômio  $p_n(x) = 3x^7 - 12x^4 + 10x^3 + 14x - 1$  tem três mudanças de sinais e uma continuação de sinais. De fato, as três mudanças acontecem, a primeira, entre os termos  $3x^7$  e  $-12x^4$ , a segunda, entre os termos  $-12x^4$  e  $10x^3$  e a terceira, entre os termos  $-22$  e  $-1$ . A permanência de sinais acontece entre os termos  $10x^3$  e  $-22$ .

**Teorema 2.3: (Regra de sinais de Descartes)**

1) O número de raízes positivas de uma equação polinomial  $p_n(x) = 0$  é igual ao número de mudanças de sinais em  $p_n(x)$  ou este número menos um número par.

2) O número de raízes negativas de uma equação polinomial  $p_n(x) = 0$  é igual ao número de mudanças de sinais em  $p_n(-x)$  ou este número menos um número par.

**Exemplo 2.12:** O polinômio da equação polinomial  $p_5(-x) = -8x^5 + 12x^4 + 10x^3 + 17x^2 + 18x + 5$  tem quatro mudanças de sinais, e portanto, tem 4, 2 ou 0 raízes positivas. O polinômio  $p_5(-x) = -8x^5 + 12x^4 + 10x^3 + 17x^2 + 18x + 5$  tem uma mudança de sinal, e portanto, uma raiz real negativa.

Observa-se que o critério de Descartes proporciona somente uma possível contagem das raízes positivas ou negativas de uma equação polinomial. Porém, existe um critério mais preciso, embora mais trabalhoso, para saber com precisão o número de raízes reais. Isto será visto na próxima subseção.

**2.3.3 Seqüência de Sturm**

**Definição 2.6:** Seja  $p_n(x) = 0$  uma equação polinomial com  $n$  raízes simples. Denote por  $s_0(x) = p_n(x)$ ,  $s_1(x) = D_x(p_n(x))$ ,  $s_2(x)$  o negativo do resto de dividir  $s_0(x)$  por  $s_1(x)$ ,  $s_3(x)$  o negativo do resto de dividir  $s_1(x)$  por  $s_2(x)$ , e assim por diante, até chegar em  $s_k(x) = c \neq 0$ , sendo  $c$  uma constante. A seqüência de polinômios  $\{s_0(x), s_1(x), \dots, s_k(x)\}$  é denominada a **seqüência de Sturm** de  $p_n(x)$ .

No caso que a equação polinomial  $p_n(x) = 0$  tenha uma raiz múltipla  $x = x_0$ , pode-se conseguir ainda uma seqüência de polinômios não nulos como segue, seja  $t_0(x) = p_n(x)$ ,  $t_1(x) = D_x(p_n(x))$ ,  $t_2(x)$  o negativo do resto de dividir  $t_0(x)$  por  $t_1(x)$ ,  $t_3(x)$  o negativo do resto de dividir  $t_1(x)$  por  $t_2(x)$ , e assim por diante, até chegar em  $t_k(x)$  igual ao máximo comum divisor de  $t_0(x)$  por  $t_1(x)$ . Neste caso a seqüência da Sturm está dada por:

$$s_0(x) = \frac{t_0(x)}{t_k(x)}, \quad s_1(x) = \frac{t_1(x)}{t_k(x)}, \quad \dots, \quad s_{k-1}(x) = \frac{t_{k-1}(x)}{t_k(x)}, \quad s_k(x) = 1.$$

**Exemplo 2.13:** Seja  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1$ . A seqüência de Sturm de  $p_n(x)$  está formada

por  $s_0(x) = x^4 - 2x^3 - 6x^2 + 2x - 1$ ,  $s_1(x) = 4x^3 - 6x^2 - 12x + 2$ ,  $s_2(x) = \frac{15}{4}x^2 + \frac{3}{4}$ ,

$s_3(x) = \frac{64}{5}x - \frac{16}{5}$  e  $s_4(x) = -\frac{63}{64}$ .

**Exemplo 2.14:** Seja  $p_n(x) = x^4 - 5x^3 + 9x^2 - 7x + 2$ . Tenta-se construir a seqüência de Sturm de  $p_n(x)$ :  
 $t_0(x) = x^4 - 5x^3 + 9x^2 - 7x + 2$ ,  $t_1(x) = 4x^3 - 15x^2 + 18x - 7$ ,  $t_2(x) = \frac{3}{16}(x^2 - 2x + 1)$ ,

$t_3(x) = 0$ . Assim, a seqüência de Sturm é:  $s_0(x) = \frac{t_0(x)}{t_2(x)} = \frac{16}{3}(x^2 - 3x + 2)$ ,  
 $s_1(x) = \frac{t_1(x)}{t_2(x)} = \frac{16}{3}(4x - 7)$  e  $s_2(x) = 1$ .

**Teorema 2.4:** (Teorema de Sturm) O número de raízes reais da equação polinomial  $p_n(x) = 0$  no intervalo  $[a, b]$  é igual ao valor absoluto da diferença entre o número de mudanças de sinal dos valores da seqüência de Sturm em  $x = a$  e  $x = b$ , sempre que  $f(a) \neq 0$  e  $f(b) \neq 0$ .

**Exemplo 2.15:** Seja a equação polinomial  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$ . Sabe-se, pelo exemplo 2.13, que a seqüência de Sturm de  $p_n(x)$  está formada por  $s_0(x) = x^4 - 2x^3 - 6x^2 + 2x - 1$ ,  
 $s_1(x) = 4x^3 - 6x^2 - 12x + 2$ ,  $s_2(x) = \frac{15}{4}x^2 + \frac{3}{4}$ ,  $s_3(x) = \frac{64}{5}x - \frac{16}{5}$  e  $s_4(x) = -\frac{63}{64}$ . Deseja-se conhecer o número de raízes que se encontram no intervalo  $[-2, 2]$ . Faz-se uma tabela com os valores da seqüência em  $x = -2$  e  $x = 2$ :

	$x = -2$	$x = 2$
$s_0(x)$	3	-21
$s_1(x)$	-30	-14
$s_2(x)$	63/4	63/4
$s_3(x)$	-144/5	112/5
$s_4(x)$	-63/64	-63/64

Em  $x = -2$  existem 3 mudanças de sinal nos valores da seqüência de Sturm, e em  $x = 2$  existem 2 mudanças de sinal. Isto significa que no intervalo  $[-2, 2]$ , existe  $3 - 2 = 1$  raiz.

Uma estratégia interessante para localizar e enumerar as raízes de uma equação polinomial é estabelecer intervalos onde tais raízes estão localizadas mediante as cotas de Lagrange e depois utilizar a seqüência de Sturm para contar as raízes. Isto é ilustrado no exemplo a seguir.

**Exemplo 2.16:** Seja a equação polinomial  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$ . Primeiro, vão ser determinados os intervalos para as raízes positivas e negativas, mediante as cotas de Lagrange. Uma cota

superior para as raízes positivas é  $L = 1 + \sqrt[4]{\frac{6}{1}} = 7$ . Agora, tem-se que  $p_n(-x) = x^4 + 2x^3 - 6x^2 - 2x - 1$

e uma cota inferior para as raízes negativas é  $-L_2 = -\left(1 + \sqrt[4]{\frac{6}{1}}\right) = -3,4495$ . Sabe-se, pelo exemplo

2.13, que a seqüência de Sturm de  $p_n(x)$  está formada por  $s_0(x) = x^4 - 2x^3 - 6x^2 + 2x - 1$ ,

$s_1(x) = 4x^3 - 6x^2 - 12x + 2$ ,  $s_2(x) = \frac{15}{4}x^2 + \frac{3}{4}$ ,  $s_3(x) = \frac{64}{5}x - \frac{16}{5}$  e  $s_4(x) = -\frac{63}{64}$ . Faz-se uma

tabela com os valores da seqüência em  $x = -4$ ,  $x = 0$  e  $x = 7$ :

	$x = -4$	$x = 0$	$x = 7$
$s_0(x)$	279	-1	1434
$s_1(x)$	-302	2	996
$s_2(x)$	$\frac{243}{4}$	$\frac{3}{4}$	$\frac{369}{2}$
$s_3(x)$	$-\frac{272}{5}$	$-\frac{16}{5}$	$\frac{432}{5}$
$s_4(x)$	$-\frac{63}{64}$	$-\frac{63}{64}$	$-\frac{63}{64}$

Em  $x = -4$  existem 3 mudanças de sinal nos valores da seqüência de Sturm, em  $x = 0$  existem 2 mudanças de sinal e em  $x = 7$ , existe 1 mudança de sinal. Isto significa que no intervalo  $[-4,0]$ , existe  $3 - 2 = 1$  raiz e no intervalo  $[0,7]$  existe  $2 - 1 = 1$  raiz.

Resumindo, pode-se dizer que existem duas raízes reais e duas raízes complexas (quatro ao todo), de dentre as duas raízes reais, uma é positiva e uma é negativa.

Até aqui, o problema de resolver uma equação polinomial está parcialmente abordado. Resta ainda calcular as raízes. Na próxima seção serão detalhados métodos para aproximar raízes de equações transcendentais, mas estes métodos podem ser utilizados também para aproximar raízes de equações polinomiais.

### ATIVIDADE 3

1) Determine as cotas de Lagrange para as possíveis raízes reais das seguintes equações polinomiais:

$$p_n(x) = x^6 - 3x^4 - 5x^2 - 8x - 1 = 0,$$

$$p_n(x) = x^7 + x^6 - x^5 - x^4 - x^3 - 2x^2 + x - 1 = 0,$$

$$p_n(x) = -x^6 - 2x^5 + x^2 + x - 1 = 0,$$

$$p_n(x) = -x^7 - 2x^2 - x - 3 = 0,$$

$$p_n(x) = x^7 - x + 2 = 0.$$

2) Mediante a regra de sinais de Descartes, forneça estimativas sobre o número de raízes positivas e negativas das equações polinomiais da questão 1)

3) Mediante as cotas de Lagrange e a seqüência de Sturm de cada polinômio, determine o número exato de raízes positivas e negativas das equações polinomiais da questão 1)

## 2.4 Equações Transcendentais

Lembre que uma equação transcendental é uma equação da forma  $f(x) = 0$ , onde  $f(x)$  contém funções trigonométricas, exponenciais, logarítmicas, hiperbólicas, etc.

Neste caso, uma maneira geral de localizar a raiz é elaborando uma tabela

$x$	$f(x)$
$x_1$	$f(x_1)$
$\vdots$	$\vdots$
$x_n$	$f(x_n)$

onde os valores  $x_1, \dots, x_n$  são escolhidos, geralmente em forma crescente. Se a função  $f(x)$  é contínua, procura-se uma mudança de sinal entre os valores da função  $f(x)$ , e de acordo com o Teorema do Valor Intermediário, existirá uma raiz no intervalo  $[x_i, x_j]$ , sendo que  $f(x_i)f(x_j) < 0$ .

Nesta seção estamos interessados em métodos iterativos que calculem uma seqüência de aproximantes  $\{x_1, x_2, \dots, x_k\}$  de uma raiz da equação  $f(x) = 0$ . Estes métodos, bem como a idéia que fundamenta cada um deles, serão expostos nas seguintes subseções.

### 2.4.1 Método da Bissecção

Seja a equação  $f(x) = 0$  com  $f(x)$  contínua. Suponha que são conhecidos dois valores  $a_0$  e  $b_0$ , com  $a_0 < b_0$ , tais que  $f(a_0)f(b_0) < 0$ . De acordo com o Teorema do Valor Intermediário, existe pelo menos uma raiz no intervalo  $(a_0, b_0)$ . A idéia do método da bissecção é calcular um aproximante de alguma raiz que se encontre no referido intervalo como sendo o ponto médio do intervalo,  $x_0 = \frac{a_0 + b_0}{2}$ . Depois, pode-se determinar se existe mudança de sinal do valor da função  $f(x)$  no intervalo  $[a_0, x_0]$  ou  $[x_0, b_0]$ ; escolhe-se o intervalo  $[a_0, x_0]$  se  $f(a_0)f(x_0) < 0$  ou o intervalo  $[x_0, b_0]$  se  $f(x_0)f(b_0) < 0$ . Chamando de  $[a_1, b_1]$  o intervalo onde acontece a mudança de sinal, calcula-se um novo aproximante como  $x_1 = \frac{a_1 + b_1}{2}$ , e procede-se de maneira similar com os intervalos  $[a_1, x_1]$  e  $[x_1, b_1]$ . O erro relativo máximo nesta iteração está dado por  $E_1 = \left| \frac{x_1 - x_0}{x_1} \right|$ . E assim por diante.

Um passo deste processo pode ser resumido como segue: dado o intervalo  $I_k = [a_k, b_k]$  que contém uma raiz da equação  $f(x) = 0$ , bisseca-se tal intervalo no ponto  $x_k = \frac{a_k + b_k}{2}$  e escolhe-se aquela metade do intervalo que contém a raiz, mediante o seguinte teste

$$I_{k+1} = [a_{k+1}, b_{k+1}] = \begin{cases} [a_k, x_k], & \text{se } f(a_k) \cdot f(x_k) < 0, \\ [x_k, b_k], & \text{se } f(x_k) \cdot f(b_k) < 0, \end{cases}$$

se, eventualmente,  $f(x_k) = 0$ , então  $x_k$  será o valor da raiz procurada.

Para facilitar o processo aritmético do método, recomenda-se elaborar uma tabela com a seguinte estrutura:

$k$	$a_k$	$b_k$	$x_k$	$E_k$	sinal de $f(a_k)$	sinal de $f(x_k)$
0	$a_0$	$b_0$	$x_0$		+ ou -	+ ou -
1	$a_1$	$b_1$	$x_1$	$E_1$	+ ou -	+ ou -
2	$a_2$	$b_2$	$x_2$	$E_2$	+ ou -	+ ou -
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$r$	$a_r$	$b_r$	$x_r$	$E_r$		

onde  $x_k = \frac{a_k + b_k}{2}$  e  $E_k = \left| \frac{x_k - x_{k-1}}{x_k} \right|$ . Cada passo deste método requer uma nova avaliação da função  $f(x)$ , que é  $f(x_k)$ .

Este método constrói uma seqüência de intervalos  $I_0 \supset I_1 \supset I_2 \supset \dots \supset I_r$  tal que cada intervalo contém uma raiz da equação  $f(x) = 0$ . Depois de repetir o procedimento  $r$  vezes, é obtido um intervalo  $I_r$  de comprimento  $b_r - a_r = \frac{b_0 - a_0}{2^r}$ .

Para determinar quantas iterações do método,  $r$ , são necessárias, no mínimo, para conseguir um aproximante com um erro (absoluto) menor que  $\varepsilon$ , pode-se utilizar a relação  $\frac{b_0 - a_0}{2^r} < \varepsilon$  ou  $r > \frac{\ln(b_0 - a_0) - \ln(\varepsilon)}{\ln(2)}$ .

O número mínimo de iterações requerido para aproximar uma raiz no intervalo  $[0,1]$  para um  $\varepsilon > 0$  é dado na tabela a seguir:

$\varepsilon$	$r$
$10^{-2}$	7
$10^{-3}$	10
$10^{-4}$	14
$10^{-5}$	17
$10^{-6}$	20
$10^{-7}$	24

Logo, o método da bisseção requer um grande número de iterações para atingir um grau razoável de precisão para a raiz.

**Exemplo 2.17:** Considere a equação  $2x - \cos(x) = 0$ . Sendo  $f(x) = 2x - \cos(x)$  tem-se que  $f(0) = -1$  e  $f(1) = 1,459697694$ . Desde que a função  $f(x)$  é contínua, existe, pelo menos, uma raiz da equação  $2x - \cos(x) = 0$  no intervalo  $[0,1]$ . Será aplicado o método da bisseção até  $k = 6$  para aproximar tal raiz.

$k$	$a_k$	$b_k$	$x_k$	$E_k$	senal de $f(a_k)$	senal de $f(x_k)$
0	0	1	0,5		-	-
1	0	0,5	0,25	1	-	-
2	0,25	0,5	0,375	0,333333	-	-
3	0,375	0,5	0,4375	0,142857	-	-
4	0,4375	0,5	0,46875	0,066667	-	+
5	0,4375	0,46875	0,453125	0,034483	-	+
6	0,4375	0,453125	0,4453125	0,017544		

Aqui, têm sido calculados os valores de:

$$\begin{aligned}
 f(0,5) &= 2 \cdot 0,5 - \cos(0,5) = 0,122417, \\
 f(0,25) &= 2 \cdot 0,25 - \cos(0,25) = -0,468912, \\
 f(0,375) &= 2 \cdot 0,375 - \cos(0,375) = -0,180507, \\
 f(0,4375) &= 2 \cdot 0,4375 - \cos(0,4375) = -0,030814, \\
 f(0,46875) &= 2 \cdot 0,46875 - \cos(0,46875) = 0,045366, \\
 f(0,453125) &= 2 \cdot 0,453125 - \cos(0,453125) = 0,007167.
 \end{aligned}$$

Assim, o aproximante de uma raiz de  $2x - \cos(x) = 0$  no intervalo  $[0,1]$  está dado por  $x_6 = 0,4453125$  com um erro relativo não maior que  $0,017544$  ou  $1,75\%$ .

**Exemplo 2.18:** Do exemplo 2.15, sabe-se que a equação  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$  tem uma raiz positiva no intervalo  $[0,7]$ . Faz-se a seguinte tabela para localizar melhor esta raiz:

$x$	$p_n(x)$
0	-1
1	-6
2	-21
3	-22
4	39
5	234
6	659
7	1434

(para fins práticos, as avaliações funcionais terminariam após detectar a mudança de sinal, ou seja, quando  $x = 4$ ).

Observa-se que  $p_n(3) = -22$  e  $p_n(4) = 39$ . Como  $p_n(x)$  é contínua (na verdade, todo polinômio é), existe uma raiz da equação  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$  no intervalo  $[3,4]$ . Será aplicado o método da bissecção até  $k = 6$  para aproximar tal raiz.

$k$	$a_k$	$b_k$	$x_k$	$E_k$	sinal de $p_n(a_k)$	sinal de $f(x_k)$
0	3	4	3,5		-	-
1	3,5	4	3,75	0,066667	-	+
2	3,5	3,75	3,625	0,034483	-	+
3	3,5	3,625	3,5625	0,017544	-	+
4	3,5	3,5625	3,53125	0,008850	-	-
5	3,53125	3,5625	3,546875	0,004405	-	-
6	3,546875	3,5625	3,5546875	0,002198		

Aqui, têm sido calculados os valores de:

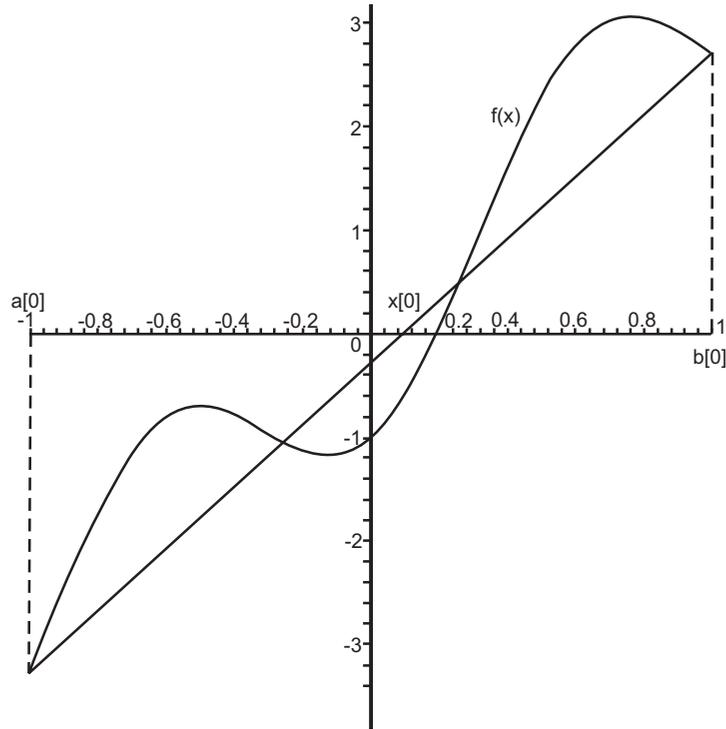
$$\begin{aligned}
 p_n(3,5) &= -3,1875, \\
 p_n(3,75) &= 14,410156, \\
 p_n(3,625) &= 4,812744, \\
 p_n(3,5625) &= 0,622086, \\
 p_n(3,53125) &= -1,329223, \\
 p_n(3,546875) &= -0,365336.
 \end{aligned}$$

Assim, o aproximante da raiz de  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$  no intervalo  $[3,4]$  está dado por  $x_6 = 3,5546875$  com um erro relativo não maior que  $0,002198$  ou  $0,21\%$ .

#### 2.4.2 Método da Falsa Posição

Considere a equação  $f(x) = 0$  com  $f(x)$  contínua. A idéia do método da falsa posição ou *regula falsi*, como também é conhecido, é, a semelhança do método da bissecção, dividir um intervalo  $(a_0, b_0)$ , com  $a_0 < b_0$ , tais que  $f(a_0)f(b_0) < 0$ . A idéia do método da falsa posição é calcular um aproximante de alguma raiz que se encontre no referido intervalo como sendo a média ponderada dos extremos do intervalo  $a_0$  e  $b_0$ , com pesos  $f(b_0)$  e  $-f(a_0)$ , respectivamente; assim,  $x_0 = \frac{a_0 f(b_0) - b_0 f(a_0)}{f(b_0) - f(a_0)}$ . Depois, pode-se determinar se existe mudança de sinal do valor da função  $f(x)$  no intervalo  $[a_0, x_0]$  ou  $[x_0, b_0]$ ; escolhe-se o intervalo  $[a_0, x_0]$  se  $f(a_0)f(x_0) < 0$  ou o intervalo  $[x_0, b_0]$  se  $f(a_0)f(x_0) > 0$ . Chamando de  $[a_1, b_1]$  o intervalo onde acontece a mudança de sinal, calcula-se um novo aproximante como  $x_1 = \frac{a_1 f(b_1) - b_1 f(a_1)}{f(b_1) - f(a_1)}$ , e procede-se de maneira similar à anterior com os intervalos  $[a_1, x_1]$  e  $[x_1, b_1]$ . O erro relativo máximo nesta iteração está dado por  $E_1 = \left| \frac{x_1 - x_0}{x_1} \right|$ . E o processo segue adiante.

A idéia do método da falsa posição tem uma característica geométrica, ilustrada na figura a seguir.



Considerando o gráfico da função  $f(x)$  e os pontos  $a_0$  e  $b_0$  tais que  $f(a_0)f(b_0) < 0$ , então o ponto  $x_0$  é a abscissa do intercepto da secante entre os pontos  $(a_0, f(a_0))$  e  $(b_0, f(b_0))$ .

Um passo deste processo pode ser resumido como segue: dado o intervalo  $I_k = [a_k, b_k]$  que contem uma raiz da equação  $f(x) = 0$ , divide-se tal intervalo no ponto  $x_k = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}$  e escolhe-se aquela parte do intervalo que contem a raiz, mediante o seguinte teste

$$I_{k+1} = [a_{k+1}, b_{k+1}] = \begin{cases} [a_k, x_k], & \text{se } f(a_k) \cdot f(x_k) < 0, \\ [x_k, b_k], & \text{se } f(a_k) \cdot f(x_k) > 0, \end{cases}$$

se, eventualmente,  $f(x_k) = 0$ , então  $x_k$  será o valor da raiz procurada.

Para facilitar o processo aritmético do método, recomenda-se elaborar uma tabela com a seguinte estrutura:

$k$	$a_k$	$b_k$	$x_k$	$E_k$	senal de $f(a_k)$	senal de $f(x_k)$
0	$a_0$	$b_0$	$x_0$		+ ou -	+ ou -
1	$a_1$	$b_1$	$x_1$	$E_1$	+ ou -	+ ou -
2	$a_2$	$b_2$	$x_2$	$E_2$	+ ou -	+ ou -
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$r$	$a_r$	$b_r$	$x_r$	$E_r$		

onde  $x_k = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}$  e  $E_k = \left| \frac{x_k - x_{k-1}}{x_k} \right|$ .

**Exemplo 2.19:** Considere a equação  $2x - \cos(x) = 0$ . Sendo  $f(x) = 2x - \cos(x)$  tem-se que  $f(0) = -1$  e  $f(1) = 1,459697694$ . Será aplicado o método da falsa posição até  $k = 3$  para aproximar tal raiz.

$k$	$a_k$	$b_k$	$x_k$	$E_k$	sinal de $p_n(a_k)$	sinal de $p_n(x_k)$
0	0	1	0,406554		-	-
1	0,406554	1	0,446512	0,089490	-	-
2	0,446512	1	0,449879	0,007484	-	-
3	0,449879	1	0,450158	0,000620		

Aqui, têm sido calculados os valores de:

$$\begin{aligned} f(0,406554) &= 2 \cdot 0,406554 - \cos(0,406554) = -0,105381, \\ f(0,446512) &= 2 \cdot 0,446512 - \cos(0,446512) = -0,008934, \\ f(0,449879) &= 2 \cdot 0,449879 - \cos(0,449879) = -0,000741. \end{aligned}$$

Assim, o aproximante de uma raiz de  $2x - \cos(x) = 0$  no intervalo  $[0,1]$  está dado por  $x_3 = 0,450158$  com um erro relativo não maior que  $0,000620$  ou  $0,062\%$ .

**Exemplo 2.20:** Do exemplo 2.18, sabe-se que a equação  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$  tem uma raiz positiva no intervalo  $[3,4]$ . Será aplicado o método da falsa posição até  $k = 3$  para aproximar tal raiz.

$k$	$a_k$	$b_k$	$x_k$	$E_k$	sinal de $p_n(a_k)$	sinal de $f(x_k)$
0	3	4	3,360656		-	-
1	3,360656	4	3,495241	0,038505	-	-
2	3,495241	4	3,536400	0,011639	-	-
3	3,536400	4	3,548149	0,003311		

Aqui, têm sido calculados os valores de:

$$\begin{aligned} p_n(3,360656) &= -10,398675, \\ p_n(3,495241) &= -3,462470, \\ p_n(3,536400) &= -1,014109. \end{aligned}$$

Assim, o aproximante da raiz de  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$  no intervalo  $[3,4]$  está dado por  $x_3 = 3,548149$  com um erro relativo não maior que  $0,003311$  ou  $0,33\%$ .

### 2.4.3 Método das Cordas

O método das cordas está fundamentado no fato seguinte: sendo  $x_{j-1}$  e  $x_j$  dois aproximantes de uma raiz da equação  $f(x) = 0$ , a função  $f(x)$  é substituída pela reta que une os pontos  $(x_{j-1}, f(x_{j-1}))$  e  $(x_j, f(x_j))$ , e desse modo o intercepto dessa reta com o eixo  $x$  é o próximo aproximante  $x_{j+1}$ , ou seja,  $x_{j+1} = \frac{x_{j-1}f(x_j) - x_jf(x_{j-1})}{f(x_j) - f(x_{j-1})}$  que tem semelhança com a fórmula utilizada no método da falsa posição.

Observe que o aproximante  $x_{j+1}$  também pode ser escrito como  $x_{j+1} = x_j - \frac{x_j - x_{j-1}}{f(x_j) - f(x_{j-1})} f(x_j)$ .

Os resultados podem ser organizados, por exemplo, imitando o formato a seguir:

$k$	$x_k$	$E_k$	$f(x_k)$	$x_k - x_{k-1}$	$f(x_k) - f(x_{k-1})$
0	$x_0$		$f(x_0)$		
1	$x_1$	$E_1$	$f(x_1)$	$x_1 - x_0$	$f(x_1) - f(x_0)$
2	$x_2$	$E_2$	$f(x_2)$	$x_2 - x_1$	$f(x_2) - f(x_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$r-1$	$x_{r-1}$	$E_{r-1}$	$f(x_{r-1})$	$x_{r-1} - x_{r-2}$	$f(x_{r-1}) - f(x_{r-2})$
$r$	$x_r$	$E_r$			

onde  $x_0$  e  $x_1$  são dados e é utilizada a fórmula  $x_k = x_{k-1} - \frac{x_{k-1} - x_{k-2}}{f(x_{k-1}) - f(x_{k-2})} f(x_{k-1})$  para  $k \geq 2$ .

**Exemplo 2.21:** Considere a equação  $2x - \cos(x) = 0$ . Sendo  $f(x) = 2x - \cos(x)$  tem-se que  $f(0) = -1$  e  $f(1) = 1,459697694$ . Será aplicado o método das cordas até  $k = 5$  para aproximar tal raiz.

$k$	$x_k$	$E_k$	$f(x_k)$	$x_k - x_{k-1}$	$f(x_k) - f(x_{k-1})$
0	0		-1		
1	1	1	1,459698	1	2,459698
2	0,406554	1,459698	-0,105381	-0,593446	-1,565079
3	0,446512	0,089490	-0,008934	0,039958	0,096447
4	0,450214	0,008221	0,000073	0,003701	0,009007
5	0,450184	0,000067			

Assim, o aproximante de uma raiz de  $2x - \cos(x) = 0$  no intervalo  $[0,1]$  está dado por  $x_5 = 0,450184$  com um erro relativo não maior que 0,000067 ou 0,0067%.

Na verdade, é desejável iniciar o método com valores onde a função muda de sinal; mas isto não é necessário, como mostram os resultados seguintes: supõe-se  $x_0 = 0$  e  $x_1 = 0,4$ . Observe que  $f(0) = -1$  e  $f(0,4) = -0,121061$ . Logo, aplicando o método até  $k = 4$

$k$	$x_k$	$E_k$	$f(x_k)$	$x_k - x_{k-1}$	$f(x_k) - f(x_{k-1})$
0	0		-1		
1	0,4	1	-0,120161	0,4	0,878939
2	0,455094	0,121061	0,011969	0,055094	0,1330030
3	0,450137	0,011012	-0,000113	-0,004957	-0,012081
4	0,450184	0,000103			

Assim, o aproximante de uma raiz de  $2x - \cos(x) = 0$  no intervalo  $[0,1]$  está dado por  $x_4 = 0,450184$  com um erro relativo não maior que 0,000103 ou 0,0103%.

**Exemplo 2.22:** Do exemplo 2.18, sabe-se que a equação  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$  tem uma raiz positiva no intervalo  $[3,4]$ . Será aplicado o método das cordas até  $k = 5$  para aproximar tal raiz.

$k$	$x_k$	$E_k$	$p_n(x_k)$	$x_k - x_{k-1}$	$p_n(x_k) - p_n(x_{k-1})$
0	3		-22		
1	4	0,25	39	1	61
2	3,360656	0,190244	-49,398675	-0,639344	-49,398675
3	3,495241	0,038505	-3,462470	0,134585	6,936205
4	3,562424	0,018859	0,617247	0,067183	4,079718
5	3,552260	0,002861			

Assim, o aproximante da raiz de  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$  no intervalo  $[3,4]$  está dado por  $x_5 = 3,552260$  com um erro relativo não maior que 0,002861 ou 0,29%

#### 2.4.4 Método de Newton-Raphson

O método de Newton-Raphson está baseado, assim como o método das cordas, na substituição da função  $f(x)$  por uma reta. Especificamente, se  $x_j$  é um aproximante de uma raiz da equação  $f(x) = 0$ , onde  $f(x)$  é derivável, então substitui-se  $f(x)$  pela reta que passa pelo ponto  $(x_j, f(x_j))$  e com coeficiente angular igual à derivada da função em  $x_j$ ,  $f'(x_j)$ . O aproximante  $x_{j+1}$  está dado pelo intercepto dessa reta com o eixo  $x$ , que é  $x_{j+1} = x_j - \frac{f(x_j)}{f'(x_j)}$ . Pode-se dizer que este método é o caso limite do método das cordas quando  $x_{j-1} \rightarrow x_j$ .

Os resultados podem ser organizados, por exemplo, imitando o formato a seguir:

$k$	$x_k$	$E_k$	$f(x_k)$	$f'(x_k)$
0	$x_0$		$f(x_0)$	$f'(x_0)$
1	$x_1$	$E_1$	$f(x_1)$	$f'(x_1)$
2	$x_2$	$E_2$	$f(x_2)$	$f'(x_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$r-1$	$x_{r-1}$	$E_{r-1}$	$f(x_{r-1})$	$f'(x_{r-1})$
$r$	$x_r$	$E_r$		

onde  $x_0$  é dado e é utilizada a fórmula  $x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})}$  para  $k \geq 1$ .

**Exemplo 2.23:** Considere a equação  $2x - \cos(x) = 0$ . Sendo  $f(x) = 2x - \cos(x)$  tem-se que  $f'(x) = 2 + \sin(x)$ . Será aplicado o método de Newton até  $k = 3$  com  $x_0 = 1$  para aproximar tal raiz.

$k$	$x_k$	$E_k$	$f(x_k)$	$f'(x_k)$
0	1		1,459698	2,841471
1	0,486288	1,056394	0,088502	2,467347
2	0,450419	0,079636	0,000572	2,435342
3	0,450184	0,000522		

Assim, o aproximante de uma raiz de  $2x - \cos(x) = 0$  no intervalo  $[0,1]$  está dado por  $x_3 = 0,450184$  com um erro relativo não maior que 0,000522 ou 0,0522%.

Pode-se observar que a rapidez de convergência deste método é superior aos anteriores.

**Exemplo 2.24:** Do exemplo 2.18, sabe-se que a equação  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$  tem uma raiz positiva no intervalo  $[3,4]$ . Será aplicado o método de Newton até  $k = 5$  com  $x_0 = 3$  para aproximar tal raiz.

$k$	$x_k$	$E_k$	$f(x_k)$	$f'(x_k)$
0	3		-22	20
1	4,1	0,268293	51,0741	127,624
2	3,699808	0,108166	63,016938	63,016938
3	3,567137	0,037192	0,919716	63,016938
4	3,552858	0,004019	0,009945	63,016938
5	3,552700	0,000044		

Assim, o aproximante da raiz de  $p_n(x) = x^4 - 2x^3 - 6x^2 + 2x - 1 = 0$  no intervalo  $[3,4]$  está dado por  $x_5 = 3,552700$  com um erro relativo não maior que  $0,000044$  ou  $0,0044\%$ .

**Exemplo 2.25:** O método de Newton pode ser utilizado para gerar uma equação recursiva que converge à raiz quadrada de um número  $N > 0$ . Basta observar que uma equação em que uma das soluções é  $\sqrt{N}$  pode ser escrita como  $f(x) = x^2 - N = 0$ . Dado  $x_0$ , a equação que gera os aproximantes é

$$x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})} = x_{k-1} - \frac{x_{k-1}^2 - N}{2x_{k-1}} = \frac{1}{2} \left( x_{k-1} + \frac{N}{x_{k-1}} \right), \text{ ou seja, } x_k = \frac{1}{2} \left( x_{k-1} + \frac{N}{x_{k-1}} \right), \text{ para}$$

$k \geq 1$ . Assumindo  $N = 3$  e começando com  $x_0 = 1$ , tem-se que  $x_1 = \frac{1}{2} \left( 1 + \frac{3}{1} \right) = 2$ ,  $x_2 = \frac{1}{2} \left( 2 + \frac{3}{2} \right) = 1,75$ ,

$$x_3 = \frac{1}{2} \left( 1,75 + \frac{3}{1,75} \right) = 1,732142857, \quad x_4 = \frac{1}{2} \left( 1,732142857 + \frac{3}{1,732142857} \right) = 1,73205081 \text{ e}$$

$$x_5 = \frac{1}{2} \left( 1,73205081 + \frac{3}{1,73205081} \right) = 1,732050808. \text{ Este último valor tem um erro relativo não}$$

maior que  $0,1155 \times 10^{-8}$ , ou seja, pelo menos nove dígitos significativos exatos.

#### ATIVIDADE 4

1) Mediante uma tabela de valores da função  $f(x)$  localize uma possível raiz real da equação  $f(x) = 0$ . A seguir, pelo método da bisseção, calcule um aproximante desta raiz, com um erro relativo menor que  $0,1$ , sendo:

a)  $f(x) = x^5 + x^2 - 1 = 0$ ,

b)  $f(x) = 3x - \cos(x) = 0$ ,

c)  $f(x) = x + \ln(x) = 0$ .

- 2) Utilizando as mesmas funções  $f(x)$  do exercício 1, calcule o aproximante de uma raiz real da equação  $f(x) = 0$  mediante o método da falsa posição, com um erro relativo não maior que 0,01.
- 3) Utilizando as mesmas funções  $f(x)$  do exercício 1, calcule o aproximante de uma raiz real da equação  $f(x) = 0$  mediante o método das cordas, com um erro relativo não maior que 0,01.
- 4) Utilizando as mesmas funções  $f(x)$  do exercício 1, calcule o aproximante de uma raiz real da equação  $f(x) = 0$  mediante o método de Newton-Raphson, com um erro relativo não maior que 0,01.

# ANOTAÇÕES

---

Lined area for notes, consisting of multiple horizontal blue lines.

Sistema de Equações  
Lineares Algébricas

3  
Capítulo

### 3.1 Introdução

Muitos problemas de Matemática Numérica são modelados em termos de um Sistema de Equações Lineares Algébricas. Isso vale em geral para o tratamento numérico de equações funcionais lineares que ocorrem, entre outras, como equações diferenciais parciais ou ordinárias e equações integrais que surgem em diversos problemas da Física e Engenharia.

Considere um sistema de  $m$  equações lineares algébricas com  $n$  incógnitas

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{22}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\dots \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned} \quad (1)$$

onde  $a_{ij}$  são valores escalares conhecidos,  $b_i, i = 1, \dots, m$  são valores conhecidos e  $x_i, i = 1, \dots, n$  são as incógnitas a serem determinadas.

Introduz-se a seguinte notação e definições:

<b>A</b>	matriz de ordem $m \times n$
<b><math>a_{ij}</math></b>	elemento na $i$ -ésima linha e $j$ -ésima coluna da matriz <b>A</b>
<b><math>A^{-1}</math></b>	inversa de <b>A</b>
<b><math>\det(\mathbf{A}) =  \mathbf{A} </math></b>	determinante de <b>A</b>
<b><math>M_{ij}</math></b>	menor de $a_{ij}$ em <b>A</b>
<b><math>C_{ij}</math></b>	cofator de $a_{ij}$ em <b>A</b>
<b>0</b>	matriz nula
<b>I</b>	matriz identidade
<b>D</b>	matriz diagonal de ordem $n$
<b>L</b>	matriz triangular inferior de ordem $n$
<b>U</b>	matriz triangular superior de ordem $n$
<b>P</b>	matriz de permutação
<b>x</b>	vetor coluna com elementos $x_i, i = 1, \dots, n$
<b>b</b>	vetor coluna com elementos $b_i, i = 1, \dots, n$

O sistema de equações (1) pode ser escrito de maneira matricial como

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \quad (2)$$

onde,

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{e} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Os métodos de solução de um sistema de equações algébricas lineares (1) podem ser classificados de maneira ampla em dois tipos:

(I) Métodos Diretos: Estes métodos produzem a solução exata após um número finito de passos (sem levar em conta os erros de arredondamento).

(II) Métodos Iterativos: Estes métodos originam uma seqüência de soluções aproximadas, que converge quando o número de passos tende a infinito.

Os métodos diretos e iterativos estudados aqui serão aplicados para sistemas quadrados, isto é, no caso que  $m = n$ , ou seja, o número de equações coincide com o número de incógnitas.

### 3.2 Métodos Diretos

O sistema de equações (1) pode ser resolvido diretamente pela regra de Cramer.

#### 3.2.1 Regra de Cramer:

Tem-se que  $x_i = \frac{|B_i|}{|A|}$ ,  $i = 1, 2, \dots, n$ , (3)

onde  $|B_i|$  é o determinante da matriz obtida por substituir a  $i$ -ésima coluna de  $A$  pelo vetor  $b$ .

O vetor  $x$  está dado por

$$x = \frac{1}{|A|} \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \vdots & C_{nm} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{|A|} \cdot adj(A) \cdot b = A^{-1} \cdot b \quad (4)$$

onde  $C_{ij}$  é o cofator correspondente ao elemento  $a_{ij}$  da matriz  $A$ . Esta expressão é equivalente à equação (3). A regra de Cramer fornece um método de obter a solução imediatamente, mas estas soluções envolvem  $n + 1$  determinantes de ordem  $n$ . Esta característica torna a regra de Cramer impraticável na resolução de sistemas.

Exemplo 1: Resolva o sistema abaixo utilizando a regra de Cramer.

$$\begin{cases} x_1 + 2x_2 - x_3 = 2 \\ 3x_1 + 6x_2 + x_3 = 1 \\ 3x_1 + 3x_2 + 2x_3 = 3 \end{cases}$$

A solução das equações podem ser escritas como  $x_i = \frac{|B_i|}{|A|}$ ,  $i = 1, 2, 3, \dots$  onde

$$|B_1| = \begin{vmatrix} 2 & 2 & -1 \\ 1 & 6 & 1 \\ 3 & 3 & 2 \end{vmatrix}, |B_2| = \begin{vmatrix} 1 & 2 & -1 \\ 3 & 1 & 1 \\ 3 & 3 & 2 \end{vmatrix}, |B_3| = \begin{vmatrix} 1 & 2 & 2 \\ 3 & 6 & 1 \\ 3 & 3 & 3 \end{vmatrix}, |A| = \begin{vmatrix} 1 & 2 & -1 \\ 3 & 6 & 1 \\ 3 & 3 & 2 \end{vmatrix}$$

ou

$$|B_1| = 35, |B_2| = -13, |B_3| = -15, |A| = 12$$

que fornece a solução:

$$x_1 = \frac{35}{12}, x_2 = \frac{-13}{12} \text{ e } x_3 = \frac{-15}{12}.$$

Exemplo 2: Resolva o sistema abaixo utilizando a equação 4:

$$\begin{cases} x_1 + 2x_2 + 2x_3 = 6 \\ x_1 + 2x_2 + x_3 = 1 \\ 2x_1 + x_2 + x_3 = 4 \end{cases}$$

Tem-se que

$$|A| = \begin{vmatrix} 1 & 2 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{vmatrix} = -3 \text{ e } \text{adj}(A) = \begin{pmatrix} 1 & 0 & -2 \\ 1 & -3 & 1 \\ -3 & 3 & 0 \end{pmatrix}$$

Assim, utilizando a equação (4), obtem-se

$$x = \frac{1}{-3} \begin{pmatrix} 1 & 0 & -2 \\ 1 & -3 & 1 \\ -3 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2/3 \\ -7/3 \\ 5 \end{pmatrix}.$$

### 3.2.2 Eliminação Gaussiana

Outro método direto é o método de eliminação gaussiana. A idéia principal consiste em transformar o sistema  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  em um sistema equivalente cuja matriz associada é do tipo triangular superior. Isto pode ser feito mediante algumas operações elementares de matrizes: como troca de linhas e soma de uma linha com um múltiplo escalar de outra linha. Se acontecer troca de linhas, o processo de eliminação denomina-se com pivoteamento parcial e se não acontecer, o processo denomina-se sem pivoteamento.

A idéia do método é transformar a matriz aumentada

$$(A : b) = \begin{pmatrix} a_{11} & a_{12} \dots & a_{1n} & \vdots & b_1 \\ a_{21} & a_{22} \dots & a_{2n} & \vdots & b_2 \\ \vdots & \vdots \ddots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} \dots & a_{nn} & \vdots & b_n \end{pmatrix}$$

na matriz

$$(A^{(n)} : b^{(n)}) = \begin{pmatrix} a_{11}^{(n)} & a_{12}^{(n)} & \dots & a_{1n}^{(n)} & \vdots & b_1^{(n)} \\ 0 & a_{22}^{(n)} & & a_{2n}^{(n)} & \vdots & b_2^{(n)} \\ & & \ddots & & & \\ 0 & 0 & \dots & a_{nn}^{(n)} & \vdots & b_n^{(n)} \end{pmatrix}$$

mediante a troca de linhas e soma de uma linha com um múltiplo escalar de outra linha. A última matriz representa um sistema triangular superior de equações lineares que pode ser resolvido facilmente mediante a técnica denominada retro-substituição.

Exemplo 3: Resolver o sistema de equações lineares

$$\begin{cases} 10x_1 + x_2 - 5x_3 = 1 \\ -20x_1 + 3x_2 + 20x_3 = 2 \\ 5x_1 + 3x_2 + 5x_3 = 6 \end{cases}$$

mediante a eliminação gaussiana sem pivotamento.

A matriz aumentada do sistema é transformada na forma triangular superior:

$$\begin{aligned}
 & l_2 = l_2 - \left(\frac{-20}{10}\right)l_1 \\
 & l_3 = l_3 - \left(\frac{5}{10}\right)l_1 \\
 & \left( \begin{array}{cccc} 10 & 1 & -5 & 1 \\ -20 & 3 & 20 & 2 \\ 5 & 3 & 5 & 6 \end{array} \right) \xrightarrow{l_2 = l_2 - \left(\frac{-20}{10}\right)l_1, l_3 = l_3 - \left(\frac{5}{10}\right)l_1} \left( \begin{array}{cccc} 10 & 1 & -5 & 1 \\ 0 & 5 & 10 & 4 \\ 0 & 2,5 & 7,5 & 5,5 \end{array} \right) \\
 & \xrightarrow{l_3 = l_3 - \left(\frac{5/2}{5}\right)l_2} \left( \begin{array}{cccc} 10 & 1 & -5 & 1 \\ 0 & 5 & 10 & 4 \\ 0 & 0 & 2,5 & 3,5 \end{array} \right)
 \end{aligned}$$

e isto significa que o sistema original é equivalente ao sistema triangular superior

$$\begin{cases} 10x_1 + x_2 - 5x_3 = 1 \\ 5x_2 + 10x_3 = 4 \\ (2,5)x_3 = 3,5 \end{cases}$$

que, utilizando a retro-substituição, fornece a solução

$$\begin{aligned}
 x_3 &= (7/2) / (5/2) = 7/5 \\
 x_2 &= (-10x_3 + 4) / 5 = -2 \\
 x_1 &= (-x_2 + 5x_3 + 1) / 10 = 1
 \end{aligned}$$

Os elementos **10** na primeira matriz e **5** na segunda matriz são denominados os pivôs da matriz correspondente.

Quando é utilizado o procedimento de pivoteamento parcial, é recomendável, por questões de minimizar o erro cometido, fazer uma troca de equações de maneira que o pivô seja aquele com maior valor absoluto entre o elemento diagonal  $a_{ii}$  e os elementos correspondentes da mesma coluna  $a_{ij}$  com  $j > i$ .

Exemplo 4: Resolver o sistema de equações lineares

$$\begin{cases} 10x_1 + x_2 - 5x_3 = 1 \\ -20x_1 + 3x_2 + 20x_3 = 2 \\ 5x_1 + 3x_2 + 5x_3 = 6 \end{cases}$$

mediante a eliminação gaussiana com pivoteamento parcial.

A matriz aumentada do sistema é transformada na forma triangular superior:

$$\begin{aligned}
 & l_2 = l_2 - \left(\frac{10}{-20}\right)l_1 \\
 & l_3 = l_3 - \left(\frac{5}{-20}\right)l_1 \\
 & \left( \begin{array}{cccc} 10 & 1 & -5 & 1 \\ -20 & 3 & 20 & 2 \\ 5 & 3 & 5 & 6 \end{array} \right) \xrightarrow{l_1 \leftrightarrow l_2} \left( \begin{array}{cccc} -20 & 3 & 20 & 2 \\ 10 & 1 & -5 & 1 \\ 5 & 3 & 5 & 6 \end{array} \right) \xrightarrow{l_3 = l_3 - \left(\frac{5}{-20}\right)l_1} \left( \begin{array}{cccc} -20 & 3 & 20 & 2 \\ 0 & 2,5 & 5 & 2 \\ 0 & 3,75 & 10 & 6,5 \end{array} \right) \\
 & \xrightarrow{l_2 \leftrightarrow l_3} \left( \begin{array}{cccc} -20 & 3 & 20 & 2 \\ 0 & 3,75 & 10 & 6,5 \\ 0 & 2,5 & 5 & 2 \end{array} \right) \xrightarrow{l_3 = l_3 - \left(\frac{5/2}{15/4}\right)l_2} \left( \begin{array}{cccc} -20 & 3 & 20 & 2 \\ 0 & 3,75 & 10 & 13/2 \\ 0 & 0 & -5/3 & -7/3 \end{array} \right)
 \end{aligned}$$

e isto significa que o sistema original é equivalente ao sistema triangular superior

$$\begin{cases} -20x_1 + 3x_2 + 20x_3 = 1 \\ 15/4 x_2 + 10x_3 = 13/2 \\ -5/3 x_3 = -7/3 \end{cases}$$

que, utilizando a retro-substituição, fornece a solução

$$x_3 = (-7/3) / (-5/3) = 7/5$$

$$x_2 = (-10x_3 + 13/2) / (15/4) = -2$$

$$x_1 = (-3x_2 - 20x_3 + 1) / (-20) = 1$$

### 3.2.3 Eliminação de Decomposição LU

A idéia deste método consiste em conseguir uma fatoração da matriz  $A$  como  $A = LU$ , onde  $L$  é uma matriz triangular inferior e  $U$ , uma matriz triangular superior. Então, neste caso, o sistema  $A \cdot x = b$  é um equivalente a dois sistemas triangulares

$$\begin{cases} Ly = b \\ Ux = y \end{cases}$$

O primeiro sistema pode ser resolvido pela técnica denominada substituição para frente e o segundo sistema, como antes, pode ser resolvido por retro-substituição.

O método de decomposição LU também pode utilizar o pivoteamento parcial. Neste caso, a montagem das matrizes  $L$  e  $U$  é feita utilizando as matrizes de permutação, mas isto não será visto aqui.

A técnica de fatoração da matriz  $A$  é similar ao processo de eliminação. Isto pode ser ilustrado no exemplo a seguir.

Exemplo 5: Resolver o sistema de equações lineares

$$\begin{cases} 10x_1 + x_2 - 5x_3 = 1 \\ -20x_1 + 3x_2 + 20x_3 = 2 \\ 5x_1 + 3x_2 + 5x_3 = 6 \end{cases}$$

mediante a decomposição LU sem pivoteamento.

A matriz do sistema é transformada na forma triangular superior:

$$\begin{aligned} \begin{pmatrix} 10 & 1 & -5 \\ -20 & 3 & 20 \\ 5 & 3 & 5 \end{pmatrix} &\xrightarrow{l_2=l_2-\frac{-20}{10}l_1} \begin{pmatrix} 10 & 1 & -5 \\ 0 & 5 & 10 \\ 5 & 3 & 5 \end{pmatrix} \xrightarrow{l_3=l_3-\frac{5}{10}l_1} \begin{pmatrix} 10 & 1 & -5 \\ 0 & 5 & 10 \\ 0 & 2,5 & 7,5 \end{pmatrix} \\ &\xrightarrow{l_3=l_3-\frac{5/2}{5}l_2} \begin{pmatrix} 10 & 1 & -5 \\ 0 & 5 & 10 \\ 0 & 0 & 2,5 \end{pmatrix} \end{aligned}$$

Então, a matriz  $U$  é a última matriz obtida

$$U = \begin{pmatrix} 10 & 1 & -5 \\ 0 & 5 & 10 \\ 0 & 0 & 5/2 \end{pmatrix}$$

e a matriz  $L$  está formada por elementos diagonais unitários e os elementos restantes pelos MULTIPLICADORES,

$$l_{21} = \frac{-20}{10} = 2, l_{31} = \frac{5}{10} = \frac{1}{2} \text{ e } l_{32} = \frac{5/2}{5} = \frac{1}{2}, \text{ ou seja,}$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1/2 & 1/2 & 1 \end{pmatrix}$$

O sistema  $Ly = b$  pode ser escrito como

$$\begin{aligned} y_1 &= 1 \\ -2y_1 + y_2 &= 2 \\ 1/2 y_1 + 1/2 y_2 + y_3 &= 6 \end{aligned}$$

cuja solução é

$$\begin{aligned} y_1 &= 1 \\ y_2 &= 2y_1 + 2 = 4 \\ y_3 &= -1/2 y_1 - 1/2 y_2 + 6 = 7/2 \end{aligned}$$

O sistema  $Ux = y$  pode ser escrito como

$$\begin{aligned} 10x_1 + x_2 - 5x_3 &= 1 \\ 5x_2 + 10x_3 &= 4 \\ 5/2 x_3 &= 7/2 \end{aligned}$$

que fornece a solução

$$\begin{aligned} x_3 &= \frac{(7/2)}{(5/2)} = 7/5 \\ x_2 &= \frac{(-10x_3 + 4)}{5} = -2 \\ x_1 &= \frac{(-x_2 + 5x_3 + 1)}{(-10)} = 1 \end{aligned}$$

### 3.3 Métodos Iterativos

O sistema

$$Ax = b$$

pode ser resolvido numericamente mediante a construção de um método iterativo. Suponha que a matriz  $A$  pode ser escrita como a diferença de duas outras matrizes  $A = S - T$ . Neste caso, o sistema original pode ser escrito como

$$Sx = Tx + b$$

que origina o método iterativo  $S\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)} + \mathbf{b}$  (equação 5), onde em cada iteração é gerado o vetor  $\mathbf{x}^{(k+1)}$  a partir do vetor  $\mathbf{x}^{(k)}$  mediante a equação recursiva (5) começando com um vetor inicial  $\mathbf{x}^{(0)}$ .

A equação recursiva (5), no caso que  $S$  seja uma matriz não singular, pode ser escrita como  $\mathbf{x}^{(k+1)} = S^{-1}T\mathbf{x}^{(k)} + S^{-1}\mathbf{b}$  (equação 6).

### 3.3.1 Método de Jacobi

Quando a matriz  $S$  é a parte diagonal de  $A$ , o método recursivo (equação 5) é denominado método de Jacobi.

Exemplo 6: Resolver o sistema

$$\begin{cases} 3x_1 + 2x_2 = 5 \\ 2x_2 + x_3 = 3 \\ x_1 + 2x_3 = 3 \end{cases}$$

mediante o método de Jacobi.

Neste caso,

$$A = \begin{pmatrix} 3 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}, S = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, T = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}$$

Logo,

$$S^{-1} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{pmatrix}, S^{-1}T = \begin{pmatrix} 0 & -2/3 & 0 \\ 0 & 0 & -1/2 \\ -1/2 & 0 & 0 \end{pmatrix}, S^{-1}\mathbf{b} = \begin{pmatrix} 5/3 \\ 3/2 \\ 3/2 \end{pmatrix}$$

e portanto a equação iterativa (6) fica

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & -2/3 & 0 \\ 0 & 0 & -1/2 \\ -1/2 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1^k \\ x_2^k \\ x_3^k \end{pmatrix} + \begin{pmatrix} 5/3 \\ 3/2 \\ 3/2 \end{pmatrix}$$

ou,

$$x_1^{(k+1)} = \frac{-2x_2^{(k)} + 5}{3}$$

$$x_2^{(k+1)} = \frac{-x_3^{(k)} + 3}{2}$$

$$x_3^{(k+1)} = \frac{-x_1^{(k)} + 3}{2}$$

Estas últimas três equações recursivas podem ser obtidas diretamente do sistema original, botando em evidência  $x_1$  da primeira equação,  $x_2$  da segunda equação e  $x_3$  da terceira equação. Este procedimento é muito prático e evita o tratamento matricial da equação (6).

Os critérios de parada mais comumente utilizados são a determinação do maior erro absoluto  $\|x^{(k+1)} - x^{(k)}\|$ , a determinação do maior erro relativo  $\frac{\|x^{(k+1)} - x^{(k)}\|}{\|x^{(k+1)}\|}$  e o número pré-determinado de iterações.

Neste caso,  $\|a\|$  é qualquer norma vetorial.

A seguinte tabela mostra os resultados das primeiras cinco iterações para  $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$j$	$x_1^{(j)}$	$x_2^{(j)}$	$x_3^{(j)}$
0	0	0	0
1	$5/3$	$3/2$	$3/2$
2	$2/3$	$3/4$	$2/3$
3	$7/6$	$7/6$	$7/6$
4	$8/9$	$11/12$	$11/12$
5	$19/18$	$25/24$	$19/18$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

e pode-se observar que a solução vai convergindo para o vetor  $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$  que é a solução exata do sistema.

**Exemplo 7:** Resolver o sistema

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 6 \\ 2x_2 + 3x_3 = 5 \\ 2x_1 + 2x_3 = 4 \end{cases}$$

mediante o método de Jacobi.

Neste caso,

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 3 \\ 2 & 0 & 2 \end{pmatrix}, S = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, T = \begin{pmatrix} 0 & -2 & -1 \\ 0 & 0 & -3 \\ -2 & 0 & 0 \end{pmatrix}$$

Logo,

$$S^{-1} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{pmatrix}, S^{-1}T = \begin{pmatrix} 0 & -2/3 & -1/3 \\ 0 & 0 & -3/2 \\ -1 & 0 & 0 \end{pmatrix}, S^{-1}b = \begin{pmatrix} 2 \\ 5/2 \\ 2 \end{pmatrix}$$

e portanto a equação iterativa (6) fica

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & -2/3 & -1/3 \\ 0 & 0 & -3/2 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1^k \\ x_2^k \\ x_3^k \end{pmatrix} + \begin{pmatrix} 2 \\ 5/2 \\ 2 \end{pmatrix}$$

ou,

$$x_1^{(k+1)} = \frac{-2x_2^{(k)} + x_3^{(k)} + 6}{3}$$

$$x_2^{(k+1)} = \frac{-3x_3^{(k)} + 5}{2}$$

$$x_3^{(k+1)} = \frac{-2x_1^{(k)} + 4}{2}$$

A seguinte tabela mostra os resultados das primeiras cinco iterações para  $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$j$	$x_1^{(j)}$	$x_2^{(j)}$	$x_3^{(j)}$
0	0	0	0
1	2	$5/2$	2
2	$-1/3$	$-1/2$	0
3	$7/3$	$5/2$	$7/3$
4	$-4/9$	-1	$-1/3$
5	$25/9$	3	$22/9$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

e pode-se observar que a solução vai convergindo para o vetor  $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$  que é a solução exata do sistema.

### 3.3.2 Método de Gauss-Seidel

Quando a matriz  $S$  é a parte triangular inferior de  $A$ , o método recursivo (equação 5) é denominado método de Gauss-Seidel.

**Exemplo 8:** Resolver o sistema

$$\begin{cases} 3x_1 + 2x_2 = 5 \\ 2x_2 = 3 \\ x_1 + 2x_3 = 3 \end{cases}$$

mediante o método de Gauss-Seidel.

Neste caso,

$$A = \begin{pmatrix} 3 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}, S = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}, T = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Logo,

$$S^{-1} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ -1/6 & 0 & 1/2 \end{pmatrix}, S^{-1}T = \begin{pmatrix} 0 & -2/3 & 0 \\ 0 & 0 & -1/2 \\ 0 & 1/3 & 0 \end{pmatrix}, S^{-1}b = \begin{pmatrix} 5/3 \\ 3/2 \\ 2/3 \end{pmatrix}$$

e portanto a equação iterativa (6) fica

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & -2/3 & 0 \\ 0 & 0 & -1/2 \\ 0 & 1/3 & 0 \end{pmatrix} \begin{pmatrix} x_1^k \\ x_2^k \\ x_3^k \end{pmatrix} + \begin{pmatrix} 5/3 \\ 3/2 \\ 2/3 \end{pmatrix}$$

ou,

$$x_1^{(k+1)} = \frac{-2x_2^{(k)} + 5}{3}$$

$$x_2^{(k+1)} = \frac{-x_3^{(k)} + 3}{2}$$

$$x_3^{(k+1)} = \frac{-x_1^{(k)} + 3}{2}$$

Pode-se observar que o método de Gauss-Seidel utiliza os valores de  $x_i^{(k+1)}$  que foram calculados na mesma iteração. Isto aumenta a rapidez de convergência do método.

A seguinte tabela mostra os resultados das primeiras cinco iterações para  $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$j$	$x_1^{(j)}$	$x_2^{(j)}$	$x_3^{(j)}$
0	0	0	0
1	5/3	3/2	2/3
2	2/3	7/6	7/6
3	8/9	11/12	19/18
4	19/18	35/36	35/36
5	55/54	73/72	107/108
⋮	⋮	⋮	⋮

e pode-se observar que a solução vai convergindo para o vetor  $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$  que é a solução exata do sistema.

**Exemplo 9:** Resolver o sistema

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 6 \\ 2x_2 + 3x_3 = 5 \\ 2x_1 + 2x_3 = 4 \end{cases}$$

mediante o método de Jacobi.

Neste caso,

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 3 \\ 2 & 0 & 2 \end{pmatrix}, \quad S = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & -2 & -1 \\ 0 & 0 & -3 \\ -2 & 0 & 0 \end{pmatrix}$$

Logo,

$$S^{-1} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ -1/3 & 0 & 1/2 \end{pmatrix}, \quad S^{-1}T = \begin{pmatrix} 0 & -2/3 & -1/3 \\ 0 & 0 & -3/2 \\ 0 & 2/3 & 1/3 \end{pmatrix}, \quad S^{-1}b = \begin{pmatrix} 2 \\ 5/2 \\ 0 \end{pmatrix}$$

e portanto a equação iterativa (6) fica

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & -2/3 & -1/3 \\ 0 & 0 & -3/2 \\ 0 & 2/3 & 1/3 \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{pmatrix} + \begin{pmatrix} 2 \\ 5/2 \\ 0 \end{pmatrix}$$

ou,

$$x_1^{(k+1)} = \frac{-2x_2^{(k)} + x_3^{(k)} + 6}{3}$$

$$x_2^{(k+1)} = \frac{-3x_3^{(k)} + 5}{2}$$

$$x_3^{(k+1)} = \frac{-2x_1^{(k)} + 4}{2}$$

A seguinte tabela mostra os resultados das primeiras cinco iterações para  $x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$j$	$x_1^{(j)}$	$x_2^{(j)}$	$x_3^{(j)}$
0	0	0	0
1	2	$5/2$	0
2	$1/3$	$5/2$	$5/3$
3	$-2/9$	0	$20/9$
4	$34/27$	$-5/6$	$20/27$
5	$187/81$	$25/18$	$-25/81$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

e pode-se observar que os aproximantes calculados não convergem para a solução exata  $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$  do sistema.

### 3.3.3 Convergência dos Métodos Iterativos

Uma condição necessária e suficiente do método iterativo (equação 6)  $\mathbf{x}^{(k+1)} = \mathbf{S}^{-1}\mathbf{T}\mathbf{x}^{(k)} + \mathbf{S}^{-1}\mathbf{b}$  (equação 6) é que os autovalores da matriz  $\mathbf{S}^{-1}\mathbf{T}$  sejam, em valor absoluto, menores que 1.

Uma condição suficiente para a convergência do método é que a matriz A seja diagonalmente dominante, isto é, que para cada  $i$ ,

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|$$

# ANOTAÇÕES

---

Lined area for taking notes, consisting of multiple horizontal blue lines.

Interpolação

4  
Capítulo

A idéia da interpolação é o cálculo de um valor entre outros que são conhecidos ou tabelados utilizando os valores vizinhos. Mas este cálculo faz parte de um processo mais amplo que consiste em construir uma função que ajusta de maneira aproximada certos valores (escalares ou vetoriais) dados. Este processo mais amplo denomina-se ajuste ou regressão. Neste sentido, a interpolação é o caso particular do ajuste em que a função construída deve passar exatamente pelos valores dados.

Outra idéia que também está relacionada com a interpolação é a aproximação de uma função complicada por funções mais simples, geralmente polinômios ou funções periódicas (séries de Fourier).

## 4.1 Introdução

### Definição 4.1: (Polinômio Interpolador)

Seja uma função,  $f(x)$ , e um conjunto de pontos,  $(x_1, f(x_1))$ ,  $(x_2, f(x_2))$ ,  $\dots$ ,  $(x_n, f(x_n))$ . Um polinômio  $p(x)$  é denominado um **polinômio interpolador** desse conjunto de pontos se os valores de  $p(x)$  e/ou suas derivadas até de certa ordem coincidem com os de  $f(x)$  e/ou as derivadas respectivas em  $x = x_1$ ,  $x = x_2$ ,  $\dots$ ,  $x = x_n$ . Assim,  $p(x)$  é um polinômio interpolador de  $f(x)$  nos pontos  $(x_1, f(x_1))$ ,  $(x_2, f(x_2))$ ,  $\dots$ ,  $(x_n, f(x_n))$  se  $p(x_i) = f(x_i)$ ,  $\frac{d}{dx} p(x_i) = \frac{d}{dx} f(x_i)$ ,  $\frac{d^k}{dx^k} p(x_i) = \frac{d^k}{dx^k} f(x_i)$ ,  $\dots$ ,  $\frac{d^k}{dx^k} p(x_i) = \frac{d^k}{dx^k} f(x_i)$ , para certo  $k$ , para todo  $i = 1, 2, \dots, n$ .

Um polinômio que satisfaz as condições  $p(x_i) = f(x_i)$  denomina-se um **polinômio interpolador de Lagrange**.

Um polinômio que satisfaz as condições  $p(x_i) = f(x_i)$  e  $\frac{d}{dx} p(x_i) = \frac{d}{dx} f(x_i)$  denomina-se um **polinômio interpolador de Hermite**.

## 4.2 Interpolação de Van der Monde

Suponha que são dados os pontos  $(x_1, f(x_1))$ ,  $(x_2, f(x_2))$ ,  $\dots$ ,  $(x_n, f(x_n))$  que podem ser organizados de forma tabular como

$x$	$f(x)$
$x_1$	$y_1$
$x_2$	$y_2$
$\vdots$	$\vdots$
$x_n$	$y_n$

sendo  $y_1 = f(x_1)$ ,  $y_2 = f(x_2)$ ,  $\dots$ ,  $y_n = f(x_n)$ .

De maneira direta, quer-se construir um polinômio de grau  $n-1$ ,

$$p(x) = a_{n-1}x^{n-1} + \dots + a_1x + a_0,$$

que satisfaz as condições  $p(x_i) = y_i$ , para todo  $i = 1, 2, \dots, n$ , ou seja, o gráfico de  $y = p(x)$  deve passar

pelos pontos  $(x_i, y_i)$ , para todo  $i = 1, 2, \dots, n$ . Observe que os coeficientes do polinômio  $p(x)$  não são conhecidos, mas podem ser encontrados utilizando as condições  $p(x_i) = y_i$ :

$$\begin{aligned} p(x_1) &= a_{n-1}x_1^{n-1} + \dots + a_1x_1 + a_0 = y_1, \\ p(x_2) &= a_{n-1}x_2^{n-1} + \dots + a_1x_2 + a_0 = y_2, \\ &\vdots \\ p(x_n) &= a_{n-1}x_n^{n-1} + \dots + a_1x_n + a_0 = y_n; \end{aligned}$$

que produz um sistema linear de equações representado de forma matricial como:

$$\begin{bmatrix} x_1^{n-1} & \dots & x_1 & 1 \\ x_2^{n-1} & \dots & x_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_n^{n-1} & \dots & x_n & 1 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

sendo a matriz do sistema uma matriz de Van Der Monde. O número de condicionamento das matrizes de Van Der Monde pode ser muito grande, originando problemas numéricos sérios para o cálculo dos coeficientes do polinômio. Pelo motivo anterior, esta abordagem não é de nosso interesse. Porém, tem-se um método alternativo semelhante, que será visto na seguinte seção.

### 4.3 Interpolação de Lagrange

Suponha que são dados os pontos  $(x_1, f(x_1))$ ,  $(x_2, f(x_2))$ , ...,  $(x_n, f(x_n))$  que podem ser organizados de forma tabular como

$x$	$f(x)$
$x_1$	$y_1$
$x_2$	$y_2$
$\vdots$	$\vdots$
$x_n$	$y_n$

sendo  $y_1 = f(x_1)$ ,  $y_2 = f(x_2)$ , ...,  $y_n = f(x_n)$ .

A idéia da interpolação de Lagrange é construir um polinômio  $p(x)$  de grau  $n-1$  que satisfaz as condições  $p(x_i) = y_i$ , para todo  $i = 1, 2, \dots, n$ , utilizando a igualdade

$$p(x) = y_1\ell_1(x) + y_2\ell_2(x) + \dots + y_n\ell_n(x),$$

onde  $\ell_1(x)$ ,  $\ell_2(x)$ , ...,  $\ell_n(x)$  são polinômios de grau  $n-1$ , denominados os polinômios fundamentais de Lagrange, que, para satisfazer as condições  $p(x_i) = y_i$ , são tais que

$$\begin{aligned} \ell_1(x_1) &= 1, & \ell_2(x_1) &= 0, & \dots, & \ell_n(x_1) &= 0, \\ \ell_1(x_2) &= 0, & \ell_2(x_2) &= 1, & \dots, & \ell_n(x_2) &= 0, \\ & \vdots & & & & & \\ \ell_1(x_n) &= 0, & \ell_2(x_n) &= 0, & \dots, & \ell_n(x_n) &= 1, \end{aligned}$$

pois, dessa maneira, por exemplo,

$$p(x_2) = y_1 \ell_1(x_2) + y_2 \ell_2(x_2) + \dots + y_n \ell_n(x_2) = y_1 \cdot 0 + y_2 \cdot 1 + \dots + y_n \cdot 0 = y_2.$$

Baseado no texto anterior, pode-se dizer que a interpolação de Lagrange é uma média ponderada dos valores

$y_1, y_2, \dots, y_n$ , com pesos  $\ell_1(x), \ell_2(x), \dots, \ell_n(x)$ , respectivamente.

Procede-se a construir os polinômios  $\ell_1(x), \ell_2(x), \dots, \ell_n(x)$ . O polinômio  $\ell_1(x)$  é de grau  $n-1$  e

satisfaz as condições  $\ell_1(x_1) = 1, \ell_1(x_2) = 0, \dots, \ell_1(x_n) = 0$ . As últimas  $n-2$  condições implicam

que  $\ell_1(x)$  deve ser da forma  $\ell_1(x) = \alpha_1(x-x_2) \cdots (x-x_n)$ , sendo  $\alpha_1$  uma constante que pode ser obtida

aplicando a primeira condição,  $\ell_1(x_1) = 1$ . Assim,  $1 = \alpha_1(x_1-x_2) \cdots (x_1-x_n)$ , implicando que

$$\ell_1(x) = \frac{(x-x_2)(x-x_3) \cdots (x-x_n)}{(x_1-x_2)(x_1-x_3) \cdots (x_1-x_n)}.$$

Similarmente, obtém-se que

$$\ell_2(x) = \frac{(x-x_1)(x-x_3) \cdots (x-x_n)}{(x_2-x_1)(x_2-x_3) \cdots (x_2-x_n)},$$

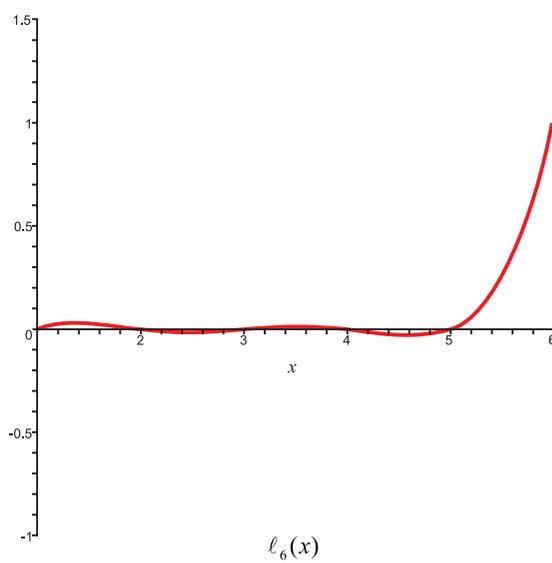
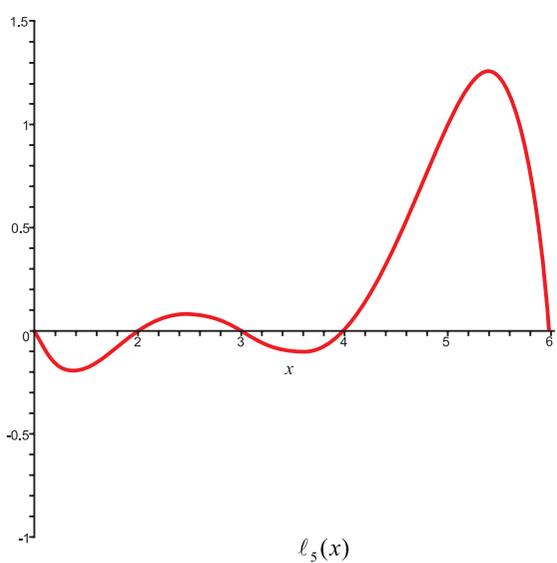
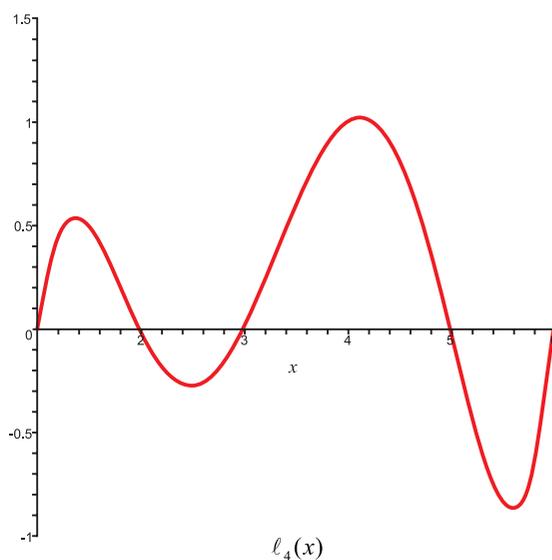
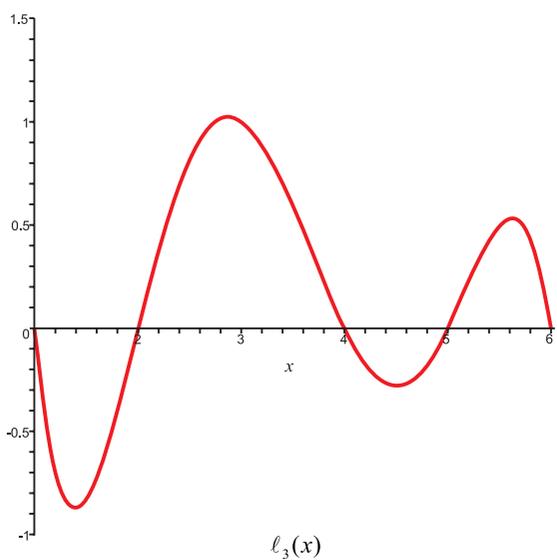
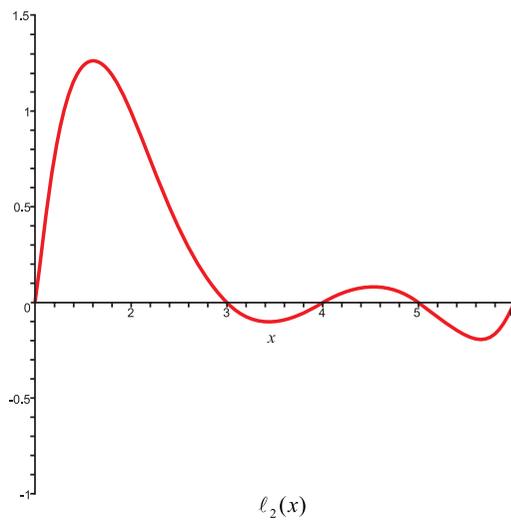
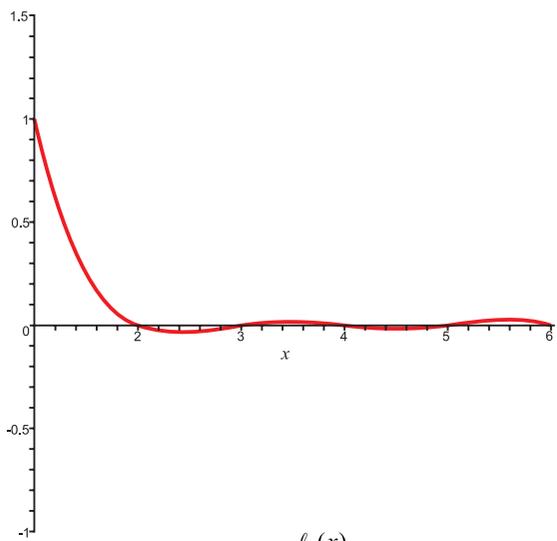
$$\ell_i(x) = \frac{(x-x_1)(x-x_2) \cdots (x-x_{i-1})(x-x_{i+1}) \cdots (x-x_n)}{(x_i-x_1)(x_i-x_2) \cdots (x_i-x_{i-1})(x_i-x_{i+1}) \cdots (x_i-x_n)}, \quad i = 2, \dots, n-1,$$

$$\ell_n(x) = \frac{(x-x_1)(x-x_2) \cdots (x-x_{n-1})}{(x_n-x_1)(x_n-x_2) \cdots (x_n-x_{n-1})}.$$

Por exemplo, para a tabela

$x$	$f(x)$
1	$y_1$
2	$y_2$
3	$y_3$
4	$y_4$
5	$y_5$
6	$y_6$

tem-se os polinômios interpoladores mostrados na figura a seguir



É interessante observar que existe simetria entre os polinômios  $l_1(x)$  e  $l_6(x)$ ,  $l_2(x)$  e  $l_5(x)$ ,  $l_3(x)$  e  $l_4(x)$ .

**Exemplo 4.1:** Sejam dados pontos organizados na forma

$x$	$f(x)$
1	-1
2	2
3	1
4	2

Quer-se determinar um polinômio cujo gráfico passa pelos dados e deseja-se saber o valor aproximado de  $f(x)$  quando  $x = 2,43$ . Os polinômios interpoladores são

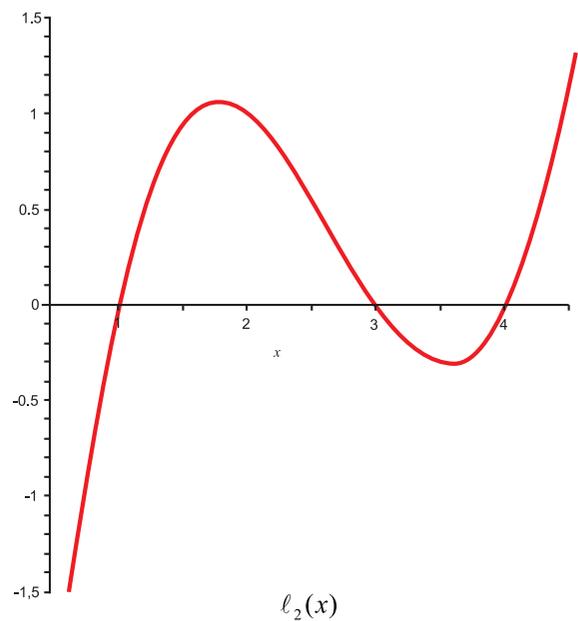
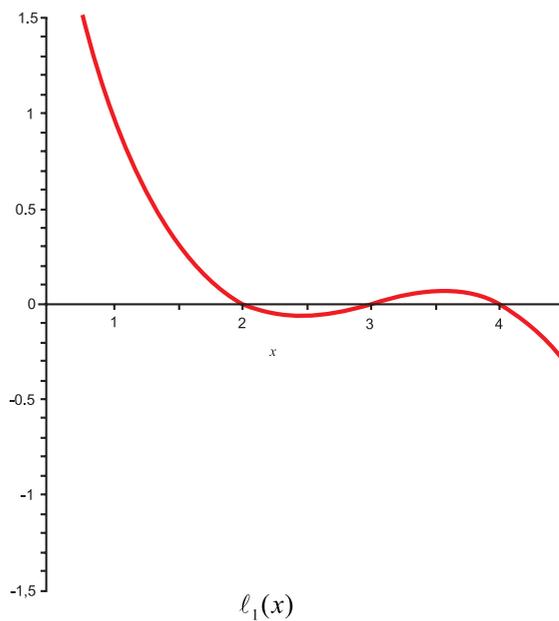
$$\ell_1(x) = \frac{(x-2)(x-3)(x-4)}{(1-2)(1-3)(1-4)} = -\frac{1}{6}(x-2)(x-3)(x-4) = -\frac{1}{6}x^3 + \frac{3}{2}x^2 - \frac{13}{3}x + 4,$$

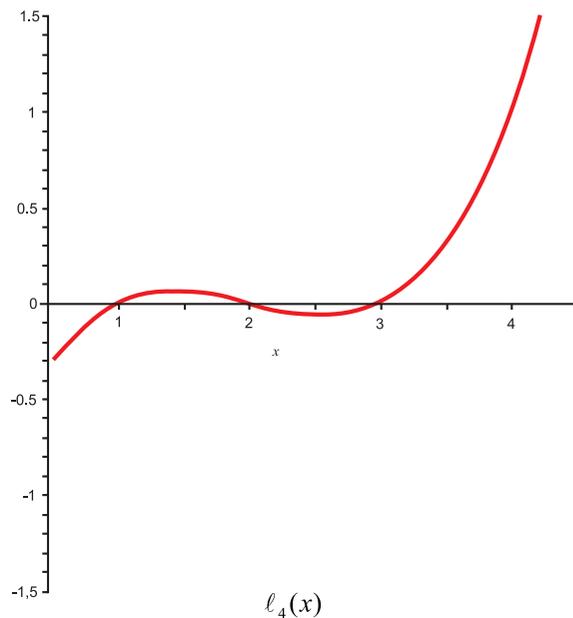
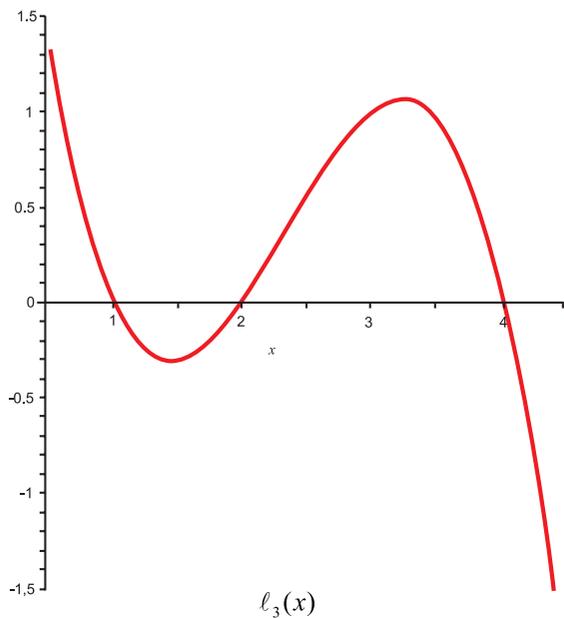
$$\ell_2(x) = \frac{(x-1)(x-3)(x-4)}{(2-1)(2-3)(2-4)} = \frac{1}{2}(x-1)(x-3)(x-4) = \frac{1}{2}x^3 - 4x^2 + \frac{19}{2}x - 6,$$

$$\ell_3(x) = \frac{(x-1)(x-2)(x-4)}{(3-1)(3-2)(3-4)} = -\frac{1}{3}(x-1)(x-2)(x-4) = -\frac{1}{2}x^3 + \frac{7}{2}x^2 - 7x + 4,$$

$$\ell_4(x) = \frac{(x-1)(x-2)(x-3)}{(4-1)(4-2)(4-3)} = \frac{1}{6}(x-1)(x-2)(x-3) = \frac{1}{6}x^3 - x^2 + \frac{11}{6}x - 1,$$

e cujos gráficos são mostrados a seguir:

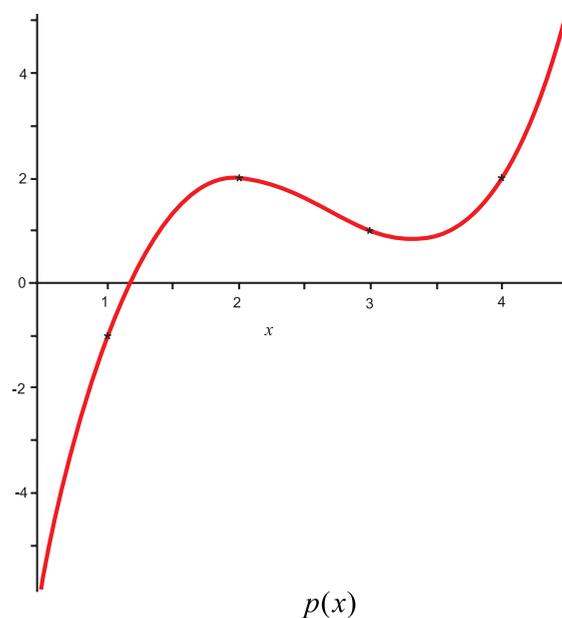




Assim, constrói-se o polinômio interpolador de acordo com

$$\begin{aligned}
 p(x) &= y_1 l_1(x) + y_2 l_2(x) + y_3 l_3(x) + y_4 l_4(x) \\
 &= (-1) \left( -\frac{1}{6}x^3 + \frac{3}{2}x^2 - \frac{13}{3}x + 4 \right) + 2 \left( \frac{1}{2}x^3 - 4x^2 + \frac{19}{2}x - 6 \right) \\
 &\quad + 1 \left( -\frac{1}{2}x^3 + \frac{7}{2}x^2 - 7x + 4 \right) + 2 \left( \frac{1}{6}x^3 - x^2 + \frac{11}{6}x - 1 \right) \\
 &= x^3 - 8x^2 + 20x - 14
 \end{aligned}$$

cujo gráfico é mostrado na figura seguinte:



Por outro lado, o valor aproximado de  $f(x)$  quando  $x = 2,43$ , será  $p(2,43) = 1,709707$ .

#### Observações 4.1:

1. Supondo que  $x_1 < x_2 < \dots < x_n$ , e conhecida a função  $f(x)$ , pode-se estimar o erro de aproximação,  $E(x)$ , de  $p(x)$  em relação a  $f(x)$ , mediante a expressão

$$|E(x)| \leq |(x - x_1)(x - x_2) \cdots (x - x_n)| \cdot \frac{M_n}{n!},$$

sendo  $M_n = \max \{|f^{(n)}(x)| : x \in [x_1, x_n]\}$ .

2. O método de Lagrange possui a desvantagem, entre outras, de produzir um polinômio de grau  $n - 1$ , pois se  $n$  é muito grande (como costuma ser), devem ser construídos muitos polinômios. Na seção a seguir, descreve-se um método onde o grau do polinômio construído pode ser controlado, obtendo-se desta maneira um polinômio que não necessariamente satisfaz as condições  $p(x_i) = y_i$ , mas o erro é conhecido de acordo com o grau do polinômio construído.

**Exemplo 4.2:** Suponha que são tabelados certos valores da função  $\cos(x)$ , como mostra a tabela a seguir:

$x$	$f(x) = \cos(x)$
0,1	$\cos(0,1) = 0,9950041653$
0,3	$\cos(0,3) = 0,9553364891$
0,4	$\cos(0,4) = 0,9210609940$
0,5	$\cos(0,5) = 0,8775825619$
0,6	$\cos(0,6) = 0,8253356149$
0,9	$\cos(0,9) = 0,6216099683$

Pode-se verificar que os polinômios fundamentais de Lagrange são:

$$l_1(x) = -104,1666667 (x - 0,3)(x - 0,4)(x - 0,5)(x - 0,6)(x - 0,9),$$

$$l_2(x) = 1388,888889 (x - 0,1)(x - 0,4)(x - 0,5)(x - 0,6)(x - 0,9),$$

$$l_3(x) = -333,333333 (x - 0,1)(x - 0,3)(x - 0,5)(x - 0,6)(x - 0,9),$$

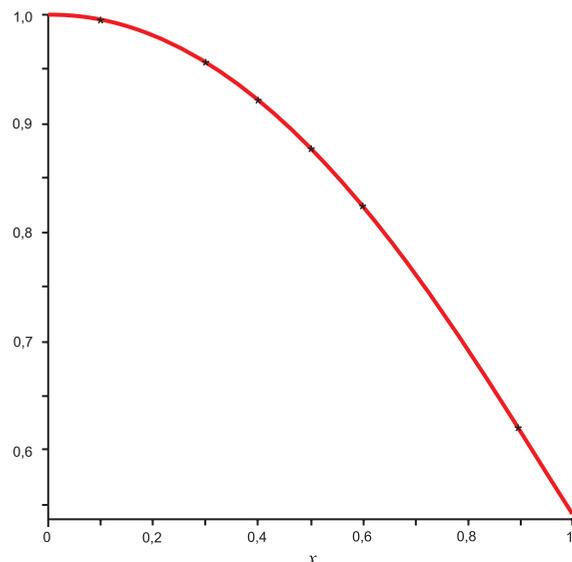
$$l_4(x) = 3125 (x - 0,1)(x - 0,3)(x - 0,4)(x - 0,6)(x - 0,9),$$

$$l_5(x) = -111,111111 (x - 0,1)(x - 0,3)(x - 0,4)(x - 0,5)(x - 0,9), \text{ e}$$

$$l_6(x) = 34,72222222 (x - 0,1)(x - 0,3)(x - 0,4)(x - 0,5)(x - 0,6);$$

formando o polinômio de interpolação

$p(x) = -0,003732 x^5 + 0,045688 x^4 - 0,002186 x^3 - 0,499378226 x^2 - 0,0000851 x + 1,000004123$ ,  
cujo gráfico, junto com os pontos da tabela, é mostrado a seguir



Polinômio  $p(x)$  que interpola os valores de  $\cos(x)$  de acordo com os pontos da tabela

Agora, o erro de aproximação de  $p(x)$  em relação a  $f(x)$  pode ser estimado como

$$|E(x)| \leq |(x-0,1)(x-0,3)(x-0,4)(x-0,5)(x-0,6)(x-0,9)| \cdot \frac{1}{6!}$$

$$= \frac{|(x-0,1)(x-0,3)(x-0,4)(x-0,5)(x-0,6)(x-0,9)|}{720}$$

para qualquer  $x \in [0,1, 0,9]$ , pois

$$M_6 = \max \{|f^{(6)}(x)| : x \in [0,1, 0,9]\} = \max \{|-\cos(x)| : x \in [0,1, 0,9]\} \leq 1.$$

Outro fato importante a ser notado neste exemplo é que alguns dos coeficientes do polinômio  $p(x)$  são pequenos em relação aos outros, reforçando o fato que é necessária uma abordagem que controle o grau do polinômio que está sendo construído.

**Exemplo 4.3:** Considere os dados tabelados da função  $f(x) = \frac{1}{x}$  mostrados a seguir:

$x$	$f(x)$
1	1
2	0,5
4	0,25
5	0,2
6	0,166667

Deseja-se saber o valor de  $f(3)$  mediante polinômios de Lagrange de vários graus e compará-los em relação à exatidão.

Procede-se, primeiro, a utilizar a interpolação linear, ou seja, um polinômio de Lagrange de primeiro grau. Observe que o  $x = 3$  está entre os valores  $x = 2$  e  $x = 4$ , e então constrói-se o polinômio

$$p_1(x) = \frac{x-4}{2-4} \cdot 0,5 + \frac{x-2}{4-2} \cdot 0,25, \text{ que dá}$$

$$p_1(3) = \frac{3-4}{2-4} \cdot 0,5 + \frac{3-2}{4-2} \cdot 0,25 = \frac{1}{2} \cdot 0,5 + \frac{1}{2} \cdot 0,25 = 0,25 + 0,125 = 0,375$$

Agora, formula-se a interpolação quadrática, ou seja, um polinômio de Lagrange de segundo grau. Embora exista outra possibilidade, constrói-se o polinômio a seguir, baseado nos pontos  $x = 2$ ,  $x = 4$  e  $x = 5$ ,

$$p_2(x) = \frac{(x-4)(x-5)}{(2-4)(2-5)} \cdot 0,5 + \frac{(x-2)(x-5)}{(4-2)(4-5)} \cdot 0,25 + \frac{(x-2)(x-4)}{(5-2)(5-4)} \cdot 0,2, \text{ que dá}$$

$$\begin{aligned} p_2(3) &= \frac{(3-4)(3-5)}{(2-4)(2-5)} \cdot 0,5 + \frac{(3-2)(3-5)}{(4-2)(4-5)} \cdot 0,25 + \frac{(3-2)(3-4)}{(5-2)(5-4)} \cdot 0,2 \\ &= \frac{1}{3} \cdot 0,5 + 1 \cdot 0,25 - \frac{1}{3} \cdot 0,2 = 0,35. \end{aligned}$$

Trabalha-se agora com a interpolação cúbica, ou seja, um polinômio de Lagrange de terceiro grau. Embora existam outras possibilidades, constrói-se o polinômio, baseado nos pontos  $x = 2$ ,  $x = 4$ ,  $x = 5$  e  $x = 6$ ,

$$\begin{aligned} p_3(x) &= \frac{(x-4)(x-5)(x-6)}{(2-4)(2-5)(2-6)} \cdot 0,5 + \frac{(x-2)(x-5)(x-6)}{(4-2)(4-5)(5-6)} \cdot 0,25 + \frac{(x-2)(x-4)(x-6)}{(5-2)(5-4)(5-6)} \cdot 0,2 \\ &\quad + \frac{(x-2)(x-4)(x-5)}{(6-2)(6-4)(6-5)} \cdot 0,166667 \end{aligned}$$

que produz

$$\begin{aligned} p_3(3) &= \frac{(3-4)(3-5)(3-6)}{(2-4)(2-5)(2-6)} \cdot 0,5 + \frac{(3-2)(3-5)(3-6)}{(4-2)(4-5)(4-6)} \cdot 0,25 + \frac{(3-2)(3-4)(3-6)}{(5-2)(5-4)(5-6)} \cdot 0,2 \\ &\quad + \frac{(3-2)(3-4)(3-5)}{(6-2)(6-4)(6-5)} \cdot 0,166667 \\ &= \frac{1}{4} \cdot 0,5 + \frac{3}{2} \cdot 0,25 - 1 \cdot 0,2 + \frac{1}{4} \cdot 0,166667 = 0,341667. \end{aligned}$$

Enfim, o polinômio interpolador de Lagrange de quarto grau utiliza todos os dados da tabela:

$$\begin{aligned} p_4(x) &= \frac{(x-2)(x-4)(x-5)(x-6)}{(1-2)(1-4)(1-5)(1-6)} \cdot 1 + \frac{(x-1)(x-4)(x-5)(x-6)}{(2-1)(2-4)(2-5)(2-6)} \cdot 0,5 \\ &\quad + \frac{(x-1)(x-2)(x-5)(x-6)}{(4-1)(4-2)(4-5)(5-6)} \cdot 0,25 + \frac{(x-1)(x-2)(x-4)(x-6)}{(5-1)(5-2)(5-4)(5-6)} \cdot 0,2 \\ &\quad + \frac{(x-1)(x-2)(x-4)(x-5)}{(6-1)(6-2)(6-4)(6-5)} \cdot 0,166667, \end{aligned}$$

que fornece o valor

$$\begin{aligned}
p_4(3) &= \frac{(3-2)(3-4)(3-5)(3-6)}{(1-2)(1-4)(1-5)(1-6)} \cdot 1 + \frac{(3-1)(3-4)(3-5)(3-6)}{(2-1)(2-4)(2-5)(2-6)} \cdot 0,5 \\
&+ \frac{(3-1)(3-2)(3-5)(3-6)}{(4-1)(4-2)(4-5)(5-6)} \cdot 0,25 + \frac{(3-1)(3-2)(3-4)(3-6)}{(5-1)(5-2)(5-4)(5-6)} \cdot 0,2 \\
&+ \frac{(3-1)(3-2)(3-4)(3-5)}{(6-1)(6-2)(6-4)(6-5)} \cdot 0,166667 \\
&= -\frac{1}{10} \cdot 1 + \frac{1}{2} \cdot 0,5 + 1 \cdot 0,25 - \frac{1}{2} \cdot 0,2 + \frac{1}{10} \cdot 0,166667 \\
&= 0,316667.
\end{aligned}$$

Como o valor de  $f(3)$  é conhecido e igual a  $0,333333$ , então comparam-se os resultados obtidos:

$$\begin{aligned}
|f(3) - p_1(3)| &= |0,333333 - 0,375| = 0,041667, \\
|f(3) - p_2(3)| &= |0,333333 - 0,35| = 0,016667, \\
|f(3) - p_3(3)| &= |0,333333 - 0,341667| = 0,008334, \\
|f(3) - p_4(3)| &= |0,333333 - 0,316667| = 0,016667.
\end{aligned}$$

Aqui, pode-se observar que  $p_3(3)$  é a aproximação mais exata. Uma maneira de comprovar isto é que de acordo com a observação 4.1, o erro cometido no caso de  $p_3(x)$  estará limitado por

$$\left| (3-2)(3-4)(3-5)(3-6) \cdot \frac{M_4}{4!} \right| \leq \frac{6}{24} \cdot 24 = 6,$$

pois  $M_4 = \max \{ |f^{(4)}(x)| : x \in [1,6] \} = \max \left\{ \left| \frac{24}{x^5} \right| : x \in [1,6] \right\} = 24$ ; e o erro cometido no caso de

$p_4(x)$  estará limitado por,

$$\left| (3-1)(3-2)(3-4)(3-5)(3-6) \cdot \frac{M_5}{5!} \right| \leq \frac{12}{120} \cdot 120 = 12,$$

pois  $M_5 = \max \{ |f^{(5)}(x)| : x \in [1,6] \} = \max \left\{ \left| -\frac{120}{x^6} \right| : x \in [1,6] \right\} = 120$ . Pode-se concluir que nem

sempre quanto maior o grau do polinômio interpolador maior a precisão da aproximação.

#### 4.4 Interpolação de Newton

Como antes, são dados certos pontos tabulares mostrados na tabela a seguir, e pretende-se encontrar um polinômio  $p(x)$  tal que  $y_i = p(x_i)$ , para todo  $i = 1, 2, \dots, n$ :

$x$	$y$
$x_1$	$y_1$
$x_2$	$y_2$
$\vdots$	$\vdots$
$x_n$	$y_n$

A idéia é escrever o polinômio  $p(x)$  na forma

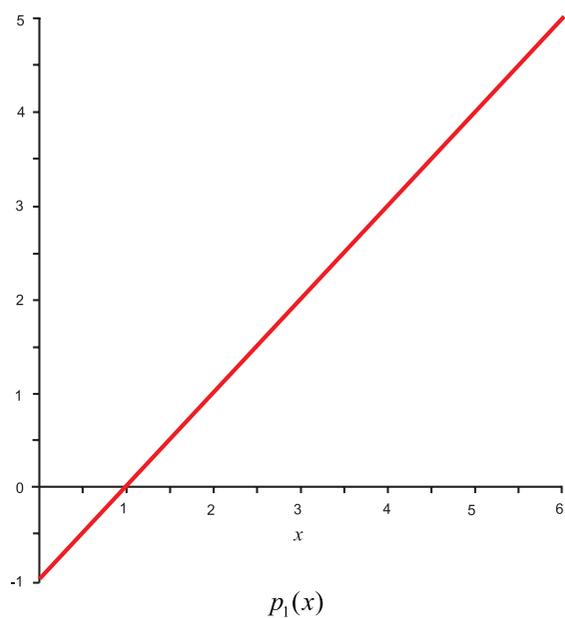
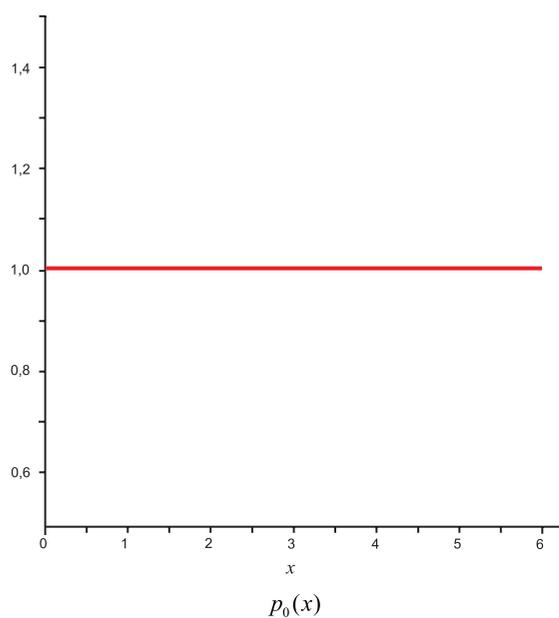
$$p(x) = \alpha_0 + \alpha_1(x - x_1) + \alpha_2(x - x_1)(x - x_2) + \cdots + \alpha_{n-1}(x - x_1)(x - x_2) \cdots (x - x_{n-1}).$$

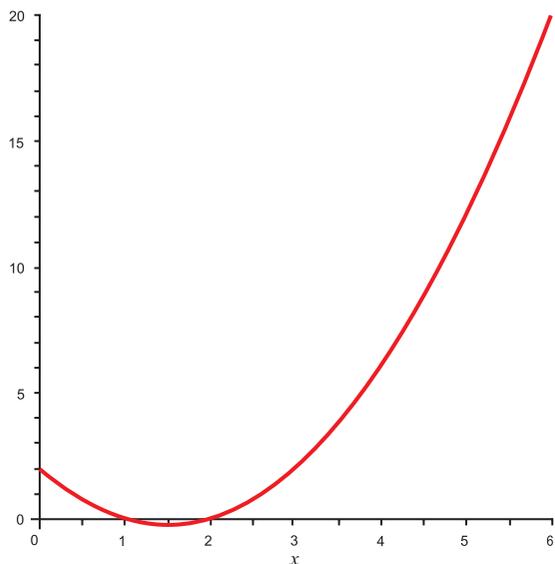
Observe que, no caso da interpolação de Lagrange, o polinômio interpolador foi construído mediante os polinômios fundamentais de Lagrange  $\ell_1(x)$ ,  $\ell_2(x)$ , ...,  $\ell_n(x)$ . No caso da interpolação de Newton, o mesmo polinômio interpolador é construído utilizando os polinômios  $p_0(x) = 1$ ,  $p_1(x) = x - x_1$ ,  $p_2(x) = (x - x_1)(x - x_2)$ , ...,  $p_{n-1}(x) = (x - x_1)(x - x_2) \cdots (x - x_{n-1})$ .

Por exemplo, para a tabela

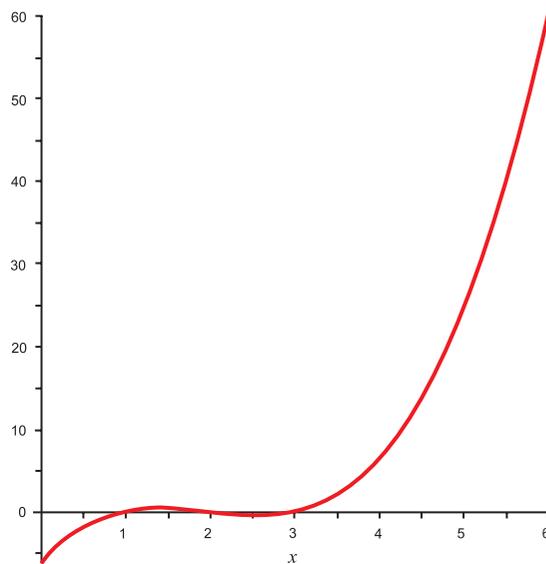
$x$	$f(x)$
1	$y_1$
2	$y_2$
3	$y_3$
4	$y_4$
5	$y_5$
6	$y_6$

tem-se os polinômios mostrados na figura a seguir

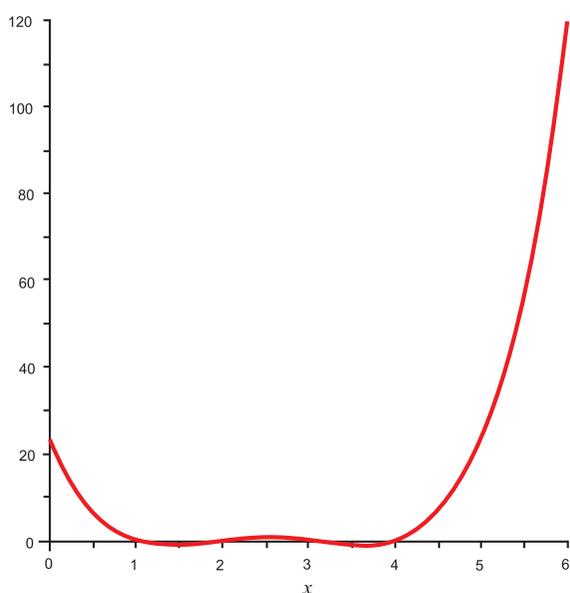




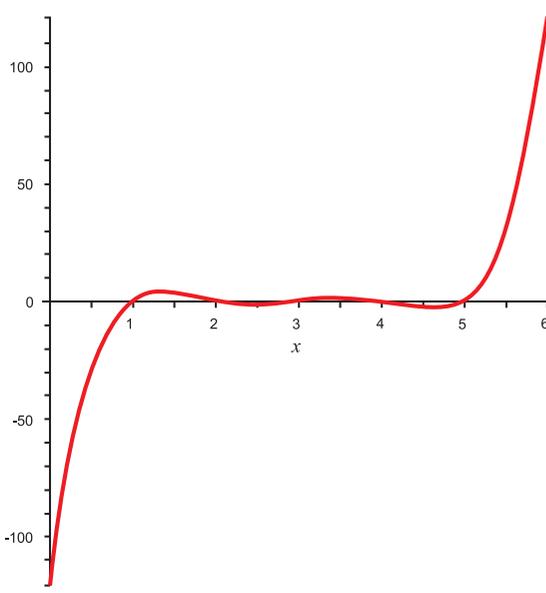
$p_2(x)$



$p_3(x)$



$p_4(x)$



$p_5(x)$

Após algum trabalho algébrico, impondo a condição que  $y_i = p(x_i)$ , verifica-se que os coeficientes  $\{\alpha_0, \alpha_1, \dots, \alpha_{n-1}\}$  fazem parte de um conjunto mais amplo descrito a seguir:

Diferenças divididas de ordem 0 :

$$\alpha_0 = y_1 \equiv [y_1], [y_2] = y_2, \dots, [y_n] = y_n;$$

Diferenças divididas de ordem 1 :

$$\alpha_1 = \frac{[y_2] - [y_1]}{x_2 - x_1} \equiv [y_1, y_2], [y_2, y_3] = \frac{[y_3] - [y_2]}{x_3 - x_2}, \dots, [y_{n-1}, y_n] = \frac{[y_n] - [y_{n-1}]}{x_n - x_{n-1}},$$

Diferenças divididas de ordem 2

$$\alpha_2 = \frac{[y_2, y_3] - [y_1, y_2]}{x_3 - x_1} \equiv [y_1, y_2, y_3], [y_2, y_3, y_4] = \frac{[y_3, y_4] - [y_2, y_3]}{x_4 - x_2}, \dots,$$

$$[y_{n-2}, y_{n-1}, y_n] = \frac{[y_{n-1}, y_n] - [y_{n-2}, y_{n-1}]}{x_n - x_{n-2}},$$

⋮

Diferença dividida de ordem  $n - 1$ :

$$\alpha_{n-1} = \frac{[y_2, y_3, \dots, y_n] - [y_1, y_2, \dots, y_{n-1}]}{x_n - x_1} \equiv [y_1, y_2, \dots, y_n]$$

Os coeficientes necessários para a construção do polinômio  $p(x)$ , junto com as diferenças divididas, podem ser conseguidos de maneira sistemática mediante uma tabela de acordo com o seguinte formato

$x$	$y$	ordem 1	ordem 2	ordem 3	⋯	ordem $n - 1$
$x_1$	$[y_1] = \alpha_0$					
		$[y_1, y_2] = \alpha_1$				
$x_2$	$[y_2]$		$[y_1, y_2, y_3] = \alpha_2$			
		$[y_2, y_3]$		$[y_1, y_2, y_3, y_4] = \alpha_3$		
$x_3$	$[y_3]$		$[y_2, y_3, y_4]$		⋮	
		$[y_3, y_4]$		$[y_2, y_3, y_4, y_5]$		$[y_1, y_2, \dots, y_n] = \alpha_{n-1}$
$x_4$	$[y_4]$		$[y_3, y_4, y_5]$			
		$[y_4, y_5]$		⋮		
$x_5$	$[y_5]$		⋮	$[y_{n-3}, y_{n-2}, y_{n-1}, y_n]$		
		⋮	$[y_{n-2}, y_{n-1}, y_n]$			
⋮	⋮	$[y_{n-1}, y_n]$				
$x_n$	$[y_n]$					

**Exemplo 4.4:** Considere os valores tabulares dados no exemplo 4.1. Procede-se a construir a tabela de diferenças divididas:

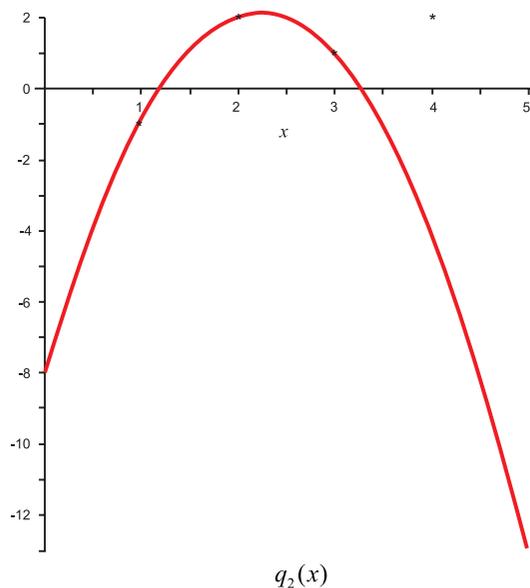
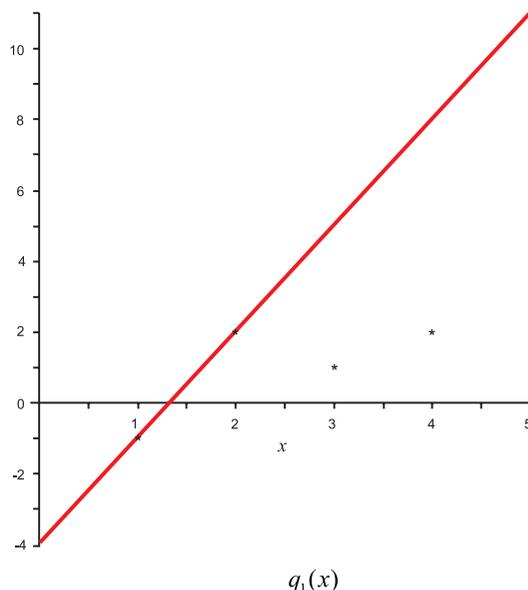
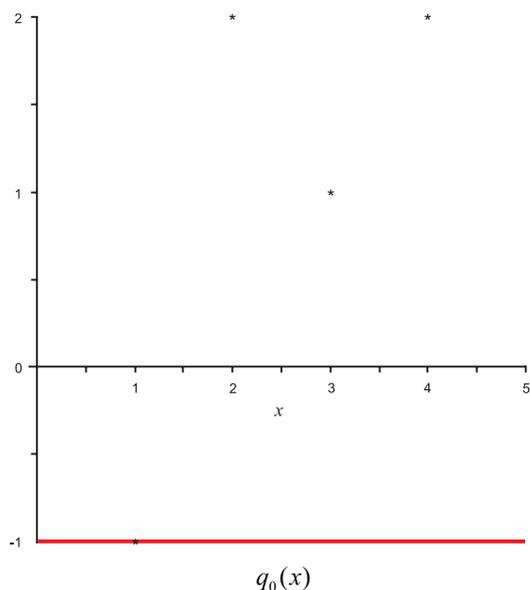
$x$	$f(x)$	ordem 1	ordem 2	ordem 3
1	-1			
		$\frac{2 - (-1)}{2 - 1} = 3$		
2	2		$\frac{-1 - 3}{3 - 1} = -2$	
		$\frac{1 - 2}{3 - 2} = -1$		$\frac{1 - (-2)}{4 - 1} = 1$
3	1		$\frac{1 - (-1)}{4 - 2} = 1$	
		$\frac{2 - 1}{4 - 3} = 1$		
4	2			

Como antes, quer-se determinar um polinômio cujo gráfico passa pelos pontos dados. Neste caso, o polinômio interpolador de Newton é

$$\begin{aligned} p(x) &= -1 + 3(x-1) - 2(x-1)(x-2) + (x-1)(x-2)(x-3) \\ &= -1 + 3x - 3 - 2x^2 + 6x - 4 + x^3 - 6x^2 + 11x - 6 \\ &= x^3 - 8x^2 + 20x - 14, \end{aligned}$$

coincidindo com o polinômio interpolador de Lagrange, construído na seção anterior.

Mas, neste mesmo exemplo, podem ser considerados outros polinômios de menor grau que  $p(x)$ ,  $q_0(x) = -1$ ,  $q_1(x) = -1 + 3(x-1)$  e  $q_2(x) = -1 + 3(x-1) - 2(x-1)(x-2)$ , cujos gráficos mostram-se a seguir



Observe que os polinômios  $q_0(x)$ ,  $q_1(x)$  e  $q_2(x)$  passam por um, dois e três pontos da tabela, respectivamente.

Se fosse requerido aproximar o valor de  $f(x)$  quando  $x = 1,4$ , mediante interpolação linear, seria utilizada  $q_1(x) = -1 + 3(x-1)$ , resultando em  $f(1,4) \approx q_1(1,4) = -1 + 3(1,4 - 1) = 0,2$ .

Por outro lado, se fosse requerido aproximar o valor de  $f(x)$  quando  $x = 2,43$ , mediante interpolação linear, seria utilizada  $r_1(x) = 2 - (x - 2)$ , resultando em  $f(2,43) \approx r_1(2,43) = 2 - (2,43 - 2) = 1,57$ . Os coeficientes  $\alpha_0 = 2$  e  $\alpha_1 = -1$  do polinômio  $r_1(x)$  foram escolhidos da tabela de diferenças divididas, observando que o valor  $x = 2,43$  encontra-se entre os valores  $x = 2$  e  $x = 3$  da tabela.

Cabe mencionar que existem versões mais sofisticadas da interpolação de Newton, aplicável a casos especiais, quando, por exemplo, os pontos  $x_1, x_2, \dots, x_n$  são igualmente espaçados, ou quando a tabela de diferenças leva à construção de outros polinômios interpoladores, embora estes assuntos não serão tratados aqui.

*Ajuste por Quadros  
Mínimos*

**5**  
Capítulo

Com frequência, como resultado de uma série de medições, é requerido determinar o valor de alguma grandeza ou constante física. Este valor deve ser calculado supondo hipóteses simplificadoras, ou seja, suposições a priori sobre a grandeza ou constante (por exemplo, comportamento linear, quadrático, exponencial, etc). Desde que as medições fornecem valores discretos, é necessário encontrar o valor que melhor ajuste as medições, utilizando algum critério ótimo. Este, entre outros, é o objetivo principal de uma área da matemática, conhecida como teoria da aproximação.

Em geral, um dos objetivos da teoria da aproximação é determinar uma função, dentre uma classe de funções, que melhor represente um conjunto de dados. Esta parte da teoria é denominada ajuste discreto.

Outro objetivo da teoria da aproximação é determinar uma função mais simples, dentre uma classe de funções, que sirva para aproximar os valores de uma função dada (hipoteticamente, complicada). Esta parte da teoria é denominada ajuste contínuo.

Neste capítulo, serão abordados os dois tipos de ajuste mencionados, utilizando o critério denominado de quadrados mínimos.

Vale a pena mencionar que a teoria de aproximação é bastante utilizada no tratamento de dados experimentais.

### 5.1 Ajuste Discreto Linear por Quadrados Mínimos

Seja uma função  $f(x)$  e um conjunto de pontos,  $(x_1, f(x_1))$ ,  $(x_2, f(x_2))$ , ...,  $(x_n, f(x_n))$ . Também, são escolhidas certas funções  $g_1(x)$ ,  $g_2(x)$ , ...,  $g_m(x)$ , ( $m \in \mathbf{N}$ ). O objetivo do ajuste discreto linear por quadrados mínimos é obter coeficientes  $\alpha_1, \alpha_2, \dots, \alpha_m$  tais que a função

$$g(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \dots + \alpha_m g_m(x)$$

satisfaz a condição que o somatório

$$\sum_{i=1}^n (f(x_i) - g(x_i))^2 = (f(x_1) - g(x_1))^2 + (f(x_2) - g(x_2))^2 + \dots + (f(x_n) - g(x_n))^2$$

tem o menor valor possível.

Observe que a função  $g(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \dots + \alpha_m g_m(x)$  é uma combinação linear das funções  $g_1(x)$ ,  $g_2(x)$ , ...,  $g_m(x)$ . Por este motivo, o ajuste é dito linear. Aqui, cabe salientar que são possíveis outras formas de ajuste discreto (racional, exponencial, logarítmico, etc).

A seguir, detalhar-se-á, sob algumas hipóteses, o processo de cálculo dos coeficientes  $\alpha_1, \alpha_2, \dots, \alpha_m$ . Primeiro, define-se a função real de várias variáveis

$$F(\alpha_1, \alpha_2, \dots, \alpha_m) = \sum_{i=1}^n (f(x_i) - g(x_i))^2 = \sum_{i=1}^n (f(x_i) - \alpha_1 g_1(x_i) - \alpha_2 g_2(x_i) - \dots - \alpha_m g_m(x_i))^2,$$

e procede-se a minimizá-la mediante o critério da primeira derivada, fazendo

$$\frac{\partial}{\partial \alpha_j} F(\alpha_1, \alpha_2, \dots, \alpha_m) = 0, \quad j = 1, 2, \dots, m.$$

Para cada  $j$ , tem-se que

$$\frac{\partial}{\partial \alpha_j} F(\alpha_1, \alpha_2, \dots, \alpha_m) = -2 \sum_{i=1}^n (f(x_i) - \alpha_1 g_1(x_i) - \alpha_2 g_2(x_i) - \dots - \alpha_m g_m(x_i)) \cdot (g_j(x_i)),$$

que, após impor a condição da derivada igual a zero, produz

$$\sum_{i=1}^n (f(x_i) - \alpha_1 g_1(x_i) - \alpha_2 g_2(x_i) - \dots - \alpha_m g_m(x_i)) \cdot (g_j(x_i)) = 0, \text{ para cada } j = 1, 2, \dots, m.$$

A partir da igualdade anterior, forma-se o sistema de equações lineares de  $m$  equações com  $m$  incógnitas:

$$\begin{cases} \left[ \sum_{i=1}^n g_1(x_i)g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_1(x_i)g_2(x_i) \right] \alpha_2 + \dots + \left[ \sum_{i=1}^n g_1(x_i)g_m(x_i) \right] \alpha_m = \sum_{i=1}^n f(x_i)g_1(x_i) \\ \left[ \sum_{i=1}^n g_2(x_i)g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_2(x_i)g_2(x_i) \right] \alpha_2 + \dots + \left[ \sum_{i=1}^n g_2(x_i)g_m(x_i) \right] \alpha_m = \sum_{i=1}^n f(x_i)g_2(x_i) \\ \vdots \\ \left[ \sum_{i=1}^n g_m(x_i)g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_m(x_i)g_2(x_i) \right] \alpha_2 + \dots + \left[ \sum_{i=1}^n g_m(x_i)g_m(x_i) \right] \alpha_m = \sum_{i=1}^n f(x_i)g_m(x_i) \end{cases}$$

denominadas as equações normais do ajuste, que podem ser escritas de maneira matricial como

$$\begin{pmatrix} \sum_{i=1}^n g_1(x_i)g_1(x_i) & \sum_{i=1}^n g_1(x_i)g_2(x_i) & \dots & \sum_{i=1}^n g_1(x_i)g_m(x_i) \\ \sum_{i=1}^n g_2(x_i)g_1(x_i) & \sum_{i=1}^n g_2(x_i)g_2(x_i) & \dots & \sum_{i=1}^n g_2(x_i)g_m(x_i) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n g_m(x_i)g_1(x_i) & \sum_{i=1}^n g_m(x_i)g_2(x_i) & \dots & \sum_{i=1}^n g_m(x_i)g_m(x_i) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n f(x_i)g_1(x_i) \\ \sum_{i=1}^n f(x_i)g_2(x_i) \\ \vdots \\ \sum_{i=1}^n f(x_i)g_m(x_i) \end{pmatrix}.$$

Mediante a resolução deste sistema podem ser calculados os coeficientes requeridos.

Uma maneira de quantificar a precisão do ajuste é calcular o somatório

$$R = \sum_{i=1}^n (f(x_i) - g(x_i))^2.$$

No caso de existirem várias escolhas de funções  $g(x)$ , o melhor ajuste estará dado por aquela função que tenha o menor valor de  $R$ .

**Exemplo 5.1:** Dada a tabela

$x$	$f(x)$
1	4
2	6
3	8
4	9

é desejado determinar, pelo critério de quadrados mínimos, uma função da forma  $g(x) = \alpha_1 x^2$  que ajuste os dados mostrados.

Identificando tem-se que  $n = 4$ ,  $m = 1$  e  $g_1(x) = x^2$ . O sistema de equações normais do ajuste é por uma só equação

$$\left[ \sum_{i=1}^4 g_1(x_i)g_1(x_i) \right] \cdot \alpha_1 = \sum_{i=1}^4 f(x_i)g_1(x_i),$$

ou,

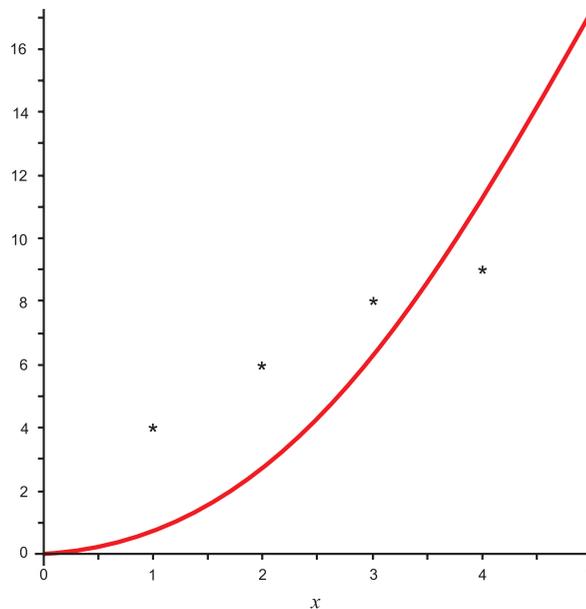
$$\left[ \sum_{i=1}^4 x_i^4 \right] \cdot \alpha_1 = \sum_{i=1}^4 x_i^2 f(x_i).$$

Agora, são calculados os somatórios

$$\sum_{i=1}^4 x_i^4 = x_1^4 + x_2^4 + x_3^4 + x_4^4 = 1^4 + 2^4 + 3^4 + 4^4 = 1 + 16 + 81 + 256 = 354 \text{ e}$$

$$\sum_{i=1}^4 x_i^2 f(x_i) = x_1^2 f(x_1) + x_2^2 f(x_2) + x_3^2 f(x_3) + x_4^2 f(x_4) = 1^2 \cdot 4 + 2^2 \cdot 6 + 3^2 \cdot 8 + 4^2 \cdot 9 = 244.$$

Então,  $\alpha_1 = \frac{244}{354} = 0,689266$ . Logo, a função de ajuste é  $g(x) = 0,689266 x^2$ . A figura a seguir ilustra graficamente o ajuste



O valor de  $R$  é

$$\begin{aligned} R &= \sum_{i=1}^4 (f(x_i) - g(x_i))^2 = (f(x_1) - g(x_1))^2 + (f(x_2) - g(x_2))^2 + (f(x_3) - g(x_3))^2 + (f(x_4) - g(x_4))^2 \\ &= (4 - 0,689266)^2 + (6 - 2,757062)^2 + (8 - 6,203390)^2 + (9 - 11,028249)^2 \\ &= 10,960963 + 10,516646 + 3,227808 + 4,113792 \\ &= 28,819209 \end{aligned}$$

### 5.1.1 Ajuste de uma Reta por Quadrados Mínimos

Considere os dados tabelados

$x$	$f(x)$
$x_1$	$y_1$
$x_2$	$y_2$
$\vdots$	$\vdots$
$x_n$	$y_n$

Quer-se ajustar uma função  $g(x) = \alpha_1 x + \alpha_2$  a esses dados. Observe que o gráfico da função  $g(x)$  determina uma reta. Neste caso,  $m = 2$ ,  $g_1(x) = x$  e  $g_2(x) = 1$ . Então as equações normais do ajuste estão dadas por

$$\begin{cases} \left[ \sum_{i=1}^n g_1(x_i)g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_1(x_i)g_2(x_i) \right] \alpha_2 = \sum_{i=1}^n y_i g_1(x_i) \\ \left[ \sum_{i=1}^n g_2(x_i)g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_2(x_i)g_2(x_i) \right] \alpha_2 = \sum_{i=1}^n y_i g_2(x_i) \end{cases}$$

que produz o sistema

$$\begin{cases} \left[ \sum_{i=1}^n x_i^2 \right] \alpha_1 + \left[ \sum_{i=1}^n x_i \right] \alpha_2 = \sum_{i=1}^n x_i y_i \\ \left[ \sum_{i=1}^n x_i \right] \alpha_1 + \left[ \sum_{i=1}^n 1 \right] \alpha_2 = \sum_{i=1}^n y_i \end{cases} \quad \text{ou,} \quad \begin{cases} \left[ \sum_{i=1}^n x_i^2 \right] \alpha_1 + \left[ \sum_{i=1}^n x_i \right] \alpha_2 = \sum_{i=1}^n x_i y_i \\ \left[ \sum_{i=1}^n x_i \right] \alpha_1 + n \alpha_2 = \sum_{i=1}^n y_i \end{cases}$$

O sistema anterior pode ser resolvido fazendo

$$\alpha_1 = \frac{\begin{vmatrix} \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \\ \sum_{i=1}^n y_i & n \end{vmatrix}}{\begin{vmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{vmatrix}} = \frac{n \left[ \sum_{i=1}^n x_i y_i \right] - \left[ \sum_{i=1}^n x_i \right] \cdot \left[ \sum_{i=1}^n y_i \right]}{n \left[ \sum_{i=1}^n x_i^2 \right] - \left[ \sum_{i=1}^n x_i \right]^2},$$

e da segunda equação,

$$\alpha_2 = \frac{\left[ \sum_{i=1}^n y_i \right] - \left[ \sum_{i=1}^n x_i \right] \alpha_1}{n}.$$

**Exemplo 5.2:** Considere a tabela do exemplo 5.1:

$x$	$f(x)$
1	4
2	6
3	8
4	9

Agora, deseja-se determinar, pelo critério de quadrados mínimos, uma função da forma  $g(x) = \alpha_1 x + \alpha_2$  que ajuste esses dados.

Como cálculos preliminares tem-se que

$$\sum_{i=1}^4 x_i^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1^2 + 2^2 + 3^2 + 4^2 = 30,$$

$$\sum_{i=1}^4 x_i y_i = x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4 = 1 \cdot 4 + 2 \cdot 6 + 3 \cdot 8 + 4 \cdot 9 = 76,$$

$$\sum_{i=1}^4 x_i = x_1 + x_2 + x_3 + x_4 = 1 + 2 + 3 + 4 = 10,$$

$$\sum_{i=1}^4 y_i = y_1 + y_2 + y_3 + y_4 = 4 + 6 + 8 + 9 = 27,$$

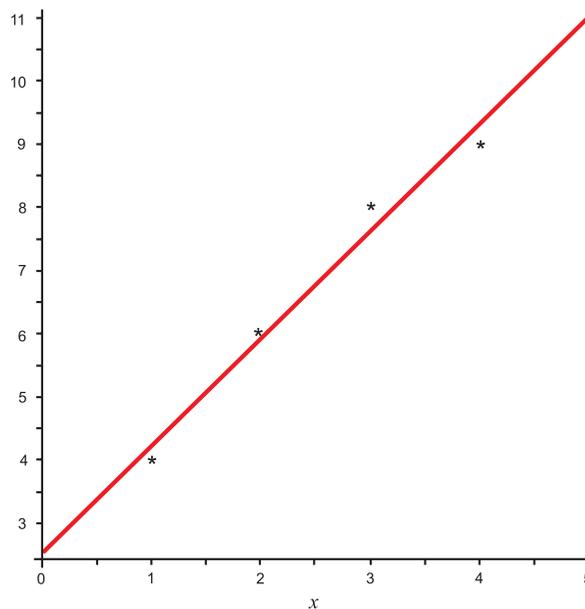
que produzem as equações normais

$$\begin{cases} 30 \alpha_1 + 10 \alpha_2 = 76, \\ 10 \alpha_1 + 4 \alpha_2 = 27, \end{cases}$$

cujas soluções são

$$\alpha_1 = \frac{\begin{vmatrix} 76 & 10 \\ 27 & 4 \end{vmatrix}}{\begin{vmatrix} 30 & 10 \\ 10 & 4 \end{vmatrix}} = \frac{4 \cdot 76 - 27 \cdot 10}{4 \cdot 30 - 10 \cdot 10} = \frac{34}{20} = 1,7 \text{ e } \alpha_2 = \frac{27 - 10 \cdot 1,7}{4} = \frac{10}{4} = 2,5.$$

Estes resultados geram a função de ajuste  $g(x) = 1,7x + 2,5$ . A figura a seguir ilustra graficamente o ajuste



Neste caso, o ajuste produz  $R$  igual a

$$\begin{aligned}
 R &= \sum_{i=1}^4 (f(x_i) - g(x_i))^2 = (f(x_1) - g(x_1))^2 + (f(x_2) - g(x_2))^2 + (f(x_3) - g(x_3))^2 + (f(x_4) - g(x_4))^2 \\
 &= (4 - 4,2)^2 + (6 - 5,9)^2 + (8 - 7,6)^2 + (9 - 9,3)^2 \\
 &= 0,04 + 0,01 + 0,16 + 0,09 \\
 &= 0,3
 \end{aligned}$$

**Exemplo 5.3:** Agora, ajusta-se aos dados dos exemplos anteriores, uma função quadrática  $g(x) = \alpha_1 x^2 + \alpha_2 x + \alpha_3$ . Tem-se que  $m = 3$ ,  $g_1(x) = x^2$ ,  $g_2(x) = x$  e  $g_3(x) = 1$ . As equações normais do ajuste estão dadas por

$$\begin{cases}
 \left[ \sum_{i=1}^n g_1(x_i)g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_1(x_i)g_2(x_i) \right] \alpha_2 + \left[ \sum_{i=1}^n g_1(x_i)g_3(x_i) \right] \alpha_3 = \sum_{i=1}^n y_i g_1(x_i) \\
 \left[ \sum_{i=1}^n g_2(x_i)g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_2(x_i)g_2(x_i) \right] \alpha_2 + \left[ \sum_{i=1}^n g_2(x_i)g_3(x_i) \right] \alpha_3 = \sum_{i=1}^n y_i g_2(x_i), \\
 \left[ \sum_{i=1}^n g_3(x_i)g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_3(x_i)g_2(x_i) \right] \alpha_2 + \left[ \sum_{i=1}^n g_3(x_i)g_3(x_i) \right] \alpha_3 = \sum_{i=1}^n y_i g_3(x_i),
 \end{cases}$$

que produz o sistema

$$\begin{cases}
 \left[ \sum_{i=1}^n x_i^4 \right] \alpha_1 + \left[ \sum_{i=1}^n x_i^3 \right] \alpha_2 + \left[ \sum_{i=1}^n x_i^2 \right] \alpha_3 = \sum_{i=1}^n x_i^2 y_i, \\
 \left[ \sum_{i=1}^n x_i^3 \right] \alpha_1 + \left[ \sum_{i=1}^n x_i^2 \right] \alpha_2 + \left[ \sum_{i=1}^n x_i \right] \alpha_3 = \sum_{i=1}^n x_i y_i, \\
 \left[ \sum_{i=1}^n x_i^2 \right] \alpha_1 + \left[ \sum_{i=1}^n x_i \right] \alpha_2 + n \alpha_3 = \sum_{i=1}^n y_i.
 \end{cases}$$

Realizam-se os cálculos preliminares

$$\sum_{i=1}^4 x_i^4 = x_1^4 + x_2^4 + x_3^4 + x_4^4 = 1^4 + 2^4 + 3^4 + 4^4 = 354,$$

$$\sum_{i=1}^4 x_i^3 = x_1^3 + x_2^3 + x_3^3 + x_4^3 = 1^3 + 2^3 + 3^3 + 4^3 = 100,$$

$$\sum_{i=1}^4 x_i^2 y_i = x_1^2 y_1 + x_2^2 y_2 + x_3^2 y_3 + x_4^2 y_4 = 1^2 \cdot 4 + 2^2 \cdot 6 + 3^2 \cdot 8 + 4^2 \cdot 9 = 244,$$

$$\sum_{i=1}^4 x_i^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1^2 + 2^2 + 3^2 + 4^2 = 30,$$

$$\sum_{i=1}^4 x_i y_i = x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4 = 1 \cdot 4 + 2 \cdot 6 + 3 \cdot 8 + 4 \cdot 9 = 76,$$

$$\sum_{i=1}^4 x_i = x_1 + x_2 + x_3 + x_4 = 1 + 2 + 3 + 4 = 10,$$

$$\sum_{i=1}^4 y_i = y_1 + y_2 + y_3 + y_4 = 4 + 6 + 8 + 9 = 27,$$

que produzem as equações normais

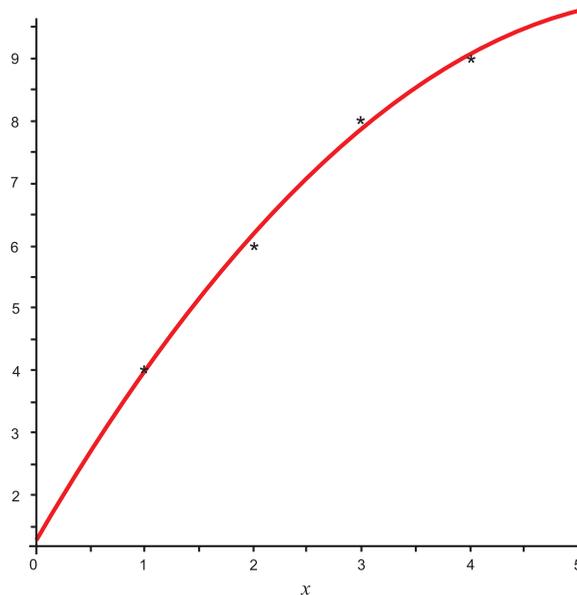
$$\begin{cases} 354 \alpha_1 + 100 \alpha_2 + 30 \alpha_3 = 244, \\ 100 \alpha_1 + 30 \alpha_2 + 10 \alpha_3 = 76, \\ 30 \alpha_1 + 10 \alpha_2 + 4 \alpha_3 = 27. \end{cases}$$

cuja solução é

$$\alpha_1 = \frac{\begin{vmatrix} 244 & 100 & 30 \\ 76 & 30 & 10 \\ 27 & 10 & 4 \end{vmatrix}}{\begin{vmatrix} 354 & 100 & 30 \\ 100 & 30 & 10 \\ 30 & 10 & 4 \end{vmatrix}} = \frac{-20}{80} = -0,25, \quad \alpha_2 = \frac{\begin{vmatrix} 354 & 244 & 30 \\ 100 & 76 & 10 \\ 30 & 27 & 4 \end{vmatrix}}{\begin{vmatrix} 354 & 100 & 30 \\ 100 & 30 & 10 \\ 30 & 10 & 4 \end{vmatrix}} = \frac{236}{80} = 2,95 \text{ e}$$

$$\alpha_3 = \frac{\begin{vmatrix} 354 & 100 & 244 \\ 100 & 30 & 76 \\ 30 & 10 & 27 \end{vmatrix}}{\begin{vmatrix} 354 & 100 & 30 \\ 100 & 30 & 10 \\ 30 & 10 & 4 \end{vmatrix}} = \frac{100}{80} = 1,25.$$

Estes resultados geram a função de ajuste  $g(x) = -0,25x^2 + 2,95x + 1,25$ . A figura a seguir ilustra graficamente o ajuste



Neste caso, o ajuste produz  $R$  igual a

$$\begin{aligned}
 R &= \sum_{i=1}^4 (f(x_i) - g(x_i))^2 = (f(x_1) - g(x_1))^2 + (f(x_2) - g(x_2))^2 + (f(x_3) - g(x_3))^2 + (f(x_4) - g(x_4))^2 \\
 &= (4 - 3,95)^2 + (6 - 6,15)^2 + (8 - 7,85)^2 + (9 - 9,05)^2 \\
 &= 0,0025 + 0,0225 + 0,0225 + 0,0025 \\
 &= 0,05
 \end{aligned}$$

Observando os valores de  $R$  encontrados para os ajustes dos exemplos 5.1, 5.2 e 5.3, pode-se concluir que o melhor ajuste corresponde à função quadrática  $g(x) = -0,25x^2 + 2,95x + 1,25$ .

## 5.2 Ajuste Discreto Não Linear por Quadrados Mínimos

Muitas vezes os dados tabelados

$x$	$f(x)$
$x_1$	$y_1$
$x_2$	$y_2$
$\vdots$	$\vdots$
$x_n$	$y_n$

$y_1 = f(x_1)$ ,  $y_2 = f(x_2)$ , ...,  $y_n = f(x_n)$ , devem ser ajustados seguindo funções  $g(x)$  que não são combinações lineares de outras. Neste caso, com frequência, é possível realizar uma “linearização” dos dados, de maneira que o ajuste pode ser reduzido ao caso linear.

**Exemplo 5.4:** Considere a tabela do exemplo 5.1:

$x$	$f(x)$
1	4
2	6
3	8
4	9

Agora, deseja-se ajustar a esses dados uma função da forma  $g(x) = \beta_2 \exp(\beta_1 x)$ , sendo  $\beta_1$  e  $\beta_2$  coeficientes a determinar. Observe que  $\ln(g(x)) = \ln(\beta_2 \exp(\beta_1 x)) = \ln(\beta_2) + \beta_1 x$ . Assim, pode-se ajustar uma reta  $\alpha_1 x + \alpha_2$  à função  $\ln(f(x))$ , e identificando os coeficientes, ter-se-á que  $\alpha_1 = \beta_1$  e  $\alpha_2 = \ln(\beta_2)$ , ou seja,  $\beta_1 = \alpha_1$  e  $\beta_2 = \exp(\alpha_2)$ . Então procede-se ao ajuste de uma reta  $\alpha_1 x + \alpha_2$  aos dados

$x$	$y = \ln(f(x))$
1	$\ln(4) = 1,386294$
2	$\ln(6) = 1,791759$
3	$\ln(8) = 2,079442$
4	$\ln(9) = 2,197225$

Aproveitam-se os valores para  $\sum_{i=1}^4 x_i^2$  e  $\sum_{i=1}^4 x_i$  fornecidos no exemplo 5.2 e refazem-se os somatórios a seguir,

$$\sum_{i=1}^4 x_i y_i = x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4 = 1 \cdot \ln(4) + 2 \cdot \ln(6) + 3 \cdot \ln(8) + 4 \cdot \ln(9) = 19,997036$$

$$\sum_{i=1}^4 y_i = y_1 + y_2 + y_3 + y_4 = \ln(4) + \ln(6) + \ln(8) + \ln(9) = 7,454720,$$

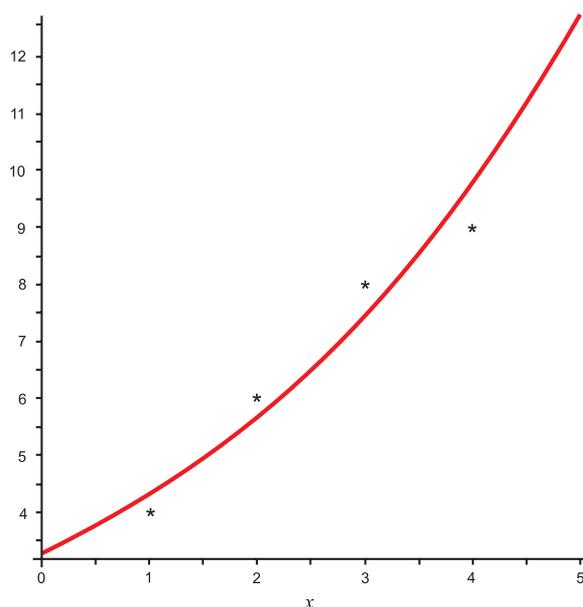
obtendo as equações normais

$$\begin{cases} 30 \alpha_1 + 10 \alpha_2 = 19,997036 \\ 10 \alpha_1 + 4 \alpha_2 = 7,454720, \end{cases}$$

cuja solução é

$$\alpha_1 = \frac{\begin{vmatrix} 19,997036 & 10 \\ 7,454720 & 4 \end{vmatrix}}{\begin{vmatrix} 30 & 10 \\ 10 & 4 \end{vmatrix}} = 0,272047 \quad \text{e} \quad \alpha_2 = \frac{7,454720 - 10 \cdot 0,272047}{4} = 1,183562.$$

Assim,  $\beta_1 = \alpha_1 = 0,272047$  e  $\beta_2 = \exp(\alpha_2) = \exp(1,183562) = 3,265986$ . Portanto, a função de ajuste é  $g(x) = 3,265986 \cdot \exp(0,272047 x)$ . A figura a seguir ilustra o ajuste



Agora, calcula-se o valor de  $R$ :

$$\begin{aligned} R &= \sum_{i=1}^4 (f(x_i) - g(x_i))^2 = (f(x_1) - g(x_1))^2 + (f(x_2) - g(x_2))^2 + (f(x_3) - g(x_3))^2 + (f(x_4) - g(x_4))^2 \\ &= (4 - 4,287094)^2 + (6 - 5,627450)^2 + (8 - 7,386866)^2 + (9 - 9,696363)^2 \\ &= 0,082423 + 0,138794 + 0,375933 + 0,484922 \\ &= 1,082072 \end{aligned}$$

A seguir, mostram-se algumas funções  $g(x)$  que devem ajustar os dados  $(x, f(x))$ , junto com o respectivo processo de linearização e o cálculo dos coeficientes

Função $g(x)$	“Linearização”	Ajuste e coeficientes
$g(x) = \beta_2 \cdot \exp(\beta_1 x)$	$(x, \ln(f(x)))$	$\alpha_1 x + \alpha_2, \beta_1 = \alpha_1, \beta_2 = \exp(\alpha_2)$
$g(x) = \beta_2 \cdot (\beta_1)^x$	$(x, \ln(f(x)))$	$\alpha_1 x + \alpha_2, \beta_1 = \exp(\alpha_1), \beta_2 = \exp(\alpha_2)$
$g(x) = \beta_2 \cdot x^{\beta_1}$	$(x, \ln(f(x)))$	$\alpha_1 \ln(x) + \alpha_2, \beta_1 = \alpha_1, \beta_2 = \exp(\alpha_2)$
$g(x) = \frac{1}{\beta_1 x + \beta_2}$	$\left(x, \frac{1}{f(x)}\right)$	$\alpha_1 x + \alpha_2, \beta_1 = \alpha_1, \beta_2 = \alpha_2$
$g(x) = \ln(\beta_1 x + \beta_2)$	$(x, \exp(f(x)))$	$\alpha_1 x + \alpha_2, \beta_1 = \alpha_1, \beta_2 = \alpha_2$

**Exemplo 5.5:** Considere a tabela do exemplo 5.1:

$x$	$f(x)$
1	4
2	6
3	8
4	9

Deseja-se ajustar a esses dados uma função da forma  $g(x) = \beta_2 \cdot x^{\beta_1}$ , sendo  $\beta_1$  e  $\beta_2$  coeficientes a determinar. Observe que  $\ln(g(x)) = \beta_1 \ln(x) + \ln(\beta_2)$ . Assim, pode-se ajustar uma função  $\alpha_1 \ln(x) + \alpha_2$  à função  $\ln(f(x))$ , e comparando os coeficientes tem-se que  $\alpha_1 = \beta_1$  e  $\alpha_2 = \ln(\beta_2)$ , ou seja,  $\beta_1 = \alpha_1$  e  $\beta_2 = \exp(\alpha_2)$ . Então procede-se ao ajuste da função  $\alpha_1 \ln(x) + \alpha_2$  aos dados

$x$	$y = \ln(f(x))$
1	$\ln(4) = 1,386294$
2	$\ln(6) = 1,791759$
3	$\ln(8) = 2,079442$
4	$\ln(9) = 2,197225$

Agora,  $m = 2$ ,  $g_1(x) = \ln(x)$  e  $g_2(x) = 1$ . Lembre que as equações normais do ajuste estão dadas por

$$\begin{cases} \left[ \sum_{i=1}^n g_1(x_i) g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_1(x_i) g_2(x_i) \right] \alpha_2 = \sum_{i=1}^n y_i g_1(x_i) \\ \left[ \sum_{i=1}^n g_2(x_i) g_1(x_i) \right] \alpha_1 + \left[ \sum_{i=1}^n g_2(x_i) g_2(x_i) \right] \alpha_2 = \sum_{i=1}^n y_i g_2(x_i) \end{cases}$$

que produz o sistema

$$\begin{cases} \left[ \sum_{i=1}^n [\ln(x_i)]^2 \right] \alpha_1 + \left[ \sum_{i=1}^n \ln(x_i) \right] \alpha_2 = \sum_{i=1}^n y_i \ln(x_i) \\ \left[ \sum_{i=1}^n \ln(x_i) \right] \alpha_1 + n \alpha_2 = \sum_{i=1}^n y_i \end{cases}$$

Realizam-se os cálculos preliminares

$$\begin{aligned} \sum_{i=1}^4 [\ln(x_i)]^2 &= [\ln(x_1)]^2 + [\ln(x_2)]^2 + [\ln(x_3)]^2 + [\ln(x_4)]^2 \\ &= [\ln(1)]^2 + [\ln(2)]^2 + [\ln(3)]^2 + [\ln(4)]^2 = 3,609214, \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^4 \ln(x_i) &= \ln(x_1) + \ln(x_2) + \ln(x_3) + \ln(x_4) \\ &= \ln(1) + \ln(2) + \ln(3) + \ln(4) = 3,178054, \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^n y_i \ln(x_i) &= y_1 \ln(x_1) + y_2 \ln(x_2) + y_3 \ln(x_3) + y_4 \ln(x_4) \\ &= \ln(4) \cdot \ln(1) + \ln(6) \cdot \ln(2) + \ln(8) \cdot \ln(3) + \ln(9) \cdot \ln(4) = 6,572453, \end{aligned}$$

$$\sum_{i=1}^n y_i = y_1 + y_2 + y_3 + y_4 = \ln(4) + \ln(6) + \ln(8) + \ln(9) = 7,454720,$$

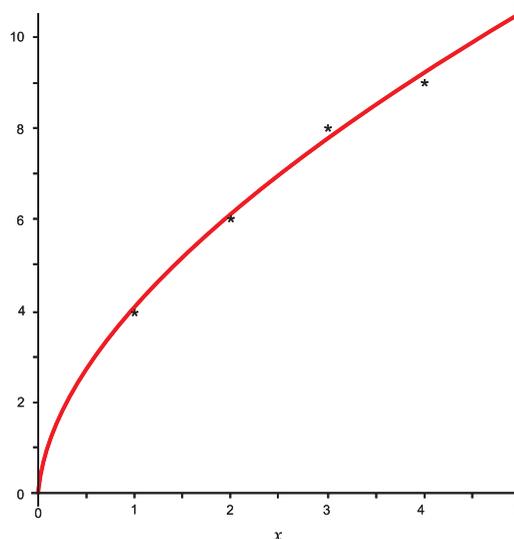
que produzem as equações normais

$$\begin{cases} 3,609214 \alpha_1 + 3,178054 \alpha_2 = 6,572453, \\ 3,178054 \alpha_1 + 4 \alpha_2 = 7,454720, \end{cases}$$

cuja solução é

$$\alpha_1 = 0,599127 \text{ e } \alpha_2 = 1,387666.$$

Mas  $\beta_1 = \alpha_1 = 0,599127$  e  $\beta_2 = \exp(\alpha_2) = 4,005489$ . Estes resultados geram a função de ajuste  $g(x) = 4,005489 \cdot x^{0,599127}$ . A figura a seguir ilustra graficamente o ajuste



Neste caso, o ajuste produz  $R$  igual a

$$\begin{aligned} R &= \sum_{i=1}^4 (f(x_i) - g(x_i))^2 = (f(x_1) - g(x_1))^2 + (f(x_2) - g(x_2))^2 + (f(x_3) - g(x_3))^2 + (f(x_4) - g(x_4))^2 \\ &= (4 - 4,005489)^2 + (6 - 6,067513)^2 + (8 - 7,735915)^2 + (9 - 9,191065)^2 \\ &= 0,000030 + 0,004558 + 0,069741 + 0,036506 \\ &= 0,110835 \end{aligned}$$

podendo-se concluir este ajuste é ainda melhor que o ajuste por uma função quadrática.

### 5.3 Ajuste Contínuo Linear por Quadrados Mínimos

Seja  $f(x)$  uma função integrável sobre um intervalo  $[a, b]$  e sejam escolhidas certas funções  $g_1(x)$ ,  $g_2(x)$ ,  $\dots$ ,  $g_m(x)$ , ( $m \in \mathbf{N}$ ) integráveis no mesmo intervalo. O objetivo do ajuste contínuo linear por quadrados mínimos é obter coeficientes  $\alpha_1$ ,  $\alpha_2$ ,  $\dots$ ,  $\alpha_m$  tais que a função

$$g(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \dots + \alpha_m g_m(x)$$

satisfaz a condição que a integral

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

tem o menor valor possível.

De maneira análoga à seção 5.1, define-se a função a função real de várias variáveis

$$F(\alpha_1, \alpha_2, \dots, \alpha_m) = \int_a^b (f(x) - g(x))^2 dx = \int_a^b (f(x) - \alpha_1 g_1(x) - \alpha_2 g_2(x) - \dots - \alpha_m g_m(x))^2 dx,$$

e procede-se a minimizá-la mediante o critério da primeira derivada, fazendo

$$\frac{\partial}{\partial \alpha_j} F(\alpha_1, \alpha_2, \dots, \alpha_m) = 0, \quad j = 1, 2, \dots, m.$$

Para cada  $j$ , tem-se que

$$\frac{\partial}{\partial \alpha_j} F(\alpha_1, \alpha_2, \dots, \alpha_m) = -2 \int_a^b (f(x) - \alpha_1 g_1(x) - \alpha_2 g_2(x) - \dots - \alpha_m g_m(x)) \cdot (g_j(x)) dx,$$

que, após impor a condição da derivada igual a zero, produz

$$\int_a^b (f(x) - \alpha_1 g_1(x) - \alpha_2 g_2(x) - \dots - \alpha_m g_m(x)) \cdot (g_j(x)) dx = 0, \text{ para cada } j = 1, 2, \dots, m.$$

A partir da igualdade anterior, forma-se o sistema de equações lineares de  $m$  equações com  $m$  incógnitas:

$$\begin{cases} \left[ \int_a^b g_1(x)g_1(x) dx \right] \alpha_1 + \left[ \int_a^b g_1(x)g_2(x) dx \right] \alpha_2 + \cdots + \left[ \int_a^b g_1(x)g_m(x) dx \right] \alpha_m = \int_a^b f(x)g_1(x) dx \\ \left[ \int_a^b g_2(x)g_1(x) dx \right] \alpha_1 + \left[ \int_a^b g_2(x)g_2(x) dx \right] \alpha_2 + \cdots + \left[ \int_a^b g_2(x)g_m(x) dx \right] \alpha_m = \int_a^b f(x)g_2(x) dx \\ \vdots \\ \left[ \int_a^b g_m(x)g_1(x) dx \right] \alpha_1 + \left[ \int_a^b g_m(x)g_2(x) dx \right] \alpha_2 + \cdots + \left[ \int_a^b g_m(x)g_m(x) dx \right] \alpha_m = \int_a^b f(x)g_m(x) dx \end{cases}$$

denominadas as equações normais do ajuste, que podem ser escritas de maneira matricial como

$$\begin{pmatrix} \int_a^b g_1(x)g_1(x) dx & \int_a^b g_1(x)g_2(x) dx & \cdots & \int_a^b g_1(x)g_m(x) dx \\ \int_a^b g_2(x)g_1(x) dx & \int_a^b g_2(x)g_2(x) dx & \cdots & \int_a^b g_2(x)g_m(x) dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_a^b g_m(x)g_1(x) dx & \int_a^b g_m(x)g_2(x) dx & \cdots & \int_a^b g_m(x)g_m(x) dx \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \int_a^b f(x)g_1(x) dx \\ \int_a^b f(x)g_2(x) dx \\ \vdots \\ \int_a^b f(x)g_m(x) dx \end{pmatrix}.$$

Mediante a resolução deste sistema podem ser calculados os coeficientes requeridos.

Uma maneira de quantificar a precisão do ajuste é calcular a integral

$$R = \int_a^b (f(x_i) - g(x_i))^2 dx$$

Como no caso discreto, se existirem várias escolhas de funções  $g(x)$ , o melhor ajuste estará dado por aquela função que tenha o menor valor de  $R$ .

**Exemplo 5.6:** Deseja-se ajustar a função  $f(x) = x^3$  mediante uma função quadrática  $g(x) = \alpha_1 x^2 + \alpha_2 x + \alpha_3$ , no intervalo  $[0,1]$ . Tem-se que  $m = 3$ ,  $g_1(x) = x^2$ ,  $g_2(x) = x$  e  $g_3(x) = 1$ . Então o sistema de equações normais pode ser escrito como

$$\begin{cases} \left[ \int_a^b g_1(x)g_1(x) dx \right] \alpha_1 + \left[ \int_a^b g_1(x)g_2(x) dx \right] \alpha_2 + \left[ \int_a^b g_1(x)g_3(x) dx \right] \alpha_3 = \int_a^b f(x)g_1(x) dx \\ \left[ \int_a^b g_2(x)g_1(x) dx \right] \alpha_1 + \left[ \int_a^b g_2(x)g_2(x) dx \right] \alpha_2 + \left[ \int_a^b g_2(x)g_3(x) dx \right] \alpha_3 = \int_a^b f(x)g_2(x) dx \\ \left[ \int_a^b g_3(x)g_1(x) dx \right] \alpha_1 + \left[ \int_a^b g_3(x)g_2(x) dx \right] \alpha_2 + \left[ \int_a^b g_3(x)g_3(x) dx \right] \alpha_m = \int_a^b f(x)g_3(x) dx \end{cases}$$

Agora, calculam-se as integrais necessárias

$$\int_0^1 g_1(x)g_1(x) dx = \int_0^1 x^4 dx = \frac{1}{5}x^5 \Big|_{x=0}^{x=1} = \frac{1}{5}, \quad \int_0^1 g_1(x)g_2(x) dx = \int_0^1 x^3 dx = \frac{1}{4}x^4 \Big|_{x=0}^{x=1} = \frac{1}{4},$$

$$\int_0^1 g_1(x)g_3(x) dx = \int_0^1 x^2 dx = \frac{1}{3}x^3 \Big|_{x=0}^{x=1} = \frac{1}{3}, \quad \int_0^1 g_2(x)g_2(x) dx = \int_0^1 x^2 dx = \frac{1}{3}x^3 \Big|_{x=0}^{x=1} = \frac{1}{3},$$

$$\int_0^1 g_2(x)g_3(x) dx = \int_0^1 x dx = \frac{1}{2}x^2 \Big|_{x=0}^{x=1} = \frac{1}{2}, \quad \int_0^1 g_3(x)g_3(x) dx = \int_0^1 dx = x \Big|_{x=0}^{x=1} = 1,$$

$$\int_0^1 f(x)g_1(x) dx = \int_0^1 x^5 dx = \frac{1}{6}x^6 \Big|_{x=0}^{x=1} = \frac{1}{6}, \quad \int_0^1 f(x)g_2(x) dx = \int_0^1 x^4 dx = \frac{1}{5}x^5 \Big|_{x=0}^{x=1} = \frac{1}{5},$$

$$\int_0^1 f(x)g_3(x) dx = \int_0^1 x^3 dx = \frac{1}{4}x^4 \Big|_{x=0}^{x=1} = \frac{1}{4},$$

e forma-se o sistema

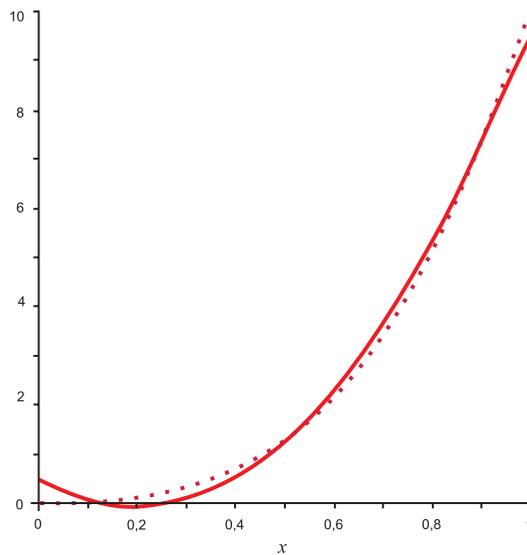
$$\begin{cases} \frac{1}{5} \alpha_1 + \frac{1}{4} \alpha_2 + \frac{1}{3} \alpha_3 = \frac{1}{6}, \\ \frac{1}{4} \alpha_1 + \frac{1}{3} \alpha_2 + \frac{1}{2} \alpha_3 = \frac{1}{5}, \\ \frac{1}{3} \alpha_1 + \frac{1}{2} \alpha_2 + \alpha_3 = \frac{1}{4}, \end{cases}$$

cuja solução é

$$\alpha_1 = \frac{\begin{vmatrix} \frac{1}{5} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 1 \end{vmatrix}}{\begin{vmatrix} \frac{1}{5} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 1 \end{vmatrix}} = \frac{1440}{2160} = \frac{3}{2}, \quad \alpha_2 = \frac{\begin{vmatrix} \frac{1}{5} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{4} & 1 \end{vmatrix}}{\begin{vmatrix} \frac{1}{5} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 1 \end{vmatrix}} = \frac{-3600}{2160} = -\frac{3}{5},$$

$$\alpha_3 = \frac{\begin{vmatrix} \frac{1}{5} & \frac{1}{4} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} \end{vmatrix}}{\begin{vmatrix} \frac{1}{5} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 1 \end{vmatrix}} = \frac{43200}{2160} = \frac{1}{20}.$$

Estes coeficientes geram a função de ajuste  $g(x) = \frac{3}{2}x^2 - \frac{3}{5}x + \frac{1}{20}$ . A figura a seguir ilustra graficamente o ajuste



A linha pontilhada corresponde ao gráfico da função  $f(x) = x^3$  e a linha contínua, à função de ajuste

$g(x) = \frac{3}{2}x^2 - \frac{3}{5}x + \frac{1}{20}$ . O valor de  $R$ , neste exemplo, é igual a

$$\begin{aligned}
 R &= \int_0^1 (f(x) - g(x))^2 dx = \int_0^1 \left(x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}\right)^2 dx \\
 &= \left(\frac{1}{7}x^7 - \frac{1}{2}x^6 + \frac{69}{100}x^5 - \frac{19}{40}x^4 + \frac{17}{100}x^3 - \frac{3}{100}x^2 + \frac{1}{400}x\right) \Bigg|_{x=0}^{x=1} \\
 &= \frac{1}{2800} = 0,000357.
 \end{aligned}$$

Integração Numérica

6  
Capítulo

Com frequência, é necessário avaliar a integral definida de uma função cuja anti-derivada não tem uma forma explícita, não é simples de obter ou não existe. Uma primeira abordagem deste problema já foi realizada ao tratar as somas de Riemann. Porém, aqui, o método básico utilizado para aproximar a integral definida  $\int_a^b f(x)dx$ , denominado quadratura numérica, consiste basicamente em fazer a aproximação

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

para alguns coeficientes  $a_i$  e alguns  $x_i \in [a, b]$ .

Um bom começo para o estudo dos métodos de quadratura pode ser dado utilizando os polinômios de interpolação de Lagrange. Procede-se como segue:

Sejam  $f(x)$  uma função integrável no intervalo  $[a, b]$  e  $\{x_1, x_2, \dots, x_n\}$  uma partição do intervalo  $[a, b]$  tal que  $a = x_1 < x_2 < \dots < x_n = b$ . Sabe-se que o polinômio interpolador de Newton para  $f(x)$  é

$$p(x) = f(x_1) \ell_1(x) + f(x_2) \ell_2(x) + \dots + f(x_n) \ell_n(x),$$

onde os  $\ell_i(x)$  são os polinômios fundamentais de Lagrange.

Assim,

$$f(x) = f(x_1) \ell_1(x) + f(x_2) \ell_2(x) + \dots + f(x_n) \ell_n(x) + (x - x_1) \dots (x - x_n) \frac{f^{(n)}(\xi_x)}{n!},$$

para algum  $\xi_x \in [a, b]$  (que depende de  $x$ ). Então integra-se a última igualdade

$$\begin{aligned} \int_a^b f(x) dx &= \int_a^b \left[ f(x_1) \ell_1(x) + f(x_2) \ell_2(x) + \dots + f(x_n) \ell_n(x) + (x - x_1) \dots (x - x_n) \frac{f^{(n)}(\xi_x)}{n!} \right] dx \\ &= f(x_1) \int_a^b \ell_1(x) dx + f(x_2) \int_a^b \ell_2(x) dx + \dots + f(x_n) \int_a^b \ell_n(x) dx + \int_a^b (x - x_1) \dots (x - x_n) \frac{f^{(n)}(\xi_x)}{n!} dx \end{aligned}$$

Portanto, pode ser feita a aproximação

$$\int_a^b f(x) dx \approx f(x_1) \int_a^b \ell_1(x) dx + f(x_2) \int_a^b \ell_2(x) dx + \dots + f(x_n) \int_a^b \ell_n(x) dx$$

sendo o erro

$$\int_a^b (x - x_1) \dots (x - x_n) \frac{f^{(n)}(\xi_x)}{n!} dx.$$

No que segue, deduzem-se e mostram-se aproximações da integral  $\int_a^b f(x)dx$  utilizando a abordagem que se acaba de detalhar.

## 6.1 Regra do Trapézio

Aplica-se, agora, a metodologia descrita no início do capítulo no caso em que  $n = 2$ . O polinômio interpolador de Lagrange é

$$p(x) = \frac{x - x_2}{x_1 - x_2} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2),$$

onde  $x_1 = a$  e  $x_2 = b$ . Logo, a aproximação de  $\int_a^b f(x)dx$  é

$$\int_a^b \left( \frac{x-b}{a-b} f(a) + \frac{x-a}{b-a} f(b) \right) dx = \left( f(a) \frac{(x-b)^2}{2(a-b)} + f(b) \frac{(x-a)^2}{2(b-a)} \right) \Big|_{x=a}^{x=b}$$

$$= f(b) \frac{b-a}{2} - f(a) \frac{a-b}{2} = f(a) \frac{b-a}{2} + f(b) \frac{b-a}{2}.$$

Quanto ao erro, pode-se afirmar que este é

$$\frac{1}{2} \int_a^b f''(\xi_x)(x-a)(x-b) dx = \frac{1}{2} f''(\xi) \int_a^b (x-a)(x-b) dx$$

$$= \frac{1}{2} f''(\xi)(b-a)^3 \int_0^1 u(u-1) du$$

$$= -\frac{1}{12} f''(\xi)(b-a)^3$$

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

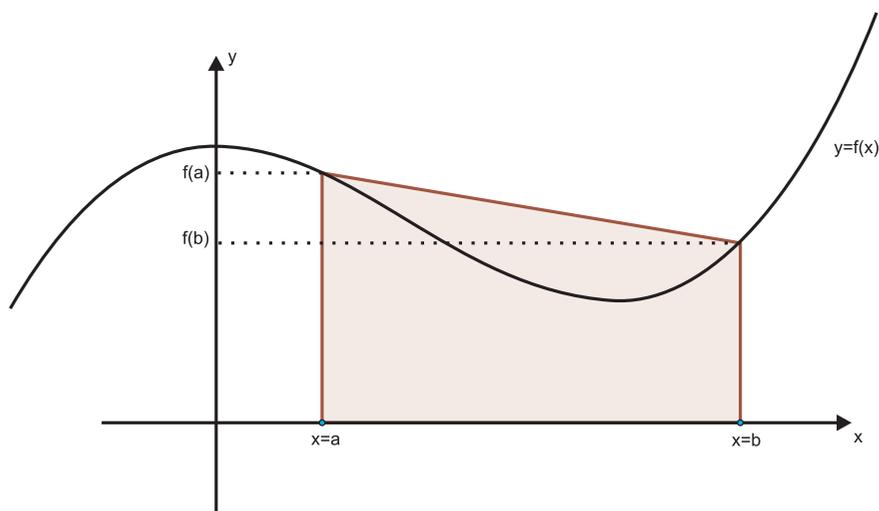
Para efeitos de generalização, utiliza-se a notação  $a = x_1$ ,  $b = x_2$ ,  $f(a) = f(x_1) = f_1$ ,  $f(b) = f(x_2) = f_2$  e  $b - a = x_2 - x_1 = h$ . Assim, este processo de aproximação pode ser estabelecido como

$$\int_a^b f(x) dx \approx \frac{h}{2} \cdot (f_1 + f_2)$$

sendo dada uma estimativa do erro de aproximação,  $E(f)$ , por

$$|E(f)| \leq \frac{h^3}{12} M_2,$$

onde  $M_2 = \max\{|f''(x)| : x \in [a, b]\}$ . Esta aproximação é denominada regra do trapézio, pois geometricamente, no caso que  $f(x)$  seja positiva no intervalo  $[a, b]$ , o valor da integral  $\int_a^b f(x) dx$  é aproximado mediante o valor numérico da área de um trapézio como é mostrado na figura a seguir



Observe que se  $f''(x) = 0, \forall x \in [a, b]$ , ou seja, a segunda derivada de  $f(x)$  é identicamente nula sobre o intervalo  $[a, b]$ , então o erro cometido pela regra do trapézio é também nulo. Assim, a regra do trapézio fornece um resultado exato quando  $f(x)$  é um polinômio de grau um ou uma função constante.

## 6.2 Regra de Simpson

Nesta seção, desenvolve-se a metodologia descrita no início do capítulo no caso em que  $n = 3$ . Logo, o polinômio interpolador de Lagrange é

$$p(x) = \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} f(x_1) + \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} f(x_2) + \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)} f(x_3),$$

onde  $a = x_1 < x_2 < x_3 = b$ . Para simplificar, considere que os pontos  $x_1, x_2$  e  $x_3$  são igualmente espaçados, isto é,  $x_2 - x_1 = x_3 - x_2 = h$ . Denote, também,  $f(x_1) = f_1, f(x_2) = f_2$  e  $f(x_3) = f_3$ . Logo, a aproximação de  $\int_a^b f(x) dx$  é

$$\begin{aligned} & \frac{1}{h^2} \left( \frac{f_1}{2} \cdot \int_a^b (x-x_2)(x-x_3) dx - f_2 \cdot \int_a^b (x-x_1)(x-x_3) dx + \frac{f_3}{2} \cdot \int_a^b (x-x_1)(x-x_2) dx \right) \\ &= \frac{1}{h^2} \left( \frac{f_1}{2} \cdot \int_a^{a+2h} (x-a-h)(x-a-2h) dx - f_2 \cdot \int_a^{a+2h} (x-a)(x-a-2h) dx + \frac{f_3}{2} \cdot \int_a^{a+2h} (x-a)(x-a-h) dx \right) \\ &= \frac{1}{h^2} \left( \frac{f_1}{2} \cdot \int_0^{2h} (u-h)(u-2h) du - f_2 \cdot \int_0^{2h} u \cdot (u-2h) du + \frac{f_3}{2} \cdot \int_0^{2h} u \cdot (u-h) du \right) \\ &= \frac{1}{h^2} \left( \frac{f_1}{2} \cdot \left( \frac{1}{3} u^3 - \frac{3}{2} h u^2 + 2h^2 u \right) \Big|_{u=0}^{u=2h} - f_2 \cdot \left( \frac{1}{3} u^3 - 2h u^2 \right) \Big|_{u=0}^{u=2h} + \frac{f_3}{2} \cdot \left( \frac{1}{3} u^3 - \frac{1}{2} h u^2 \right) \Big|_{u=0}^{u=2h} \right) \\ &= \frac{1}{h^2} \left( \frac{f_1}{2} \cdot \frac{2}{3} h^3 + f_2 \cdot \frac{4}{3} h^3 + \frac{f_3}{2} \cdot \frac{2}{3} h^3 \right) = \frac{h}{3} \cdot (f_1 + 4f_2 + f_3) \end{aligned}$$

e quanto ao erro, pode-se afirmar que este é  $\frac{1}{6} \int_a^b f'''(\xi_x)(x-a)(x-a-h)(x-a-2h) dx$ . Mas, mediante abordagens mais específicas obtém-se a expressão  $\frac{h^5}{90} f^{(4)}(\zeta)$ , para algum  $\zeta \in [a, b]$ , para o erro.

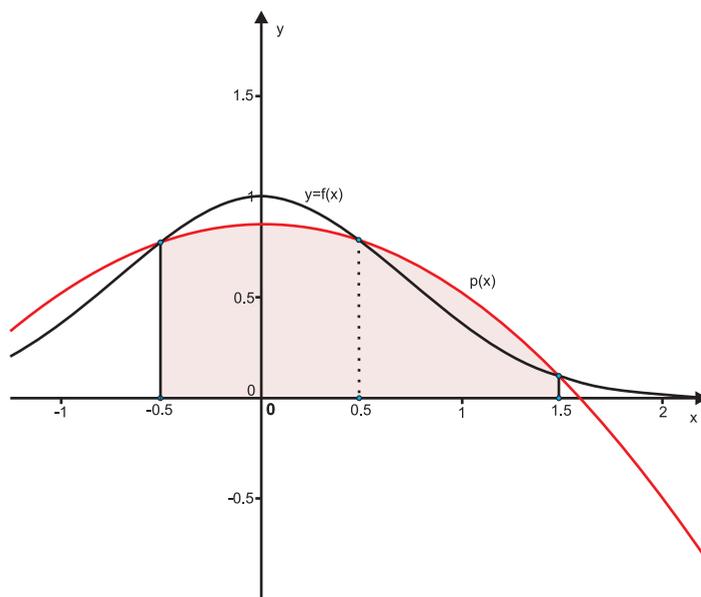
Assim, o atual processo de aproximação pode ser estabelecido como

$$\int_a^b f(x) dx \approx \frac{h}{3} \cdot (f_1 + 4f_2 + f_3)$$

sendo dada uma estimativa do erro de aproximação,  $E(f)$ , por

$$|E(f)| \leq \frac{h^5}{90} M_4,$$

onde  $M_4 = \max\{|f^{(4)}(x)| : x \in [a, b]\}$ . Esta aproximação é denominada regra de Simpson. Graficamente, no caso que  $f(x)$  é não negativa para todo  $x \in [a, b]$ , a regra de Simpson aproxima o valor da integral  $\int_a^b f(x) dx$  mediante a área limitada pela parte da parábola  $p(x)$ , compreendida no intervalo  $[a, b]$ ; por exemplo, considere o intervalo  $[-0,5, 0,5]$  que interpola os três pontos do gráfico de  $f(x)$  correspondentes a  $x_1 = -0,5, x_2 = 0,5$  e  $x_3 = 1,5$  como é mostrado na figura a seguir



Observe que se  $f^{(4)}(x) = 0, \forall x \in [a, b]$ , ou seja, a quarta derivada de  $f(x)$  é identicamente nula sobre o intervalo  $[a, b]$ , então o erro cometido pela regra de Simpson é também nulo. Assim, a regra de Simpson fornece um resultado exato quando  $f(x)$  é um polinômio de grau três ou menos.

**Exemplo 6.1:** Considere a função  $f(x)$  sobre o intervalo  $[0, 1]$ . A regra do trapézio dá a aproximação

$$\int_0^1 f(x) dx \approx \frac{1}{2} \cdot (f(0) + f(1)),$$

enquanto que a regra de Simpson fornece

$$\int_0^1 f(x) dx \approx \frac{1}{6} \cdot (f(0) + 4f(0,5) + f(1)).$$

Na tabela a seguir, mostram-se as aproximações obtidas mediante a regra do trapézio para algumas funções  $f(x)$  específicas, no intervalo  $[0, 1]$ , utilizando a aproximação descrito no início do exemplo. A linha vermelha indica o gráfico da função e a linha azul delimita o trapézio cuja área está sendo calculada para aproximar a integral:

Função $f(x)$	$2x^3 + 1$	$\sqrt{8x+1}$	$e^{-x}$
Valor exato	1,5	2,166667	0,632121
Regra do trapézio	2	2	0,683940
Gráfico	Gráfico 1*	Gráfico 2*	Gráfico 3*

\* Os gráficos estão representados a seguir.

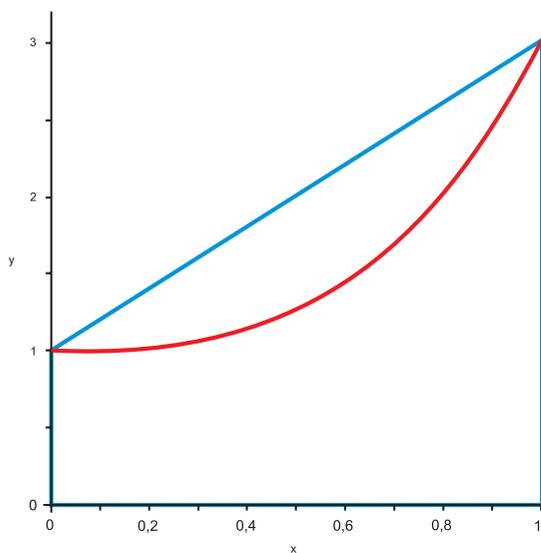


Gráfico 1

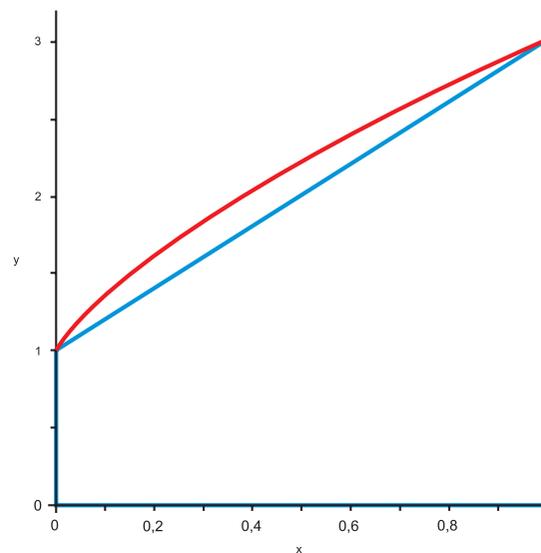


Gráfico 2

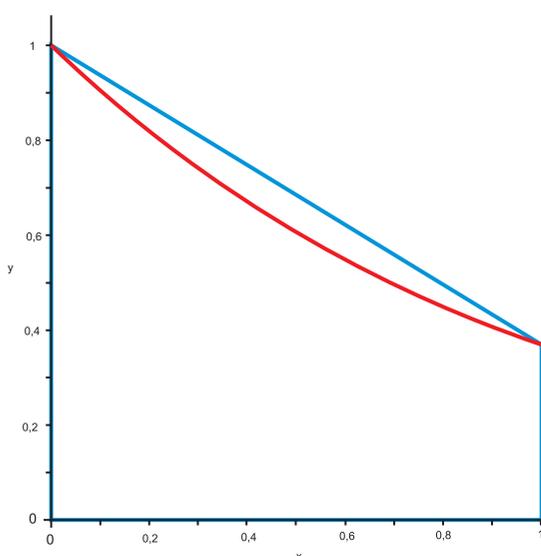


Gráfico 3

Agora, mostram-se as aproximações pela regra de Simpson. Agora, a linha vermelha indica o gráfico da função e a linha azul delimita a região abaixo da parábola compreendida no intervalo  $[0,1]$  cuja área está sendo calculada para aproximar a integral:

Função $f(x)$	$2x^3 + 1$	$\sqrt{8x+1}$	$e^{-x}$
Valor exato	1,5	2,166667	0,632121
Regra de Simpson	1,5	2,157379	0,632334
Gráfico	Gráfico 1*	Gráfico 2*	Gráfico 3*

\* Os gráficos estão representados a seguir.

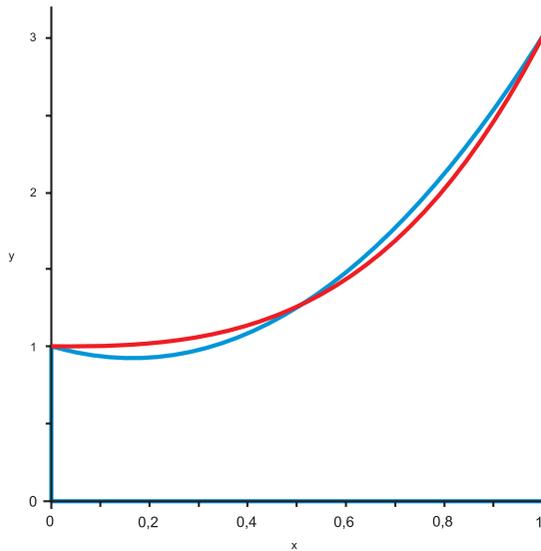


Gráfico 1

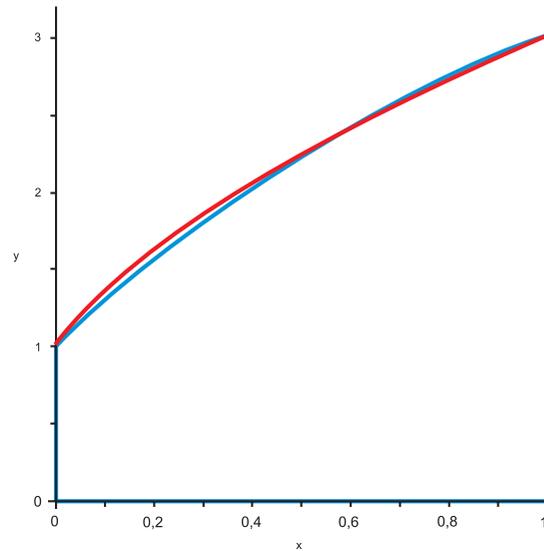


Gráfico 2

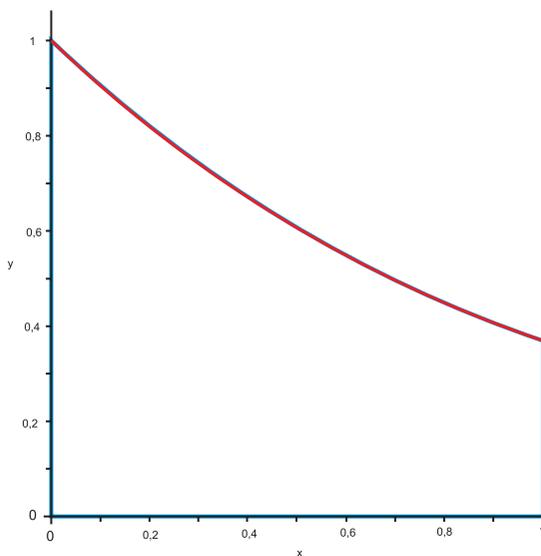


Gráfico 3

Comparando os resultados numéricos, pode-se observar a maior precisão da regra de Simpson.

As regras do trapézio e de Simpson são exemplos de uma classe maior de aproximações da integral  $\int_a^b f(x)dx$  denominada fórmulas fechadas de Newton-Cotes. Sem mostrar as provas, detalham-se a seguir algumas fórmulas de Newton-Cotes:

Número de pontos da partição	Aproximação	Estimativa do erro
$n = 2$ (trapézio)	$\frac{h}{2} \cdot (f_1 + f_2)$	$\frac{h^3}{12} M_2$
$n = 3$ (Simpson)	$\frac{h}{3} \cdot (f_1 + 4f_2 + f_3)$	$\frac{h^5}{90} M_4$
$n = 4$ (três oitavos)	$\frac{3h}{8} \cdot (f_1 + 3f_2 + 3f_3 + f_4)$	$\frac{3h^5}{80} M_4$
$n = 5$ (Boole)	$\frac{2h}{45} \cdot (7f_1 + 32f_2 + 12f_3 + 32f_4 + 7f_5)$	$\frac{8h^7}{945} M_6$

Aqui,  $h$  denota o espaçamento da partição  $a = x_1 < x_2 < \dots < x_n = b$  (igualmente espaçada),  $f_j = f(x_j)$  e  $M_j = \max\{|f^{(j)}(x)| \mid x \in [a, b]\}$ .

### 6.3 Técnicas Compostas de Integração Numérica

As fórmulas de Newton-Cotes, mostradas na seção anterior, são inadequadas quando o intervalo de integração é grande e/ou quando o integrando é uma função com natureza oscilatória. Nesta seção, descrevem-se técnicas em que são aplicadas as fórmulas de Newton-Cotes de ordem pequena (trapézio e Simpson) por partes numa partição com muitos pontos. Na prática, estas técnicas são utilizadas com mais frequência.

**Exemplo 6.2:** Para ilustrar o que foi descrito no parágrafo anterior, considere a função  $f(x) = x \exp(x^2)$  no intervalo  $[0,1]$ . Aplicando a regra do trapézio, consegue-se a aproximação

$$\int_0^1 x \exp(x^2) dx \approx \frac{1}{2} (f(0) + f(1)) = \frac{1}{2} e = 1,359141,$$

sendo o valor exato

$$\int_0^1 x \exp(x^2) dx = \frac{1}{2} \exp(x^2) \Big|_{x=0}^{x=1} = \frac{1}{2} (e - 1) = 0,859141;$$

o que mostra a pouca precisão da aproximação da integral.

Agora, subdivide-se o intervalo  $[0,1]$  em  $[0,1/2]$  e  $[1/2,1]$ , e aplica-se a regra do trapézio em cada subintervalo. Tem-se, então,

$$\int_0^{1/2} x \exp(x^2) dx \approx \frac{1}{4} (f(0) + f(1/2)) = \frac{1}{8} e^{1/4} = 0,160503,$$

$$\int_{1/2}^1 x \exp(x^2) dx \approx \frac{1}{4} (f(1/2) + f(1)) = \frac{1}{8} e^{1/4} + \frac{1}{4} e = 0,840074,$$

que dá

$$\int_0^1 x \exp(x^2) dx = \int_0^{1/2} x \exp(x^2) dx + \int_{1/2}^1 x \exp(x^2) dx \approx 0,160503 + 0,840074 = 1,000577,$$

resultado que melhora a precisão anterior. O processo de subdividir os intervalos pode ser estendido. Se o intervalo  $[0,1]$  é subdividido em, por exemplo, quatro subintervalos iguais, tem-se

$$\begin{aligned}
\int_0^1 x \exp(x^2) dx &= \int_0^{1/4} x \exp(x^2) dx + \int_{1/4}^{1/2} x \exp(x^2) dx + \int_{1/2}^{3/4} x \exp(x^2) dx + \int_{3/4}^1 x \exp(x^2) dx \\
&\approx \frac{1}{8}(f(0) + f(1/4)) + \frac{1}{8}(f(1/4) + f(1/2)) + \frac{1}{8}(f(1/2) + f(3/4)) + \frac{1}{8}(f(3/4) + f(1)) \\
&= \frac{1}{8}(f(0) + 2f(1/4) + 2f(1/2) + 2f(3/4) + f(1)) \\
&= \frac{1}{16}e^{1/16} + \frac{1}{8}e^{1/4} + \frac{3}{16}e^{9/16} + \frac{1}{8}e = 0,895892,
\end{aligned}$$

conseguindo-se ainda maior precisão.

### 6.3.1 Regra Composta do Trapézio

No caso geral, suponha uma função  $f(x)$  com derivada contínua até de segunda ordem no intervalo  $[a, b]$  e uma partição  $a = x_1 < x_2 < \dots < x_n = b$ , ( $n > 1$ ), igualmente espaçada com espaçamento  $h = \frac{b-a}{n-1}$ . Então a regra composta do trapézio para a partição é formada aplicando a regra do trapézio,  $(n-1)$  vezes, uma vez a cada subintervalo  $[x_1, x_2]$ , ...,  $[x_{n-1}, x_n]$ , e pode ser escrita como

$$\int_a^b f(x) dx = \frac{h}{2}(f_1 + 2(f_2 + \dots + f_{n-1}) + f_n),$$

tendo como estimativa de erro o valor

$$\frac{(n-1)h^3}{12}M_2,$$

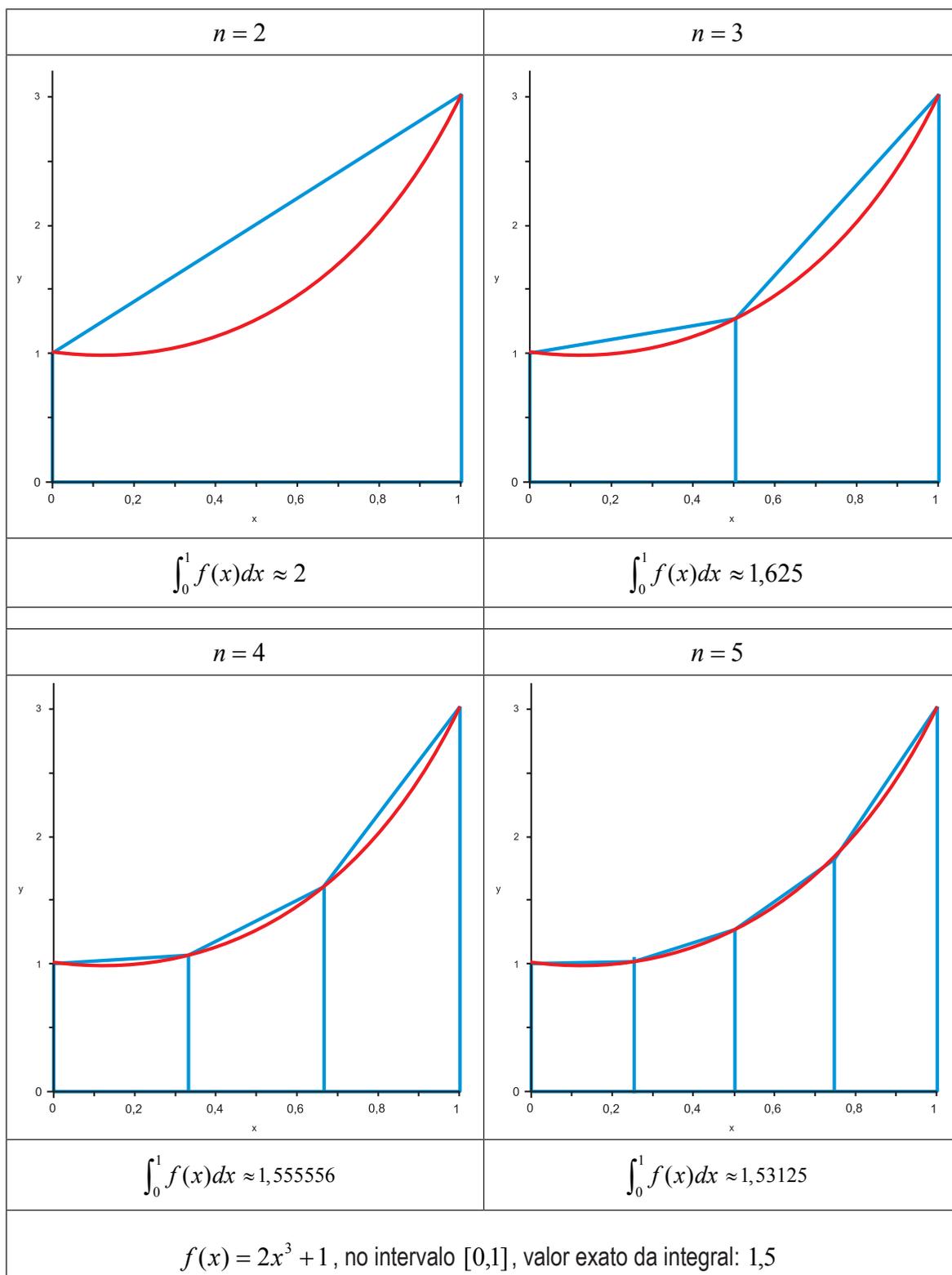
segundo a notação da seção anterior.

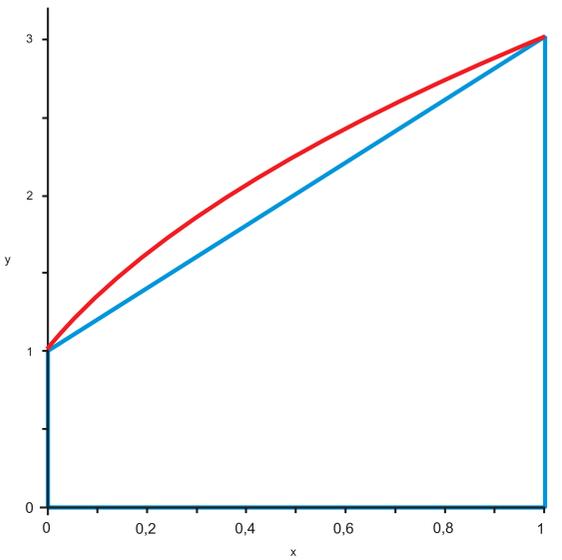
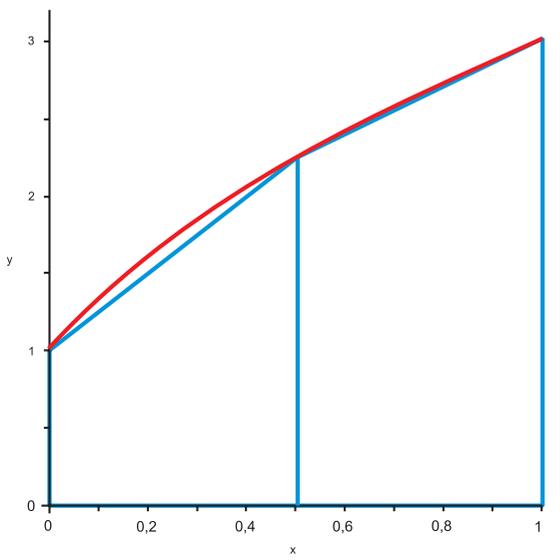
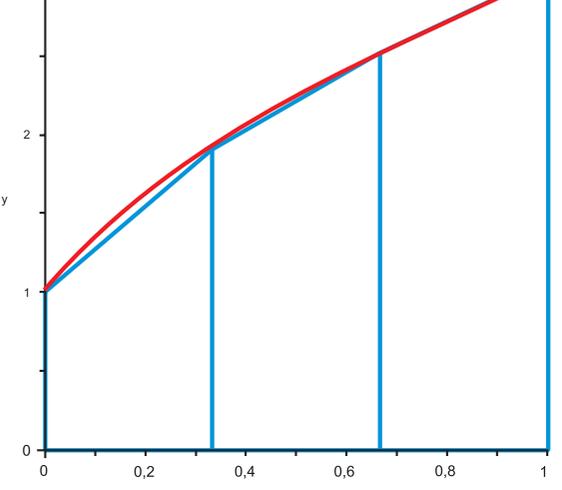
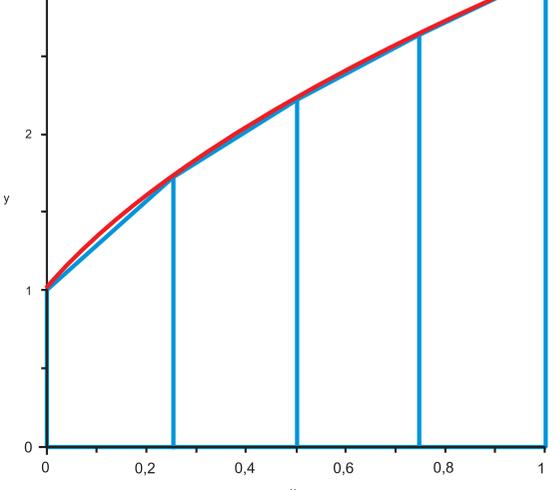
**Exemplo 6.3:** Aplica-se a regra composta do trapézio com  $n = 5$  pontos, à função  $f(x) = 2x^3 + 1$  sobre o intervalo  $[0,1]$ , para aproximar a integral  $\int_0^1 f(x) dx$  (observe que  $h = \frac{1}{4}$ ):

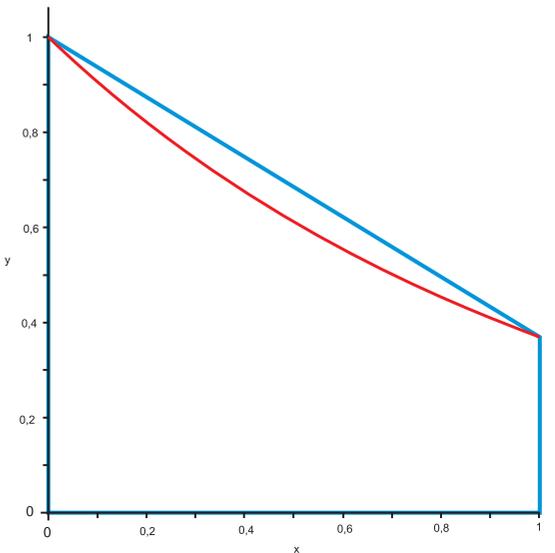
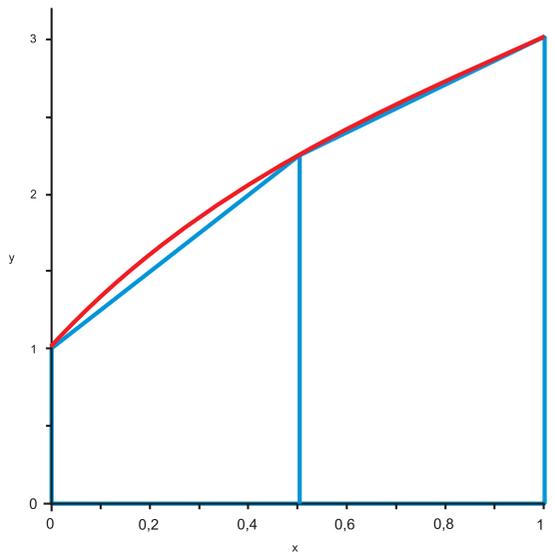
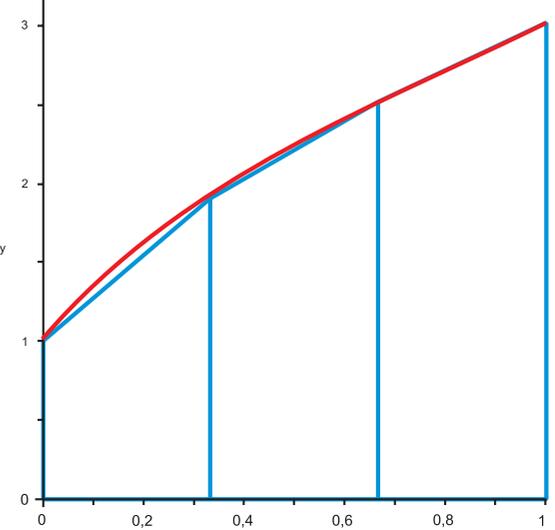
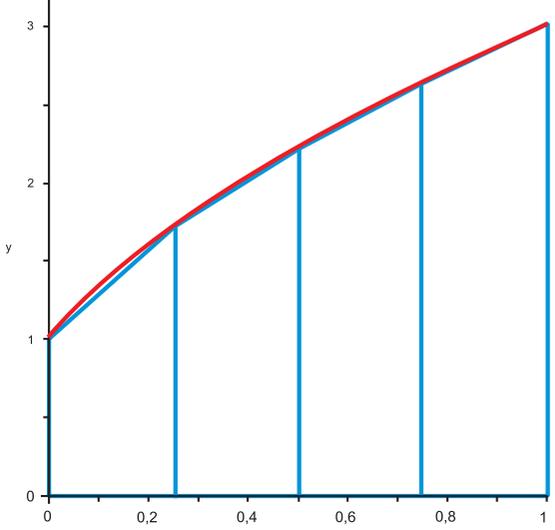
$$\int_0^1 (2x^3 + 1) dx \approx \frac{1}{8}(f_1 + 2(f_2 + f_3 + f_4) + f_5) = \frac{1}{8}\left(1 + 2\left(\frac{65}{64} + \frac{9}{8} + \frac{91}{64}\right) + 3\right) = \frac{49}{32} = 1,53125.$$

Lembre que o valor exato da integral é  $\int_0^1 (2x^3 + 1) dx = 1,5$ .

A seguir, nas seguintes tabelas, mostram-se diversas aproximações das integrais das funções indicadas mediante a regra composta do trapézio.



$n = 2$	$n = 3$
	
$\int_0^1 f(x)dx \approx 2$	$\int_0^1 f(x)dx \approx 2,118034$
$n = 4$	$n = 5$
	
$\int_0^1 f(x)dx \approx 2,143822$	$\int_0^1 f(x)dx \approx 2,153468$
$f(x) = \sqrt{8x+1}$ , no intervalo $[0,1]$ , valor exato da integral: 2,166667	

$n = 2$	$n = 3$
	
$\int_0^1 f(x)dx \approx 0,683940$	$\int_0^1 f(x)dx \approx 0,645235$
$n = 4$	$n = 5$
	
$\int_0^1 f(x)dx \approx 0,637963$	$\int_0^1 f(x)dx \approx 0,635409$
$f(x) = e^{-x}, \text{ no intervalo } [0,1], \text{ valor exato da integral: } 0,632121$	

### 6.3.2 Regra Composta de Simpson

Expõe-se um exemplo:

**Exemplo 6.4:** Agora, considere a função  $f(x) = x \exp(x^2)$  no intervalo  $[0,2]$ . Aplicando a regra de Simpson, consegue-se a aproximação

$$\int_0^2 x \exp(x^2) dx \approx \frac{1}{3}(f(0) + 4f(1) + f(2)) = \frac{4}{3}e + \frac{2}{3}e^4 = 40,023142,$$

sendo o valor exato

$$\int_0^2 x \exp(x^2) dx = \frac{1}{2} \exp(x^2) \Big|_{x=0}^{x=2} = \frac{1}{2}(e^4 - 1) = 26,799075;$$

o que mostra a pouca precisão da aproximação da integral.

A semelhança do exemplo 6.2, subdivide-se o intervalo  $[0,2]$  em  $[0,1]$  e  $[1,2]$ , e aplica-se a regra de Simpson em cada subintervalo. Tem-se, então,

$$\int_0^1 x \exp(x^2) dx \approx \frac{1}{6}(f(0) + 4f(1/2) + f(1)) = \frac{1}{3}e^{1/4} + \frac{1}{6}e = 0,881055,$$

$$\int_1^2 x \exp(x^2) dx \approx \frac{1}{6}(f(1) + 4f(3/2) + f(2)) = \frac{1}{6}e + e^{9/4} + \frac{1}{3}e^4 = 28,140166,$$

que dá

$$\int_0^2 x \exp(x^2) dx = \int_0^1 x \exp(x^2) dx + \int_1^2 x \exp(x^2) dx \approx 0,881055 + 28,140166 = 29,021221,$$

resultado que melhora a precisão anterior. Se o intervalo  $[0,2]$  é subdividido em quatro subintervalos iguais, tem-se

$$\begin{aligned} \int_0^2 x \exp(x^2) dx &= \int_0^{1/2} x \exp(x^2) dx + \int_{1/2}^1 x \exp(x^2) dx + \int_1^{3/2} x \exp(x^2) dx + \int_{3/2}^2 x \exp(x^2) dx \\ &\approx \frac{1}{12}(f(0) + 4f(1/4) + f(1/2)) + \frac{1}{12}(f(1/2) + 4f(3/4) + f(1)) \\ &\quad + \frac{1}{12}(f(1) + 4f(5/4) + f(3/2)) + \frac{1}{12}(f(3/2) + 4f(7/4) + f(2)) \\ &= \frac{1}{12}(f(0) + 4f(1/4) + 2f(1/2) + 4f(3/4) + 2f(1) + 4f(5/4) + 2f(3/2) + 4f(7/4) + f(2)) \\ &= \frac{1}{12}(e^{1/16} + e^{1/4} + 3e^{9/16} + 2e + 5e^{25/16} + 3e^{9/4} + 7e^{49/16} + 2e^4) = 27,019168, \end{aligned}$$

conseguindo-se ainda maior precisão.

Considere então o caso geral: suponha uma função  $f(x)$  com derivada contínua até de quarta ordem no intervalo  $[a, b]$  e uma partição  $a = x_1 < x_2 < \dots < x_n = b$ , ( $n > 1$ ), igualmente espaçada com espaçamento

$h = \frac{b-a}{n-1}$ . Então a regra composta de Simpson para a partição é formada aplicando a regra de Simpson,  $(\frac{n-1}{2})$

vezes, uma vez a cada subintervalo  $[x_1, x_3], \dots, [x_{n-2}, x_n]$ , e pode ser escrita como

$$\int_a^b f(x) dx = \frac{h}{3}(f_{\text{úüü}} + 4(f_1 + f_3 + \dots + f_{n-1}) + 2(f_2 + f_4 + \dots + f_{n-2}) + f_n),$$

tendo como estimativa de erro o valor

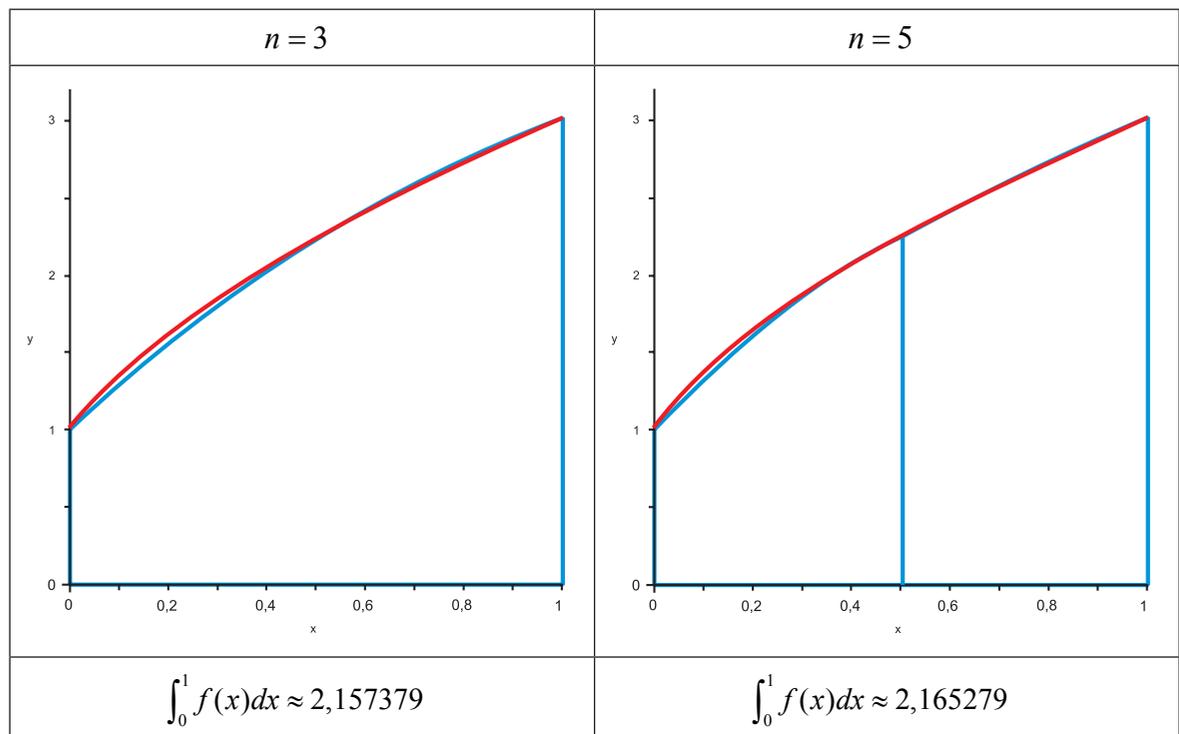
$$\frac{(n-1)h^5}{180} M_2.$$

Observe que para realizar este processo, o número de pontos da partição,  $n$ , tem que ser ímpar.

**Exemplo 6.5:** Aplica-se a regra composta de Simpson com  $n = 9$  pontos, à função  $f(x) = \sqrt{8x+1}$  sobre o intervalo  $[0,1]$ , para aproximar a integral  $\int_0^1 f(x)dx$  (observe que  $h = \frac{1}{8}$ ):

$$\begin{aligned} \int_0^1 (2x^3 + 1) dx &\approx \frac{1}{24} (f_1 + 4(f_2 + f_4 + f_6 + f_8) + 2(f_3 + f_5 + f_7) + f_9) \\ &= \frac{1}{24} (1 + 4(\sqrt{2} + 2 + \sqrt{6} + 2\sqrt{2}) + 2(\sqrt{3} + \sqrt{5} + \sqrt{7}) + 3). \\ &= 2,166511 \end{aligned}$$

Também, nas seguintes tabelas, mostram-se diversas aproximações das integrais das funções indicadas mediante a regra composta de Simpson.



$n = 7$	$n = 9$
$\int_0^1 f(x)dx \approx 2,166269$	$\int_0^1 f(x)dx \approx 2,166511$
$f(x) = \sqrt{8x+1}$ , no intervalo $[0,1]$ , valor exato da integral: 2,166667	

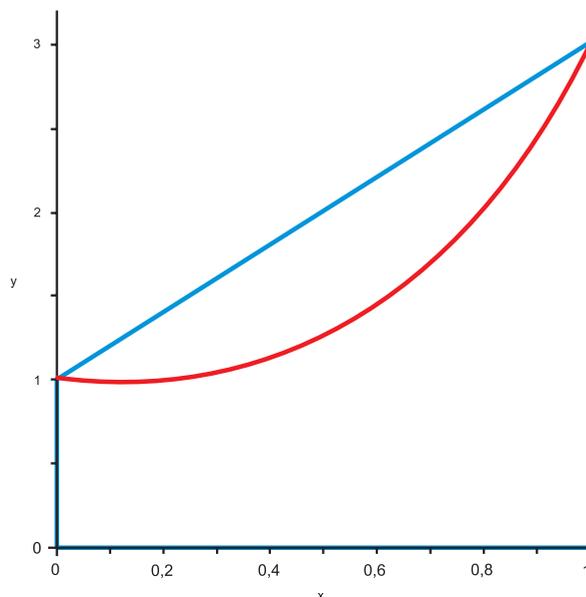
$n = 3$	$n = 5$	$n = 7$	$n = 9$
$\int_0^2 f(x)dx \approx 40,023142$	$\int_0^2 f(x)dx \approx 29,021222$	$\int_0^2 f(x)dx \approx 27,404267$	$\int_0^2 f(x)dx \approx 27,019168$
$f(x) = x \exp(x^2)$ , no intervalo $[0,2]$ , valor exato da integral: 26,799075			

$n = 3$	$n = 5$	$n = 7$	$n = 9$
$\int_0^1 f(x)dx \approx 0,632334$	$\int_0^1 f(x)dx \approx 0,632134$	$\int_0^1 f(x)dx \approx 0,632123$	$\int_0^1 f(x)dx \approx 0,632121$
$f(x) = e^{-x}$ , no intervalo $[0,1]$ , valor exato da integral: 0,632121			

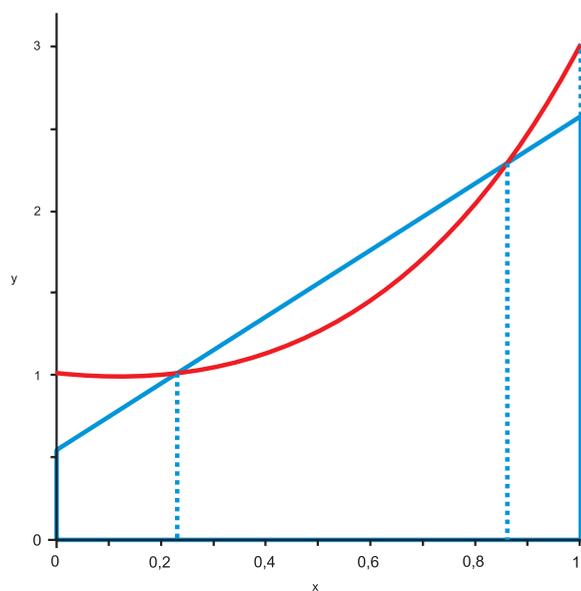
## 6.4 Quadratura Gaussiana

As fórmulas de Newton-Cotes foram deduzidas mediante a integração do polinômio de interpolação de Lagrange. Observe que a restrição que impõe o uso de pontos da partição igualmente espaçados é conveniente para o uso repetido destas fórmulas, mas, existem situações em que estes pontos escolhidos não são adequados.

**Exemplo 6.6:** Veja a situação quando  $f(x) = 2x^3 + 1$  no intervalo  $[0,1]$ . Aplicando uma vez a regra do trapézio tem-se, graficamente, o que se mostra na figura



A regra do trapézio aproxima a integral da função considerando sempre os extremos do intervalo. Mas esta não é a melhor das escolhas. Considere, por exemplo, outra escolha de pontos que não são os extremos do intervalo e que parece mais adequada, ilustrada graficamente na figura a seguir



A idéia básica da quadratura gaussiana consiste na escolha de um conjunto de pontos  $\{x_1, x_2, \dots, x_n\}$  e de coeficientes  $a_i, i = 1, 2, \dots, n$ , de maneira que a aproximação

$$\int_a^b f(x) dx \approx \sum_{i=1}^n a_i f(x_i) = a_1 f(x_1) + a_2 f(x_2) + \dots + a_n f(x_n)$$

seja exata no caso em que  $f(x)$  seja um polinômio de grau até  $2n - 1$ . Para o cálculo prático dos valores necessários, pode-se escolher, entre outras, a base formada pelas funções  $f(x) = x^m$ , para  $m = 0, 1, \dots, 2n - 1$ .

Especificamente, por exemplo, considere que é desejado determinar os valores de  $a_1, a_2, x_1$  e  $x_2$  de maneira que a aproximação

$$\int_{-1}^1 f(x) dx \approx a_1 f(x_1) + a_2 f(x_2),$$

fornece o valor exato da integral no caso em que  $f(x) = 1, x, x^2$  ou  $x^3$ .

Então,

$$a_1 \cdot 1 + a_2 \cdot 1 = \int_{-1}^1 1 dx = 2,$$

$$a_1 \cdot x_1 + a_2 \cdot x_2 = \int_{-1}^1 x dx = 0, \quad (f(x) = x \text{ é uma função ímpar sobre } [-1,1]),$$

$$a_1 \cdot x_1^2 + a_2 \cdot x_2^2 = \int_{-1}^1 x^2 dx = \frac{2}{3},$$

$$a_1 \cdot x_1^3 + a_2 \cdot x_2^3 = \int_{-1}^1 x^3 dx = 0, \quad (f(x) = x^3 \text{ é uma função ímpar sobre } [-1,1]).$$

Juntando as últimas quatro igualdades tem-se

$$\begin{cases} a_1 + a_2 = 2, \\ a_1 \cdot x_1 + a_2 \cdot x_2 = 0 \\ a_1 \cdot x_1^2 + a_2 \cdot x_2^2 = \frac{2}{3} \\ a_1 \cdot x_1^3 + a_2 \cdot x_2^3 = 0 \end{cases}$$

que é um sistema de quatro equações não lineares com quatro incógnitas. Após um trabalho algébrico, consegue-se a única solução  $a_1 = a_2 = 1, x_1 = -\frac{1}{\sqrt{3}}$  e  $x_2 = \frac{1}{\sqrt{3}}$ , fornecendo então a aproximação

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

que é exata para polinômios até de terceiro grau.

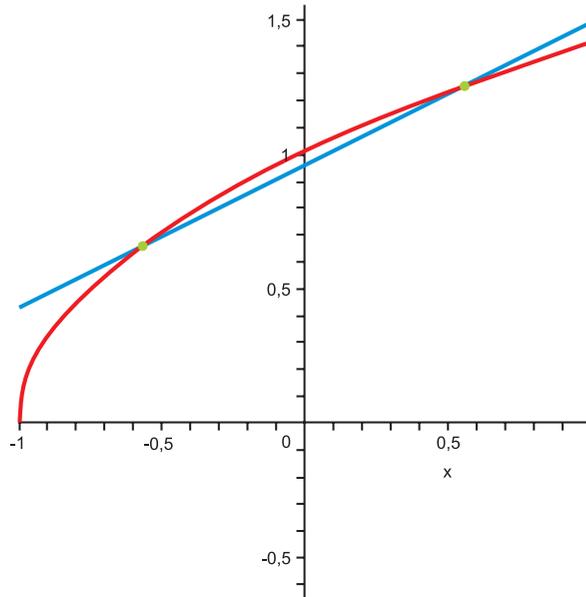
**Exemplo 6.7:** Considere a função  $f(x) = \sqrt{x+1}, x \in [-1,1]$ . Então

$$\int_{-1}^1 \sqrt{x+1} dx = \frac{2}{3} (x+1)^{3/2} \Big|_{x=-1}^{x=1} = \frac{2}{3} \cdot 2^{3/2} = \frac{4}{3} \sqrt{2} = 1,885618.$$

Por outro lado, utilizando a quadratura gaussiana com dois pontos, tem-se

$$\int_{-1}^1 \sqrt{x+1} dx \approx \sqrt{-\frac{1}{\sqrt{3}}+1} + \sqrt{\frac{1}{\sqrt{3}}+1} = 0,650115 + 1,255926 = 1,906041.$$

A figura a seguir mostra a localização dos elementos da quadratura



Agora, considerando a regra do trapézio, tem-se

$$\int_{-1}^1 \sqrt{x+1} dx \approx \frac{2}{2} (\sqrt{-1+1} + \sqrt{1+1}) = 0 + \sqrt{2} = 1,414214,$$

resultado que foi obtido calculando a área do triângulo (caso extremo de um trapézio em que uma das bases tem comprimento zero) mostrado na figura

Os gráficos mostrados e os resultados numéricos mostram a melhor exatidão da quadratura gaussiana.

Os resultados a seguir mostram os coeficientes  $a_i$  e os valores  $x_i$  para o cálculo aproximado da integral

$\int_{-1}^1 f(x) dx$  mediante quadratura gaussiana

$$\int_{-1}^1 f(x) dx = a_1 f(x_1) + a_2 f(x_2) + \dots + a_n f(x_n).$$

Observe que o intervalo de integração é  $[-1,1]$ .

Número de pontos, $n$	Pontos, $x_i$	Coeficientes, $a_i$
1	0	2
2	$-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$	1, 1
3	$-\sqrt{\frac{3}{5}}, 0, \sqrt{\frac{3}{5}}$	$\frac{5}{9}, \frac{8}{9}, \frac{5}{9}$
4	$\pm \sqrt{\frac{3-2\sqrt{6/5}}{7}}$	$\frac{1}{2} + \frac{1}{6}\sqrt{\frac{5}{6}}$
	$\pm \sqrt{\frac{3+2\sqrt{6/5}}{7}}$	$\frac{1}{2} + \frac{1}{6}\sqrt{\frac{5}{6}}$

	0	$\frac{128}{255}$
5	$\pm \frac{1}{3} \sqrt{5 - 2\sqrt{10/7}}$	$\frac{161}{450} + \frac{13}{30} \sqrt{\frac{7}{90}}$
	$\pm \frac{1}{3} \sqrt{5 + 2\sqrt{10/7}}$	$\frac{161}{450} - \frac{13}{30} \sqrt{\frac{7}{90}}$

Estes valores podem ser utilizados inclusive para integrais do tipo  $\int_a^b f(x) dx$  com intervalo de integração  $[a, b]$  arbitrário, utilizando a mudança de variável  $u = \frac{b-a}{2}x + \frac{a+b}{2}$ , que realiza a mudança de intervalo de integração para  $[-1, 1]$ , mediante a seguinte igualdade

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

**Exemplo 6.8:** Requer-se aproximar  $\int_0^1 \sqrt{8x+1} dx$  utilizando a quadratura gaussiana com  $n = 2$  pontos.

Então, tem-se que

$$\begin{aligned} \int_0^1 \sqrt{8x+1} dx &= \frac{1}{2} \int_{-1}^1 \sqrt{8\left(\frac{1}{2}x + \frac{1}{2}\right) + 1} dx = \frac{1}{2} \int_{-1}^1 \sqrt{4x+5} dx \\ &\approx \frac{1}{2} \left[ \sqrt{-\frac{4}{\sqrt{3}} + 5} + \sqrt{\frac{4}{\sqrt{3}} + 5} \right] = \frac{1}{2} [1,640305 + 2,703590] = 2,171947 \end{aligned}$$

Se fossem requeridos  $n = 3$  pontos, se teria

$$\begin{aligned} \int_0^1 \sqrt{8x+1} dx &= \frac{1}{2} \int_{-1}^1 \sqrt{4x+5} dx \\ &\approx \frac{1}{2} \left[ \frac{1}{9} \cdot \sqrt{-4\sqrt{\frac{1}{5}} + 5} + \frac{1}{9} \cdot \sqrt{5} + \frac{1}{9} \cdot \sqrt{4\sqrt{\frac{1}{5}} + 5} \right] = \frac{1}{2} [0,766106 + 1,987616 + 1,580981] \\ &= 2,167351 \end{aligned}$$

Lembre que o valor exato da integral é  $\int_0^1 \sqrt{8x+1} dx = 2,166667$ .