

UNIVERSIDADE FEDERAL DE OURO PRETO - UFOP  
INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS - ICEB  
DEPARTAMENTO DE COMPUTAÇÃO - DECOM

---

---

NOTAS DE AULA

**CÁLCULO NUMÉRICO (BCC760)**

---

---

2020-1

ELABORADO POR PROFESSORES DA DISCIPLINA BCC760  
<[HTTP://WWW.DECOM.UFOP.BR/BCC760/](http://www.decom.ufop.br/bcc760/)>

## PREFÁCIO

Esta apostila consiste em notas de aula da disciplina Cálculo Numérico, oferecida pelo Departamento de Computação (DECOM) da Universidade Federal de Ouro Preto (UFOP). O material presente nas notas não é original e não substitui a consulta dos livros bases, tem por finalidade apenas auxiliar no desenvolvimento do conteúdo da disciplina e apoio didático aos discentes. O material é uma reescrita das notas de aulas do Professor José Álvaro Tadeu Ferreira e contém textos adaptados e exemplos das seguintes obras:

1. CAMPOS filho, F. F. *Algoritmos Numéricos: uma abordagem moderna de Cálculo Numérico*. 3<sup>a</sup>. ed. Rio de Janeiro: LTC, 2018. v. 1.
2. GILAT, A.; SUBRAMANIAM, V. *Métodos Numéricos para Engenheiros e Cientistas: Uma introdução com aplicações usando o MATLAB*. Porto Alegre: Bookman, 2008.
3. FRANCO, N. M. B. *Cálculo Numérico*. 1<sup>a</sup>. ed. São Paulo: Pearson Prentice Hall, 2006. v. 1.
4. SPERANDIO, D.; MENDES, J. T.; SILVA, L. H. M. e. *Cálculo Numérico*. 2<sup>a</sup>. ed. São Paulo: Pearson, 2014. v. 1.

Todo e qualquer erro encontrado neste material é de responsabilidade (e culpa) dos autores.

A modelagem matemática é utilizada por engenheiros e cientistas para investigar problemas em diversas áreas. Para resolver um modelo matemático, na maioria dos casos, podem ser utilizados métodos numéricos. Neste contexto, o objetivo geral da disciplina é apresentar os principais fundamentos de um grupo de métodos numéricos utilizados para resolver computacionalmente problemas reais e práticos.

Utilizar a matemática e os algoritmos na resolução de problemas práticos permite desenvolver diferentes habilidades, tais como: compreender um problema; projetar uma solução algorítmica geral para ele; implementar o algoritmo em uma linguagem de programação; testar o programa e realizar simulações para validar a solução e a implementação realizada; compreender a construção de métodos numéricos; analisar em que condições se pode ter a garantia de que os resultados obtidos estão próximos dos reais; entre outras habilidades.

Nos tópicos que compõem a ementa desta disciplina, os métodos serão apresentados exibindo as vantagens e desvantagens de cada um deles através da comparação entre os métodos que são aplicados a uma mesma classe de problema. O propósito é fornecer condições para que os alunos possam identificar e aplicar os métodos numéricos estudados na resolução de problemas da sua área de formação, escolhendo e desenvolvendo algoritmos para as técnicas mais apropriadas para o problema.

De forma sucinta, a ementa deste curso é a seguinte: resolução de sistemas de equações lineares; interpolação polinomial; integração numérica e raízes de equações algébricas e transcendentais. E a divisão dos tópicos propostos e a ordem de apresentação de todo conteúdo têm como finalidade facilitar todo o processo de ensino aprendizagem. Assumindo que os estudantes matriculados tenham um conhecimento prévio de Cálculo Diferencial e Integral e Álgebra linear, uma breve revisão e a apresentação dos principais conceitos abordados nesta disciplina são apresentados na Parte I. As partes seguintes apresentam o conteúdo programático da disciplina seguindo a ordem: nas Partes II e III são apresentados os métodos numéricos para resolução de sistemas lineares; na Parte IV métodos de interpolação polinomial; na Parte V métodos de integração numérica e; por fim, a Parte VI apresenta métodos para a resolução de equações não-lineares.

Ouro Preto, 2020

# Sumário

<b>I</b>	<b>Apresentação do Curso e Revisão</b>	<b>1</b>
<b>1</b>	<b>Apresentação da Disciplina</b>	<b>3</b>
1.1	Resolvendo um problema por meio do Cálculo Numérico . . . . .	4
1.2	Conceitos Básicos . . . . .	5
1.3	Erros em Soluções Numéricas . . . . .	6
1.4	Exercícios . . . . .	8
<b>2</b>	<b>Fundamentos da Álgebra Linear</b>	<b>11</b>
2.1	Matrizes e Operações Matriciais . . . . .	11
2.2	Exercícios . . . . .	15
<b>II</b>	<b>Resolução de Sistemas de Equações Lineares Simultâneas - Métodos Diretos</b>	<b>19</b>
<b>3</b>	<b>Sistemas Triangulares</b>	<b>21</b>
3.1	Sistemas de Equações Lineares Simultâneas . . . . .	21
3.2	Métodos Diretos . . . . .	23
3.2.1	Resolução de Sistemas Triangulares . . . . .	23
3.3	Exercícios . . . . .	26
<b>4</b>	<b>Eliminação de Gauss</b>	<b>31</b>
4.1	Definições . . . . .	31
4.2	Método de Eliminação de Gauss . . . . .	32
4.2.1	Descrição do Algoritmo . . . . .	33
4.2.2	Avaliação do Resíduo . . . . .	37
4.3	Exercícios . . . . .	39
<b>5</b>	<b>Eliminação de Gauss com Pivotação</b>	<b>43</b>
5.1	Método de Eliminação de Gauss com Pivotação . . . . .	43
5.1.1	Pivotação Parcial . . . . .	44
5.2	Exercícios . . . . .	46
<b>6</b>	<b>Fatoração LU</b>	<b>49</b>
6.1	Fatoração LU da matriz A . . . . .	50
6.2	Exercícios . . . . .	53
<b>7</b>	<b>Decomposição LU</b>	<b>57</b>
7.1	Decomposição LU . . . . .	57
7.2	Decomposição LU com pivotação parcial . . . . .	59
7.3	Algoritmo e Complexidade . . . . .	61
7.4	Exercícios . . . . .	62

### III Resolução de Sistemas de Equações Lineares Simultâneas - Métodos Iterativos 65

<b>8 Métodos Iterativos</b>	<b>67</b>
8.1 Introdução	67
8.1.1 Revisão da Álgebra: Norma Matricial	68
8.1.2 Convergência dos Métodos Iterativos	70
8.1.3 Critério de Parada	71
8.2 Método de Jacobi	71
8.2.1 Critério de Convergência	73
8.3 Exercícios	75
<b>9 Métodos de Gauss Seidel</b>	<b>79</b>
9.1 Método de Gauss Seidel	79
9.1.1 Critério de Convergência	81
9.2 Complexidade dos métodos iterativos	83
9.3 Considerações Finais	83
9.4 Exercícios	84

### IV Interpolação Polinomial 89

<b>10 Interpolação Polinomial</b>	<b>91</b>
10.1 Introdução	91
10.1.1 Interpolação Linear	94
10.1.2 Erro de Truncamento	94
10.2 Polinômios Interpoladores de Lagrange	97
10.3 Exercícios	101
<b>11 Método de Diferenças Divididas</b>	<b>105</b>
11.1 Introdução	105
11.2 Polinômio Interpolador de Diferenças Divididas	109
11.3 Exercícios	114
<b>12 Método de Diferenças Finitas</b>	<b>117</b>
12.1 Método das Diferenças Finitas	117
12.1.1 Diferenças Finitas	117
12.1.2 Polinômio Interpolador de Diferenças Finitas	120
12.2 Complexidade dos métodos de interpolação	123
12.3 Escolha dos pontos de interpolação	124
12.4 Considerações Finais	124
12.5 Exercícios	125

### V Integração Numérica 129

<b>13 Integração Numérica</b>	<b>131</b>
13.1 Introdução	131
13.2 Fórmulas de Newton-Cotes	134
13.3 Regra dos Trapézios	136
13.3.1 Regra dos trapézios generalizada	137

13.3.2 Erro de truncamento na regra dos trapézios . . . . .	138
13.4 Exercícios . . . . .	141
<b>14 Primeira Regra de Simpson</b>	<b>145</b>
14.1 Introdução . . . . .	145
14.1.1 Primeira regra de Simpson generalizada . . . . .	147
14.1.2 Erro de Truncamento na primeira regra de Simpson . . . . .	148
14.2 Exercícios . . . . .	151
<b>15 Segunda Regra de Simpson</b>	<b>155</b>
15.1 Introdução . . . . .	155
15.1.1 Segunda regra de Simpson generalizada . . . . .	157
15.1.2 Erro de Truncamento na segunda regra de Simpson . . . . .	158
15.2 Considerações Finais . . . . .	161
15.3 Exercícios . . . . .	162
<b>VI Raízes de Equações Algébricas e Transcendentes</b>	<b>165</b>
<b>16 Raízes de Equações</b>	<b>167</b>
16.1 Introdução . . . . .	167
16.2 Determinação das raízes . . . . .	171
16.2.1 Isolamento das raízes . . . . .	171
16.3 Estudo Especial de Equações Polinomiais . . . . .	176
16.3.1 Delimitação das raízes reais . . . . .	176
16.3.2 Enumeração das raízes reais . . . . .	179
16.4 Exercícios . . . . .	180
<b>17 Método da Bisseção</b>	<b>183</b>
17.1 Método da Bisseção . . . . .	183
17.2 Exercícios . . . . .	188
<b>18 Método da Falsa Posição</b>	<b>191</b>
18.1 Método da Falsa Posição . . . . .	191
18.1.1 Considerações Finais . . . . .	196
18.2 Exercícios . . . . .	196
<b>19 Método de Newton Raphson</b>	<b>199</b>
19.1 Método de Newton-Raphson . . . . .	199
19.1.1 Considerações Finais . . . . .	204
19.2 Exercícios . . . . .	204
<b>VII Atividades Avaliativas</b>	<b>207</b>
<b>Referências Bibliográficas</b>	<b>209</b>

## **Parte I**

# **Apresentação do Curso e Revisão**





## Apresentação da Disciplina

### Sumário da Aula

1.1	Resolvendo um problema por meio do Cálculo Numérico . . . . .	4
1.2	Conceitos Básicos . . . . .	5
1.3	Erros em Soluções Numéricas . . . . .	6
1.4	Exercícios . . . . .	8

O cálculo numérico é um ramo da matemática aplicada que trata de técnicas matemáticas usadas na resolução de problemas matemáticos das mais variadas áreas do conhecimento, que não podem ser resolvidos ou que são difíceis de serem resolvidos analiticamente. As técnicas estudadas envolvem apenas operações aritméticas elementares e, normalmente, geram uma sequência de aproximações numéricas que convergem para a solução exata do problema.

*“Uma solução numérica é um valor numérico aproximado para a solução, ou seja, um número” [2].* E, por se tratar de uma aproximação, é necessário avaliar a qualidade da solução obtida, ou seja, o quão próxima ela esta solução está da solução exata do problema, mesmo quando ela não é conhecida. Em muitos métodos numéricos, os cálculos são executados de maneira iterativa até que uma precisão desejada seja obtida.

Nesta aula, com o intuito de facilitar a compreensão dos métodos numéricos estudados, serão definidos alguns conceitos e princípios gerais do cálculo numérico, termos que serão mencionados durante toda a disciplina.



## 1.1 Resolvendo um problema por meio do Cálculo Numérico

O processo de solução de um problema qualquer na ciência e na engenharia é influenciado pelas ferramentas, ou métodos matemáticos, disponíveis para resolvê-lo. O processo, conforme ilustrado na Figura 1.1, pode ser dividido em quatro etapas: definição do problema, modelo matemático, solução numérica e análise dos resultados.

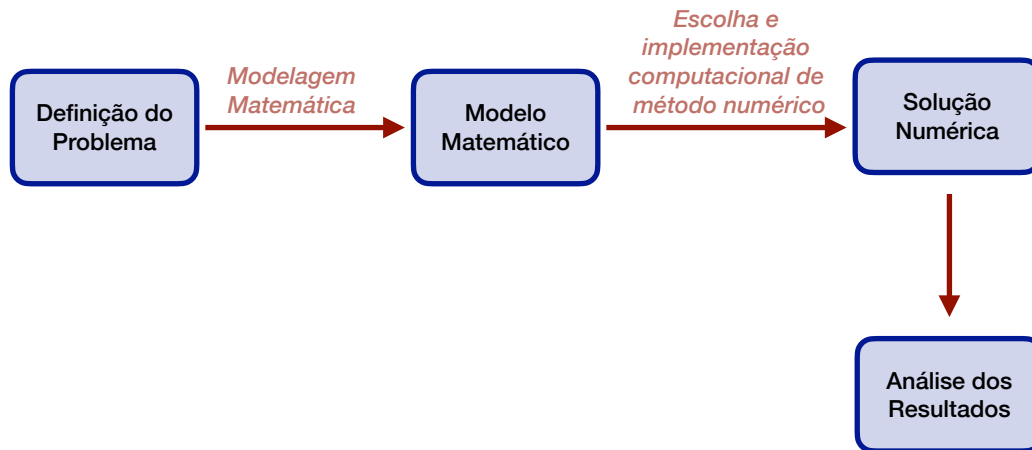


Figura 1.1: Etapas da resolução de um problema.

**1. Definição do Problema:** nesta etapa define-se qual é o problema a ser resolvido, listam-se as variáveis envolvidas e identificam-se as restrições na forma de condições iniciais e/ou de contorno.

**2. Modelo Matemático:** esta etapa consiste em transformar o problema em uma formulação matemática, ou seja, um modelo matemático usado para representar o problema na forma de equações, que serão resolvidas.

Os modelos usados (escolhidos) para resolver o problema precisam ser consistentes com os métodos que serão usados em seguida na resolução das equações. Se o emprego de métodos analíticos é esperado para a solução, as equações que governam o problema devem permitir uma solução analítica. Se necessário, a formulação tem que ser simplificada para que as equações possam ser resolvidas analiticamente. Se são usados métodos numéricos na solução, os modelos e as equações podem ser mais complicados. Ainda assim, podem existir algumas limitações. Por exemplo, se a formulação é tal que uma solução numérica requer um longo tempo de processamento, pode ser necessário simplificar a formulação para que a solução seja obtida em um tempo razoável. Um exemplo disso é a previsão do tempo. O problema a ser resolvido é grande, e os métodos numéricos utilizados são muito complicados. A simulação numérica da previsão do tempo, contudo, não pode ir além do período no qual a previsão é necessária [2, p.22].

**3. Solução Numérica:** nesta etapa, se o problema é resolvido numericamente, deve-se escolher o método numérico mais apropriado para resolvê-lo. Para cada tipo de problema matemático, existem várias técnicas numéricas disponíveis na literatura distinguindo-se entre si quanto a precisão, tempo de processamento e dificuldade de programação. Selecionado o método numérico, ele deve ser descrito por intermédio de um algoritmo que será implementado em um computador.

**4. Análise dos Resultados:** como as soluções numéricas são uma aproximação e o programa de computador que executa o método numérico pode conter erros, é preciso verificar a adequação da solução encontrada na etapa anterior e, se necessário, determinar uma nova solução. Em alguns casos, como os modelos matemáticos podem produzir soluções que não possuem sentido físico ou químico, como tempo negativo ou concentração complexa, por exemplo, é imprescindível identificar quais, dentre todas as soluções fornecidas pelo modelo, são válidas.

## 1.2 Conceitos Básicos

Nesta disciplina serão tratados problemas numéricos. Um **problema numérico** é aquele em que os dados de entrada e os dados de saída (resultados) são números ou conjuntos numéricos finitos.

**Método numérico** é um conjunto de procedimentos usados para transformar um modelo matemático em um problema numérico ou em um conjunto de procedimentos que serão usados para resolver um problema numérico [4]. A escolha do método mais adequado para resolver o problema deve levar em consideração: a precisão desejada para os resultados, a velocidade de convergência do método e o esforço computacional despendido (tempo de processamento, quantidade de memória necessária para a resolução do problema).

**Algoritmo** é uma descrição sequencial dos passos que caracterizam um método numérico. O algoritmo fornece uma descrição completa das operações, aritméticas e lógicas, por meio das quais um conjunto de dados de entrada é transformado em dados de saída. Em [4], algoritmo ainda é definido como “uma sequência de  $n$  passos, cada um envolvendo um número finito de operações. Ao fim desses  $n$  passos, o algoritmo deve fornecer valores ao menos ‘próximos’ daqueles procurados.” Nos casos em que número  $n$  não é conhecido, utiliza-se  $n$  como um limitante superior.

Uma das ideias gerais do cálculo numérico é o conceito de **iteração** ou **aproximação sucessiva**, que significa a repetição de um processo. Existe um grande número de métodos numéricos que são processos iterativos. Como o próprio nome diz, esses métodos se caracterizam pela aplicação de um procedimento de forma repetida com o objetivo de obter em cada repetição, ou iteração, uma aproximação para a solução do problema que seja mais precisa do que aquela obtida na iteração anterior.

Um método iterativo é dito **estacionário** se cada solução aproximada é obtida da solução anterior utilizando sempre o mesmo processo. Uma importante classe dos métodos iterativos é a dos **métodos iterativos estacionários de grau um**, nos quais o resultado obtido em cada iteração é uma função, somente, do resultado da iteração anterior. Nestes métodos, dado um problema  $P$  e uma estimativa inicial  $S^0$ , é gerada uma sequência de aproximações  $\{S^{(k)}\}$ ,  $k = 1, 2, \dots$  para a sua solução  $S$ , tal que:

$$S^{(k)} = \Phi(P, S^{k-1}), k = 1, 2, 3, \dots$$

sendo que a expressão  $\Phi(\cdot)$  é a função de iteração do método iterativo.

**Definição 1.1.** Um método iterativo é dito estacionário se a função de iteração é sempre a mesma em todas as iterações. Caso ela se modifique é dito não estacionário.

**Definição 1.2.** Um método iterativo é dito de grau  $g$  se são necessárias  $g$  estimativas anteriores da solução do problema para obter-se uma estimativa, ou seja, a função de iteração é da forma:

$$S^{(k)} = \Phi(P, S^{k-1}, S^{k-2}, \dots, S^{k-g}), k = g, g+1, g+2, \dots$$

por exemplo:

$$g = 1 \Rightarrow S^0 = \Phi(P) \text{ e } S^k = \Phi(P, S^{k-1}), k = 1, 2, \dots$$

$$g = 2 \Rightarrow S^0 = \Phi(P), S^1 = \Phi(P, S^0) \text{ e } S^k = \Phi(P, S^{k-1}, S^{k-2}), k = 2, 3, \dots$$

A seguir são apresentados elementos construtivos presentes nos processos iterativos estacionários de grau um, independentemente do problema a ser resolvido:

**Solução inicial:** ou “estimativa inicial”, consiste em uma primeira aproximação para a solução desejada do problema numérico. Como em cada iteração é necessário utilizar o resultado da iteração anterior, faz-se necessária uma estimativa inicial para a solução do problema, que pode ser conseguida de diferentes formas, conforme o problema que se deseja resolver.

**Função de iteração:** equação por meio da qual, a partir da estimativa inicial, são realizadas as iterações do algoritmo para se chegar na solução desejada. Por meio da função de iteração constrói-se uma sequência de estimativas para a solução do problema.

**Critério de parada:** é o procedimento responsável por finalizar o método iterativo. Obviamente não se pode repetir um processo numérico de forma indefinida. É preciso pará-lo em um determinado instante. Para isto, deve ser utilizado um critério, que vai depender do problema a ser resolvido, por meio do qual é tomada a decisão quanto à finalização do processo. Este critério de parada envolve a precisão desejada na solução do problema e/ou um número máximo de iterações.

**Convergência:** É preciso que o método iterativo gere uma sequência que convirja para a solução do problema. Isto significa que o resultado obtido em uma iteração deve estar mais próximo da solução do que o anterior. Observe-se que nem sempre se tem a garantia de que essa convergência ocorrerá.

Como nos métodos iterativos é importante saber se a sequência gerada de soluções aproximadas está convergindo ou não para a solução exata do problema, é preciso conhecer e entender bem o conceito de **convergência**. Em geral, dada uma sequência de vetores  $x^{(k)} \in E$  e uma norma sobre  $E$ , onde  $E$  é um espaço vetorial, dizemos que a sequência  $\{x^{(k)}\}$  converge para  $x \in E$  se

$$\|x^{(k)} - x\| \rightarrow 0, \text{ quando } k \rightarrow \infty.$$

## 1.3 Erros em Soluções Numéricas

Soluções numéricas podem ser muito precisas, mas em geral são inexatas. A seguir são apresentados os principais erros que podem acontecer durante o processo de busca da solução de um modelo matemático.

Os erros podem ter diversas origens e merecem um tratamento especial para que o resultado final não esteja muito distante da solução do problema real. As principais fontes de erros são: nos dados de entrada; na modelagem matemática; nos arredondamentos durante a computação; nos erros de truncamentos; e nos erros humanos e de máquinas.

### Erro de Modelagem

São erros que ocorrem na tentativa de representar um fenômeno da natureza através de um modelo matemático. A modelagem do problema real em uma expressão matemática nem sempre é fácil, pois é preciso que a expressão represente bem o fenômeno que ocorre no mundo físico. Uma modelagem incorreta pode acontecer quando a expressão matemática não reflete perfeitamente o fenômeno físico estudado ou quando os dados experimentais obtidos possuem baixa exatidão. Nestes casos, novos testes devem ser realizados para validar o quanto os resultados são sensíveis as alterações dos dados fornecidos e correções no modelo matemático devem ser realizadas.

### Erro de arredondamento

Os erros de arredondamento surgem devido a algumas propriedades básicas da aritmética real não valerem quando são executadas no computador. Na matemática, alguns números possuem a mantissa muito longa ou são representados por infinitos dígitos, como o caso dos números irracionais. Para representar esses números

no computador é preciso encurtá-los, visto que a memória da máquina, ou a quantidade de bits disponível para representá-los, é finita.

Um número pode ser encurtado fazendo o corte (descarte dos algarismos a mais) ou utilizando a regra de arredondamento ABNT NBR 5891:1977 <sup>1</sup>. No corte, os algarismos na mantissa além do comprimento que pode ser armazenado são simplesmente deixados de fora, enquanto, no arredondamento, o último algarismo armazenado é arredondado.

Para uma simples ilustração, considere o número  $2/3 = 0,6666\ldots$ . Na forma decimal com quatro algarismos significativos, obtém-se  $2/3 \approx 0,6666$  fazendo o corte e  $2/3 \approx 0,6667$  seguindo as regras de arredondamento. Ambas as representações para o número  $2/3$  possuem **erro de arredondamento**.

## Erro de truncamento

Os erros de truncamento ocorrem quando os métodos numéricos usados na solução de um problema matemático adotam um procedimento matemático finito.

Um exemplo simples em que ocorre o erro de truncamento é na avaliação numérica de  $\text{sen}(x)$  usando a expansão em série de Taylor:

Numericamente, a expansão da série de Taylor da função  $f(x) = \text{sen}(x)$  é expressa por:

$$\text{sen}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \cdots$$

para calcular  $\text{sen}(\pi/6)$  de forma exata, utilizando a série de Taylor, é necessário usar os infinitos termos da série. Utilizando os dois primeiros termos, o valor encontrado é:

$$\text{sen}(\pi/6) = \pi/6 - \frac{(\pi/6)^3}{3!} = 0,4996742$$

Nessa aproximação, a diferença da solução exata para a solução encontrada ( $\text{Erro} = |0,5 - 0,4996742| = 0,0003258$ ) é denominada **erro de truncamento**.

O erro de truncamento é independente do erro de arredondamento e existe mesmo quando as operações matemáticas são exatas. Este erro é dependente do método numérico adotado ou do algoritmo usado na solução do problema.

## Erro absoluto e erro relativo

A solução numérica é uma aproximação podendo incluir erros de arredondamento e, dependendo do método numérico utilizado, erros de truncamento. Juntos, os erros de arredondamento e truncamento resultam no **erro absoluto** (ou erro total) incluído na solução numérica. Esse erro é a diferença entre a solução verdadeira (exata) e a solução numérica conforme apresentado na Equação (1.1).

$$\text{Erro absoluto} = |\text{Valor Exato} - \text{valor Aproximado}| \quad (1.1)$$

O razão entre o erro absoluto e a solução exata é chamado de **erro relativo**:

$$\text{Erro Relativo} = \left| \frac{\text{Valor Exato} - \text{Valor Aproximado}}{\text{Valor Exato}} \right| \quad (1.2)$$

Essa grandeza adimensional e independente de escalas indica quão grande é o erro em relação à solução exata. Geralmente, o erro relativo é apresentado em forma de porcentagem. Assim, ao dizer que a solução

---

<sup>1</sup>**Regra de arredondamento:** no corte, quando o algarismo imediatamente seguinte ao último algarismo a ser conservado for inferior a 5, o último algarismo a ser conservado permanecerá sem modificação. Quando o algarismo imediatamente seguinte ao último algarismo a ser conservado for superior ou igual a 5, for seguido de no mínimo um algarismo diferente de zero, o último algarismo a ser conservado deverá ser aumentado de uma unidade.

encontrada apresenta 3% de erro relativo significa que este erro é de 0,03. Se a magnitude do erro do valor aproximado não excede  $0,5 \times 10^{-t}$ , dizemos que este valor possui  $t$  casas decimais corretas do valor exato. Por exemplo, considere o valor exato 0,001234 e o valor aproximado 0,001230. Como o erro é 0,000004, dizemos que o valor aproximado possui 5 casas decimais corretas.

O erro absoluto e o erro relativo, Equações 1.1 e 1.2, não podem ser determinados em problemas cuja solução verdadeira não é conhecida. Contudo, essas grandezas podem ser úteis na verificação da precisão de diversos métodos numéricos. Nos casos onde esses erros não puderem ser calculados, a precisão de uma solução numérica poderá ser avaliada de outra maneira. A forma como a precisão será calculada vai depender do método numérico utilizado. Em alguns casos, o erro numérico pode ser colocado em limites conhecidos ou aplicados outros métodos que permitem uma estimativa da ordem de grandeza do erro. Em aplicações práticas, soluções numéricas também podem ser comparadas com resultados experimentais.

## 1.4 Exercícios



E. 1. Explique os seguintes conceitos:

- a) Problema numérico.
- b) Método numérico.
- c) Algoritmo.

E. 2. Explique o que é erro de arredondamento e erro de truncamento. Apresente exemplos.

E. 3. Sejam  $x = 0,714286$  e  $y = 0,714239$  aproximações para  $5/7$ . Quantas casas decimais corretas tem  $x$  e  $y$ ? Determine o erro absoluto e o erro relativo em  $x$  e  $y$ .

E. 4. Quantos termos da série de Taylor da função  $f$  definida por  $f(x) = \text{sen}(x)$  devem ser utilizados para calcular  $\text{sen}(3\pi/4)$ , de modo que o erro absoluto cometido na aproximação seja menor que 0,00001?





A series of horizontal lines for writing, spanning the width of the page. The lines are evenly spaced and extend from the left margin to the right edge of the page.



# Fundamentos da Álgebra Linear

## Sumário da Aula

2.1	Matrizes e Operações Matriciais . . . . .	11
2.2	Exercícios . . . . .	15

Esta seção é baseada em Anton e Rorres [Anton e Rorres\[6\]](#). Consulte a referência para mais detalhes.

## 2.1 Matrizes e Operações Matriciais

As matrizes ordenam e simplificam diversos problemas, sendo comumente utilizadas para resolver questões nas diversas áreas das ciências e engenharias.

**Definição 2.1.** Uma **matriz** é um agrupamento retangular de números. Dizemos que os números nesse agrupamento são as **entradas** da matriz.

O tamanho de uma matriz, ou a ordem, é descrito em termos do número de linhas e de colunas que ela possui. Uma matriz  $m \times n$  (“lê-se  $m$  por  $n$ ”), possui  $m$  linhas e  $n$  colunas e é representada pela seguinte notação:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (2.1)$$



Utiliza-se letras maiúsculas para denotar matrizes e sua ordem fica indicada no lado inferior direito da letra, ou seja,  $A_{m \times n}$ . Os elementos de uma matriz são, comumente, representadas pela mesma letra que a domina em minúsculo, ou seja  $a_{ij}$ . A linha e a coluna onde o elemento está posicionado na matriz é indicado, respectivamente, pelos índices  $i$  e  $j$ . Quando for desejada uma notação mais compacta, a matriz precedente pode ser escrita como

$$[a_{ij}]_{m \times n} \quad \text{ou} \quad [a_{ij}]$$

Vetores linha e coluna são tipos especiais de matriz e, normalmente, são denotados por letras minúsculas em vez de letras maiúsculas. Para essas matrizes, é desnecessário usar índices duplos para as entradas. Assim, um vetor linha  $1 \times n$  arbitrário  $v$  e um vetor coluna  $m \times 1$  arbitrário  $t$  podem ser escritos como:

$$v = [v_1 \ v_2 \ \cdots \ v_n] \quad \text{ou} \quad t = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{bmatrix} \quad (2.2)$$

Uma matriz  $A$  com  $n$  linhas e  $n$  colunas é chamada de **matriz quadrada de ordem  $n$** . As entradas em destaque ( $a_{11}, a_{22}, \dots, a_{nn}$ ) formam a **diagonal principal** da matriz  $A$ .

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

As notações apresentadas são usadas em Cálculo Numérico para abreviar o trabalho durante a resolução de sistemas de equações lineares. Para outras aplicações é preciso definir a “aritmética de matrizes” na qual algumas operações matemáticas são aplicadas em matrizes. Essas operações incluem a multiplicação por um escalar, a adição, a subtração e a multiplicação. A divisão não é uma operação permitida na álgebra matricial.

**Definição 2.2. Igualdade de Matrizes:** Sejam as matrizes  $A = [a_{ij}]$  e  $B = [b_{ij}]$ . Elas são definidas como sendo **iguais** se tiverem o mesmo tamanho e suas entradas correspondentes forem iguais.

Em notação matricial, se  $A = [a_{ij}]$  e  $B = [b_{ij}]$  possuem o mesmo tamanho e elementos idênticos, então  $(A)_{ij} = (B)_{ij}$  ou, então,  $a_{ij} = b_{ij}$ , para quaisquer valores de  $i$  e  $j$ .

**Definição 2.3. Adição e Subtração de Matrizes:** Sejam  $A$  e  $B$  matrizes do mesmo tamanho. Então, definimos a **soma**  $A + B$  a matriz resultante da soma de todas as entradas de  $B$  às entradas correspondentes de  $A$ , e a **diferença**  $A - B$  a matriz resultante da subtração das entradas de  $B$  às entradas correspondentes de  $A$ .

Em notação matricial, se  $A = [a_{ij}]$  e  $B = [b_{ij}]$  possuem o mesmo tamanho, então:

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij} \quad \text{e} \quad (A - B)_{ij} = (A)_{ij} - (B)_{ij} = a_{ij} - b_{ij}$$

para quaisquer valores de  $i$  e  $j$ .

**Exemplo 1.** Sejam as matrizes:

$$A = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix}, B = \begin{bmatrix} 0 & 2 \\ 3 & 7 \end{bmatrix}, C = \begin{bmatrix} 9 & -6 & -2 \\ 1 & 3 & 0 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix}$$

- As matrizes  $A$  e  $D$  são iguais, pois elas possuem o mesmo tamanho e todas as suas entradas correspondentes são iguais. As matrizes  $A$  e  $B$  são diferentes, pois apesar de possuírem o mesmo tamanho,

as suas entradas correspondentes se diferem. Já as matrizes  $A$  e  $C$ , como possuem tamanhos distintos não são iguais.

- Como as matrizes  $A$  e  $B$  possuem o mesmo tamanho, definimos as operações  $A + B$  e  $A - B$  como:

$$A + B = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 7 & 6 \end{bmatrix} \text{ e } A - B = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & -8 \end{bmatrix}$$

Note que expressões como  $A + C$  ou  $B - C$  não estão definidas.

**Definição 2.4. Multiplicação por um Escalar:** Seja uma matriz  $A$  e um escalar  $c$ . Define-se o **produto**  $cA$  como a matriz resultante da multiplicação de cada entrada da matriz  $A$  pela constante  $c$ . Dizemos também que  $cA$  é um **múltiplo escalar** da matriz  $A$ .

Em notação matricial, se  $A = [a_{ij}]$ , então  $(cA)_{ij} = c(A)_{ij} = ca_{ij}$ , para todos os valores de  $i$  e  $j$ .

**Exemplo 2.** Sejam as matrizes

$$A = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 \\ 7 \end{bmatrix} \text{ e } C = \begin{bmatrix} 9 & -6 & -2 \\ 1 & 3 & 0 \end{bmatrix}$$

Temos os seguintes múltiplos escalares:

$$2A = \begin{bmatrix} 4 & 6 \\ 8 & -2 \end{bmatrix}, \quad (-1)B = -B = \begin{bmatrix} -2 \\ -7 \end{bmatrix} \text{ e } 2,2C = \begin{bmatrix} 19,8 & -13,2 & -4,4 \\ 2,2 & 6,6 & 0 \end{bmatrix}$$

**Definição 2.5. Multiplicação de Matrizes:** Sejam as matrizes  $A$  do tipo  $m \times p$  e  $B$  do tipo  $p \times n$ . O **produto** de  $A$  por  $B$  é a matriz  $C = [c_{ij}]_{m \times n}$ , cujo elemento da linha  $i$  e coluna  $j$  é obtido multiplicando da linha  $i$  de  $A$  pelos elementos correspondentes da coluna  $j$  de  $B$  e, posteriormente, somando os produtos obtidos.

**Exemplo 3.** Sejam as matrizes  $A = \begin{bmatrix} 0 & 1 \\ 4 & -1 \\ 2 & 3 \end{bmatrix}_{3 \times 2}$  e  $B = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}_{2 \times 2}$ . A matriz  $C = AB$  será uma matriz de tamanho  $3 \times 2$ , ou seja:

$$C = AB = \begin{bmatrix} 0 & 1 \\ 4 & -1 \\ 2 & 3 \end{bmatrix}_{3 \times 2} \cdot \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}_{2 \times 2} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}_{3 \times 2}.$$

Para calcular a entrada da linha 1 e da coluna 1 da matriz  $C$  (termo  $c_{11}$ ), na Expressão 2.3 foram destacadas a linha 1 da matriz  $A$  e a coluna 1 da matriz  $B$ . O termo é então calculado fazendo a operação  $c_{11} = a_{11} * b_{11} + a_{12} * b_{21} = (0) * (2) + (1) * (0) = 0$

$$C = AB = \begin{bmatrix} 0 & 1 \\ 4 & -1 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} \quad (2.3)$$

Para calcular a entrada da linha 3 e da coluna 2 da matriz  $C$  (termo  $c_{32}$ ), na Expressão 2.4 foram destacadas a linha 3 da matriz  $A$  e a coluna 2 da matriz  $B$ . O termo é então calculado fazendo a operação  $c_{32} = a_{31} * b_{12} + a_{32} * b_{22} = (2) * (1) + (3) * (3) = 11$

$$C = AB = \begin{bmatrix} 0 & 1 \\ 4 & -1 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & c_{12} \\ c_{21} & c_{22} \\ c_{31} & 11 \end{bmatrix} \quad (2.4)$$

Os demais termos  $c_{ij}$  são calculados de forma análoga. A matriz resultante do produto das matrizes  $A$  e  $B$  será:

$$C = AB = \begin{bmatrix} 0 & 3 \\ 8 & 1 \\ 4 & 11 \end{bmatrix}$$

É importante destacar que a definição de multiplicação de matrizes exige que, para ser possível formar o produto entre duas matrizes, o número de colunas do primeiro fator tem que ser igual ao número de linhas do segundo fator, como mostra a Figura 2.1. Se essa condição não for satisfeita, o produto não será definido. A matriz resultante terá número de linhas igual ao da matriz do primeiro fator e o número de colunas igual ao da matriz do segundo fator.

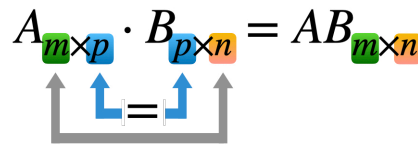


Figura 2.1: Condição de existência para a multiplicação de matrizes.

**Definição 2.6. Transposta de uma Matriz:** Seja  $A$  uma matriz qualquer  $m \times n$ . A matriz **transposta de  $A$** , denotada por  $A^t$ , é definida como uma matriz  $n \times m$  resultante da troca das linhas com as colunas da matriz  $A$ . Em outras palavras, a primeira coluna da matriz  $A^t$  é a primeira linha da matriz  $A$ ; a segunda coluna da matriz  $A^t$  é a segunda linha da matriz  $A$ ; e assim por diante.

Em notação matricial,  $(A^t)_{ij} = (A)_{ji}$ , para todos os valores de  $i$  e  $j$ .

**Exemplo 4.** Sejam as matrizes

$$A = \begin{bmatrix} 0 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 \\ 7 \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} 9 & -6 & -2 \\ 1 & 3 & 0 \end{bmatrix}$$

as suas transpostas são:  $A^t = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$ ,  $B^t = \begin{bmatrix} 2 & 7 \end{bmatrix}$  e  $C^t = \begin{bmatrix} 9 & 1 \\ -6 & 3 \\ -2 & 0 \end{bmatrix}$

**Definição 2.7. Traço de uma Matriz:** Seja  $A$  uma matriz quadrada qualquer. O **traço de  $A$** , denotada por  $tr(A)$ , é definido pela soma dos elementos da diagonal principal de  $A$ .

**Exemplo 5.** Dadas as matrizes

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 & -2 & 0 & 4 \\ 3 & 4 & -1 & 1 \\ 4 & -1 & 2 & 3 \\ 0 & 2 & 5 & 7 \end{bmatrix}$$

Temos que

$$tr(A) = a_{11} + a_{22} + \cdots + a_{nn} \quad \text{e} \quad tr(B) = 1 + 4 + 2 + 7 = 14.$$

**Observação:**

**O traço somente é definido em matrizes quadradas!**

## 2.2 Exercícios



**E. 1.** Sejam as matrizes  $A = [a_{ij}]_{4 \times 5}$ ,  $B = [b_{ij}]_{4 \times 5}$ ,  $C = [c_{ij}]_{5 \times 2}$ ,  $D = [d_{ij}]_{4 \times 2}$  e  $E = [e_{ij}]_{5 \times 4}$ . Em cada parte, determine se a expressão matricial dada está definida. Para as que estão definidas, informe qual o tamanho da matriz resultante.

- |               |                 |
|---------------|-----------------|
| a) $AB$       | b) $AB + B$     |
| c) $E^t A$    | d) $AC + D$     |
| e) $E(A + B)$ | f) $(A^t + E)D$ |
| g) $AE + B$   | h) $E(AC)$      |

**E. 2.** Sejam as matrizes

$$A = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 5 & 2 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{bmatrix}, \quad E = \begin{bmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{bmatrix}$$

Calcule, se possível, as expressões seguintes:

- |               |                |                         |
|---------------|----------------|-------------------------|
| a) $D + E$    | b) $-7C$       | c) $2B - C$             |
| d) $5A$       | e) $AB$        | f) $BA$                 |
| g) $A - A$    | h) $4E - 2D$   | i) $4tr(7B)$            |
| j) $2A^t + C$ | l) $D^t - E^t$ | m) $(CD)E$              |
| n) $tr(D)$    | o) $3(D + 2E)$ | p) $tr(C^t A^t + 2E^t)$ |

**E. 3.** Determine se as afirmações seguintes são verdadeiras ou falsas. Justifique sua resposta.

- ( ) A matriz  $\begin{bmatrix} 1 & 4 & 0 \\ -9 & 5 & 3 \end{bmatrix}$  não tem diagonal principal.
- ( ) Uma matriz  $m \times n$  tem  $m$  vetores coluna e  $n$  vetores linha.
- ( ) Se  $A$  e  $B$  forem matrizes  $2 \times 2$ , então  $AB = BA$ .
- ( ) Dada qualquer matriz  $A$ , vale  $(A^t)^t = A$ .
- ( ) Se  $A$  e  $B$  forem matrizes quadradas de mesma ordem, então  $tr(AB) = tr(A) * tr(B)$
- ( ) Se  $A$  e  $B$  são matrizes quadradas de mesma ordem, então  $(AB)^t = A^t B^t$
- ( ) Dada qualquer matriz quadrada  $A$ , vale que  $tr(A^t) = tr(A)$ .
- ( ) Se  $A$  for uma matriz  $6 \times 4$  e  $B$  uma matriz  $m \times n$  tal que  $B^t A^t$  é uma matriz  $2 \times 6$ , então  $m = 4$  e  $n = 2$ .
- ( ) Se  $A$  for uma matriz  $n \times n$  e  $c$  um escalar, então  $tr(cA) = ctr(A)$ .
- ( ) Se  $A, B$  e  $C$  forem matrizes do mesmo tamanho tais que  $A - C = B - C$ , então  $A = B$ .
- ( ) Se  $A, B$  e  $C$  forem matrizes quadradas do mesmo tamanho tais que  $AC = BC$ , então  $A = B$ .
- ( ) Se a soma de matrizes  $AB + BA$  estiver definida, então  $A$  e  $B$  devem ser quadradas de mesmo tamanho.





A series of horizontal lines for writing, starting from the first line below the icon and continuing down to the bottom of the page.





## **Parte II**

# **Resolução de Sistemas de Equações Lineares Simultâneas - Métodos Diretos**







## Sistemas Triangulares

### Sumário da Aula

3.1	Sistemas de Equações Lineares Simultâneas . . . . .	21
3.2	Métodos Diretos . . . . .	23
3.2.1	Resolução de Sistemas Triangulares . . . . .	23
3.3	Exercícios . . . . .	26

### 3.1 Sistemas de Equações Lineares Simultâneas

A resolução de sistemas de equações lineares simultâneas é um problema numérico que ocorre com muita frequência em aplicações científicas nas quais faz-se necessária a simulação de situações do mundo real. É etapa fundamental na resolução de problemas que envolvem, por exemplo, equações diferenciais parciais, determinação de caminhos ótimos em redes (grafos), regressão, sistemas não lineares, interpolação de pontos, dentre outros. Em vários problemas da Engenharia há a necessidade da resolução de sistemas de equações lineares, a título de exemplo, considere a determinação do potencial em redes elétricas, o cálculo da tensão em estruturas metálicas na construção civil, o cálculo da razão de escoamento em um sistema hidráulico com derivações, a previsão da concentração de reagentes sujeitos a reações químicas simultâneas.

Chamamos de **equação linear** toda equação da forma:

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b$$

onde  $x_1, x_2, \dots, x_n$  são as variáveis;  $a_1, a_2, \dots, a_n$  são os coeficientes das variáveis, e  $b$  é o termo independente. A solução de uma equação linear é composta pelos valores das variáveis que satisfazem à equação, também chamados de *raízes da equação*.

Um **sistema de equações lineares** é um conjunto de  $n$  equações lineares:

$$AX = B \Rightarrow \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases} \Rightarrow \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

em que  $A$  é a matriz dos coeficientes das incógnitas,  $X$  o vetor coluna das incógnitas e  $B$  o vetor coluna dos termos independentes. A notação  $AX = B$  facilita o uso algébrico do sistema linear. A segunda notação é chamada de forma desenvolvida e a terceira é a forma matricial que possibilita a implementação computacional.

As  $n$  equações de um sistema de equações lineares podem ser armazenada através de uma estrutura de dados matricial única, também conhecida como matriz aumentada (ou matriz expandida) do sistema linear. A **matriz aumentada** do sistema é obtida acrescentando à matriz dos coeficientes o vetor  $B$  dos termos independentes:

$$[A|B] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & | & b_2 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & | & b_n \end{bmatrix}$$

A solução de um sistema de equações lineares  $AX = B$  é um vetor  $S$  que satisfaz, de forma simultânea, a todas as equações do sistema, ou seja,  $AS = B$ .

Com relação ao número de soluções, um sistema de equações lineares simultâneas pode ser classificado como:

**(I) - Compatível determinado:** admite uma única solução.

**(II) - Compatível indeterminado:** admite um número infinito de soluções.

**(III) - Incompatível:** não admite solução.

Um sistema de equações lineares será compatível determinado se, e somente se, a matriz dos coeficientes for não singular, ou seja, quando o determinante da matriz dos coeficientes for não nulo ( $\det(A) \neq 0$ ). Caso contrário será compatível indeterminado ou incompatível.

Se todos os termos independentes são nulos, isto é, se  $B = [0, \dots, 0]$ , o sistema é dito homogêneo. Todo sistema homogêneo é compatível, pois admite pelo menos a solução trivial  $x = [0, \dots, 0]$ .

De uma forma mais ampla, pode-se considerar que resolver um sistema de equações é mais do que determinar um vetor  $x$ , uma vez que ele pode não existir ou não ser único. Resolver um sistema consiste em diagnosticar em qual das três situações ele se enquadra.



## 3.2 Métodos Diretos

Os métodos de solução de sistemas de equações lineares estão agrupados em dois tipos de métodos numéricos:

- (a) **Métodos Diretos:** são aqueles que, exceto por erros de arredondamento, fornecem a solução exata de um sistema de equações lineares, caso ela exista, por meio de um número finito de operações aritméticas. Nos métodos diretos, a solução é obtida com a realização de operações algébricas nas equações.
- (b) **Métodos Iterativos:** são aqueles que para resolver o sistema  $AX = B$  é atribuído um candidato para a solução e posteriormente gerada uma sequência recursiva de vetores colunas,  $x^{(k)}, k = 1, 2, \dots$ , cujo limite, se existir, será a solução procurada. Em outras palavras, uma solução inicial aproximada é assumida e utilizada em um processo iterativo para que soluções mais precisas sejam obtidas de forma sucessiva.

Os **métodos diretos**, geralmente, são utilizados na resolução de sistemas de equações lineares densos e de pequeno porte. Entenda-se por sistema denso aquele no qual os componentes da matriz dos coeficientes são em maioria não nulos e um sistema de pequeno porte aquele que possui uma dimensão pequena. Pertencem à classe dos métodos diretos todos os que são estudados nos cursos de 1 e 2 graus como, por exemplo, a **Regra de Cramer** e **Método de Gauss-Jordan**. Nas próximas seções são apresentados os métodos diretos de Eliminação de Gauss e Decomposição LU. Em geral, estes métodos transformam o sistema linear original em um sistema equivalente, cuja solução é obtida através da resolução de sistemas lineares triangulares.

### 3.2.1 Resolução de Sistemas Triangulares

#### Sistema Triangular Inferior

Considere o sistema linear  $LX = C$ , onde a matriz dos coeficientes  $L = [l_{ij}]$ ,  $i, j = 1, 2, \dots, n$  é uma matriz triangular inferior, ou seja, é uma matriz quadrada que tem seus coeficientes  $(l_{ij}) = 0$  para  $i < j$ . O sistema  $LX = C$  é escrito como:

$$\begin{cases} l_{11}x_1 & = c_1 \\ l_{21}x_1 + l_{22}x_2 & = c_2 \\ & \vdots \\ l_{n1}x_1 + l_{n2}x_2 + \dots + l_{nn}x_n & = c_n \end{cases}$$

A solução do sistema  $LX = C$  é obtida por meio de **substituições sucessivas**:

1. **Cálculo de  $x_1$ :**  $l_{11}x_1 = c_1 \longrightarrow x_1 = \frac{c_1}{l_{11}}$
2. **Cálculo de  $x_2$ :**  $l_{21}x_1 + l_{22}x_2 = c_2 \longrightarrow x_2 = \frac{c_2 - l_{21}x_1}{l_{22}}$
3. **Cálculo de  $x_3$ :**  $l_{31}x_1 + l_{32}x_2 + l_{33}x_3 = c_3 \longrightarrow x_3 = \frac{c_3 - l_{31}x_1 - l_{32}x_2}{l_{33}}$
- $\vdots$
4. **Cálculo de  $x_n$ :**  $l_{n1}x_1 + l_{n2}x_2 + \dots + l_{nn}x_n = c_n \longrightarrow x_n = \frac{c_n - l_{n1}x_1 - \dots - l_{nn-1}x_{n-1}}{l_{nn}}$

De forma geral, para construir o algoritmo que calcula a solução do sistema linear triangular inferior, considere a linha genérica para o cálculo da variável  $x_i$ :

$$l_{i1}x_1 + l_{i2}x_2 + l_{i3}x_3 \dots + l_{ii}x_i = c_i$$

isolando  $x_i$ , temos:

$$x_i = \frac{c_i - l_{i1}x_1 - l_{i2}x_2 - l_{i3}x_3 - \cdots - l_{ii-1}x_{i-1}}{l_{ii}} = \frac{c_i - (l_{i1}x_1 + l_{i2}x_2 + l_{i3}x_3 + \cdots + l_{ii-1}x_{i-1})}{l_{ii}}$$

ou ainda:

$$x_i = \frac{c_i - \sum_{j=1}^{i-1} l_{ij}x_j}{l_{ii}}, \quad i = 2, \dots, n \quad (3.1)$$

Temos, assim, o seguinte algoritmo:

```

1:  $x_1 = \frac{c_1}{l_{11}}$ ;
2: para ( $i = 2$ ) até  $n$  faça
     $c_i - \sum_{j=1}^{i-1} l_{ij}x_j$ 
3:    $x_i = \frac{\quad}{l_{ii}};$ 
4: fim para

```

Algoritmo 1: Substituições Sucessivas.

**Exemplo 6.** Calcule  $\begin{cases} 2x_1 & = 6 \\ x_1 + 4x_2 & = 7 \\ x_1 - x_2 + x_3 & = 2 \end{cases}$  utilizando o algoritmo de substituições sucessivas:

**Solução:** dado o Algoritmo 1, temos:

$$x_1 = \frac{b_1}{l_{11}} = \frac{6}{2} = 3$$

$$x_2 = \frac{b_2 - l_{21}x_1}{l_{22}} = \frac{7 - (1)(3)}{4} = 1$$

$$x_3 = \frac{b_3 - (l_{31}x_1 + l_{32}x_2)}{l_{33}} = \frac{2 - [(1)(3) + (-1)(1)]}{1} = 0$$

Portanto, a solução do sistema é  $x = \begin{bmatrix} 3 & 1 & 0 \end{bmatrix}^t$ .

## Sistema Triangular Superior

Considere o sistema linear  $UX = D$ , onde a matriz dos coeficientes  $U = [u_{ij}]$ ,  $i, j = 1, 2, \dots, n$  é uma matriz triangular superior, ou seja, é uma matriz quadrada que tem seus coeficientes  $(u_{ij}) = 0$  para  $i > j$ . O sistema  $UX = D$  é escrito como:

$$\begin{cases} u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n & = d_1 \\ u_{22}x_2 + \cdots + u_{2n}x_n & = d_2 \\ \vdots & \\ u_{nn}x_n & = d_n \end{cases}$$

A solução do sistema  $UX = D$  pode ser obtida por meio de **substituições retroativas**:

$$1. \text{ Cálculo da variável } x_n: u_{nn}x_n = d_n \longrightarrow x_n = \frac{d_n}{u_{nn}}$$

$$2. \text{ Cálculo da variável } x_{n-1}: u_{n-1,n-1}x_{n-1} + u_{n-1,n}x_n = d_{n-1} \longrightarrow x_{n-1} = \frac{d_{n-1} - u_{n-1,n}x_n}{u_{n-1,n-1}}$$

$\vdots$

3. **Cálculo da variável  $x_2$ :**  $u_{22}x_2 + u_{23}x_3 + \cdots + u_{2n}x_n = d_2 \longrightarrow x_2 = \frac{d_2 - u_{23}x_3 - \cdots - u_{2n}x_n}{u_{22}}$

4. **Cálculo da variável  $x_1$ :**  $u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n = d_1 \longrightarrow x_1 = \frac{d_1 - u_{12}x_2 - u_{13}x_3 - \cdots - u_{1n}x_n}{u_{11}}$

De forma geral, para construir o algoritmo que calcula a solução do sistema linear, considere a linha genérica para o cálculo da variável  $x_i$  :

$$u_{ii}x_i + u_{i+1,i}x_{i+1} + \cdots + u_{in}x_n = d_i$$

isolando  $x_i$ , temos:

$$x_i = \frac{d_i - u_{i+1,i}x_{i+1} - \cdots - u_{in}x_n}{u_{ii}} = \frac{d_i - (u_{i+1,i}x_{i+1} - \cdots - u_{in}x_n)}{u_{ii}}$$

ou ainda,

$$x_i = \frac{d_i - \sum_{j=(i+1)}^n u_{ij}x_j}{u_{ii}}, \quad i = 2, \dots, n \quad (3.2)$$

Temos, assim, o seguinte algoritmo:

1:  $x_n = \frac{d_n}{u_{nn}};$   
 2: **para** ( $i = n - 1$ ) **até** 1 **faça**  
      $d_i - \sum_{j=(i+1)}^n u_{ij}x_j$   
 3:  $x_i = \frac{\quad}{u_{ii}};$   
 4: **fim para**

Algoritmo 2: Substituições Retroativas

**Exemplo 7.** Calcule  $\begin{cases} 3x_1 + x_2 - 5x_3 = 4 \\ 2x_2 - x_3 = 2 \\ 3x_3 = 0 \end{cases}$  utilizando o algoritmo de substituições retroativas:

**Solução:** usando o Algoritmo 2, temos:

$$x_3 = \frac{d_3}{u_{33}} = \frac{0}{3} = 0$$

$$x_2 = \frac{d_2 - u_{23}x_3}{u_{22}} = \frac{2 - (-1)(0)}{2} = 1$$

$$x_1 = \frac{d_1 - (u_{12}x_2 + u_{13}x_3)}{u_{11}} = \frac{4 - [(1)(1) + (-5)(0)]}{3} = 1$$

Portanto, a solução do sistema é  $x = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^t$ .



### Observação

O esforço computacional ( $E_c$ ) de um algoritmo é a quantidade de operações elementares necessárias para resolver um determinado problema.

Considerando que  $\sum_{i=1}^n i = \frac{n(n+1)}{2}$ , o esforço computacional dos algoritmos de substituições sucessivas (retroativas) encontra-se resumido na Tabela 3.1.

Substituições sucessivas		Substituições retroativas	
Operações	Complexidade	Operações	Complexidade
Adições	$\frac{n^2 + 3n}{2} - 2$	Adições	$\frac{n^2 + 3n}{2} - 1$
Multiplicações	$\frac{n^2 - n}{2}$	Multiplicações	$\frac{n^2 - n}{2}$
Divisões	$n$	Divisões	$n$

Tabela 3.1: Complexidade dos métodos de substituições sucessivas e retroativas, sendo  $n$  é a ordem do sistema linear. Fonte: [1].

“Como o número de operações aritméticas das substituições sucessivas e retroativas é descrito por polinômios, esses algoritmos são ditos polinomiais.” [1]. Para a resolução de um sistema de equações triangular inferior de ordem  $n = 10$ , por exemplo, estão envolvidas um total de 10 operações de divisão, 63 operações de adição (ou subtração) e 45 operações de multiplicação. O que sugere um esforço computacional total de 118 operações.

O fato dos métodos de resolução de sistemas triangulares apresentarem um esforço computacional baixo para a obtenção de uma solução, instiga o estudo por técnicas de resolução de sistemas que se baseiam em transformar o sistema  $AX = B$  em um sistema equivalente triangular, que permite a aplicação dos algoritmos de substituições sucessivas ou retroativas para resolvê-lo. Esta ideia básica será usada nas próximas seções para descrever os métodos diretos de eliminação de Gauss e decomposição LU, que basicamente aplicam operações elementares sobre linhas da matriz  $A$  do sistema linear original, sempre que possível, para obter um sistema linear triangular equivalente.

## 3.3 Exercícios



E. 1. Resolva os sistemas lineares utilizando o método de substituição retroativa ou sucessiva:

$$\text{a) } \begin{cases} 2x_1 & & & = & 4 \\ 3x_1 & +5x_2 & & = & 1 \\ x_1 & -6x_2 & +8x_3 & = & 48 \\ -x_1 & +4x_2 & -3x_3 & +9x_4 & = & 6 \end{cases} \quad \text{b) } \begin{cases} 5x_1 & -2x_2 & +6x_3 & +x_4 & = & 1 \\ & +3x_2 & +7x_3 & -4x_4 & = & -2 \\ & & +4x_3 & +5x_4 & = & 28 \\ & & & +2x_4 & = & 8 \end{cases}$$

E. 2. Implementar, em qualquer linguagem de programação, os algoritmos de substituições sucessivas e retroativas.

E. 3. Calcular a solução dos sistemas triangulares a seguir, utilizando um dos programas implementados no exercício anterior.

$$\text{a) } \begin{cases} 4x_1 & & & = & 2 \\ -4x_1 & +5x_2 & & = & 3 \\ 1x_1 & 4x_2 & -3x_3 & = & -4 \end{cases} \quad \text{b) } \begin{cases} 3x_1 & -2x_2 & +5x_3 & = & 1 \\ & +4x_2 & -1x_3 & = & 3 \\ & & 6x_3 & = & 12 \end{cases}$$

*[Handwritten mark]*



[illegible]





# 4

## Eliminação de Gauss

### Sumário da Aula

4.1	Definições . . . . .	31
4.2	Método de Eliminação de Gauss . . . . .	32
4.2.1	Descrição do Algoritmo . . . . .	33
4.2.2	Avaliação do Resíduo . . . . .	37
4.3	Exercícios . . . . .	39

### 4.1 Definições

**Definição 4.1. Matrizes Equivalentes:** duas matrizes são ditas equivalentes quando é possível partir de uma delas e chegar à outra por meio de um número finito de transformações elementares.

**Definição 4.2. Sistemas Equivalentes:** dois sistemas de equações lineares são equivalentes quando admitem a mesma solução.

Considere  $[A|B]$  a matriz aumentada de um sistema de equações  $AX = B$ , tal que o determinante de  $A$  é não nulo, e  $[T|C]$  uma matriz equivalente a ela. Então, os sistemas de equações  $AX = B$  e  $TX = C$  são equivalentes e, portanto, possuem a mesma solução.

**Definição 4.3. Transformações Elementares:** as transformações elementares constituem um conjunto de operações que podem ser efetuadas sobre as linhas ou colunas de uma matriz. No que se refere à resolução

de sistemas de equações lineares, estas transformações são, normalmente, aplicadas apenas sobre as linhas da matriz dos coeficientes ou da matriz aumentada dependendo do método utilizado. São três tipos de operações:

- (i) multiplicação de uma linha por uma constante não-nula:

$$l_i \leftarrow \alpha \times l_i, \alpha \in \mathbb{R}, \alpha \neq 0, i = 1, 2, \dots, n$$

- (ii) troca de posição entre duas linhas:

$$l_i \leftrightarrow l_j; i, j = 1, 2, \dots, n; i \neq j$$

- (iii) adição de um múltiplo de uma linha a outra linha:

$$L_i \leftarrow l_i + \beta \times l_j, \beta \in \mathbb{R}, \beta \neq 0, i, j = 1, 2, \dots, n; i \neq j$$

---

**Observação**

---

Da Álgebra Linear sabe-se que as operações elementares não alteram a solução do sistema linear, uma vez que uma operação elementar sobre uma linha equivale a uma operação sobre a respectiva equação.

---

## 4.2 Método de Eliminação de Gauss

O Método da Eliminação de Gauss é um dos mais conhecidos e utilizados para a resolução de sistemas de equações lineares densos de pequeno a médio porte. A ideia fundamental do método é transformar o sistema linear  $AX = B$ , por meio das transformações elementares em matrizes, em um sistema equivalente triangular superior  $UX = D$ . Está dividido em duas fases:

- **Fase de Eliminação:** consiste em efetuar transformações elementares sobre as linhas da matriz aumentada do sistema de equações  $AX = B$  até que, depois de  $(n - 1)$  passos, se obtenha um sistema triangular superior equivalente  $UX = D$ .
- **Fase de Substituição:** consiste em resolver o sistema triangular superior,  $UX = D$ , por meio do algoritmo de **substituições retroativas**.

$$\underbrace{\left[ \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array} \right]}_{[A \mid B]} \xrightarrow{\text{Transf. Elemen.}} \underbrace{\left[ \begin{array}{cccc|c} u_{11} & u_{12} & \cdots & u_{1n} & d_1 \\ 0 & u_{22} & \cdots & u_{2n} & d_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & u_{nn} & d_n \end{array} \right]}_{[U \mid D]}$$



### 4.2.1 Descrição do Algoritmo

Para a descrição do método, considere o sistema linear  $AX = B$  de ordem  $n$ , onde a matriz  $A$  tem todas as submatrizes principais não singulares, ou seja,  $\det(A_k) \neq 0$ ,  $k = 1, 2, \dots, n$ . Dados  $i, j = 1, 2, \dots, n$ ,  $a_{ij}^{(1)} = a_{ij}$  e  $b_i^{(1)} = b_i$ , a matriz aumentada do sistema é dada por:

$$\left[ \begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} & \cdots & a_{3n}^{(1)} & b_3^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & a_{n3}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{array} \right].$$

#### Fase de Eliminação:

Para transformar o sistema  $AX = B$  em um sistema triangular superior  $UX = C$  são necessários  $(n - 1)$  passos.

**Passo 1 - Eliminação na primeira coluna:** neste passo, devem ser zerados todos os elementos da primeira coluna que estão abaixo da diagonal principal, ou seja, eliminação da incógnita  $x_1$  das linhas  $2, 3, \dots, n$ .

- **Determinação do pivô:** “suponha”  $a_{11}^{(1)} \neq 0$ .
- **Cálculo de multiplicadores para cada termo que será zerado (elementos abaixo do pivô):**

$$m_{i1}^{(1)} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}, \quad 2 < i < n$$

- **Eliminação da incógnita  $x_1$  nas linhas  $2, 3, \dots, n$ :** aplica-se as seguintes operações elementares

$$L_i \leftarrow L_i - m_{i1}^{(1)} L_1, \quad 2 < i < n$$

- **Matriz aumentada resultante do Passo 1:**

$$\left[ \begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & a_{32}^{(2)} & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array} \right]$$

**Passo 2 - Eliminação na segunda coluna:** neste passo, devem ser zerados todos os elementos da segunda coluna que estão abaixo da diagonal principal, ou seja, eliminação da incógnita  $x_2$  das linhas  $3, 4, \dots, n$ .

- **Determinação do pivô:** “suponha”  $a_{22}^{(2)} \neq 0$ .
- **Cálculo de multiplicadores para cada termo que será zerado (elementos abaixo do pivô):**

$$m_{i2}^{(2)} = \frac{a_{i2}^{(2)}}{a_{22}^{(2)}}, \quad 3 < i < n$$

- **Eliminação da incógnita  $x_2$  nas linhas  $3, \dots, n$ :** aplica-se as seguintes operações elementares

$$L_i \leftarrow L_i - m_{i2}^{(2)} L_2, \quad 3 < i < n$$

- Matriz aumentada resultante do Passo 2:

$$\left[ \begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ & & & \ddots & & \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} & b_n^{(3)} \end{array} \right]$$

e assim, sucessivamente, até chegar ao **Passo (n-1)**.

**Passo (n-1): Eliminação na coluna (n-1)** neste passo, deve ser zerado o elemento da coluna (n-1) que está abaixo da diagonal principal, ou seja, eliminação da incógnita  $x_{n-1}$  da linha  $n$ .

- Determinação do pivô: “**suponha**”  $a_{n-1n-1}^{(n-1)} \neq 0$ .
- Cálculo do multiplicador para o termo abaixo do pivô:

$$m_{nn-1}^{(n-1)} = \frac{a_{nn-1}^{(n-1)}}{a_{n-1n-1}^{(n-1)}}$$

- Eliminação da incógnita  $x_{n-1}$  na linha  $n$ : aplica-se a seguinte operação elementar

$$L_n \leftarrow L_n - m_{nn-1}^{(n-1)} L_{n-1}$$

- Matriz aumentada resultante do Passo (n-1):

$$\left[ \begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n-1}^{(1)} & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n-1}^{(2)} & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n-1}^{(3)} & a_{3n}^{(3)} & b_3^{(3)} \\ & & & \ddots & & & \\ 0 & 0 & 0 & & a_{1n-1}^{(n-1)} & a_{nn}^{(n-1)} & b_n^{(n-1)} \\ 0 & 0 & 0 & & 0 & a_{nn}^{(n)} & b_n^{(n)} \end{array} \right]$$

De modo geral, o **Passo k** do método de Eliminação de Gauss é obtido por:

- Determinação do Pivô: “**suponha**”  $a_{kk}^{(k)} \neq 0$ .
- Cálculo de multiplicadores para cada termo que será zerado (elementos abaixo do pivô):

$$m_{ik}^{(k)} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \quad k+1 < i < n.$$

- Eliminação da incógnita  $x_k$  nas linhas  $k+1, \dots, n$ : aplica-se as seguintes operações elementares

$$L_i \leftarrow L_i - m_{ik}^{(k)} L_k, \quad k+1 < i < n.$$

- Obtenção da matriz aumentada resultante do Passo k.

## Fase de Substituição:

O sistema triangular superior resultante:

$$\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n-1}^{(1)} & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n-1}^{(2)} & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n-1}^{(3)} & a_{3n}^{(3)} \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & a_{1n-1}^{(n-1)} & a_{nn}^{(n-1)} \\ 0 & 0 & 0 & \cdots & 0 & a_{nn}^{(n)} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ b_3^{(3)} \\ \vdots \\ b_{n-1}^{(n-1)} \\ b_n^{(n)} \end{bmatrix}$$

é equivalente ao sistema linear original e, portanto, os sistemas possuem a mesma solução. Assim, nesta fase o sistema triangular superior resultante deverá ser resolvido por meio de **substituições retroativas** (Algoritmo 2).

### Observação:

No algoritmo de eliminação de Gauss é necessário que o pivô seja um elemento não nulo, ou seja,  $a_{kk}^{(k)} \neq 0$ . Se no processo de eliminação um dos pivôs se anular ( $a_{kk}^{(k)} = 0$ ), deve-se efetuar a troca da linha  $k$  por outra linha que esteja abaixo dela, de modo que o novo candidato a pivô seja um elemento não nulo. Assim, para dar continuidade ao método deve-se trocar a linha  $k$  por qualquer linha  $r$ , em que  $k < r \leq n$ , de tal forma que o novo  $a_{kk}^{(k)} \neq 0$ .

**Exemplo 8.** Resolva o sistema linear

$$\begin{cases} 3x_1 + 2x_2 + 3x_3 = 3 \\ 9x_1 + 8x_2 - 8x_3 = 6 \\ -6x_1 + 4x_2 + 3x_3 = 0 \end{cases}$$

utilizando método de eliminação de Gauss. Efetue os cálculos retendo quatro casas decimais.

**Solução:** Seja a matriz aumentada do sistema linear é  $[A|B] = \begin{bmatrix} 3 & 2 & -3 & | & 3 \\ 9 & 8 & -8 & | & 6 \\ -6 & 4 & 3 & | & 0 \end{bmatrix}$ . Aplicando a etapa de eliminação de Gauss, tem-se:

**Passo 1 - Eliminação na primeira coluna:** dado que  $a_{11}^{(1)} \neq 0$ :

- **Pivô:**  $a_{11}^{(1)} = 3$

$$m_{21} = \frac{a_{21}^{(1)}}{a_{11}^{(1)}} = \frac{9}{3} = 3$$

- **Cálculo dos multiplicadores:**

$$m_{31} = \frac{a_{31}^{(1)}}{a_{11}^{(1)}} = \frac{-6}{3} = -2$$

- **Eliminação da incógnita  $x_1$  nas linhas 2 e 3:** operações elementares  $L_2^{(2)} \leftarrow L_2^{(1)} - 3L_1^{(1)}$   
 $L_3^{(2)} \leftarrow L_3^{(1)} - (-2)L_1^{(1)}$

- **Matriz aumentada resultante:**  $\begin{bmatrix} 3 & 2 & -3 & | & 3 \\ 0 & 2 & 1 & | & -3 \\ 0 & 8 & -3 & | & 6 \end{bmatrix}$



**Passo 2 - Eliminação na segunda coluna:** dado que  $a_{22}^{(2)} \neq 0$ , logo:

- **Pivô:**  $a_{22}^{(2)} = 2$
- **Cálculo de multiplicadores:**  $m_{32} = \frac{a_{32}^{(2)}}{a_{22}^{(2)}} = \frac{8}{2} = 4$
- **Eliminação da incógnita  $x_2$  na linha 3:** operação elementar  $L_3^{(3)} \leftarrow L_3^{(2)} - 4L_2^{(2)}$
- **Matriz aumentada resultante:** 
$$\left[ \begin{array}{ccc|c} 3 & 2 & -3 & 3 \\ 0 & 2 & 1 & -3 \\ 0 & 0 & -7 & 18 \end{array} \right]$$

Da etapa de eliminação, obtém-se o seguinte sistema triangular equivalente ao sistema linear original:

$$\begin{cases} 3x_1 + 2x_2 + 3x_3 = 3 \\ \phantom{3x_1} 2x_2 + x_3 = -3 \\ \phantom{3x_1} \phantom{2x_2} -7x_3 = 18 \end{cases}$$

que, resolvido por substituições retroativas, tem-se:

$$x_3 = \frac{d_3}{u_{33}} = \frac{18}{-7} = -2,5714$$

$$x_2 = \frac{d_2 - (u_{23}x_3)}{u_{22}} = \frac{-3 - [(1)(-2,5714)]}{2} = -0,2143$$

$$x_1 = \frac{d_1 - (u_{12}x_2 + u_{13}x_3)}{u_{11}} = \frac{3 - [(2)(-0,2143) + (-3)(-2,5714)]}{3} = -1,4285$$

Portanto, a solução do sistema linear é  $x = \begin{bmatrix} -1,4285 & -0,2143 & -2,5714 \end{bmatrix}^t$ .

**Exemplo 9.** Resolva o sistema linear

$$\begin{cases} 3x_1 + 2x_2 + \phantom{-3x_3} + x_4 = 3 \\ 9x_1 + 8x_2 - 3x_3 + 4x_4 = 6 \\ -6x_1 + 4x_2 - 8x_3 \phantom{+4x_4} = -16 \\ 3x_1 - 8x_2 + 3x_3 - 8x_4 = 22 \end{cases}$$

utilizando método de eliminação de Gauss. Efetue os cálculos retendo quatro casas decimais.

**Solução:** Este sistema será resolvido utilizando uma tabela (ou **dispositivo prático**) para sumarizar os cálculos.

Linhas	Multiplicadores	Coeficientes	T. indep.	Operações
$L_1^{(1)}$	<b>pivô</b> ( $a_{11} \neq 0$ )	<u>3</u> <u>2</u> <u>0</u> <u>1</u>	<u>3</u>	
$L_2^{(1)}$	$m_{21} = \frac{a_{21}^{(1)}}{a_{11}^{(1)}} = \frac{9}{3} = 3$	9   8   -3   4	6	
$L_3^{(1)}$	$m_{31} = \frac{a_{31}^{(1)}}{a_{11}^{(1)}} = \frac{-6}{3} = -2$	-6   4   -8   0	-16	
$L_4^{(1)}$	$m_{41} = \frac{a_{41}^{(1)}}{a_{11}^{(1)}} = \frac{3}{3} = 1$	3   -8   3   -8	22	
$L_2^{(2)}$	<b>pivô</b> ( $a_{22} \neq 0$ )	<u>0</u> <u>2</u> <u>-3</u> <u>1</u>	<u>-3</u>	$L_2^{(2)} = L_2^{(1)} - 3L_1^{(1)}$
$L_3^{(2)}$	$m_{32} = \frac{a_{32}^{(2)}}{a_{22}^{(2)}} = \frac{8}{2} = 4$	0   8   -8   2	-10	$L_3^{(2)} = L_3^{(1)} + 2L_1^{(1)}$
$L_4^{(2)}$	$m_{42} = \frac{a_{42}^{(2)}}{a_{22}^{(2)}} = \frac{-10}{2} = -5$	0   -10   3   -9	19	$L_4^{(2)} = L_4^{(1)} - L_1^{(1)}$
$L_3^{(3)}$	<b>pivô</b> ( $a_{33} \neq 0$ )	<u>0</u> <u>0</u> <u>4</u> <u>-2</u>	<u>2</u>	$L_3^{(3)} = L_3^{(2)} - 4L_2^{(2)}$
$L_4^{(3)}$	$m_{43} = \frac{a_{43}^{(3)}}{a_{33}^{(3)}} = \frac{-12}{4} = -3$	0   0   -12   -4	4	$L_4^{(3)} = L_4^{(2)} + 5L_2^{(2)}$
$L_4^{(4)}$		<u>0</u> <u>0</u> <u>0</u> <u>-10</u>	<u>10</u>	$L_4^{(4)} = L_4^{(3)} + 3L_3^{(3)}$

Tabela 4.1: Dispositivo prático com a sumarização dos cálculos da etapa de eliminação.

O sistema triangular resultante da etapa de eliminação, que encontra-se em destaque no dispositivo prático (Tabela 4.1) é:

$$\begin{cases} 3x_1 & +2x_2 & & +x_4 & = 3 \\ & 2x_2 & -3x_3 & +x_4 & = -3 \\ & & 4x_3 & -2x_4 & = 2 \\ & & & -10x_4 & = 10 \end{cases}$$

Aplicando o algoritmo de substituições retroativas, tem-se:

$$x_4 = \frac{d_4}{u_{44}} = \frac{10}{-10} = -1$$

$$x_3 = \frac{d_3 - u_{34}x_4}{u_{33}} = \frac{2 - (-2)(-1)}{4} = 0$$

$$x_2 = \frac{d_2 - (u_{23}x_3 + u_{24}x_4)}{u_{22}} = \frac{-3 - [(-3)(0) + (1)(-1)]}{2} = -1$$

$$x_1 = \frac{d_1 - (u_{12}x_2 + u_{13}x_3 + u_{14}x_4)}{u_{11}} = \frac{3 - [(2)(-1) + (0)(0) + (1)(-1)]}{3} = 2$$

Portanto, a solução do sistema linear original é  $x = \begin{bmatrix} 2 & -1 & 0 & -1 \end{bmatrix}^t$ .

#### 4.2.2 Avaliação do Resíduo

O **resíduo** (erro) produzido pela solução do sistema  $AX = B$  pode ser avaliado pela expressão:

$$\xi = \max_{1 \leq i \leq n} |r_i|$$

onde  $r_i$ ,  $i = 1, 2, \dots, n$  é a  $i$ -ésima componente do vetor resíduo  $R$ , calculado por:

$$R = B - AX$$

**Exemplo 10.** Resolva o sistema linear utilizando o método de eliminação de Gauss e retendo nos cálculos três casas decimais.

$$\begin{cases} 4,5x_1 + 1,8x_2 + 2,4x_3 = 19,62 \\ 3x_1 + 5,2x_2 + 1,2x_3 = 12,36 \\ 0,8x_1 + 2,4x_2 + 3,6x_3 = 9,2 \end{cases}$$

**Solução:** os cálculos da etapa de eliminação estão apresentados na tabela abaixo:

Linhas	Multiplicadores	Coeficientes			T. indep.	Operações
$L_1^{(1)}$		4,5	1,8	2,4	19,62	
$L_2^{(1)}$	$m_{21} = \frac{3}{4,5} = 0,667$	3,0	5,2	1,2	12,36	
$L_3^{(1)}$	$m_{31} = \frac{0,8}{4,5} = 0,178$	0,8	2,4	3,6	9,20	
$L_2^{(2)}$		0	3,999	-0,401	-0,727	$L_2^{(2)} = L_2^{(1)} - 0,667L_1^{(1)}$
$L_3^{(2)}$	$m_{32} = \frac{2,080}{3,999} = 0,520$	0	2,080	3,173	5,708	$L_3^{(2)} = L_3^{(1)} - 0,178L_1^{(1)}$
$L_3^{(3)}$		0	0	3,382	6,086	$L_3^{(3)} = L_3^{(2)} - 0,520L_2^{(2)}$

#### Observação:

Quando aplicada a transformação elementar para a eliminação do termo na posição linha dois coluna um (termo  $a_{21}$ ), o cálculo realizado  $(3,0 + (-0,667) \times (4,5))$  produz o resultado  $(-0,0015)$  que, considerando três casas decimais, vai a  $(-0,002)$  e não 0, como desejado. Este reflexo na eliminação foi causado pelo erro de arredondamento gerado no cálculo do multiplicador  $m_{21}$ . Como, no final terá utilidade apenas a parte triangular superior da matriz dos coeficientes, então, nas posições nas quais devem ocorrer a eliminação, os cálculos podem "deixar de serem feitos". Este procedimento é interessante porque reduz o esforço computacional.

O sistema triangular superior resultante da etapa de eliminação é:

$$\begin{cases} 4,5x_1 + 1,8x_2 + 2,4x_3 = 19,62 \\ 3,999x_2 - 0,401x_3 = -0,727 \\ 3,382x_3 = 6,086 \end{cases}$$

que, aplicando o algoritmo de substituições retroativas no sistema linear acima, resulta na seguinte solução  $x = [3,400 \quad -0,001 \quad 1,800]^t$ . O resíduo, calculado pela expressão  $R = B - AX$ , é:

$$R = \begin{bmatrix} 19,62 \\ 12,36 \\ 9,20 \end{bmatrix} - \begin{bmatrix} 4,5 & 1,8 & 2,4 \\ 3,0 & 5,2 & 1,2 \\ 0,8 & 2,4 & 3,6 \end{bmatrix} \cdot \begin{bmatrix} 3,400 \\ -0,001 \\ 1,800 \end{bmatrix} = \begin{bmatrix} 0,0018 \\ 0,0052 \\ 0,0024 \end{bmatrix}$$

e o erro cometido:

$$\xi = \max_{1 \leq i \leq n} |r_i| = \max_{1 \leq i \leq 3} \{|0,0018|, |0,0052|, |0,0024|\} = 0,0052.$$

Portanto, a solução do sistema linear é  $x = [3,400 \quad -0,001 \quad 1,800]^t$  com erro  $\xi = 0,0052$ .

## 4.3 Exercícios



**E. 1.** Resolva o sistema linear pelo método de eliminação de Gauss retendo nos cálculos quatro casas decimais.

$$\begin{bmatrix} -2 & 3 & 1 & 5 \\ 5 & 1 & -1 & 0 \\ 1 & 6 & 3 & -1 \\ 4 & 5 & 2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 6 \end{bmatrix}$$

**E. 2.** Resolva os sistemas lineares pelo método de Eliminação de Gauss. O que é possível afirmar em relação a quantidade de soluções desses sistemas?

a)  $\begin{bmatrix} -1 & -1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \\ 4 \end{bmatrix}$

b)  $\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$





A series of horizontal lines for writing, starting from the first line below the icon and continuing down to the bottom of the page.







## Eliminação de Gauss com Pivotação

### Sumário da Aula

5.1	Método de Eliminação de Gauss com Pivotação . . . . .	43
5.1.1	Pivotação Parcial . . . . .	44
5.2	Exercícios . . . . .	46

### 5.1 Método de Eliminação de Gauss com Pivotação

Conforme visto na seção anterior, o método da eliminação de Gauss requer na fase de eliminação o cálculo dos multiplicadores. Este fato pode ocasionar problemas se o pivô estiver próximo de zero ou for nulo. Isto porque trabalhar com pivô nulo é impossível, pois ocorrerá uma divisão por zero ao calcular o multiplicador  $m_{ij}$ , e o pivô próximo de zero pode conduzir a resultados imprecisos, visto que tal fato dá origem a multiplicadores bem maiores do que a unidade o que, por sua vez, provoca uma ampliação dos erros de arredondamento.

A **ampliação de erros de arredondamento** ocorre quando se multiplica um número muito grande por outro que já contém erro de arredondamento. Por exemplo, admita-se que um número  $x'$  tenha erro de arredondamento  $\xi$ . Este número pode, então, ser escrito na forma:

$$x' = x + \xi.$$

Se  $x'$  é multiplicado por um valor  $m$ , tem-se que  $mx' = mx + m\xi$ , de forma que  $m\xi$  representa o erro de



arredondamento. Se  $m$  for grande este erro pode ser muito maior que o erro original. Diz-se, então, que o erro  $\xi$  foi amplificado.

Para contornar este problema, ou seja, para minimizar o efeito dos erros de arredondamento e assegurar a estabilidade numérica no método de eliminação de Gauss será adotada uma **estratégia de pivotação**, que é um processo de **escolha do pivô**.

### 5.1.1 Pivotação Parcial

Será considerado a estratégia de **pivotação parcial** para a escolha do pivô. A estratégia consiste em escolher como pivô o maior elemento em módulo da coluna, cujos elementos serão eliminados. Em termos gerais:

- (i) no passo  $k$  da fase de eliminação, o pivô será o elemento de **maior valor absoluto dentre os coeficientes**  $a_{ik}^{(k)}$  da coluna em que está sendo aplicada a eliminação:

$$a_{rk}^{(k)} = \max_{k \leq i \leq n} |a_{ik}^{(k)}|$$

$r$  é a linha em que se encontra o elemento máximo.

- (ii) se necessário, **efetuar a troca da linha  $k$  pela linha  $r$** , ou seja  $a_{kk}^{(k)} = a_{rk}^{(k)}$ .

**Exemplo 11.** Resolva o sistema linear utilizando método de eliminação de Gauss com pivotação retendo nos cálculos quatro casas decimais:

$$\begin{cases} 2x_1 - 5x_2 + 3x_3 + x_4 = 5 \\ 3x_1 - 7x_2 + 3x_3 - x_4 = -1 \\ 5x_1 - 9x_2 + 6x_3 + 2x_4 = 7 \\ 4x_1 - 6x_2 + 3x_3 + x_4 = 8 \end{cases}$$

**Solução:** Resolvendo o sistema utilizando o **dispositivo prático** para os cálculos:

Linhas	Multiplicadores	Coeficientes				T. indep.	Operações
$L_1^{(1)}$	<b>Maior valor absoluto <math>\Rightarrow</math></b>	2	-5	3	1	5	
$L_2^{(1)}$		3	-7	3	-1	-1	
$L_3^{(1)}$		<b>5</b>	-9	6	2	7	
$L_4^{(1)}$		4	-6	3	1	8	
$L_1^{(2)}$	<b>pivô <math>\Rightarrow</math></b>	<b>5</b>	-9	6	2	7	$L_1^{(2)} = L_3^{(1)}$
$L_2^{(2)}$	$m_{21} = \frac{3}{5} = 0,6$	3	-7	3	-1	-1	
$L_3^{(2)}$	$m_{31} = \frac{2}{5} = 0,4$	2	-5	3	1	5	$L_3^{(2)} = L_1^{(3)}$
$L_4^{(2)}$	$m_{41} = \frac{4}{5} = 0,8$	4	-6	3	1	8	
$L_2^{(3)}$	<b>pivô <math>\Rightarrow</math></b>	0	<b>-1,6</b>	-0,6	-2,2	-5,2	$L_2^{(3)} = L_2^{(2)} - 0,6L_1^{(2)}$
$L_3^{(3)}$	$m_{32} = \frac{-1,4}{-1,6} = 0,875$	0	-1,4	0,6	0,2	2,2	$L_3^{(3)} = L_3^{(2)} - 0,4L_1^{(2)}$
$L_4^{(3)}$	$m_{42} = \frac{1,2}{-1,6} = -0,75$	0	1,2	-1,8	-0,6	2,4	$L_4^{(3)} = L_4^{(2)} - 0,8L_1^{(2)}$
$L_3^{(4)}$	<b>Maior valor absoluto <math>\Rightarrow</math></b>	0	0	1,125	2,125	6,75	$L_3^{(4)} = L_3^{(3)} - 0,875L_2^{(3)}$
$L_4^{(4)}$		0	0	<b>-2,25</b>	-2,25	-1,5	$L_4^{(3)} = L_4^{(3)} + 0,75L_2^{(3)}$
$L_3^{(5)}$	<b>pivô <math>\Rightarrow</math></b>	0	0	<b>-2,25</b>	-2,25	-1,5	$L_3^{(5)} = L_4^{(4)}$
$L_4^{(5)}$	$m_{43} = \frac{1,125}{-2,25} = -0,5$	0	0	1,125	2,125	6,75	$L_4^{(5)} = L_3^{(4)}$
$L_4^{(6)}$		0	0	0	<b>1</b>	6	$L_4^{(4)} = L_4^{(3)} + 0,5L_3^{(3)}$

O sistema triangular resultante da etapa de eliminação é:

$$\begin{cases} 5x_1 & -9x_2 & +6x_3 & +2x_4 & = 7 \\ & -1,6x_2 & -0,6x_3 & -2,2x_4 & = -5,2 \\ & & -2,25x_3 & -2,25x_4 & = -1,5 \\ & & & x_4 & = 6 \end{cases}$$

Resolvendo o sistema triangular por substituições retroativas, temos:

$$x_4 = \frac{d_4}{u_{44}} = \frac{6}{1} = 6$$

$$x_3 = \frac{d_3 - u_{34}x_4}{u_{33}} = \frac{-1,5 - (-2,25)(6)}{-2,25} = -5,333$$

$$x_2 = \frac{d_2 - (u_{23}x_3 + u_{24}x_4)}{u_{22}} = \frac{-5,2 - [(-0,6)(-5,333) + (-2,2)(6)]}{-1,6} = -3$$

$$x_1 = \frac{d_1 - (u_{12}x_2 + u_{13}x_3 + u_{14}x_4)}{u_{11}} = \frac{5 - [(9)(-3) + (6)(-5,333) + (2)(6)]}{5} = 0$$

Logo, a solução é  $x = \begin{bmatrix} 0 & -3 & -5,333 & 6 \end{bmatrix}^t$ .

O pseudocódigo do método de eliminação com pivotação está ilustrado no Algoritmo 3.

```

1: Entrada:  $A, b, n$ .
2: para  $(k = 1)$  até  $(n - 1)$  faça
3:   Encontre  $r \geq k$  tal que  $A(r, k) = \max_{k \leq i \leq n} \text{abs}(A(i, j))$ ; {Escolha do elemento pivô}
4:   Troque a linha  $k$  com a linha  $r$ ;
5:   para  $(i = k + 1)$  até  $n$  faça
6:      $m(i, k) = -A(i, k)/A(k, k)$ ;
7:     para  $(j = k)$  até  $n$  faça
8:        $A(i, j) = A(i, j) + m(i, k)A(k, j)$ ;
9:     fim para
10:     $b(i) = b(i) + m(i, k)b(k)$ ;
11:  fim para
12: fim para
13:  $U = A$ ;  $d = b$ 
14: Algoritmo Substituições Retroativas( $U, d$ );
15: Saída: vetor solução  $x$ .

```

Algoritmo 3: Eliminação de Gauss com pivotação parcial. (abs: valor absoluto).

A Tabela 7.1 apresenta o resumo do esforço computacional do Algoritmo de Eliminação de Gauss.

Operações	Complexidade
Adições	$\frac{1}{3}n^3 + \frac{2}{3}n$
Multiplicações	$\frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n$
Divisões	$n - 1$

Tabela 5.1: Complexidade da eliminação de Gauss considerando um sistema linear de ordem  $n$ .  
Fonte: [1].

## 5.2 Exercícios



E. 1. Resolva o sistema linear pelo método de eliminação de Gauss com pivotação e verifique a exatidão da solução encontrada através do cálculo do resíduo. Retenha nos cálculos quatro casas decimais.

$$\begin{cases} x_1 & +4x_2 & +52x_3 & = 57 \\ 22x_1 & +110x_2 & -3x_3 & = 134 \\ 13x_1 & +2x_2 & +14x_3 & = 38 \end{cases}$$





A series of horizontal lines for writing, starting from the first line below the icon and continuing down to the bottom of the page.





## Fatoração LU

### Sumário da Aula

6.1 Fatoração LU da matriz A . . . . .	50
6.2 Exercícios . . . . .	53

Em muitas situações, é necessário resolver vários sistemas de equações lineares que possuem em comum a matriz dos coeficientes e termos independentes diferentes, ou seja, casos em que:

$$Ax = B_i, i = 1, 2, \dots, m$$

Nestes casos, indica-se resolver os sistemas por meio de uma técnica de **fatoração da matriz A**. A técnica consiste em decompor a matriz dos coeficientes em um produto de dois ou mais fatores e, em seguida, resolver uma sequência de sistemas de equações lineares que conduzirá à solução do sistema original. A vantagem da utilização da técnica de fatoração é que a resolução de qualquer sistema de equações lineares que tenha  $A$  como matriz dos coeficientes é quase imediata, sendo necessário alterar apenas o vetor  $B$ .



Dentre as técnicas de fatoração mais utilizadas, destaca-se a **Decomposição LU**. Nesta técnica, a matriz  $A$  é decomposta no produto de duas matrizes  $L$  e  $U$  ( $A = LU$ ), onde  $L$  é uma matriz triangular inferior e  $U$  uma matriz triangular superior:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ & & \ddots & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}}_A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ & & \ddots & & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix}}_L \cdot \underbrace{\begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ & & \ddots & & \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}}_U$$

## 6.1 Fatoração LU da matriz A

O Teorema 6.1 apresenta em que condições é possível decompor uma matriz quadrada  $A = (a_{ij})$  no produto de uma matriz triangular inferior por uma matriz triangular superior.

**Definição 6.1.** Chamamos de “menores principais” de ordem  $k$  de uma matriz  $A = (a_{ij})$  o  $\Delta_k = \det(A_k)$ , onde  $A_k = (a_{ij})$  é formada pelas  $k$  primeiras linhas e  $k$  primeiras colunas da matriz  $A$ .

**Teorema 6.1. (Teorema LU)** Sejam  $A = (a_{ij})$  uma matriz quadrada de ordem  $n$ . Se os menores principais de  $A$ ,  $\Delta_i \neq 0$ ,  $i = 1, 2, \dots, n-1$ . Então,  $A$  se decompõe de forma única, no produto de uma triangular inferior  $L = (l_{ij})$ , com  $l_{11} = l_{22} = l_{33} = \dots = l_{nn} = 1$ , e uma única matriz triangular superior  $U = (u_{ij})$ , tal que  $LU = A$ .

A **fatoração LU da matriz A** nos fatores  $L$  e  $U$  pode ser feita por meio das seguintes fórmulas gerais:

$$\begin{cases} u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}, & i \leq j \\ l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj}}{u_{jj}}, & i > j \end{cases}$$

ou, de forma similar, pelo Algoritmo 4 de Fatoração LU.

```

1: para  $m = 1$  to  $n$  faça
2:   para  $j = m, m+1$  to  $n$  faça
3:      $u_{mj} = a_{mj} - \sum_{k=1}^{m-1} l_{mk} u_{kj}$ ;
4:   fim para
5:   para  $i = m+1$  to  $n$  faça
6:      $l_{mm} = 1$ 
7:      $l_{im} = \frac{a_{im} - \sum_{k=1}^{m-1} l_{ik} u_{km}}{u_{mm}}$ ;
8:   fim para
9: fim para
10: para  $m = n$  faça
11:    $l_{nn} = 1$ 
12:    $u_{nn} = a_{nn} - \sum_{k=1}^{n-1} l_{nk} u_{kn}$ ;
13: fim para

```

Algoritmo 4: Fatoração LU

A fatoraão de uma matriz no produto  $LU$  onde  $L$  tem 1 na diagonal principal tamb m   conhecida como **M todo de Doolittle**.

**Exemplo 12.** Use o algoritmo 4 para decompor a matriz  $A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & 0 \\ 3 & 1 & 2 \end{bmatrix}$  no produto  $LU$ .

**Solu o:** Para que a matriz  $A$  satisfaa as condi es da decomposi o LU, conforme o Teorema 6.1   necess rio que  $\det(A_1) \neq 0$  e  $\det(A_2) \neq 0$ . Temos que  $\det(A_1) = 2 \neq 0$  e  $\det(A_2) = [(2 \times 2) - (1 \times 0)] = 4 \neq 0$ . Logo  $A$  satisfaz o Teorema e atende as condi es de Decomposi o LU.

Aplicando o Algoritmo 4 temos:

$(m = 1)$

Linha 1 da matriz  $U$ :

$$\begin{aligned} j = 1 & \Rightarrow u_{11} = a_{11} - \sum_{k=1}^0 l_{1k} u_{k1} = a_{11} = 2 \\ j = 2 & \Rightarrow u_{12} = a_{12} - \sum_{k=1}^0 l_{1k} u_{k2} = a_{12} = 0 \\ j = 3 & \Rightarrow u_{13} = a_{13} - \sum_{k=1}^0 l_{1k} u_{k3} = a_{13} = 1 \end{aligned}$$

Coluna 1 da matriz  $L$ :

$$\begin{aligned} i = 1 & \Rightarrow l_{11} = 1 \\ i = 2 & \Rightarrow l_{21} = \frac{a_{21} - \sum_{k=1}^0 l_{2k} u_{k1}}{u_{11}} = \frac{a_{21}}{u_{11}} = \frac{1}{2} \\ i = 3 & \Rightarrow l_{31} = \frac{a_{31} - \sum_{k=1}^0 l_{3k} u_{k1}}{u_{11}} = \frac{a_{31}}{u_{11}} = \frac{3}{2} \end{aligned}$$

$(m = 2)$

Linha 2 da matriz  $U$ :

$$\begin{aligned} j = 2 & \Rightarrow u_{22} = a_{22} - \sum_{k=1}^1 l_{2k} u_{k2} = a_{12} - [l_{21} u_{12}] = 2 - \left(\frac{1}{2}(0)\right) = 2 \\ j = 3 & \Rightarrow u_{23} = a_{23} - \sum_{k=1}^1 l_{2k} u_{k3} = a_{13} - [l_{21} u_{13}] = 0 - \left(\frac{1}{2}(1)\right) = -1/2 \end{aligned}$$

Coluna 2 da matriz  $L$ :

$$\begin{aligned} i = 2 & \Rightarrow l_{22} = 1 \\ i = 3 & \Rightarrow l_{32} = \frac{a_{32} - \sum_{k=1}^1 l_{3k} u_{k2}}{u_{22}} = \frac{a_{32} - [l_{31} u_{12}]}{u_{22}} = \frac{1 - (3/2)(0)}{2} = 1/2 \end{aligned}$$

$(m = n = 3)$

Linha 3 da matriz  $U$ :

$$j = 3 \Rightarrow u_{33} = a_{33} - \sum_{k=1}^2 l_{3k} u_{k3} = a_{13} - [l_{31} u_{13} + l_{32} u_{23}] = 2 - \left[ \frac{3}{2}(1) + \frac{1}{2} \frac{-1}{2} \right] = \frac{3}{4}$$

Coluna 3 da matriz  $L$ :

$$i = 3 \Rightarrow l_{33} = 1$$



Logo, as matrizes  $L$  e  $U$  resultantes da etapa de fatoração são:

$$U = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -1/2 \\ 0 & 0 & 3/4 \end{bmatrix} \quad \text{e} \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & 1/2 & 1 \end{bmatrix}$$

A fatoração da matriz  $A$  no produto das matrizes  $LU$  também pode ser obtida através da **fase de eliminação do método de eliminação de Gauss**. Veja o exemplo seguinte.

**Exemplo 13.** Seja a matriz

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

Aplicando a fase de eliminação de Gauss em  $A$ , tem-se:

Linhas	Multiplicadores	Coefficientes	Operações
$L_1^{(1)}$		2   0   1	
$L_2^{(1)}$	$m_{21} = \frac{1}{2} = 0,5$	1   2   0	
$L_3^{(1)}$	$m_{31} = \frac{3}{2} = 1,5$	3   1   2	
$L_2^{(2)}$		$m_{21}$ 2   -1/2	$L_2^{(2)} = L_2^{(1)} - 0,5L_1^{(1)}$
$L_3^{(2)}$	$m_{32} = \frac{1}{2} = 0,5$	$m_{31}$ 1   1/2	$L_3^{(2)} = L_3^{(1)} - 3/2L_1^{(1)}$
$L_3^{(3)}$		$m_{31}$ $m_{32}$ 3/4	$L_3^{(3)} = L_3^{(2)} - 1/2L_2^{(2)}$

Para **escrever as matrizes  $L$  e  $U$**  da fatoração  $LU$  da matriz  $A$ , basta **analisar as linhas marcadas em azul** da tabela. A matriz  $U$  é a matriz triangular superior resultante da etapa de eliminação de Gauss:

$$U = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -1/2 \\ 0 & 0 & 3/4 \end{bmatrix}$$

e  $L$  é uma matriz triangular inferior, na qual os elementos da diagonal principal são unitários e, abaixo da diagonal principal, encontram-se os multiplicadores da etapa  $k$  da fase de eliminação:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & 1/2 & 1 \end{bmatrix}$$

**Observação:**

Vale ressaltar que a decomposição  $LU$  também fornece um dos algoritmos mais eficientes para o cálculo do determinante de uma matriz.

## Cálculo do Determinante

Se a matriz  $A$  puder ser decomposta como produto de dois fatores  $L$  e  $U$ , onde  $L$  é uma matriz triangular inferior com elementos diagonais unitários e  $U$  uma matriz triangular superior, então o **determinante de  $A$**

pode ser calculado por:

$$\det(A) = \det(LU) = \det(L) \times \det(U)$$

Sabe-se que o **determinante de uma matriz triangular é igual ao produto dos elementos da sua diagonal principal**, assim,  $\det(L) = 1$  e

$$\det(U) = \prod_{i=1}^n u_{ii} = u_{11}u_{22} \cdots u_{nn}$$

Ao **utilizar o procedimento de pivotação parcial** na etapa de eliminação de Gauss, para o cálculo do determinante, deve-se levar em consideração o número total de trocas de linhas realizadas. Assim, o cálculo do determinante da matriz  $U$  deverá ser feito como:

$$\det(U) = (-1)^k \prod_{i=1}^n u_{ii}$$

sendo  $k$  o número mínimo de trocas de linhas necessárias durante a fase de eliminação. **Assim, o determinante da matriz  $A$  é:**

$$\det(A) = \det(L) \times \det(U) = (-1)^k \prod_{i=1}^n u_{ii}$$

**Exemplo 14.** Para o cálculo do determinante da matriz  $A$  do Exemplo 13, basta calcular o determinante das matrizes triangulares  $L$  e  $U$  encontradas.

Assim, sendo:

$$\det(L) = 1 \quad \text{e} \quad \det(U) = (-1)^0 \prod_{i=1}^3 u_{ii} = 2 \times 2 \times \frac{3}{4} = 1 \times 3 = 3.$$

obtem-se que  $\det(A) = \det(L) \times \det(U) = 1 \times 3 = 3$ .

## 6.2 Exercícios



E. 1. Seja

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

- Verifique se  $A$  satisfaz as condições de decomposição LU (Teorema 6.1).
- Decompor  $A$  em LU.
- Através da decomposição LU, calcule o determinante de  $A$ .





A series of horizontal lines for writing, starting from the first line below the icon and continuing down to the bottom of the page.





# Decomposição LU

## Sumário da Aula

7.1	Decomposição LU . . . . .	57
7.2	Decomposição LU com pivotação parcial . . . . .	59
7.3	Algoritmo e Complexidade . . . . .	61
7.4	Exercícios . . . . .	62

## 7.1 Decomposição LU

A decomposição LU também pode ser aplicada na resolução de sistemas de equações lineares. Para tanto, considere um sistema linear  $Ax = B$ , cuja matriz dos coeficientes é uma matriz não singular ( $\det(A) \neq 0$ ) e satisfaz às condições do Teorema 6.1.

Para resolver o sistema linear  $Ax = B$ , utilizando decomposição LU, basta executar a seguinte sequência de passos:

- (i) Obtenção da fatoração  $LU$  da matriz  $A$ . O sistema  $Ax = B$  será reescrito como  $LUx = B$ .
- (ii) Tomando  $Ux = y$ , obtém-se  $Ly = B$ .  
**O cálculo do sistema  $Ax = B$  é então substituído pela resolução de dois sistemas triangulares:**
- (iii) Resolve-se o sistema triangular inferior  $Ly = B$  por meio de substituições sucessivas.
- (iv) Resolve-se o sistema triangular superior  $Ux = y$  por meio de substituições retroativas, obtendo, então, a solução do sistema de equações  $Ax = B$ .

**Exemplo 15.** Resolva o sistema linear utilizando o método de decomposição  $LU$ .

$$\begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & 0 \\ 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

Do exemplo 13, temos que a fatora  o  $LU$  da matriz  $A$    :

$$U = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -1/2 \\ 0 & 0 & 3/4 \end{bmatrix} \quad \text{e} \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & 1/2 & 1 \end{bmatrix}$$

Resolvendo os sistemas triangulares formados:

- $Ly = B$  (**sistema triangular inferior**)

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

Resolvendo por substitui  es sucessivas, obt  m-se  $y = \begin{bmatrix} 2 & 2 & -3 \end{bmatrix}^t$

- $Ux = y$  (**sistema superior**)

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -1/2 \\ 0 & 0 & 3/4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ -3 \end{bmatrix}$$

Resolvendo por substitui  es retroativas, obt  m-se  $x = \begin{bmatrix} 3 & 0 & 2 \end{bmatrix}^t$

Portanto, a solu  o do sistema linear     $x = \begin{bmatrix} 3 & 0 & 2 \end{bmatrix}^t$ .

**Exemplo 16.** Resolva o sistema de equa  es lineares pelo m  todo de Decomposi  o  $LU$ :

$$\begin{cases} x_1 - 3x_2 + 2x_3 = 11 \\ -2x_1 + 8x_2 - x_3 = -15 \\ 4x_1 - 6x_2 + 5x_3 = 29 \end{cases}$$

Os c  lculos da etapa de elimina  o de Gauss na matriz  $A$  do sistema linear est  o sumarizados na tabela seguinte:

Linhas	Multiplicadores	Coeficientes	Opera��es
$L_1^{(1)}$		<b>1</b> <b>-3</b> <b>2</b>	
$L_2^{(1)}$	$m_{21} = -2$	-2   8   -1	
$L_3^{(1)}$	$m_{31} = 4$	4   -6   5	
$L_2^{(2)}$		$m_{21}$ <b>2</b> <b>3</b>	$L_2^{(2)} = L_2^{(1)} + 2L_1^{(1)}$
$L_3^{(2)}$	$m_{32} = 3$	$m_{31}$ 6   -3	$L_3^{(2)} = L_3^{(1)} - 4L_1^{(1)}$
$L_3^{(3)}$		$m_{31}$ $m_{32}$ <b>-12</b>	$L_3^{(3)} = L_3^{(2)} - 3L_2^{(2)}$

Na tabela, os elementos  $L_i^{(k)}$  da primeira coluna indicam as linhas  $i$  da matriz  $A$  em cada passo  $(k)$  da etapa de elimina  o de Gauss e n  o a matriz triangular inferior  $L$ .

Da tabela, a fatoração  $LU$  da matriz  $A$  é dada por:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 1 & -3 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & -12 \end{bmatrix}$$

O sistema triangular inferior,  $Ly = B$ , será resolvido por substituições sucessivas:

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 11 \\ -15 \\ 29 \end{bmatrix} \Rightarrow y = \begin{bmatrix} 11 \\ 7 \\ -36 \end{bmatrix}$$

E o sistema triangular superior,  $Ux = y$ , será resolvido por substituições retroativas:

$$\begin{bmatrix} 1 & -3 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & -12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 11 \\ 7 \\ -36 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

Portanto, a solução do sistema linear é  $x = \begin{bmatrix} 2 & -1 & 3 \end{bmatrix}^t$ .



## 7.2 Decomposição LU com pivotação parcial

A estratégia de pivotação parcial também pode ser usada no método da Decomposição LU. Para tanto, faz-se necessário utilizar um vetor de permutação,  $P$ , que é gerado atribuindo-se um número de ordem a cada equação que compõe o sistema.

**Exemplo 17.** Resolva o sistema de equações lineares utilizando o Método da Decomposição LU com pivotação parcial e considerando, quando for o caso, duas casas decimais:

$$\begin{cases} 4x_1 & -x_2 & & -x_4 & = & 6 \\ x_1 & -2x_2 & +x_3 & & = & 8 \\ & 4x_2 & -4x_3 & +x_4 & = & -7 \\ 5x_1 & & +5x_3 & -10x_4 & = & -40 \end{cases}$$

O vetor de permutação é  $P = \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}^t$  e os cálculos da fase de eliminação para fazer a fatoração  $LU$  de  $A$  estão sumarizados na tabela seguinte:



Linhas	Multiplicadores	Coeficientes				P	Operações
$L_1^{(1)}$	<b>pivô</b>	4	-1	0	-1	1	$L_1$
$L_2^{(1)}$		1	-2	1	0	2	$L_2$
$L_3^{(1)}$		0	4	-4	1	3	$L_3$
$L_4^{(1)}$		<b>5</b>	0	5	-10	4	$L_4$
$L_1^{(2)}$	$m_{21} = 0,2$ $m_{31} = 0$ $m_{41} = 0,8$	<b>5</b>	<b>0</b>	<b>5</b>	<b>-10</b>	<b>4</b>	$L_4^{(1)}$
$L_2^{(2)}$		1	-2	1	0	2	$L_2^{(1)}$
$L_3^{(2)}$		0	4	-4	1	3	$L_3^{(1)}$
$L_4^{(2)}$		4	-1	0	-1	1	$L_1^{(1)}$
$L_2^{(3)}$	<b>pivô</b>	$m_{21}$	-2	0	2	2	$L_2^{(2)} - 0,2L_1^{(2)}$
$L_3^{(3)}$		$m_{31}$	<b>4</b>	-4	1	3	$L_3^{(2)} - 0L_1^{(2)}$
$L_4^{(3)}$		$m_{41}$	-1	-4	7	1	$L_4^{(2)} - 0,8L_1^{(2)}$
$L_2^{(4)}$	$m_{32} = -0,5$ $m_{42} = -0,25$	$m_{31}$	<b>4</b>	<b>-4</b>	<b>1</b>	<b>3</b>	$L_3^{(3)}$
$L_3^{(4)}$		$m_{21}$	-2	0	2	2	$L_2^{(3)}$
$L_4^{(4)}$		$m_{41}$	-1	-4	7	1	$L_1^{(3)}$
$L_3^{(5)}$	<b>pivô</b>	$m_{21}$	$m_{32}$	-2	2,5	2	$L_3^{(4)} + 0,5L_2^{(4)}$
$L_4^{(5)}$		$m_{41}$	$m_{42}$	<b>-5</b>	7,25	1	$L_4^{(4)} + 0,25L_2^{(4)}$
$L_3^{(6)}$	$m_{43} = 0,4$	$m_{41}$	$m_{42}$	<b>-5</b>	<b>7,25</b>	<b>1</b>	$L_4^{(5)}$
$L_4^{(6)}$		$m_{21}$	$m_{32}$	-2	2,5	2	$L_3^{(5)}$
$L_4^{(7)}$		$m_{21}$	$m_{32}$	$m_{43}$	<b>-0,4</b>	<b>2</b>	$L_4^{(6)} - 0,4L_3^{(6)}$

Observe que é feita, de imediato, a troca de posição entre as linhas um e quatro, e a mesma troca deve ser feita no vetor de permutação. Obtém-se, então,  $P^{(1)} = \begin{bmatrix} 4 & 2 & 3 & 1 \end{bmatrix}^t$  e realiza-se a eliminação dos elementos da primeira coluna.

No passo dois, que consiste na eliminação dos elementos da segunda coluna, é verificado que o pivô está na terceira linha. Logo, é necessário fazer a troca de posição entre as linhas dois e três. Esta mesma transformação deve ser realizada no vetor de permutação, obtendo, então,  $P^{(2)} = \begin{bmatrix} 4 & 3 & 2 & 1 \end{bmatrix}^t$ .

No passo três, o pivô está na quarta linha. Logo, é necessário fazer a troca de posição entre as linhas três e quatro. Efetuando a mesma transformação no vetor de permutação, obtém-se  $P^{(3)} = \begin{bmatrix} 4 & 3 & 1 & 2 \end{bmatrix}$ . Temos, a seguir, as matrizes  $L$  e  $U$  resultantes da etapa de eliminação:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0,8 & -0,25 & 1 & 0 \\ 0,2 & -0,5 & 0,4 & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 5 & 0 & 5 & -10 \\ 0 & 4 & -4 & 1 \\ 0 & 0 & -5 & 7,25 \\ 0 & 0 & 0 & -0,4 \end{bmatrix}$$

**Resolução do sistema  $Ly = B$ :**

$$\text{Aplicando } P^{(3)} = \begin{bmatrix} 4 \\ 3 \\ 1 \\ 2 \end{bmatrix}^t \text{ ao vetor } B = \begin{bmatrix} 6 \\ 8 \\ -7 \\ -40 \end{bmatrix}^t, \text{ obtém-se } B_p = \begin{bmatrix} -40 \\ -7 \\ 6 \\ 8 \end{bmatrix}^t.$$

Resolvendo o sistema  $Ly = B$  por substituições sucessivas:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0,8 & -0,25 & 1 & 0 \\ 0,2 & -0,5 & 0,4 & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} -40 \\ -7 \\ 6 \\ 8 \end{bmatrix} \Rightarrow y = \begin{bmatrix} -40 \\ -7 \\ 36,25 \\ -2 \end{bmatrix}$$

**Resolução do sistema**  $Ux = y$  por substituições retroativas:

$$\begin{bmatrix} 5 & 0 & 5 & -10 \\ 0 & 4 & -4 & 1 \\ 0 & 0 & -5 & 7,25 \\ 0 & 0 & 0 & -0,4 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -40 \\ -7 \\ 36,25 \\ -2 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 2 \\ -3 \\ 0 \\ 5 \end{bmatrix}$$

O vetor resíduo é nulo. Logo temos que  $x = \begin{bmatrix} 2 & -3 & 0 & 5 \end{bmatrix}^t$  é solução exata do sistema de equações lineares.

## 7.3 Algoritmo e Complexidade

O pseudocódigo seguinte apresenta o algoritmo da decomposição LU com pivotação parcial para resolver um sistema linear  $Ax = B$ .

```

1: Entrada:  $A, b, n$ .
2:  $L = \text{identidade}(n)$ ; {Criação de uma matriz identidade de ordem}  $n$ .
3: para ( $k = 1$ ) até  $(n - 1)$  faça
4:   Encontre  $r \geq k$  tal que  $A(r, k) = \max_{k \leq i \leq n} \text{abs}(A(i, j))$ ; {Escolha do elemento pivô}
5:   Troque a linha  $k$  com a linha  $r$ ;
6:   para ( $i = k + 1$ ) até  $n$  faça
7:      $m(i, k) = -A(i, k)/A(k, k)$ ;
8:      $L(i, j) = m(i, k)$ ;
9:     para ( $j = k$ ) até  $n$  faça
10:       $A(i, j) = A(i, j) + m(i, k)A(k, j)$ ;
11:   fim para
12: fim para
13: fim para
14:  $y = \text{Algoritmo Substituições Sucessivas}(L, b)$ ;
15:  $U = A$ ;  $d = y$ 
16:  $x = \text{Algoritmo Substituições Retroativas}(U, d)$ ;
17: Saída: vetor solução  $x$ .

```

Algoritmo 5: Eliminação de Decomposição LU. (abs: valor absoluto).

A Tabela 7.1 apresenta o resumo do esforço computacional do Algoritmo de Eliminação de Gauss.

Operações	Complexidade
Adições	$\frac{1}{3}n^3 - \frac{1}{2}n^2 + \frac{7}{6}n$
Multiplicações	$\frac{1}{3}n^3 - \frac{1}{3}n$
Divisões	$n$

Tabela 7.1: Complexidade da Decomposição LU considerando um sistema linear de ordem  $n$ . Fonte: [1].

## 7.4 Exercícios



E. 1. Considere o seguinte sistema de equações lineares:

$$\begin{cases} 7x_1 + 2x_2 - 3x_3 = -12 \\ 2x_1 + 5x_2 - 3x_3 = -20 \\ x_1 - x_2 - 6x_3 = -26 \end{cases}$$

- a) Obtenha a fatoração LU da matriz A.
- b) Calcule o determinante da matriz A.
- c) Multiplique as matrizes L e U resultantes para verificar que A é produzida.
- d) Utilize a decomposição LU para resolver o sistema linear.
- e) Resolva o sistema linear utilizando a decomposição LU com pivotação para o vetor alternativo  $B_2 = \begin{bmatrix} 12 & 18 & -6 \end{bmatrix}^t$ .



This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



## **Parte III**

# **Resolução de Sistemas de Equações Lineares Simultâneas - Métodos Iterativos**





## Métodos Iterativos

### Sumário da Aula

8.1	Introdução . . . . .	67
8.1.1	<b>Revisão da Álgebra: Norma Matricial</b> . . . . .	68
8.1.2	Convergência dos Métodos Iterativos . . . . .	70
8.1.3	Critério de Parada . . . . .	71
8.2	Método de Jacobi . . . . .	71
8.2.1	Critério de Convergência . . . . .	73
8.3	Exercícios . . . . .	75

### 8.1 Introdução

Assim como os métodos diretos, métodos iterativos também podem ser aplicados na resolução de um sistema  $Ax = B$  cujo  $\det(A) \neq 0$ . São apresentados a seguir um breve resumo de resultados e definições necessários para a compreensão dos métodos iterativos aplicados na solução de sistemas de equações lineares.

Para determinar a solução de um sistema de equações lineares por meio de um método iterativo é preciso transformá-lo em um sistema de equações equivalente que possibilite a definição de um esquema iterativo. Ou seja, transformar  $Ax = B$  na seguinte **função de iteração**:

$$x^k = Mx^{k-1} + C$$



onde  $M$  é uma matriz iterativa de ordem  $n$  e  $C$  é um vetor de tamanho  $n \times 1$ .

Um método iterativo é dito **estacionário** quando a matriz  $M$  da função de iteração for fixa, ou seja, quando uma nova aproximação é obtida da anterior sempre pelo mesmo processo.

Partindo de uma aproximação inicial  $x^0$ , a função de iteração fornece uma sequência de soluções aproximadas  $x^1, x^2, \dots, x^k, \dots$ , onde cada uma delas é **obtida da solução anterior** por meio da aplicação de um mesmo procedimento, da seguinte forma:

$$\begin{aligned} x^1 &= Mx^0 + C \\ x^2 &= Mx^1 + C \\ &\vdots \\ x^k &= Mx^{k-1} + C, \quad k = 1, 2, \dots \end{aligned}$$

Pretende-se que esta sequência gerada seja convergente para a solução  $\bar{x}$  do sistema linear, ou seja:

$$\lim_{k \rightarrow \infty} x^k = \bar{x}$$

**Definição 8.1.** Se a sequência  $\{x^{(k)}\}$  convergir para um limite, qualquer que seja a aproximação inicial  $x^0$ , então o método iterativo é dito **convergente**.

**Definição 8.2.** Se os sistemas de equações  $Ax = B$  e  $(I - M)x = C$  possuírem a mesma solução, então o método iterativo é dito **consistente**.

**Proposição 8.1.** Seja  $\det(A) \neq 0$ . O método iterativo proposto é consistente se, e somente se,

$$(I - M)A^{-1}B = C$$

**Prova:** O sistema linear é consistente, se admite pelo menos uma solução.

Temos as seguintes relações:

$$(I) \quad x = Mx + C \Rightarrow x - Mx = C \Rightarrow (I - M)x = C$$

$$(II) \quad Ax = B \Rightarrow A^{-1}Ax = A^{-1}B \Rightarrow Ix = A^{-1}B \Rightarrow x = A^{-1}B$$

Substituindo a relação (II) em (I) obtemos

$$(I - M).(A^{-1}B) = C.$$

■

### 8.1.1 Revisão da Álgebra: Norma Matricial

**Definição 8.3. (Norma de vetores)** Define-se norma de um vetor  $x \in V$  (espaço vetorial):

$$\begin{aligned} \|\cdot\| : V &\rightarrow \mathbb{R} \\ x &\rightarrow \|x\| \end{aligned}$$

onde as seguintes condições são satisfeitas:

$$n_1) \|x\| \geq 0, \quad \forall x \in V; \|x\| = 0 \Leftrightarrow x = 0$$

$$n_2) \|\alpha x\| = |\alpha| \|x\|; \quad \forall \alpha \in \mathbb{R}, \quad \forall x \in V$$

$$n_3) \|x + y\| \leq \|x\| + \|y\|; \quad \forall x, y \in V$$

Considere  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ , logo:

$$\|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots + |x_n|$$

$$\|x\|_2 = \left( \sum_{i=1}^n |x_i|^2 \right)^{1/2} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| = \max \{|x_1|, |x_2|, \dots, |x_n|\}$$

De forma geral, quando  $V = \mathbb{R}^n$ , as normas  $l_p$  são definidas por:

$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad p \geq 1$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

**Exemplo 18.** Seja o vetor  $x = \begin{bmatrix} 2 & 5 & 3 & 1 \end{bmatrix}$ , logo:

$$\|x\|_1 = |2| + |5| + |3| + |1| = 11$$

$$\|x\|_2 = \sqrt{2^2 + 5^2 + 3^2 + 1^2} = \sqrt{39}$$

$$\|x\|_\infty = \max \{|2|, |5|, |3|, |1|\} = 5$$

**Definição 8.4. (Norma de matrizes)** Considere  $V = \mathbb{R}(n, n)$  o espaço vetorial de todas as matrizes quadradas de ordem  $(n \times n)$  sobre  $\mathbb{R}$ . Uma norma em  $V$  é uma aplicação indicada por  $\|\cdot\|$  tal que:

$$\begin{aligned} \|\cdot\| : \mathbb{R}(n, n) &\rightarrow \mathbb{R} \\ A &\rightarrow \|A\| \end{aligned}$$

satisfazendo as seguintes condições:

$$n_1) \|A\| \geq 0, \quad \forall A \in \mathbb{R}(n, n); \|A\| = 0 \Leftrightarrow A = 0$$

$$n_2) \|\alpha A\| = |\alpha| \|A\|; \quad \forall \alpha \in \mathbb{R}, \forall A \in \mathbb{R}(n, n)$$

$$n_3) \|A + B\| \leq \|A\| + \|B\|; \quad \forall A, B \in \mathbb{R}(n, n)$$

Considere  $A = (a_{ij})_{i,j=1,\dots,n}$ . São definidas as seguintes normas de matrizes:

$$\|A\|_1 = \|A\|_C = \max_{1 \leq j \leq n} \left( \sum_{i=1}^n |a_{ij}| \right) \rightarrow \text{Norma coluna}$$

$$\|A\|_\infty = \|A\|_L = \max_{1 \leq i \leq n} \left( \sum_{j=1}^n |a_{ij}| \right) \rightarrow \text{Norma linha}$$

$$\|A\|_2 = \sqrt{\sum_{i=1}^n \sum_{j=1}^n (a_{ij})^2} \rightarrow \text{Norma Euclidiana}$$

Para as normas  $\|\cdot\|_1$  e  $\|\cdot\|_\infty$ , vale a seguinte propriedade:

$$\|A + B\| \leq \|A\| + \|B\|, \quad \forall A, B \in \mathbb{R}(n, n)$$


---

**Exemplo 19.** Considere, por exemplo, a matriz

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

São calculadas as seguintes normas:

$$\begin{aligned} \|A\|_1 &= \|A\|_C = \max_{1 \leq j \leq n} \left( \sum_{i=1}^n |a_{ij}| \right) \\ &= \max \{ (2 + 3 + 1), (1 + 4 + 2), (3 + 2 + 5) \} = \max \{ (6), (7), (10) \} = 10 \end{aligned}$$

$$\begin{aligned} \|A\|_\infty &= \|A\|_L = \max_{1 \leq i \leq n} \left( \sum_{j=1}^n |a_{ij}| \right) \\ &= \max \{ (2 + 1 + 3), (3 + 4 + 2), (1 + 2 + 5) \} = \max \{ (6), (9), (8) \} = 9 \end{aligned}$$

$$\begin{aligned} \|A\|_2 &= \sqrt{\sum_{i=1}^n \sum_{j=1}^n (a_{ij})^2} \\ &= \sqrt{2^2 + 1^2 + 3^2 + 3^2 + (-4)^2 + 2^2 + 1^2 + 2^2 + (-5)^2} = \sqrt{73} \approx 8,5440 \end{aligned}$$

**Definição 8.5.** Considere uma norma de vetor  $x \in \mathbb{R}^n$  e uma norma de matriz  $A \in \mathbb{R}(n, n)$ . Dizemos que estas normas são **consistentes** se satisfazem a expressão:

$$\|Ax\| \leq \|A\| \|x\|, \quad \forall A \in \mathbb{R}(n, n) \text{ e } \forall x \in \mathbb{R}^n.$$

## 8.1.2 Convergência dos Métodos Iterativos

**Definição 8.6. (Sequência Convergente:)** Considere uma sequência de vetores  $x^{(i)} = (x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)})$  do espaço vetorial  $\mathbb{R}^n$ . Dizemos que a sequência converge para

$$\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) \in \mathbb{R}^n$$

se  $\|x^{(i)} - \bar{x}\| \rightarrow 0$ , quando  $i \rightarrow \infty$ , para qualquer norma em  $\mathbb{R}^n$ .

**Teorema 8.1.** A condição **necessária e suficiente** para a convergência do processo iterativo  $x^k = Mx^{k-1} + C$  é que  $\max |\lambda_i| < 1$ , onde  $\lambda_i$  são os autovalores da matriz  $M$ .

**Corolário 8.1. (Critério Geral de Convergência)** O processo iterativo  $x^k = Mx^{k-1} + C$  é **convergente se**, para qualquer norma de matrizes,  $\|M\| < 1$ .

### 8.1.3 Critério de Parada

Nos métodos iterativos escolhe-se um  $x^{(0)}$  como uma aproximação inicial para a solução do sistema linear  $Ax = B$ . Esta aproximação inicial é refinada pelo processo iterativo até obter uma nova solução que possua uma determinada precisão (número de casas decimais corretas).

O **critério de parada** usado para finalizar o processo iterativo quando se obtém  $x^k$  será tal que  $\max_{1 \leq i \leq n} |x_i^k - x_i^{k-1}|$  seja menor ou igual a uma precisão estabelecida, ou seja,

$$\max_{1 \leq i \leq n} |x_i^k - x_i^{k-1}| \leq \epsilon$$

onde  $\epsilon$  é uma precisão pré-fixada,  $x_i^k$  e  $x_i^{k-1}$  são duas aproximações consecutivas para  $\bar{x}$ . O valor de  $x^k$  é tomado como uma aproximação para a solução do sistema de equações.

Também é possível usar como critério de parada, junto com a precisão, o número máximo de iterações  $k = k_{max}$  aplicado pelo método iterativo. Assim, o método poderá ser interrompido quando atingir a precisão desejada ou quando atingir o número máximo de iterações pré-estabelecido.

Nesta disciplina serão estudados dois métodos iterativos estacionários: Jacobi e Gauss-Seidel.

## 8.2 Método de Jacobi



Considere o sistema de equações lineares  $Ax = B$ , em que  $A = (a_{ij})_{i,j=1,\dots,n}$ ,  $\det(A) \neq 0$  e com a diagonal principal  $a_{ii} \neq 0$ ,  $i = 1, 2, \dots, n$ :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n = b_n \end{cases}$$

Podemos reescrever o sistema na forma equivalente, dividindo cada linha pelo elemento da diagonal principal e explicitando  $x_1$  na primeira coluna,  $x_2$  na segunda coluna,  $x_3$  na terceira coluna, e, sucessivamente, até  $x_n$  na  $n$ -ésima coluna. Assim:

$$\begin{cases} x_1 = \frac{1}{a_{11}}(b_1 - a_{12}x_2 - a_{13}x_3 - \dots - a_{1n}x_n) \\ x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1 - a_{23}x_3 - \dots - a_{2n}x_n) \\ \vdots \\ x_n = \frac{1}{a_{nn}}(b_n - a_{n1}x_1 - a_{n2}x_2 - a_{n3}x_3 - \dots - a_{nn}x_n) \end{cases}$$

O sistema anterior é escrito na forma matricial como:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & \frac{-a_{12}}{a_{11}} & \frac{-a_{13}}{a_{11}} & \dots & \frac{-a_{1n}}{a_{11}} \\ \frac{-a_{21}}{a_{22}} & 0 & \frac{-a_{23}}{a_{22}} & \dots & \frac{-a_{2n}}{a_{22}} \\ & & \ddots & & \vdots \\ \frac{-a_{n1}}{a_{nn}} & \frac{-a_{n2}}{a_{nn}} & & \dots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}.$$

ou na forma de função de iteração  $x = Mx + C$ , onde  $M = (m_{ij})$  é a matriz iterativa:

$$m_{ij} = \begin{cases} 0 & i = j \\ \frac{-a_{ij}}{a_{ii}} & i \neq j \end{cases} \quad i, j = 1, 2, \dots, n \text{ e } C_i = \frac{b_i}{a_{ii}}, i = 1, 2, \dots, n$$

Assim, o **método iterativo de Jacobi** é então definido como:

$$x^{(k)} = Mx^{(k-1)} + C, \quad k = 1, 2, \dots$$

ou seja,

$$\begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{bmatrix} = \begin{bmatrix} 0 & \frac{-a_{12}}{a_{11}} & \frac{-a_{13}}{a_{11}} & \dots & \frac{-a_{1n}}{a_{11}} \\ \frac{-a_{21}}{a_{22}} & 0 & \frac{-a_{23}}{a_{22}} & \dots & \frac{-a_{2n}}{a_{22}} \\ & & \ddots & & \vdots \\ \frac{-a_{n1}}{a_{nn}} & \frac{-a_{n2}}{a_{nn}} & & \dots & 0 \end{bmatrix} \begin{bmatrix} x_1^{(k-1)} \\ x_2^{(k-1)} \\ \vdots \\ x_n^{(k-1)} \end{bmatrix} + \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}$$

Ainda é possível escrever o método com a notação usual para sistemas:

$$\begin{cases} x_1^{(k)} = \frac{1}{a_{11}}(b_1 - a_{12}x_2^{(k-1)} - a_{13}x_3^{(k-1)} - \dots - a_{1n-1}x_{n-1}^{(k-1)} - a_{1n}x_n^{(k-1)}) \\ x_2^{(k)} = \frac{1}{a_{22}}(b_2 - a_{21}x_1^{(k-1)} - a_{23}x_3^{(k-1)} - \dots - a_{2n-1}x_{n-1}^{(k-1)} - a_{2n}x_n^{(k-1)}) \\ \vdots \\ x_n^{(k)} = \frac{1}{a_{nn}}(b_n - a_{n1}x_1^{(k-1)} - a_{n2}x_2^{(k-1)} - a_{n3}x_3^{(k-1)} - \dots - a_{nn-1}x_{n-1}^{(k-1)} - a_{nn}x_n^{(k-1)}) \end{cases}$$

---

#### Observação

---

O método iterativo de Jacobi é escrito de forma geral por:

$$x_i^{(k)} = \frac{1}{a_{ii}} \left( b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j^{(k-1)} \right), \quad i = 1, \dots, n$$


---

**Exemplo 20.** Resolva o sistema linear pelo método de Jacobi com o limite de 6 iterações ou uma precisão menor que  $10^{-2}$ . Utilize  $x^{(0)} = [0 \ 0 \ 0]^t$  como estimativa inicial e retenha, nos cálculos, quatro casas decimais.

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

**Solução:** A função de iteração é dada por:

$$\begin{cases} x_1^k = 1 - 2x_2^{k-1} + 2x_3^{k-1} \\ x_2^k = 1 - x_1^{k-1} - x_3^{k-1} \\ x_3^k = 1 - 2x_1^{k-1} - 2x_2^{k-1} \end{cases}$$

e os resultados da aplicação do método de Jacobi são apresentados no quadro a seguir:

$k$	$x_1^k$	$x_2^k$	$x_3^k$	$\max_{1 \leq i \leq 3}  x_i^k - x_i^{k-1} $
0	0	0	0	
1	1	1	1	1
2	1	-1	-3	4
3	-3	3	1	4
4	-3	3	1	0

Portanto, a solução do sistema linear é  $x \approx x^4 = [-3 \ 3 \ 1]^t$ .

### 8.2.1 Critério de Convergência

Embora a ordem das equações em um sistema linear não exerça qualquer influência com relação à existência de solução, quando se trata da utilização de um método iterativo esta ordem será relevante, uma vez que define-se uma função de iteração.

É **condição suficiente** para que o método iterativo de Jacobi **gere uma sequência que converge** para a solução de um sistema de equações  $Ax = B$ , qualquer que seja a aproximação inicial  $x^0$ , que:

a) o **critério das linhas** seja satisfeito, isto é, se:

$$|a_{ii}| > \sum_{\substack{j=1 \\ i \neq j}}^n |a_{ij}|, \quad i = 1, 2, \dots, n$$

Quanto mais próxima de zero estiver a relação  $\frac{\sum_{j=1}^n |a_{ij}|}{|a_{ii}|}$ , mais rápida será a convergência.

b) o **critério das colunas** seja satisfeito, isto é, se:

$$|a_{jj}| > \sum_{\substack{i=1 \\ i \neq j}}^n |a_{ij}|, \quad j = 1, 2, \dots, n$$

Quanto mais próxima de zero estiver a relação  $\frac{\sum_{i=1}^n |a_{ij}|}{|a_{jj}|}$  mais rápida será a convergência.

Observe que estes dois critérios envolvem **condições que são apenas suficientes**. Assim, se pelo menos uma delas for satisfeita, então está assegurada a convergência. Entretanto, se nenhuma das duas for satisfeita nada se pode afirmar. Considere agora a Definição 8.7.

**Definição 8.7.** Seja uma matriz  $A = (a_{ij})_{i,j=1,\dots,n}$ . Diz-se que  $A$  é **estritamente diagonal dominante** se:

$$|a_{ii}| > \sum_{\substack{j=1 \\ i \neq j}}^n |a_{ij}|, \quad i = 1, 2, \dots, n.$$

Caso

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ i \neq j}}^n |a_{ij}|, \quad i = 1, 2, \dots, n$$

a matriz  $A$  é dita **diagonalmente dominante**.

**Exemplo 21.** Considere a matriz

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

Da Definição 8.7, tem-se que:

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ i \neq j}}^3 |a_{ij}|, \quad i = 1, 2, 3$$

$$\text{Para } i = 1: |a_{11}| > |a_{12}| + |a_{13}| \rightarrow |3| > |-1| + |1|$$

$$\text{Para } i = 2: |a_{22}| > |a_{21}| + |a_{23}| \rightarrow |4| > |1| + |1|$$

$$\text{Para } i = 3: |a_{33}| > |a_{31}| + |a_{32}| \rightarrow |-5| \geq |2| + |3|$$

Portanto, conclui-se que a matriz  $A$  é diagonalmente dominante.

---

Observação

---

**Da Definição 8.7 é possível concluir que se a matriz  $A$  for estritamente diagonalmente dominante, então o critério de linhas é satisfeito. Assim, para testar se o método de Jacobi converge basta verificar se a matriz dos coeficientes  $A$  é estritamente diagonalmente dominante.**

---

**Exemplo 22.** Resolva o sistema linear pelo método de Jacobi com precisão menor ou igual a  $10^{-3}$  e  $k_{max} = 10$ . Utilize  $x^{(0)} = [5, 7 \ 2, 5 \ -0, 8]^t$  como estimativa inicial e retenha, nos cálculos, quatro casas decimais.

$$\begin{cases} 10x_1 + 3x_2 - 2x_3 = 57 \\ 2x_1 + 8x_2 - x_3 = 20 \\ x_1 + x_2 + 5x_3 = -4 \end{cases}$$

**Solução:** Da Definição 8.7, tem-se que a matriz  $A$  é estritamente diagonal dominante; isto é:

$$\text{Para } i = 1: |a_{11}| > |a_{12}| + |a_{13}| \rightarrow |10| > |3| + |-2|$$

Para  $i = 2$ :  $|a_{22}| > |a_{21}| + |a_{23}| \rightarrow |8| > |2| + |-1|$

Para  $i = 3$ :  $|a_{33}| > |a_{31}| + |a_{32}| \rightarrow |5| > |1| + |1|$

Portanto, conclui-se que o processo iterativo de Jacobi convergirá.

A função de iteração é dada por:

$$\begin{cases} x_1^k = \frac{1}{10}(-3x_2^{k-1} + 2x_3^{k-1} + 57) \\ x_2^k = \frac{1}{8}(-2x_1^{k-1} + 1x_3^{k-1} + 20) \\ x_3^k = \frac{1}{5}(-x_1^{k-1} - x_2^{k-1} - 4) \end{cases}$$

Os resultados da aplicação do método de Jacobi estão sumiraziados no quadro a seguir:

$k$	$x_1^k$	$x_2^k$	$x_3^k$	$\max_{1 \leq i \leq 3}  x_i^k - x_i^{k-1} $
0	5,7	2,5	-0,8	
1	4,79	0,975	-2,44	1,64
2	4,9195	0,9975	-1,953	0,487
3	5,0102	1,0260	-1,9834	0,0907
4	4,9955	0,9995	-2,0072	0,0265
5	4,9987	1,0002	-1,9990	0,0082
6	5,0001	1,0004	-1,9998	0,0007

Portanto, a solução do sistema linear é  $x \approx x^6 = \begin{bmatrix} 5,0001 & 1,0004 & -1,9998 \end{bmatrix}^t$ .

## 8.3 Exercícios



E. 1. Considere o seguinte sistema linear:

$$\begin{cases} 10x_1 + 2x_2 + 3x_3 = 28 \\ x_1 + 10x_2 + 6x_3 = 7 \\ 2x_1 - 2x_2 - 10x_3 = -17 \end{cases}$$

- Verifique a condição de convergência do método de Jacobi.
- Resolva o sistema linear e pelo método iterativo de Jacobi com  $x^{(0)} = [0 \ 0 \ 0]^t$  e  $\epsilon = 0,01$ . Reordene as equações, se for preciso, de modo que a convergência esteja garantida.

E. 2. Considere o seguinte sistema linear:

$$\begin{bmatrix} 6 & 2 & -2 \\ 1 & 4 & 1 \\ 2 & 2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 6 \\ 9 \end{bmatrix}$$

- Verifique a condição de convergência do método de Jacobi.
- Resolva o sistema linear e pelo método iterativo de Jacobi com  $x^{(0)} = [0 \ 0 \ 0]^t$  e  $\epsilon = 0,01$ . Reordene as equações, se for preciso, de modo que a convergência esteja garantida.





[illegible]





# Métodos de Gauss Seidel

## Sumário da Aula

9.1	Método de Gauss Seidel . . . . .	79
9.1.1	Crítério de Convergência . . . . .	81
9.2	Complexidade dos métodos iterativos . . . . .	83
9.3	Considerações Finais . . . . .	83
9.4	Exercícios . . . . .	84

## 9.1 Método de Gauss Seidel

Considere o sistema de equações lineares  $Ax = B$ , em que  $A = (a_{ij})_{i,j=1,\dots,n}$ ,  $\det(A) \neq 0$  e com a diagonal principal  $a_{ii} \neq 0$ ,  $i = 1, 2, \dots, n$ :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n &= b_n \end{cases}$$

O sistema pode ser reescrito na forma equivalente, dividindo cada linha pelo elemento da diagonal principal e explicitando  $x_1$  na primeira coluna,  $x_2$  na segunda coluna,  $x_3$  na terceira coluna, e, sucessivamente, até  $x_n$  na n-ésima coluna:

$$\left\{ \begin{array}{l} x_1 = \frac{1}{a_{11}}(b_1 - a_{12}x_2 - a_{13}x_3 - \dots - a_{1n}x_n) \\ x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1 - a_{23}x_3 - \dots - a_{2n}x_n) \\ \vdots \\ x_n = \frac{1}{a_{nn}}(b_n - a_{n1}x_1 - a_{n2}x_2 - a_{n3}x_3 - \dots - a_{nn}x_n) \end{array} \right.$$

ou ainda na forma matricial:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & \frac{-a_{12}}{a_{11}} & \frac{-a_{13}}{a_{11}} & \dots & \frac{-a_{1n}}{a_{11}} \\ \frac{-a_{21}}{a_{22}} & 0 & \frac{-a_{23}}{a_{22}} & \dots & \frac{-a_{2n}}{a_{22}} \\ & & \ddots & & \vdots \\ \frac{-a_{n1}}{a_{nn}} & \frac{-a_{n2}}{a_{nn}} & & \dots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}$$

O método iterativo de Gauss-Seidel é dado por:

$$\left\{ \begin{array}{l} x_1^{(k)} = \frac{1}{a_{11}}(b_1 - a_{12}x_2^{(k-1)} - a_{13}x_3^{(k-1)} - \dots - a_{1n}x_n^{(k-1)}) \\ x_2^{(k)} = \frac{1}{a_{22}}(b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k-1)} - \dots - a_{2n}x_n^{(k-1)}) \\ x_3^{(k)} = \frac{1}{a_{33}}(b_3 - a_{31}x_1^{(k)} - a_{32}x_2^{(k)} - \dots - a_{3n}x_n^{(k-1)}) \\ \vdots \\ x_n^{(k)} = \frac{1}{a_{nn}}(b_n - a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - a_{n3}x_3^{(k)} - \dots - a_{nn}x_n^{(k-1)}) \end{array} \right.$$

cuja forma Matricial é:

$$\begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \frac{-a_{21}}{a_{22}} & 0 & 0 & \dots & 0 \\ & \ddots & & & \\ \frac{-a_{n1}}{a_{nn}} & \frac{-a_{n2}}{a_{nn}} & \dots & & 0 \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{bmatrix} + \begin{bmatrix} 0 & \frac{-a_{12}}{a_{11}} & \frac{-a_{13}}{a_{11}} & \dots & \frac{-a_{1n}}{a_{11}} \\ 0 & 0 & \frac{-a_{23}}{a_{22}} & \dots & \frac{-a_{2n}}{a_{22}} \\ & & \ddots & & \vdots \\ 0 & 0 & & \dots & 0 \end{bmatrix} \begin{bmatrix} x_1^{(k-1)} \\ x_2^{(k-1)} \\ \vdots \\ x_n^{(k-1)} \end{bmatrix} + \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}$$



**Observação:**

De forma geral, o método iterativo de Gauss-Seidel é escrito como:

$$x_i(k) = \frac{\left( b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^n a_{ij}x_j^{(k-1)} \right)}{a_{ii}}, \quad i = 1, \dots, n$$

**Exemplo 23.** Resolva o sistema linear pelo método de Gauss-Seidel com o limite de 6 iterações ou uma precisão menor que  $10^{-2}$ . Utilize  $x^{(0)} = [0 \ 0 \ 0]$  como estimativa inicial e retenha, nos cálculos, quatro casas decimais.

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

**Solução:** A função de iteração é:

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

e os resultados da aplicação do método de Gauss-Seidel são apresentados no quadro a seguir:

$k$	$x_1^k$	$x_2^k$	$x_3^k$	$\max_{1 \leq i \leq 3}  x_i^k - x_i^{k-1} $
0	0	0	0	
1	0,5	0,1667	0,2667	0,5
2	0,55	0,2389	0,2622	0,0722
3	0,5117	0,2502	0,2523	0,038
4	0,5011	0,2504	0,2501	0,0009

Portanto, a solução do sistema linear é  $x = [0,5011 \ 0,2504 \ 0,2501]^t$  com uma precisão  $\epsilon = 0,0009$ .

### 9.1.1 Critério de Convergência

O Método de Gauss-Seidel converge se:

a) O **critério de Sassenfeld** for satisfeito, ou seja,

$$\max_{1 \leq i \leq n} \beta_i < 1,$$

onde os  $\beta_i$  são calculados por recorrência através de:

$$\beta_i = \sum_{j=1}^{i-1} |m_{ij}| \beta_j + \sum_{j=i+1}^n |m_{ij}|$$

b) O **critério das linhas** for satisfeito.

c) A matriz dos coeficientes for **estritamente diagonalmente dominante**.

---

Observações:

---

1. Se a matriz dos coeficientes  $A$  for estritamente diagonal dominante, então, qualquer que seja a estimativa inicial  $x^{(0)}$ , tanto o método de Jacobi como o de Gauss-Seidel convergem.
2. Se  $A$  é uma matriz diagonal dominante e se o sistema linear é irredutível<sup>1</sup>, então, para qualquer estimativa inicial, os métodos iterativos convergem.
3. Dado um sistema linear  $Ax = b$ , pode acontecer do método de Jacobi aplicado a ele gerar uma sequência de soluções que converge para a solução do sistema linear, enquanto o método de Gauss-Seidel gerar uma sequência que diverge e vice-versa.
4. A convergência dos métodos iterativos não dependem da aproximação inicial. Evidentemente, quanto melhor for a aproximação inicial menor será o número de iterações necessárias para atingir uma determinada precisão. Como não conhecemos a priori a solução é comum utilizar o vetor nulo ( $x^{(0)} = \vec{0}$ ) como sendo o vetor inicial.

---

**Exemplo 24.** Resolva o sistema linear pelo método de Gauss Seidel com precisão menor ou igual a  $10^{-3}$  e  $k_{max} = 10$ . Utilize  $x^{(0)} = [5, 7 \quad 2, 5 \quad -0, 8]^t$  como estimativa inicial e retenha, nos cálculos, quatro casas decimais.

$$\begin{cases} 10x_1 + 3x_2 - 2x_3 &= 57 \\ 2x_1 + 8x_2 - x_3 &= 20 \\ x_1 + x_2 + 5x_3 &= -4 \end{cases}$$

**Solução:** Da Definição 8.7, tem-se que a matriz  $A$  é estritamente diagonal dominante; isto é:

$$\text{Para } i = 1: |a_{11}| > |a_{12}| + |a_{13}| \rightarrow |10| > |3| + |-2|$$

$$\text{Para } i = 2: |a_{22}| > |a_{21}| + |a_{23}| \rightarrow |8| > |2| + |-1|$$

$$\text{Para } i = 3: |a_{33}| > |a_{31}| + |a_{32}| \rightarrow |5| > |1| + |1|$$

Portanto, conclui-se que o processo iterativo de Gauss-Seidel convergirá.

A função de iteração é dada por:

$$\begin{cases} x_1^k &= \frac{1}{10}(-3x_2^{k-1} + 2x_3^{k-1} + 57) \\ x_2^k &= \frac{1}{8}(-2x_1^k + 1x_3^{k-1} + 20) \\ x_3^k &= \frac{1}{5}(-x_1^k - x_2^k - 4) \end{cases}$$

Os resultados da aplicação do método de Gauss-Seidel estão sumarizados no quadro a seguir:

---

<sup>1</sup>Um sistema linear é dito **redutível** quando for possível trabalhar com um número menor de equações que o número dado no sistema linear original e, conseqüentemente, determinar a solução do sistema para algumas incógnitas. Quando isso não for possível, o sistema é dito **irredutível**

$k$	$x_1^k$	$x_2^k$	$x_3^k$	$\max_{1 \leq i \leq 3}  x_i^k - x_i^{k-1} $
0	5,7	2,5	-0,8	
1	4,79	1,2025	-1,9985	1,2975
2	4,9396	1,0153	-1,9909	0,1872
3	4,9972	1,0018	-1,9998	0,0576
4	4,9995	1,0002	-1,9999	0,0023
5	4,9999	1,0002	-2,0000	0,0004

Portanto, a solução do sistema linear é  $x \approx x^5 = \begin{bmatrix} 4,9999 & 1,0002 & -2,0000 \end{bmatrix}^t$ .

## 9.2 Complexidade dos métodos iterativos

A análise da complexidade (quantidade de operações requeridas) em um método iterativo, em cada iteração, é bastante simples. O que não é trivial é determinar o número total de operações realizadas por um programa de resolução de sistemas de equações lineares por meio de um método iterativo, pois este depende do critério de parada adotado. Para evitar que se entre em loop, realizando operações quando não ocorre convergência, ou quando não se alcança a precisão estabelecida, sempre deve ser adotado como critério de parada, além da precisão desejada, um número máximo de iterações. No pior caso, este será o número de vezes que as iterações serão executadas.

A complexidade computacional dos algoritmos de Jacobi e Gauss-Seidel é apresentada na Tabela 9.1, onde  $k$  indica as iterações do algoritmo e  $n$  a ordem do sistema linear.

Operações	Complexidade
Adições	$kn^2 + kn + k$
Multiplicações	$kn^2 - kn$
Divisões	$(k + 1)n + k$

Tabela 9.1: Complexidade computacional dos métodos de Jacobi e Gauss-Seidel.

## 9.3 Considerações Finais

Ao resolver um sistema linear é preciso definir se a melhor estratégia a se usar é a aplicação de um método direto ou iterativo. A questão é que não se pode garantir, a priori, qual tipo de método é mais eficiente, pois é preciso fazer uma análise criteriosa das características da matriz dos coeficientes  $A$  e do porte do sistema a ser resolvido.

### Métodos Diretos:

Os métodos diretos possuem a vantagem de serem mais gerais e robustos do que os métodos iterativos, podendo ser utilizados na resolução de qualquer tipo de sistema de equações. São processos finitos e, teoricamente, obtêm a solução de qualquer sistema de equações não singular.

De forma geral, devem ser utilizados quando a matriz dos coeficientes  $A$  for densa, ou seja, apresentar poucos elementos não nulos e quando o sistema linear é de pequeno porte (dimensão pequena).

Estes métodos apresentam problemas com a propagação dos erros de arredondamento, e uma forma de minimizar este efeito é a utilização de técnicas de pivotamento.

Em problemas práticos, em que a matriz  $A$  do sistema se mantém inalterada e o vetor  $b$  sofre modificações, o método de decomposição LU é mais indicado, visto que ele faz a fatoração da matriz  $A$  uma única vez.



## Métodos Iterativos:

Os métodos iterativos são indicados para resolver sistemas lineares de grande porte (centenas de equações) e quando a matriz  $A$  for esparsa (possui muitos elementos nulos).

Estes métodos convergem apenas sob determinadas condições. Com a convergência assegurada, estes métodos se mostram eficientes pois: independem da solução inicial, não alteram a estrutura da matriz  $A$  durante a aplicação, e em geral, minimizam a propagação dos erros de arredondamento no resultado final.

A Tabela 9.2 apresenta de forma sucinta uma comparação entre os métodos diretos e iterativos levando em consideração alguns indicadores.

Indicador	Método Direto	Método Iterativo
Aplicação	Para a resolução de sistemas de equações densos de pequeno a médio porte.	Para a resolução de sistemas de equações de grande porte.
Esparsidade	Destroi a esparsidade da matriz dos coeficientes durante a fase de eliminação.	Preserva a esparsidade da matriz da matriz dos coeficientes.
Convergência	Se a matriz dos coeficientes não é singular, então a solução é sempre obtida.	Há garantia de se obter a solução somente sob certas condições
Número de operações	É possível determinar a priori o número de operações necessárias.	Não é possível determinar a priori a complexidade.
Erro de arredondamento	Propaga os erros durante os cálculos. A ampliação pode ser minimizada usando técnicas de pivotação.	Os erros de arredondamento não afetam as soluções obtidas em cada iteração. Apenas a solução final pode conter erro.

Tabela 9.2: Comparação entre métodos diretos e iterativos.

## 9.4 Exercícios



E. 1. Considere o seguinte sistema linear:

$$\begin{cases} x_1 + 10x_2 + 6x_3 = 7 \\ 10x_1 + 2x_2 + 3x_3 = 28 \\ 2x_1 - 2x_2 - 10x_3 = -17 \end{cases}$$

a) Verifique a condição de convergência do método de Jacobi.

b) Resolva o sistema linear pelo método iterativo de Gauss-Seidel com  $x^{(0)} = [0 \ 0 \ 0]^t$  e  $\epsilon = 0,01$ . Reordene as equações, se for preciso, de modo que a convergência esteja garantida.

E. 2. Considere o seguinte sistema de equações lineares:

$$\begin{cases} -9x_1 + 2x_2 + 6x_3 = 11 \\ 2x_1 + 4x_2 + x_3 = 4 \\ -x_1 + x_2 - 3x_3 = -2 \end{cases}$$

a) Resolva o sistema linear pelo método iterativo de Jacobi com  $x^{(0)} = [0 \ 1,5 \ 1]^t$  e  $\epsilon = 0,01$  ou  $k_{max} = 6$

- b) Resolva o sistema linear pelo método iterativo de Gauss-Seidel com  $x^{(0)} = [0 \quad 1,5 \quad 1]^t$  e  $\epsilon = 0,01$  ou  $k_{max} = 6$
- c) Compare o número de iterações e a precisão de ambos os métodos. O que é possível concluir?



[illegible]





## **Parte IV**

# **Interpolação Polinomial**



# 10

## Interpolação Polinomial

### Sumário da Aula

10.1 Introdução . . . . .	91
10.1.1 Interpolação Linear . . . . .	94
10.1.2 Erro de Truncamento . . . . .	94
10.2 Polinômios Interpoladores de Lagrange . . . . .	97
10.3 Exercícios . . . . .	101

### 10.1 Introdução

Neste tópico será apresentado a aproximação de função de uma variável real por outras funções mais simples, de modo que operações em geral possam ser realizadas com mais facilidade. Em geral, dispõe-se de dados que são fornecidos em um conjunto discreto de valores, dentro de um contínuo de possibilidades. Entretanto, pode ser necessário fazer estimativas em pontos que estão entre os valores discretos fornecidos. Ocorre, também, a situação na qual se faz necessária uma versão simplificada de uma função complicada. Ambas as aplicações são conhecidas como ajuste de curvas. Há duas abordagens gerais para o ajuste de curvas, as quais se distinguem com base na quantidade de erro associada com os dados.

Primeiro, quando os dados exibirem um grau significativo de erro, a estratégia será determinar uma única curva que represente a tendência geral dos dados. Como cada ponto individual poderá estar incorreto, não será feito qualquer esforço para passar a curva por todos os pontos. Em vez disto, a curva é escolhida para seguir o padrão dos pontos considerados como um grupo. Uma abordagem desta natureza é chamada



de regressão por mínimos quadrados. Segundo, quando se souber que os dados são muito precisos, a abordagem básica é ajustar uma curva ou uma série de curvas que **passam diretamente por cada um dos pontos**. Esta abordagem é o objeto de estudo deste tópico e é conhecida como **interpolação polinomial**.



Interpolar uma função  $y = f(x)$  em um intervalo finito  $[a, b]$  consiste em substituí-la, ou aproximá-la, por outra função  $y = g(x)$ . A necessidade de se utilizar este procedimento ocorre, basicamente, quando a função:

- a) não é conhecida na sua forma analítica, mas apenas por meio de um conjunto de pontos  $(x_i, y_i), i = 0, 1, \dots, n$ . Esta situação ocorre com muita frequência, na prática, quando se trabalha com dados obtidos de forma experimental;
- b) é conhecida na sua forma analítica, porém, operações como a diferenciação e a integração são difíceis (ou mesmo impossíveis) de realizar, ou seja, a função é de difícil tratamento. Teoricamente, a função  $y = g(x)$  pode ser qualquer uma e o caso mais comumente considerado é aquele em que  $g(x)$  pertence à classe das funções polinomiais.

A aproximação de funções por polinômios é uma das ideias mais antigas da análise numérica e é fácil entender a razão: os polinômios são facilmente computáveis, suas derivadas e integrais são, novamente, polinômios, seus zeros podem ser determinados com facilidade, etc. **O uso de polinômios interpoladores é importante, por exemplo, para a obtenção de valores intermediários em tabelas, na integração numérica, no cálculo de raízes de equações e na resolução de equações diferenciais ordinárias.**

Considere uma função  $f(x)$  definida em pontos distintos de um intervalo  $[a, b]$  e denotada por  $y_i = f(x_i)$ , onde  $i = 0, 1, \dots, n$ . **Interpolar esta função** consiste em aproximar  $f(x)$  por um polinômio  $P(x)$  de grau menor ou igual a  $n$ , tal que este polinômio coincida com a função em todos os seus  $(n + 1)$  pontos distintos, isto é:

$$P(x_i) = f(x_i) = y_i, i = 0, 1, \dots, n$$

As funções polinomiais interpoladoras são as mais populares não só por suas propriedades algébricas, mas, sobretudo, pela justificativa fornecida pelo teorema de aproximação de Weierstrass que, de fato, garante a existência de um polinômio capaz de aproximar uma função  $f(x)$  tão bem quanto se queira.

**Teorema 10.1. (Weierstrass)** Se  $f(x)$  é uma função contínua em um intervalo fechado  $[a, b]$ , então, dado  $\epsilon > 0$ , existe alguma função polinomial,  $P(x)$ , de ordem  $n = n(\epsilon)$ , tal que

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

Apesar de justificar a existência da função polinomial interpoladora, este teorema não é construtivo, isto é, não fornece modos ou critérios para a sua obtenção. Neste texto apresentam-se alguns dos procedimentos mais usuais para a obtenção de funções polinomiais interpoladoras.

**Teorema 10.2. (Existência e unicidade do polinômio interpolador)** Seja  $f(x)$  uma função definida em  $(x_i, y_i), i = 0, 1, \dots, n; (n + 1)$  pontos com abscissas distintas em um intervalo  $[a, b]$ , então existe um único polinômio  $P(x)$  de grau menor ou igual a  $n$  tal que

$$P(x_i) = f(x_i) = y_i, i = 0, 1, \dots, n.$$

**Prova:** O objetivo é aproximar uma função  $y = f(x)$  por um polinômio  $P(x)$  ou seja, deseja-se obter

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x^1 + a_0 = y$$

tal que  $p(x_i) = f(x_i) = y_i$  para todo  $i = 0, 1, 2, \dots, n$

Com esta condição, nos  $x_0, x_1, \dots, x_n$  pontos, temos:

$$\begin{cases} P(x_0) = a_n x_0^n + a_{n-1} x_0^{n-1} + \cdots + a_2 x_0^2 + a_1 x_0^1 + a_0 = y_0 \\ P(x_1) = a_n x_1^n + a_{n-1} x_1^{n-1} + \cdots + a_2 x_1^2 + a_1 x_1^1 + a_0 = y_1 \\ \vdots \\ P(x_n) = a_n x_n^n + a_{n-1} x_n^{n-1} + \cdots + a_2 x_n^2 + a_1 x_n^1 + a_0 = y_n \end{cases}$$

que consiste em um sistema de equações lineares  $Ax = B$ , em que:

$$x = \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix}, \quad B = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix} \quad \text{e} \quad A = \begin{bmatrix} x_0^n & x_0^{n-1} & \cdots & x_0^2 & x_0 & 1 \\ x_1^n & x_1^{n-1} & \cdots & x_1^2 & x_1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{n-1}^n & x_{n-1}^{n-1} & \cdots & x_{n-1}^2 & x_{n-1} & 1 \\ x_n^n & x_n^{n-1} & \cdots & x_n^2 & x_n & 1 \end{bmatrix}.$$

O determinante de A, chamado de determinante de Vandermonde, é calculado por:

$$\det(A) = \prod_{i < j} (x_i - x_j)$$

Por condição,  $x_0, x_1, \dots, x_n$  são valores distintos, então tem-se que  $\det(A) \neq 0$ , o que significa que o sistema linear admite solução única, e os coeficientes  $a_0, a_1, \dots, a_n$  do polinômio são únicos. Logo, existe um único polinômio,  $y = p(x)$ , tal que  $p(x_i) = f(x_i) = y_i$ ,  $i = 0, 1, \dots, n$  ■

O procedimento de interpolação polinomial está representado graficamente na Figura 10.1.

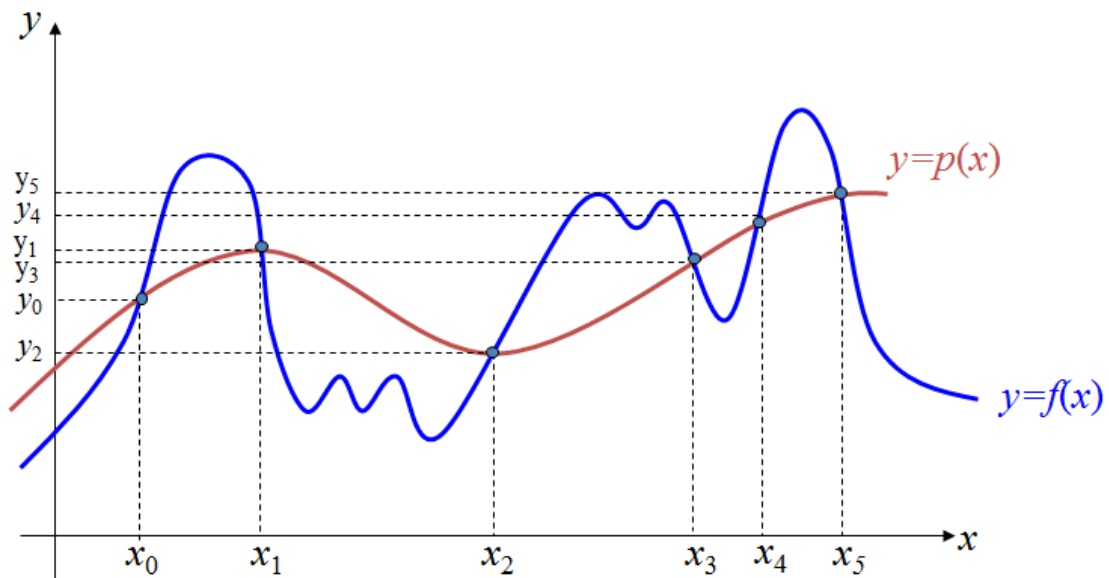


Figura 10.1: Representação Geométrica do Procedimento de Interpolação Polinomial.

### 10.1.1 Interpolação Linear

Um polinômio interpolador pode ser calculado conforme a prova do Teorema do 10.2.

**Exemplo 25.** Seja uma função  $y = f(x)$  definida pelos seguintes pontos:

$i$	0	1	2
$x_i$	-1	0	2
$y_i$	4	1	-1

O polinômio interpolador da função  $f(x)$  de grau 2 possui a seguinte forma padrão  $P_2(x) = a_2x^2 + a_1x + a_0$ . Seus coeficientes são determinados pela solução de um sistema linear cujas equações são obtidas escrevendo o polinômio explicitamente em cada ponto (substituindo cada ponto da função no polinômio), ou seja:

$$P_2(-1) = f(-1) = 4 \Rightarrow a_2(-1)^2 + a_1(-1) + a_0 = 4$$

$$P_2(0) = f(0) = 1 \Rightarrow a_2(0)^2 + a_1(0) + a_0 = 1$$

$$P_2(2) = f(2) = -1 \Rightarrow a_2(2)^2 + a_1(2) + a_0 = -1$$

Logo, tem-se o sistema linear que fornece os valores dos coeficientes do polinômio  $P_2(x)$ :

$$\begin{cases} a_2 - a_1 + a_0 = 4 \\ a_0 = 1 \\ 4a_2 + 2a_1 + a_0 = -1 \end{cases}$$

Resolvendo o sistema por meio de um método direto obtém-se o seguinte vetor solução  $a = \begin{bmatrix} 1 & -2,3333 & 0,6667 \end{bmatrix}^t$ . Logo, o polinômio interpolador obtido é:

$$P_2(x) = 0,6667x^2 - 2,333x + 1$$

---

**Observação**

---

**Nos pontos tabelados, o valor do polinômio encontrado e o valor da função devem coincidir. Se os valores forem diferentes, você terá cometido erros de cálculo!**

Na prática, a determinação do polinômio interpolador através da solução de um sistema de equações é muito trabalhosa, especialmente quando se deseja determinar polinômios de ordem mais elevada. Além disso, na solução do sistema linear podem ocorrer erros de arredondamento, o que faz a solução obtida se distanciar da solução exata.

Uma vez que o polinômio interpolador é único, é possível escrevê-lo usando outras formas de uso mais fácil que por resolução de sistemas lineares e que exigem um esforço computacional menor.

### 10.1.2 Erro de Truncamento

Embora o polinômio interpolador coincida com a função nos pontos de interpolação, ou seja,  $P(x_i) = f(x_i) = y_i$ ,  $i = 0, 1, \dots, n$ , nos pontos  $\bar{x} \neq x_i$  nem sempre é verdade que  $P(\bar{x}) = f(\bar{x})$ .

**Espera-se que  $P(\bar{x}) \approx f(\bar{x})$  para  $\bar{x} \neq x_i, i = 0, 1, \dots, n$ , e comete-se um erro nesta aproximação (veja a Figura 10.2), dado por:**

$$E(\bar{x}) = f(\bar{x}) - P(\bar{x}).$$

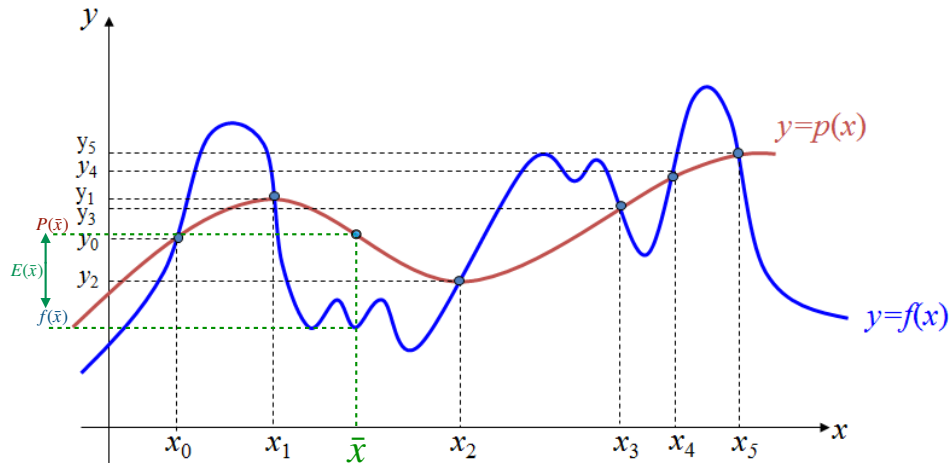


Figura 10.2: Representação Geométrica do erro de truncamento  $E(\bar{x}) = f(\bar{x}) - P(\bar{x})$ .

A expressão geral para o erro cometido quando aproximamos uma função  $f(x)$  por um polinômio interpolador  $P(x)$  é dado pelo Teorema 10.3.

**Teorema 10.3.** *Sejam  $(x_i, y_i), i = 0, 1, \dots, n$  pontos com abscissas distintas relacionados a uma função  $y = f(x)$  com  $(n + 1)$  derivadas contínuas no intervalo  $[x_0, x_n]$ . Então, para cada  $x \in [x_0, x_n]$ , existe um número  $\xi \in (x_0, x_n)$ , que depende de  $x$ , tal que*

$$E_t(x) = f(x) - P(x) = \prod_{i=0}^n (x - x_i) \frac{f^{(n+1)}(\xi)}{(n+1)!} \quad (10.1)$$

onde  $f^{(n+1)}(\cdot)$  é a derivada de ordem  $(n + 1)$  de  $y = f(x)$  e  $P(x)$  é o polinômio que a interpola nos pontos  $(x_i, y_i), i = 0, 1, \dots, n$ .

A expressão da Equação 10.1 é chamada de termo do erro ou **erro de truncamento** e corresponde ao erro que se comete, no ponto  $x$ , quando se substitui a função pelo polinômio que a interpola. A importância do Teorema 10.3 é mais teórica do que prática, uma vez que não é possível determinar o ponto  $\xi$ . Na prática, para estimar o erro cometido, quando a função é dada analiticamente, utiliza-se o Corolário 10.1.

**Corolário 10.1. (Limitante superior para o erro)** *Se  $f(x)$  e suas derivadas até a ordem  $(n + 1)$  são contínuas no intervalo  $[x_0, x_n]$ , então o limitante superior para o erro pode ser calculado por:*

$$|E_t(x)| \leq \left| \prod_{i=0}^n (x - x_i) \right| \frac{M}{(n+1)!} \quad (10.2)$$

em que  $M = \max |f^{(n+1)}(x)|$  no intervalo  $[x_0, x_n]$ .

**Exemplo 26.** Sabendo-se que os pontos a seguir são da função  $f(x) = xe^{3x}$ , calcule um limitante superior para o erro de truncamento quando se avalia  $y$  para  $x = 0,25$ .

$x_i$	0,2	0,3	0,4
$f(x_i)$	1,8221	2,4596	3,3201

**Solução:** Do Corolário 10.1, tem-se que:

$$E_t(x) \leq |(x - x_0)(x - x_1)(x - x_2)| \frac{M}{(3!)}$$

onde  $M = \max |f'''(x)|$  no intervalo  $[0, 2, 0, 4]$ . Dado  $f(x) = xe^{3x}$ , segue que:

$$f'(x) = (1 + 3x)e^{3x};$$

$$f''(x) = (6 + 9x)e^{3x};$$

$$f'''(x) = 27(1 + x)e^{3x}.$$

Assim, no intervalo  $[0, 2, 0, 4]$  tem-se que  $f'''(x) = 27(1 + x)e^{3x}$  é máxima para  $x = 0, 4$ . Logo,  $M = f'''(0, 4) = 125, 4998$ .

Segue que o limitante superior para o erro de truncamento é então calculado por:

$$E_t(0, 25) \leq |(0, 25 - 0, 2)(0, 25 - 0, 3)(0, 25 - 0, 4)| \frac{125, 4998}{(3!)} = 0, 0078$$

O que significa que o **erro máximo cometido**, ao usar o polinômio interpolador de Lagrange de grau 2 no intervalo  $[0, 2, 0, 4]$  para estimar  $f(0, 25)$ , é igual a 0, 0078.

---

### Observações

---

1. É possível calcular uma estimativa para o erro de truncamento somente quando a expressão analítica da função  $f(x)$  é conhecida, pois, de acordo com a expressão do limitante superior para o erro (Equação 10.2) é preciso calcular a  $(n + 1)$ -ésima derivada de  $f(x)$ . Nos casos em que a função for apresentada apenas por uma tabela de pontos, sabe-se que existe um erro de truncamento no ponto a ser avaliado, porém não é possível estimar exatamente o valor deste erro.
  2.  $P(x)$  não necessariamente converge para  $y = f(x)$  em  $[a, b]$  com o aumento do número de pontos de interpolação. Polinômios interpoladores de grau elevado podem produzir grandes oscilações nos extremos do intervalo. Trata-se do Fenômeno de Runge. Este fenômeno demonstra que polinômios de grau elevado são normalmente pouco recomendáveis para a interpolação porque aumentam o erro em valores próximos aos extremos do intervalo de interpolação e melhoram a aproximação em valores próximos ao centro. O problema pode ser evitado usando interpolação polinomial por partes com polinômios de grau moderado. Desta forma, pode-se tentar diminuir o erro de interpolação aumentando o número de peças de polinômios usadas, em vez de aumentar o grau do polinômio. Exemplos típicos: interpolação linear por partes (uma reta para cada par de pontos) e interpolação quadrática por partes (uma parábola para cada três pontos) e curvas *spline*.
-

## 10.2 Polinômios Interpoladores de Lagrange



Seja  $f(x)$  uma função definida em  $x_0, x_1, \dots, x_n$  ( $n+1$ ) pontos distintos de um intervalo fechado  $[a, b]$  e  $y_i = f(x_i)$ ,  $i = 1, 2, \dots, n$ . Considere o polinômio na forma:

$$P(x) = y_0 L_0(x) + y_1 L_1(x) + \dots + y_i L_i(x) + \dots + y_n L_n(x) = \sum_{i=0}^n y_i L_i(x)$$

em que  $P(x)$  é um polinômio interpolador, ou seja,  $P(x_i) = y_i$ ,  $i = 0, 1, \dots, n$ . Assim:

$$P(x_i) = y_0 L_0(x_i) + y_1 L_1(x_i) + \dots + y_i L_i(x_i) + \dots + y_n L_n(x_i) = \sum_{i=0}^n y_i L_i(x_i) = y_i,$$

para  $i = 0, 1, \dots, n$ . Os polinômios  $L_i(x_i)$  são chamados de polinômios base de interpolação de Lagrange e constituem uma base completa de funções polinomiais de mesmo grau  $n$ , em que qualquer polinômio de grau  $n$  pode ser expresso como uma combinação linear destes polinômios. Os polinômios  $L_i(x_i)$  são tais que:

$$\begin{cases} L_i(x_i) = 1 \\ L_i(x_j) = 0 \quad i, j = 0, 1, \dots, n; \text{ para } i \neq j \end{cases}$$

Para que o polinômio  $P(x)$  satisfaça esta propriedade, deve-se considerar que

$$L_i(x) = \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}, \quad i = 0, 1, 2, \dots, n$$

ou ainda

$$L_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{(x - x_j)}{(x_i - x_j)}, \quad i = 0, 1, 2, \dots, n. \quad (10.3)$$

Assim, o **método de Lagrange** para a obtenção do polinômio interpolador  $P(x)$  é dado por:

$$P(x) = \sum_{i=0}^n y_i L_i(x)$$

onde os polinômios base de Lagrange  $L_i(x)$  são calculados conforme a Equação 10.3.

**Exemplo 27.** Seja uma função  $y = f(x)$  definida nos pontos, conforme a tabela:

$i$	0	1	2
$x_i$	-1	0	2
$y_i$	4	1	-1

Determine o polinômio interpolador de Lagrange.

**Solução:** O polinômio interpolador de Lagrange de grau menor ou igual a 2 é dado por

$$P_2(x) = \sum_{i=0}^2 y_i L_i(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x).$$

As funções  $L_i(x)$  são calculadas por:

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 0)(x - 2)}{(-1 - 0)(-1 - 2)} = \frac{x^2 - 2x}{3}$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x + 1)(x - 2)}{(0 + 1)(0 - 2)} = \frac{x^2 - x - 2}{-2}$$

$$L_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x + 1)(x - 0)}{(2 + 1)(2 - 0)} = \frac{x^2 + x}{6}$$

Assim,

$$P_2(x) = 4 \left( \frac{x^2 - 2x}{3} \right) + 1 \left( \frac{x^2 - x - 2}{-2} \right) + (-1) \left( \frac{x^2 + x}{6} \right) = 0,6667x^2 - 2,333x + 1.$$

**Exemplo 28.** Seja uma função  $y = f(x)$  definida nos pontos, conforme a tabela:

$i$	0	1	2
$x_i$	0	0.5	1.0
$y_i$	1.3	2.5	0.9

Determine o polinômio interpolador de Lagrange e estime o valor de  $f(0,8)$ .

**Solução:** O polinômio interpolador de Lagrange de grau menor ou igual a 2 é dado por

$$P_2(x) = \sum_{i=0}^2 y_i L_i(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x)$$

As funções  $L_i(x)$  são calculadas por:

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 0,5)(x - 1)}{(0 - 0,5)(0 - 1)} = \frac{x^2 - 1,5x + 0,5}{0,5}$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 0)(x - 1)}{(0,5 - 0)(0,5 - 1)} = \frac{x^2 - x}{-0,25}$$

$$L_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 0)(x - 0,5)}{(1 - 0)(1 - 0,5)} = \frac{x^2 - 0,5x}{0,5}$$

Assim,

$$P_2(x) = (1,3) \left( \frac{x^2 - 1,5x + 0,5}{0,5} \right) + (2,5) \left( \frac{x^2 - x}{-0,25} \right) + (0,9) \left( \frac{x^2 - 0,5x}{0,5} \right) = -5,6x^2 + 5,2x + 1,3$$

e  $f(0,8) \approx P_2(0,8) = -5,6(0,8)^2 + 5,2(0,8) + 1,3 = 1,8760$ .

No Exemplo 28, existe um erro de truncamento cometido ao fazer a aproximação da função no ponto  $x = 0,8$ . Porém, não é possível estimar exatamente o valor deste erro, visto que a expressão analítica da função  $f(x)$  não é conhecida.

**Exemplo 29.** Considere a função  $f(x) = \cos(x)$  definida nos pontos, conforme a tabela abaixo:

$x_i$	0,9	1	1,1	1,3
$y_i$	0,62160	0,54030	0,45359	0,26749

Determine o polinômio interpolador de  $f(x)$ , usando o método de Lagrange, estime o valor de  $f(1,2)$  e um limitante superior para o erro.

**Solução:** O polinômio interpolador de Lagrange de grau menor ou igual a 3 é dado por

$$P_3(x) = \sum_{i=0}^3 y_i L_i(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x) + y_3 L_3(x).$$

As funções  $L_i(x)$  são calculadas por:

$$L_0(x) = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} = \frac{(1,2-1)(1,2-1,1)(1,2-1,3)}{(0,9-1)(0,9-1,1)(0,9-1,3)} \\ \Rightarrow L_0(1,2) = \frac{-0,002}{-0,008} = 0,25$$

$$L_1(x) = \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)} = \frac{(1,2-0,9)(1,2-1,1)(1,2-1,3)}{(1-0,9)(1-1,1)(1-1,3)} \\ \Rightarrow L_1(1,2) = \frac{-0,003}{0,003} = -1$$

$$L_2(x) = \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} = \frac{(1,2-0,9)(1,2-1)(1,2-1,3)}{(1,1-0,9)(1,1-1)(1,1-1,3)} \\ \Rightarrow L_2(1,2) = \frac{-0,006}{-0,004} = 1,5$$

$$L_3(x) = \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)} = \frac{(1,2-0,9)(1,2-1)(1,2-1,1)}{(1,3-0,9)(1,3-1)(1,3-1,1)} \\ \Rightarrow L_3(1,2) = \frac{0,006}{0,024} = 0,25$$

Assim,

$$P_3(1,2) = y_0 L_0(1,2) + y_1 L_1(1,2) + y_2 L_2(1,2) + y_3 L_3(1,2)$$

$$P_3(1,2) = 0,62160(0,25) + 0,54030(-1) + 0,45359(1,5) + 0,26749(0,25) = 0,362358.$$

Então,

$$f(1,2) \approx P_2(1,2) = 0,362358.$$

**Cálculo do limitante superior para o erro a partir da fórmula:**

$$|E_t(x)| \leq \left| \prod_{i=0}^n (x-x_i) \right| \frac{M}{(n+1)!} \quad \text{com} \quad M = \max |f^{(n+1)}(x)|, \quad x \in [x_0, x_n].$$

Para  $n = 3$ , temos:

$$E_t(x) \leq |(x-x_0)(x-x_1)(x-x_2)(x-x_3)| \frac{M}{(4!)} \quad \text{com} \quad M = \max |f^{(4)}(x)|, \quad x \in [0,9, 1,3].$$

Como  $f^{(4)}(x) = \cos(x)$  e a função é decrescente em módulo no intervalo  $[0,9, 1,3]$ , tem-se que  $f^{(4)}(x)$



assume o valor máximo em  $x = 0,9$ , ou seja:

$$M = \max |f^{(4)}(x)| = \cos(0,9) = 0,62160.$$

Logo, o limitante para o erro de truncamento no ponto interpolado  $x = 1,2$  é dado por:

$$E_t(1,2) \leq |(1,2 - 0,9)(1,2 - 1)(1,2 - 1,1)(1,2 - 1,3)| \frac{(0,62160)}{(4!)} = \frac{0,00037}{24} = 0,000015.$$

Como no Exemplo 29 a função a ser interpolada é conhecida ( $f(x) = \cos(x)$ ) é possível calcular o erro absoluto. Assim,

$$\text{Erro absoluto} = |\cos(1,2) - P(1,2)| = |0,362357 - 0,362358| = 0,000001.$$

Note que o erro absoluto é um valor menor que o limitante superior do erro de truncamento determinado.

---

#### Observação

---

O método de Lagrange para a determinação do polinômio interpolador de uma função  $f(x)$  sobre os pontos  $x_0, x_1, \dots, x_n$  possui um inconveniente. Sempre que se deseja aumentar (ou diminuir) o grau do polinômio, todo o trabalho precisa ser refeito. As fórmulas de interpolação estudadas nas próximas seções permitem a troca do grau do polinômio interpolador apenas acrescentando (ou retirando) um termo do polinômio interpolador.

---



---

#### Exercício Resolvido

---

Seja  $y = f(x)$  tabelada:

$i$	0	1	2
$x_i$	2,5	3,0	3,5
$y_i$	0,503	0,806	1,192

Determine o polinômio interpolador de Lagrange que intercepta  $f(x)$  nos pontos fornecidos e calcule o valor aproximado da função no ponto de abscissa  $x = 2,8$  utilizando 3 casas decimais.

**Solução:** Assista o vídeo a seguir:



Vídeo elaborado pela Estagiária em Docência Débora Nasser Diniz.

---

## 10.3 Exercícios



E. 1. Calcule a partir da tabela de  $y = \sin(x)$

x	y
0,3	0,2955
0,4	0,3894
0,5	0,4794

- a) um polinômio interpolador linear ( $P_1(x)$ ) para os 2 primeiros pontos da tabela.  
b) um polinômio interpolador quadrático ( $P_2(x)$ ).

utilizando os polinômios obtidos nas letras (a) e (b), determine:

- c)  $P_1(0,33)$       d)  $P_2(0,33)$       e)  $P_1(0,38)$       f)  $P_2(0,38)$

E. 2. Seja a função  $f(x) = 2e^x + 3$  definida no intervalo  $[0, 2]$  conforme a tabela a seguir:

$i$	0	1	2	3	4	5
$x_i$	0	0,2	0,5	1	1,3	2
$y_i$	5	5,4428	6,2974	8,4366	10,3385	17,7781

- a) Aproxime  $f(0,35)$  utilizando um polinômio interpolador de Lagrange de grau 1.  
b) Aproxime  $f(0,35)$  utilizando um polinômio interpolador de Lagrange de grau 2.  
c) Aproxime  $f(0,35)$  utilizando um polinômio interpolador de Lagrange de grau 3.  
d) Calcule o erro de truncamento máximo cometido nas letras anteriores.  
e) Em qual dos casos obtemos a melhor aproximação no ponto desejado? Justifique.



[illegible]





# 11

## Método de Diferenças Divididas

### Sumário da Aula

11.1 Introdução . . . . .	105
11.2 Polinômio Interpolador de Diferenças Divididas . . . . .	109
11.3 Exercícios . . . . .	114

### 11.1 Introdução

Nesta seção será apresentada a fórmula interpolatória de Newton, a qual é construída a partir do operador diferenças divididas.



**Definição 11.1.** Dada uma função  $y = f(x)$ , a sua primeira derivada é definida como:

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

sendo  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ ; um conjunto de pontos da função, então:

$$f'(x_i) = \lim_{h \rightarrow 0} \frac{f(x_i+h) - f(x_i)}{h}$$

Seja  $x_i + h = x_{i+1} \Rightarrow h = x_{i+1} - x_i$ , logo

$$f'(x_i) = \lim_{x_i \rightarrow x_{i+1}} \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \quad (11.1)$$

**Definição 11.2. (Operador Diferença Dividida)** Seja  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ ; um conjunto de pontos, com abscissas distintas, de uma função  $f(x)$ , define-se o operador de diferença dividida de primeira ordem como:

$$Dy_i = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, i = 0, 1, \dots, n-1 \quad (11.2)$$

Observe-se que este operador nada mais é do que uma aproximação do valor numérico da primeira derivada de uma função em um ponto. Pode ser demonstrado que as diferenças divididas de ordem superior são aproximações para as derivadas de ordem superior.

A diferença dividida de segunda ordem é definida como:

$$D^2y_i = \frac{Dy_{i+1} - Dy_i}{x_{i+2} - x_i}, i = 0, 1, \dots, n-2 \quad (11.3)$$

E a diferença dividida de terceira ordem é definida como:

$$D^3y_i = \frac{D^2y_{i+1} - D^2y_i}{x_{i+3} - x_i}, i = 0, 1, \dots, n-3 \quad (11.4)$$

Considerando as definições de diferenças de ordens um, dois e três, representadas, respectivamente, pelas Equações 11.2, 11.3 e 11.4, tem-se que a diferença dividida de ordem  $k$  é definida como:

$$D^k y_i = \frac{D^{k-1} y_{i+1} - D^{k-1} y_i}{x_{i+k} - x_i}, \begin{cases} k = 1, 2, \dots, n \\ i = 0, 1, 2, \dots, n-k \end{cases} \quad (11.5)$$

E a diferença de ordem zero sendo  $D^0 y_i = y_i$ ,  $i = 0, 1, \dots, n$ .

De forma geral, e usando uma notação mais simples, dada uma função  $f(x)$  contínua,  $(n+1)$  vezes diferenciável e definida em  $x_0, x_1, \dots, x_n$ ,  $(n+1)$  pontos distintos de um intervalo  $[a, b]$ :

- define-se por diferença dividida de ordem zero de  $f(x)$ :

$$f[x_i] = f(x_i), i = 0, 1, \dots, n$$

- definimos diferença dividida de ordem um de  $f(x)$  por:

$$f[x_i, x_{i+1}] = \frac{f[x_{i+1}] - f[x_i]}{(x_{i+1} - x_i)}, i = 0, 1, \dots, n$$

- define-se por diferença dividida de ordem dois de  $f(x)$ :

$$f[x_i, x_{i+1}, x_{i+2}] = \frac{f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]}{(x_{i+2} - x_i)}, \quad i = 0, 1, \dots, n$$

•  $\vdots$

- define-se por diferença dividida de ordem  $n$  de  $f(x)$ :

$$f[x_0, x_1, \dots, x_n] = \frac{f[x_1, \dots, x_n] - f[x_0, x_1, \dots, x_{n-1}]}{(x_n - x_0)}, \quad i = 0, 1, \dots, n.$$

Dada uma função  $f(x)$  e os valores que a função assume nos  $(n + 1)$  pontos distintos  $x_0, x_1, \dots, x_n$ , pode-se tabular as diferenças divididas da seguinte forma:

Pontos	(Ordem 0)	(Ordem 1)	(Ordem 2)	...	(Ordem n-1)	(Ordem n)
$x$	$D^0 f(x_i)$	$D f(x_i)$	$D^2 f(x_i)$	...	$D^{n-1} f(x_i)$	$D^n f(x_i)$
$x_0$	$f[x_0]$	$f[x_0, x_1]$	$f[x_0, x_1, x_2]$	...	$f[x_0, x_1, \dots, x_{n-1}]$	$f[x_0, x_1, \dots, x_n]$
$x_1$	$f[x_1]$	$f[x_1, x_2]$	$f[x_1, x_2, x_3]$	...	$f[x_1, x_2, \dots, x_{n-1}]$	
$x_2$	$f[x_2]$	$f[x_2, x_3]$	$f[x_2, x_3, x_4]$	...		
$x_3$	$f[x_3]$	$f[x_3, x_4]$	$f[x_3, x_4, x_4]$			
$\vdots$	$\vdots$	$\vdots$	$\vdots$			
$x_{n-2}$	$f[x_{n-2}]$	$f[x_{n-2}, x_{n-1}]$	$f[x_{n-2}, x_{n-1}, x_n]$			
$x_{n-1}$	$f[x_{n-1}]$	$f[x_{n-1}, x_n]$				
$x_n$	$f[x_n]$					

Tabela 11.1: Operadores de Diferenças Divididas.

Note que as diferenças de ordem 1 são calculadas a partir das diferenças de ordem zero, as diferenças de ordem 2, a partir das diferenças de ordem 1, e assim, sucessivamente. De forma geral, a Tabela 11.1 é construída da seguinte maneira:

- a primeira coluna é constituída pelos pontos  $x_i$ , com  $i = 0, 1, \dots, n$ ;
- a segunda coluna contém os valores de  $f(x)$  nos pontos  $x_i, i = 0, 1, \dots, n$ ;
- as colunas 3, 4, ... contém as diferenças divididas de ordem 1, 2, ...,  $n$ . Cada uma destas diferenças é uma fração cujo numerador é sempre a diferença entre duas diferenças consecutivas e de ordem imediatamente inferior, e o denominador é a diferença entre os dois extremos dos pontos envolvidos.



**Exemplo 30.** Construa a tabela de diferenças divididas da função  $f(x) = \frac{1}{x}$  definida sobre os pontos  $x_0 = 1$ ,  $x_1 = 2$ ,  $x_2 = 4$  e  $x_3 = 5$ .

**Solução:** Cálculo das diferenças divididas:

- de ordem zero:

$$\begin{aligned} f[x_0] &= f(x_0) = 1 & f[x_2] &= f(x_2) = 1/4 \\ f[x_1] &= f(x_1) = 1/2 & f[x_3] &= f(x_3) = 1/5 \end{aligned}$$

- de ordem um:

$$\begin{aligned} f[x_0, x_1] &= \frac{f[x_1] - f[x_0]}{x_1 - x_0} = \frac{(1/2 - 1)}{2 - 1} = -1/2 \\ f[x_1, x_2] &= \frac{f[x_2] - f[x_1]}{x_2 - x_1} = \frac{(1/4 - 1/2)}{4 - 2} = -1/8 \\ f[x_2, x_3] &= \frac{f[x_3] - f[x_2]}{x_3 - x_2} = \frac{(1/5 - 1/4)}{5 - 4} = -1/20 \end{aligned}$$

- de ordem dois:

$$\begin{aligned} f[x_0, x_1, x_2] &= \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = \frac{(-1/8 + 1/2)}{4 - 1} = \frac{3/8}{3} = 1/8 \\ f[x_1, x_2, x_3] &= \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1} = \frac{(-1/20 + 1/8)}{5 - 2} = \frac{3/40}{3} = 1/40 \end{aligned}$$

- de ordem três:

$$f[x_0, x_1, x_2, x_3] = \frac{f[x_1, x_2, x_3] - f[x_0, x_1, x_2]}{x_3 - x_0} = \frac{(1/40 - 1/8)}{5 - 1} = \frac{-1/10}{4} = -1/40$$

Os cálculos anteriores estão organizados na tabela seguinte:

$x$	Ordem 0	Ordem 1	Ordem 2	Ordem 3
1	1	-1/2	1/8	-1/40
2	1/2	-1/8	1/40	
4	1/4	-1/20		
5	1/5			

Tabela 11.2: Tabelamento das diferenças divididas da função  $f(x) = 1/x$  nos pontos  $x_0 = 1$ ,  $x_1 = 2$ ,  $x_2 = 4$  e  $x_3 = 5$ .

O Teorema 11.1 e os Corolários 11.1 e 11.2 são **resultados sobre Diferenças Divididas** e são usados na Seção 11.2 para definir o polinômio interpolador de diferenças divididas.

**Teorema 11.1.** As diferenças divididas de ordem  $k$  de uma função  $f(x)$  satisfazem:

$$f[x_0, x_1, \dots, x_k] = \sum_{i=0}^k \frac{f(x_i)}{(x_i - x_0) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_k)}.$$

**Corolário 11.1.** As diferenças divididas de ordem  $k$  de uma função  $f(x)$  satisfazem:

$$f[x_0, x_1, \dots, x_k] = f[x_{j_0}, x_{j_1}, \dots, x_{j_k}].$$

onde  $(j_0, j_1, \dots, j_k)$  é qualquer permutação dos inteiros  $(0, 1, \dots, k)$ .

O resultado deste corolário mostra que a diferença dividida de  $f(x)$  é uma função simétrica de seus argumentos, ou seja, independe da ordem dos pontos  $x_0, x_1, \dots, x_k$ .

**Corolário 11.2.** As diferenças divididas de ordem  $k$  de uma função  $f(x)$  satisfazem:

$$f[x_0, x_1, \dots, x_k] = \frac{f[x_0, \dots, x_{i-1}, x_{i+1}, \dots, x_k] - f[x_0, \dots, x_{j-1}, x_{j+1}, \dots, x_k]}{x_j - x_i}, \quad i \neq j.$$

O resultado deste corolário mostra que podemos retirar quaisquer dois pontos distintos para construir a diferença dividida de uma função, e não necessariamente o primeiro e o último termo.

## 11.2 Polinômio Interpolador de Diferenças Divididas

Neste método, a ideia é obter o polinômio,  $P(x)$ , que interpola uma função,  $y = f(x)$ , em um conjunto de pontos  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ , baseando-se nos resultados obtidos das diferenças divididas.

Considere uma função  $f(x)$  contínua no intervalo  $[a, b]$  e definida em  $x_0, x_1, \dots, x_n$ ,  $(n+1)$  pontos distintos. As diferenças divididas de  $f(x)$  nos pontos:

- $x_0$  e  $x$ :

$$f[x_0, x] = \frac{f[x] - f[x_0]}{(x - x_0)}, \quad x \neq x_0.$$

Logo,

$$f(x) = f(x_0) + (x - x_0)f[x_0, x] \quad (11.6)$$

- $x_0, x_1$  e  $x$ :

$$f[x_0, x_1, x] = \frac{f[x_0, x] - f[x_0, x_1]}{(x - x_1)}, \quad x \neq x_0 \text{ e } x \neq x_1 \quad (11.7)$$

Substituindo  $f[x_0, x]$  da Equação 11.6 na Equação 11.7:

$$f(x) = f[x_0] + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x]$$

- De forma sucessiva:

$$f[x_0, x_1, \dots, x_n, x] = \frac{f[x_0, x_1, \dots, x_{n-1}, x] - f[x_0, x_1, \dots, x_n]}{(x - x_n)}, \quad x \neq x_n \quad (11.8)$$

Logo,

$$\begin{aligned} f(x) &= f[x_0] + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x_2] + \dots + \\ &+ (x - x_0)(x - x_1) \dots (x - x_{n-1})f[x_0, x_1, \dots, x_n] + \\ &+ (x - x_0)(x - x_1) \dots (x - x_{n-1})(x - x_n)f[x_0, x_1, \dots, x_n, x]. \end{aligned} \quad (11.9)$$

Assim, pode-se escrever:

$$f(x) = P(x) + R(x) \quad (11.10)$$

em que

$$\begin{aligned} P(x) &= f[x_0] + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x_2] + \dots + \\ &+ (x - x_0)(x - x_1) \dots (x - x_{n-1})f[x_0, x_1, \dots, x_n] \\ R(x) &= (x - x_0)(x - x_1) \dots (x - x_{n-1})(x - x_n)f[x_0, x_1, \dots, x_n, x] \end{aligned}$$

---

### Observação

---

A equação

$$R(x) = f(x) - P(x) \quad (11.11)$$

é uma nova expressão para o erro de truncamento  $E(x)$ .

---

**Teorema 11.2.** *Seja uma função  $f(x)$  contínua e definida nos pontos  $x_0, x_1, \dots, x_n$  distintos de um intervalo  $[a, b]$ . O polinômio:*

$$P_n(x) = f[x_0] + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x_2] + \dots + (x - x_0)(x - x_1) \dots (x - x_{n-1})f[x_0, x_1, \dots, x_n] \quad (11.12)$$

*é o polinômio interpolador de grau menor ou igual a  $n$  da função  $f(x)$  sobre os pontos  $x_0, x_1, \dots, x_n$ , isto é*

$$P_n(x_i) = f(x_i), \quad i = 1, 2, \dots, n$$

**Prova:** A prova deste teorema pode ser encontrada em [7].

A Equação 11.12 é chamada de **Fórmula de Newton do Polinômio de Interpolação** ou **Polinômio Interpolador de Diferenças Divididas**

**Teorema 11.3.** *Seja  $f(x)$  uma função contínua e suficientemente diferenciável no intervalo  $[a, b]$  e definida em  $x_0, x_1, \dots, x_n$ ,  $(n + 1)$  pontos deste intervalo. Então, para  $x \in [a, b]$  e  $x \neq x_i$ ,  $i = 0, 1, \dots, n$ , temos que:*

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

**Prova:** Do Teorema 10.3:

$$E_t(x) = f(x) - P(x) = \prod_{i=0}^n (x - x_i) \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

Porém,

$$R(x) = f(x) - P(x) = (x - x_0)(x - x_1) \dots (x - x_{n-1})(x - x_n)f[x_0, x_1, \dots, x_n, x].$$

Portanto, para  $x \neq x_i$ :

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

■

As expressões do erro e do limitante superior do erro são as mesmas para todas as fórmulas interpolatórias:

$$|E_t(x)| \leq \left| \prod_{i=0}^n (x - x_i) \right| \frac{M}{(n+1)!}, \quad \text{com } M = \max |f^{(n+1)}(x)|, \quad x \in [x_0, x_n].$$

Na ausência da informação sobre  $f^{(n+1)}(x)$  da expressão anterior, uma estimativa para o erro de truncamento máximo pode ser obtida utilizando uma diferença dividida de ordem  $(n + 1)$ , caso estas não variem muito. Assim, a partir dos Teoremas 10.3 e 11.3, o erro de truncamento gerado pelo método de interpolação de diferenças divididas também pode ser calculado por:

$$E_t(x) = \prod_{i=0}^n (x - x_i) f[x_0, x_1, \dots, x_n, x]. \quad (11.13)$$

E o limitante superior por:

$$|E_t(x)| \leq \left| \prod_{i=0}^n (x - x_i) \right| \max |f[x_0, x_1, \dots, x_n, x_{n+1}]|, \quad x \in [x_0, x_n]. \quad (11.14)$$

## Exemplos:



**Exemplo 31.** A tabela a seguir apresenta valores da voltagem (V) em função da corrente elétrica (I). Utilizando interpolação polinomial, método das diferenças divididas, estime o valor de V quando  $I = 3A$ .

i	0	1	2	3
$I = x_i$	1	2	4	8
$V = y_i$	120	94	75	62

**Solução:** Neste exercício, temos um polinômio interpolador de grau menor ou igual a 3 dado por:

$$P_3(x) = f[x_0] + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x_2] + (x - x_0)(x - x_1)(x - x_2)f[x_0, x_1, x_2, x_3]$$

Inicialmente, são determinados os valores das diferenças divididas:

$i$	$I = x_i$	$V = y_i$	Ordem 1	Ordem 2	Ordem 3
0	1	120	-26	5,5	-0,64
1	2	94	-9,5	1,04	
2	4	75	-3,25		
3	8	62			

Logo:

$$P_3(x) = 120 + (x - 1)(-26) + (x - 1)(x - 2)(5,5) + (x - 1)(x - 2)(x - 4)(-0,64)$$

$$P_3(3) = 120 + (3 - 1)(-26) + (3 - 1)(3 - 2)(5,5) + (3 - 1)(3 - 2)(3 - 4)(-0,64)$$

$$P_3(3) = 80,28V$$

**Exemplo 32.** Considere a função  $f(x) = e^x + \sin(x)$  tabelada, como segue:

i	0	1	2
$x_i$	0	0,5	1,0
$y_i$	1	2,1281	3,5598

Determine o polinômio interpolador usando o polinômio interpolador de diferenças divididas, avalie  $f(0,7)$  e um limitante superior para o erro.

**Solução:** Neste exercício, temos um polinômio interpolador de grau menor ou igual a 2 dado por:

$$P_2(x) = f[x_0] + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x_2]$$

Inicialmente, são determinados os valores das diferenças divididas:

$i$	$x_i$	$f(x_i)$	Ordem 1	Ordem 2
0	0	1	2,2562	0,6072
1	0,5	2,1281	2,8634	
2	1	3,5598		

Logo:

$$P_2(x) = 1 + (x - 0)(2,2562) + (x - 0)(x - 0,5)(0,6072) = 0,6072x^2 + 1,9526x + 1$$

$$P_2(0,7) = 1 + (0,7 - 0)(2,2562) + (0,7 - 0)(0,7 - 0,5)(0,6072)$$

$$P_2(0,7) = 2,6643$$

Assim,  $f(0,7) \approx P_2(0,7) = 2,6643$ .

Para avaliar o **limitante superior para o erro**, considere:

$$|E_t(x)| \leq \left| \prod_{i=0}^n (x - x_i) \right| \frac{M}{(n+1)!}, \quad \text{com } M = \max |f^{(n+1)}(x)|, \quad x \in [x_0, x_n].$$

Assim, para  $n = 2$ :

$$|E_t(x)| \leq |(x - x_0)(x - x_1)(x - x_2)| \frac{\max |f^{(3)}(x)|}{(3!)} \quad x \in [0, 1].$$

$f^{(3)}(x) = e^x - \cos(x)$  é uma função crescente em módulo no intervalo  $[0, 1]$ . Então,  $f^{(3)}(x)$  assume o valor máximo em  $x = 1$ . Segue que:

$$M = \max |f^{(3)}(x)| = |f^{(3)}(1)| = 2,1780,$$

Assim, um limitante superior para o erro no ponto interpolador  $x = 0,7$  é dado por:

$$|E_t(0,7)| \leq |(0,7 - 0)(0,7 - 0,5)(0,7 - 1)| \frac{2,1780}{(3!)} = 0,0152.$$

Portanto,  $f(0,7) \approx P_2(0,7) = 2,6643$  com um erro máximo igual a 0,0152.

**Exemplo 33.** Dada a função tabelada por:

$x$	2	3	4	5	6	7
$f(x)$	3,69	6,70	13,65	29,68	67,24	156,66

Determine o polinômio interpolador de diferenças divididas de grau menor ou igual a 3, avalie  $f(4,5)$  e um limitante superior para o erro.

**Solução:** O polinômio interpolador de grau menor ou igual a 3 dado por:

$$P_3(x) = f[x_0] + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x_2] + (x - x_0)(x - x_1)(x - x_2)f[x_0, x_1, x_2, x_3]$$

Inicialmente, são determinados os valores das diferenças divididas. A tabela será calculada considerando todos os pontos informados de  $f(x)$ :

$i$	$x_i$	$f(x_i)$	Ordem 1	Ordem 1	Ordem 3	Ordem 4	Ordem 5
0	2	3,69	3,01	1,97	0,857	0,305	0,088
1	3	6,70	6,95	4,54	2,075	0,745	
2	4	13,65	16,03	10,765	5,055		
3	5	29,68	37,56	25,93			
4	6	67,24	89,42				
5	7	156,66					

Logo:

$$\begin{aligned}P_3(x) &= f[x_0] + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x_2] \\&\quad + (x - x_0)(x - x_1)(x - x_2)f[x_0, x_1, x_2, x_3] \\P_3(x) &= (3,69) + (x - 2)(3,01) + (x - 2)(x - 3)(1,97) + (x - 2)(x - 3)(x - 4)(0,857) \\P_3(4,5) &= (3,69) + (4,5 - 2)(3,01) + (4,5 - 2)(4,5 - 3)(1,97) + (4,5 - 2)(4,5 - 3)(4,5 - 4)(0,857) \\P_3(4,5) &= 20,2093\end{aligned}$$

Portanto,  $f(4,5) \approx P_3(4,5) = 20,209375$ .

Para avaliar o **limitante superior para o erro**, considere:

$$|E_t(x)| \leq \left| \prod_{i=0}^n (x - x_i) \right| \max |f[x_0, x_1, \dots, x_n, x_{n+1}]|, \quad x \in [x_0, x_n].$$

Assim, para  $n = 3$ :

$$|E_t(x)| \leq |(x - x_0)(x - x_1)(x - x_2)(x - x_3)| \max |f[x_0, x_1, x_2, x_3, x_4]|, \quad x \in [2, 7].$$

Para determinar o  $\max |f[x_0, x_1, x_2, x_3, x_4]|$  olhe para a coluna “**Ordem 4**” da tabela de diferenças divididas. Segue que:

$$\max |f[x_0, x_1, x_2, x_3, x_4]| = 0,745$$

Assim, o limitante superior para o erro no ponto interpolador  $x = 4,5$  é dado por:

$$E_t(4,5) \leq |(4,5 - 2)(4,5 - 3)(4,5 - 4)(4,5 - 5)| \cdot (0,745) = 0,6984.$$

Portanto,  $f(4,5) \approx P_2(4,5) = 20,209375$  com um erro máximo igual a 0,6984.



### Exercício Resolvido

---

Seja  $y = f(x)$  tabelada:

$i$	0	1	2
$x_i$	2,5	3,0	3,5
$y_i$	0,503	0,806	1,192

Determine o polinômio interpolador de Diferenças Divididas que intercepta  $f(x)$  nos pontos fornecidos e calcule o valor aproximado da função no ponto de abscissa  $x = 2,8$  utilizando 3 casas decimais.

**Solução:** Assista o vídeo a seguir:



Vídeo elaborado pela Estagiária em Docência Débora Nasser Diniz.

---

### 11.3 Exercícios



E. 1. Em um experimento foram obtidos os seguintes dados:

x	0	1,8	5	6	8,2	9,2	12
y	26	16,415	5,375	3,5	2,015	2,54	8

Obtenha um polinômio interpolador de grau 2 usando o método das diferenças divididas. Determine o valor de  $y$  em  $x = 3,5$  e calcule uma estimativa para o erro de truncamento.

E. 2. Seja a função de distribuição de probabilidade normal padrão definida por

$$N(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{t^2}{2}\right) dt$$

cujos valores são mostrados na tabela:

$z$	$N(z)$
0,0	0,50000
0,5	0,69146
1,0	0,84134
1,5	0,93319
2,0	0,97725
2,5	0,99379
3,0	0,99865

Calcule  $P_n(0,3)$  utilizando polinômios interpoladores de diferenças divididas de graus  $n = 1, 2, 3$  e 4.

[illegible]

This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.





# 12

## Método de Diferenças Finitas

### Sumário da Aula

12.1 Método das Diferenças Finitas . . . . .	117
12.1.1 Diferenças Finitas . . . . .	117
12.1.2 Polinômio Interpolador de Diferenças Finitas . . . . .	120
12.2 Complexidade dos métodos de interpolação . . . . .	123
12.3 Escolha dos pontos de interpolação . . . . .	124
12.4 Considerações Finais . . . . .	124
12.5 Exercícios . . . . .	125

## 12.1 Método das Diferenças Finitas

Nesta seção será apresentada a fórmula interpolatória de **Newton-Gregory** ou **polinômios interpoladores de diferenças finitas**, a qual é construída baseada nas diferenças finitas para um conjunto de pontos equidistantes.

### 12.1.1 Diferenças Finitas

Seja uma função  $f(x)$  contínua em um intervalo  $[a, b]$ . Sejam  $x_0, x_1, \dots, x_n$ ,  $n + 1$  pontos distintos do intervalo  $[a, b]$  tais que  $x_{i+1} - x_i = h$ ,  $i = 0, 1, \dots, n - 1$ , isto é, os pontos são equidistantes.



**Definição 12.1. (Diferença finita de ordem zero)** A diferença finita de ordem zero de uma função  $f(x)$  definida nos pontos  $x \in [a, b]$  é dada por:

$$\Delta^0 f(x) = f(x) \quad (12.1)$$

Em um ponto  $x_i$ , temos que  $\Delta^0 f(x_i) = f(x_i) = y_i$ .

**Definição 12.2. (Diferença finita de ordem n)** A diferença finita de ordem  $n$  de uma função  $f(x)$  definida nos pontos  $x \in [a, b]$  é dada por:

$$\Delta^n f(x) = \Delta^{n-1} f(x+h) - \Delta^{n-1} f(x) \quad (12.2)$$

A diferença finita de ordem  $n$  da Equação 12.2 em um ponto  $x_i$ :

$$\Delta^n f(x_i) = \Delta^{n-1} f(x_i+h) - \Delta^{n-1} f(x_i).$$

Dado que  $x_{i+1} - x_i = h \Rightarrow x_i + h = x_{i+1}$ :

$$\Delta^n f(x_i) = \Delta^{n-1} f(x_{i+1}) - \Delta^{n-1} f(x_i) \quad (12.3)$$

Assim, da Definição 12.2 podem ser obtidas as seguintes diferenças:

$$n = 0 \Rightarrow \Delta^0 f(x_i) = f(x_i)$$

$$n = 1 \Rightarrow \Delta^1 f(x_i) = \Delta^0 f(x_{i+1}) - \Delta^0 f(x_i)$$

$$n = 2 \Rightarrow \Delta^2 f(x_i) = \Delta^1 f(x_{i+1}) - \Delta^1 f(x_i)$$

$$n = 3 \Rightarrow \Delta^3 f(x_i) = \Delta^2 f(x_{i+1}) - \Delta^2 f(x_i)$$

$\vdots$

$$n = k \Rightarrow \Delta^k f(x_i) = \Delta^{k-1} f(x_{i+1}) - \Delta^{k-1} f(x_i)$$

---

### Observação

---

**O Operador Diferença Finita para  $n = 1$  é :**

$$\Delta^1 f(x) = \Delta^0 f(x+h) - \Delta^0 f(x) = f(x+h) - f(x)$$

**ou ainda  $\Delta f(x) = f(x+h) - f(x)$ . Ou seja, a aplicação do operador  $\Delta$  de ordem 1 a uma função  $f(x)$  resulta na variação da função sob os pontos  $(x)$  e  $(x+h)$ . Além disso,  $\frac{\Delta f(x)}{h}$  é uma aproximação para a derivada de  $f(x)$ .**

---

Dada uma função  $f(x)$  e conhecidos os valores que a função assume nos pontos distintos  $x_0, x_1, \dots, x_n$ , **todos igualmente espaçados** ( $x_{i+1} - x_i = h$ ), a Tabela 11.1 de diferenças finitas pode ser construída de forma que:

a) a primeira coluna é constituída pelos pontos  $x_i$ ,  $i = 0, 1, \dots, n$ ;

- b) a segunda coluna contém os valores de  $f(x)$  nos pontos  $x_i$ ,  $i = 0, 1, \dots, n$ ;
- c) as colunas  $3, 4, \dots, n+2$  contém as diferenças finitas de ordem  $1, 2, \dots, n$ . Cada uma destas diferenças é simplesmente a diferença entre duas diferenças finitas consecutivas e de ordem imediatamente inferior.

Pontos	(Ordem 0)	(Ordem 1)	(Ordem 2)	...	(Ordem n-1)	(Ordem n)
$x$	$\Delta^0 f(x_i)$	$\Delta f(x_i)$	$\Delta^2 f(x_i)$	...	$\Delta^{n-1} f(x_i)$	$\Delta^n f(x_i)$
$x_0$	$\Delta^0 f(x_0) = f(x_0)$	$\Delta^1 f(x_0)$	$\Delta^2 f(x_0)$	...	$\Delta^{n-1} f(x_0)$	$\Delta^n f(x_0)$
$x_1$	$\Delta^0 f(x_1) = f(x_1)$	$\Delta^1 f(x_1)$	$\Delta^2 f(x_1)$	...	$\Delta^{n-1} f(x_1)$	
$x_2$	$\Delta^0 f(x_2) = f(x_2)$	$\Delta^1 f(x_2)$	$\Delta^2 f(x_2)$	...		
$x_3$	$\Delta^0 f(x_3) = f(x_3)$	$\Delta^1 f(x_3)$	$\Delta^2 f(x_3)$			
$\vdots$	$\vdots$	$\vdots$	$\vdots$			
$x_{n-2}$	$\Delta^0 f(x_{n-2}) = f(x_{n-2})$	$\Delta^1 f(x_{n-2})$	$\Delta^2 f(x_{n-2})$			
$x_{n-1}$	$\Delta^0 f(x_{n-1}) = f(x_{n-1})$	$\Delta^1 f(x_{n-1})$				
$x_n$	$\Delta^0 f(x_n) = f(x_n)$					

Tabela 12.1: Operadores de Diferenças Finitas.

**Exemplo 34.** Construa a tabela de diferenças finitas da função  $f(x) = x^2$  definida sobre os pontos  $x_0 = 1$ ,  $x_1 = 2$ ,  $x_2 = 3$  e  $x_3 = 4$ .

**Solução:** Cálculo das diferenças finitas:

$x$	$\Delta^0 f(x_i)$	$\Delta^1 f(x_i)$	$\Delta^2 f(x_i)$	$\Delta^3 f(x_i)$
1	1	$4 - 1 = 3$	$5 - 3 = 2$	$2 - 2 = 0$
2	4	$9 - 4 = 5$	$7 - 5 = 2$	
3	9	$16 - 9 = 7$		
4	16			

O Teorema 12.1 mostra que é possível relacionar as diferenças divididas com as diferenças finitas.

**Teorema 12.1.** Seja  $f(x)$  uma função contínua e  $(n+1)$  vezes diferenciável no intervalo  $[a, b]$ . Sejam  $x_0, x_1, \dots, x_n$  os  $n+1$  pontos distintos e equidistantes deste intervalo. Então:

$$f[x_0, x_1, \dots, x_n] = \frac{\Delta^n f(x_0)}{h^n n!} \quad (12.4)$$

**Prova:** (Prova por indução finita)

I) Caso base ( $n=1$ ). Por definição:

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f(x_0 + h) - f(x_0)}{h} = \frac{\Delta f(x_0)}{h}.$$

II) Supondo válido para  $n = k-1$ .

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

III) Será provado que é válido para  $(n=k)$ . Usando a definição e hipótese de indução:

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

■

**Corolário 12.1.** *Seja  $f(x)$  uma função contínua e  $n+1$  vezes diferenciável no intervalo  $[a, b]$ . Sejam  $x_0, x_1, \dots, x_n$   $n+1$  pontos distintos e equidistantes deste intervalo. Então, existe  $\xi \in [a, b]$  tal que:*

$$\Delta^{n+1}f(x) = h^{n+1}f^{(n+1)}(\xi)$$

A partir do Teorema 12.1, que relaciona as diferenças divididas com as diferenças finitas, e do Corolário 12.1 são anunciados os **polinômios de diferenças finitas**.

### 12.1.2 Polinômio Interpolador de Diferenças Finitas

Os polinômios de diferenças finitas, ou **fórmula interpolatória de Newton-Gregory** é obtida substituindo a relação do Teorema 12.1 no polinômio interpolador de diferenças divididas (Equação 11.12):

$$\begin{aligned} P_n(x) = & \Delta^0 f(x_0) + (x - x_0) \frac{\Delta^1 f(x_0)}{1!h^1} + (x - x_0)(x - x_1) \frac{\Delta^2 f(x_0)}{2!h^2} \\ & + \dots + (x - x_0)(x - x_1) \dots (x - x_{n-1}) \frac{\Delta^n f(x_0)}{n!h^n} \end{aligned} \quad (12.5)$$

A expressão do erro do polinômio interpolador em termos das diferenças finitas é dada pelo Teorema 12.2.

**Teorema 12.2.** *Seja  $f(x)$  uma função contínua e  $n+1$  vezes diferenciável no intervalo  $[a, b]$ . Sejam  $x_0, x_1, \dots, x_n$   $n+1$  pontos distintos e equidistantes de  $[a, b]$  e*

$$\begin{aligned} P_n(x) = & \Delta^0 f(x_0) + (x - x_0) \frac{\Delta^1 f(x_0)}{1!h^1} + (x - x_0)(x - x_1) \frac{\Delta^2 f(x_0)}{2!h^2} \\ & + \dots + (x - x_0)(x - x_1) \dots (x - x_{n-1}) \frac{\Delta^n f(x_0)}{n!h^n} \end{aligned}$$

o polinômio interpolador de  $f(x)$  nestes pontos. Então, o erro de truncamento é dado por:

$$E(x) = (x - x_0)(x - x_1) \dots (x - x_n) \frac{\Delta^{n+1}f(x)}{h^{n+1}(n+1)!} \quad (12.6)$$

O limitante superior para o erro pode ser calculado por:

$$|E(x)| \leq |(x - x_0)(x - x_1) \dots (x - x_n)| \frac{\max |\Delta^{n+1}f(x)|}{h^{n+1}(n+1)!}, \quad x \in [x_0, x_n]. \quad (12.7)$$

Observação

Dado que a fórmula geral para o erro de truncamento (Equação 10.1) pode ser aplicada em todas as fórmulas interpolatórias (Lagrange, Diferenças Divididas e Diferenças Finitas) e considerando o resultado apresentado no corolário 12.1, a estimativa para o erro (limitante superior para o erro) no método das diferenças finitas, quando se conhece a expressão analítica de  $f(x)$ , é calculada por:

$$|E(x)| \leq |(x - x_0)(x - x_1) \cdots (x - x_n)| \frac{\max |f^{(n+1)}(x)|}{(n+1)!}, \quad x \in [x_0, x_n]. \quad (12.8)$$

**Exemplo 35.** Considere a função  $f(x) = \frac{x}{x+1}$  tabelada nos pontos:

$x_i$	0	1	2
$f(x_i)$	0	1/2	2/3

Determine o polinômio interpolador de diferenças finitas utilizando a Equação 12.5, avalie  $f(1,3)$  e determine um limitante superior para o erro de truncamento da estimativa realizada.

**Solução:** Para os três pontos tabelados o polinômio interpolador de diferenças finitas de grau menor ou igual a dois possui a seguinte forma:

$$P_2(x) = \Delta^0 f(x_0) + (x - x_0) \frac{\Delta^1 f(x_0)}{1!h^1} + (x - x_0)(x - x_1) \frac{\Delta^2 f(x_0)}{2!h^2}.$$

E a tabela de diferenças finitas:

$x$	$\Delta^0 f(x_i)$	$\Delta^1 f(x_i)$	$\Delta^2 f(x_i)$
0	0	1/2	-1/3
1	1/2	1/6	
2	2/3		

Dado  $h = 1$ , o polinômio é:  $P_2(x) = 0 + (x - 0) \frac{(1/2)}{1!(1)^1} + (x - 0)(x - 1) \frac{(-1/3)}{2!(1)^2} = -\frac{1}{6}x^2 + \frac{2}{3}x$ .

A estimativa para  $f(1,3)$  será calculada por  $P_2(1,3)$ , ou seja:

$$f(1,3) \approx P_2(1,3) = -\frac{1}{6}(1,3)^2 + \frac{2}{3}(1,3) = 0,5850.$$

O limitante superior para o erro, será calculado por  $|E(x)| \leq |(x - x_0)(x - x_1) \cdots (x - x_n)| \frac{\max |f^{(n+1)}(x)|}{(n+1)!}$ ,  $x \in [x_0, x_n]$ . Para  $n = 2$ :

$$f^3(x) = \frac{6}{(1+x)^4}, \quad \text{e} \quad \max |f^3(x)| = 6, \quad \text{em} \quad x = 0.$$

Assim,

$$|E(1,3)| \leq |(1,3 - 0)(1,3 - 1)(1,3 - 2)| \frac{6}{3!} = 0,2730.$$

O polinômio interpolador de diferenças finitas também pode ser escrito considerando a seguinte **mu-**  
**dança de variável:**

$$z = \frac{x - x_0}{h} \rightarrow x = x_0 + hz$$

como os pontos são equidistantes, segue que:

$$\begin{aligned} x - x_0 &= hz \\ x - x_1 &= x - (x_0 + h) = x - x_0 - h = hz - h = h(z - 1) \\ x - x_2 &= x - (x_0 + 2h) = x - x_0 - 2h = hz - 2h = h(z - 2) \\ x - x_3 &= x - (x_0 + 3h) = x - x_0 - 3h = hz - 3h = h(z - 3) \\ &\vdots \\ x - x_{n-1} &= h[z - (n - 1)] \end{aligned}$$

Efetuando as substituições na Equação 12.5 obtém-se o polinômio interpolador de diferenças finitas sob a variável  $z$ :

$$\begin{aligned} P_n(z) = & \Delta^0 f(x_0) + z\Delta f(x_0) + \frac{z(z-1)}{2!}\Delta^2 f(x_0) + \frac{z(z-1)(z-2)}{3!}\Delta^3 f(x_0) \\ & + \dots + \frac{z(z-1)\dots[z-(n-1)]}{n!}\Delta^n f(x_0) \end{aligned} \quad (12.9)$$

**Exemplo 36.** Considere a função tabelada nos seguintes pontos:

$x_i$	0,1	0,2	0,3	0,4
$f(x_i)$	1,01	1,05	1,12	1,23

Determine o polinômio interpolador de diferenças finitas utilizando a Equação 12.9 e avalie  $f(0,35)$ .

**Solução:** O ponto  $x = 0,35$ , corresponde na variável  $z$ , a:

$$z = \frac{x - x_0}{h} = \frac{0,35 - 0,1}{0,1} = 2,5.$$

O polinômio interpolador de diferenças finitas de grau menor ou igual a três possui a forma:

$$P_3(x_0 + hz) = \Delta^0 f(x_0) + z\Delta f(x_0) + \frac{z(z-1)}{2!}\Delta^2 f(x_0) + \frac{z(z-1)(z-2)}{3!}\Delta^3 f(x_0).$$

Dada a tabela de diferenças finitas:

$x$	$\Delta^0 f(x_i)$	$\Delta^1 f(x_i)$	$\Delta^2 f(x_i)$	$\Delta^3 f(x_i)$
0,1	1,01	0,04	0,03	0,01
0,2	1,05	0,07	0,04	
0,3	1,12	0,11		
0,4	1,23			

$$\begin{aligned} P_3(x_0 + hz) &= 1,01 + z(0,04) + \frac{z(z-1)}{2!}0,03 + \frac{z(z-1)(z-2)}{3!}0,01 \\ &= 0,0017z^3 + 0,0099z^2 + 0,0284z + 1,0100 \end{aligned}$$

Logo,  $f(0,35) \approx P_3(0,35) = 1,1694$ .



### Exercício Resolvido

Seja  $y = f(x)$  tabelada:

$i$	0	1	2
$x_i$	2,5	3,0	3,5
$y_i$	0,503	0,806	1,192

Determine o polinômio interpolador de Diferenças Finitas que intercepta  $f(x)$  nos pontos fornecidos e calcule o valor aproximado da função no ponto de abscissa  $x = 2,8$  utilizando 3 casas decimais.

**Solução:** Assista o vídeo a seguir:



Vídeo elaborado pela Estagiária em Docência Débora Nasser Diniz.

## 12.2 Complexidade dos métodos de interpolação

É importante, quando se avalia a eficiência de um algoritmo qualquer, saber como ele se comporta em relação ao número de operações aritméticas em função do tamanho da sua entrada. Esta é a análise de complexidade de tempo do algoritmo. Quando se avalia a quantidade de memória necessária em função do tamanho da entrada, tem-se a análise de complexidade de espaço. Existe uma vasta teoria sobre técnicas de avaliação formal destas complexidades. Neste texto considera-se, estritamente, o número de operações aritméticas.

Os métodos de interpolação de Lagrange, Diferenças Divididas e Diferenças Finitas realizam, cada um, um número específico de operações aritméticas conforme apresentado na Tabela 12.2. Considere  $n$  o grau do polinômio interpolador.

Método	Adições	Multiplicações	Divisões
<b>Lagrange</b>	$2n^2 + 3n + 1$	$2n^2 + 3n + 1$	$n + 1$
<b>Diferenças Divididas</b>	$2n^2 + 5n + 1$	$n$	$\frac{n^2}{2} + \frac{n}{2}$
<b>Diferenças Finitas</b>	$n^2 + 5n + 3$	$n$	$n + 1$

Tabela 12.2: Complexidade dos algoritmos de interpolação.



## 12.3 Escolha dos pontos de interpolação

Outro ponto importante na interpolação consiste em como fazer a escolha dos pontos a serem utilizados. Nos exemplos que foram apresentados ao longo do texto, todos os pontos da tabela foram considerados para interpolação.

No entanto, na prática é comum que sejam fornecidos uma quantidade  $m$  de pontos, em que  $m > n + 1$ , para que seja construído um polinômio interpolador de grau no máximo  $n$ . Nestes casos, será necessário escolher  $n + 1$  pontos entre os  $m$  valores disponíveis em uma tabela. Esta escolha é importante e deve ser feita observando:

- o aumento do erro de arredondamento, que pode ser observado quando utilizado polinômios interpoladores de grau muito elevado, o chamado Fenômeno de Runge; e
- o intervalo de interpolação, pois o ponto a ser estimado precisa pertencer ao intervalo de interpolação. Ou seja, deve ser evitado que o valor a ser estimado esteja fora do intervalo de pontos usados para a construção do polinômio.

## 12.4 Considerações Finais

Em relação aos métodos de interpolação:

- Os métodos que utilizam diferenças (divididas ou finitas) são eficientes quando se deseja aumentar (ou diminuir) o grau do polinômio obtido, pois basta, simplesmente, acrescentar (ou retirar) termos. Logo, para cálculos exploratórios, estes métodos, em geral, são preferíveis. No método de Lagrange a alteração do grau do polinômio exige que os cálculos sejam, todos, refeitos.
- O método de Lagrange ocupa menos memória, uma vez que não é necessário o cálculo e o armazenamento de uma tabela de diferenças divididas ou finitas. Além disso, é um pouco mais fácil de ser implementado.
- A desvantagem na utilização do método das diferenças finitas é a exigência de que as abscissas dos pontos utilizados para a interpolação devam ser, necessariamente, equidistantes.
- Nos métodos que utilizam diferenças divididas ou diferenças finitas, a estimativa do erro de truncamento pode ser facilmente integrada ao algoritmo, uma vez que podem ser calculados utilizando uma diferença. No método de Lagrange, a estimativa do erro de truncamento pode ser obtida somente se a função interpolada for conhecida analiticamente.

## 12.5 Exercícios



E. 1. Os dados a seguir representam o crescimento bacteriológico em uma cultura líquida durante alguns dias.

Dia	0	4	8	12	16	20
Quantidade, $\times 10^6$	63,38	74,67	82,74	91,69	101,60	112,58

Encontre uma equação linear de grau 3, pelo método das diferenças finitas, que descreve a tendência dos dados da tabela acima e que permite prever a quantidade de bactérias em 10 e 18 dias.

E. 2. Seja a função de distribuição de probabilidade normal padrão definida por

$$N(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{t^2}{2}\right) dt$$

cujos valores são mostrados na tabela:

$z$	$N(z)$
0,0	0,50000
0,5	0,69146
1,0	0,84134
1,5	0,93319
2,0	0,97725
2,5	0,99379
3,0	0,99865

Calcule  $P_n(0, 3)$  utilizando polinômios interpoladores de diferenças finitas de graus  $n = 1, 2, 3$  e 4.



[illegible]





## **Parte V**

# **Integração Numérica**



# 13

## Integração Numérica

### Sumário da Aula

13.1 Introdução . . . . .	131
13.2 Fórmulas de Newton-Cotes . . . . .	134
13.3 Regra dos Trapézios . . . . .	136
13.3.1 Regra dos trapézios generalizada . . . . .	137
13.3.2 Erro de truncamento na regra dos trapézios . . . . .	138
13.4 Exercícios . . . . .	141

### 13.1 Introdução

No Cálculo Diferencial e Integral estuda-se o conceito de integral definida e como calculá-la por meio de processos analíticos. Os resultados obtidos correspondem a áreas ou volumes de figuras geométricas, dependendo do tipo de integral. Uma função  $y = f(x)$  pode ser integrada sobre o intervalo  $[a, b]$  por

$$I = \int_a^b f(x)dx = F(b) - F(a) \quad (13.1)$$

em que  $F(x)$  é a primitiva de  $f(x)$ , ou seja,  $F'(x) = f(x)$ .

Graficamente, considerando a função  $f(x) \geq 0$ , para todo  $x \in [a, b]$ , pode-se relacionar a integral definida da Equação 13.1 com a área  $A$ , entre a curva  $f(x)$  e o eixo das abscissas, conforme a Figura 13.1.



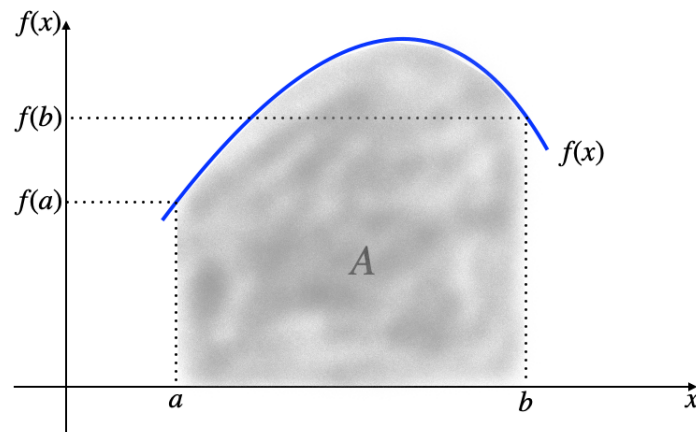


Figura 13.1: Representação gráfica da Equação 13.1.

As aplicações mais óbvias das integrais definidas se encontram no cálculo de comprimentos, áreas, volumes, massa, centro de massa, distância percorrida, tempo decorrido, etc.

Algumas funções, tais como  $f(x) = \sqrt{x^3 + 1}$ ,  $f(x) = \text{sen}(x^2)$ ,  $f(x) = \cos(e^x)$  e  $f(x) = \text{sen}(x)/x$  **não possuem a primitiva**  $F(x)$  que possa ser expressa por uma combinação finita de funções algébricas e/ou transcendentais. Também existem casos, onde são **conhecidos somente valores discretos** de  $f(x)$  ou que a **obtenção da forma analítica de**  $F(x)$  **pode ser complexa e trabalhosa**. Nestes casos, é necessário o uso de métodos numéricos para obter aproximações para o valor da integral definida  $I$ . Os métodos numéricos consistem em aproximar a função  $f(x)$  por um polinômio interpolador e determinar analiticamente a integral desse polinômio sobre o intervalo  $[a, b]$ .

Considere alguns exemplos práticos em que é necessário a utilização de métodos numéricos para o cálculo de integrais definidas:

1. O cálculo do comprimento de uma curva  $f(x)$  em um intervalo  $[a, b]$ . Se a função  $f(x)$  for diferenciável, esse problema remete a uma integral. Por exemplo, no cálculo do perímetro de uma elipse, que exige a avaliação da expressão

$$p = 4b \int_0^{\pi/2} \sqrt{1 - k^2 \text{sen}^2(t)} dt$$

ocorre que a integral  $\int \sqrt{1 - k^2 \text{sen}^2(t)} dt$  é conhecida como integral elíptica do primeiro tipo, e não admite uma primitiva que resulte da combinação finita de funções elementares. Em outras palavras, não há uma fórmula fechada para o perímetro da elipse.

2. Distribuição de probabilidades, calculada normalmente pela função:

$$P_{\tau, \sigma}(t) = \frac{1}{\sigma \sqrt{2\pi}} e^{\left(-\frac{(t - \tau)^2}{2\sigma^2}\right)}$$

Para determinar a probabilidade de que um evento ocorra dentro de um intervalo  $[a, b]$  é necessário calcular a integral

$$\int_a^b P_{\tau, \sigma}(t) dt.$$

Acontece que  $e^{-x^2}$  é uma função cuja primitiva não pode ser expressa como uma combinação finita de funções elementares. Em probabilidade, como é muito frequente o uso dessa integral, adotam-se tabelas com precisão limitada, mas razoável, que servem para a maioria dos propósitos. Essas

tabelas podem ser facilmente montadas com a utilização dos métodos de integração numérica que serão tratados neste texto.

3. Quando há a necessidade de se trabalhar com dados experimentais. Nesta situação, não há funções matemáticas que descrevem um fenômeno físico, mas apenas tabelas de dados que devem ser integrados para se analisar o problema. O tratamento é feito, essencialmente, de forma numérica.

De forma resumida, a aplicação de métodos numéricos na resolução de uma integral definida é vantajosa quando:

- (i) a determinação da primitiva  $F(x)$  é muito difícil;
- (ii) a função a integrar não admite uma primitiva  $F(x)$  que possa ser escrita como uma combinação finita de funções elementares;
- (iii) a função a ser integrada não é conhecida na sua forma analítica, mas, apenas em um conjunto de pontos  $(x_i, y_i), i = 0, 1, \dots, n$ .

Integrar numericamente uma função  $f(x)$  definida em um conjunto de pontos  $(x_i, y_i), i = 0, 1, \dots, n$  distintos sobre um intervalo  $[a, b]$  consiste em aproximar a função  $f(x)$  pelo polinômio  $P_n(x)$  que a interpola e integrar  $P_n(x)$  no lugar de  $f(x)$ . Ou seja:

$$\int_{x_0}^{x_n} f(x)dx \approx \int_{x_0}^{x_n} P_n(x)dx \quad (13.2)$$

A Figura 13.2 representa graficamente a aproximação realizada na Equação 13.2.

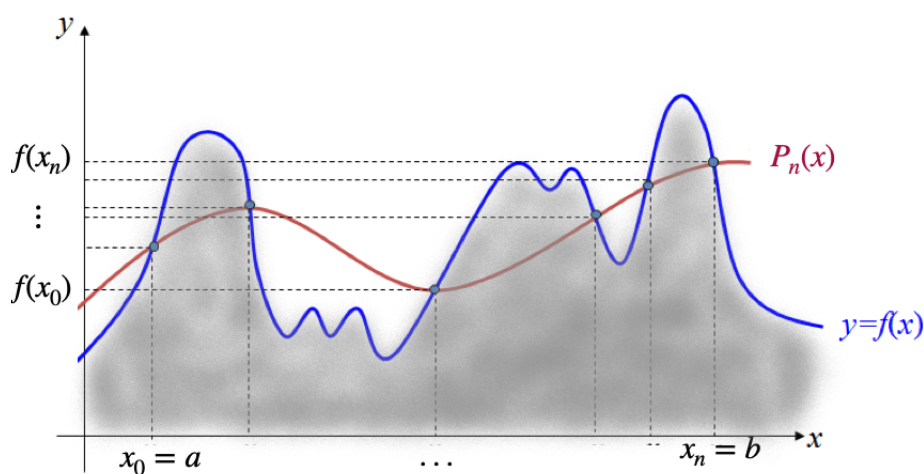


Figura 13.2: Representação gráfica da Equação 13.2.

Por razões históricas, as fórmulas de integração numérica também são denominadas “**quadratura numérica**”, pois foi com o problema da quadratura do círculo que Arquimedes fez os primeiros cálculos usando a noção de integral. As fórmulas de quadratura numérica podem ser do tipo:

**(a) fechado:** quando todos os pontos de  $f(x)$  estão no intervalo de integração  $[a, b]$ , e  $x_0 = a$  e  $x_n = b$  são os extremos.

**(b) aberto:** quando todos os pontos estão no intervalo,  $[a, b]$ , de integração, porém a função integrada  $y = f(x)$  não é avaliada em ambas as extremidades do intervalo, mas em pontos próximos. São utilizadas quando a função integrada apresenta descontinuidades nos extremos do intervalo de integração, ou seja, têm utilidade na análise de integrais impróprias.



## 13.2 Fórmulas de Newton-Cotes

Neste texto serão estudadas as Fórmulas de Newton-Cotes do tipo fechado. Estas fórmulas permitem calcular, por aproximação, uma integral definida substituindo a função a ser integrada por um polinômio de diferenças finitas que a interpola em um conjunto de pontos  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$ .

Considere uma função definida em um conjunto de pontos  $(x_i, y_i)$ ,  $i = 0, 1, \dots, n$  todos distintos e equidistantes no intervalo  $[a, b]$ , ou seja,  $x_{i+1} - x_i = h$ ,  $i = 0, 1, \dots, n$ , sendo  $h$  a distância entre os pontos consecutivos. Seja também  $P_n(x)$  o polinômio interpolador da função  $f(x)$ , tal que  $x_0 = a$  e  $x_n = b$ . Considerando que  $f(x)$  é suficientemente diferenciável, da interpolação polinomial:

$$f(x) = P_n(x) + E_n(x)$$

onde:

$$\begin{aligned} P_n(x) = & \Delta^0 f(x_0) + (x - x_0) \frac{\Delta^1 f(x_0)}{1!h^1} + (x - x_0)(x - x_1) \frac{\Delta^2 f(x_0)}{2!h^2} \\ & + \dots + (x - x_0)(x - x_1) \dots (x - x_{n-1}) \frac{\Delta^n f(x_0)}{n!h^n} \\ E_n(x) = & (x - x_0)(x - x_1) \dots (x - x_n) \frac{f^{(n+1)}(\xi)}{(n+1)!}, \quad x_0 \leq \xi \leq x_n \end{aligned}$$

Assim,

$$\int_{x_0}^{x_n} f(x) dx \approx \int_{x_0}^{x_n} P_n(x) dx \quad (13.3)$$

e o erro cometido na integração numérica calculado por:

$$\int_{x_0}^{x_n} E_n(x) dx$$

onde  $E_n$  é determinado conforme o Teorema 13.1.

**Teorema 13.1.** Se a função  $f(x)$  possui:

- a)  $(n+1)$  derivadas contínuas no intervalo  $[x_0, x_n]$  e os pontos  $x_i = x_0 + ih$ ,  $i = 0, 1, \dots, n$  subdividem o intervalo de integração num número ímpar de intervalos, então existe  $\xi$  tal que:

$$E_n = \frac{h^{n+2} f^{(n+1)}(\xi)}{(n+1)!} \int_0^n z(z-1)(z-2) \dots (z-n) dz, \quad x_0 \leq \xi \leq x_n$$

- b)  $(n+2)$  derivadas contínuas no intervalo  $[x_0, x_n]$  e os pontos  $x_i = x_0 + ih$ ,  $i = 0, 1, \dots, n$  subdividem o intervalo de integração num número par de intervalos, então existe  $\xi$  tal que:

$$E_n = \frac{h^{n+3} f^{(n+2)}(\xi)}{(n+2)!} \int_0^n \left(z - \frac{n}{2}\right) z(z-1)(z-2) \dots (z-n) dz, \quad x_0 \leq \xi \leq x_n$$

Como os pontos escolhidos são todos equidistantes, por simplicidade, será usado o polinômio interpolador de diferenças finitas (fórmula de Newton-Gregory) reescrito para a variável  $z$ :

$$z = \frac{x - x_0}{h} \Rightarrow x = x_0 + zh$$

nesta variável os pontos de interpolação são sempre  $0, 1, \dots, n$ . Sejam:

- o polinômio interpolador:

$$\begin{aligned} P_n(z) = & y_0 + z\Delta y_0 + \frac{z(z-1)}{2!}\Delta^2 y_0 + \frac{z(z-1)(z-2)}{3!}\Delta^3 y_0 \\ & + \dots + \frac{z(z-1)\dots[z-(n-1)]}{n!}\Delta^n y_0 \end{aligned} \quad (13.4)$$

- o erro de interpolação na variável  $z$ :

$$E_n(z) = \frac{z(z-1)\dots(z-n)}{(n+1)!}h^{n+1}f^{(n+1)}(\xi_z), \quad 0 \leq \xi_z \leq n.$$

Assim, o erro na integração será escrito como:

$$E_n = \int_{x_0}^{x_n} E_n(x)dx = h \int_0^n E_n(z)dz = h \int_0^n \frac{z(z-1)\dots(z-n)}{(n+1)!}h^{n+1}f^{(n+1)}(\xi_z)dz.$$

Para avaliar a  $I = \int_a^b f(x)dx$ , será feita a substituição de  $f(x)$  por  $P(z)$  (Equação 13.4). Observe que:

- da relação:

$$x = x_0 + hz \Rightarrow dx = h.dz$$

- nos extremos do intervalo:

$$\text{para } x = x_0 \Rightarrow z = \frac{x_0 - x_0}{h} \Rightarrow z = 0$$

$$\text{para } x = x_n \Rightarrow z = \frac{x_n - x_0}{h} = \frac{nh}{h} \Rightarrow z = n$$

Portanto, o cálculo da integral será efetivamente substituído por  $I = \int_0^n P_n(z)h dz$ . E, visto que  $h$  é uma constante:

$$I = h \int_0^n P_n(z)dz \quad (13.5)$$

A expressão 13.5 constitui-se em uma família de regras de integração ou de fórmulas de quadratura. De acordo com o valor atribuído a  $n$ , determina-se o grau do polinômio interpolador e se obtêm diferentes regras de integração.

### 13.3 Regra dos Trapézios

Seja  $f(x)$  uma função definida em dois pontos consecutivos  $x_0$  e  $x_1$  no intervalo  $[a, b]$ . Nesta regra, a fórmula para integrar  $f(x)$  é obtida por meio da **integração do polinômio interpolador de diferenças finitas de grau um**, ou seja,  $n = 1$ .

Assim, a integral  $I = h \int_0^1 P_n(z) dz$  é substituída:

$$I = h \int_0^1 [y_0 + z \Delta f(x_0)] dz.$$

Integrando, analiticamente, o polinômio e substituindo  $\Delta f(x_0) = f(x_1) - f(x_0) = y_1 - y_0$ :

$$I = h \left[ z f(x_0) + \frac{z^2}{2} \Delta f(x_0) \right]_0^1 = h \left[ f(x_0) + \frac{\Delta f(x_0)}{2} \right] = h \left[ y_0 + \frac{\Delta f(x_0)}{2} \right] = \frac{h}{2} [y_0 + y_1].$$

Logo, a fórmula simples da regra dos trapézios é então definida por:

$$I = \int_{x_0}^{x_n} f(x) dx \approx \frac{h}{2} [y_0 + y_1] \rightarrow \text{Regra dos Trapézios} \quad (13.6)$$

Observe que a  $I$ , conforme a Figura 13.3, corresponde à área do trapézio formada entre o polinômio interpolador  $P_1(x)$  e o eixo das abscissas, o que justifica a denominação **regra dos trapézios**.

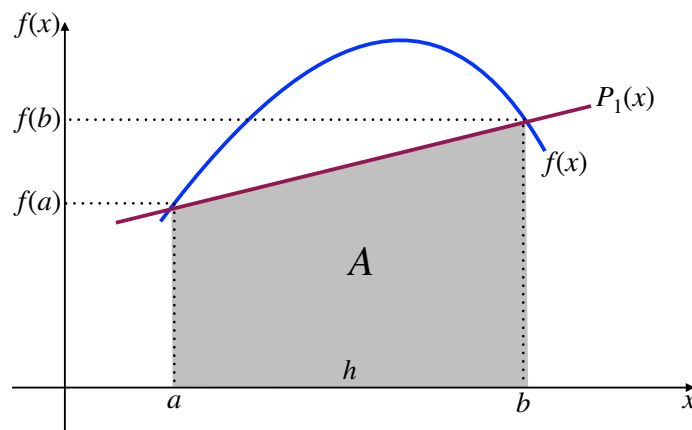


Figura 13.3: Representação gráfica da regra dos Trapézios.

**Exemplo 37.** Calcule, usando a regra dos trapézios, o valor aproximado de

$$\int_1^7 \frac{1}{x} dx.$$

**Solução:** Tabelando a função  $f(x)$  nos pontos  $x_0 = 1$  e  $x_1 = 7$ :

$x_i$	1	7
$f(x_i)$	1	0,1429

E aplicando a regra dos trapézios:

$$\int_1^7 \frac{1}{x} dx \approx \frac{h}{2} [f(x_0) + f(x_1)] = \frac{6}{2} [1 + 0,1429] = 3,4286$$

A Figura 13.4 ilustra a aproximação do cálculo  $I = \int_1^7 \frac{1}{x} dx$  realizada no exemplo 37.

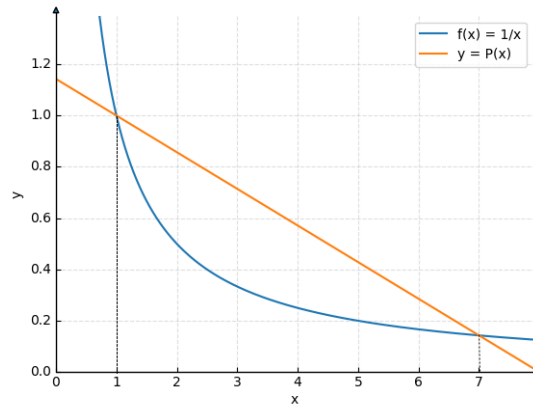
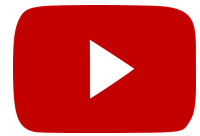


Figura 13.4: Representação gráfica do exemplo 37.

O resultado obtido pela fórmula simples da Regra dos Trapézios não tem, muitas vezes, a precisão desejada. Uma maneira de obter aproximações com erro menor consiste em subdividir o intervalo de integração em  $m$  partes do mesmo tamanho e aplicar a fórmula simples de forma repetida.

### 13.3.1 Regra dos trapézios generalizada



A regra dos trapézios generalizada consiste na subdivisão do intervalo de integração  $[a, b]$  em  $m$  subintervalos iguais, cada qual de tamanho  $h = (x_m - x_0)/m$ ,  $x_0 = a$ ,  $x_m = b$  e aplicação da fórmula simples da regra dos trapézios em cada subintervalo, ou seja, a cada dois pontos consecutivos. A Figura 13.5 ilustra este procedimento.

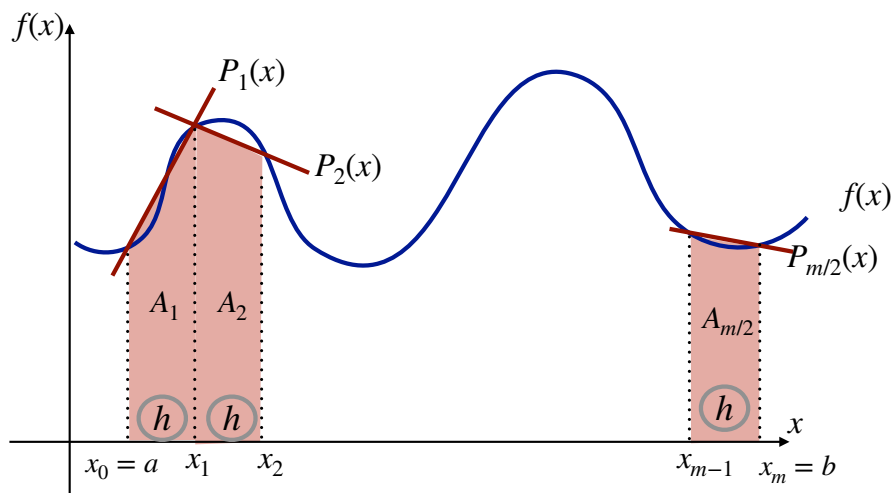


Figura 13.5: Representação gráfica da regra dos trapézios generalizada.

Assim, obtém-se a seguinte aproximação da integral:

$$I = \int_{x_0}^{x_m} f(x)dx \approx \frac{h}{2}[y_0 + y_1] + \frac{h}{2}[y_1 + y_2] + \cdots + \frac{h}{2}[y_{m-1} + y_m]$$

que resulta na **fórmula composta da regra dos trapézios**:

$$I = \int_{x_0}^{x_m} f(x)dx \approx \frac{h}{2}[y_0 + 2y_1 + 2y_2 + \cdots + 2y_{m-1} + y_m] \quad (13.7)$$

### 13.3.2 Erro de truncamento na regra dos trapézios

Como o intervalo de integração na fórmula simples é 1 (que é ímpar), será usada a parte (a) do Teorema 13.1 para escrever a fórmula do erro. Logo,

$$E_n = \frac{h^{n+2} f^{(n+1)}(\xi)}{(n+1)!} \int_0^n z(z-1)(z-2) \cdots (z-n) dz$$

e para  $n = 1$ :

$$E_1 = \frac{h^3 f^{(2)}(\xi)}{(2)!} \int_0^1 z(z-1) dz, \quad x_0 \leq \xi \leq x_1$$

Visto que  $\int_0^1 z(z-1) dz = -\frac{1}{6}$ , o erro de será determinado pela expressão (13.8).

$$E_1 = -\frac{h^3}{12} f''(\xi), \quad \xi \in [x_0, x_1] \quad (13.8)$$

Este erro é um erro de truncamento, pois o grau do polinômio interpolador foi truncado em um em função do número de pontos utilizados.

O erro total da fórmula composta da regra dos trapézios é obtido a partir da soma dos erros cometidos em cada subintervalo, ou seja:

$$E_t = \sum_{i=1}^n -\frac{h^3}{12} f''(\xi_i), \quad \xi_i \in [x_{i-1}, x_i], \quad i = 1, \dots, n$$

Como  $x_{i-1} \leq \xi_i \leq x_i$ ,  $i = 1, \dots, n$  e  $f''(x)$  é uma função contínua por hipótese, então existe  $\xi \in [x_0, x_n]$  tal que:

$$\sum_{i=1}^n f''(\xi_i) = n f''(\xi), \quad x_0 \leq \xi \leq x_n$$

Assim, a expressão para o erro de truncamento na regra dos trapézios generalizada é:

$$E_t = -\frac{h^3}{12} n f''(\xi), \quad \xi \in [x_0, x_n]$$

Dado que  $m$  é o número de subintervalos e sendo  $h = \frac{b-a}{m}$ , temos:

$$E_t = -\frac{(b-a)^3}{12m^2} f''(\xi), \quad \xi \in [x_0, x_m] \quad (13.9)$$

### Limitante superior para o erro

Como o argumento  $\xi$  na fórmula 13.9 não é uma grandeza numérica conhecida no intervalo  $[x_0, x_1]$ , não é possível calcular  $E_1$  de forma exata na regra dos trapézios, mas é possível fazer uma estimativa para esse valor através do **limitante superior para o erro** (Equação 13.10):

$$|E_t| \leq \frac{(b-a)^3}{12m^2} \max \{|f''(x)|, x \in [x_0, x_n]\} \quad (13.10)$$

**Exemplo 38.** Calcule o valor aproximado da integral  $\int_1^7 \frac{1}{x} dx$  usando a fórmula composta da regra dos trapézios para 6 subintervalos e estime o limitante superior para o erro cometido.

**Solução:** A constante  $h$  é calculada por:  $h = \frac{(b-a)}{m} = \frac{7-1}{6} = 1$ . Tabelaando a função  $f(x)$ , para  $h = 1$ :

$x_i$	1	2	3	4	5	6	7
$f(x_i)$	1	0,5	0,3333	0,25	0,2	0,1667	0,1429

e aplicando a regra dos trapézios considerando 6 subintervalos:

$$\begin{aligned} \int_1^7 \frac{1}{x} dx &\approx \frac{h}{2} [y_0 + 2y_1 + 2y_2 + 2y_3 + 2y_4 + 2y_5 + y_6] \\ &= \frac{1}{2} [1 + 2(0,5) + 2(0,3333) + 2(0,25) + 2(0,2) + 2(0,1667) + 0,1429] \\ &= 2,0215. \end{aligned}$$

O limitante superior para o erro é calculado por:

$$|E_t| \leq \frac{(b-a)^3}{12m^2} \max \{|f''(x)|, x \in [1, 7]\}$$

Como a função  $f''(x) = 2/x^3$ , então:  $\max \{|f''(x)|, x \in [1, 7]\} = |f''(1)| = 2$ , pois a função  $|f''(x)| = \frac{2}{x^3}$  é decrescente no intervalo  $[1, 7]$ . Assim:

$$|E_t| \leq \frac{(6)^3}{12(6)^2} (2) = 1$$

Portanto,  $I = \int_1^7 \frac{1}{x} dx \approx 2,0215$  com o limitante superior para o erro igual a 1.

**Exemplo 39.** Seja  $f(x) = \ln(x+2) - 1$ , estime o valor de  $\int_2^{3,2} f(x) dx$ , utilizando a regra dos trapézios generalizada, de modo que o erro de truncamento máximo seja menor ou igual a 0,0004.

**Solução:** O erro de truncamento máximo é calculado pela fórmula

$$E_t \leq \left| \frac{(b-a)^3}{12m^2} \right| \cdot \max |f^{(2)}(x)|$$

A derivada de segunda ordem da função  $f(x)$  é  $f^{(2)}(x) = -\frac{1}{(x+2)^2}$ . No intervalo  $[2; 3,2]$ ,  $f^{(2)}(x)$  assume o maior valor em módulo para  $x = 2$ , logo:

$$\max |f^{(2)}(x)| = |f^{(2)}(2)| = 0,0625.$$

Substituindo na fórmula do limitante superior do erro de truncamento:

$$E_t \leq \left| \frac{(3,2-2)^3}{12m^2} \right| \cdot 0,0625 \leq 0,0004 \Rightarrow m \geq 4,7 \Rightarrow m \geq 5$$



Considerando que o intervalo de integração será dividido em  $m = 5$  partes iguais, obtemos os seguintes dados:

- o valor de  $h$ :  $h = \frac{3,2 - 2}{5} = 0,24$
- a fórmula composta da regra dos trapézios:  $I_1 = \frac{h}{2}[y_0 + 2y_1 + 2y_2 + 2y_3 + 2y_4 + y_5] = \frac{h}{2} \cdot \sum_{i=0}^5 c_i y_i$

Será utilizado um dispositivo prático (tabela) para organizar os cálculos: na primeira coluna da tabela está indicado os índices  $i = 0, 1, \dots, m$ , na segunda coluna as abscissas  $x_i$  (pontos igualmente espaçados), na terceira coluna os valores de  $f(x_i)$  e, por fim, na quarta coluna as constantes  $c_i$  (coeficientes de que cada termo  $y_i$ ). Assim:

$i$	$x_i$	$y_i$	$c_i$
0	2	0,3863	1
1	2,24	0,4446	2
2	2,48	0,4996	2
3	2,72	0,5518	2
4	2,96	0,6014	2
5	3,2	0,6487	1

onde  $\sum_{i=0}^5 c_i y_i = 5,2298$ . Substituindo todos os valores obtidos na fórmula composta da regra dos trapézios:

$$I = \frac{h}{2} \cdot \sum_{i=0}^5 c_i y_i = \frac{0,24}{2} \cdot 5,2298 = 0,6276$$

---

**Observação:**

---

**Utilizando o Cálculo Diferencial e Integral e quatro casas decimais, a resolução analítica da integral do Exemplo 39 é**

$$I = \int_2^{3,2} [\ln(x+2) - 1] dx = \{(x+2) \cdot [\ln(x+2) - 1] - x\}_2^{3,2} = 0,6278.$$

**E o erro total cometido ao utilizar a regra composta dos trapézios e cinco subintervalos para calcular aproximadamente  $I$  é**

$$\text{Erro Total} = |\text{Valor Exato} - \text{Valor Aproximado}| = |0,6278 - 0,6276| = 0,0002.$$

**Note que o erro total é menor ou igual ao erro de truncamento máximo fornecido.**

---

## 13.4 Exercícios



E. 1. Seja a integral

$$\int_1^5 \sqrt{x-1} dx$$

- a) Calcule a integral pela regra dos trapézios simples.
- b) Encontre o polinômio interpolador de diferenças finitas de grau 1 que passa pelos pontos  $x_0 = 1$  e  $x_1 = 5$  da função  $f(x) = \sqrt{x-1}$ .
- c) Calcule analiticamente a integral do polinômio obtido na letra (b) no intervalo  $[1, 5]$ .
- d) Os resultados obtidos nas letras a) e c) são iguais? Justifique a sua resposta.

E. 2. Calcule o valor aproximado da integral:

$$\int_1^4 \sqrt{x} dx$$

usando a fórmula composta da regra dos trapézios para 2, 4 e 6 subintervalos e um limitante superior para o erro. Compare os resultados obtidos.



[illegible]





# 14

## Primeira Regra de Simpson

### Sumário da Aula

14.1 Introdução . . . . .	145
14.1.1 Primeira regra de Simpson generalizada . . . . .	147
14.1.2 Erro de Truncamento na primeira regra de Simpson . . . . .	148
14.2 Exercícios . . . . .	151

### 14.1 Introdução



Seja  $f(x)$  uma função definida em três pontos consecutivos  $x_0$ ,  $x_1$  e  $x_2$  no intervalo  $[a, b]$ . Na primeira regra de Simpson, também conhecida como **Regra do 1/3 de Simpson**, a fórmula para integrar  $f(x)$  é obtida por meio da integração do polinômio interpolador de diferenças finitas de grau dois.

Assim, para  $n = 2$ , o cálculo da integral  $I = h \int_0^n P_n(z) dz$  será substituído por:

$$I = h \int_0^2 \left[ y_0 + z \Delta f(x_0) + \frac{z(z-1)}{2!} \Delta^2 f(x_0) \right] dz$$

que, resolvida analiticamente, resulta em:

$$\begin{aligned}
 I &= h \left[ z f(x_0) + \frac{z^2}{2} \Delta f(x_0) + \left( \frac{z^3}{6} + \frac{z^2}{4} \right) \Delta^2 f(x_0) \right]_0^2 \\
 &= h \left[ 2\Delta^0 f(x_0) + 2\Delta f(x_0) + \frac{1}{3}\Delta^2 f(x_0) \right]
 \end{aligned}
 \tag{14.1}$$

sendo  $\Delta^0 f(x_0) = y_0$ ,  $\Delta f(x_0) = f(x_1) - f(x_0) = y_1 - y_0$  e  $\Delta^2 f(x_0) = y_0 - 2y_1 + y_2$  a Equação 14.1 é escrita de forma simplificada por:

$$I = \int_{x_0}^{x_n} f(x).dx \approx \frac{h}{3} [y_0 + 4y_1 + y_2] \rightarrow \text{Primeira Regra de Simpson} \tag{14.2}$$

Graficamente, este procedimento está representado na Subfigura 14.1 (b). A Figura 14.1 apresenta uma comparação gráfica para a aproximação da integral de uma função  $f(x)$  utilizando a fórmula simples da regra dos trapézios (Subfigura 14.1 (a)) e a fórmula simples da primeira regra de Simpson (Subfigura 14.1 (b)).

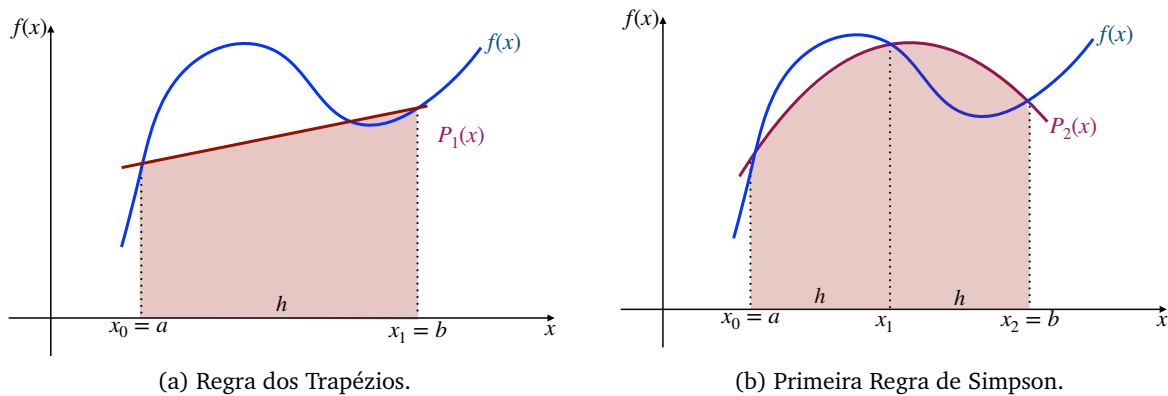


Figura 14.1: Representação gráfica da aplicação da integração numérica.

**Exemplo 40.** Calcule o valor aproximado de  $\int_1^7 \frac{1}{x} dx$  usando a primeira regra de Simpson.

**Solução:** o valor do espaçamento  $h$  é calculado por  $h = \frac{(b-a)}{m} = \frac{7-1}{2} = 3$ . Tabelando a função  $f(x)$  nos pontos  $x_0 = 1$ ,  $x_1 = 4$  e  $x_2 = 7$ :

$x_i$	1	4	7
$f(x_i)$	1	0,25	0,1429

Assim, usando a primeira regra de Simpson a integral será calculada aproximadamente por:

$$\int_1^7 \frac{1}{x} dx \approx \frac{h}{3} [f(x_0) + 4f(x_1) + f(x_2)] = \frac{3}{3} [1 + 4(0,25) + 0,1429] = 2,1429$$

A Figura 14.2 ilustra a aproximação do cálculo da  $I = \int_1^7 \frac{1}{x} dx$  realizada no exemplo 40.

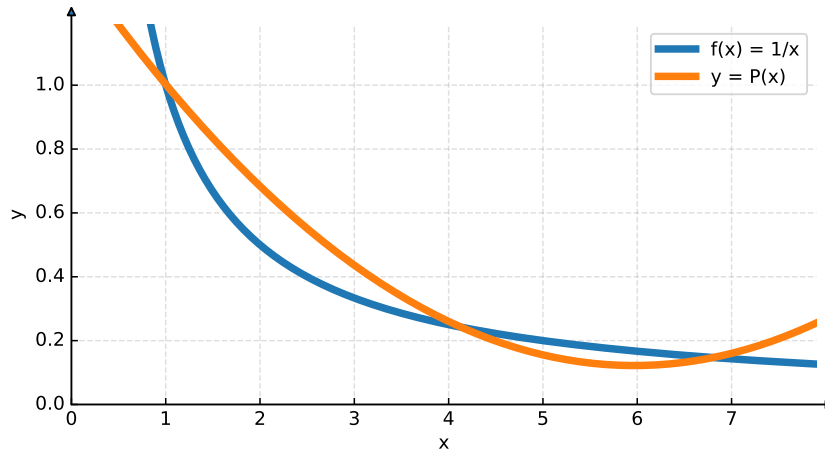


Figura 14.2: Representação gráfica do exemplo 40.

### 14.1.1 Primeira regra de Simpson generalizada

Na primeira regra de Simpson (**Regra do 1/3 de Simpson**) são necessários 3 pontos para fazer a interpolação quadrática, o que significa que o intervalo de integração deve ser dividido em 2 subintervalos. A **primeira regra de Simpson generalizada** consiste na subdivisão do intervalo de integração  $[a, b]$  em  $m$  subintervalos iguais, cada qual de tamanho  $h = (x_m - x_0)/m$ ,  $x_0 = a$ ,  $x_m = b$ , em que  $m$  é um número par e  $x_0 = a$  e  $x_n = b$ . Então, aplica-se a fórmula simples da primeira regra de Simpson em cada dois subintervalos consecutivos. A Figura 14.3 ilustra este procedimento.

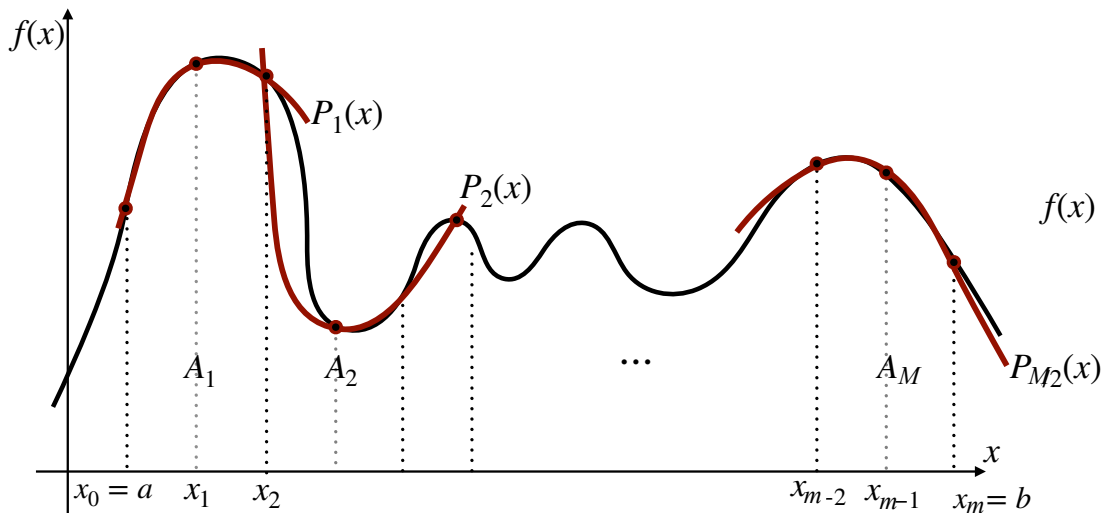


Figura 14.3: Representação gráfica da primeira regra de Simpson generalizada.

Assim, temos a seguinte aproximação da integral:

$$I = \int_{x_0}^{x_m} f(x)dx \approx \frac{h}{3}[y_0 + 4y_1 + y_2] + \frac{h}{3}[y_2 + 4y_3 + y_4] + \cdots + \frac{h}{3}[y_{m-2} + 4y_{m-1} + y_m]$$



que resulta em:

$$I = \int_{x_0}^{x_m} f(x)dx \approx \frac{h}{3}[y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{m-2} + 4y_{m-1} + y_m] \quad (14.3)$$

A Equação 14.3 é chamada de **fórmula composta da primeira regra de Simpson** ou **Regra do 1/3 de Simpson generalizada**.

### 14.1.2 Erro de Truncamento na primeira regra de Simpson

Como o intervalo de integração é 2 (que é par), vamos usar a parte (b) do Teorema 13.1:

$$E_n = \frac{h^{n+3} f^{(n+2)}(\xi)}{(n+2)!} \int_0^n z(z-1)(z-2) \cdots (z-n) dz \quad \xi \in [x_0, x_n]$$

para  $n = 2$ :

$$E_2 = \frac{h^5 f^{(4)}(\xi)}{(4)!} \int_0^2 z(z-1)(z-2) dz, \quad x_0 \leq \xi \leq x_2$$

Como  $\int_0^2 z(z-1)(z-2) dz = -\frac{4}{15}$ , temos que o erro será calculado pela expressão (14.4). Este erro é de truncamento, pois o grau do polinômio interpolador foi truncado em dois em função do número de pontos utilizados.

$$E_2 = -\frac{h^5}{90} f^{(4)}(\xi), \quad \xi \in [x_0, x_2] \quad (14.4)$$

O erro total na fórmula generalizada é obtido a partir da soma dos erros cometidos em cada dois subintervalos subsequentes, ou seja:

$$E_t = \sum_{i=1}^{m/2} -\frac{h^5}{90} f^{(4)}(\xi_i), \quad \xi_i \in [x_{2i-2}, x_{2i}], \quad i = 1, \dots, m/2$$

Como  $x_{2i-2} \leq \xi_i \leq x_{2i}$ ,  $i = 1, \dots, m/2$  e  $f^{(4)}(x)$  é uma função contínua por hipótese, então existe  $\xi \in [x_0, x_m]$  tal que:

$$\sum_{i=1}^{m/2} f^{(4)}(\xi_i) = \frac{m}{2} f^{(4)}(\xi), \quad x_0 \leq \xi \leq x_m$$

Assim, a expressão para o erro de truncamento cometido na primeira regra de Simpson generalizada é calculada por:

$$E_t = -\frac{h^5}{90} \frac{m}{2} f^{(4)}(\xi), \quad \xi \in [x_0, x_m]$$

$m$  é o número de subintervalos e dado que  $h = \frac{b-a}{m}$ , a fórmula anterior será reescrita por:

$$E_t = -\frac{(b-a)^5}{180m^4} f^{(4)}(\xi), \quad \xi \in [x_0, x_m]$$

### Limitante superior para o erro

Veja que o argumento  $\xi$  não é uma grandeza numérica conhecida no intervalo  $[x_0, x_1]$ , então não é possível determinar exatamente o valor do erro, mas é possível estimar **o limitante superior para o erro**:

$$|E_t| \leq \frac{(b-a)^5}{180m^4} \max \left\{ \left| f^{(4)}(x) \right|, \quad x \in [x_0, x_m] \right\} \quad (14.5)$$

**Exemplo 41.** Calcule o valor aproximado da integral  $\int_1^7 \frac{1}{x} dx$  usando a fórmula composta da primeira regra de Simpson para 6 subintervalos e estime o limitante superior para o erro.

**Solução:** o valor da constante  $h = \frac{(x_n - x_0)}{m} = \frac{7-1}{6} = 1$ . Tabelaando a função  $f(x)$ , para  $h = 1$ :

$x_i$	1	2	3	4	5	6	7
$f(x_i)$	1	0,5	0,3333	0,25	0,2	0,1667	0,1429

Usando a fórmula composta da primeira regra de Simpson:

$$\begin{aligned} \int_1^7 \frac{1}{x} dx &\approx \frac{h}{3} [y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + 4y_5 + y_6] \\ &\approx \frac{1}{3} [1 + 4(0,5) + 2(0,3333) + 4(0,25) + 2(0,2) + 4(0,1667) + 0,1429] \\ &\approx 1,9587 \end{aligned}$$

O limitante superior para o erro será determinado pela fórmula:

$$|E_t| \leq \frac{(b-a)^5}{180m^4} \max \left\{ \left| f^{(4)}(x) \right|, \quad x \in [1, 7] \right\}$$

A função  $f^{(4)}(x) = \frac{24}{x^5}$ , então:

$$\max \left\{ \left| f^{(4)}(x) \right|, \quad x \in [1, 7] \right\} = |f^{(4)}(1)| = 24.$$

pois a função  $|f^{(4)}(x)| = 24/x^5$  é decrescente em todo o intervalo  $[1, 7]$ . Assim:

$$|E_t| \leq \frac{(6)^5}{180(6)^4} (24) = 0,8.$$

Logo, Temos  $I = \int_1^7 \frac{1}{x} dx \approx 1,9587$  com o limitante superior para o erro igual a 0,8.

---

#### Observação

---

**Utilizando o Cálculo Diferencial e Integral e quatro casas decimais:**

$$I = \int_1^7 \frac{1}{x} dx = \{\ln(x)\}_1^7 = 1,9459$$

**O erro total cometido**  $E = |1,9459 - 1,9587| = 0,0128 \leq 0,8$ , onde 0,8 é o limitante superior para o erro de truncamento.

---

**Exemplo 42.** Verifique usando a primeira regra de Simpson composta com passo de integração igual a  $h = 0,25$  que  $\pi = 4 \int_0^1 \frac{1}{1+x^2} dx$ .

**Solução:** determinando o número de subintervalos:

$$m = \frac{(x_n - x_0)}{h} = \frac{1 - 0}{0,25} = 4$$

É importante que  $m$  seja par. Tabelando a função  $f(x)$ , com  $h = 0,25$ , temos:

$i$	$x_i$	$y_i$	$c_i$
0	0,00	1,0000	1
1	0,25	0,9412	4
2	0,50	0,8000	2
3	0,75	0,6400	4
4	1,00	0,5000	1

Assim, usando a primeira regra de Simpson composta, temos:

$$I_2 = \frac{0,25}{3} [1,000 + 4(0,9412) + 2(0,8000) + 4(0,6400) + 0,5000] = 0,7854$$

Logo,

$$\pi \approx 4.I_2 = 3,1416$$

**Exemplo 43.** O PROCON tem recebido reclamações com relação ao peso dos pacotes de açúcar de 5 kg. Com a finalidade de verificar a validade das reclamações, foi coletada uma amostra de 100 pacotes. Com isto, chegou-se a conclusão de que para determinar a probabilidade de um pacote de açúcar pesar menos do que 5 kg deve ser avaliada a expressão a seguir:

$$F = 0,5 + \frac{1}{\sqrt{2\pi}} \int_0^{1,8} e^{-\frac{x^2}{2}} dx$$

Estime essa probabilidade e o erro de truncamento máximo cometido utilizando a primeira regra de Simpson generalizada. Divida o intervalo de integração em 6 partes e faça os cálculos retendo 4 casas decimais.

**Solução:** o primeiro passo será calcular o valor da integral  $\int_0^{1,8} e^{-\frac{x^2}{2}} dx$ . Para tanto, considere

$$h = \frac{b - a}{m} = \frac{1,8 - 0}{6} = 0,3.$$

Tabelando a função  $f(x)$  para  $h = 0,3$ :

$i$	$x_i$	$y_i$	$c_i$
0	0,0	1	1
1	0,3	0,9560	4
2	0,6	0,8353	2
3	0,9	0,6670	4
4	1,2	0,4868	2
5	1,5	0,3247	4
6	1,8	0,1979	1

Logo:

$$I_2 = \frac{h}{3} \sum_{i=0}^6 c_i \cdot y_i \Rightarrow I_2 = \frac{0,3}{3} 11,6325 \Rightarrow I = 1,1633$$

Assim, a probabilidade do pacote de açúcar pesar menos que 5 kg é

$$F = 0,5 + \frac{1}{\sqrt{2\pi}} \cdot 1,1633 \Rightarrow F = 0,9640.$$

Para o cálculo do limitante superior do erro de truncamento, considere a expressão:

$$|E_2| \leq \left| \frac{(b-a)^5}{180m^4} \right| \cdot \max |f^{(IV)}(x)|$$

$f(x) = e^{-\frac{x^2}{2}}$  possui a seguinte derivada de quarta ordem:  $f^{(4)}(x) = e^{-\frac{x^2}{2}} \cdot (x^4 - 6x^2 + 3)$ . Assim,

$$\max |f^{(4)}(x)| = |f^{(4)}(0)| = 3.$$

Logo

$$E_2 \leq \left| \frac{(1,8-0)^5}{180(6)^4} \right| 3 \Rightarrow E_2 \leq 0,000243.$$

## 14.2 Exercícios



**E. 1.** Calcule o valor aproximado da integral:

$$\int_1^4 \sqrt{x} dx$$

usando a fórmula composta da primeira regra de Simpson para 2, 4 e 6 subintervalos e um limitante superior para o erro. Compare os resultados obtidos.



[illegible]





# 15

## Segunda Regra de Simpson

### Sumário da Aula

15.1 Introdução . . . . .	155
15.1.1 Segunda regra de Simpson generalizada . . . . .	157
15.1.2 Erro de Truncamento na segunda regra de Simpson . . . . .	158
15.2 Considerações Finais . . . . .	161
15.3 Exercícios . . . . .	162

### 15.1 Introdução



Seja  $f(x)$  uma função definida em três pontos consecutivos  $x_0, x_1$  e  $x_2$  no intervalo  $[a, b]$ . Na segunda regra de Simpson, também conhecida como **Regra do 3/8 de Simpson**, a fórmula para integrar  $f(x)$  é obtida por meio da integração do polinômio interpolador de diferenças finitas de grau três.



Assim, para  $n = 3$ , o cálculo da integral  $I = h \int_0^n P_n(z) dz$  será substituído por:

$$I = h \int_0^3 \left[ y_0 + z \Delta f(x_0) + \frac{z(z-1)}{2!} \Delta^2 f(x_0) + \frac{z(z-1)(z-2)}{3!} \Delta^3 f(x_0) \right] dz$$

que, resolvida analiticamente, resulta em:

$$I = h \left[ z f(x_0) + \frac{z^2}{2} \Delta f(x_0) + \left( \frac{z^3}{6} + \frac{z^2}{4} \right) \Delta^2 f(x_0) + \left( \frac{z^4}{24} - \frac{z^3}{6} + \frac{z^2}{6} \right) \Delta^3 f(x_0) \right]_0^3. \quad (15.1)$$

Uma vez que  $\Delta^0 f(x_0) = y_0$ ,  $\Delta f(x_0) = y_1 - y_0$ ,  $\Delta^2 f(x_0) = y_0 - 2y_1 + y_2$  e  $\Delta^3 f(x_0) = y_3 - 3y_2 + 3y_1 - y_0$ , e substituindo essas diferenças na Equação 15.1 temos:

$$I = \int_{x_0}^{x_n} f(x).dx \approx \frac{3h}{8} [y_0 + 3y_1 + 3y_2 + y_3] \rightarrow \text{Segunda Regra de Simpson} \quad (15.2)$$

Este procedimento é ilustrado na Figura 15.1.

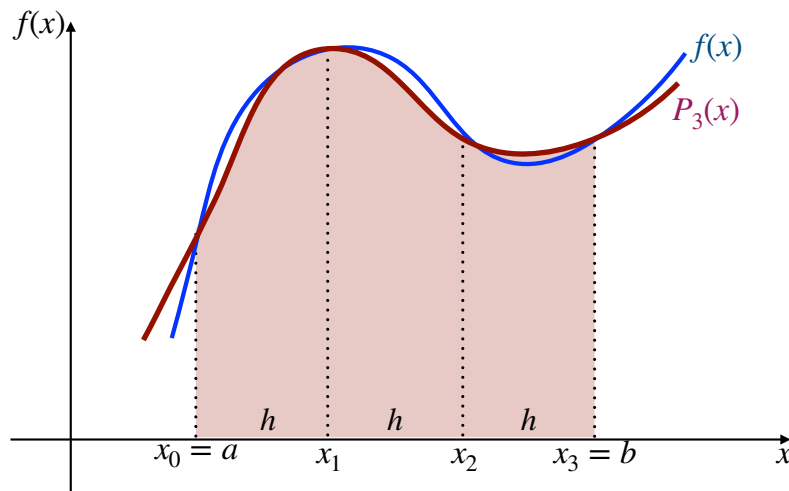


Figura 15.1: Representação gráfica da segunda regra de Simpson.

**Exemplo 44.** Calcule o valor aproximado de  $\int_1^7 \frac{1}{x} dx$  usando a segunda regra de Simpson.

**Solução:** Neste caso, o valor do espaçamento  $h$  é  $h = \frac{(b-a)}{n} = \frac{7-1}{3} = 2$ . Tabelando a função  $f(x)$  nos pontos  $x_0 = 1$ ,  $x_1 = 3$ ,  $x_2 = 5$  e  $x_3 = 7$ :

$x_i$	1	3	5	7
$f(x_i)$	1	0,3333	0,2	0,1429

Assim, usando a segunda regra de Simpson:

$$\int_1^7 \frac{1}{x} dx \approx \frac{3h}{8} [y_0 + 3y_1 + 3y_2 + y_3] = \frac{3(2)}{8} [(1) + 3(0,3333) + 3(0,2) + (0,1429)] = 2,0571.$$

A Figura 15.2 ilustra a aproximação do cálculo da  $I = \int_1^7 \frac{1}{x} dx$  realizada no Exemplo 44.

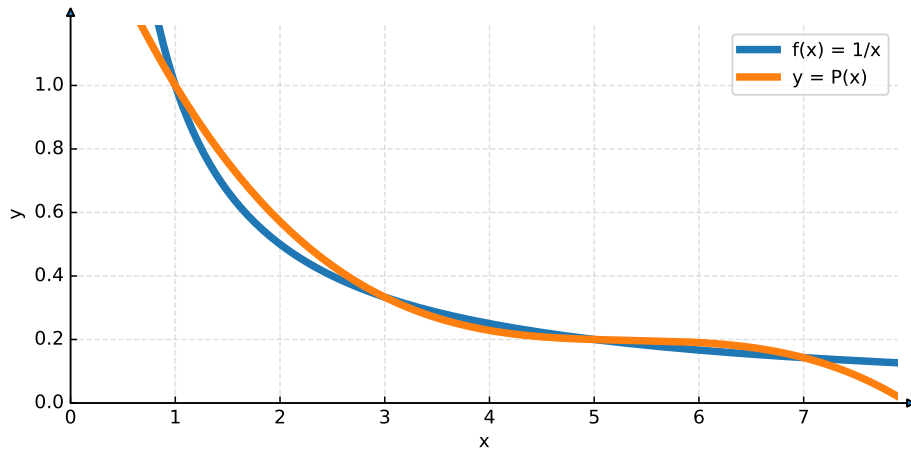


Figura 15.2: Representação gráfica do exemplo 44.

### 15.1.1 Segunda regra de Simpson generalizada

Na segunda regra de Simpson (**Regra dos 3/8 de Simpson**) são necessários 4 pontos para fazer a interpolação, o que significa que o intervalo de integração deverá ser dividido em 3 subintervalos.

Assim, a **segunda regra de Simpson generalizada** consiste na subdivisão do intervalo de integração  $[a, b]$  em  $m$  subintervalos iguais, cada qual de tamanho  $h = (x_m - x_0)/m$ ,  $x_0 = a$ ,  $x_m = b$ , onde  $m$  deve ser um número **múltiplo de três**. Então, aplica-se a fórmula simples da segunda regra de Simpson a cada três subintervalos ou a cada quatro pontos consecutivos. A Figura 15.3 ilustra este procedimento.

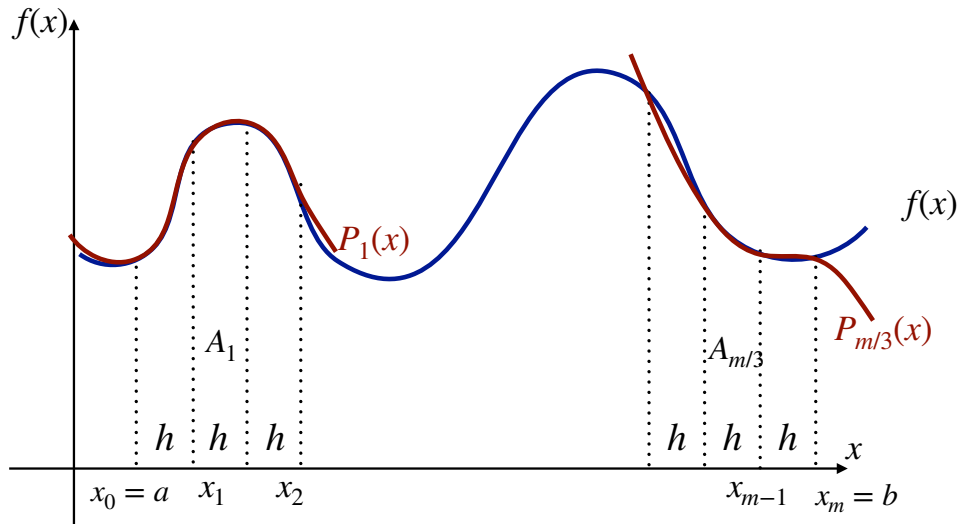


Figura 15.3: Representação gráfica da segunda regra de Simpson generalizada.

Assim, a aproximação da integral será:

$$I = \int_{x_0}^{x_n} f(x) dx = \frac{3h}{8} [y_0 + 3y_1 + 3y_2 + y_3] + \frac{3h}{8} [y_3 + 3y_4 + 3y_5 + y_6] + \cdots + \frac{3h}{8} [y_{n-3} + 3y_{n-2} + 3y_{n-1} + y_n]$$

que resulta em:

$$I = \int_{x_0}^{x_n} f(x)dx \approx \frac{3h}{8}[y_0 + 3y_1 + 3y_2 + 2y_3 + 3y_4 + 3y_5 + 2y_6 \cdots + 2y_{n-3} + 3y_{n-2} + 3y_{n-1} + y_n]. \quad (15.3)$$

A Equação 15.3 é chamada de **fórmula composta da segunda regra de Simpson** ou **Regra dos 3/8 de Simpson generalizada**.

### 15.1.2 Erro de Truncamento na segunda regra de Simpson

Como o intervalo de integração é igual a 3 (que é ímpar), será usada a parte (a) do Teorema 13.1:

$$E_n = \frac{h^{n+2}f^{(n+1)}(\xi)}{(n+1)!} \int_0^n z(z-1)(z-2)\cdots(z-n)dz \quad \xi \in [x_0, x_n]$$

para  $n = 3$ :

$$E_3 = \frac{h^5 f^{(4)}(\xi)}{(4)!} \int_0^3 z(z-1)(z-2)(z-3)dz, \quad x_0 \leq \xi \leq x_3$$

Resolvendo a integral da expressão anterior, temos que  $\int_0^3 z(z-1)(z-2)(z-3)dz = -\frac{9}{10}$ . Assim, o erro de truncamento é então calculado pela expressão (15.4).

$$E_3 = -\frac{3h^5}{80}f^{(4)}(\xi), \quad \xi \in [x_0, x_3] \quad (15.4)$$

O erro total na fórmula generalizada é obtido a partir da soma dos erros cometidos em cada três subintervalos subsequentes, ou seja:

$$E_t = \sum_{i=1}^{m/3} -\frac{3h^5}{80}f^{(4)}(\xi_i), \quad \xi_i \in [x_{3i-3}, x_{3i}], \quad i = 1, \dots, m/3$$

Como  $x_{3i-3} \leq \xi_i \leq x_{3i}$ ,  $i = 1, \dots, m/3$  e  $f^{(4)}(x)$  é uma função contínua por hipótese, então existe  $\xi \in [x_0, x_m]$  tal que:

$$\sum_{i=1}^{m/3} f^{(4)}(\xi_i) = \frac{m}{3}f^{(4)}(\xi), \quad x_0 \leq \xi \leq x_m$$

Assim, a expressão para o erro de truncamento cometido na primeira regra de Simpson generalizada é calculada por:

$$E_t = -\frac{3h^5}{80} \frac{m}{3} f^{(4)}(\xi), \quad \xi \in [x_0, x_m]$$

dados que  $m$  é o número de subintervalos e  $h = \frac{b-a}{m}$ :

$$E_t = -\frac{(b-a)^5}{80m^4}f^{(4)}(\xi), \quad \xi \in [x_0, x_m]$$

### Limitante superior para o erro

Observe que o argumento  $\xi$  não é uma grandeza numérica conhecida no intervalo  $[x_0, x_1]$ , então não é possível determinar exatamente o valor do erro, mas é possível estimar o limitante superior para ele.

$$|E_t| \leq \frac{(b-a)^5}{80m^4} \max \left\{ \left| f^{(4)}(x) \right|, \quad x \in [x_0, x_m] \right\} \quad (15.5)$$

**Exemplo 45.** Calcule o valor aproximado da integral  $\int_1^7 \frac{1}{x} dx$  usando a fórmula composta da segunda regra de Simpson para 6 subintervalos e estime um limitante superior para o erro.

**Solução:** seja o valor da constante  $h = \frac{(x_n - x_0)}{n} = \frac{7-1}{6} = 1$ . Tabelando a função  $f(x)$ , considerando  $h = 1$ :

$x_i$	1	2	3	4	5	6	7
$f(x_i)$	1	0,5	0,3333	0,25	0,2	0,1667	0,1429

Utilizando a segunda regra de Simpson:

$$\begin{aligned} \int_1^7 \frac{1}{x} dx &\approx \frac{3h}{8} [y_0 + 3y_1 + 3y_2 + 2y_3 + 3y_4 + 3y_5 + y_6] \\ &= \frac{3}{8} [1 + 3(0,5) + 3(0,3333) + 2(0,25) + 3(0,2) + 3(0,1667) + 0,1429] \\ &= 1,9665 \end{aligned}$$

O limitante superior para o erro será calculado pela expressão

$$|E_t| \leq \frac{(b-a)^5}{80m^4} \max \left\{ \left| f^{(4)}(x) \right|, x \in [1, 7] \right\}$$

A função  $f^{(4)}(x) = 24/x^5$ , então:

$$\max \left\{ \left| f^{(4)}(x) \right|, x \in [1, 7] \right\} = |f^{(4)}(1)| = 24$$

dado que  $|f^{(4)}(x)| = 24/x^5$  é decrescente no intervalo  $[1, 7]$ . Então:

$$|E_t| \leq \frac{(6)^5}{80(6)^4} (24) = 1,8.$$

Logo,  $I = \int_1^7 \frac{1}{x} dx \approx 1,9665$  com o limitante superior para o erro igual a 1,8.

#### Observação

Os resultados da tabela a seguir correspondem ao cálculo da  $\int_1^7 \frac{1}{x} dx$  utilizando cada uma das regras estudadas, 6 subintervalos e a estimativa do limitante superior para cada caso. Na tabela, também é apresentado o erro total considerando o valor exato da integral  $I = 1,9459$ .

Regra	Valor Aproximado	Limitante Superior	Erro Total
Regra dos Trapézios	2,0215	1	0,0756
1 Regra de Simpson	1,9459	0,8	0,0128
2 Regra de Simpson	1,9665	1,8	0,0206

**Exemplo 46.** Calcule o valor aproximado da integral  $\int_1^7 \ln(x+9)dx$  usando a fórmula composta da segunda regra de Simpson para 6 subintervalos e estime o limitante superior para o erro.

**Solução:** a constante  $h = \frac{(x_n - x_0)}{n} = \frac{7-1}{6} = 1$ . Tabelando a função  $f(x)$ , considerando  $h = 1$ :

$x_i$	$y_i$	$c_i$	$c_i y_i$
1	2,3026	1	2,3026
2	2,3979	3	7,1937
3	2,4849	3	7,4547
4	2,5649	2	5,1298
5	2,6391	3	7,9173
6	2,7081	3	8,1243
7	2,7726	1	2,7726
soma:			40,8950

Assim, usando a segunda regra de Simpson, temos:

$$\begin{aligned} \int_1^7 \ln(x+9)dx &\approx \frac{3h}{8}[y_0 + 3y_1 + 3y_2 + 2y_3 + 2y_4 + 3y_5 + y_6] \\ &= \frac{3}{8} \sum_{i=0}^6 c_i y_i = \frac{3}{8} 40,8950 = 15,3356 \end{aligned}$$

O limitante superior para o erro será calculado pela expressão:

$$|E_t| \leq \frac{(b-a)^5}{80n^4} \max \left\{ \left| f^{(4)}(x) \right|, \quad x \in [1, 7] \right\},$$

onde

$$\max \left\{ \left| f^{(4)}(x) \right|, \quad x \in [1, 7] \right\} = |f^{(4)}(1)| = 0,0006.$$

Então:

$$|E_t| \leq \frac{(6)^5}{80(6)^5} (0,0006) = 0,000045.$$

Portanto,  $I = \int_1^7 \ln(x+9)dx \approx 15,3356$  com o limitante superior para o erro igual a 0,000045.

**Exemplo 47.** Um tanque esférico de raio  $R = 5m$  está cheio com água. A água será drenada através de um orifício de raio  $r = 0,1m$  situado no fundo do tanque. A variação do nível,  $h$ , da água com o tempo,  $t$ , em segundos, é dada pela relação:

$$dt = \frac{R^2 - h^2}{r^2 \sqrt{2g(h+R)}} dh \quad (15.6)$$

onde  $g = 9,81m/s^2$  é a aceleração devida à gravidade. Utilize a Segunda Regra de Simpson, para estimar o tempo para que o nível da água chegue a  $1m$  do fundo. Divida o intervalo de integração em nove partes e faça os cálculos com duas casas decimais.

**Solução:** Fazendo as substituições do enunciado na equação 15.6:

$$dt = \frac{25 - h^2}{0,01 \sqrt{19,6(h+5)}} dh$$

como o raio do tanque é  $5m$ , inicialmente o nível da água, em relação ao fundo sera  $10m$ . Portanto, a integral

a ser calculada é:

$$I = \int_{10}^1 \frac{25 - h^2}{0,01 \sqrt{19,6(h+5)}} dh$$

O intervalo de integração deve ser dividido em 9 partes, assim a fórmula da segunda regra de Simpson para calcular a aproximação da integral é:

$$I = \frac{3h}{8} [y_0 + 3y_1 + 3y_2 + 2y_3 + 3y_4 + 3y_5 + 2y_6 + 3y_7 + 3y_8 + y_9]$$

A constante  $h = \frac{(b-a)}{9} = \frac{1-10}{9} = -1$ . Tabelando a função  $f(x)$ , considerando  $h = -1$ :

$i$	$x_i$	$y_i$	$c_i$
0	10	-437,19	1
1	9	-337,89	3
2	8	-244,20	3
3	7	-156,41	2
4	6	-74,88	3
5	5	0,0	3
6	4	67,73	2
7	3	127,71	3
8	2	179,19	3
9	1	221,20	1

onde  $\sum_{i=0}^9 c_i y_i = -1.443,56$ . Assim:

$$I = \frac{3h}{8} \sum_{i=0}^9 c_i y_i \Rightarrow \frac{3(-1)}{8} (-1.443,56) \Rightarrow I = 541,34.$$

## 15.2 Considerações Finais

Algumas considerações em relação as fórmulas de Newton-Cotes.

- vamos chamar de ordem de convergência a velocidade com a qual uma sucessão converge para o seu limite.
- Uma regra de integração tem grau de exatidão  $g$  se integrar, exatamente, todos os polinômios de grau menor ou igual a  $g$  e se existir pelo menos um polinômio de grau  $g+1$  que não é integrado exatamente por esta regra.
- Comparando-se as expressões dos erros, verifica-se que as regras de Simpson têm ordem de convergência  $h^4$ , enquanto que a regra dos trapézios possui ordem de convergência  $h^2$ . Assim, as regras de Simpson produzem resultados que convergem para o valor real da integral com a mesma velocidade, e mais rapidamente do que a regra dos trapézios, quando  $h \rightarrow 0$ .
- Embora a primeira regra de Simpson tenha sido obtida por meio da integração do polinômio interpolador de grau dois, ela é exata, também, para polinômios de grau três, visto que, na fórmula do erro, aparece a derivada quarta da função. Pode ser demonstrado que, quando o grau,  $n$ , do polinômio é par, então as fórmulas de Newton-Cotes do tipo fechado têm grau de exatidão  $(n+1)$ .

- Para obter o resultado de uma integral com uma determinada precisão, pode-se impor que o erro, em módulo, seja menor que  $0,5 \times 10^{-k}$ , onde  $k$  é o número de casas decimais corretas que se deseja e, assim, determinar em quantas partes deverá se dividido o intervalo de integração. Outra alternativa é aumentar, sucessivamente, o número de pontos e comparar dois resultados consecutivos até que seja obtida a precisão desejada. Este segundo procedimento é o mais comumente utilizado.

## 15.3 Exercícios



E. 1. Calcule o valor aproximado da integral:

$$\int_1^4 \sqrt{x} dx$$

usando a fórmula composta da segunda regra Simpson para 3, 6 e 9 subintervalos e um limitante superior para o erro. Compare os resultados obtidos.

E. 2. Calcule

$$\int_2^5 \frac{1}{x \ln(x)} dx$$

com  $m = 12$  pelas regras abaixo:

- trapézio.
- primeira regra de Simpson.
- segunda regra de Simpson.
- Comparar os três resultados obtidos com o valor exato  $\int_2^5 \frac{1}{x \ln(x)} dx = \ln(\ln(5)) - \ln(\ln(2)) \approx 0,84240$ .





A series of horizontal lines for writing, starting from the first line below the icon and continuing down to the bottom of the page.





## **Parte VI**

# **Raízes de Equações Algébricas e Transcendentes**



# 16

## Raízes de Equações

### Sumário da Aula

16.1 Introdução . . . . .	167
16.2 Determinação das raízes . . . . .	171
16.2.1 Isolamento das raízes . . . . .	171
16.3 Estudo Especial de Equações Polinomiais . . . . .	176
16.3.1 Delimitação das raízes reais . . . . .	176
16.3.2 Enumeração das raízes reais . . . . .	179
16.4 Exercícios . . . . .	180

### 16.1 Introdução

A necessidade de determinar valores  $x = \xi$  que satisfaçam a uma equação da forma  $f(x) = 0$  ocorre com bastante frequência em uma grande variedade de problemas nas diversas áreas científicas. Estes valores são chamados de raízes da equação  $f(x) = 0$  ou zeros da função  $y = f(x)$ , onde  $f(x)$  pode ser um polinômio em  $x$  ou uma função transcendente.



Em raros casos é possível obter as raízes exatas de  $f(x) = 0$ , como ocorre, por exemplo, supondo-se que  $f(x)$  seja um polinômio quadrático, biquadrático, cúbico ou fatorável em termos dos polinômios com as características anteriores. Contudo, para polinômios irreduzíveis de grau superior a dois ou funções transcendentais, para as quais não existem um método geral para obter os seus zeros, faz-se necessária a utilização de métodos numéricos.

Por meio de métodos numéricos, é possível obter uma solução aproximada, que em alguns casos é tão próxima da solução exata quanto se deseja. A maioria dos procedimentos numéricos fornecem uma sequência de aproximações, cada uma das quais mais precisa que a anterior, de tal modo que a repetição do procedimento fornece uma aproximação que se difere da solução exata por alguma tolerância pré-fixada.

Serão apresentados métodos iterativos para a determinação de aproximações para raízes isoladas de  $f(x) = 0$  e um estudo especial às equações polinomiais em virtude da importância que elas possuem em algumas aplicações práticas. Para tanto, considere as definições seguintes.

**Definição 16.1.** Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função qualquer. Dizemos que  $\xi \in [a, b]$  é um zero (ou raiz) de  $f(x) = 0$  se  $f(\xi) = 0$ .

**Definição 16.2. (raiz aproximada)** Sendo  $\varepsilon$  uma precisão, diz-se que um ponto  $x_k$  é uma aproximação para a raiz  $\xi$ , de uma equação  $f(x) = 0$ , se satisfizer uma das condições:

- (i)  $|f(x_k)| < \varepsilon$
- (ii)  $|x_k - \xi| < \varepsilon$

Conforme mostrado na Figura 16.1 as duas condições apresentadas na Definição 16.2 não são equivalentes. A Figura 16.1 (a) apresenta uma situação em que a condição (i) é satisfeita e a condição (ii) não. Na Figura 16.1 (b) é mostrado o contrário, ou seja, uma situação em que a condição (ii) é satisfeita e a condição (i) não.

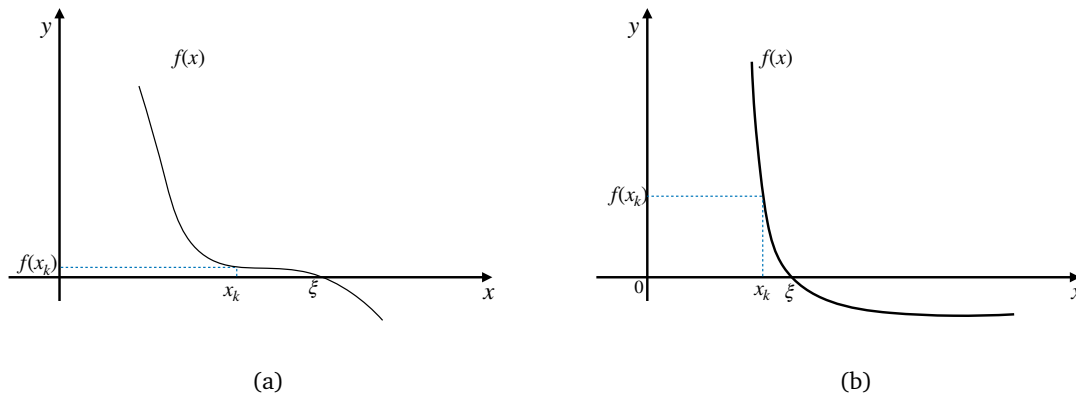


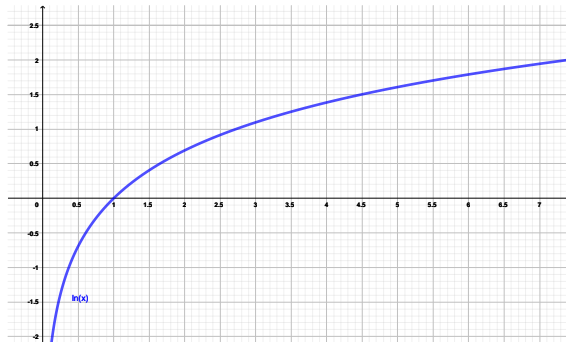
Figura 16.1: Representação gráfica da Definição 16.2.

**Teorema 16.1.** Dada a equação  $f(x) = 0$ , se  $f(x)$  é contínua e assume valores de sinais opostos nos extremos do intervalo  $[a, b]$ , ou seja, se  $f(a) \times f(b) < 0$ , então existe pelo menos um ponto  $\xi \in [a, b]$ , tal que  $f(\xi) = 0$ .

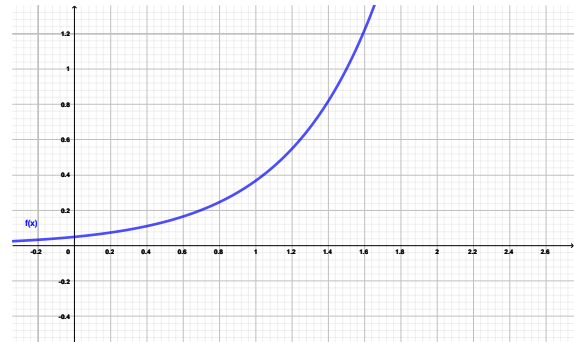
**Prova:** Veja [8]. ■

**Exemplo 48.** Seja  $f : [0, \infty] \rightarrow \mathbb{R}$ , onde  $f(x) = \ln(x)$ . O gráfico da função  $f(x) = \ln(x)$  é dado pela Figura 16.2 (a). Note que  $f(0,5) \times f(1,5) < 0$ . Logo, existe uma raiz de  $f(x)$  no intervalo  $[0,5 \ 1,5]$ . Além disso, como a função é crescente, temos que a curva intercepta o eixo das abscissas num único ponto. Então,  $\xi = 1$  é a única raiz de  $f(x) = 0$ .

**Exemplo 49.** Seja  $f : [0, \infty] \rightarrow \mathbb{R}$ , onde  $f(x) = e^{2x-3}$ . O gráfico da função  $f(x) = e^{2x-3}$  é dado pela Figura 16.2 (b). Note que, neste caso, a curva não intercepta o eixo das abscissas (eixo  $x$ ). Logo, não existe  $\xi$  tal que  $f(\xi) = 0$ .



(a) Representação gráfica da raiz da função  $f(x) = \ln(x)$  do Exemplo 48.



(b) Representação gráfica da função  $f(x) = e^{2x-3}$  do Exemplo 49.

Figura 16.2: Representação gráfica das funções apresentadas nos exemplos 48 e 49 .

**Definição 16.3. (raiz múltipla)** Um ponto  $\xi \in [a, b]$  é uma raiz de multiplicidade  $m$  da equação  $f(x) = 0$  se  $f(x) = (x - \xi)^m g(x)$ , com  $g(x) \neq 0$  em  $[a, b]$ .

**Exemplo 50.** Seja  $f : \mathbb{R} \rightarrow \mathbb{R}$ , onde  $f(x) = x^3 - 3x - 2$ .

O gráfico da função  $f(x)$  é dado pela Figura 16.3. Note que a curva intercepta o eixo das abscissas nos pontos  $\xi_1 = -1$  e  $\xi_2 = 2$ . Observe que  $\xi_1 = -1$  é uma raiz de multiplicidade 2, pois  $f(x) = x^3 - 3x - 2 = (x + 1)^2(x - 2)$ .

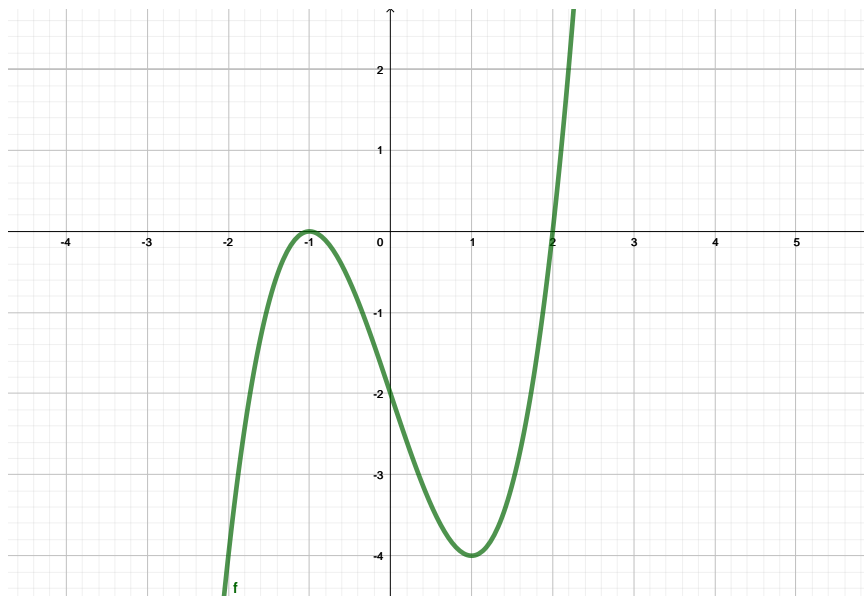


Figura 16.3: Representação gráfica da função  $f(x) = x^3 - 3x - 2$ .

A raiz  $\xi$  de uma equação  $f(x) = 0$  possui multiplicidade  $m$  se:

$$f(\xi) = f'(\xi) = f''(\xi) = \dots = f^{m-1}(\xi) = 0 \text{ e } f^m(\xi) \neq 0$$

onde  $f^{(k)}(\xi)$ ,  $k = 1, 2, \dots, m$  é a derivada de ordem  $k$  da função  $f(x)$  calculada no ponto  $\xi$ .

**Exemplo 51.** Sabe-se que  $\xi = 2$  é uma raiz da equação  $f(x) = x^4 - 5x^3 + 6x^2 + 4x - 8 = 0$ . Assim:

$$\begin{aligned} f(2) &= 0 \\ f'(x) &= 4x^3 - 15x^2 + 12x + 4 \Rightarrow f'(2) = 0 \\ f''(x) &= 12x^2 - 30x + 12 \Rightarrow f''(2) = 0 \\ f'''(x) &= 24x - 30 \Rightarrow f'''(2) = 18 \neq 0 \end{aligned}$$

Logo,  $\xi = 2$  é uma raiz com multiplicidade 3. A Figura 16.4 ilustra o comportamento da função.

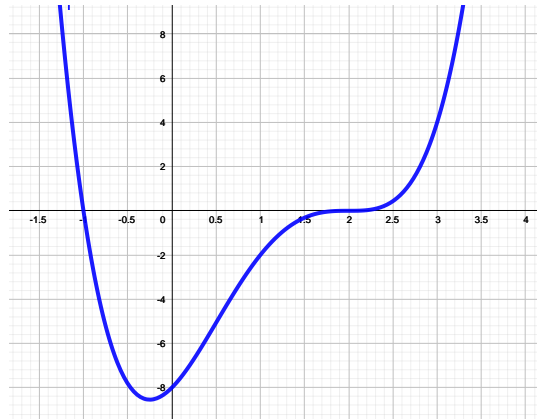


Figura 16.4: Representação gráfica da função  $f(x) = x^4 - 5x^3 + 6x^2 + 4x - 8$ .

**Exemplo 52.** Verifique que  $\xi = 0$  é uma raiz da equação  $f(x) = \text{sen}^2(x) - x\text{sen}(x) + 0,25x^2 = 0$ . Assim:

$$\begin{aligned} f'(x) &= 2\text{sen}(x)\cos(x) - \text{sen}(x) - x\cos(x) + 0,5x \Rightarrow f'(0) = 0 \\ f''(x) &= -2\text{sen}^2(x) + x\text{sen}(x) + 2\cos^2(x) - 2\cos(x) + 0,5 \Rightarrow f''(0) = 0,5 \end{aligned}$$

Portanto,  $\xi = 0$  é uma raiz com multiplicidade 2. A Figura 16.5 ilustra o comportamento da função.

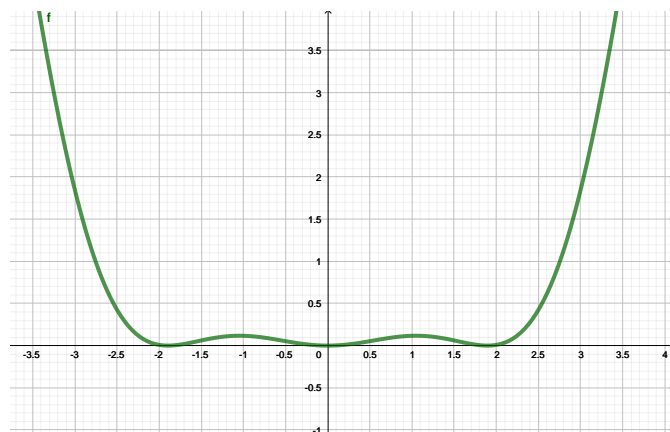


Figura 16.5: Representação gráfica da função  $f(x) = \text{sen}^2(x) - x\text{sen}(x) + 0,25x^2$ .

Observe que a equação  $f(x) = \text{sen}^2(x) - x\text{sen}(x) + 0,25x^2 = 0$  possui outras duas raízes múltiplas cujos valores aproximados são  $1,8954943$  e  $-1,8954943$ .

## 16.2 Determinação das raízes

A determinação das raízes de uma equação envolve duas fases:

**Fase I – Isolamento das raízes:** nesta fase, o objetivo é determinar intervalos que contenham, cada um, uma única raiz.

**Fase II - Refinamento:** trata-se da utilização de métodos numéricos, com precisão pré-fixada, para calcular cada uma das raízes.

### 16.2.1 Isolamento das raízes

**Teorema 16.2. (Teorema do valor intermediário)** Seja  $f(x)$  uma função contínua em um intervalo  $[a, b]$  e  $k$  um número real compreendido entre  $f(a)$  e  $f(b)$ , então existe pelo menos um valor real  $c$ , pertencente ao intervalo aberto  $]a, b[$  tal que  $f(c) = k$ .

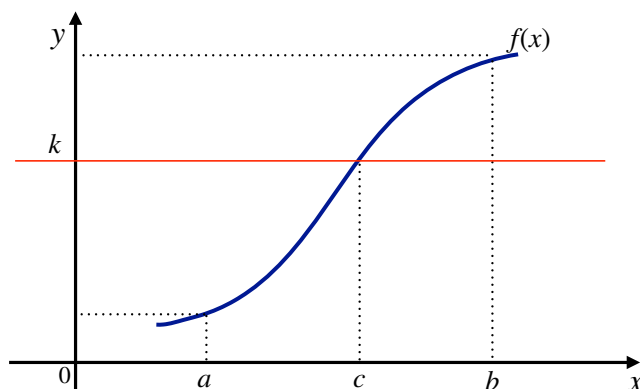
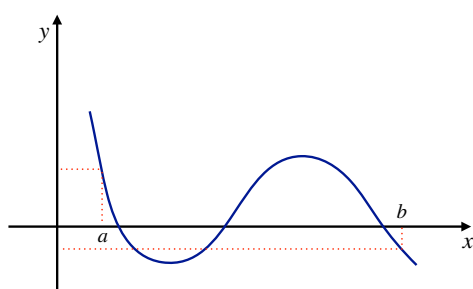


Figura 16.6: Ilustração do Teorema do valor intermediário

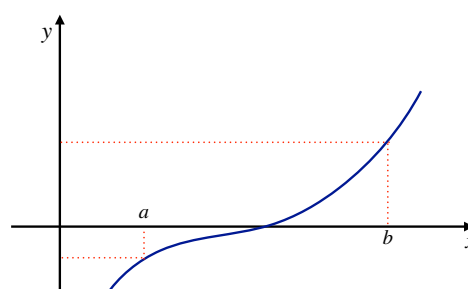
A Figura 16.6 apresenta a ilustração do Teorema 16.2, que tem especial importância na etapa de isolamento das raízes de determinadas funções, principalmente as funções em que não é possível obter os seus zeros por meio de processos algébricos. Assim, considere o corolário seguinte.

**Corolário 16.1. (Cauchy-Bolzano)** Seja  $y = f(x)$  uma função contínua em um intervalo  $[a, b]$ .

- (i) Se  $f(a) \times f(b) < 0$ , então existe um número ímpar de raízes da equação  $f(x) = 0$  no intervalo  $(a, b)$ . Além disso, se  $f'(x)$  preservar o sinal em  $(a, b)$  então a raiz  $\xi$  é única.



(a) Número ímpar de raízes.

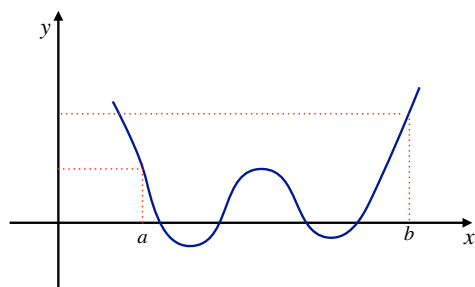


(b) Raízes com multiplicidades ímpar.

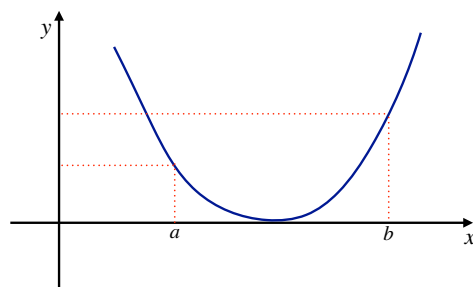
Figura 16.7



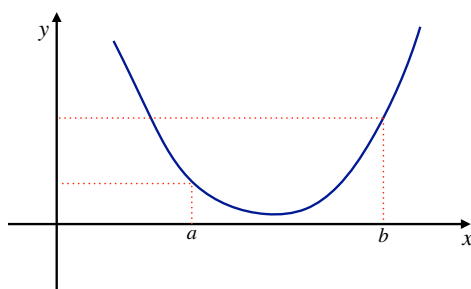
(ii) Se  $f(a) \times f(b) > 0$ , então existe um número par de raízes, ou nenhuma raiz real, da equação  $f(x) = 0$  no intervalo  $(a, b)$ .



(a) Número par de raízes.



(b) Raízes com multiplicidades par.



(c) Nenhuma raiz real.

Figura 16.8

O resultado apresentado no Corolário 16.1 permite concluir que uma das formas de localizar intervalos que contém as raízes da função  $f(x)$  é através da utilização de uma tabela de pontos  $[x_i, f(x_i)]$ ,  $i = 1, 2, \dots, n$ . Para ilustrar este resultado, considere o Exemplo 53.

**Exemplo 53.** Sabe-se que a equação  $f(x) = x^5 - 6x^4 - 14x^3 + 72x^2 + 44x - 180 = 0$  possui três raízes reais situadas no intervalo  $(0, 7)$ . Para isolar as raízes positivas da equação será gerada uma tabela de pontos considerando um passo  $h = 1$ .

$x$	0	1	2	3	4	5	6	7
$f(x)$	-180	-83	20	-21	-260	-535	-348	1255

Tendo em vista que  $f(1) \times f(2) < 0$ ,  $f(2) \times f(3) < 0$  e  $f(6) \times f(7) < 0$  e considerando o Corolário 16.1, conclui-se que a equação dada tem uma raiz em cada destes intervalos:  $(1, 2)$ ,  $(2, 3)$  e  $(6, 7)$ .

Um outra maneira para localizar as raízes de uma equação  $f(x) = 0$  é por meio da utilização de métodos gráficos. São apresentados dois procedimentos gráficos com o objetivo de detectar intervalos que contenham uma única raiz:

## Procedimento I

Este procedimento consiste em esboçar o gráfico de  $y = f(x)$  e identificar onde a curva da função intercepta o eixo das abscissas. Neste(s) ponto(s) teremos a(s) raiz(raízes) de  $f(x)$ .

**Exemplo 54.** Seja a equação  $f(x) = x^3 - 9x + 3 = 0$ . O esboço gráfico de  $f(x)$  é dado pela Figura 16.9.

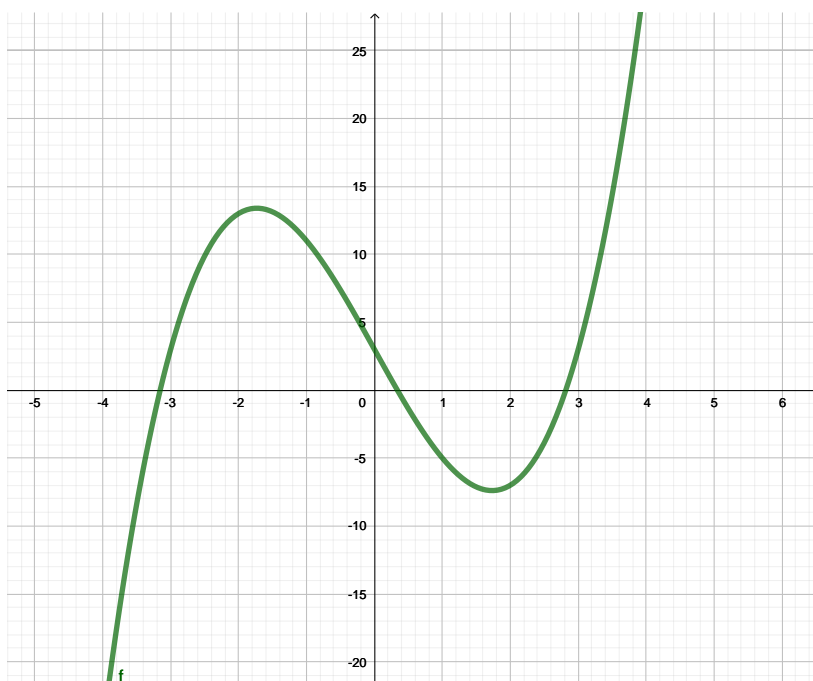


Figura 16.9: Isolamento das raízes da equação  $f(x) = x^3 - 9x + 3 = 0$ .

Conforme mostra a Figura 16.9, a equação  $f(x) = x^3 - 9x + 3 = 0$  possui três raízes isoladas nos intervalos  $(-4, -3)$ ;  $(0, 1)$  e  $(2, 3)$ .

## Procedimento II

Este procedimento consiste transformar a equação  $f(x) = 0$  na equação  $g(x) = h(x)$ , com o objetivo de obter duas funções  $y_1 = g(x)$  e  $y_2 = h(x)$  que sejam conhecidas e mais fáceis de serem traçadas do que a função  $f(x)$ . As raízes da equação original serão dadas pelos pontos onde a curva de  $y_1 = g(x)$  intercepta a curva de  $y_2 = h(x)$ .

**Exemplo 55.** A equação  $f(x) = x^3 - 9x + 3 = 0$  pode ser reescrita como  $x^3 = 9x - 3$ , onde  $y_1 = g(x) = x^3$  e  $y_2 = h(x) = 9x - 3$ . O esboço com os gráficos de  $g(x)$  e  $h(x)$  se encontra na Figura 16.10. Veja que as funções  $g(x) = x^3$  e  $h(x) = 9x - 3$  se interceptam nos intervalos  $(-4, -3)$ ;  $(0, 1)$  e  $(2, 3)$ .

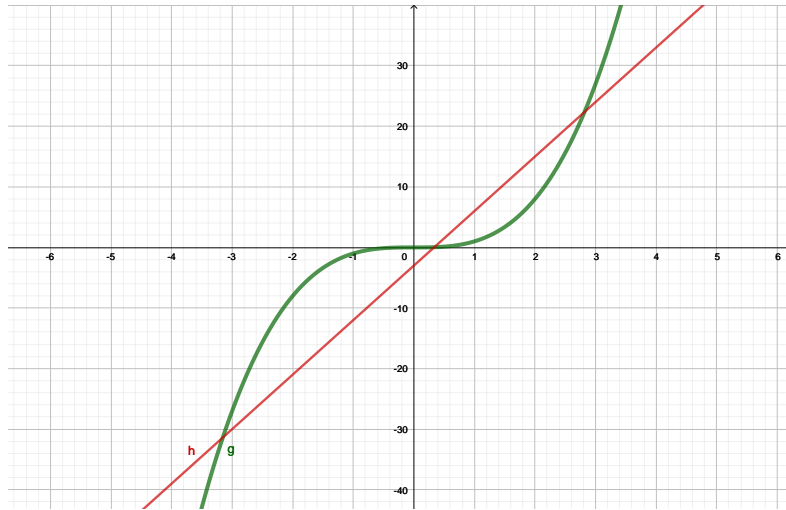


Figura 16.10: Isolamento das raízes da equação  $f(x) = x^3 - 9x + 3 = 0$ , usando os gráficos de  $g(x)$  e  $h(x)$ .

**Exemplo 56.** Seja a equação  $f(x) = e^x + x^2 - 2 = 0$ . Ela pode ser reescrita como  $e^x = 2 - x^2$ , onde  $y_1 = g(x) = e^x$  e  $y_2 = h(x) = -x^2 + 2$ . O esboço com os gráficos de  $g(x)$  e  $h(x)$  está ilustrado na Figura 16.11. Veja que as funções  $g(x) = e^x$  e  $h(x) = -x^2 + 2$  se interceptam nos intervalos  $(-2, -1)$  e  $(0, 1)$ .

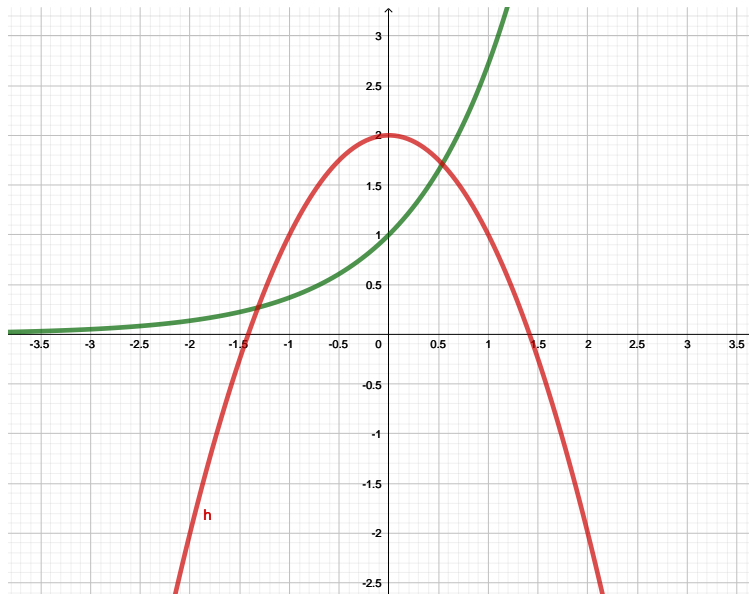


Figura 16.11: Isolamento das raízes da equação  $f(x) = e^x + x^2 - 2 = 0$ .

**Exemplo 57.** Seja  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Para determinar as raízes da equação  $f(x) = (x+1)^2 e^{(x^2-2)} - 1 = 0$ , é preciso reescrevê-la como:

$$(x+1)^2 e^{(x^2-2)} = 1 \Rightarrow (x+1)^2 = \frac{1}{e^{(x^2-2)}} \Rightarrow (x+1)^2 = e^{(-x^2+2)},$$

em que  $y_1 = g(x) = (x+1)^2$  e  $y_2 = h(x) = e^{(-x^2+2)}$ . O esboço com os gráficos de  $g(x)$  e  $h(x)$  está ilustrado na Figura 16.12. Veja que as funções  $y_1 = g(x) = (x+1)^2$  e  $y_2 = h(x) = e^{(-x^2+2)}$  se interceptam apenas nos intervalos  $[-2, -1]$  e  $[0, 2]$ .

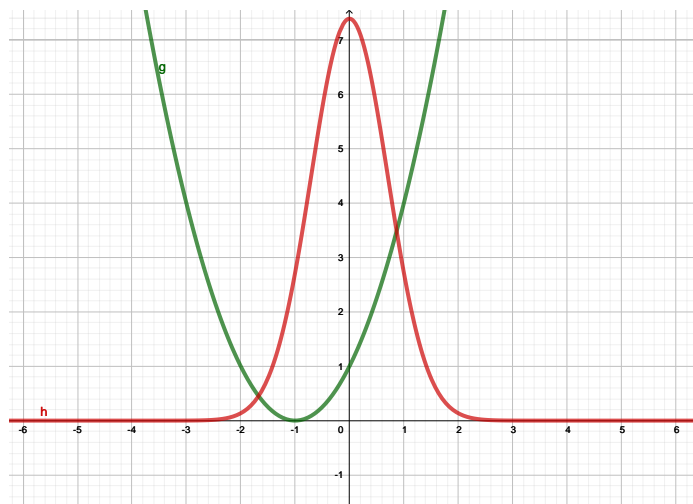


Figura 16.12: Isolamento das raízes da equação  $f(x) = (x+1)^2 e^{(x^2-2)} - 1 = 0$ .

Para melhorar a precisão dos resultados obtidos, pode-se analisar os valores de  $f(x)$  para  $x = -3, -2, \dots, 2, 3$ , conforme a tabela a seguir:

$x$	-3	-2	-1	0	1	2	3
$f(x)$	4385,5	6,4	-1	-0,9	0,5	65,5	17545,1

Assim, considerando o esboço das funções  $g(x) = (x+1)^2$  e  $h(x) = e^{(-x^2+2)}$  e o tabelamento de  $f(x) = (x+1)^2 e^{(x^2-2)} - 1$ , pode-se concluir que a equação dada possui zeros nos intervalos  $[-2, -1]$  e  $[0, 1]$ .

## 16.3 Estudo Especial de Equações Polinomiais

Considerando a existência de vários teoremas da Álgebra que fornecem informações relevantes sobre as equações polinomiais, será tratada, nesta seção, de forma especial, a execução da fase de isolamento das raízes para este tipo de equação. Posteriormente, serão apresentados métodos numéricos para o cálculo das raízes reais de uma equação qualquer.



**Definição 16.4.** Toda equação da forma:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 = 0$$

com  $a_i \in \mathbb{R}$ ,  $\forall i = 0, 1, \dots, n$  é dita polinomial. O número natural  $n$  é dito o grau da equação.

Então, os seguintes resultados são válidos:

- a) Uma equação polinomial de grau  $n$  tem exatamente  $n$  raízes, reais ou complexas, desde que uma raiz com multiplicidade  $k$  seja contada  $k$  vezes.
- b) Se os coeficientes de uma equação polinomial forem reais, então as suas raízes complexas ocorrerão em pares conjugados. Em outras palavras, se  $a + bi$  for uma raiz complexa de  $P(x)$ , então  $a - bi$  também será raiz de  $P(x)$ .
- c) Uma equação polinomial de grau ímpar, com coeficientes reais, tem, no mínimo, uma raiz real.
- d) Toda equação polinomial de grau par, cujo termo independente é negativo, tem pelo menos uma raiz real positiva e outra negativa.
- e) Toda equação polinomial de grau ímpar, tem pelo menos uma raiz real com o sinal contrário ao do termo independente.
- f) Se  $x_1, x_2, \dots, x_n$  forem as raízes de  $P(x)$ , então  $P(x)$  pode ser expresso na forma fatorada por:

$$P(x) = a_n(x - x_1)(x - x_2) \cdots (x - x_n)$$

- g) Se os valores numéricos de dois polinômios de grau menor ou igual a  $n$  coincidem para mais do que  $n$  valores distintos de  $x$ , os polinômios são idênticos.

### 16.3.1 Delimitação das raízes reais

#### Limite Superior das Raízes Positivas

Para determinar o **Limite Superior das Raízes Positivas (LSRP)** considere o Teorema 16.3.

**Teorema 16.3. (Lagrange)** Seja  $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 = 0$  uma equação polinomial de grau  $n$  na qual  $a_n > 0$  e  $a_0 \neq 0$ . Para o limite superior das suas raízes positivas, caso existam, pode ser tomado o número:

$$L = 1 + \sqrt[n-k]{\frac{M}{a_n}} \quad (16.1)$$

Onde  $k$  é o grau do primeiro termo negativo e  $M$  o módulo do menor coeficiente negativo.

**Exemplo 58.** Considere a equação

$$f(x) = x^5 - 2x^4 - 7x^3 + 9x^2 + 8x - 6 = 0$$

Para determinar o **limite superior das raízes positivas (LSRP)**, é necessário que  $a_n = a_5 = 1 > 0$  e  $a_0 = -6 \neq 0$ . Assim, dado que  $k = 4$  e  $M = 7$ :

$$L_{sup} = 1 + \sqrt[5-4]{\frac{7}{1}} = 1 + \sqrt[1]{\frac{7}{1}} = 1 + 7 = 8$$

Logo, a equação  $f(x) = x^5 - 2x^4 - 7x^3 + 9x^2 + 8x - 6 = 0$  **não possui raízes positivas maiores que 8.**

## Limite Inferior das Raízes Negativas

Para determinar o **Limite Inferior das Raízes Negativas (LIRN)** considere os seguintes passos:

- (i) Toma-se a equação auxiliar  $f^*(x) = f(-x) = 0$ .
- (ii) Aplica-se o teorema de Lagrange em  $f^*(x) = 0$  para determinar  $L_{sup}^*$ , que é o limite superior das raízes positivas de  $f^*(x)$ .
- (iii) O limite inferior das raízes negativas de  $f(x) = 0$  será dado por  $L_{inf} = -L_{sup}^*$ .

Para mostrar que a afirmativa (iii) é verdadeira, considere

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 = 0$$

uma equação com as raízes  $r_1, r_2, r_3, \dots, r_n$ . Portanto,  $f(x)$  pode ser reescrita na forma fatorada, por:

$$f(x) = a_n (x - r_1)(x - r_2)(x - r_3) \cdots (x - r_n) = 0.$$

Substituindo  $x$  por  $-x$  em  $f(x)$ , tem-se:

$$f(-x) = a_n (-x - r_1)(-x - r_2)(-x - r_3) \cdots (-x - r_n) = 0$$

que possui como raízes os termos  $-r_1, -r_2, -r_3, \dots, -r_n$ . Sendo algum  $r_i, i = 1, 2, \dots, n$ ; a maior raiz positiva de  $f(-x) = 0$ , então  $-r_i$  será a menor raiz negativa de  $f(x) = 0$ . O que confirma a afirmativa (iii).

**Exemplo 59.** Considere a equação

$$f(x) = x^5 - 2x^4 - 7x^3 + 9x^2 + 8x - 6 = 0.$$

Para determinar **limite inferior das raízes negativas**, é preciso:

- (i) escrever a equação auxiliar  $f^*(x) = f(-x) = 0$ :

$$f^*(x) = f(-x) = (-x)^5 - 2(-x)^4 - 7(-x)^3 + 9(-x)^2 + 8(-x) - 6 = 0.$$

Logo:

$$f^*(x) = -x^5 - 2x^4 + 7x^3 + 9x^2 - 8x - 6 = 0.$$

- (ii) Aplicar o teorema de Lagrange em  $f^*(x) = 0$  para determinar  $L_{sup}^*$ :

De acordo com o teorema de Lagrange  $a_5$  deve ser maior que zero. Para garantir que esta condição seja satisfeita, basta multiplicar  $f^*(x) = 0$  por  $(-1)$ , assim:

$$f^*(x) = x^5 + 2x^4 - 7x^3 - 9x^2 + 8x + 6 = 0.$$

Assim:  $a_5 = 1 > 0$ ,  $a_0 = 6 \neq 0$   $k = 3$  e  $M = 9$ . Logo:

$$L_{sup}^* = 1 + \sqrt[5-3]{\frac{9}{1}} = 1 + \sqrt[2]{\frac{9}{1}} = 1 + 3 = 4.$$

- (iii) O limite inferior das raízes negativas de  $f(x) = 0$  será dado por  $L_{inf} = -L_{sup}^*$ . Assim:

$$L_{inf} = -L_{sup}^* = -4.$$

Logo, a equação  $f(x) = x^5 - 2x^4 - 7x^3 + 9x^2 + 8x - 6 = 0$  não possui raízes negativas menores que  $-4$ .

### 16.3.2 Enumeração das raízes reais

A regra dos sinais de Descartes, primeiramente descrita por René Descartes é um teorema que determina o número de raízes positivas e negativas de um polinômio.

**Teorema 16.4. (Regra de sinais de Descartes)** Considere um polinômio  $P(x)$  cujo os termos de um polinômio são coeficientes reais e estão colocados em ordem decrescente de grau. Então o número de raízes positivas do polinômio é ou igual ao número de permutações de sinal ou menor que este número por uma diferença par.

Em outras palavras, o número de permutações (ou variações de sinais) é igual ao número de raízes positivas acrescido do número de raízes imaginárias (que sempre acontecem ao pares em polinômios de coeficientes reais). Para determinar o número de raízes negativas, basta trocar  $x$  por  $-x$  e calcular o número de raízes positivas de  $f(-x) = 0$ , este valor será o número de raízes negativas de  $f(x) = 0$ .

**Exemplo 60.** Seja a equação

$$f(x) = x^5 - 2x^4 - 7x^3 + 9x^2 + 8x - 6 = 0.$$

A regra dos sinais de Descartes pode ser utilizada para enumerar as raízes reais da seguinte forma:

- **O número de raízes positivas:** considere as seguintes variações de sinais:  $+1, -2, -7, +9, +8, -6$ . Logo existem 3 ou 1 raiz (raízes) positiva(s).
- **O número de raízes negativas:** Seja a função auxiliar do exemplo anterior

$$f^*(x) = -x^5 - 2x^4 + 7x^3 + 9x^2 - 8x - 6 = 0.$$

Dada que a sequência dos coeficientes é  $-1, -2, +7, +9, -8, -6$ , a equação tem 2 ou nenhuma raiz negativa.

Dessa forma, para a equação  $f(x) = x^5 - 2x^4 - 7x^3 + 9x^2 + 8x - 6 = 0$  temos as seguintes possibilidades:

Raízes Reais Positivas	Raízes Reais Negativas	Raízes Complexas
3	2	0
3	0	2
1	2	2
1	0	4



## 16.4 Exercícios



**E. 1. Localize graficamente os intervalos que contém raízes das seguintes equações.**

$$a) f(x) = x^2 + x - 5 = 0 \qquad d) f(x) = e^x - x - 2 = 0$$

$$b) f(x) = e^x - x^2 + 4 = 0 \qquad e) f(x) = xe^x - 1 = 0$$

$$c) f(x) = x^2 - \ln(x) = 0 \qquad f) f(x) = x^2 - \frac{1}{x} = 0$$

**E. 2. Isole os zeros da equação  $f(x) = x \ln(x) - 3.2 = 0$**

*Dica: Use a tabela de valores para  $f(x)$  e analise as mudanças de sinais.*

$x$	1	2	3	4	5	10
$f(x)$						

**E. 3. Calcule os limitantes e a quantidade de raízes reais existentes das equações polinomiais: seguintes. Utilize o tabelamento da função de  $f(x)$  para analisar as variações de sinal e isolar as raízes dessas equações.**

$$a) f(x) = 2x^3 - 5x^2 - x + 3 = 0$$

$$c) f(x) = -x^5 + 2x^4 + 4x^3 - x^2 + x + 1 = 0$$

$$b) f(x) = x^6 + 5x^5 + x^4 + 3x^3 - x^2 + x - 2 = 0$$

$$d) f(x) = 3x^4 - 5x^3 - x^2 - 8x + 4 = 0$$



This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



# 17

## Método da Bisseção

### Sumário da Aula

17.1 Método da Bisseção . . . . .	183
17.2 Exercícios . . . . .	188

### 17.1 Método da Bisseção

Seja  $y = f(x)$  uma função contínua em um intervalo  $[a, b]$  que contém uma, e só uma, raiz,  $\xi$ , da equação  $f(x) = 0$ . O método da bisseção é baseado no teorema do valor intermediário, que afirma que em uma função contínua no intervalo  $[a, b]$  que satisfaz a condição  $f(a)f(b) < 0$ , ou seja, valores de  $f(a)$  e  $f(b)$  possuem sinais opostos, existe um  $\xi \in [a, b]$  tal que  $f(\xi) = 0$ , isto é, existe pelo menos uma raiz no intervalo  $[a, b]$ , como mostra a Figura 17.1.



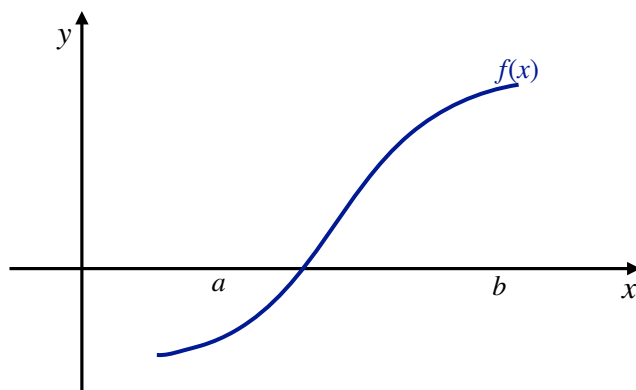


Figura 17.1: Raiz única de  $f(x)$  no intervalo  $[a, b]$ .

A ideia geral deste método consiste em dividir o intervalo  $[a, b]$ , de forma iterativa, ao meio e obter dois subintervalos, estando a raiz em um dos subintervalos calculados. Para verificar se a raiz está contida na primeira ou na segunda metade do intervalo inicial, utiliza-se o teorema de Cauchy-Bolzano. Em seguida, aplica-se novamente o processo para aquela metade que contém a raiz, ou seja, o subintervalo em que a função,  $y = f(x)$ , tem valores numéricos com sinais opostos nos seus extremos. Este processo é repetido até que um critério de parada seja satisfeito, e uma sequência de aproximações  $x_0, x_1, \dots, x_k$  para a raiz  $\xi$  é obtida. A Figura 17.2 ilustra o procedimento.

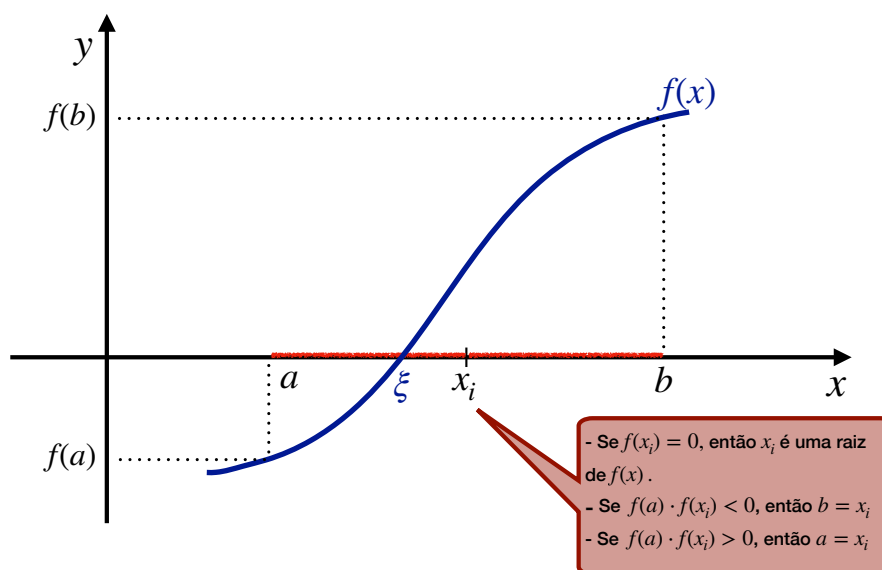


Figura 17.2: Representação gráfica do método da bisseção.

## Função de iteração

Considerando que em cada iteração é atualizado o ponto  $a$  ou  $b$ , tem-se que a função de iteração do método da bisseção é dado por:

$$x_i = \frac{a_i + b_i}{2}, \quad i = 0, 1, 2, \dots, k. \quad (17.1)$$

## Critério de parada

Dada uma precisão  $\varepsilon$ , o processo iterativo é finalizado quando se obtém um intervalo cujo tamanho é menor ou igual a  $\varepsilon$ , ou seja,  $|b_i - a_i| \leq \varepsilon$ . Assim, qualquer ponto contido neste intervalo pode ser tomado como uma estimativa para a raiz. Também é comum utilizar como critério de parada um número máximo de iterações.

## Critério de convergência

Se  $y = f(x)$  for contínua em  $[a, b]$ ,  $f(a)f(b) < 0$  e além disso  $f'(x)$  preservar o sinal neste intervalo, então o método da bisseção gera uma sequência que converge para uma raiz de  $f(x) = 0$ .

## Estimativa do número de iterações

O método da bisseção permite que seja estimado, a priori, o número mínimo de iterações para calcular uma raiz  $\xi$  com uma precisão  $\varepsilon$  a partir de um intervalo  $[a, b]$ . As iterações geram uma sequência de intervalos encaixados da forma:

$$\{[a, b], [a_1, b_1], [a_2, b_2], [a_3, b_3], \dots, [a_k, b_k]\}.$$

Como cada intervalo gerado tem tamanho igual a metade do intervalo anterior, temos:

$$(b_1 - a_1) = \frac{b - a}{2}, \quad (b_2 - a_2) = \frac{b_1 - a_1}{2} \Rightarrow (b_2 - a_2) = \frac{b - a}{2^2}$$

Assim,  $(b_3 - a_3) = \frac{b - a}{2^3}$  e de forma análoga chega-se até  $(b_k - a_k) = \frac{b - a}{2^k}$ . Como deseja-se obter  $k$  tal que  $b_k - a_k \leq \varepsilon$ :

$$\frac{(b - a)}{2^k} \leq \varepsilon \Rightarrow k \geq \frac{\log(b - a) - \log(\varepsilon)}{\log(2)}$$

Logo,  $k$  é o número mínimo de iterações que devem ser realizadas para obter uma aproximação para  $\xi$  com precisão  $\varepsilon$  desejada.

## Pseudocódigo

A sequência de aproximações  $x_0, x_1, \dots, x_k$  para a raiz  $\xi$  da equação  $f(x) = 0$  é obtida da seguinte maneira:

- i) Determine o intervalo inicial  $[a, b]$  tal que  $f(a)f(b) < 0$  e  $f'(x)$  preserve o sinal neste intervalo;
- ii) Calcule  $x_i = \frac{a + b}{2}$  (ponto médio do intervalo  $[a, b]$ );
- ii) Se  $f(x_i) = 0$ , então  $x_i$  é uma raiz de  $f(x) = 0$ ;
- iv) Se  $f(a)f(x_i) < 0$ , então a raiz está no primeiro subintervalo, ou seja  $\xi \in [a, x_0]$ . Assim,  $b = x_0$ .
- v) Se  $f(a)f(x_i) > 0$ , então a raiz está no segundo subintervalo, ou seja  $\xi \in [x_0, b]$ . Assim,  $a = x_0$ .

E assim, sucessivamente, até que se determine uma raiz aproximada  $\bar{x}$  da raiz exata  $\xi$  da equação  $f(x) = 0$  com precisão  $\varepsilon$  desejada.

Sejam:  $\varepsilon > 0$ , um intervalo inicial  $[a, b]$  que contenha a raiz ( $f(a)f(b) < 0$ ) e  $k$  o número máximo de iterações. O Pseudocódigo do método da bisseção está escrito no Algoritmo 6.

```

1: Entrada:  $[a, b], f(\cdot), \varepsilon$  e  $k$ .
2:  $erro = \infty$ ;
3:  $fa = f(a), fb = f(b)$ ;
4:  $i = 0$ ;
5: enquanto ( $erro \geq \varepsilon$ ) e ( $i \leq k$ ) faça
6:    $x_i = \frac{a+b}{2}$ ;
7:    $fx = f(x_i)$ ;
8:   se ( $fa)(fx) < 0$  então
9:      $b = x_i$ ;
10:     $fb = f(b)$ ;
11:   senão
12:      $a = x_i$ ;
13:      $fa = f(a)$ ;
14:   fim se
15:    $erro = |b - a|$ ;
16:    $i = i + 1$ ;
17: fim enquanto
18: Saída:  $x_i$ .

```

Algoritmo 6: Método da Bisseção.

**Exemplo 61.** Seja a equação  $f(x) = x^3 - 9x + 3 = 0$ . Estime a raiz de  $f(x) = 0$  contida no intervalo  $(0, 1)$  considerando uma precisão  $\varepsilon = 0,065$ .

- Seja o intervalo inicial  $[a_0, b_0] = [0, 1]$ . Existe um número ímpar de raízes neste intervalo, pois  $f(0)f(1) = (3)(-5) < 0$ .
- Como  $f'(x) = 3x^2 - 9 < 0$  em todo o intervalo dado, a raiz  $\xi \in [0, 1]$  é única.
- Cálculo do número mínimo de iterações:

$$k \geq \frac{\log(b-a) - \log(\varepsilon)}{\log(2)} = \frac{\log(1-0) - \log(0,065)}{\log(2)} = 3,9434$$

Portanto  $k \geq 3,9434$ , isto é, deve-se executar no mínimo  $k = 4$  iterações para obter a raiz aproximada  $\bar{x}$  com a precisão  $\varepsilon$  desejada.

- O cálculo da sequência de raízes aproximadas está apresentado na tabela a seguir:

$k$	$a_k$	$b_k$	$ b_k - a_k $	$x_k$	$f(x_k)$	$f(a_k)f(x_k)$
0	0	1	1	0,5	-1,375	(+)(-)
1	0	0,5	0,5	0,25	0,765	(+)(+)
2	0,25	0,5	0,25	0,375	-0,322	(+)(-)
3	0,25	0,375	0,125	0,313	0,213	(+)(+)
4	0,313	0,375	0,062	0,344	-0,0553	

Veja que o critério de parada ( $|b_4 - a_4| \leq \varepsilon$ ) foi satisfeito na iteração quatro. Logo, a solução aproximada é  $\xi \approx 0,344$ .

**Exemplo 62.** Usando o método da bisseção, resolva a equação  $f(x) = x^2 + \ln(x) = 0$  com  $\varepsilon = 0,01$ .

- Será isolada graficamente uma vizinhança para a(s) raiz (raízes) da equação: para tanto,  $f(x) = 0$  deve ser reescrita como  $g(x) = h(x)$ , com  $g(x) = x^2$  e  $h(x) = -\ln(x)$ . O esboço para os gráficos  $g(x)$  e  $h(x)$  é:

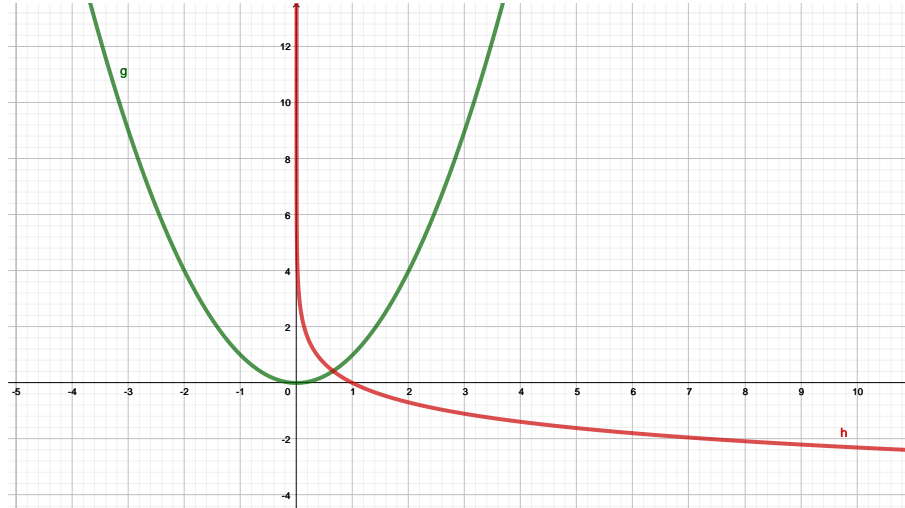


Figura 17.3: O esboço com os gráficos das funções  $g(x) = x^2$  e  $h(x) = -\ln(x)$ .

- Observe na Figura 17.3 que existe uma raiz  $\xi$  que pertence ao intervalo  $[0, 1, 1]$ . Esta raiz encontra-se na interseção dos gráficos de  $g(x)$  e  $h(x)$ .
- Para garantir que neste intervalo a raiz é única, deve ser aplicado o Teorema de Cauchy-Bolzano:
  - $f(0,1)f(1) = (-2,29263)(1) < 0$ , portanto existe pelo menos uma raiz neste intervalo.
  - $f'(x) = 2x + \frac{1}{x} > 0$  em todo o intervalo  $[0, 1, 1]$ . Logo a raiz  $\xi \in [0, 1, 1]$  é única.
- Cálculo do número mínimo de iterações:

$$k \geq \frac{\log(b-a) - \log(\varepsilon)}{\log(2)} = \frac{\log(1-0,1) - \log(0,01)}{\log(2)} = 6,4919$$

Portanto  $k \geq 6,4919$ , isto é, deve-se executar no mínimo  $k = 7$  iterações para obter a raiz aproximada  $\bar{x}$  com a precisão  $\varepsilon$  desejada.

- O cálculo da sequência de raízes aproximadas está apresentado na tabela:

$k$	$a_k$	$b_k$	$ b_k - a_k $	$x_k$	$f(x_k)$	$f(a_k)f(x_k)$
0	0,1	1	0,9000	0,55	-0,2953	(-)(-)
1	0,55	1	0,4500	0,6625	0,3457	(-)(+)
2	0,55	0,7750	0,2250	0,6063	0,0272	(-)(+)
3	0,55	0,6625	0,1125	0,6344	-0,1329	(-)(-)
4	0,6063	0,6625	0,0563	0,6384	-0,0527	(-)(-)
5	0,6344	0,6625	0,0281	0,6484	-0,0127	(-)(-)
6	0,6484	0,6625	0,0141	0,6555	0,0072	(-)(+)
7	0,6484	0,6555	0,0070	0,6520	-0,0027	

Como o critério de parada foi satisfeito, ou seja,  $|b_4 - a_4| \leq \varepsilon$ , a solução aproximada é  $\xi \approx 0,6520$ .



---

Observação

---

O método da bisseção exige pouco esforço computacional e sempre gera uma sequência convergente. Porém, a convergência é lenta. Notadamente, se o intervalo inicial tiver  $|b - a|$  um tamanho muito maior que uma precisão, o número de iterações tende ser muito grande. Como exemplo, considere o tamanho de intervalo  $|b - a| = 2$  e uma precisão  $\varepsilon = 10^{-6}$ . Para estes valores, a estimativa do número de iterações será:

$$k \geq \frac{\log(b - a) \log(\varepsilon)}{\log(2)} \Rightarrow k \geq 20,9 \Rightarrow k = 21$$

---

## 17.2 Exercícios



E. 1. Para cada uma das equações seguintes, calcule pelo menos uma raiz com precisão de 0,001 pelo método da bisseção.

a)  $f(x) = 2x^3 - 5x^2 - x + 3 = 0$

b)  $f(x) = e^x - x^2 + 4 = 0$



[illegible]



# 18

## Método da Falsa Posição

### Sumário da Aula

18.1 Método da Falsa Posição . . . . .	191
18.1.1 Considerações Finais . . . . .	196
18.2 Exercícios . . . . .	196

### 18.1 Método da Falsa Posição

Seja  $y = f(x)$  uma função contínua em um intervalo  $[a, b]$  que contém uma, e só uma, raiz,  $\xi$ , da equação  $f(x) = 0$ . O método da falsa posição, **ou regula falsi**, também se fundamenta no teorema do valor intermédio. Assim, sendo  $f(x)$  contínua no intervalo  $[a, b]$  e a condição  $f(a)f(b) < 0$  satisfeita, então existe um  $\xi \in [a, b]$  tal que  $f(\xi) = 0$ , isto é, existe pelo menos uma raiz no intervalo  $[a, b]$ .



A ideia geral do método da falsa posição consiste em dividir o intervalo  $[a, b]$ , de forma iterativa, no ponto em que a reta que passa pelos pontos  $[a, f(a)]$  e  $[b, f(b)]$  intercepta o eixo das abscissas. A Figura 18.1 ilustra o procedimento.

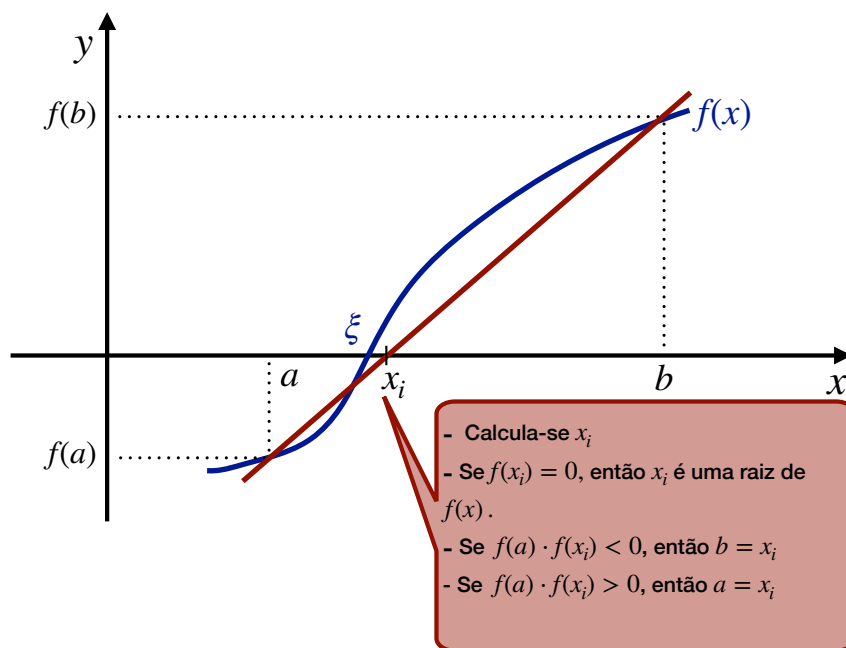
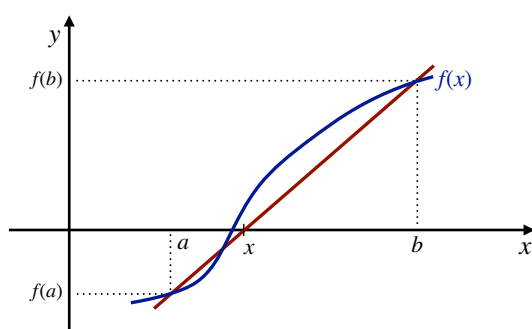


Figura 18.1: Representação gráfica do método da falsa posição.

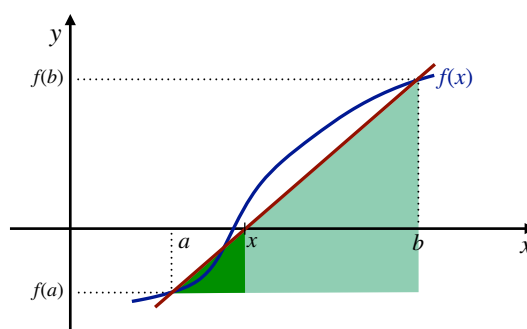
Para verificar se a raiz está contida no primeiro ou no segundo subintervalo obtido, utiliza-se o teorema de Cauchy-Bolzano. Em seguida, aplica-se novamente o processo no subintervalo que contém a raiz, ou seja, no subintervalo em que a função,  $y = f(x)$ , tem valores numéricos com sinais opostos nos seus extremos. Este processo é repetido até que um critério de parada seja satisfeito, e uma sequência de aproximações  $x_0, x_1, \dots, x_k$  para a raiz  $\xi$  é obtida.

## Função de iteração

Para determinar a função de iteração, basta considerar a equação da reta que passa pelos pontos  $[a, f(a)]$  e  $[b, f(b)]$ , conforme a Figura 18.2 (a).



(a) Reta que passa pelos pontos.



(b) Determinação função de iteração por semelhança de triângulos.

Figura 18.2: Obtenção da função de iteração do método da falsa posição.

Por semelhança de triângulos, caso ângulo ângulo (AA), os dois triângulos formados na Figura 18.2 (b) são semelhantes. Assim, temos que os seus lados são proporcionais, ou seja, a razão entre lados correspondentes

é constante. Logo:

$$\frac{x-a}{b-a} = \frac{P(x) - f(a)}{f(b) - f(a)}.$$

Que pode ser reescrita como:

$$x - a = \frac{[P(x) - f(a)][b - a]}{f(b) - f(a)}.$$

Dado que  $x$  é um ponto no eixo das abscissas,  $P(x) = 0$ . Assim:

$$x = a - \frac{f(a)(b-a)}{f(b) - f(a)} \quad \text{ou} \quad x = \frac{af(b) - bf(a)}{f(b) - f(a)}.$$

Neste texto, será como função de iteração do método da falsa posição a Equação 18.1.

$$x_k = \frac{af(b) - bf(a)}{f(b) - f(a)}, \quad k = 0, 1, \dots \quad (18.1)$$

onde, em cada iteração, atualiza-se  $a$  ou  $b$ .

Note que no método da falsa posição procura-se gerar, em cada iteração, uma aproximação para a raiz cuja imagem seja a menor possível, ou seja, uma aproximação tal que  $|f(x_k)| \geq \varepsilon$ , sem se preocupar com a diminuição do tamanho do intervalo  $[a, b]$  que a contém.

## Critério de parada

Dada uma precisão  $\varepsilon$ , o processo iterativo é finalizado quando se obtém um ponto  $x_k$ ,  $k = 0, 1, \dots$ ; tal que  $|f(x_k)| \leq \varepsilon$  e, então, este ponto é tomado como uma estimativa para a raiz de  $f(x) = 0$ . Também é comum utilizar como critério de parada um número máximo de iterações.

## Critério de convergência

Se  $y = f(x)$  for contínua em  $[a, b]$ ,  $f(a)f(b) < 0$  e além disso,  $f'(x)$  preservar o sinal neste intervalo, então o método da falsa posição gera uma sequência que converge para uma raiz de  $f(x) = 0$ .

## Pseudocódigo

A sequência de aproximações  $x_0, x_1, \dots, x_k$  para a raiz  $\xi$  da equação  $f(x) = 0$  é obtida da seguinte maneira:

i) Determine o intervalo inicial  $[a, b]$  tal que  $f(a)f(b) < 0$  e  $f'(x)$  preserve o sinal neste intervalo;

ii) Calcule  $x_i = \frac{af(b) - bf(a)}{f(b) - f(a)}$ ;

iii) Se  $f(x_i) = 0$ , então  $x_i$  é uma raiz de  $f(x) = 0$ ;

iv) Se  $f(a)f(x_i) < 0$ , então a raiz está no primeiro subintervalo, ou seja  $\xi \in [a, x_0]$ . Assim,  $b = x_0$ .

v) Se  $f(a)f(x_i) > 0$ , então a raiz está no segundo subintervalo, ou seja  $\xi \in [x_0, b]$ . Assim,  $a = x_0$ .

E assim, sucessivamente, até que se determine uma raiz aproximada  $\bar{x}$  da equação  $f(x) = 0$  com precisão  $\varepsilon$  desejada. Sejam  $\varepsilon > 0$ , um intervalo inicial  $[a, b]$  que contenha a raiz ( $f(a)f(b) < 0$ ) e  $k$  o número máximo de iterações. O Pseudocódigo do método da falsa posição está descrito no Algoritmo 7.

```

1: Entrada:  $[a, b], f(\cdot), \varepsilon$  e  $k_{max}$ .
2:  $erro = \infty$ ;
3:  $fa = f(a), fb = f(b)$ ;
4:  $i = 0$ ;
5: enquanto  $(erro \geq \varepsilon)$  e  $(i \leq k_{max})$  faça
6:    $x_i = \frac{af(b) - bf(a)}{f(b) - f(a)}$ ;
7:    $fx = f(x_i)$ ;
8:   se  $(fa)(fx) < 0$  então
9:      $b = x_i$ ;
10:     $fb = f(b)$ ;
11:   senão
12:      $a = x_i$ ;
13:      $fa = f(a)$ ;
14:   fim se
15:    $erro = |f(x_i)|$ ;
16:    $i = i + 1$ ;
17: fim enquanto
18: Saída:  $x_i$ .

```

Algoritmo 7: Método da Falsa Posição.

**Exemplo 63.** Seja a equação  $f(x) = x^3 - 9x + 3 = 0$ . Estime a raiz de  $f(x) = 0$  contida no intervalo  $(0, 1)$  considerando uma precisão  $\varepsilon = 0,065$  e o método da falsa posição.

**Solução:**

- Seja o intervalo inicial  $[a_0, b_0] = [0, 1]$ . Existe um número ímpar de raízes neste intervalo, pois  $f(0)f(1) = (3)(-5) < 0$ .
- Como  $f'(x) = 3x^2 - 9 < 0$  em todo o intervalo dado, a raiz  $\xi \in [0, 1]$  é única.
- O cálculo da sequência de raízes aproximadas está apresentado na tabela a seguir:

$k$	$a_k$	$b_k$	$x_k$	$f(x_k)$	$ f(x_k) $	$f(a_k)f(x_k)$
0	0	1	0,3750	-0,3223	0,3223	(+)(-)
1	0	0,375	0,3386	-0,0088	0,0088	(+)(-)

Como o critério de parada foi satisfeito, ou seja,  $|f(x_2)| \leq \varepsilon$ , a solução aproximada da equação  $f(x) = x^3 - 9x + 3 = 0$  é  $\xi \approx 0,3386$ .

**Exemplo 64.** Usando o método da falsa posição, calcule a maior raiz positiva da equação  $f(x) = x^4 - 14x^2 + 24x - 10 = 0$  com  $\varepsilon = 0,01$ .

**Solução:**

- Isolando a(s) raiz (raízes) da equação:

**a) Limitante das raízes reais utilizando o Teorema de Lagrange:**

**a.1)** Limite superior das raízes positivas:  $n = 4, a_4 > 0, a_0 \neq 0, k = 2, M = 14$

$$L_{sup} = 1 + \sqrt[2]{\frac{14}{1}} = 4,74$$

**a.2)** Limite inferior das raízes negativas: função auxiliar  $f^*(x) = f(-x) = x^4 - 14x^2 - 24x - 10 = 0$ , então  $n = 4, a_4 > 0, a_0 \neq 0, k = 2, M = 24$

$$L_{sup}^* = 1 + \sqrt[2]{\frac{24}{1}} = 5,89 \Rightarrow L_{inf} = -L_{sup}^* = -5,89$$

**b) Enumeração das raízes reais utilizando a regra dos sinais de Descartes:**

**b.1)** Raízes positivas:  $+1, -14, +24, -10 \Rightarrow 3 \text{ ou } 1$ .

**b.2)** Raízes negativas: função auxiliar  $f^*(x) = f(-x) = x^4 - 14x^2 - 24x - 10 = 0$ . Logo  $+1, -14, -24, -10 \Rightarrow 1 \text{ raiz}$ .

**b) Separação das raízes reais utilizando tabelamento da função:**

$x$	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5
$f(x)$	638	145	-74	-127	-98	-47	-10	1	-2	17	118	385

Observando a tabela, podemos concluir que existem raízes pertencentes aos seguintes intervalos  $\xi_1 \in [-5, -4]$ ,  $\xi_2 \in [0, 1]$ ,  $\xi_3 \in [1, 2]$  e  $\xi_4 \in [2, 3]$ . A maior raiz positiva da equação  $f(x) = 0$  está no intervalo  $[2, 3]$ .

- Aplicando o Teorema de Cauchy-Bolzano é possível verificar que no intervalo  $[2, 3]$  a raiz é única, veja:
  - $f(2)f(3) = (-2)(17) < 0$ , portanto existe pelo menos uma raiz neste intervalo.
  - $f'(x) = 4x^3 - 28x + 24 > 0$  em todo o intervalo  $[2, 3]$ . Logo a raiz  $\xi \in [2, 3]$  é única.
- O cálculo da sequência de raízes aproximadas aplicando o método da falsa posição está apresentado na tabela a seguir:

$k$	$a_k$	$b_k$	$x_k$	$f(x_k)$	$ f(x_k) $	$f(a_k)f(x_k)$
0	2	3	2,1053	-1,8797	1,8797	$(-)(-)$
1	2,1053	3	2,1953	-1,5621	1,5621	$(-)(-)$
2	2,1943	3	2,2621	-1,1639	1,1639	$(-)(-)$
3	2,2621	3	2,3094	-0,7964	0,7964	$(-)(-)$
4	2,3094	3	2,3403	-0,5130	0,5130	$(-)(-)$
5	2,3403	3	2,3597	-0,3176	0,3176	$(-)(-)$
6	2,3597	3	2,3714	-0,1918	0,1918	$(-)(-)$
7	2,3714	3	2,3784	-0,1141	0,1141	$(-)(-)$
8	2,3784	3	2,3826	-0,0672	0,0672	$(-)(-)$
9	2,3826	3	2,3850	-0,0394	0,0394	$(-)(-)$
10	2,3850	3	2,3864	-0,0230	0,0230	$(-)(-)$
11	2,3864	3	2,3872	-0,0134	0,0134	$(-)(-)$
12	2,3872	3	2,3877	-0,0078	0,0078	$(-)(-)$



O critério de parada foi satisfeito na 12ª iteração. A solução aproximada é  $\xi \approx 2,3872$ .

### 18.1.1 Considerações Finais

A grande vantagem do Método da falsa posição é que ele é uma técnica robusta, que converge independentemente da forma do gráfico de  $y = f(x)$  no intervalo  $[a, b]$ .

Entretanto, quando a convergência para a raiz só se faz a partir de um extremo do intervalo  $[a, b]$  e a imagem desse ponto fixo tem um valor muito elevado, a convergência é lenta. Este fato pode ser verificado analisando-se mais cuidadosamente a expressão 18.1.

Admita que o ponto fixo seja  $b$ . Para mostrar a parcela de acréscimo dado ao extremo esquerdo  $a$ , que nesta situação é variável, adicione e subtraia a parcela  $a \times f(a)$  no numerador da expressão 18.1:

$$x = \frac{af(b) - bf(a) - af(a) + af(a)}{f(b) - f(a)} = \frac{a[f(b) - f(a)] - (b - a)f(a)}{f(b) - f(a)}.$$

Assim,

$$x = a - \frac{f(a)}{f(b) - f(a)}(b - a). \quad (18.2)$$

Sendo, por hipótese,  $b$  fixo e  $f(b)$  elevado, a expressão  $\frac{f(a)}{f(b) - f(a)}(b - a)$ , que representa o acréscimo, terá um valor pequeno, acarretando convergência tão mais lenta quanto maior for o valor de  $f(b)$ .

Quando se considera  $a$  como ponto fixo tem-se:

$$x = b - \frac{f(b)}{f(b) - f(a)}(b - a) \quad (18.3)$$

que é obtido somando e subtraindo a parcela  $b \times f(b)$  no numerador da expressão 18.1.

Uma forma de evitar que um extremo fique fixo durante o processo iterativo, situação que ocorre quando  $f(x_k) \times f(x_{k-1}) > 0$ , é substituir a reta que passa pelos pontos  $[a, f(a)]$  e  $[b, f(b)]$  por uma reta de inclinação menor. Por exemplo, se em duas iterações consecutivas o extremo  $b$  ficar fixo, substitui-se  $f(b)$  por  $f(b)/2$ .

## 18.2 Exercícios

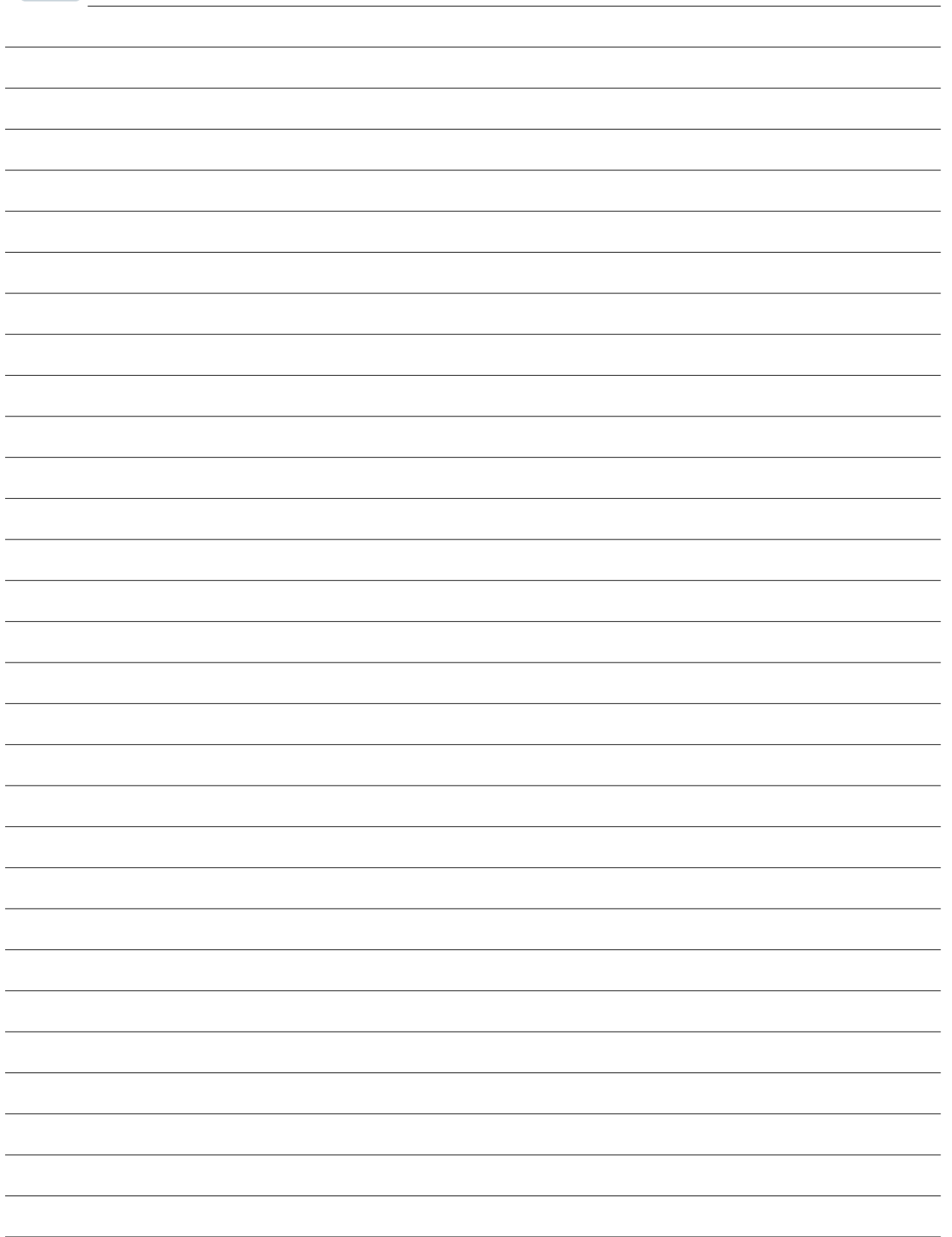


**E. 1. Para cada uma das equações seguintes, calcule pelo menos uma raiz com precisão de 0,001 pelo método da Falsa-Posição.**

a)  $f(x) = 2x^3 - 5x^2 - x + 3 = 0$

b)  $f(x) = e^x - x^2 + 4 = 0$







# 19

## Método de Newton Raphson

### Sumário da Aula

19.1 Método de Newton-Raphson . . . . .	199
19.1.1 Considerações Finais . . . . .	204
19.2 Exercícios . . . . .	204

### 19.1 Método de Newton-Raphson

Seja  $y = f(x)$  uma função contínua em um intervalo  $[a, b]$  que contém uma, e só uma, raiz,  $\xi$ , da equação  $f(x) = 0$  e no qual  $f'(x)$  e  $f''(x)$  preservam o sinal e não se anulam. O Método de Newton-Raphson consiste em:

- atribuir uma estimativa inicial  $x_0 \in [a, b]$  para uma raiz de  $f(x) = 0$ ;
- gerar uma sequência de estimativas,  $x_k + 1, k = 0, 1, 2, 3, \dots$ ; onde cada ponto é a interseção da reta tangente a  $y = f(x)$ , em  $[x_k, f(x_k)]$ , com o eixo das abscissas.



A Figura 19.1 ilustra o procedimento.

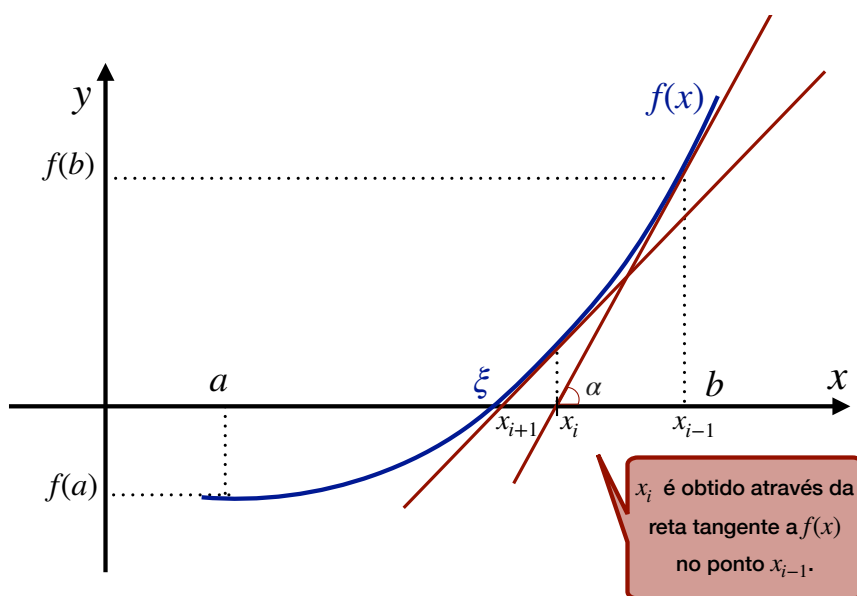


Figura 19.1: Representação gráfica do método de Newton-Rapshon.

## Função de iteração

Para determinar a função de iteração e obtenção de  $x_{i+1}$ , considere a Figura 19.2.

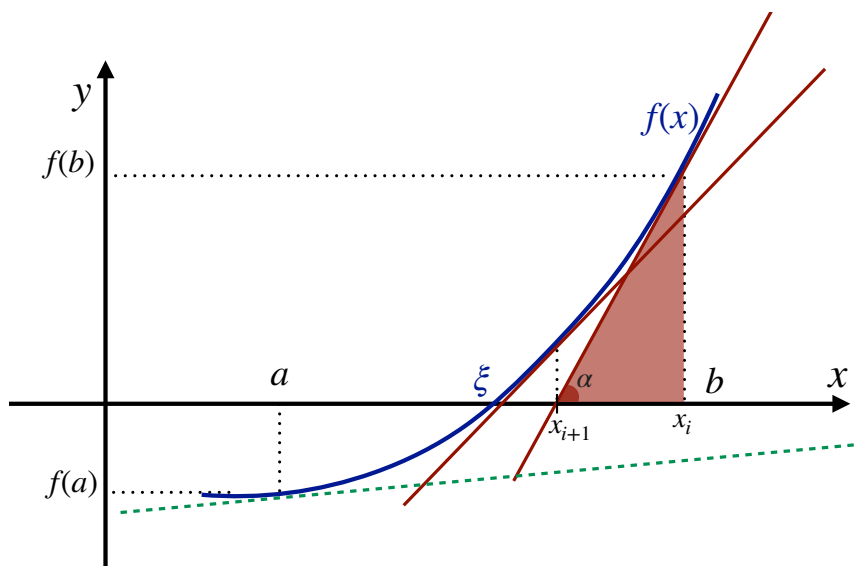


Figura 19.2: Obtenção da função de iteração do método de Newton-Rapshon.

O coeficiente angular da reta tangente à curva  $f(x)$  no ponto  $P(x_i, f(x_i))$ , dado por  $tg(\alpha) = f'(x_i)$ , também pode ser obtido utilizando as relações trigonométricas no triângulo retângulo:

$$tg(\alpha) = \frac{f(x_i)}{x_i - x_{i+1}} \Rightarrow f'(x_i) = \frac{f(x_i)}{x_i - x_{i+1}}.$$

Assim, considerando o resultado anterior, conclui-se que a função de iteração do método de Newton-Rapshon é dada pela Equação 19.1.

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \dots \quad (19.1)$$

## Critério de parada

Dada uma precisão  $\varepsilon$ , o processo iterativo é finalizado quando se obtém um ponto  $x_{k+1}$ ,  $k = 0, 1, \dots$ ; tal que  $|x_{k+1} - x_k| \leq \varepsilon$  ou  $|f(x_k)| \leq \varepsilon$  e, então, este ponto é tomado como uma estimativa para a raiz de  $f(x) = 0$ . Também é comum utilizar como critério de parada um número máximo de iterações.

## Critério de convergência

Note na 19.2, que a sequência produzida pela Equação 19.1 convergirá para a raiz  $\xi$  se o valor inicial for  $x_0 = b$ . Na mesma figura, para  $x_0 = a$ , o processo pode não convergir, pois  $x_1 \notin [a, b]$ . Assim, para a escolha do valor inicial de modo a garantir a convergência para a raiz deve ser considerado o Teorema 19.1.

**Teorema 19.1.** Se  $f(a)f(b) < 0$ , e  $f'(x)$  e  $f''(x)$  forem não nulas e preservarem o sinal em  $[a, b]$  tal que  $f(x_0)f''(x_0) > 0$  é possível construir, pelo método de Newton-Rapshon uma sequência  $x_i$  que convirja para a raiz  $\xi$  de  $f(x) = 0$

Por este teorema, temos que o valor inicial  $x_0$  deve ser um ponto no qual a função tenha o mesmo sinal de sua derivada de segunda ordem, ou seja, se  $f''(x_0) > 0$ , então  $x_0$  é tal que  $f(x_0) > 0$ , e por outro lado, se  $f''(x_0) < 0$ , então  $f(x_0) < 0$ .

## Pseudocódigo

A sequência de aproximações  $x_0, x_1, \dots, x_k$  para a raiz  $\xi$  da equação  $f(x) = 0$  é obtida considerando o Pseudocódigo do método da Newton-Rapshon descrito no Algoritmo 8.

```

1: Entrada:  $x_0, f(\cdot), f'(\cdot), \varepsilon$  e  $k_{max}$ .
2:  $erro = \infty$ ;
3:  $x = x_0$ ;
4:  $i = 1$ ;
5: enquanto ( $erro \geq \varepsilon$ ) e ( $i \leq k_{max}$ ) faça
6:    $fx = f(x)$ ;
7:    $Dfx = f'(x)$ ;
8:    $x_{atual} = x - \frac{fx}{Dfx}$ ;
9:    $erro = |x_{atual} - x|$ ;
10:   $x = x_{atual}$ ;
11:   $i = i + 1$ ;
12: fim enquanto
13: Saída:  $x_i$ .

```

Algoritmo 8: Método de Newton

### observação

O erro também pode ser calculado usando  $|f(x)|$ .

**Exemplo 65.** Seja a equação  $f(x) = x^3 - 9x^2 + 3 = 0$ . Estime a raiz de  $f(x) = 0$  contida no intervalo  $(0, 1)$  considerando um precisão  $\varepsilon = 0,065$  e o método iterativo de Newton-Raphson.

**Solução:**

- Dado o intervalo inicial  $[a_0, b_0] = [0, 1]$ , existe um número ímpar de raízes neste intervalo, pois  $f(0)f(1) = (3)(-5) < 0$ .
- Como  $f'(x) = 3x^2 - 18x < 0$  em todo o intervalo dado, a raiz  $\xi \in [0, 1]$  é única.
- Para que a convergência seja garantida é necessário verificar também o sinal da derivada de segunda ordem. Veja que a  $f''(x) = 6x - 18 < 0$  em todo o intervalo  $[0, 1]$ .
- Para a escolha do  $x_0$ , deve-se analisar qual dos extremos do intervalo  $[0, 1]$  satisfaz a condição  $f(x_0)f''(x_0) > 0$ . Considerando o critério de convergência, segue que  $x_0 = 1$ .
- O cálculo da sequência de raízes aproximadas está apresentado na tabela a seguir:

$k$	$x_k$	$f(x_k)$	$f'(x_k)$	$ x_k - x_{k-1} $
0	1	-5	-15	— — —
1	0,6667	-0,7041	-10,6671	0,3333
2	0,6007	-0,0308	-9,7300	0,0660
3	0,5975			0,003

Como o critério de parada foi satisfeito, ou seja,  $|x_k - x_{k-1}| \leq \varepsilon$ , a solução aproximada é  $\xi \approx 0,5975$ .

**Exemplo 66.** Um objeto de massa  $m$  é solto de uma altura  $S_0$  em relação ao solo. Após  $t$  segundos a sua altura é dada pela expressão:

$$S(t) = S_0 - \frac{mg}{k}t + \frac{m^2g}{k^2} \left(1 - e^{-\frac{kt}{m}}\right) \quad (19.2)$$

onde  $k$  é o coeficiente de resistência do ar e  $g$  a aceleração da gravidade. Sendo  $m = 1kg$ ,  $S_0 = 30m$ ,  $k = 0,5kg/s$  e  $g = 9,8m/s^2$ , estime o tempo que o objeto leva para chegar ao solo utilizando o método de Newton-Raphson, com precisão 0,001 e um máximo de 5 iterações.

**Solução:**

- Resolver este problema consiste em determinar o tempo  $t$  para o qual  $S(t) = 0$ . Efetuando as substituições na Equação 19.2 tem-se:

$$S(t) = 30 - \frac{9,8}{0,5}t + \frac{9,8}{(0,5)^2} (1 - e^{-0,5t})$$

que, fazendo as simplificações adequadas, resulta na seguinte equação a ser resolvida:

$$S(t) = 69,2 - 19,6t - 39,2e^{-0,5t} = 0.$$

- Uma vizinhança para a(s) raiz(es) será isolada graficamente. Assim, transformando  $S(t) = 0$  em  $g(t) = h(t)$ , onde  $g(t) = 69,2 - 19,6t$  e  $h(t) = 39,2e^{-0,5t}$  obtém-se o esboço para os gráficos de  $g(t)$  e  $h(t)$  (ver Figura 19.3):

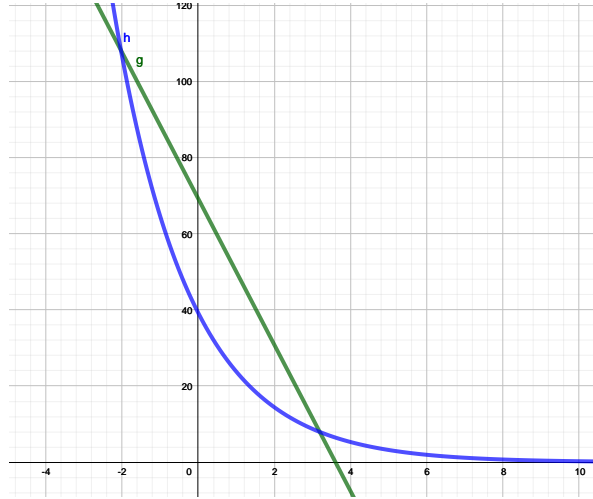


Figura 19.3: O esboço com os gráficos das funções  $g(t) = 69,2 - 19,6t$  e  $h(t) = 39,2e^{-0,5t}$ .

- Observando a Figura 19.3, nota-se que a função  $S(t)$  possui duas raízes. Como para o problema a raiz negativa não tem sentido, já que  $t$  representa o tempo, ela não será considerada. Assim, podemos concluir que existe uma raiz  $\xi$  que pertence ao intervalo  $[3, 4]$ . Esta raiz encontra-se na interseção dos gráficos de  $g(t)$  e  $h(t)$ .
- No intervalo  $[3, 4]$  a raiz é única, conforme a aplicação do Teorema de Cauchy-Bolzano:
  - $S(3)S(4) = (1,653)(-14,505) < 0$ , portanto existe pelo menos uma raiz neste intervalo.
  - $S'(t) = 19,6(e^{-0,5t} - 1) < 0$  em todo o intervalo  $[3, 4]$ . Logo a raiz  $\xi \in [3, 4]$  é única.
- Para que a convergência seja garantida é necessário verificar também o sinal da derivada de segunda ordem. Assim:

$$S''(t) = -9,8(e^{-0,5t}) < 0$$

em todo o intervalo  $[3, 4]$ .

- Para a escolha do  $t_0$ , deve-se analisar qual dos extremos do intervalo  $[3, 4]$  satisfaz a condição  $f(x_0)f''(x_0) > 0$ . Considerando o critério de convergência,  $t_0 = 4$ .
- O cálculo da sequência de raízes aproximadas aplicando o método de Newton-Rapshon está apresentado na tabela a seguir:

$k$	$t_k$	$S(t_k)$	$S'(t_k)$	$ t_k - t_{k-1} $
0	4	-14,505	-16,947	— — —
1	3,144	-0,561	-15,530	0,856
2	3,108	-0,004	-15,457	0,000
3	3,1088			0

Como o critério de parada foi satisfeito, ou seja,  $|t_3 - t_2| \leq \varepsilon$ , a estimativa aproximada para o objeto atingir o solo é de  $\bar{t} \approx 3,108$  segundos.



### 19.1.1 Considerações Finais

O Método de Newton-Raphson tem convergência muito boa (quadrática) o que, por consequência, proporciona um número pequeno de iterações. Entretanto, o método apresenta as seguintes desvantagens:

- (i) Exige a análise do sinal de  $f'(x)$  e  $f''(x)$ .
- (ii) Exige o cálculo do valor da primeira derivada em cada iteração.
- (iii) Se  $f'(x_k)$  for muito elevada a convergência será lenta.
- (iv) Se  $f'(x_k)$  for próximo de zero pode ocorrer *overflow*.

Para contornar o item (i), que é necessário para a escolha da estimativa inicial, é comum calcular somente o valor da função e o da sua segunda derivada nos extremos  $a$  e  $b$ , tomando como  $x_0$  o ponto que satisfizer a condição  $f(x_0)f''(x_0) > 0$ .

## 19.2 Exercícios



**E. 1. Para cada uma das equações seguintes, calcule pelo menos uma raiz com precisão de 0,001 pelo método de Newton-Raphson.**

a)  $f(x) = 2x^3 - 5x^2 - x + 3 = 0$

b)  $f(x) = e^x - x^2 + 4 = 0$



This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.



**Parte VII**

**Atividades Avaliativas**



# Referências Bibliográficas

- 1 CAMPOS filho, F. F. *Algoritmos Numéricos: uma abordagem moderna de Cálculo Numérico*. 3ª. ed. Rio de Janeiro: LTC, 2018. v. 1.
- 2 GILAT, A.; SUBRAMANIAM, V. *Métodos Numéricos para Engenheiros e Cientistas: Uma introdução com aplicações usando o MATLAB*. Porto Alegre: Bookman, 2008.
- 3 FRANCO, N. M. B. *Cálculo Numérico*. 1ª. ed. São Paulo: Pearson Prentice Hall, 2006. v. 1.
- 4 SPERANDIO, D.; MENDES, J. T.; SILVA, L. H. M. e. *Cálculo Numérico*. 2ª. ed. São Paulo: Pearson, 2014. v. 1.
- 5 CAMPOS filho, F. F. *Algoritmos Numéricos*. 2ª. ed. Rio de Janeiro: LTC, 2018. v. 1.
- 6 ANTON, H.; RORRES, C. *Algebra Linear: com aplicações*. 10ª. ed. Porto Alegre: Bookman, 2012. v. 1.
- 7 ARENALES, S.; DAREZZO, A. *Cálculo Numérico: Aprendizagem com apoio de software*. 2ª. ed. São Paulo: Cengage Learning, 2016. v. 1.
- 8 GUIDORIZZI, H. L. *Um curso de Cálculo*. 5ª. ed. Rio de Janeiro: LTC, 2001. v. 1.