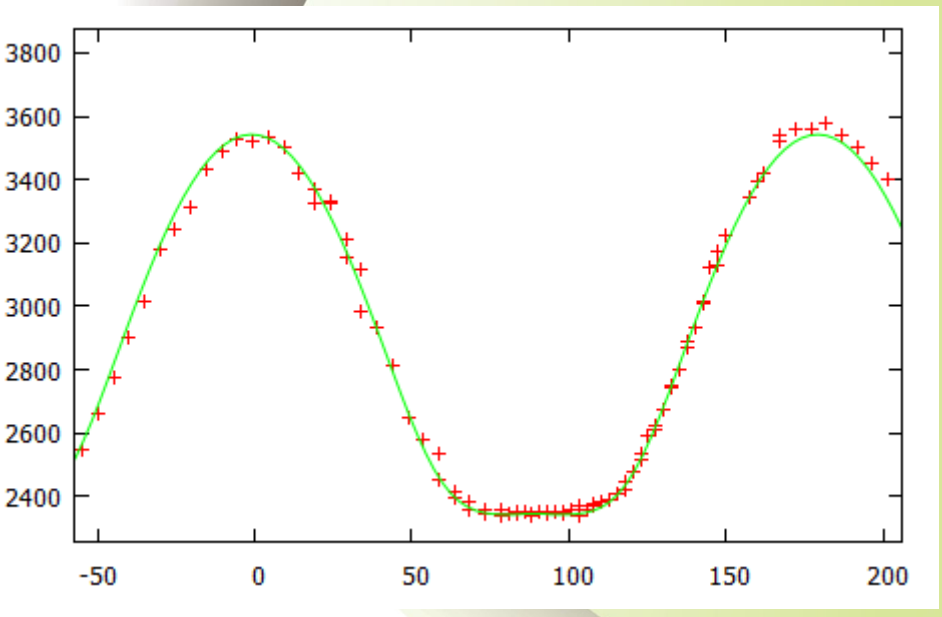
MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA CATARINENSE
CAMPUS LUZERNA

Engenharia de controle e Automação

Cálculo Numérico



Prof: Ricardo Kerschbaumer

ALUNO: _____

Luzerna, Setembro de 2023

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA CATARINENSE

Cálculo Numérico

Este material foi desenvolvido pelo professor Ricardo Kerschbaumer para ser utilizado na disciplina de Cálculo Numérico do curso de Engenharia de Controle e Automação. O objetivo deste documento é servir como material de apoio as aulas, bem como de material de estudo para que os alunos possam tirar suas dúvidas nos períodos extra classe.

Grande parte do conteúdo deste material foi retirado das referências bibliográficas e seus direitos pertencem aos seus autores.

Sumário

1.LISTA DE FIGURAS

2.LISTA DE TABELAS

AULA 1 - MATEMÁTICA COMPUTACIONAL

1.1 Objetivo:

O objetivo desta primeira aula é apresentar aos alunos os principais conceitos sobre a matemática computacional, mostrando assim como os computadores podem ser usados para resolver problemas matemáticos. Também será abordada a aritmética de ponto flutuante, pois para que se possa entender de onde vem os erros é necessário entender como o computador armazena os números na memória.

1.2 Natureza da matemática numérica

A matemática numérica teve seu início com o advento das máquinas mecânicas de calcular. Com o surgimento dos computadores o interesse pela matemática numérica aumentou muito, principalmente para aplicações militares, científicas e tecnológicas. A matemática numérica surgiu principalmente porque os métodos matemáticos utilizados com os números reais não são válidos nos computadores. Isso acontece porque o computador tem uma precisão finita, números como por exemplo o $\pi = 3.1415...$ que na natureza tem um número infinito de casas, no computador são truncados.

Assim a matemática numérica ou matemática computacional é a área da matemática que se preocupa com o desenvolvimento, emprego e estudo de métodos numéricos, podendo ser dividida em:

Matemática Computacional: estudo da matemática do ponto de vista computacional.

Matemática Numérica: parte da matemática computacional que se preocupa com o desenvolvimento de algoritmos para a resolução aproximada de problemas utilizando os quatro operadores básicos.

Matemática Simbólica: Busca a solução analítica de um problema matemático, por exemplo, a solução analítica da integral:

$$\int x^2 \cdot dx = \frac{x^3}{3} \tag{1}$$

Matemática Gráfica: Trabalha com modelos gráficos buscando soluções na forma gráfica.

Matemática Intervalar: trata dados na forma de intervalos, buscando controlar os limites de erro da matemática numérica.

1.3 Algoritmos

Para que possamos resolver um problema através de um computador é necessário ter toda a sequência de passos e operações estabelecida de modo formal para a resolução do problema. Pois apenas desta forma o computador será capaz de realizar a tarefa. Para realizar esta sequência de

passos é que necessitamos de um algoritmo. Um algoritmo é uma sequência ordenada de instruções que leva a solução de um problema específico. Um algoritmo deve ter um número finito de instruções e deve ser executado com uma quantidade limitada de esforço. Um algoritmo pode ou não depender de dados de entrada, mas deve necessariamente fornecer dados de saída. Além disso deve ter um único ponto de início e, pelo menos, um ponto de parada.

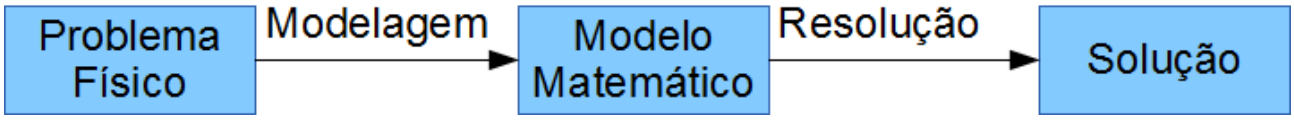
Os algoritmos numéricos são uma categoria de algoritmos voltados ao processamento numérico. Um algoritmo numérico de boa qualidade deve ter as seguintes características:

- Inexistência de erro lógico
- Inexistência de erro operacional
- Quantidade finita de cálculos
- Existência de um critério de exatidão
- Independência de máquina
- Com precisão finita os limites de erro devem convergir a zero
- Eficiência

1.4 Modelagem matemática

Um modelo matemático, de forma geral, pode ser definido como uma formulação ou equação que expressa as características essenciais de um sistema ou processo físico em termos matemáticos. Este processo é ilustrado pela 1.

Figura 1: Processo de elaboração de um modelo matemático



No processo de modelagem matemática de um problema físico são executadas as seguintes etapas:

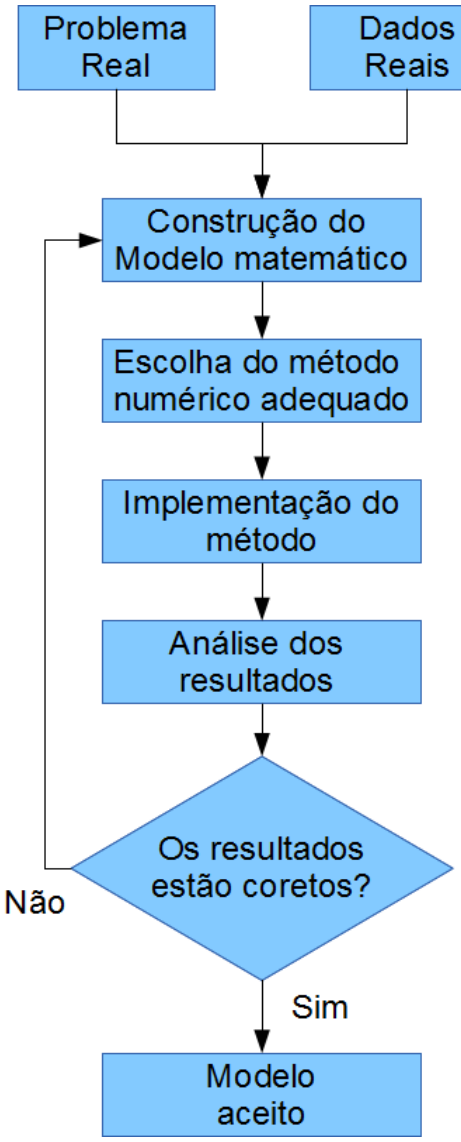
Modelagem: Fase de obtenção de um modelo matemático que descreve o comportamento do problema que se quer estudar.

Resolução: Fase de obtenção da solução do modelo matemático através da aplicação de métodos numéricos.

Levando em consideração que ambas as etapas estão susceptíveis a erros é necessário um processo de refinamento dos modelos obtidos e dos métodos numéricos aplicados a estes modelos de forma a tornar a solução o mais precisa possível, mantendo os erros dentro de uma faixa tolerável.

A 2 apresenta um diagrama deste processo de forma mais detalhada, enfatizando a necessidade do refinamento do modelo e dos métodos numéricos aplicados.

Figura 2: Fluxograma de elaboração de um modelo matemático



Para que um modelo seja considerado válido é necessário que ele seja testado e seus resultados comparados com os resultados do problema real. Se os resultados obtidos do modelo estiverem dentro de uma margem de erro aceitável o modelo é considerado válido.

1.5 Números em ponto flutuante

Como já foi visto no decorrer do curso a representação de números inteiros é feita no computador utilizando-se a representação em complemento de dois. Porem esta notação não é válida para números reais, onde a parte fracionária é diferente de zero.

A representação de números reais no computador pode ser feita de várias formas diferentes, dentre elas podemos citar a representação em ponto fixo e a representação em ponto flutuante. A representação em ponto fixo é normalmente utilizada em microcontroladores e processadores

digitais de sinal, pois exige muito menos poder de processamento. O lado negativa da representação de números em ponto fixo é a baixa precisão, pois poucos dígitos significativos são utilizados.

A forma mais utilizada de representação de números reais em sistemas computadorizados é a representação em ponto flutuante. Existem várias formas de representar números em ponto flutuante. Para que aja compatibilidade entre os vários fabricantes de equipamentos e softwares foi criada uma padronização para a representação destes números. O padrão mais adotado atualmente é o IEEE-754, que determina a arquitetura de números de precisão simples com 32 bits e números de dupla precisão com 64 bits.

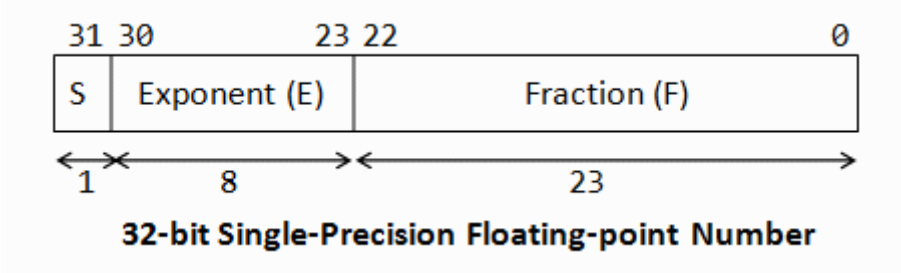
1.5.1 Representação de números de ponto flutuante de precisão simples

Estes números utilizam 32 bits para armazenar os valores reais. Estes 32 bits são definidos da seguinte forma:

- O bit mais significativo representa o sinal (S), utilizando 1 para números negativos e 0 para números positivos.
- Os próximos 8 bits representam o expoente (E).
- Os últimos 23 bits representam a parte fracionária (F) ou mantissa.

A 3 mostra esta representação.

Figura 3: Ponto flutuante de precisão simples



Forma normalizada

Analisando o exemplo a seguir podemos entender melhor este sistema.

Veja o seguinte padrão de 32 bits: 1 1000 0001 011 0000 0000 0000 0000 0000

Onde:

- S = 1
- E = 1000 0001
- F = 011 0000 0000 0000 0000 0000

Na forma normalizada a parte fracionária recebe um “1” na forma 1.F, assim a parte fracionária fica: 1,011 0000 0000 0000 0000 0000 ou 1, $0\times2^{-1}+1\times2^{-2}+1\times2^{-3}+0\times2^{-4}...$ = 1,375 em

decimal.

O bit de sinal S = 1 indica um número negativo, assim temos -1,375 em decimal.

Na forma normalizada o expoente é representado por E-127, isso é necessário para que se possa representar expoentes negativos e positivos. Como temos 8 bits para o expoente é possível representar números entre -127 e 128. Em nosso exemplo temos E – 127 ou 129 – 127 = 2 em decimal. Assim o número fica $-1{,}375 \times 2^2$ = -5,5 em decimal.

Resumindo temos:

$$N = -1^S \times 1, F \times 2^{E-127} \quad (2)$$

A forma normalizada tem um sério problema, com o "1" utilizado na parte fracionária é impossível representar o número 0.

Forma de-normalizada

A forma de-normalizada é utilizada para representar o número 0 e alguns outros números. Quando E = 0, o número está na forma de-normalizada, um 0 é utilizado no lugar do 1 que precedia a parte fracionária e o expoente é sempre -126. Assim o número 0 pode ser representado com E = 0 e F = 0, pois $0{,}0 \times 2^{-126} = 0$.

Também é possível representar números muito pequenos positivos e negativos na forma de-normalizada, com E = 0. Por exemplo, se S = 1, E = 0 e F = 011 0000 0000 0000 0000. A parte fracionária fica 0, $0 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3} + 0 \times 2^{-4} ...$ = 0,375 em decimal. Como S = 1 o número fica negativo e o expoente fica -126 pois E = 0. Assim temos $-0{,}375 \times 2^{-126} = -4{,}4 \times 10^{-39}$ em decimal, o que é um número muito pequeno.

Resumindo temos:

$$N = -1^S \times 0, F \times 2^{-126} \quad (3)$$

A seguir temos alguns exemplos:

Exemplo 1: Suponha o seguinte número representado em ponto flutuante segundo padrão IEEE-754 0 10000000 110 0000 0000 0000 0000 0000.

S = 0 ⇒ número positivo
E = 1000 0000B = 128D (forma normalizada)
F = 1,11B = $1 + 1 \times 2^{-1} + 1 \times 2^{-2}$ = 1,75D
o número é $+1.75 \times 2^{(128-127)}$ = +3,5D

Exemplo 2: Suponha o seguinte número representado em ponto flutuante segundo padrão IEEE-754 1 01111110 100 0000 0000 0000 0000 0000.

S = 1 ⇒ número negativo
E = 0111 1110B = 126D (forma normalizada)

F = 1,1B = $1+1\times 2^{-1}$ = 1,5D
o número é $-1.5\times 2^{(126-127)}$ = -0,75D

Exemplo 3: Suponha o seguinte número representado em ponto flutuante segundo padrão IEEE-754

1 01111110 000 0000 0000 0000 0000 0001.

S = 1 ⇒ número negativo
E = 0111 1110B = 126D (forma normalizada)
F = 1,000 0000 0000 0000 0000 0001B = $1+1\times 2^{-23}$
o número é $-(2^{-23})\times 2^{(126-127)}$ = -0.5000000596046447753906 25D a representação em decimal não é exata.

Exemplo 4: Suponha o seguinte número representado em ponto flutuante segundo padrão IEEE-754

1 00000000 000 0000 0000 0000 0000 0001.

S = 1 ⇒ número positivo
E = 0 (forma de-normalizada)
F = 0,000 0000 0000 0000 0000 0001B = 1×2^{-23}
o número é $-(2^{-23})\times 2^{(-126)}\approx -1{,}4\times 10^{-45}$ em decimal.

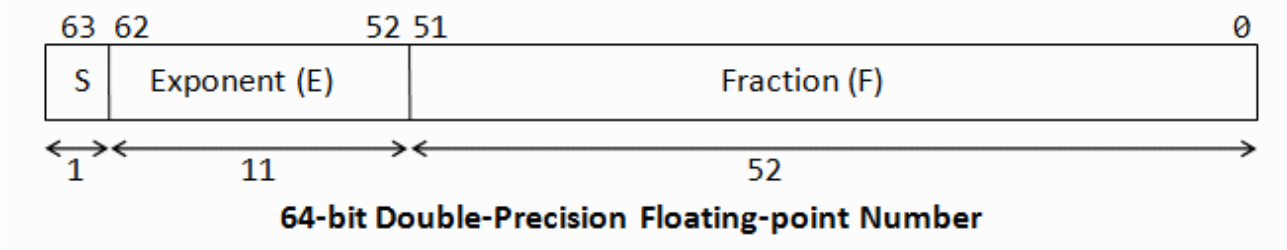
1.5.2 Representação de números de ponto flutuante de precisão dupla

Estes números utilizam 64 bits para armazenar os valores reais. Estes 64 bits são definidos da seguinte forma:

- O bit mais significativo representa o sinal (S), utilizando 0 para números negativos e 1 para números positivos.
- Os próximos 11 bits representam o expoente (E).
- Os últimos 52 bits representam a parte fracionária (F) ou mantissa.

A 4 mostra esta representação.

Figura 4: Ponto flutuante de precisão dupla



Não entraremos em detalhes sobre a representação de números de precisão dupla, mas as regras

são as mesmas utilizadas nos números de precisão simples.

1.5.3 Comparação entre números de precisão simples e dupla

A Erro: Origem da referência não encontrada faz um comparativo entre as duas representações.

Tabela 1: Comparativo entre precisão simples e dupla

Precisão	Normalizado N(min)	Normalizado N(max)
Simples	0080 0000H 0 00000001 00000000000000000000000B E = 1, F = 0 $N_{min} = 1{,}0\,B \times 2^{-126}$ $(\approx 1{,}17549435 \times 10^{-38})$	7F7F FFFFH 0 11111110 11111111111111111111111B E = 254 $N_{max} = 1{,}...1\,B \times 2^{127}$ $(\approx 3{,}4028235 \times 10^{38})$
Dupla	0010 0000 0000 0000H $N_{min} = 1{,}0\,B \times 2^{-1022}$ $(\approx 2{,}2250738585072014 \times 10^{-308})$	7FEF FFFF FFFF FFFFH $N_{max} = 1{,}...1\,B \times 2^{1023}$ $(\approx 1{,}7976931348623157 \times 10^{308})$

Outra característica que deve ser levada em conta é o esforço computacional necessário para realizar as operações com números de ponto flutuante. Devido à complexidade destes números o computador demora muito mais para operar um número em ponto flutuante se comparado a um número inteiro. Da mesma forma números de precisão dupla exigem muito mais processamento do que números de precisão simples.

1.5.4 Valores especiais

Zero: Não pode ser representado na forma normalizada e é representado na forma de-normalizada com E = 0 e F = 0.

Infinito: O valor de mais infinito ou menos infinito é representado por E = 255 em 32 bits ou E = 2047 em 64 bits, F=0 S = 0 ou 1.

Not a Number (NaN): Denota números que não podem ser representados na forma real como por exemplo 0 / 0, e são representados por E = 255 em 32 bits ou E = 2047 em 64 bits e F $\neq$ 0.

AULA 2 - ERROS

2.1 Objetivo:

O objetivo desta aula é mostrar aos alunos que nenhum resultado obtido através de cálculos eletrônicos ou métodos numéricos tem valor se não tivermos conhecimento e controle sobre os possíveis erros envolvidos no processo.

A análise dos resultados obtidos através de um método numérico representa uma etapa fundamental no processo de obtenção das soluções numéricas.

2.2 Conceito de exatidão e Conceito de precisão

2.2.1 Conceito de exatidão

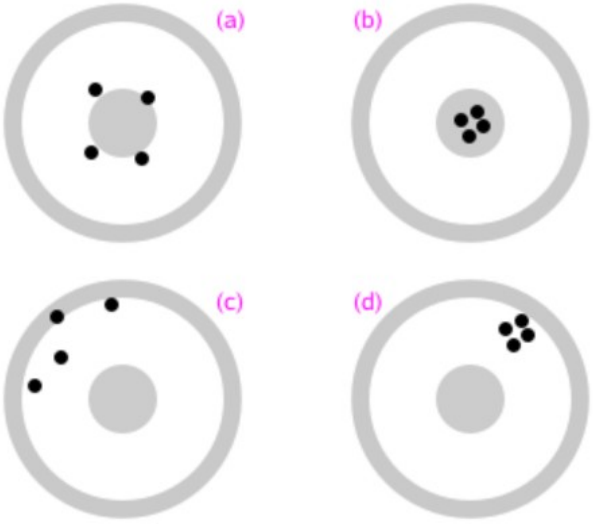
A exatidão é definida pelo desvio máximo, em relação ao valor verdadeiro x, de uma série de medidas, portanto, a expressão os resultados da medição são exatos significa resultados próximos do valor verdadeiro.

2.2.2 Conceito de precisão

Precisão é uma medida da reprodutibilidade, geralmente expressa pela maior diferença entre os valores medidos e a média desses valores, portanto, a expressão os resultados da medição são precisos significa resultados próximos entre si.

A 5 mostra este conceito aplicado a tiros disparados contra um alvo.

Figura 5: Precisão e exatidão



No alvo (a) os tiros são exatos mas não são precisos, no alvo (b) os tiros são exatos e precisos, no alvo (c) os tiros não são nem exatos e nem precisos e no alvo (d) os tiros são precisos mas não

são exatos.

Embora sejam sinônimos na linguagem do dia a dia, exatidão e precisão são, conforme visto, coisas distintas no conceito técnico.

2.3 Número Aproximado

Um número $\bar{x}$ é dito uma aproximação para o número exato x se existe uma pequena diferença entre eles. Geralmente, nos cálculos os números exatos não são conhecidos e deste modo são substituídos por suas aproximações. Dizemos que $\bar{x}$ é um número aproximado por falta do valor exato x se $\bar{x} < x$. Se $\bar{x} > x$ temos uma aproximação por excesso.

Exemplo:

Como $1{,}41 < \sqrt{2} < 1{,}42$ temos que 1,41 é uma aproximação de $\sqrt{2}$ por falta e 1,42 é uma aproximação de $\sqrt{2}$ por excesso.

2.4 Erros Absolutos e Relativos

2.4.1 Erro Absoluto

A diferença entre um valor exato x e sua aproximação $\bar{x}$ é dita **erro absoluto** o qual denotamos por e_x

$$e_x = x - \bar{x} \tag{4}$$

2.4.2 Cota para o Erro

Na prática, o valor exato é quase sempre não conhecido. Como o erro é definido por e $e_x = x - \bar{x}$ consequentemente também será não conhecido.

Uma solução para este problema é, ao contrário de determinar o erro, determinar uma cota para o erro. Isso permitirá que, mesmo não conhecendo o erro, saber que ele está entre dois valores conhecidos. Dizemos que um número $e > 0$ é uma cota para o erro e_x se $|e_x| > e$

$$\therefore \quad |e_x| > e \Leftrightarrow |x - \bar{x}| < e \Leftrightarrow \bar{x} - e < x < \bar{x} + e$$

Assim, mesmo não conhecendo o valor exato, podemos afirmar que ele está entre $\bar{x} - e$ e $\bar{x} + e$ que são valores conhecidos. É evidente que uma cota e só tem algum valor prático se $e \approx 0$.

2.5 Erro Relativo

Considere: x = 100; $\bar{x}$ = 100.1 e y = 0.0006; $\bar{y}$ = 0.0004.
Assim e_x = 0.1 e e_y = 0.0002.

Como $|e_y|$ é muito menor que $|e_x|$ poderíamos “imaginar” que a aproximação $\bar{y}$ de y é melhor que a $\bar{x}$ de x. Numa análise mais cuidadosa percebemos que as grandezas dos números

envolvidos são muito diferentes.

Inspirados nessa observação definimos:

$$E_x = \frac{e_x}{|x|} \tag{5}$$

que é denominado erro relativo. Temos então para os dados acima:

$$E_x = \frac{e_x}{|x|} = \frac{0{,}1}{100} = 0{,}001$$

$$E_y = \frac{e_y}{|y|} = \frac{0{,}0002}{0{,}0006} = 0{,}333333$$

Agora podemos concluir que a aproximação $\bar{x}$ de x é melhor que a aproximação $\bar{y}$ de y, pois $|E_x| < |E_y|$

2.6 Fontes de Erros

2.6.1 Erros Inerentes

São os erros que existem nos dados e são causados por erros inerentes aos equipamentos utilizados na captação dos dados.

2.6.2 Erros de Truncamento

São os erros causados quando num processo é necessário um algorítmico infinito mas somos obrigados a usar apenas uma parte finita deste algoritmo.
Exemplo:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots = \sum_{n=1}^{\infty} \frac{x^n}{n!}$$

Podemos assim usar: $1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}$ para uma aproximação de e^x Observe que para isso truncamos uma série infinita utilizando apenas uma parte finita dela. No exemplo utilizamos para a aproximação apenas quatro termos da série.

Usando a aproximação acima para de terminar a constante *e* fazendo *x = 1* temos:

$\bar{e} = 1 + 1 + \frac{1}{3} + \frac{1}{6} = 2.66666$ que é uma aproximação muito pobre para *e*.

2.6.3 Erros de Arredondamento

Erros de Arredondamento são os erros originados pela representação dos números reais utilizando-se apenas um número finito de casas decimais. Como se sabe, desde a mais simples calculadora até o mais potente computador, utiliza apenas um número finito de casas decimais para

representar um número real. (o número real é denominado número de ponto flutuante nas linguagens de programação). Dizemos então que os equipamentos eletrônicos utilizam nos cálculos a chamada aritmética finita.

2.7 Propagação dos Erros

2.7.1 Propagação dos Erros Absolutos

Seja $\bar{x}$ uma aproximação para x e $\bar{y}$ uma aproximação para y, ou seja, $e_x = x - \bar{x}$ e $e_y = y - \bar{y}$ Então temos:

Soma:

$$e_{x+y} = e_x + e_y \tag{6}$$

Demonstração:

$$e_{x+y} = (x+y) - (\bar{x}+\bar{y}) = (x-\bar{x}) + (y-\bar{y}) = e_x + e_y$$

Subtração:

$$e_{x-y} = e_x - e_y \tag{7}$$

Demonstração:

$$e_{x-y} = (x-y) - (\bar{x}-\bar{y}) = x - \bar{x} - (y-\bar{y}) = e_x - e_y$$

Multiplicação:

$$e_{x\cdot y} = \bar{x}\, e_y + \bar{y}\, e_x \tag{8}$$

Demonstração:

$$xy = (\bar{x}+e_x)(\bar{y}+e_y) = \bar{x}\,\bar{y} + \bar{y}\, e_x + \bar{x}\, e_y + e_x e_y$$

Como e_x e e_y são supostamente pequenos, o produto $e_x e_y$ torna-se desprezível com relação aos outros termos, e assim podemos escrever:

$$xy = \bar{x}\,\bar{y} \approx \bar{x}\, e_y + \bar{y}\, e_x \Rightarrow e_{x\cdot y} = (x\,y) - (\bar{x}\,\bar{y}) \approx \bar{x}\, e_y + \bar{y}\, e_x$$

Divisão:

$$e_{\left(\frac{x}{y}\right)} = \frac{e_x}{\bar{y}} - \frac{\bar{x}}{(\bar{y})^2} e_y \tag{9}$$

Demonstração:

Como $x = \bar{x} + e_x$ e $y = \bar{y} + e_y$ temos:

$$\frac{x}{y} = \frac{(\bar{x}+e_x)}{(\bar{y}+e_y)} = (\bar{x}+e_x)\frac{1}{\bar{y}\left(1+\frac{e_y}{\bar{y}}\right)}$$

mas $\forall a \in R$ com $|a| < 1$ vale a igualdade

$$\frac{1}{1-a}=1+a+a^2+...+a^n...(série\ geométrica)$$

e como $\frac{e_y}{\bar{y}}$ é próximo a zero, ou seja $|\frac{e_y}{\bar{y}}|<1$ podemos fazer $a=\frac{-e_y}{\bar{y}}$ na igualdade acima e temos

$$\frac{1}{1+\frac{e_y}{\bar{y}}}=1-\frac{e_y}{\bar{y}}+(\frac{e_y}{\bar{y}})^2-(\frac{e_y}{\bar{y}})^3+...$$

assim temos

$$\frac{1}{1+\frac{e_y}{\bar{y}}}\approx 1-\frac{e_y}{\bar{y}}$$

pois como $\frac{e_y}{\bar{y}}$ os fatores é pequeno os fatores $(\frac{e_y}{\bar{y}})^2, (\frac{e_y}{\bar{y}})^3$... são desprezíveis.

substituindo esta aproximação na equação acima temos

$$\frac{x}{y}\approx\frac{(\bar{x}+e_x)}{\bar{y}}(1-\frac{e_y}{\bar{y}})=\frac{\bar{x}\,\bar{y}+\bar{y}\,e_x-\bar{x}\,e_y-e_x\,e_y}{(\bar{y})^2}\Rightarrow\frac{x}{y}-\frac{\bar{x}}{\bar{y}}\approx\frac{e_x}{\bar{y}}-e_y\frac{\bar{x}}{(\bar{y})^2}$$

pois $e_x e_y\approx 0$ assim é bastante razoável considerar

$$e_{(\frac{x}{y})}=\frac{e_x}{\bar{y}}-\frac{\bar{x}}{(\bar{y})^2}e_y$$

2.8 Propagação de erros

2.8.1 Soma e Subtração

Na soma e na subtração a propagação de erros se da segundo a seguinte fórmula.

$$E_{(x\pm y)}=\frac{\bar{x}}{\bar{x}\pm\bar{y}}E_x\pm\frac{\bar{y}}{\bar{x}\pm\bar{y}} \quad (10)$$

2.8.2 Multiplicação

Na multiplicação a propagação de erros se da segundo a seguinte fórmula.

$$E_{x\,y}=E_x+E_y \quad (11)$$

2.8.3 Divisão

Na divisão a propagação de erros se da segundo a seguinte fórmula.

$$E_{\frac{x}{y}}=E_x-E_y \quad (12)$$

2.9 Exercícios

1) Elabore um algoritmo que calcula a constante de Euler com 1, 2, …, 20 termos e plota em um gráfico o erro relativo e absoluto para cada uma das aproximações. A constante de Euler é calculada segundo a fórmula $e = \sum_{n=0}^{\infty} \frac{1}{n!}$ e seu valor real é 2,7182818284590452...

AULA 3 - RESOLUÇÃO DE SISTEMAS LINEARES - GAUSS

3.1 Objetivo:

Esta aula tem como objetivo mostrar aos alunos como resolver sistemas de equações lineares utilizando o método de Gauss, bem como fazer a retro substituição. Serão discutidos os métodos e apresentados os algoritmos, possibilitando assim a implementação e teste destes métodos no computador.

3.2 Introdução

Os métodos apresentados nesta aula tem o objetivo de resolver sistemas de equações lineares, onde as equações estão dispostas em forma de matriz, como a seguir.

$$\begin{Bmatrix} a_{11}x_1 & +a_{12}x_2 & +... & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +... & +a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +... & +a_{nn}x_n & = & b_n \end{Bmatrix} \tag{13}$$

ou, equivalentemente:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \quad \forall\, i = 1,2,...,n \tag{14}$$

Estes métodos se aplicam a sistemas onde o número de equações é igual ao número de incógnitas.

Na forma matricial, um sistema linear é representado por $Ax = b$ onde:

$$A = \begin{bmatrix} a_{11} & a_{12} & ... & A_{1n} \\ a_{21} & a_{22} & ... & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & ... & A_{nn} \end{bmatrix}$$ é a matriz dos coeficientes

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$ é o vetor das variáveis e

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$ é o vetor dos termos independentes.

É comum também apresentar o sistema $Ax=b$ pela sua matriz aumentada, isto é, por:

$$[A|b]=\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array}\right]$$ Matriz aumentada do sistema.

O vetor solução de um sistema $Ax=b$ é denotado como $\bar{x}=[\bar{x}_1,\ \bar{x}_2,\ \dots,\ \bar{x}_n]^T$ ou como vetor que contém os elementos $\bar{x}_j,\ j=1,\ \dots,\ n$ que satisfazem a todas as equações do sistema.

3.3 Classificação de sistemas lineares

Os sistemas lineares podem se classificados de acordo com o número de soluções que ele apresenta:

a) Compatível e determinado: Quando houver uma única solução.

b) Compatível e indeterminado: Quando houverem infinitas soluções.

c) Incompatível: Quando o sistema não admitir solução.

3.4 Sistemas triangulares

Sistemas triangulas são aqueles em que os elementos a baixo ou acima da diagonal principal da matriz de coeficientes são nulos. Aqui trabalharemos apenas com o sistema triangular superior, uma vez que um sistema triangular inferior pode ser facilmente convertido em um sistema triangular superior. Assim sendo todo sistema $Ax=b$ em que $a_{ij}=0 \forall\ j<i,$ ou seja, o sistema tem a forma:

$$\left\{\begin{array}{rrrcrcl} a_{11}x_1 & +a_{12}x_2 & a_{13}x_3 & + \dots & +a_{1n}x_n & = & b_1 \\ & +a_{22}x_2 & a_{23}x_3 & + \dots & +a_{2n}x_n & = & b_2 \\ & & a_{33}x_3 & + \dots & +a_{3n}x_n & = & b_3 \\ & & & & \vdots & \vdots & \vdots \\ & & & & +a_{nn}x_n & = & b_n \end{array}\right\} \quad (15)$$

Tais sistemas são resolvidos por substituições retroativas, através de equações da forma:

$$x_i=\frac{b_i-\sum_{j=i+1}^{n} a_{ij}x_j}{a_{ii}} \quad \forall\ i=n,\ \dots,\ 1 \quad (16)$$

3.4.1 Discussão da solução

1. Se $a_{ii}\neq 0 \quad \forall\ i$ o sistema é **compatível e determinado**.
2. Se $a_{ii}=0$ para algum i há dois casos possíveis:

a) Se $b_i - \sum_{j=i+1}^{n} a_{ij} x_j = 0$ o sistema é **compatível e indeterminado**

b) Se $b_i - \sum_{j=i+1}^{n} a_{ij} x_j \neq 0$ o sistema é **incompatível**

3.4.2 Algoritmo

A seguir está o pseudocódigo do procedimento que resolve um sistema triangular superior por intermédio de substituições retroativas. Supõe-se neste procedimento que $aii \neq 0 \ \forall i$, isto é, que os elementos diagonais da matriz dos coeficientes do sistema são todos não-nulos.

```
substituição_retroativa(n,A,b,x)
x_n ← b_n / a_nn;
para i de n - 1 até 1 passo -1 faça
      soma ← 0;
      para j de i + 1 até n faça
            soma ← soma + a_ij × x_j;
      fim-para;
      x_i ← (b_i – soma) / a_ii;
fim-para
retorne x
```

3.5 Métodos numéricos

É notório que a maioria dos sistemas lineares não estão na forma triangular, assim são necessários outros métodos para a resolução destes sistemas. Os métodos numéricos empregados na resolução de sistemas lineares são divididos em dois grupos: métodos diretos e métodos interativos. Iniciaremos estudando os métodos diretos.

3.6 Métodos diretos

São métodos que produzem a solução exata de um sistema, com exceção dos erros de arredondamento, depois de um número finito de operações aritméticas. Com esses métodos é possível determinar, a priori, o tempo máximo gasto para resolver um sistema, uma vez que sua complexidade é conhecida.

A clássica Regra de Cramer, ensinada no ensino médio, é um método direto. Entretanto, pode-se mostrar que o número máximo de operações aritméticas envolvidas na resolução de um sistema $n \times n$ por este método é $(n+1)(n!n-1)+n$ Assim, um computador que efetua uma operação aritmética em 10^{-8} segundos gastaria cerca de 36 dias para resolver um sistema de ordem n = 15. A complexidade exponencial desse algoritmo inviabiliza sua utilização em casos práticos. O estudo de métodos mais eficientes torna-se, portanto, necessário, uma vez que, em geral, os casos práticos exigem a resolução de sistemas lineares de porte mais elevado.

Apresentaremos, a seguir, métodos mais eficientes, cuja complexidade é polinomial, para resolver sistemas lineares. Antes, porém, introduziremos uma base teórica necessária à apresentação de tais métodos.

Transformações elementares: Denominam-se transformações elementares as seguintes operações efetuadas sobre as equações (ou linhas da matriz aumentada) de um sistema linear:

1. Trocar duas equações:
   $L_i \Leftarrow L_j$;
   $L_j \Leftarrow L_i$;
2. Multiplicar uma equação por uma constante não nula:
   $L_j \Leftarrow c \times L_j \quad c \in \mathbb{R}, \quad c \neq 0$
3. adicionar a uma equação um múltiplo de uma outra equação:
   $L_i \Leftarrow L_j + c \times L_i \quad c \in \mathbb{R}$

Sistemas equivalentes: Dois sistemas $Ax = b$ e $\tilde{A}x = \tilde{b}$ se dizem equivalentes se a solução de um for também solução do outro.

Teorema: Seja $Ax = b$ um sistema linear. Aplicando-se somente transformações elementares sobre as equações de $Ax = b$ obtemos um novo sistema $\tilde{A}x = \tilde{b}$ sendo que $Ax = b$ e $\tilde{A}x = \tilde{b}$ são equivalentes.

3.7 Método de Gauss

O método de Gauss consiste em operar transformações elementares sobre as equações de um sistema $Ax = b$ até que, depois de $n-1$ passos, se obtenha um sistema triangular superior, $Ux = c$, equivalente ao sistema dado, sistema esse que é resolvido por substituições retroativas.

$$\underbrace{\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array}\right]}_{Ax=b} \rightarrow \underbrace{\left[\begin{array}{cccc|c} a'_{11} & a'_{12} & \dots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \dots & a'_{2n} & b'_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a'_{nn} & b'_n \end{array}\right]}_{Ux=c}$$

3.7.1 Descrição do método

Para descrevermos o método, consideraremos o sistema linear 4×4 abaixo.

$$\left[\begin{array}{rrrrcr} 7x1 & +4x2 & -2x3 & +x4 & = & 14{,}308 \\ 3x1 & +11x2 & +4x3 & -5x4 & = & 25{,}744 \\ -2x1 & +3x2 & +8x3 & +2x4 & = & -3{,}872 \\ 10x1 & -5x2 & +x3 & -3x4 & = & 36{,}334 \end{array}\right] \tag{17}$$

A resolução deste sistema pelo método de Gauss envolve duas fases distintas. A primeira, chamada de fase de eliminação, consiste em transformar o sistema dado em um sistema triangular superior. A segunda, chamada de fase de substituição, consiste em resolver o sistema triangular superior através de substituições retroativas.

Para aplicar a primeira fase, utilizemos o quadro abaixo, onde cada grupo de linhas representa um passo (ou estágio) da obtenção do sistema triangular superior. Trabalharemos com 3 dígitos com arredondamento na apresentação em ponto flutuante.

Tabela 2: Fase de eliminação

Linha	Multiplicadores	Coeficientes das incógnitas				Termos Ind.	Transformações Elementares
$L_1^{(0)}$		7	4	-2	1	14,308	
$L_2^{(0)}$	$m_{21} = -3/7 = -0,429$	3	11	4	-5	25,744	
$L_3^{(0)}$	$m_{31} = 2/7 = 0,286$	-2	3	8	2	-3,872	
$L_4^{(0)}$	$m_{41} = -10/7 = -1,429$	10	-5	1	-3	36,334	
$L_2^{(1)}$		0	9,284	4,858	-5,429	19,606	$L_2^{(1)} \leftarrow -0,429 \times L_1^{(0)} + L_2^{(0)}$
$L_3^{(1)}$	$m_{32} = -4,144/9,284 = -0,446$	0	4,144	7,428	2,286	0,220	$L_3^{(1)} \leftarrow 0,286 \times L_1^{(0)} + L_3^{(0)}$
$L_4^{(1)}$	$m_{42} = 10,716/9,284 = 1,154$	0	-10,716	3,858	-4,429	15,888	$L_4^{(1)} \leftarrow -1,429 \times L_1^{(0)} + L_4^{(0)}$
$L_3^{(2)}$		0	0	5,261	4,707	-8,524	$L_3^{(2)} \leftarrow -0,446 \times L_2^{(1)} + L_3^{(1)}$
$L_4^{(2)}$	$m_{43} = -9,464/5,261 = -1,799$	0	0	9,464	-10,694	38,513	$L_4^{(2)} \leftarrow 1,154 \times L_2^{(1)} + L_4^{(1)}$
$L_4^{(3)}$		0	0	0	-19,162	53,848	$L_4^{(3)} \leftarrow -1,799 \times L_3^{(2)} + L_4^{(2)}$

Detalhando a Tabela 2 observamos que constam 3 passos:

Passo $k=1$:

Pivô: $a_{11}^{(0)}=7$

Linha pivotal: $L_1^{(0)}$

Objetivo: zerar os elementos abaixo do pivô $a_{11}^{(0)}$.

Ao final do primeiro passo obtemos o sistema $A^{(1)}x=b^{(1)}$ equivalente ao sistema inicial, em que:

$$[A^{(1)}|b^{(1)}]=\left[\begin{array}{cccc|c} 7 & 4 & -2 & 1 & 14,308 \\ 0 & 9,284 & 4,858 & -5,429 & 19,606 \\ 0 & 4,144 & 7,428 & 2,286 & 0,220 \\ 0 & -10,716 & 3,858 & -4,429 & 15,888 \end{array}\right]$$

Passo $k=2$:

Pivô: $a_{22}^{(1)}=9,284$

Linha pivotal: $L_2^{(1)}$

Objetivo: zerar os elementos abaixo do pivô $a_{22}^{(1)}$.

Ao final do primeiro passo obtemos o sistema $A^{(2)}x=b^{(2)}$ equivalente ao sistema inicial, em que:

$$[A^{(2)}|b^{(2)}]=\left[\begin{array}{cccc|c} 7 & 4 & -2 & 1 & 14,308 \\ 0 & 9,284 & 4,858 & -5,429 & 19,606 \\ 0 & 0 & 5,261 & 4,707 & -8,524 \\ 0 & 0 & 9,464 & -10,694 & 38,513 \end{array}\right]$$

Passo $k=3$:

Pivô: $a_{33}^{(2)}=5,261$

Linha pivotal: $L_3^{(2)}$

Objetivo: zerar os elementos abaixo do pivô $a_{33}^{(2)}$.

Ao final do primeiro passo obtemos o sistema $A^{(3)}x=b^{(3)}$ equivalente ao sistema inicial, em que:

$$[A^{(2)}|b^{(2)}]=\left[\begin{array}{cccc|c} 7 & 4 & -2 & 1 & 14{,}308 \\ 0 & 9{,}284 & 4{,}858 & -5{,}429 & 19{,}606 \\ 0 & 0 & 5{,}261 & 4{,}707 & -8{,}524 \\ 0 & 0 & 0 & -19{,}162 & 53{,}848 \end{array}\right]$$

Portanto, ao final de 3 passos, o sistema $Ax=b,$ expresso por (17), foi transformado no seguinte sistema triangular superior $A^{(3)}x=b^{(3)}$.

$$\left[\begin{array}{ccccc} 7 & +4 & -2 & +1 & = & 14{,}308 \\ & +9{,}284 & +4{,}858 & -5{,}429 & = & 19{,}606 \\ & & +5{,}261 & +4{,}707 & = & -8{,}524 \\ & & & -19{,}162 & = & 53{,}848 \end{array}\right]$$

Terminada a fase de eliminação, passamos, agora, à fase de substituição, resolvendo o sistema anterior através das seguintes substituições retroativas:

$$x_4=\frac{-53{,}848}{19{,}162}=-2{,}810$$

$$x_3=\frac{-8{,}524-4{,}707\times(-2{,}810)}{5{,}261}=0{,}892$$

$$x_2=\frac{19{,}606+5{,}429\times(-2{,}810)-4{,}858\times 0{,}894}{9{,}284}=0{,}001$$

$$x_1=\frac{14{,}308-4\times 0{,}001+2\times 0{,}894-2{,}810}{7}=2{,}700$$

Portanto a solução do sistema é:

$$\bar{x}=\left[\begin{array}{c} 2{,}700 \\ 0{,}001 \\ 0{,}894 \\ -2{,}810 \end{array}\right]$$

3.7.2 Algoritmo para o método de Gauss

Apresentamos, a seguir, o pseudocódigo do procedimento relativo à fase de eliminação do método de Gauss. A ele se segue o procedimento de substituição retroativa descrito anteriormente. Esse algoritmo supõe que os elementos diagonais (a_{kk}) são não-nulos. Na hipótese de existir algum $(a_{kk}=0)$ esse elemento deve ser colocado em outra posição fora da diagonal principal, por intermédio de operações de troca de linhas e/ou colunas.

```
eliminação(n,A,b)
para k de 1 até n - 1 faça
    para i de k + 1 até n faça
        m ← -aik / akk;
        para j de k até n faça
            aij ← aij + m × akj;
        fim-para;
        bi ← bi + m × bk;
    fim-para
fim-para
retorne A e b;
```

3.7.3 Avaliação do resíduo

O erro ε produzido por uma solução $\bar{x}$ do sistema $Ax=b$ pode ser avaliado pela expressão:

$$\varepsilon = \max_{1 \le i \le n} |r_i| \tag{18}$$

Sendo r_i a i-ésima componente do vetor de resíduo *R*, o qual é dado por:

$$R = b - A\bar{x} \tag{19}$$

Para o exemplo considerado, o vetor de resíduo é:

$$R = b - A\bar{x} = \begin{bmatrix} 14{,}308 \\ 25{,}744 \\ -3{,}872 \\ 36{,}334 \end{bmatrix} - \begin{bmatrix} 7 & 4 & -2 & 1 \\ 3 & 11 & 4 & -5 \\ -2 & 3 & 8 & 2 \\ 10 & -5 & 1 & -3 \end{bmatrix} \begin{bmatrix} 2{,}700 \\ 0{,}001 \\ 0{,}894 \\ -2{,}810 \end{bmatrix} = \begin{bmatrix} 0{,}002 \\ 0{,}007 \\ -0{,}007 \\ 0{,}015 \end{bmatrix}$$

Assim, o erro ε cometido vale:

$$\varepsilon = \max_{1 \le i \le n} |r_i| = \max_{1 \le i \le n} \{|0{,}002|, |0{,}007|, |-0{,}007|, |0{,}015|\} = 0{,}015$$

3.8 Exercícios

1) Faça um programa utilizando a linguagem que desejar que aplique o algoritmo da substituição retroativa.
2) Altere o programa do exercício anterior para que ele determine o sistema é "compatível e determinado", "compatível e indeterminado" ou "incompatível"
3) Utilizando o programa do exercício anterior determine a solução para os seguintes sistemas.

a)
$$\begin{bmatrix} 3x1 & +4x2 & -5x3 & +x4 & = & -10 \\ & +x2 & +x3 & -2x4 & = & -1 \\ & & +4x3 & -5x4 & = & 3 \\ & & & 2x4 & = & 2 \end{bmatrix}$$

Solução: $\bar{x} = [1,\ -1,\ 2,\ 1]^T$

b)
$$\begin{bmatrix} 3x1 & +4x2 & -5x3 & +x4 & = & -10 \\ & & +x3 & -2x4 & = & 0 \\ & & +4x3 & -5x4 & = & 3 \\ & & & 2x4 & = & 2 \end{bmatrix}$$

Solução: indeterminado

c)
$$\begin{bmatrix} 3x1 & +4x2 & -5x3 & +x4 & = & -10 \\ & & +x3 & -2x4 & = & -1 \\ & & +4x3 & -5x4 & = & 3 \\ & & & 2x4 & = & 2 \end{bmatrix}$$

Solução: incompatível

d)
$$\begin{bmatrix} x1 & +x2 & +x3 & +x4 & = & 4 \\ & +x2 & +3x3 & +x4 & = & 3 \\ & & +x3 & +x4 & = & 2 \\ & & & +x4 & = & 1 \end{bmatrix}$$

Solução: $\bar{x} = [3,\ -1,\ 1,\ 1]^T$

4) Faça um programa utilizando a linguagem que desejar que aplique o algoritmo de Gauss e em seguida o algoritmo da substituição retroativa.
5) Altere o programa do exercício anterior para que ele determine o sistema é "compatível e determinado", "compatível e indeterminado" ou "incompatível"
6) Utilizando o programa do exercício anterior determine a solução para os seguintes sistemas.

a)
$$\begin{bmatrix} 2x1 & +3x2 & -x3 & = & 5 \\ 4x1 & +4x2 & -3x3 & = & 3 \\ 2x1 & -3x2 & +x3 & = & -1 \end{bmatrix}$$

Solução: $\bar{x} = [1,\ 2,\ 3]^T$

b)
$$\begin{bmatrix} 2x1 & +3x2 & +x3 & -x4 & = & 6{,}9 \\ -x1 & +x2 & -4x3 & +x4 & = & -6{,}6 \\ x1 & +x2 & +x3 & +x4 & = & 10{,}2 \\ 4x1 & -5x2 & +x3 & -2x4 & = & -12{,}3 \end{bmatrix}$$

Solução: $\bar{x} = [0{,}9,\ 2{,}1,\ 3{,}0,\ 4{,}2]^T$

c)
$$\begin{bmatrix} 4x1 & +3x2 & +2x3 & +x4 & = & 10 \\ x1 & +2x2 & +3x3 & +4x4 & = & 5 \\ x1 & -x2 & -x3 & -x4 & = & -1 \\ x1 & +x2 & +x3 & +x4 & = & 3 \end{bmatrix}$$

Solução: Indeterminado

d)
$$\begin{bmatrix} x1 & +2x2 & +3x3 & +4x4 & = & 10 \\ 2x1 & +x2 & +2x3 & +3x4 & = & 7 \\ 3x1 & +2x2 & +x3 & +2x4 & = & 6 \\ 4x1 & +3x2 & +2x3 & +x4 & = & 5 \end{bmatrix}$$

Solução: $\bar{x} = [0,\ 1,\ 0,\ 2]^T$

AULA 4 - RESOLUÇÃO DE SISTEMAS LINEARES – PIVOTAMENTO, JORDAN E DETERMINANTES

4.1 Objetivo:

Nesta aula os alunos conhecerão os problemas relacionados ao método de Gauss utilizado na resolução de sistemas de equações lineares. Aprenderão também a utilizar o método do pivotamento para contornar estes problemas. Esta aula também tem o objetivo de apresentar aos alunos o método de Jordan para resolução de sistemas lineares. E finalmente também será estudado um método para encontrar o determinante de uma matriz, baseado nos métodos estudados anteriormente.

4.2 Introdução

No método de Gauss, os multiplicadores do passo k da fase de eliminação são calculados pela expressão:

$$m_{ik}^{(k)} = \frac{-a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \forall\, i = k+1, \ldots, n \tag{20}$$

Observe que o pivô do k-ésimo passo da fase de eliminação é sempre $a_{kk}^{(k-1)}$, isto é, o elemento diagonal da matriz $A^{(k-1)}$ do sistema transformado $A^{(k-1)}x = b^{(k-1)}$ obtido no passo anterior.

Desvantagens do método de Gauss:

1. Não pode ser aplicado quando o pivô for nulo ($a_{kk} = 0$);
2. Os erros de arredondamento cometidos durante um passo da obtenção do sistema triangular se propagam para os passos seguintes, podendo comprometer a validade da solução obtida.

Para contornar o problema (1) e minimizar o problema (2), a ideia é usar uma estratégia de pivotamento, conforme a seguir se descreve.

4.3 O Método de Gauss com Pivotamento Parcial

Esta estratégia de pivotamento consiste em:

1. No início da etapa *k* da etapa de eliminação, escolher para pivô o maior elemento, em módulo, dentre os coeficientes:

$$a_{ik}^{(k-1)}; i = k, k+1, \ldots, n$$

2. Trocar as linhas *k* e *i* se necessário

Exemplo

Resolver o sistema a seguir, avaliando o erro cometido em cada caso:

$$\begin{bmatrix} 0{,}0002\,x_1 & +2\,x_2 & = & 5 \\ 2\,x_1 & +2\,x_2 & = & 6 \end{bmatrix} \tag{21}$$

(a) Pelo método de Gauss

(b) Pelo método de Gauss com pivotamento parcial

Resolução do item (a):

A Tabela 3 apresenta a fase de eliminação do Método de Gauss aplicado ao sistema linear (21).

Tabela 3: Fase de eliminação do método de Gauss

Linha	Multiplicadores	Coeficientes das incógnitas		Termos Ind.	Transformações Elementares
$L_1^{(0)}$		$0,0002$	2	5	
$L_2^{(0)}$	$m_{21} = -2/0,0002 = -10^4$	2	2	6	
$L_2^{(1)}$		0	-19998	-49994	$L_2^{(1)} \leftarrow -10000 \times L_1^{(0)} + L_2^{(0)}$

Tendo triangularizado a matriz dos coeficientes do sistema (21), passemos à fase de resolução do sistema triangular (22), o qual é equivalen+te ao sistema dado:

$$\begin{bmatrix} 0{,}0002\,x_1 & +2\,x_2 & = & 5 \\ & -19998\,x_2 & = & -49994 \end{bmatrix} \tag{22}$$

Cuja solução é:

$$\bar{x} = \begin{bmatrix} 0{,}0001 \\ 2{,}4999 \end{bmatrix}$$

Avaliemos o resíduo R e o erro ε produzido por esta solução.

$$R = b - A\,\bar{x} = \begin{bmatrix} 5 \\ 6 \end{bmatrix} - \begin{bmatrix} 4{,}9998 \\ 5 \end{bmatrix} = \begin{bmatrix} 0{,}0002 \\ 1 \end{bmatrix}$$

$$\varepsilon = \max_{1 \leqslant i \leqslant n} |r_i| = \max_{1 \leqslant i \leqslant n} \{|0{,}0002|, |1|\} = 1$$

Resolução do item (b):

A Tabela 4 apresenta a fase de eliminação do método de Gauss, com pivotamento parcial, aplicado ao sistema linear (21).

Tendo triangularizado a matriz dos coeficientes do sistema (21), passemos à fase de resolução do sistema triangular (5.14), que é equivalente ao sistema dado:

$$\begin{bmatrix} 2x_1 & +2x_2 & = & 6 \\ & 1.9998x_2 & = & 4.9994 \end{bmatrix} \tag{23}$$

Cuja solução é:

Tabela 4: Método de Gauss com pivotamento

Linha	Multiplicadores	Coeficientes das incógnitas		Termos Ind.	Transformações Elementares
$L_1^{(0)}$		$0,0002$	2	5	
$L_2^{(0)}$		2	2	6	
$L_1^{(0)'}$		2	2	6	$L_1^{(0)'} \leftarrow L_2^{(0)}$
$L_2^{(0)'}$	$m_{21} = -0,0002/2 = -0,0001$	$0,0002$	2	5	$L_2^{(0)'} \leftarrow L_1^{(0)}$
$L_2^{(1)}$		0	$1,9998$	$4,9994$	$L_2^{(1)} \leftarrow -0,0001 \times L_1^{(0)'} + L_2^{(0)'}$

$$\bar{x} = \begin{bmatrix} 0,5001 \\ 2,4999 \end{bmatrix}$$

Avaliemos o resíduo R e o erro ε produzido por esta solução.

$$R = b - A\bar{x} = \begin{bmatrix} 5 \\ 6 \end{bmatrix} - \begin{bmatrix} 4,9999 \\ 6,0018 \end{bmatrix} = \begin{bmatrix} 0,0001 \\ -0,0018 \end{bmatrix}$$

$$\varepsilon = \max_{1 \leqslant i \leqslant n} |r_i| = \max_{1 \leqslant i \leqslant n} \{|0,0001|, |-0,0018|\} = 0,0018$$

Tais resultados mostram, claramente, a melhora obtida com a técnica de pivotamento.

Observamos, finalmente, que a escolha do maior elemento em módulo entre os candidatos a pivô faz com que os multiplicadores, em módulo, estejam entre zero e um, o que minimiza a ampliação dos erros de arredondamento.

Apresentamos, a seguir, o pseudocódigo do procedimento de Gauss com pivotamento parcial para resolver sistemas lineares. Neste procedimento, $\bar{A}$ é a matriz aumentada do sistema, isto é, $\bar{A} = [A|b]$.

```
eliminação_pivotamento(n, A, b)
para k de 1 até n - 1 faça
    w ← |akk|;
    r ← k ;
    para i de k até n faça
        se |aik| > w então
            w ← |aik|;
            r ← i ;
        fim-se
    fim-para;
    para j de k até n faça
        aux ← akj;
        akj ← arj;
```

```
            a_rj ← aux;
        fim-para;
        aux ← b_k;
        b_k ← b_r;
        b_r ← aux;
        para i de k +1 até n faça
            m ← -a_ik / a_kk;
            para j de k até n faça
                a_ij ← a_ij + m × a_kj;
            fim-para;
            b_i ← b_i + m × b_k;
        fim-para
fim-para
retorne A e b;
```

4.4 O método de Gauss-Jordan

O método de Gauss-Jordan é uma variação do método de eliminação de Gauss. A maior diferença é que, quando uma variável é eliminada no método de eliminação de Gauss-Jordan, ela também é eliminada de todas as outras equações, não só as posteriores. Além disso, todas as linhas são normalizadas pela divisão pelo seu elemento pivô. Então o passo de eliminação resulta na matriz identidade e não em uma matriz triangular. Vaja a seguir.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \vdots & b_1 \\ a_{21} & a_{22} & a_{23} & \vdots & b_2 \\ a_{n1} & a_{n2} & a_{33} & \vdots & b_3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & \vdots & b_1^{(n)} \\ 0 & 1 & 0 & \vdots & b_2^{(n)} \\ 0 & 0 & 1 & \vdots & b_3^{(n)} \end{bmatrix}$$

$$\begin{bmatrix} x_1 & & & \vdots & b_1^{(n)} \\ & x_2 & & \vdots & b_2^{(n)} \\ & & x_3 & \vdots & b_3^{(n)} \end{bmatrix}$$

Consequentemente, não é necessário usar a substituição regressiva para obter o resultado. O método é mais bem ilustrado por um exemplo.

Exemplo:

$$\begin{bmatrix} 3x_1 & -0{,}1x_2 & -0{,}2x_3 & = & 7{,}85 \\ 0{,}1x_1 & +7x_2 & -0{,}3x_3 & = & -19{,}3 \\ 0{,}3x_1 & -0{,}2x_2 & +10x_3 & = & 71{,}4 \end{bmatrix}$$

Primeiramente expresse os coeficientes e as constantes na forma de uma matriz.

$$\begin{bmatrix} 3 & -0,1 & -0,2 & 7,85 \\ 0,1 & +7 & -0,3 & -19,3 \\ 0,3 & -0,2 & +10 & 71,4 \end{bmatrix}$$

Então, normalize a primeira linha dividindo-a pelo elemento pivô, 3, para obter.

$$\begin{bmatrix} 1 & -0,0333333 & -0,066667 & 2,61667 \\ 0,1 & +7 & -0,3 & -19,3 \\ 0,3 & -0,2 & +10 & 71,4 \end{bmatrix}$$

O termo x_1 pode ser eliminado da segunda linha subtraindo 0,1 vezes a primeira linha da segunda linha. De maneira semelhante, subtraindo 0,3 vezes a primeira linha da terceira linha, eliminaremos o termo x_1 da terceira linha:

$$\begin{bmatrix} 1 & -0,0333333 & -0,066667 & 2,61667 \\ 0 & +7.00333 & -0,293333 & -19,5617 \\ 0 & -0,19000 & +10,0200 & 70,6150 \end{bmatrix}$$

Em seguida, normalizamos a segunda linha, dividindo-a por 7,00333.

$$\begin{bmatrix} 1 & -0,0333333 & -0,066667 & 2,61667 \\ 0 & 1 & -0,0418848 & -2,79320 \\ 0 & -0,19000 & +10,0200 & 70,6150 \end{bmatrix}$$

A redução do termo x_2 da primeira e da terceira linhas, resulta em.

$$\begin{bmatrix} 1 & 0 & -0,0680629 & 2,52356 \\ 0 & 1 & -0,0418848 & -2,79320 \\ 0 & 0 & +10,01200 & 70,0843 \end{bmatrix}$$

A terceira linha é então normalizada dividindo-se por 10,012.

$$\begin{bmatrix} 1 & 0 & -0,0680629 & 2,52356 \\ 0 & 1 & -0,0418848 & -2,79320 \\ 0 & 0 & 1 & 7,0000 \end{bmatrix}$$

Finalmente, o termo x_3 pode ser reduzido da primeira e da segunda equação.

$$\begin{bmatrix} 1 & 0 & 0 & 3,0000 \\ 0 & 1 & 0 & -2,5000 \\ 0 & 0 & 1 & 7,0000 \end{bmatrix}$$

Assim, a matriz dos coeficientes foi transformada em uma matriz identidade, e a solução é obtida no vetor do lado direito. Observe que não foi necessária a substituição retroativa para se obter a solução.

A seguir temos um algoritmo para implementar o método de Gauss-Jordan. É importante salientar que o algoritmo recebe como entrada a matriz expandida, contendo os termos independentes na última coluna.

```
eliminação_gauss-jordan(n, A)
para k de 1 até n faça
    para i de 1 até n faça
        m ← -aik / akk;
        p ← akk;
        para j de k até n + 1 faça
            se i = k então
                aij ← aij / p;
            senão
                aij ← aij + m × akj;
            fim-se
        fim-para;
    fim-para;
fim-para
retorne A;
```

4.5 Cálculo de determinantes

De modo análogo ao que foi feito com sistemas, pode-se definir transformações elementares para matrizes e também definir matrizes equivalentes A e B quando B pode ser obtida de A por transformações elementares nas linhas ou colunas. Pode-se provar que se A e B são equivalentes então $det(A)=det(B)$.

Como nas matrizes triangulares ou diagonais o determinante é o produto dos elementos diagonais usa-se, para o cálculo de determinante, o método de Gauss ou de Gauss-Jordan.

Exemplo:

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

é transformada pelo método de Gauss-Jordan em:

$$D=\begin{bmatrix}1 & 0 & 0\\ 0 & -3 & 0\\ 0 & 0 & 0.3333\end{bmatrix}$$

$$\text{logo,}\quad det(A)=det(D)=1\cdot(-3)\cdot 0.33333=-1$$

A seguir temos um algoritmo baseado no método de Gauss-Jordan que encontra o determinante de uma matriz.

```
determinante(n, A)
para k de 1 até n faça
    para i de 1 até n faça
        m ← -a_ik / a_kk;
        para j de k até n faça
            se i ≠ k então
                a_ij ← a_ij + m × a_kj;
            fim-se
        fim-para;
    fim-para;
fim-para
det ← 1;
para k de 1 até n faça
    det ← det × a_kk;
fim-para
retorne det;
```

4.6 Exercícios

1) Faça um programa utilizando a linguagem que desejar aplicando o algoritmo de Gauss com pivotamento e em seguida o algoritmo da substituição retroativa para encontrar o vetor solução dos seguintes sistemas lineares.

a) $\begin{bmatrix} 2x1 & +3x2 & -x3 & = & 5 \\ 4x1 & +4x2 & -3x3 & = & 3 \\ 2x1 & -3x2 & +x3 & = & -1 \end{bmatrix}$

Solução: $\bar{x} = [1,\ 2,\ 3]^T$

b) $\begin{bmatrix} 2x1 & +3x2 & +x3 & -x4 & = & 6{,}9 \\ -x1 & +x2 & -4x3 & +x4 & = & -6{,}6 \\ x1 & +x2 & +x3 & +x4 & = & 10{,}2 \\ 4x1 & -5x2 & +x3 & -2x4 & = & -12{,}3 \end{bmatrix}$

Solução: $\bar{x} = [0{,}9,\ 2{,}1,\ 3{,}0,\ 4{,}2]^T$

c) $\begin{bmatrix} x1 & +2x2 & +3x3 & +4x4 & = & 10 \\ 2x1 & +x2 & +2x3 & +3x4 & = & 7 \\ 3x1 & +2x2 & +x3 & +2x4 & = & 6 \\ 4x1 & +3x2 & +2x3 & +x4 & = & 5 \end{bmatrix}$

Solução: $\bar{x} = [0,\ 1,\ 0,\ 2]^T$

2) Faça agora um programa aplicando o algoritmo de Gauss-Jordan. Encontre a solução para os sistemas lineares da questão anterior utilizando este programa.
3) Utilizando o algoritmo de Gauss-Jordan e faça um programa que calcula o determinante das seguintes matrizes.

a) $\begin{bmatrix} 2 & 3 & -1 \\ 4 & 4 & -3 \\ 2 & -3 & 1 \end{bmatrix}$

det = -20

b) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 2 \\ -1 & 3 & -2 & 1 \\ 3 & -1 & 1 & 2 \end{bmatrix}$

det = 63

AULA 5 -RESOLUÇÃO DE SISTEMAS LINEARES - JACOBI

5.1 Objetivo:

Nesta aula será apresentado um método interativo para a resolução de sistemas lineares. O método estudado é o método de Jacobi. Estudaremos seu funcionamento e será apresentado um algoritmo com sua implementação.

5.2 Introdução

A solução $\bar{x}$ de um sistema linear $A\bar{x}=b$ pode ser obtida utilizando-se um método interativo, que consiste em calcular uma sequencia $x^{(1)}, x^{(2)}, \ldots, x^{(k)}, \ldots$ de aproximações de $\bar{x}$, sendo dada uma aproximação inicial $x^{(0)}$. Para tanto, deve-se transformar o sistema dado num sistema equivalente, da forma

$$x = F\,x + d \tag{24}$$

Onde *F* é uma matriz de *n x n* e *x* e *d* são vetores de *n x 1*.

Partindo-se de uma aproximação inicial $x^{(0)}=(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})^T$ obtêm-se

$$\begin{aligned} x^{(1)} &= F\,x^{(0)} + d \\ x^{(2)} &= F\,x^{(1)} + d \\ &\vdots \\ x^{(k+1)} &= F\,x^{(k)} + d \\ &\vdots \end{aligned}$$

Seja $\|x^{(k)} - x\| = \max_{1 \le i \le n} \{(x_i^{(k)} - x_i)\}$

Se $\lim \|x^{(k)} - x\| = 0$ então $x^{(1)}, x^{(2)}, \ldots, x^{(k)}, \ldots$ converge quando $k \Rightarrow \infty$.

5.3 Método de Jacobi

Seja o sistema linear $A\,x = b$ em sua forma expandida:

$$\begin{Bmatrix} a_{11}x_1 & +a_{12}x_2 & +\ldots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\ldots & +a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +\ldots & +a_{nn}x_n & = & b_n \end{Bmatrix}$$

Explicitemos x_1 na primeira equação, x_2 na segunda equação e assim sucessivamente.

$$x_1 = \frac{b_1 - (a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n)}{a_{11}}$$
$$x_2 = \frac{b_2 - (a_{21}x_1 + a_{23}x_3 + ... + a_{2n}x_n)}{a_{22}}$$
$$\vdots$$
$$x_n = \frac{b_n - (a_{n1}x_1 + a_{n2}x_2 + ... + a_{n\,n-1}x_{n-1})}{a_{n\,n}}$$

O método de Jacobi consiste na seguinte sequência de passos:

(i) Escolher uma aproximação inicial $x^{(0)} = [x_1^{(0)}, x_2^{(0)}, ..., x_n^{(0)}]^T$ arbitrária;

(ii) Gerar aproximações sucessivas $x^{(k)}$ a partir de $x^{(k-1)}$ com base nas seguintes equações de interação:

$$x_1^{(k)} = \frac{b_1 - (a_{12}x_2^{(k-1)} + a_{13}x_3^{(k-1)} + ... + a_{1n}x_n^{(k-1)})}{a_{11}}$$
$$x_2^{(k)} = \frac{b_2 - (a_{21}x_1^{(k-1)} + a_{23}x_3^{(k-1)} + ... + a_{2n}x_n^{(k-1)})}{a_{22}}$$
$$\vdots$$
$$x_n^{(k)} = \frac{b_n - (a_{n1}x_1^{(k-1)} + a_{n2}x_2^{(k-1)} + ... + a_{n\,n-1}x_{n-1}^{(k-1)})}{a_{n\,n}}$$

Sinteticamente, cada componente $x_i^{(k)}$ é determinada com base na seguinte equação:

$$x_i^{(k)} = \frac{b_i - \sum_{j=1\ j\neq 1}^{n} a_{ij}x_j^{(k-1)}}{a_{ii}} \quad \forall\, i = 1,2,...,n \tag{25}$$

Na forma matricial:

$$x^{(k)} = J \cdot x^{(k-1)} + D \quad \forall\, k = 1,2... \tag{26}$$

sendo *J* e *D* definidas de acordo com (27) e (28). A matriz *J* é conhecida como "Matriz de interação de Jacobi".

$$J = \begin{bmatrix} 0 & -a_{12}/a_{11} & -a_{13}/a_{11} & ... & -a_{1n}/a_{11} \\ -a_{21}/a_{22} & 0 & -a_{23}/a_{22} & ... & -a_{2n}/a_{22} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -a_{n1}/a_{nn} & -a_{n2}/a_{nn} & -a_{n3}/a_{nn} & ... & 0 \end{bmatrix} \tag{27}$$

$$D = \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \vdots \\ b_n/a_{nn} \end{bmatrix} \tag{28}$$

(iii) Interromper o processo quando um dos critérios abaixo for satisfeito:

1. $\underset{1 \leq i \leq n}{máx} |x_i^{(k)} - x_i^{(k-1)}| < \varepsilon$ *onde* ε *é a tolerância permitida*
2. $k > MaxIter$ *onde MaxIter é o número máximo de iterações*

Exemplo:

Resolver o sistema (29), a seguir, pelo método de Jacobi usando como aproximação inicial $x^{(0)} = [0,\ 0,\ 0]^T$ e com critério de parada $\underset{1 \leq i \leq 3}{máx} |x_i^{(k)} - x_i^{(k-1)}| < 0{,}001$ ou $k > 10$ iterações:

$$\begin{bmatrix} 10x1 & +2x2 & +x3 & = & 7 \\ x1 & -15x2 & +x3 & = & 32 \\ 2x1 & +3x2 & +10x3 & = & 6 \end{bmatrix} \tag{29}$$

(a) Equações de iteração:

$x^{(k)} = J \cdot x^{(k-1)} + D$, onde:

$$J = \begin{bmatrix} 0 & -2/10 & -1/10 \\ 1/15 & 0 & 1/15 \\ -2/10 & -3/10 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 7/10 \\ -32/15 \\ 6/10 \end{bmatrix}$$

Assim:

$$x_1^{(k)} = \frac{7 - 2x_2^{(k-1)} - x_3^{(k-1)}}{10}$$
$$x_2^{(k)} = \frac{32 - x_1^{(k-1)} - x_3^{(k-1)}}{-15}$$
$$x_3^{(k)} = \frac{6 - 2x_1^{(k-1)} - 3x_2^{(k-1)}}{10}$$

(b) Determinação da solução do sistema:

Tabela 5: Determinação da solução

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	Erro $= \max\limits_{1 \leq i \leq 3} \lvert x_i^{(k)} - x_i^{(k-1)} \rvert$
0	0	0	0	-
1	0,7000	-2,1333	0,6000	2,1333
2	1,0667	-2,0467	1,1000	0,5000
3	0,9993	-1,9889	1,0007	0,0993
4	0,9977	-2,0000	0,9968	0,0111
5	1,0003	-2,0004	1,0005	0,0037
6	1,0000	-1,9999	1,0000	0,0004

Portanto, $\bar{x} = [1{,}0000, -1{,}9999, 1{,}0000]^T$ é a solução do sistema (29) com precisão $\varepsilon = 0{,}001$.

5.3.1 Convergência do método de Jacobi

O método de Jacobi é convergente sempre que a matriz A é estrita ou irredutivelmente uma matriz estritamente diagonal dominante.

Uma matriz é dita estritamente diagonal dominante se, para todas as linhas da matriz, o módulo do valor da matriz na diagonal é maior que a soma dos módulos dos demais valores (não-diagonais) daquela linha. Mais precisamente, a matriz A é estritamente diagonal dominante quando:

$$|a_{ii}| > \sum_{j=1\ j\neq i}^{n} |a_{ij}| \quad \forall i = 1,2,\ldots,n \tag{30}$$

Onde A_{ij} denota o termo da i-ésima linha e j-ésima coluna da matriz. O mesmo raciocínio se aplica para as colunas, e para uma matriz ser estritamente dominante **basta que seja por linhas ou por colunas**.

5.3.2 Algoritmo do Método de Jacobi

A seguir é apresentado o pseudocódigo do método de Jacobi. *Tol* é a tolerância admitida, *MaxIter* é o número máximo de iterações permitida e *x* é o vetor solução, o qual começa com uma aproximação inicial.

```
jacobi(n, A, MaxIter, Tol, x, b)
k ← 0;
erro ← ∞ ;
enquanto (k < MaxIter e erro ≥ Tol) faça
    erro ← 0;
    para i de 1 até n faça
        xant_i ← x_i;
    fim-para;
    para i de 1 até n faça
        soma ← 0;
        para j de 1 até n faça
            se j ≠ i então
                soma ← soma + a_ij * xant_j;
            fim-se
        fim-para
        x_i ← (b_i – soma)/a_ii;
        se (|x_i – xant_i|>erro) então
            erro ← |x_i - xant_i|;
        fim-se
    fim-para;
    k ← k+1;
fim-enquanto;
se (erro < Tol) então
    retorne x;
senão
Imprima: Não houve convergência em MaxIter iterações
```

5.4 Exercícios

1) Faça um programa utilizando a linguagem que desejar aplicando o algoritmo de Jacobi para encontrar o vetor solução dos seguintes sistemas lineares.

$$x^{(0)}=[0,\ 0,\ 0,\ 0]^T \quad \text{e} \quad \varepsilon<10^{-2}$$

a)
$$\begin{bmatrix} 4x_1 & +x_2 & +x_3 & +x_4 & = & 7 \\ 2x_1 & -8x_2 & +x_3 & -x_4 & = & -6 \\ x_1 & +2x_2 & -5x_3 & +x_4 & = & -1 \\ x_1 & +x_2 & +x_3 & -4x_4 & = & -1 \end{bmatrix}$$

Solução: $\bar{x}=[1{,}0017\ 1{,}0025\ 1{,}0021\ 1{,}0019]^T$

$$x^{(0)}=[1,\ 3,\ 1,\ 3]^T \quad \text{e} \quad \varepsilon<10^{-2}$$

b)
$$\begin{bmatrix} 5x_1 & -x_2 & +2x_3 & -x_4 & = & 5 \\ x_1 & +9x_2 & -3x_3 & +4x_4 & = & 26 \\ & 3x_2 & -7x_3 & +2x_4 & = & -7 \\ -2x_1 & +2x_2 & -3x_3 & +10x_4 & = & 33 \end{bmatrix}$$

Solução: $\bar{x}=[1{,}0032\ 1{,}9974\ 3{,}0028\ 3{,}99719]^T$

2) Altere o programa do exercício anterior de forma que ele indique se a matriz faz com que o método de Jacobi convirja.

AULA 6 -RESOLUÇÃO DE SISTEMAS LINEARES – GAUSS-SEIDEL E CONVERGÊNCIA

6.1 Objetivo:

Nesta aula será apresentado outro método interativo para a resolução de sistemas lineares. O método estudado é o método de Gauss-Seidel. Estudaremos seu funcionamento e em seguida será apresentado um algoritmo com sua implementação. Também veremos nesta aula vários critérios de convergência dos métodos interativos.

6.2 O Método de Gauss-Seidel

Este método difere do anterior apenas com relação às equações de iteração, as quais são:

$$x_1^{(k)} = \frac{b_1 - \left(a_{12}x_2^{(k-1)} + a_{13}x_3^{(k-1)} + \ldots + a_{1n}x_n^{(k-1)}\right)}{a_{11}}$$

$$x_2^{(k)} = \frac{b_2 - \left(a_{21}x_1^{(k)} + a_{23}x_3^{(k-1)} + \ldots + a_{2n}x_n^{(k-1)}\right)}{a_{22}}$$

$$x_3^{(k)} = \frac{b_3 - \left(a_{31}x_1^{(k)} + a_{32}x_2^{(k)} + \ldots + a_{3n}x_n^{(k-1)}\right)}{a_{22}}$$

$$\vdots$$

$$x_n^{(k)} = \frac{b_n - \left(a_{n1}x_1^{(k)} + a_{n2}x_2^{(k)} + \ldots + a_{n\,n-1}x_{n-1}^{(k)}\right)}{a_{n\,n}}$$

Sinteticamente:

$$x_i^{(k)} = \frac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k-1)}}{a_{ii}} \quad \forall\, i = 1,2,\ldots,n \tag{31}$$

Na forma matricial, o Método de Gauss-Seidel pode ser posto na forma:

$$x^{(k)} = Lx^{(k)} + Ux^{(k-1)} + D \tag{32}$$

Sendo:

$$L = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ -a_{21}/a_{22} & 0 & 0 & \ldots & 0 \\ -a_{31}/a_{33} & -a_{32}/a_{33} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -a_{n1}/a_{nn} & -a_{n2}/a_{nn} & -a_{n3}/a_{nn} & \ldots & 0 \end{bmatrix}$$

$$U=\begin{bmatrix} 0 & -a_{12}/a_{11} & -a_{13}/a_{11} & \dots & -a_{1n}/a_{11} \\ 0 & 0 & -a_{23}/a_{22} & \dots & -a_{2n}/a_{22} \\ 0 & 0 & 0 & \dots & -a_{3n}/a_{33} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

$$D=\begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \vdots \\ b_n/a_{nn} \end{bmatrix}$$

A equação (32) pode ser escrita na forma $x^{(k)}=Gx^{(k-1)}+D.$ De fato, a partir de (32), podemos escrever:

$$x^{(k)}-Lx^{(k)}=Ux^{(k-1)}+D$$

$$(I-L)x^{(k)}=Ux^{(k-1)}+D$$

$$x^{(k)}=\underbrace{(I-L)^{-1}U}_{G}x^{(k-1)}+\underbrace{(I-L)^{-1}D}_{\bar{D}}$$

$$x^{(k)}=Gx^{(k-1)}+\bar{D}$$

A matriz G, dada pela equação (33), é a chamada “matriz de iteração de Gauss-Seidel”.

$$G=(I-L)^{-1}U \tag{33}$$

Exemplo:

Resolver o sistema abaixo (que é o mesmo sistema (29) usado para exemplificar o Método de Jacobi) pelo Método de Gauss-Seidel usando como aproximação inicial $x^{(0)}=[0,\ 0,\ 0]^T$ e com critério de parada $\underset{1\leq i\leq 3}{máx}|x_i^{(k)}-x_i^{(k-1)}|<0{,}001$ ou $k > 10$ iterações:

$$\begin{bmatrix} 10x1 & +2x2 & +x3 & = & 7 \\ x1 & -15x2 & +x3 & = & 32 \\ 2x1 & +3x2 & +10x3 & = & 6 \end{bmatrix}$$

(a) Equações de iteração:

$x^{(k)}=Lx^{(k)}+Ux^{(k-1)}+D,$ onde:

$$L=\begin{bmatrix} 0 & 0 & 0 \\ 1/15 & 0 & 0 \\ -2/10 & -3/10 & 0 \end{bmatrix},\quad U=\begin{bmatrix} 0 & -2/10 & -1/10 \\ 0 & 0 & -1/15 \\ 0 & 0 & 0 \end{bmatrix},\quad D=\begin{bmatrix} 7/10 \\ -32/15 \\ 6/10 \end{bmatrix}$$

Assim:

$$x_1^{(k)} = \frac{7 - 2x_2^{(k-1)} - x_3^{(k-1)}}{10}$$
$$x_2^{(k)} = \frac{32 - x_1^{(k)} - x_3^{(k-1)}}{-15}$$
$$x_3^{(k)} = \frac{6 - 2x_1^{(k)} - 3x_2^{(k)}}{10}$$

(b) Determinação da solução do sistema:

Tabela 6: Solução por Gauss-Seidel

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	Erro $= \max_{1 \le i \le 3} \lvert x_i^{(k)} - x_i^{(k-1)} \rvert$
0	0	0	0	-
1	0,7000	-2,0867	1,0860	2,0867
2	1,0087	-1,9937	0,9964	0,3087
3	0,9991	-2,0003	1,0003	0,0096
4	1,0000	-2,0000	1,0000	0,0009

Portanto, $\bar{x} = [1{,}0000\,, -2{,}000\,, 1{,}0000\,]^T$ é a solução do sistema (29) com precisão $\varepsilon = 0{,}001$.

6.2.1 Algoritmo do Método de Gauss-Seidel

A seguir é apresentado o pseudocódigo do Método de Gauss-Seidel. *Tol* é a tolerância admitida, *MaxIter* é o número máximo de iterações permitida e *x* é o vetor solução, o qual começa com uma aproximação inicial.

```
gauss-seidel(n, A, MaxIter, Tol, x, b)
k ← 0;
erro ← ∞ ;
enquanto (k < MaxIter e erro ≥ Tol) faça
    erro ← 0;
    para i de 1 até n faça
        xant_i ← x_i;
        soma ← 0;
        para j de 1 até n faça
            se j ≠ i então
                soma ← soma + a_ij * x_j;
            fim-se
        fim-para
        x_i ← (b_i – soma)/a_ii;
        se (|x_i – xant_i|>erro) então
            erro ← |x_i - xant_i|;
        fim-se
```

```
    fim-para;
    k ← k+1;
fim-enquanto;
se (erro < Tol) então
    retorne x;
senão
Imprima: Não houve convergência em MaxIter iterações
```

6.3 Convergência dos Métodos Iterativos

Para os métodos iterativos de Jacobi e Gauss-Seidel são válidos os seguintes critérios de convergência:

6.3.1 Critério das colunas

O critério das colunas é condição suficiente para que um sistema linear convirja usando um método iterativo, sendo que:

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

Além do mais, quanto mais próximo de zero estiver a relação $\max\limits_{1\leq j\leq n} \dfrac{\sum\limits_{i=1\, i\neq j}^{n} |a_{ij}|}{|a_{jj}|}$, mais rápida será a convergência.

6.3.2 Critério das Linhas

É condição suficiente para que um sistema linear convirja usando um método iterativo que:

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

Além do mais, quanto mais próximo de zero estiver a relação $\max\limits_{1\leq i\leq n} \dfrac{\sum\limits_{j=1\, j\neq i}^{n} |a_{ij}|}{|a_{ii}|}$, mais rápida será a convergência.

6.3.3 Critério de Sassenfeld

Seja:

$$\beta_i = \frac{\sum\limits_{j=1}^{i-1} (|a_{ij}| \cdot \beta_j) + \sum\limits_{j=i+1}^{n} |a_{ij}|}{|a_{ii}|} \tag{34}$$

É condição suficiente para que um método iterativo convirja, que:

$$\beta = \max_{1 \leq i \leq n} \beta_i < 1$$

Além disso, quanto menor for β mais rápida será a convergência.

6.3.4 Critério do Raio Espectral

É condição necessária e suficiente para que um método iterativo convirja que $\rho(F) < 1$, isto é, que o raio espectral (maior autovalor, em módulo) da matriz de iteração do método seja menor que a unidade. Além disso, quanto mais próximo de zero for $\rho(F)$ mais rápida será a convergência.

Exemplo

Verificar se há garantia de convergência do sistema a seguir usando um método iterativo.

$$\begin{bmatrix} 3x1 & & +x3 & = & 3 \\ x1 & -x2 & & = & 1 \\ 2x1 & +x2 & +3x3 & = & 9 \end{bmatrix} \tag{35}$$

(a) Critério das colunas:

$$|a_{11}| = |3| = 3 \not> |a_{21}| + |a_{31}| = |1| + |2| = 3$$

$$|a_{22}| = |-1| = 1 \not> |a_{12}| + |a_{32}| = |0| + |1| = 1$$

$$|a_{33}| = |3| = 3 > |a_{13}| + |a_{23}| = |1| + |0| = 1$$

Como o critério das colunas não é verificado para as colunas 1 e 2 (bastava que não fosse satisfeito para uma única coluna), concluímos que esse critério não garante convergência se usarmos um método iterativo.

(b) Critério das linhas:

$$|a_{11}| = |3| = 3 > |a_{12}| + |a_{13}| = |0| + |1| = 1$$

$$|a_{22}| = |-1| = 1 \not> |a_{21}| + |a_{23}| = |1| + |0| = 1$$

$$|a_{33}| = |3| = 3 \not> |a_{31}| + |a_{32}| = |2| + |1| = 3$$

Como o critério das linhas não é verificado para as linhas 2 e 3, concluímos que não há garantia de convergência, por esse critério, se usarmos um método iterativo.

(c) Critério de Sassenfeld:

$$\beta_1 = \frac{|a_{12}| + |a_{13}|}{|a_{11}|} = \frac{|0| + |1|}{|3|} = 1/3$$

$$\beta_2 = \frac{|a_{21}| \times \beta_1 + |a_{23}|}{|a_{22}|} = \frac{|1| \times 1/3 + |0|}{|-1|} = 1/3$$

$$\beta_3 = \frac{|a_{31}| \times \beta_1 + |a_{32}| \times \beta_2}{|a_{33}|} = \frac{|2| \times 1/3 + |1| \times 1/3}{|3|} = 1/3$$

$$\beta = \max_{1 \leqslant i \leqslant 3} \beta_i = máx\{1/3,\ 1/3,\ 1/3,\} = 1/3$$

Como $\beta = 1/3 < 1$ resulta que o critério de Sassenfeld foi satisfeito. Portanto, pode-se aplicar um método iterativo ao sistema (35), uma vez que há garantia de convergência do mesmo.

6.4 Exercícios

1) Faça um programa utilizando a linguagem que desejar aplicando o algoritmo de Gauss-Seidel para encontrar o vetor solução dos seguintes sistemas lineares.

$$x^{(0)}=[0,\ 0,\ 0,\ 0]^T \quad \text{e} \quad \varepsilon<10^{-2}$$

a)
$$\begin{bmatrix} 4x_1 & +x_2 & +x_3 & +x_4 & = & 7 \\ 2x_1 & -8x_2 & +x_3 & -x_4 & = & -6 \\ x_1 & +2x_2 & -5x_3 & +x_4 & = & -1 \\ x_1 & +x_2 & +x_3 & -4x_4 & = & -1 \end{bmatrix}$$

Solução: $\bar{x}=[1{,}0011\ 1{,}0005\ 1{,}0000\ 1{,}0004]^T$

$$x^{(0)}=[1,\ 3,\ 1,\ 3]^T \quad \text{e} \quad \varepsilon<10^{-2}$$

b)
$$\begin{bmatrix} 5x_1 & -x_2 & +2x_3 & -x_4 & = & 5 \\ x_1 & +9x_2 & -3x_3 & +4x_4 & = & 26 \\ & 3x_2 & -7x_3 & +2x_4 & = & -7 \\ -2x_1 & +2x_2 & -3x_3 & +10x_4 & = & 33 \end{bmatrix}$$

Solução: $\bar{x}=[0{,}9992\ 2{,}0004\ 3{,}0000\ 3{,}9998]^T$

2) Altere o programa do exercício anterior de forma que ele indique se o método interativo utilizado converge para a matriz dada. Utilize o critério das colunas, o critério das linhas e o critério de Sassenfeld.
3) Aplique o critério de Sassenfeld na matriz a seguir e determine se ela converge para os métodos interativos.

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

Solução: $\beta_1=0{,}7\ \beta_2=0{,}44\ \beta_3=0{,}358\ \beta_4=0{,}274$ *assim* $\beta=0{,}7<1$, *o sistema converge.*

AULA 7 - RESOLUÇÃO DE EQUAÇÕES ALGÉBRICAS E TRANSCENDENTAIS - BISSEÇÃO

7.1 Objetivo:

Nesta aula pretende-se ensinar aos alunos o que são equações algébricas e transcendentais, bem como o que é a sua solução. Para tanto nesta e na próxima aula, serão estudados alguns métodos de resolução destes sistemas, iniciando pelo método da bisseção.

7.2 Equações Algébricas e Transcendentais

Equações algébricas são as equações em que as incógnitas são submetidas apenas às chamadas operações algébricas, ou seja, soma, subtração, multiplicação, divisão, potenciação inteira e radiciação. Como por exemplo:

$$f(x)=3x-4 \quad \text{e} \quad f(x)=x^2+2x-3$$

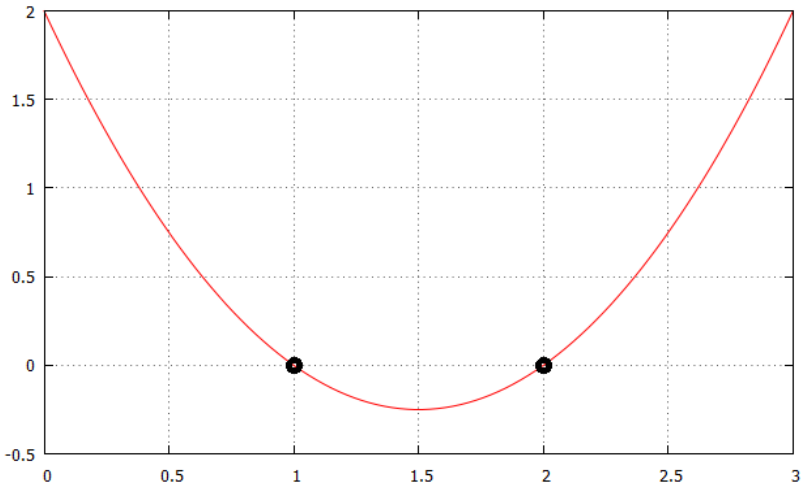
Uma equação transcendental é uma equação que contém alguma função que não é redutível a uma fração entre polinômios, e cuja solução não pode ser expressa através de funções elementares. De modo geral, uma equação transcendental não possui uma solução exata expressa através de funções conhecidas, sendo necessário recorrer ao cálculo numérico para obter uma solução. Como por exemplo:

$$f(x)=x\,e^x-2 \quad \text{e} \quad f(x)=x\sin(x)-x$$

7.3 Resolução de Algébricas e Transcendentais

Resolver uma equação $f(x)=0$ significa encontrar números ξ_i, denominados raízes, tais que $f(\xi_i)=0$. Geometricamente, conforme mostra a 6, as raízes representam os pontos de interseção do gráfico de f com o eixo x.

Figura 6: Raizes de uma equação



Existem diversas técnicas para encontrar as raízes de equações algébricas, a fórmula de Bhaskara é uma delas, e serve para determinar a solução de equações do segundo grau. Já para equações transcendentais não existe solução analítica.

Assim são necessários outros métodos que sirvam para encontras as soluções de qualquer tipo de equação.

A determinação de raízes envolve as seguintes fases:

Fase I - Isolamento

Nesta fase o objetivo é o de determinar um intervalo $[a;b]$, o menor possível, que contenha uma única raiz. Para cumprir este objetivo os métodos que apresentaremos a seguir apoiam-se em dois resultados do Cálculo Diferencial e Integral.

Teorema de Cauchy-Bolzano: Seja f uma função contínua em um intervalo $[a;b]$. Se $f(a)\times f(b)<0$ então existe pelo menos um ponto $\xi\in[a;b]: f(\xi)=0$.

Assim se f' preservar o sinal em $[a;b]$ e o Teorema de Cauchy-Bolzano for verificado neste intervalo então a raiz ξ é única.

Portanto, para isolarmos as raízes de uma equação $f(x)=0$ comumente utilizamos um dos seguintes procedimentos:

Procedimento I:

Esboçar o gráfico de f, determinando intervalos $[x_i;x_{i+1}]$ que contenham uma única raiz. Este objetivo pode ser cumprido gerando-se uma tabela de pontos $(x_i; f(x_i))$, onde os pontos inicial e final, bem como o valor do passo considerado $(x_{i+1}-x_i)$, dependerão do problema considerado e da experiência do usuário.

Exemplo: Isolar as raízes de $f(x)=2x-\cos x=0$.

Inicialmente, geremos uma tabela de pontos $x_i, f(x_i)$.

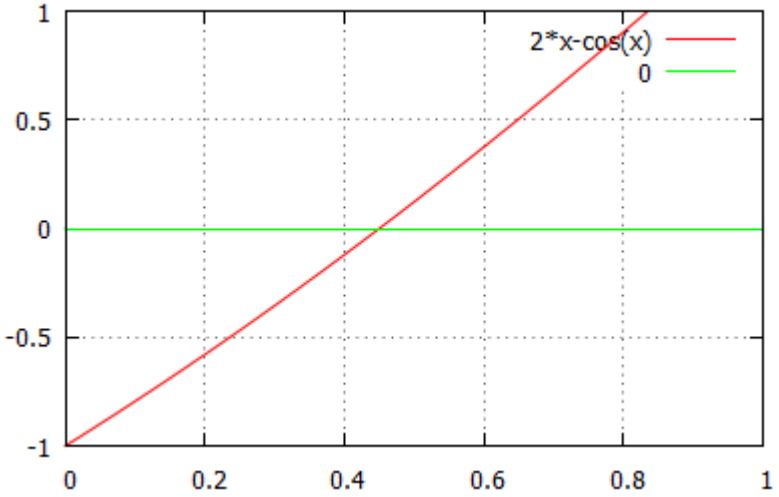
Tabela 7: Pontos de x e f(x)

x_i	$f(x_i)$
⋮	⋮
-2	-3.58
-1	-2.54
0	-1
1	1.46
2	4.42
⋮	⋮

Como $f(0)\times f(1)<0\Rightarrow\xi\in[0,1]$. sendo $f'(x)=2x+sen\,x>0\ \forall x\Rightarrow\xi$ é única.

Graficamente temos:

Figura 7: Gráfico da função 2x - cos(x)



Procedimento II:

Decompor a função f, se possível, na forma $f=g-h$, onde os gráficos de g e h sejam conhecidos e mais simples. Neste caso, os pontos de interseção dos gráficos de g e h representam as raízes de $f(x)=0$.

Com efeito, sejam x_i os pontos de interseção dos gráficos de g e h. Logo:

$$g(x_i)=h(x_i)\Rightarrow g(x_i)-h(x_i)=0 \tag{36}$$

Como $f(x)=(g-h)(x)=g(x)-h(x)\forall x$ então:

$$g(x_i)-h(x_i)=f(x_i) \tag{37}$$

Sendo $g(x_i)-h(x_i)=0$ (36), resulta pela equação (37) que $f(x_i)=0$, isto é, os valores x_i são as raízes de $f(x)=0$.

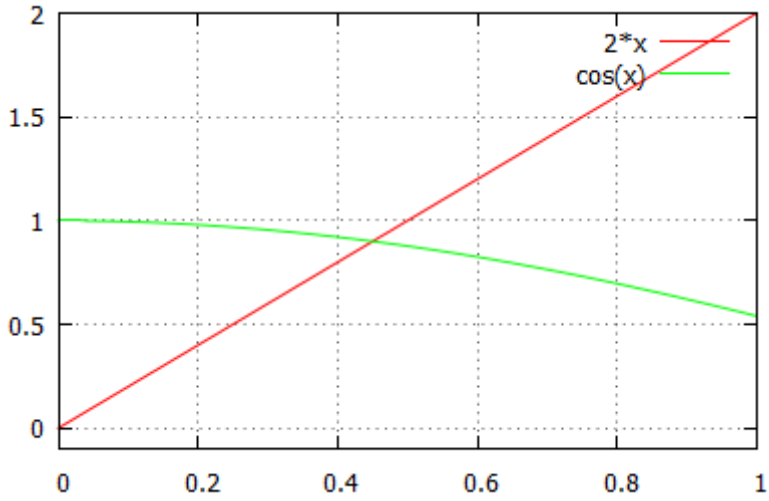
Exemplo: Isolar as raízes de $f(x)=2x-\cos(x)$.

Inicialmente, façamos a decomposição da função f dada:

$$f(x)=0\Leftrightarrow 2x-\cos(x)=0\Leftrightarrow \underbrace{2x}_{g(x)}=\underbrace{\cos(x)}_{h(x)}$$

Esbocemos, a seguir, os gráficos das funções $g(x)=2x$ e $h(x)=\cos(x)$.

Figura 8: Gráfico de 2x e cos(x)



A partir da visualização gráfica é possível determinar o ponto de intercessão onde $g(x_i)=h(x_i)$. Assim é possível determinar o intervalo onde a raiz esta contida.

Fase II - Refinamento

Uma vez isolada uma raiz em um intervalo $[a;b]$, procura-se, nesta fase, considerar uma aproximação para a raiz e "melhorá-la" sucessivamente até se obter uma aproximação com a precisão requerida.

Critérios de parada

Dizemos que x_k é uma "boa" aproximação para a raiz ξ de uma equação $f(x)=0$ se os critérios abaixo forem satisfeitos:

(i) $|f(x_k)|<\varepsilon$
(ii) $|x_k-\xi|<\varepsilon$

onde ε é a precisão (tolerância) admitida.

Observamos que estes dois critérios não são equivalentes. De fato:

Figura 9: Caso (a)

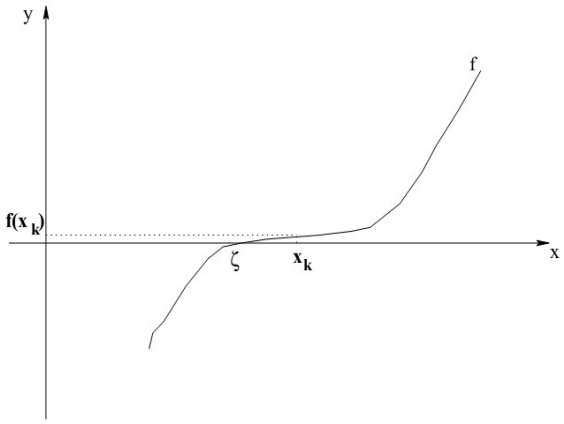
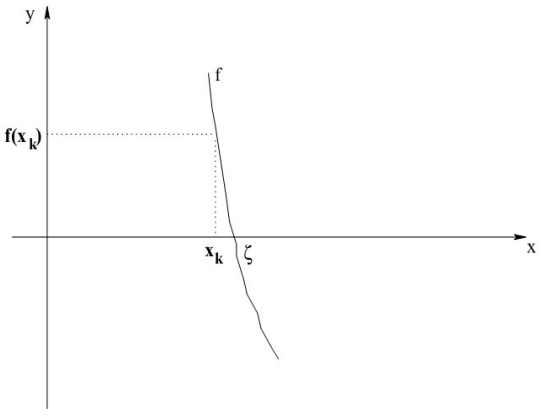


Figura 10: Caso (b)



No caso (a) temos $|f(x_k)|<\varepsilon$ mas com $|x_k-\xi|\gg\varepsilon$. No caso (b), ao contrário, temos $|x_k-\xi|<\varepsilon$ mas com $|f(x_k)|\gg\varepsilon$.

Desta forma, faz-se necessário impor os dois critérios. Por outro lado, como um determinado método pode não convergir em uma dada aplicação, é comum impor-se, também, um número máximo de iterações.

A questão que surge é: Como avaliar o critério de parada (ii) se não se conhece ξ ?

Para resolver esta questão a ideia é reduzir o intervalo $[a;b]$ que contém a raiz ξ até que sua amplitude seja inferior à precisão requerida, isto é, até que $b-a<\xi$.

Assim, sendo $b-a<\xi\Rightarrow\forall\, x_k\in[a,b]$. tem-se $|x_k-\xi|<b-a<\varepsilon$. Logo, $|x_k-\xi|<\varepsilon$. e qualquer $x_k\in[a,b]$ é uma boa aproximação para a raiz ξ.

7.4 Método da Bisseção

A ideia do Método da Bisseção é reduzir o intervalo $[a;b]$ que contém a raiz ξ dividindo-o ao meio a cada iteração.

7.4.1 O algoritmo para o método da bisseção

Apresentamos a seguir o pseudo-código do procedimento da bisseção.

```
bisseção(a, b, Tol, MaxIter)
k ← 0;
x ← (a+b)/2;

enquanto ((b – a ≥ Tol ou |f(x)| ≥ Tol) e k ≤ MaxIter) faça
    se f(a) * f(x) < 0 então
        b ← x;
    senão
        a ← x;
    fim-se
    x ← (a + b)/2;
    k ← k + 1;
fim-enquanto;

se (k ≤ MaxIter) então
    retorne x;
senão
    Imprima: Não houve convergência em MaxIter iterações
```

7.4.2 Estimativa do número de iterações

Estimemos o número de iterações necessárias para obter uma aproximação x_k com uma precisão ξ estabelecida a priori, utilizando-se o critério $|x_k-\xi|<\varepsilon$, como único critério de parada.

$$b_0 - a_0 = b - a$$
$$b_1 - a_1 = (b_0 - a_0)/2 = (b-a)/2$$
$$b_2 - a_2 = (b_1 - a_1)/2 = (b_0 - a_0)/4 = (b-a)/2^2$$
$$\vdots$$
$$b_k - a_k = (b-a)/2^k$$
$$\text{impondo } b_k - a_k < \varepsilon \text{, vem:}$$
$$\frac{b-a}{2^k} < \varepsilon \Rightarrow \frac{b-a}{\varepsilon} < 2^k > \frac{b-a}{\varepsilon} \Rightarrow \ln 2^k > \ln \frac{b-a}{\varepsilon} \Rightarrow k \ln 2 > \ln \frac{b-a}{\varepsilon}$$

Assim, o número mínimo de iterações necessárias para se calcular uma aproximação para a raiz de uma equação com precisão ε pode ser determinado pela expressão:

$$k > \frac{\ln(\frac{b-a}{\varepsilon})}{\ln 2} \tag{38}$$

Exemplo 1

Determinar o número de iterações necessárias para calcular a raiz de $f(x) = 2x - \cos(x) = 0$ no intervalo $[0;1]$ com precisão $\varepsilon < 0{,}01$, utilizando-se $|x_k - \xi| < \varepsilon$ como critério de parada.

Solução

Para o exemplo considerado temos: $a = 0$, $b = 1$, $\varepsilon < 0{,}01$. Aplicando o resultado (38), temos:

$$k > \frac{\ln(\frac{b-a}{\varepsilon})}{\ln 2} = \frac{\ln(\frac{1-0}{0{,}01})}{\ln 2} = 6{,}64$$

Como o número k de iterações é um número inteiro, resulta que $k = 7$.

Exemplo 2

Determinar com precisão $\varepsilon < 0{,}01$ e com um máximo de 10 iterações, a raiz da equação $f(x) = 2x - \cos(x) = 0$

Solução

(a) Isolamento da raiz:

Já foi visto que $\xi \in [0,1]$.

(b) Refinamento da solução:

Tabela 8: solução através do método da bisseção

k	a	b	x_k	$f(x_k)$	$b-a$	Conclusão
0	0	1	0.500	0.122	1	$\xi \in [0.000, 0.500]$
1	0	0.500	0.250	-0.469	0.500	$\xi \in [0.250, 0.500]$
2	0.250	0.500	0.375	-0.181	0.250	$\xi \in [0.375, 0.500]$
3	0.375	0.500	0.438	-0.031	0.125	$\xi \in [0.438, 0.500]$
4	0.438	0.500	0.469	0.045	0.063	$\xi \in [0.438, 0.469]$
5	0.438	0.469	0.453	0.007	0.031	$\xi \in [0.438, 0.453]$
6	0.438	0.453	0.445	-0.012	0.016	$\xi \in [0.445, 0.453]$
7	0.445	0.453	0.449	-0.002	0.008	Pare! pois $b-a < \varepsilon$ e $\lvert f(x_k)\rvert < \varepsilon$

Na iteração 7, tanto a amplitude do intervalo $[a;b]$ quanto a imagem, em módulo, de x_7 são menores que a precisão requerida, isto é, $b-a=0{,}453-0{,}445=0{,}008<\varepsilon=0{,}01$ e $|f(x_7)|=0{,}008<\varepsilon=0{,}01$. Desta forma, dizemos que $x_7=0{,}449$ é uma aproximação para a raiz ξ da equação $f(x)=2\mathrm{x}-\cos(x)=0$ com uma precisão $\varepsilon<0{,}01$.

7.4.3 Vantagens e Desvantagens do Método da Bisseção

A maior vantagem do Método da Bisseção é que, para sua convergência, não há exigências com relação ao comportamento do gráfico de f no intervalo $[a,b]$. Entretanto, ele não é eficiente devido à sua convergência lenta. Pode ser observado que $f(x)$ não decresce monotonicamente. Isto decorre do fato de que na escolha de uma aproximação $x=\frac{a+b}{2}$ não se leva em consideração os valores da função nos extremos do intervalo. No pior caso, a raiz ξ está próxima a um extremo.

O Método da Bisseção é mais usado para reduzir o intervalo antes de usar um outro método de convergência mais rápida.

7.5 Exercícios

1) Isole graficamente as raízes das seguintes equações:
   (a) $f(x)=x^3-9x+3=0$
   (b) $f(x)=x+\ln(x)=0$
   (c) $f(x)=x\ln(x)-1=0$
   (d) $f(x)=x^3+2+10^x=0$
   (e) $f(x)=\sqrt{x}-5e^{-x}=0$
2) Usando uma linguagem de programação implemente o algoritmo de bisseção e encontre as raízes para as equações do exercício anterior.

AULA 8 - RESOLUÇÃO DE EQUAÇÕES ALGÉBRICAS E TRANSCENDENTAIS - NEWTON-RAPHSON

8.1 Objetivo:

Nesta aula continuaremos estudando os métodos da resolução de equações algébricas e transcendentais, mais especificamente o método de Newton-Raphson.

8.2 O Método de Newton-Raphson

Seja f uma função contínua em [a; b] tal que:

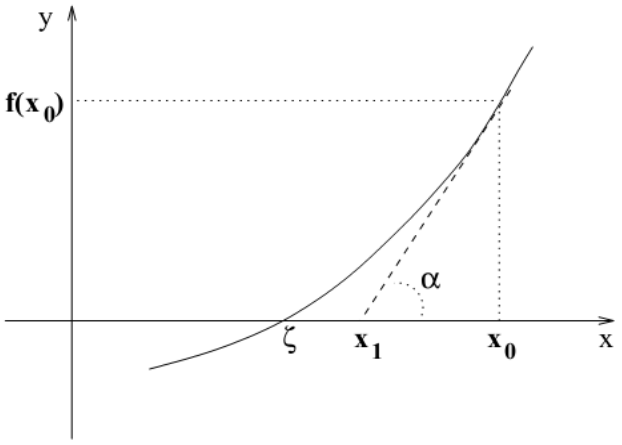
(i) $f(a)\times f(b)<0$

(ii) Existe uma única raiz $\xi\in[a\,;b]$

(iii) f' e f'' preservam o sinal e não se anulam em $[a\,;b]$

Um exemplo de uma função satisfazendo as condições acima é o da 11: A ideia do Método de Newton-Raphson é a de aproximar um arco da curva por um reta tangente traçada a partir de um ponto da curva.

Figura 11: Interpretação geométrica do Método de Newton-Raphson



Seja $x_0\in[a\,,b]$ uma aproximação inicial para a raiz. A tangente de α na Figura 11 é:

$$\tan(\alpha)=\frac{f(x_0)}{x_0-x_1}=f'(x_0)$$

De onde resulta que:

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

De forma análoga obtemos x_2, que representa a interseção da reta tangente ao gráfico de f no ponto $(x_1, f(x1))$ com o eixo dos x:

$$\tan(\beta)=\frac{f(x_1)}{x_1-x_2}=f'(x_1)$$

Isto é:

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

Genericamente:

$$x_k=x_{k-1}-\frac{f(x_{k-1})}{f'(x_{k-1})}\ \forall\ k=1,2,... \tag{39}$$

8.2.1 Escolha da aproximação inicial

Teorema

Se $f(a)\times f(b)<0$ é f' e f'' forem não nulas e preservarem o sinal em $[a,b]$, então partindo-se de uma aproximação inicial $x_0\in[a,b]$ tal que $f(x_0)\times f''(x_0)<0$ é possível gerar, pelo Método de Newton, uma sequência de aproximações x_k que convirja para a raiz $\xi\ de\ f(x)=0$.

8.2.2 Exemplo

Determinar pelo Método de Newton-Raphson, com precisão $\varepsilon<0{,}01$ em um máximo de 10 iterações, a raiz da equação $f(x)=2x-\cos(x)=0$

(a) Isolamento

Já foi visto que $\xi\in[0,1]$.

(b) Determinação de x_0

$$f'(x)=2+sen(x)\Rightarrow f'(x)>0\ \forall\ x\in[0,1]$$
$$f''(x)=\cos(x)\Rightarrow f''(x)>0\ \forall\ x\in[0,1]$$

Sendo $f(0)=-1, f(1)=1{,}46\ e\ f''(x)>0$ então podemos tomar como aproximação inicial $x_0=1$, pois $f(1)\times f''(1)>0$.

(c) Refinamento

Tabela 9: Aproximação de Newton-Raphson

k	x_k	$f(x_k)$	$f'(x_k)$	$\lvert x_k - x_{k-1}\rvert$	Conclusão
0	1	1.460	2.841	-	
1	0.486	0.088	2.467	0.514	
2	0.450	0.001	2.435	0.036	
3	0.450	0.000	2.435	0.000	Pare! pois $\lvert f(x_3)\rvert < 0.01$ e $\lvert x_3 - x_2\rvert < \varepsilon$

Logo, $x_3 = 0{,}450$ é uma aproximação para a raiz ξ da equação $f(x) = 2x\cos(x) = 0$ com uma precisão $\varepsilon < 0{,}01$.

8.2.3 Algoritmo

Apresentamos a seguir o pseudocódigo do procedimento Newton-Raphson. Este procedimento considera os critérios $|f(x_k)| < \varepsilon$ e $|x_k - x_{k-1}| < \varepsilon$ como mecanismos de parada.

```
newton-raphson(Tol, MaxIter, x)
k ← 0;
d ← ∞ ;
enquanto ((|d| ≥ Tol ou |f(x)| ≥ Tol) e k ≤ MaxIter) faça
      d ← -f(x) / f´(x);
      x ← x+d;
      k ← k + 1;
fim-enquanto;

se (k ≤ MaxIter) então
      retorne x;
senão
      Imprima: Não houve convergência em MaxIter iterações
```

8.2.4 Vantagens e desvantagens do Método de Newton

O Método de Newton-Raphson tem convergência muito boa (quadrática). Entretanto, apresenta as seguintes desvantagens:

(i) Exige o cálculo e a análise do sinal de f' e f''

(ii) Se $f'(x_{k-1})$ for muito elevado a convergência será lenta

(iii) Se $f'(x_{k-1})$ for próximo de zero pode ocorrer overflow

Para contornar o item (i), o qual é necessário para a escolha da aproximação inicial, é comum apenas calcular-se o valor da função e o de sua derivada segunda nos extremos a e b, considerando para x_0 o extremo que satisfazer a condição $f(x_0) \times f''(x_0) > 0$. Para tanto, é importante que o intervalo $[a, b]$ considerado seja suficientemente pequeno, de forma a minimizar a possibilidade de variação de sinal de f' e f''

8.3 Exercícios

1) Usando uma linguagem de programação implemente o algoritmo de Newton-Raphson e encontre as raízes para as equações a seguir.
   (a) $f(x)=x^3-9x+3=0$
   (b) $f(x)=x+\ln(x)=0$
   (c) $f(x)=x\ln(x)-1=0$
   (d) $f(x)=x^3+2+10^x=0$
   (e) $f(x)=\sqrt{x}-5e^{-x}=0$

2) Faça uma pesquisa e descubra como funciona o método da falsa posição para resolução de equações algébricas e transcendentais, em seguida implemente este método e faça alguns testes.

AULA 9 - LISTA DE EXERCÍCIOS 1

9.1 Objetivo:

Nesta aula os alunos realizarão vários exercícios com o objetivo de aprofundar os conhecimentos adquiridos nas aulas anteriores. Esta aula também serve para que as dúvidas que ainda perduram sejam resolvidas bem como para que os alunos se preparem para a avaliação que se segue.

9.2 Exercícios

1) Utilize eliminação de Gauss para resolver o seguinte sistema:

$$\begin{aligned} 10x_1 + 2x_2 - x_3 &= 27 \\ -3x_1 - 6x_2 + 2x_3 &= -61{,}5 \\ x_1 + x_2 + 5x_3 &= -21{,}5 \end{aligned}$$

Mostre todos os passos dos cálculos.

2) Dado o sistema linear Ax = b abaixo, pede-se:

$$\begin{bmatrix} 7 & -4 \\ 10 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

(a) Determinar a matriz de iteração de Jacobi.
(b) Determinar a matriz de iteração de Gauss-Seidel.
(c) Podemos resolver o sistema dado pelo Método de Jacobi? Justifique.
(d) Podemos resolver o sistema dado pelo Método de Gauss-Seidel? Justifique.
(e) Nesse exemplo, qual dos dois métodos iterativos convergirá mais rapidamente para a solução do sistema? Justifique.

3) Faça uma pesquisa sobre o método de resolução de sistemas lineares da decomposição LU, implemente este método em um programa e teste seu funcionamento.

4) Use a implementação do método de Gauss-Seidel para resolver o seguinte sistema, com tolerância $\varepsilon = 0{,}01$. Se necessário, reorganize as equações para garantir a convergência.

$$\begin{aligned} -3x_1 + x_2 + 12x_3 &= 50 \\ 6x_1 - x_2 - x_3 &= 3 \\ 6x_1 + 9x_2 + x_3 &= 40 \end{aligned}$$

5) Determine a raiz real para $f(x) = 5x^3 - 5x^2 + 6x - 2$:
(a) Graficamente.
(b) Utilizando o método da bisseção para encontrar a raiz, no intervalo $[0,1]$ itere até que o erro seja menor que 10%. Anote todos os passos.

AULA 10 - INTERPOLAÇÃO POLINOMIAL E LAGRANGE

10.1 Objetivo:

Nesta aula pretende-se ensinar aos alunos o que é a interpolação, bem como alguns métodos de realizá-la. Para tanto, nesta aula serão estudados os métodos de interpolação polinomial e de Lagrange.

10.2 Interpolação

O problema de interpolação surge quando se deseja aproximar uma função $f(x)$ por outra $g(x)$. A função $f(x)$ é de difícil manuseio ou avaliação. Normalmente desconhece-se a sua forma analítica mas o que se sabe sobre ela é um pequeno conjunto de entradas e saídas $(x_i, f(x_i))$, os quais são chamados de pontos bases. Já a função $g(x)$ possui um tratamento mais simples. Ela é chamada de função interpolante. Ela consiste na combinação linear de funções simples, as quais são:

1. Monômios $(x^k, onde\, k = 0,1,2,...,n)$. Ex.: $f(x) \approx g(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n$

2. Funções trigonométricas $(\sin kx\, e \cos kx, com\, k = 0,1,2,...,n)$. Ex.:
$$f(x) \approx g(x) = a_0 + a_1 \cos x + a_2 \cos 2x + ... + a_n \cos nx + b_1 \sin x + b_2 \sin 2x + ... + b_n \sin nk$$

3. Funções exponenciais $(a_k e^{b_k x}, com\, K = 0,1,2,...,n)$. Ex.:
$$f(x) \approx g(x) = a_0 e^{b_0 x} + a_1 e^{b_1 x} + ... + a_n e^{b_n x}$$

Para determinar $g(x)$ completamente, deve-se determinar os coeficientes presentes nessas expressões. Para esta aula, as funções interpolantes que serão utilizadas são os polinômios (que utilizam os monômios).

Observação: Em uma interpolação, deseja-se descobrir uma função $g(x)$, tal que seja possível determinar o valor de $g(\hat{x})$, sendo que $\hat{x}$ pertença ao intervalo fechado $[x_0, x_n]$ mas não esteja presente na tabela.

Exemplo: Número de habitantes

Tabela 10: Exemplo de interpolação

Ano	Habitantes
1950	325724
1960	638908
1970	1235030
1980	1814990

Caso deseje-se saber o número de habitantes no ano de 1975, isso seria um problema de interpolação, uma vez que $1975 \in [1950,1980]$. Caso deseje-se conhecer qual a população em 1983, o problema passaria a ser de extrapolação, que não será objeto de estudo nesta aula.

10.3 Interpolação Polinomial

Dado que $f(x) \approx g(x)$, onde:

$$f(x) \approx g(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n$$

e estando disponíveis apenas os pontos bases $(x_i, f(x_i))$, com $i = 0,1,2,...,n$, então pode-se montar o seguinte sistema linear:

$$\begin{aligned} f(x_0) &= a_0 + a_1 x_0 + a_2 x_0^2 + ... + a_n x_0^n \\ f(x_1) &= a_0 + a_1 x_1 + a_2 x_1^2 + ... + a_n x_1^n \\ f(x_2) &= a_0 + a_1 x_2 + a_2 x_2^2 + ... + a_n x_2^n \\ &\vdots \\ f(x_n) &= a_0 + a_1 x_n + a_2 x_n^2 + ... + a_n x_n^n \end{aligned}$$

que na forma matricial fica igual a:

$$\begin{bmatrix} 1 & x_0 & x_0^2 & ... & x_0^n \\ 1 & x_1 & x_1^2 & ... & x_1^n \\ 1 & x_2 & x_2^2 & ... & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & ... & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix}$$

que é um sistema linear de $(n+1)$ equações e $(n+1)$ incógnitas.

A matriz dos coeficientes é chamada de matriz de Vandermonde de ordem $(n+1)$ pois possui a seguinte propriedade:

$$det(A) \neq 0 \Rightarrow x_i \neq x_j, \text{ para todo } i \neq j$$

Há um teorema em interpolação polinomial que diz que, dados $(n+1)$ pontos bases distintos, existe uma única interpolante polinomial de ordem n, chamado de polinômio interpolador que satisfaz o sistema.

Exemplo:

Seja a seguinte tabela de valores da função $f(x) = e^x$, a partir da qual deseja-se obter uma aproximação para o ponto $x = 1{,}32$:

Tabela 11: Exemplo de interpolação

x	1,3	1,4	1,5
e^x	3,669	4,055	4,482

1. Determine o grau do polinômio interpolador: Como se conhecem os valores da função em três pontos, pode-se utilizar um polinômio de 2° grau.

$$p(x)=a_0+a_1x+a_2x^2$$

2. Construção do sistema: O sistema para a solução do problema é do tipo:

$$\begin{aligned} a_0 \quad a_1x_0 \quad a_2x_0^2 &= f(x_0) \\ a_0 \quad a_1x_1 \quad a_2x_1^2 &= f(x_1) \\ a_0 \quad a_1x_2 \quad a_2x_2^2 &= f(x_2) \end{aligned}$$

no qual substituímos os valores da tabela:

$$\begin{aligned} a_0 \quad a_1 1{,}3 \quad a_2 1{,}3^2 &= 3{,}669 \\ a_0 \quad a_1 1{,}4 \quad a_2 1{,}4^2 &= 4{,}055 \\ a_0 \quad a_1 1{,}5 \quad a_2 1{,}5^2 &= 4{,}482 \end{aligned}$$

3. Solução do sistema: Através de um método de solução de sistemas lineares (triangulação de Gauss por exemplo) Resolve-se o sistema, obtendo-se:

$$\begin{aligned} a_0&=2{,}383 \\ a_1&=-1{,}675 \\ a_2&=2{,}05 \end{aligned}$$

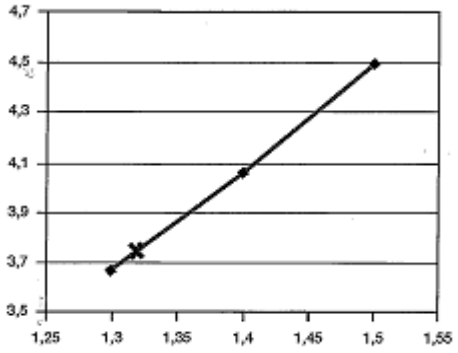
portanto, $p(x)=2{,}383-1{,}675x+2{,}05x^2$

4. Cálculo da interpolação: Substituindo o valor para o ponto $x=1{,}32$, obtêm-se:

$$\begin{aligned} p(1{,}32)&=2{,}383-1{,}675\times1{,}32+2{,}05\times1{,}32^2 \\ p(1{,}32)&=3{,}744 \end{aligned}$$

Deve-se levar em consideração que esse valor de $p(x)$ é uma aproximação, não tendo sido obtido através da função original $f(x)$. Como, neste caso, $f(x)$ é conhecida, pode-se avaliar qual foi o erro introduzido devido ao processo de interpolação.

Figura 12: Pontos da interpolação



10.4 Interpolação de Lagrange

A forma de Lagrange representa o polinômio interpretador diretamente a partir dos pontos originais. Sua praticidade é tal que se torna uma das formas mais utilizadas para a obtenção de um polinômio interpolador.

Seja f uma função tabelada em $(n+1)$ pontos distintos $\{x_0, x_1, \dots, x_n\} \in \mathbb{R}$ e sejam os polinômios de grau n dados pela forma genérica:

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

onde

$$L_i(x) = \frac{(x-x_0)(x-x_1)(x-x_2)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)(x_i-x_1)(x_i-x_2)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)}$$

são denominados polinômios de Lagrange.

Nota-se que nos polinômios de Lagrange, não são inseridos os fatores $(x-x_i)$ e (x_i-x_i), o que resultaria em um denominador nulo.

Portanto, utilizando-se esses polinômios, pode-se determinar o polinômio interpolador de f relativamente aos pontos $x_0, x_1, \dots, x_n$ da seguinte forma:

$$p(x) = L_0(x) f(x_0) + L_1(x) f(x_1) + \cdots + L_n(x) f(x_n)$$

Os polinômios $L_i(x)$ satisfazem as condições:

a) Contém os pontos da tabela, pois $p(x_i) = f(x_i)$, para $i = 0, 1, 2, \dots, n$.

b) O grau de $p(x)$ é menor ou igual a n assim o polinômio obtido é o polinômio interpolador da tabela na forma de Lagrange.

Para melhor ilustrar o formalismo, considere a tabela:

Tabela 12: Lagrange

x	x_0	x_1	x_2	x_3
$f(x)$	y_0	y_1	y_2	y_3

O polinômio de Lagrange será:

$$p(x) = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} \cdot f(x_0) + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}$$
$$\cdot f(x_1) + \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} \cdot f(x_2) + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)} \cdot f(x_3)$$

Exemplo: Seja a seguinte tabela de valores da função $f(x) = e^x$, a partir da qual se deseja obter

através de polinômios de Lagrange uma aproximação para o ponto $x=1{,}32$:

Tabela 13: Valores da função

x	1,3	1,4	1,5
e^x	3,669	4,055	4,482

1. Determinação dos polinômios de Lagrange: Uma vez que a tabela apresenta três pontos, deseja-se obter polinômios de Lagrange de 2° grau, cujas formas são:

$$L_0(x)=\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}$$
$$L_1(x)=\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}$$
$$L_2(x)=\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$$

E o polinômio interpolador de Lagrange é dado por:

$$p(x)=L_0(x)f(x_0)+L_1(x)f(x_1)+L_2(x)f(x_2)$$

2. Substituição pelos valores da tabela:

$$L_0(x)=\frac{(x-1{,}4)(x-1{,}5)}{(1{,}3-1{,}4)(1{,}3-1{,}5)} = \frac{(x-1{,}4)(x-1{,}5)}{0{,}02}$$
$$L_1(x)=\frac{(x-1{,}3)(x-1{,}5)}{(1{,}4-1{,}3)(1{,}4-1{,}5)} = \frac{(x-1{,}4)(x-1{,}5)}{-0{,}01}$$
$$L_2(x)=\frac{(x-1{,}3)(x-1{,}4)}{(1{,}5-1{,}3)(1{,}5-1{,}4)} = \frac{(x-1{,}4)(x-1{,}5)}{0{,}02}$$

3. Interpolar na tabela $x=1{,}32$ significa calcular:

$$p(1{,}32)=L_0(1{,}32)f(x_0)+L_1(1{,}32)f(x_1)+L_2(1{,}32)f(x_2)$$

ou

$$p(1{,}32)=\frac{(1{,}32-1{,}4)(1{,}32-1{,}5)}{0{,}02}3{,}669+\frac{(1{,}32-1{,}3)(1{,}32-1{,}5)}{-0{,}01}4{,}055$$
$$+\frac{(1{,}32-1{,}3)(1{,}32-1{,}4)}{0{,}02}4{,}482$$

O que resulta em:

$$p(1{,}32)=3{,}743$$

10.4.1 Algoritmo

Apresentamos a seguir o pseudocódigo do método de Lagrange.

```
Lagrange(x, n, X, Y)
p ← 0;
para i de 1 até n faça
      num ← 1;
      den ← 1;
      para j de 1 até n faça
            se (j ≠ i)
                  num ← num * (x – X[j]);
                  den ← den * (X[i] - X[j]);
            fim-se;
      fim-para;
      p ← p + num*Y[i]/den;
fim-para;
retorne p;
```

10.5 Exercícios

1) Dada a tabela a seguir, obter $f(40°)$ por interpolação de um polinômio de 3° grau.

x	30°	35°	45°	50°
$f\,x$	0,5	0,57358	0,70711	0,76604

2) Faça um programa implementando o algoritmo do método de Lagrange e teste seu funcionamento.

3) Através do polinômio interpolador de Lagrange, para os pontos da tabela calcule a aproximação de $f(x)$ para $x=2$ e $x=5{,}2$. Mostre todos os passos do cálculo.

x	0	1,5	3	4,5	6
$f\,x$	2,0	3,54	2,5	1,6	0,3

AULA 11 - AJUSTE DE CURVAS: MÍNIMOS QUADRADOS

11.1 Objetivo:

Nesta aula pretende-se ensinar aos alunos o que é o ajuste de curvas. Para isso, nesta aula será estudado o método dos mínimos quadrados

11.2 Introdução

O objetivo desta aula é obter uma função que se aproxime de um conjunto de pontos dados ou de outra função dada. Esse tipo de procedimento se faz necessário quando os pontos dados devem sofrer um processo de interpolação ou seu comportamento deve ser usado em outros cálculos. Muitas vezes o uso de uma função complexa, com cálculo lento e complicado, pode ser evitado se for utilizada uma outra função que possa substituí-la, dentro de uma determinada margem de erro, em um determinado trecho.

Também é bastante comum em engenharia a realização de testes de laboratório para a validação de sistemas reais. Os resultados são obtidos na forma de pontos cujo comportamento demonstra o relacionamento de uma variável independente com uma, ou mais, variáveis dependentes. Nestes casos é pouco provável que haja uma curva que passe exatamente por cada ponto e que descreva fielmente o sistema observado em laboratório.

O método dos Mínimos Quadrados, é o mais utilizado, nestes casos, por ter uma abordagem simples, ser preciso, e seu resultado poder abranger várias famílias de funções, principalmente os polinômios.

11.3 Método dos mínimos quadrados

O objetivo desse método é encontrar uma função $g(x)$ que mais se aproxime de outra função $f(x)$. Essa substituição pode ser necessária se:

- A função $f(x)$ descreve um fenômeno real, e deseja-se encontrar uma outra função, $g(x)$, que melhore a aproximação, mas ainda represente o comportamento do fenômeno.
- Tem-se uma função $f(x)$, mas há a necessidade de outra função, $g(x)$, que facilite o modelamento do processo.

A aproximação é feita da seguinte maneira:

Seja $f(x)$ a função original, $g(x)$ a função que irá aproximar $f(x)$, e $r(x)$ a função erro, que exprime as diferenças entre $f(x)$ e $g(x)$. Assim:

$$r(x)=f(x)-g(x)$$

Teoricamente, a melhor aproximação será aquela em que $r(x)=0$, ou, como na maioria

dos casos de aplicação desse método,

$$min\left[\sum_{x} r(x)\right] \tag{40}$$

Suponhamos que um experimento levantou os pontos $p_1, p_2, p_3, e\, p_4$. Suponhamos também que seja conhecido que o comportamento do fenômeno seja descrito por uma reta. O objetivo do método é, portanto, obter $g(x)$ tal que se comporte como uma reta, e que se aproxime ao máximo de $f(x)=p_x=\{p_1, p_2, p_3, p_4\}$.

Se aplicarmos a expressão 40, como alguns erros serão positivos e outros negativos, poderemos ter uma impressão errada das retas escolhidas, e talvez a pior reta seja aquela que apresenta o menor valor de $r(x)$.

Temos portanto, que descobrir o sinal dos erros. O uso do valor absoluto dos erros, $|r(x)|$ não é conveniente, pois a função módulo apresenta um comportamento difícil de trabalhar, principalmente na região em que $|x|$ esta próximo de zero. Um aboa escolha é elevar o erro ao quadrado. Assim, o novo critério para reduzir o erro será:

$$min\left[\sum_{x} r^2(x)\right] \tag{41}$$

que explica o nome dado ao método.

11.4 Regressão linear

Quando a função aproximadora, $g(x)$, deve ser uma reta, o Método dos Mínimos Quadrados reduz-se a uma "Regressão Linear". Esse termo é largamente usado em Engenharia, Economia, e outras áreas do conhecimento, sempre que um conjunto de pontos necessita ser aproximado por uma reta. Assim, o objetivo da Regressão Linear é aproximar uma função $f(x)$ por uma outra função, $g(x)$, da família $g(x)=a+bx$, usando o Método dos Mínimos Quadrados.

Isso significa que se deseja determinar os parâmetros a e b da reta $g(x)=a+bx$ de modo que a soma dos quadrados dos erros em cada ponto de $f(x)$ seja a menor possível, como determinado pela expressão (41).

Isso é feito através de um sistema linear dado por

$$\begin{bmatrix} \sum_{i=1}^{n} 1 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix} \tag{42}$$

Assim uma vez de posse dos valores dos pontos que compõem $f(x)=\{x_1, x_2, x_3, \ldots, x_n\}$ pode-se resolver o sistema linear e obter a e b.

Exemplo 1

Como resultado de um experimento prático, foram obtidos os seguintes valores para a função $f(x)$:

x	0	1	2	3	4
$f(x)$	0	1	1	4	4

Para determinar qual é a melhor reta que se ajusta a esses pontos deve-se achar $g(x)=a+b\,x.$ Para tanto, reorganiza-se os valores de forma a satisfazer a expressão (42).

$$n=5$$
$$x_1=0\ x_2=1\ x_3=2\ x_4=3\ x_5=4$$
$$y_1=0\ y_2=1\ y_3=1\ y_4=4\ y_5=4$$

Pode-se agora calcular os termos que serão usados no sistema linear:

$$\sum_{i=1}^{5} 1=1+1+1+1+1=5$$

$$\sum_{i=1}^{5} x_i=x_1+x_2+x_3+x_4+x_5=10$$

$$\sum_{i=1}^{5} x_i^2=x_1^2+x_2^2+x_3^2+x_4^2+x_5^2=0^2+1^2+2^2+3^2+4^2=30$$

$$\sum_{i=1}^{5} y_i=y_1+y_2+y_3+y_4+y_5=0+1+1+4+4=10$$

$$\sum_{i=1}^{5} x_i\,y_i=x_1\,y_1+x_2\,y_2+x_3\,y_3+x_4\,y_4+x_5\,y_5=0\cdot 0+1\cdot 1+2\cdot 1+3\cdot 4+4\cdot 4=31$$

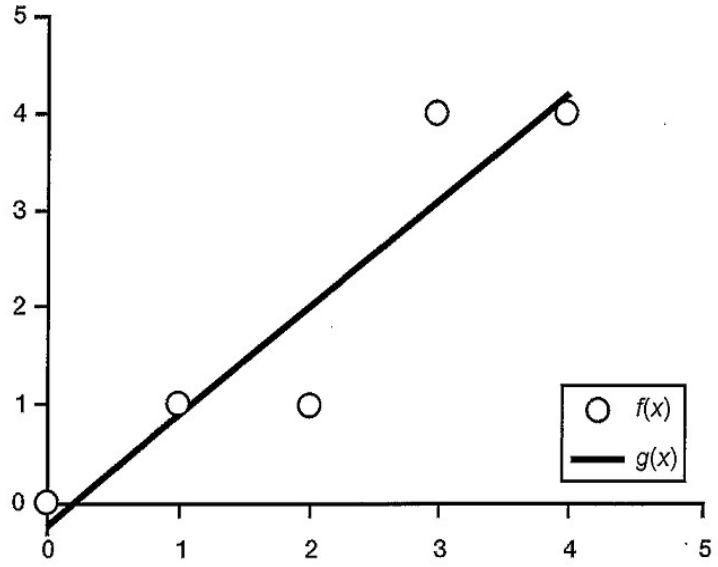
Agora, o sistema pode ser montado:

$$\begin{bmatrix} 5 & 10 \\ 10 & 30 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}=\begin{bmatrix} 10 \\ 31 \end{bmatrix}$$

Resolvendo o sistema, obtêm-se os valores $a=-1/5$ e $b=11/10.$ logo a função $g(x)$ pode ser expressa por:

$$g(x)=-\frac{1}{5}+\frac{11}{10}x$$

Figura 13: Resultados da regressão linear do exemplo



Exemplo 2

Seja a tabela relacionada aos dados experimentais de magnetização de uma barra de ferro:

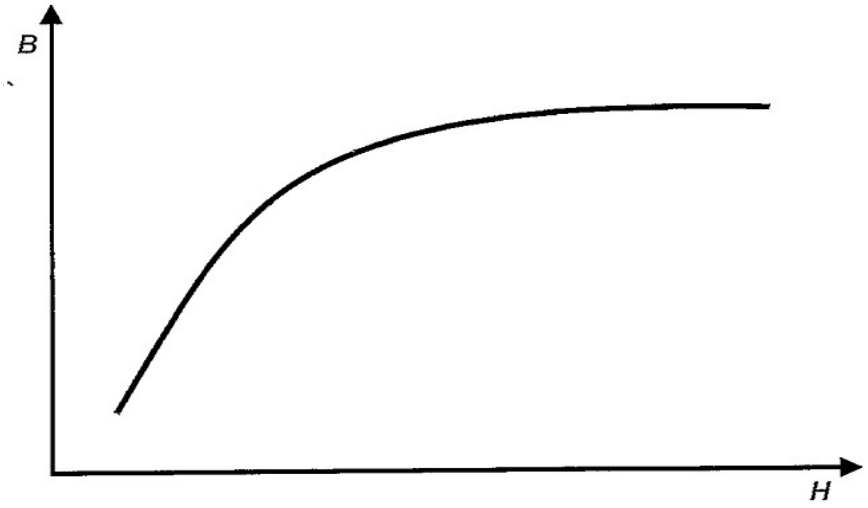
H	3,8	7,0	9,5	11,3	17,5	31,5	45,0	64,0	95,0
B	10,0	12,5	13,4	14,0	15,0	16,0	16,5	17,0	17,5

onde H é o campo magnético e B é a densidade de fluxo magnético induzido por H.

Sabemos que a densidade de fluxo magnético induzido tem um comportamento semelhante ao representado na 14, e que pode ser aproximado por:

$$B=\frac{H}{a+bH}$$

Figura 14: Comportamento de B em função de H



onde temos que determinar a e b.

Como a função não é linear nesses parâmetros, temos que transformar a função para poder aplicar a regressão linear. Uma maneira simples de linearizar essa expressão é fazer:

$$\frac{H}{B}=a+bH$$

obtendo-se assim uma função linear, ou seja, uma reta na qual $H\rightarrow x$ e $(H/B)\rightarrow y.$ Como consequência desta transformação, deve-se reescrever a tabela com os novos valores:

x	3,8	7,0	9,5	11,3	17,5	31,5	45,0	64,0	95,0
y	0,380	0,560	0,704	0,807	1,167	1,969	2,727	3,765	5,429

Onde os valores de y foram obtidos fazendo $y=H/B.$

$$\sum_{i=1}^{9} i=9$$

$$\sum_{i=1}^{9} x_i=284{,}6$$

$$\sum_{i=1}^{9} x_i^2=16725{,}88$$

$$\sum_{i=1}^{9} y_i=17{,}508$$

$$\sum_{i=1}^{9} x_i y_i=983{,}047$$

o sistema fica:

$$\begin{bmatrix} 9 & 284{,}6 \\ 284{,}6 & 16725{,}88 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}=\begin{bmatrix} 17{,}508 \\ 983{,}047 \end{bmatrix}$$

que uma vez resolvido resulta em:

$$g(x)=0{,}174+0{,}056\,x$$

Deve-se agora retornar aos parâmetros originais, H e $B.$ Para tanto, substitui-se x por H e y por $H/B.$ O resultado é:

$$\frac{H}{B}=0{,}174+0{,}056\,H$$

o que pode ser reescrito na forma desejada como:

$$B=\frac{H}{0{,}174+0{,}056\,H}$$

11.5 Linearização de funções

Existem várias funções de uma variável que podem ser expressas na forma $y = ax + b$, apesar de originalmente não apresentar um formalismo linear.

As funções que podem passar por essa transformação, através de uma substituição de variáveis, são chamadas de "linearizáveis". A vantagem de linearizar vem do fato de que se pode, então, aplicar o Método de Regressão Linear (e outros métodos e propriedades aplicáveis a funções lineares).

A Tabela 14 apresenta algumas funções e as substituições de variáveis necessárias para a linearização. Após a solução via regressão linear, deve-se voltar às formas originais.

Tabela 14: Linearização de funções

Função original $y = f(x)$	**Forma linearizada** $Y = AX + B$	**Troca de variáveis**
$y = \frac{A}{x} + B$	$y = A\frac{1}{x} + B$	$X = \frac{1}{x}, \quad Y = y$
$y = \frac{A}{x + B}$	$y = \frac{-1}{B}(xy) + \frac{A}{B}$	$X = xy, \ Y = y$
$y = \frac{1}{Ax + B}$	$\frac{1}{y} = Ax + B$	$X = x, \quad Y = \frac{1}{y}$
$y = A \ln x + B$	$y = A \ln x + B$	$X = \ln x, \ Y = y$
$y = Be^{Ax}$	$\ln y = Ax + \ln B$	$X = x, \ Y = \ln y$
$y = Bx^{A}$	$\ln y = A \ln x + \ln B$	$X = \ln x, \ Y = \ln y$
$y = (Ax + B)^2$	$\sqrt{y} = Ax + B$	$X = x, \quad Y = \sqrt{y}$
$y = Bxe^{Ax}$	$\ln\frac{y}{x} = Ax + \ln B$	$X = x, \quad Y = \ln\frac{y}{x}$
$y = \frac{1}{1 + Be^{Ax}}$	$\ln\left(\frac{1}{y} - 1\right) = Ax + \ln B$	$X = x, \quad Y = \ln\left(\frac{1}{y} - 1\right)$

11.6 Exercícios

1) Ajuste uma reta aos pontos:

x	-1	0	1	2	3	4	5	6
y	10	9	7	5	4	3	0	-1

Apresente todos os cálculos.

2) Crie um algoritmo para o método estudado e implemente este algoritmo em uma linguagem de programação. Comprove se funcionamento.

AULA 12 - MÍNIMOS QUADRADOS: GENERALIZAÇÃO

12.1 Objetivo:

Nesta aula pretende-se continuar com o ajuste de curvas. Para isso, nesta aula continuaremos estudando o método dos mínimos quadrados de forma generalizada, não apenas para equações do primeiro grau.

12.2 Introdução

O Método de Regressão Linear visto na aula anterior é uma particularização do Método dos Mínimos Quadrados para o caso em que se deseja que a função aproximadora seja uma reta. Apesar de termos apresentado um exemplo no qual a função final não era uma reta, o processo de aproximação ainda teve como base uma função linear.

Entretanto, o Método dos Mínimos Quadrados permite ajustar outros tipos de funções. De fato, a forma geral das funções ajustadas pelo método é:

$$Y = ag_0(x) + bg_1(x) + cg_2(x) + ... \tag{43}$$

onde, para ajustar uma reta, fazemos $g_0(x) = 1, g_l(x) = x,$ e todos os outros termos são anulados. As funções $g_n(x)$ podem ser as mais variadas possíveis, desde que dependentes apenas de x.

12.3 Método dos mínimos quadrados

Objetivo: Ajustar os dados de uma tabela de pontos:

Tabela 15: Dados dos pontos

x	x_0	x_1	x_2	...	x_n
y	$f(x_0)$	$f(x_1)$	$f(x_2)$	...	$f(x_n)$

através de uma função aproximadora do tipo:

$$G(x) = a_0x_0(x) + a_1x_1(x) + a_2x_2(x) + ... + a_mx_m(x) \tag{44}$$

onde os parâmetros $a_0, a_1, a_2, ..., a_n,$ são calculados resolvendo-se um sistema linear de ordem $m+1,$ chamado de Sistema Normal, cujos coeficientes são dados a seguir. Como no ponto $x = x_0$ tanto $g(x)$ como $f(x)$ têm o mesmo valor, $f(x_0),$ podemos obter o valor de b:

$$\begin{bmatrix} \sum g_0 g_0 & \sum g_0 g_1 & \cdots & \sum g_0 g_m \\ \sum g_1 g_0 & \sum g_1 g_1 & \cdots & \sum g_1 g_m \\ \vdots & \vdots & \ddots & \vdots \\ \sum g_m g_0 & \sum g_m g_1 & \cdots & \sum g_m g_m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \sum g_0 f \\ \sum g_1 f \\ \vdots \\ \sum g_m f \end{bmatrix} \tag{45}$$

E possível representar o Sistema Normal de ordem $m+1$ na forma matricial:

$$A\,\bar{X} = B$$

ou

$$\sum_{j=0}^{m} A_{ij}\bar{X}_j = B_i$$

para $i = 0, 1, 2, \ldots, m,$ onde

$$\begin{gathered} A_{ij} = \sum g_i g_j = \sum_{k=0}^{n} g_i(x_k) g_j(x_k) \\ B_i = \sum g_i f = \sum_{k=0}^{n} g_i(x_k) f(x_k) \\ \bar{X}_j = a_j \end{gathered} \tag{46}$$

Seja um caso geral que necessite obter uma aproximação dos dados experimentais de uma tabela, por um polinômio do tipo:

$$y(x) = a_0 g_0(x) + a_1 g_1(x) + a_2 g_2(x) + \ldots + a_m g_m(x) \tag{47}$$

Então:

x	x_0	x_1	x_2	...	x_n
y	$f(x_0)$	$f(x_1)$	$f(x_2)$	...	$f(x_n)$

onde $f(x)$ é uma função desconhecida. De (47) deduz-se que $a_0, a_1, a_2, \ldots, a_n$ são números reais a determinar, e $g_0(x), g_1(x), g_2(x), \ldots, g_m(x)$ são funções conhecidas. O resíduo $r_i, i = 1, 2, 3, \ldots, n,$ calculado no ponto x_i da tabela, é definido por:

$$r_i = y(x_i) - f(x_i)$$

onde $y(x_i)$ é a função aproximadora. Deseja-se reduzir ao mínimo a soma dos resíduos quadráticos, ou seja, deseja-se minimizar a função:

$$S = \sum_{i=0}^{n} r_i^2 = \sum_{i=0}^{n} [y(x_i) - f(x_i)]^2$$

A partir daí, impõe-se que a derivada parcial em relação a cada parâmetro seja nula, pois estamos procurando o mínimo da função $S(a_j)$:

$$\frac{\partial S}{\partial a_j}=0, j=0,1,2,\dots,n$$

fornecendo $m+1$ equações lineares com $m+1$ parâmetros $a_0, a_1, a_2, \dots, a_m$ que são incógnitas do Sistema Normal expresso em (46).

Dispositivo prático

A montagem do Sistema Normal é facilitada com a construção de um quadro conforme o que segue:

Tabela 16: Dispositivo prático usado nos cálculos

i	x_i	$G_0(x_i)$	$G_1(x_i)$	...	$G_m(x_i)$	$f(x_i)$
0	x_0	$G_0(x_0)$	$G_1(x_0)$	...	$G_m(x_0)$	$f(x_0)$
1	x_1	$G_0(x_1)$	$G_1(x_1)$	...	$G_m(x_1)$	$f(x_1)$
2	x_2	$G_0(x_2)$	$G_1(x_2)$	...	$G_m(x_2)$	$f(x_2)$
⋮	⋮	⋮	⋮		⋮	⋮
n	x_n	$G_0(x_n)$	$G_1(x_n)$	...	$G_m(x_n)$	$f(x_n)$
		$\sum g_0^2$	$\sum g_0 g_1$	...	$\sum g_0 g_m$	$\sum g_0 f$
		$\sum g_1 g_0$	$\sum g_1^2$	...	$\sum g_1 g_m$	$\sum g_1 f$
		$\sum g_2 g_0$	$\sum g_2 g_1$	...	$\sum g_2 g_m$	$\sum g_2 f$
		⋮	⋮		⋮	⋮
		$\sum g_m g_0$	$\sum g_m g_1$	...	$\sum g_m^2$	$\sum g_m f$
		a_0	a_1	...	a_m	

Onde $\sum g_1 g_0$ significa:

$$\sum_{k=0}^{n} g_i(x_k) g_j(x_k)=g_0(x_0) g_1(x_0)+g_0(x_1) g_1(x_1)+\dots+g_0(x_n) g_1(x_n)$$

e $\sum g_0 f$ significa:

$$\sum_{k=0}^{n} g_i(x_k) f(x_k)=g_0(x_0) f(x_0)+g_0(x_1) f(x_1)+\dots+g_0(x_n) f(x_n)$$

Exemplo 1

Foram feitas cinco medições da velocidade de um carro de Fórmula 1 e da pressão do ar na superfície do aerofólio dianteiro. Nos dados da Tabela 17, apresenta-se como a velocidade V relaciona-se com a pressão P.

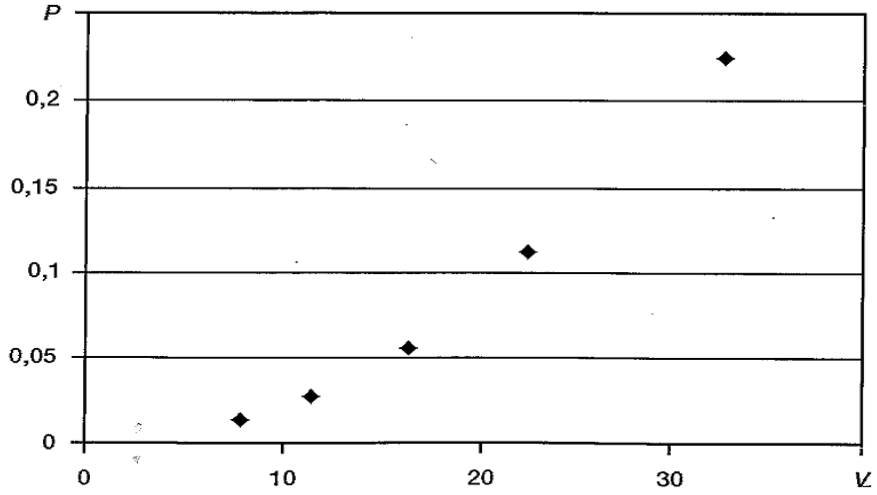
Tabela 17: Dados exemplo 1

V	7,87	11,50	16,40	22,60	32,80
P	0,014	0,028	0,056	0,112	0,225

Deseja-se ajustar esses dados por uma função aproximadora $P(V)$ que permita, mesmo sem conhecer a lei física que rege o fenômeno, calcular a pressão correspondente a uma dada velocidade. É sugerida a função $P=a+bV^2$. Logo, é necessário determinar os valores de a e b.

1. O primeiro passo é definir exatamente o que será feito.

Figura 15: Localização dos pontos



Foi solicitado que o ajuste seja feito através da função $P=a+bV^2$. Logo, os parâmetros a serem determinados são a e b, sendo:

$$P=1\cdot a+b\,V^2\Rightarrow\begin{Bmatrix} g_0(V)=1 \\ g_1(V)=V^2 \end{Bmatrix}$$

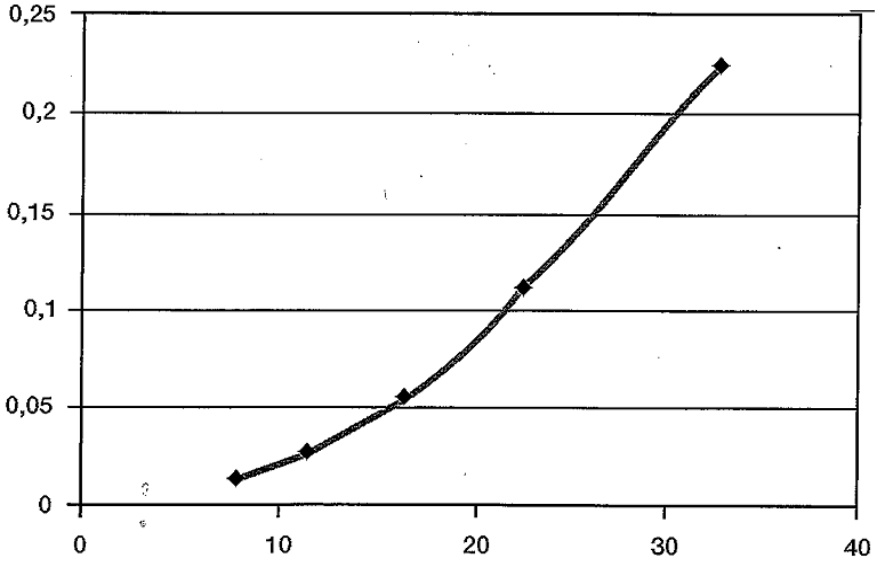
2. Em seguida é construído o dispositivo prático:

i	x_i	$G_0(V_i)$	$G_1(V_i)$	$P(V_i)$
0	7,87	1	$G_1(V_0)\,7{,}87^2=61{,}9369$	$f(V_0)=0{,}014$
1	11,50	1	$G_1(V_1)\,11{,}50^2=132{,}2500$	$f(V_1)=0{,}028$
2	16,40	1	$G_1(V_2)\,16{,}40^2=268{,}9600$	$f(V_2)=0{,}056$
3	22,60	1	$G_1(V_3)\,22{,}60^2=510{,}7600$	$f(V_3)=0{,}112$
4	32,80	1	$G_1(V_4)\,32{,}80^2=1075{,}840$	$f(V_4)=0{,}225$
		$\sum g_0^2=5$	$\sum g_0 g_1=2049{,}7469$	$\sum g_0 f=0{,}4350$
		$\sum g_1 g_0=2049{,}7469$	$\sum g_1^2=1511973{,}2070$	$\sum g_1 f=318{,}9010$
		a	b	

3. Agora, é possível escrever o Sistema Normal:

$$\begin{bmatrix} 5 & 2049{,}7469 \\ 2049{,}7469 & 1511973{,}2070 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0{,}4350 \\ 318{,}9010 \end{bmatrix}$$

Figura 16: Função aproximada



4. A última etapa é resolver o Sistema Normal, do qual obtém-se:

$$a = 0{,}001203$$
$$b = 0{,}000209$$

e finalmente:

$$P = 0{,}001203 + 0{,}000209\, V^2$$

12.4 Comentários sobre o método

Esse método soluciona o problema de obtenção de uma função aproximadora cuja expressão geral é:

$$y(x) = \sum_{k=1}^{n} a_k X_k(x)$$

onde $X_1(x), \ldots, X_m(x),$ podem ser quaisquer funções não-lineares de $x.$ A única referência à linearidade recai sobre os parâmetros multiplicativos $a_k.$ Supondo-se que existam N valores $y_1(x), \ldots, y_N(x)$ que representam a tabela de pontos que se deseja ajustar através da função aproximadora, pode-se estabelecer uma expressão que determina o erro (ou desvio) entre o valor obtido pela função aproximadora e o valor original da tabela de pontos. Para um ponto qualquer $i,$ o erro será, portanto:

$$\delta_i = \frac{y_i(x) - \sum_{k=1}^{n} a_k X_k(x)}{\sigma_i}$$

onde δ_i representa o desvio entre o valor calculado e o valor real, e σ_i é o erro (ou desvio-padrão) do i-ésimo ponto. Essa expressão é especialmente interessante em aplicações científicas ou de Engenharia. Se os valores de σ_i não forem conhecidos, pode-se adotar $\sigma_i = 1$.

Como o desvio pode ser positivo ou negativo (pois a função aproximadora pode resultar em valores, acima ou abaixo de y_i) sua soma pode ter parcelas que se anulem, o que não representaria o verdadeiro desvio global do ajuste. Assim, adota-se a expressão do desvio quadrático como sendo:

$$X^2 = \sum_{i=1}^{N} \delta_i$$

logo,

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{y_i(x) - \sum_{k=1}^{n} a_k X_k(x)}{\sigma_i} \right]^2$$

A função χ^2 é conhecida na bibliografia como o desvio global de uma função com relação a outra (em português, o som da letra grega χ pronuncia-se "qui", e a função χ^2 como "qui-quadrado").

Como se deseja que χ^2 seja o mais próximo possível de zero, deve-se proceder à minimização da função χ^2, ou seja, a busca dos mínimos quadrados. Isto é, deseja-se

$$\sum_{i=1}^{N} \frac{1}{\sigma_1^2} \left[y_i - \sum_{i=1}^{N} a_j X_j(x_i) \right] X_k(x_i) = 0$$

Para $k = 1, \ldots, M$.

Essa expressão leva à construção da matriz do Sistema Normal (expressão (45)), que, por sua vez, é resolvida pelo método apresentado.

É importante salientar que, se os pontos a serem ajustados apresentarem descontinuidades ou a função aproximadora eleita não for uma boa escolha para o ajuste, o Sistema Normal poderá não ter solução. Essa condição é reconhecível no caso de os elementos da diagonal principal do Sistema Normal serem nulos ou próximos de zero. Assim, torna-se imprescindível uma análise do problema, que pode ser feita através de um estudo da distribuição dos pontos experimentais no plano cartesiano e de esboços da função aproximadora.

12.5 Exercícios

1) Aplicando o método dos mínimos quadrados, ache o polinômio do 2° grau $y=a_0+a_1 x+a_2 x^2$ que melhor se ajuste aos pontos da tabela a seguir.

x	0,78	1,56	2,34	3,12	3,81
y	2,50	1,20	1,12	2,25	4,28

Apresente todos os cálculos e analise o resultado graficamente.

2) Crie um algoritmo para o método estudado e implemente este algoritmo em uma linguagem de programação. Comprove se funcionamento.

AULA 13 - INTEGRAÇÃO NUMÉRICA: TRAPÉZIOS

13.1 Objetivo:

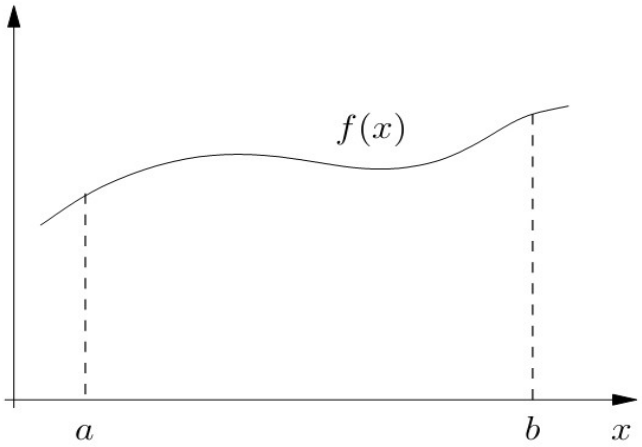
Nesta aula pretende-se estudar a integração numérica, bem como a regra dos trapézios, com o objetivo de demostrar aos alunos como é possível determinar o valor da integral definida de uma determinada função ou conjunto de pontos em um intervalo definido.

13.2 Introdução

O conceito de integral está ligado ao problema de determinar a área de uma figura plana qualquer. A definição da área de uma figura plana é feita aproximando a figura por polígonos cujas áreas podem ser calculadas pelos métodos de Geometria Elementar.

Considerando a definição da área da figura delimitada por uma função $f(x)$, pelo eixo das abscissas x e por duas retas $x=a$ e $x=b$, como ilustrado pela 17.

Figura 17: Área de uma função f(x).



Dividindo este intervalo $[a,b]$ em n subintervalos iguais de comprimento:

$$\Delta x = \frac{b-a}{n}$$

onde $x_0 = a < x_1 < x_2 < ... < x_n = b$ são os pontos dessa divisão. Em cada um desses intervalos, definem-se os pontos ξ_1 no primeiro, ξ_2 no segundo intervalo, e assim sucessivamente até ξ_n no último intervalo. Dessa forma, é possível definir uma série de retângulos de base Δx e altura $f(\xi_1), (\xi_2), \ldots, (\xi_n)$ (ver 18).

A soma das áreas dos retângulos é:

$$S_n = \sum_{i=1}^{n} f(\xi_i)\Delta x$$

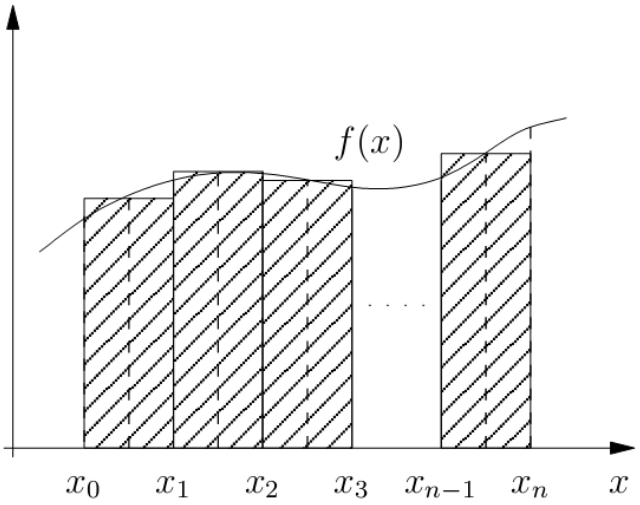


Figura 18: Retângulos definidos nos subintervalos de $[x_0=a, x_n=b]$.

Nota-se que este valor S_n é, aproximadamente, o valor da área delimitada por $f(x)$ e x, no intervalo $[a,b]$. Se a quantidade de subintervalos cresce tendendo ao infinito, então obtém-se o conceito de integral:

$$S_n=\lim_{n\to\infty}\sum_{i=1}^{n} f(\xi_i)\Delta x=\int_a^b f(x)\,dx$$

que é chamada de Integral de Riemann. O seu resultado é um valor numérico. Embora haja um conjunto de regras para calcular a chamada função primitiva $F(x)$, ou seja:

$$F(x)=\int f(x)\,dx$$

em determinados casos, esta função primitiva não é conhecida, ou a sua obtenção não é trivial. Além disso, em situações práticas nem sempre se tem a forma analítica da função a ser integrada, $f(x)$, mas é disponibilizada uma tabela de pontos que descreve o comportamento da função.

Assim, para calcular o valor da integral de $f(x)$ considerando estes casos particulares, torna-se necessário a utilização de métodos numéricos.

13.3 Regra dos Trapézios

Na Regra dos Trapézios utilizam-se apenas duas abscissas separadas por uma distância h. Assim, utiliza-se um polinômio interpolador de primeiro grau. Utilizando a fórmula de Lagrange para expressar o polinômio $P_1(x)$ que interpola $f(x)$ em x_0 e x_1 tem-se:

$$f(x) \simeq b_0\, p_0(x) + b_1\, p_1(x)$$

onde:

$$p_0(x) = x - x_1$$
$$p_1(x) = x - x_0$$
$$b_0 = \frac{f(x_0)}{x_0 - x_1}$$
$$b_1 = \frac{f(x_1)}{x_1 - x_0}$$

Como tem-se apenas dois pontos, $x_0 = a$ e $x_1 = b$. Então $x_1 - x_0 = h$. Assim:

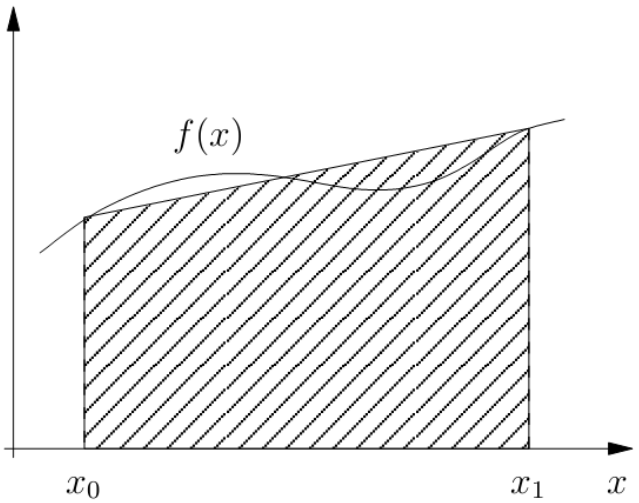
$$f(x) \simeq -\frac{f(x_0)}{h}(x - x_1) + \frac{x_1}{h}(x - x_0)$$

Integrando, no intervalo $[x_0, x_1]$, ambos os lados desta aproximação então obtém-se a fórmula geral para a Regra dos Trapézios:

$$\int_a^b f(x)\,dx \simeq \frac{h}{2}[f(x_0) + f(x_1)]$$

Esse resultado corresponde à área do trapézio de altura $h = x_1 - x_0$ e bases, $f(x_0)$ e $f(x_1)$, como ilustrado na 19.

Figura 19: Interpretação gráfica da regra dos trapézios.



E possível notar que, se o intervalo de integração é grande, a fórmula dos Trapézios fornece resultados que pouco tem a ver com o valor da integral exata. Para diminuir este erro é preciso

subdividir o intervalo de integração e aplicar a regra dos Trapézios repetidas vezes, para cada par subsequente de pontos. Chamando x_i os pontos de divisão de $[a,b]$, tal que $x_{i+1}-x_i=h$, sendo $i=0,1,2,\ldots,n-1$, tem-se:

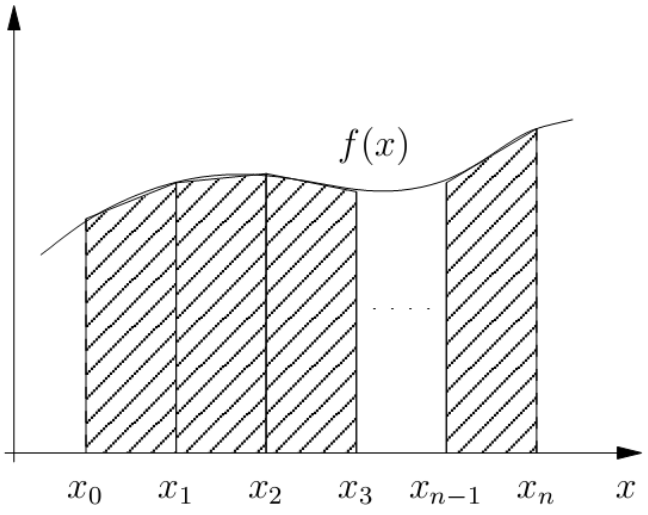
$$\int_a^b f(x)\,dx=\sum_{i=0}^{n-1}\int_{x_1}^{x_n-1} f(x)\,dx\simeq\sum_{i=0}^{n-1}\frac{h}{2}[f(x_i)+f(x_{i+1})]$$

ou, de uma forma mais simplificada:

$$\int_a^b f(x)\,dx\simeq\frac{h}{2}[f(x_0)+2f(x_1)+2f(x_2)+\ldots+2f(x_{n-1})+f(x_n)]$$

cuja interpretação geométrica está ilustrada na 20.

Figura 20: Interpretação gráfica da regra dos trapézios repetida.



Exemplo

Seja $I=\int_0^1 e^x\,dx$, calcular uma aproximação para I utilizando 10 subintervalos e a regra dos Trapézios repetida.

$$h=\frac{b-a}{n}=\frac{1-0}{10}=0{,}1$$

Desta forma, como $x_{i+1}=x_1+h$, então:

$$x_0 = 0,0$$
$$x_1 = 0,1$$
$$x_2 = 0,2$$
$$\vdots$$
$$x_9 = 0,9$$
$$x_{10} = 1,0$$

Assim:

$$I \simeq \frac{0,1}{2}[e^0 + 2e^{0,1} + 2e^{0,2} + ... + 2e^{0,9} + e^{1,0}] = 1,719713$$

13.4 Exercícios

1) Calcular os valores das integrais a seguir utilizando a Regra dos Trapézios. Comprove os resultados.

(a) $\int_0^1 \frac{\cos x}{x+1} dx$

(b) $\int_4^{4.5} \frac{1}{x^2} dx$

(c) $\int_3^6 (3x+2)\, dx$

2) Dada a função $y = f(x)$ através da tabela a seguir, calcular o valor de $I = \int_0^3 f(x)\, dx$ utilizando a Regra dos Trapézios.

i	x_i	y_i
0	0.0	5.021
1	0.5	6.146
2	1.0	6.630
3	1.5	6.940
4	2.0	7.178
5	2.5	7.364
6	3.0	7.519

3) Crie um algoritmo para o método estudado e implemente este algoritmo em uma linguagem de programação. Comprove se funcionamento.

AULA 14 - INTEGRAÇÃO NUMÉRICA: SIMPSON

14.1 Objetivo:

Nesta aula continuaremos estudando a integração numérica através da primeira e segunda regra de Simpson. Diferentemente do método dos trapézios, Simpson faz uso de funções de aproximação de segundo e terceiro grau.

14.2 Primeira Regra de Simpson

Esta primeira regra é obtida aproximando-se a função $f(x)$ por um polinômio interpolador de segundo grau. Para isto, serão necessário 3 pontos $(x_0 = a, x_1 e\, x_2 = b)$ igualmente espaçados.

$$f(x) \simeq b_0\, p_0(x) + b_1\, p_1(x) + b_2\, p_2(x)$$

onde os termos $p_i(x)$ são:

$$p_0(x) = (x - x_1)(x - x_2)$$
$$p_1(x) = (x - x_0)(x - x_2)$$
$$p_2(x) = (x - x_0)(x - x_1)$$

e os termos b_i são:

$$b_0 = \frac{f(x_0)}{(x_0 - x_1)(x_0 - x_2)}$$
$$b_1 = \frac{f(x_1)}{(x_1 - x_0)(x_1 - x_2)}$$
$$b_2 = \frac{f(x_2)}{(x_2 - x_0)(x_2 - x_1)}$$

Dessa forma, tem-se que a fórmula geral para a Primeira Regra de Simpson é obtida através de:

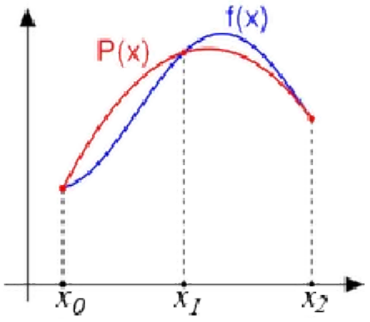
$$\int_{x_0}^{x_2} f(x)\,dx \simeq b_0 \int_{x_0}^{x_2} p_0(x)\,dx + b_1 \int_{x_0}^{x_2} p_1(x)\,dx + b_2 \int_{x_0}^{x_2} p_2(x)\,dx$$

Resolvendo estas integrais e, depois, substituindo $x_1 = x_0 + h\, e\, x_2 = x_0 + 2\mathrm{h}$, a fórmula geral fica igual a:

$$\int_{x_0}^{x_2} f(x)\,dx \simeq \frac{h}{3}[f(x_0) + 4 f(x_1) + f(x_2)]$$

cuja interpretação significa que os pontos $x_0, x_1 e\, x_2$ são interpolados pelo polinômio de Lagrange de segundo grau. A 21 apresenta graficamente este arranjo.

Figura 21: Interpretação gráfica da primeira regra de Simpson



Exemplo

Seja $f(x)$ uma função conhecida apenas nos pontos tabelados a seguir. Utilizando a primeira regra de Simpson, encontrar uma aproximação para $\int_2^4 f(x)dx$.

i	x_i	$f(x_i)$
0	2,0	41
1	3,0	130
2	4,0	297

Como, neste caso, o espaçamento h é igual a 1, então:

$$\int_2^4 f(x)dx \simeq \frac{1}{3}[f(2{,}0)+4f(3{,}0)+f(4{,}0)]$$
$$\simeq \frac{1}{3}[41+4\cdot 130+297]=286$$

Da mesma forma como foi realizado com a Regra dos Trapézios, deve-se subdividir o intervalo de integração $[a,b]$ em n subintervalos iguais de amplitude h e, a cada par de subintervalos, aplicar a Primeira Regra de Simpson. Uma observação importante é que o número de subintervalos deverá ser sempre par.

Assim, sendo $h=(b-a)/n$, os pontos serão $x_0=a, x_1, x_2, x_3, \ldots, x_n=b$. A aproximação da integral de uma função ficará:

$$\int_a^b f(x)dx \simeq \frac{h}{3}[f(x_0)+4f(x_1)+f(x_2)]+\frac{h}{3}[f(x_2)+4f(x_3)+f(x_4)]+\ldots+$$
$$\frac{h}{3}[f(x_{n-2})+4f(x_{n-1})+f(x_n)]$$

que de uma maneira mais simplificada será:

$$\int_a^b f(x)\,dx \simeq \frac{h}{3}[f(x_0)+f(x_n)+2(f(x_2)+f(x_4)+...+f(x_{n-2}))+4(f(x_1)+f(x_3)+...+f(x_{n-1}))]$$

Exemplo

Adicionando alguns pontos na tabela do exemplo anterior, tem-se:

i	x_i	$f(x_i)$
0	2,0	41
1	2,5	77,25
2	3,0	130
3	3,5	202,25
4	4,0	297

Recalcular a integral $\int_2^4 f(x)\,dx.$ utilizando a Primeira Regra de Simpson repetida.

Neste caso, o espaçamento $é\, h=0{,}5.$ Assim:

$$\int_2^4 f(x)\,dx \simeq \frac{0{,}5}{3}[f(x_0)+f(x_4)+2\,f(x_2)+4(f(x_1)+f(x_3))]$$
$$\simeq \frac{0{,}5}{3}[41+297+2\cdot 130+4(77{,}25+202{,}25)]=286$$

14.3 Segunda Regra de Simpson

De maneira análoga às anteriores, a Segunda Regra de Simpson é obtida aproximando-se a função $f(x)$ pelo polinômio interpolador de terceiro grau. Dessa forma, através da metodologia de Lagrange:

$$f(x)\simeq b_0\,p_0(x)+b_1\,p_1(x)+b_2\,p_2(x)+b_3\,p_3(x)$$

onde os termos $p_i(x)$ são:

$$p_0(x)=(x-x_1)(x-x_2)(x-x_3)$$
$$p_1(x)=(x-x_0)(x-x_2)(x-x_3)$$
$$p_2(x)=(x-x_0)(x-x_1)(x-x_3)$$
$$p_3(x)=(x-x_0)(x-x_1)(x-x_2)$$

e os termos b_i são:

$$b_0=\frac{f(x_0)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}$$
$$b_1=\frac{f(x_1)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}$$
$$b_2=\frac{f(x_2)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}$$
$$b_3=\frac{f(x_3)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}$$

Assim:

$$\int_a^b f(x)\,dx \simeq b_0\int_a^b p_0(x)\,dx + b_1\int_a^b p_1(x)\,dx + b_2\int_a^b p_2(x)\,dx + b_3\int_a^b p_3(x)\,dx \tag{48}$$

Como se utiliza um polinômio de terceiro grau, então são necessários quatro pontos a serem interpolados:

$$\begin{aligned} x_0&=a\\ x_1&=x_0+h\\ x_2&=x_0+2h\\ x_3&=x_0+3h=b \end{aligned} \tag{49}$$

Fazendo a integração indicada pela aproximação (48) e substituindo os termos dados pelas equações (49), então fica-se com a seguinte fórmula geral para a Segunda Regra de Simpson:

$$\int_{x_0=a}^{x_3=b} f(x)\,dx \simeq \frac{3\mathrm{h}}{8}\left[f(x_0)+3f(x_1)+3f(x_2)+f(x_3)\right]$$

Esta segunda regra também é conhecida como a Regra dos 3/8.

Subdividindo o intervalo $[a,b]$ em n subintervalos, onde n deverá ser múltiplo de 3, tem-se a seguinte fórmula para a aplicação repetida:

$$\int_{x_0}^{x_n} f(x)\,dx \simeq$$
$$\frac{3\mathrm{h}}{8}\left[f(x_0)+3f(x_1)+3f(x_2)+2f(x_3)+3f(x_4)+3f(x_5)+2f(x_6)+\ldots+3f(x_{n-2})+3f(x_{n-1})+f(x_n)\right]$$

Exemplo

Calcular o valor da integral $I=\int_1^4 \ln\left(x^3+\sqrt{e^x+1}\right)dx$ aplicando a regra dos 3/8 com 3 e 9 intervalos.

Para 3 intervalos tem-se que $h=\frac{4-1}{3}=1.$ assim:

$$x_0=1 \rightarrow f(1)=1{,}0744$$
$$x_1=1+1 \rightarrow f(2)=2{,}3884$$
$$x_2=2+1 \rightarrow f(3)=3{,}4529$$
$$x_3=3+1 \rightarrow f(4)=4{,}2691$$

Portanto:

$$I \simeq \frac{3\cdot 1}{8}[1{,}0744+3\cdot 2{,}3884+3\cdot 3{,}4529+4{,}2691]=8{,}5753$$

Para 9 intervalos, tem-se que:

$$h=\frac{4-1}{9}=\frac{3}{9}=\frac{1}{3}$$

Pode-se construir uma tabela utilizando este h e $x_o=a=1,$ resultando em:

i	x_i	$f(x_i)$
0	1	1.0744
1	4/3	1.5173
2	5/3	1.9655
3	2	2.3884
4	7/3	2.7768
5	8/3	3.1305
6	3	3.4529
7	10/3	3.7477
8	11/3	4.0187
9	4	4.2691

Aplicando a fórmula da segunda regra de Simpson repetida, tem-se:

$$\int_{x_0}^{x_n} f(x)\,dx \simeq$$

$$\frac{3\mathrm{h}}{8}[f(x_0)+3f(x_1)+3f(x_2)+2f(x_3)+3f(x_4)+3f(x_5)+2f(x_6)3f(x_7)+3f(x_8)+f(x_9)]$$

Por fim, substituindo os valores correspondentes, o resultado é $I=8.5619.$

14.4 Exercícios

1) Dada a função $y = f(x),$ definida através da tabela a seguir, calcular $\int_{1}^{1,6} f(x)\,dx$ aplicando:

(a) A primeira regra de Simpson.
(b) A segunda regra de Simpson.

i	x_i	y_i
0	1,0	0,099
1	1,1	0,131
2	1,2	0,163
3	1,3	0,194
4	1,4	0,2244
5	1,5	0,253
6	1,6	0,281

2) Crie um algoritmo para cada um dos métodos estudados e implemente estes algoritmos em uma linguagem de programação. Comprove se funcionamento.

AULA 15 - SOLUÇÃO DE EQUAÇÕES DIFERENCIAIS

15.1 Objetivo:

Nesta aula iremos estudar a solução numérica de equações diferenciais ordinárias. Para tanto serão estudados os métodos de Euler e Runge-Kutta.

15.2 Introdução

Uma equação diferencial ordinária (ou EDO) é uma equação que envolve as derivadas de uma função desconhecida de uma variável.

Uma equação diferencial ordinária (EDO) é uma equação que envolve $x, y, y', y'', \ldots$ A ordem de uma equação diferencial é a ordem n da maior derivada na equação. A solução de uma EDO é uma função $y(x)$ cujas derivadas satisfazem a equação. Não está garantido que tal função exista, e caso exista, normalmente ela não é única.

15.3 Método de Euler

O método de Euler, também conhecido como método da reta secante, é um dos métodos mais antigos que se conhece para a solução de equações diferenciais ordinárias. Problemas práticos não devem ser resolvidos com o método de Euler. Existem outros métodos que proporcionam resultados com uma melhor precisão e estabilidade se comparados ao método de Euler para o mesmo passo.

15.3.1 Derivação da Fórmula de Euler

Seja uma função $\frac{dy}{dx} = f(x, y)$ com a condição de contorno $y = y_n$ quando $x = x_n$. Da (50) observa-se que o valor de y_{n+1}, em $x = x_{n+1}$ é dado por:

$$y_{n+1} = y_n + \Delta y \tag{50}$$

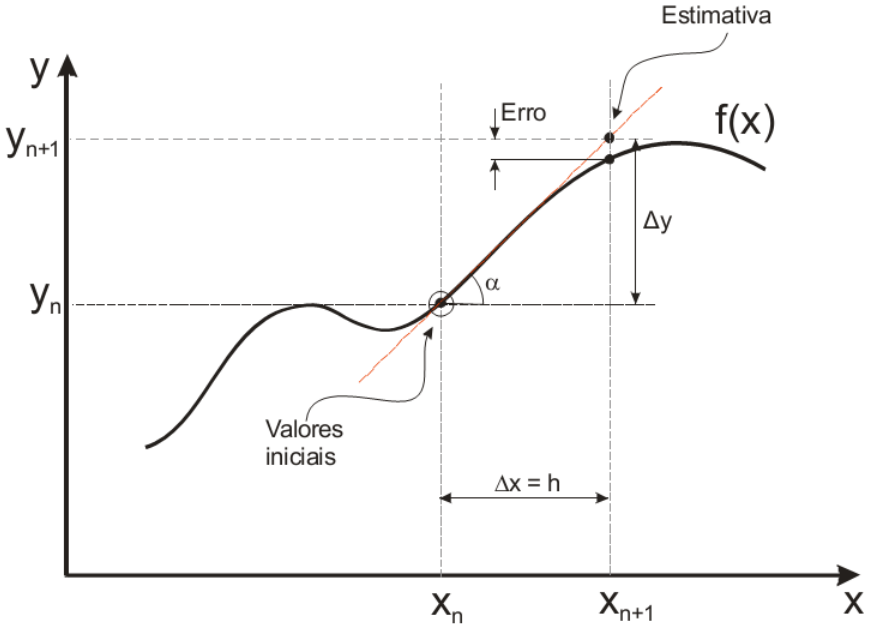
Do cálculo, pode-se escrever que:

$$dy = \frac{dy}{dx} \cdot dx \tag{51}$$

Da equação (51), encontra-se uma aproximação para Δy:

$$\Delta y \simeq \frac{dy}{dx} \cdot \Delta x \tag{52}$$

Figura 22: Método de Euler



Das equações (50) e (52), encontra-se:

$$y_{n+1} = y_n + (x_{n+1} - x_n) \cdot f(x_n, y_n) \tag{53}$$

Na figura 22, observa-se que quanto menor o valor da diferença entre x_{n+1} e x_n (desprezando os erros causados pela representação finita dos números pelos computadores), menor o erro da estimativa para y_{n+1}. Todavia, o número de computações para um intervalo aumenta à medida que a diferença entre x_{n-1} e x_n é reduzida. Define-se o passo h como sendo igual a:

$$h = x_{n+1} - x_n \tag{54}$$

Usando a equação (54) nas equações (54) e (53), tem-se:

$$x_{n+1} = x_n + h \tag{55}$$

e

$$y_{n+1} = y_n + h \cdot f(x_n, y_n) \tag{56}$$

A equação (56) é conhecida como fórmula de Euler. A solução de uma equação diferencial pelo método de Euler é realizada pelo uso recursivo das equações (55) e (56), usando as condições de contorno x_0 e y_0. O erro no método de Euler é da ordem de $O(h^2)$.

Exemplo

As seguir será apresentado o método de Euler na solução de uma equação diferencial ordinária de 1ª ordem. A equação escolhida será:

$$\frac{dy}{dx} = 1 - x + 4\,y$$

Esta equação será resolvida de $x = 0\text{s}$ a $x = 2\text{s}$, com a seguinte condição de contorno:

$$y_0 = 1$$

A solução da equação diferencial com a condição de contorno dada é conhecida:

$$y(x) = \frac{1}{4} \cdot x - \frac{3}{16} + \frac{19}{16}\, e^{4x}$$

A solução numérica é encontrada com a avaliação das equações (55) e (56):

$$x_{n+1} = x_n + h \tag{57}$$

e

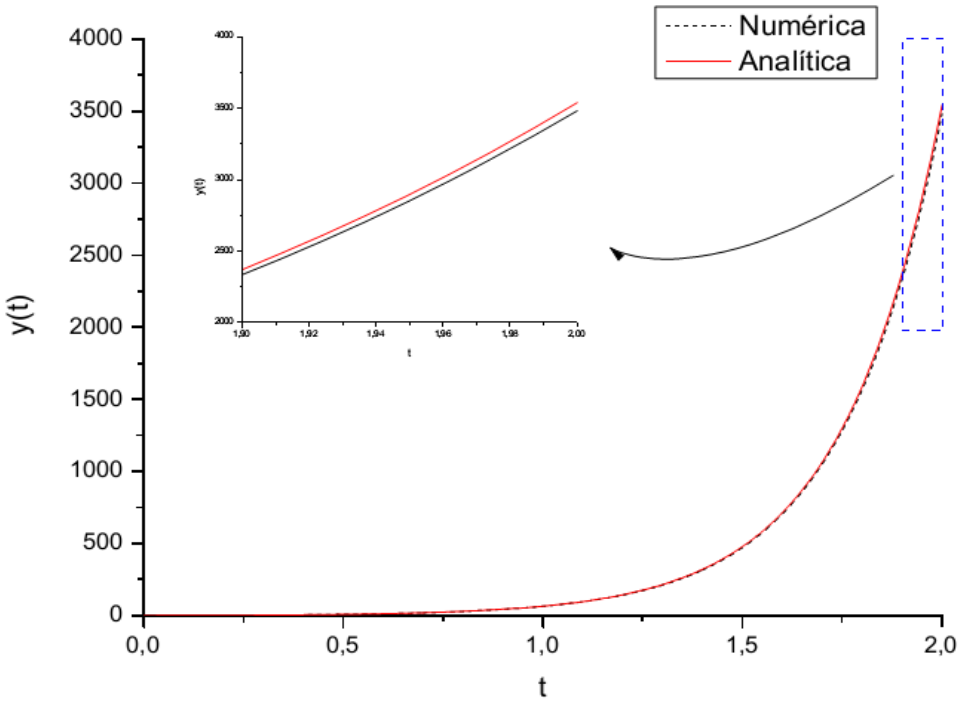
$$\begin{aligned} y_{n+1} &= y_n + h \cdot f(x_n, y_n) \\ &= y_n + h \cdot (1 - x + 4\text{y}) \end{aligned} \tag{58}$$

Com a condição de contorno da equação, temos que $x_0 = 0$ e $y_0 = 1$. Os próximos valores são calculados com o uso recursivo das equações (57) e (58). O valor do passo é escolhido considerando-se o erro desejado. Neste exemplo, escolhemos $h = 0{,}001$. A seguir é apresentada uma tabela com alguns valores calculados de y pelo método de Euler e usando a solução algébrica.

Tabela 18: Solução numérica da equação

n	x_n	y_n	y
0	0,000	1,000000	1,000000
1	0,001	1.005000	1.005010
2	0,002	1.010019	1.010038
3	0,003	1.015057	1.015086
4	0,004	1.020114	1.020153
5	0,005	1.025191	1.025239
...	...	...	...
500	0,500	8.677069	8.712004
1000	1,000	64.382558	64.897803
1500	1,500	473.559790	479.259192
2000	2,000	3484.160803	3540.200110

Figura 23: Gráfico apresentando a solução numérica



15.4 Método de Runge-Kutta

O método de Runge-Kutta pode ser entendido como um aperfeiçoamento do método de Euler, com uma melhor estimativa da derivada da função. No método de Euler a estimativa do valor de y_{n+1} é realizado com o valor de y_n e com a derivada no ponto x_n. No método de Runge-Kutta, busca-se uma melhor estimativa da derivada com a avaliação da função em mais pontos no intervalo $[x_n, x_{n+1}]$. O método de Runge-Kutta de ordem n possui um erro da ordem de $O(h^{n+1})$ A seguir será discutido o método de Runge-Kutta de 2ª ordem, ilustrado pela figura 24.

No método de Euler de passo h, a estimativa de y_{n+1} é realizada com os valores de x_n da derivada de y_n. No método de Runge-Kutta de 2ª ordem, o valor da estimativa de y_{n+1} é encontrado com o valor de y_n e com uma estimativa da derivada em um ponto mais próximo de x_{n+1}, em $x_n+\frac{h}{2}$:

$$y_{n+1}=y_n+h\cdot f\left(x_n+\frac{1}{2}h, y_{n+\frac{1}{2}}\right) \tag{59}$$

Na equação (59), $y_{n+\frac{1}{2}}$ é o valor de y em $x_n+\frac{1}{2}h$ Uma estimativa do valor de $y_{n+\frac{1}{2}}$ é encontrado com o auxílio do método de Euler:

$$y_{n+\frac{1}{2}}=y_n+\frac{h}{2}\cdot f(x_n, y_n) \tag{60}$$

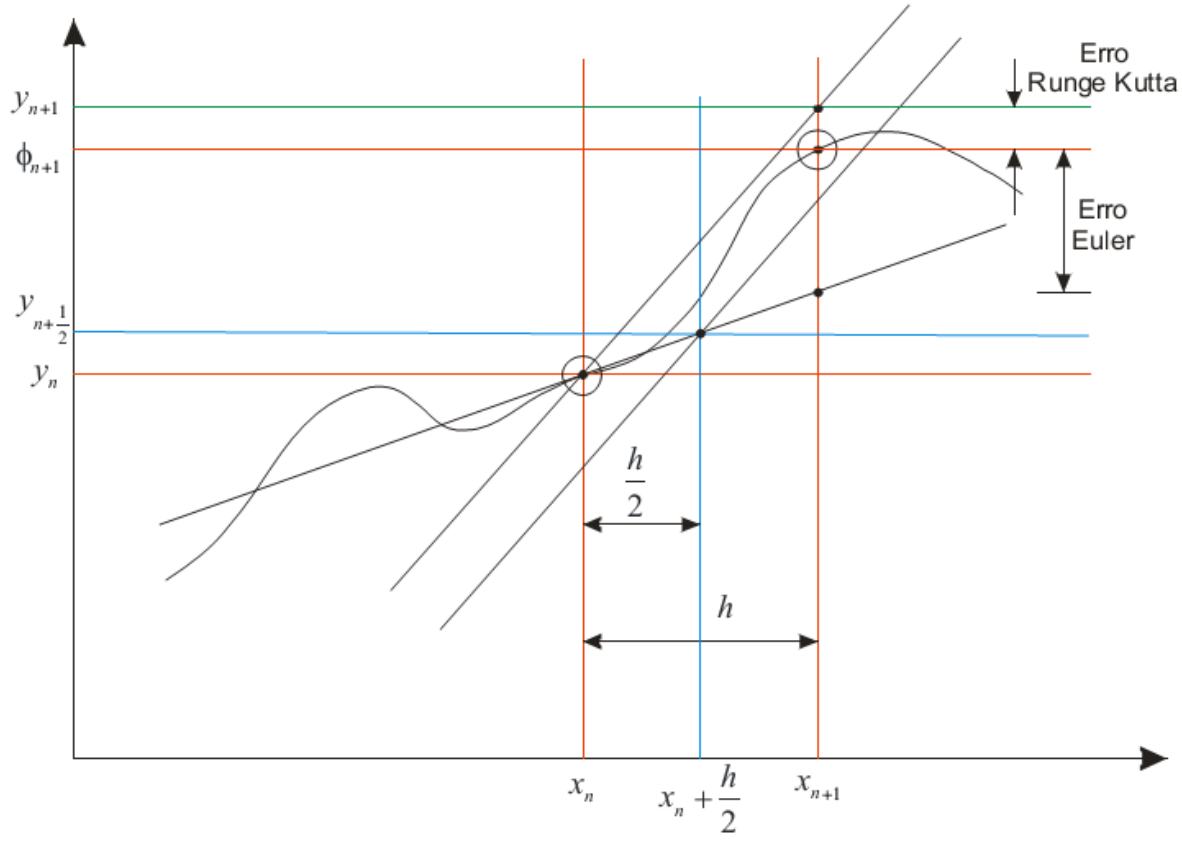
Denominando:

$$k_1 = h \cdot f(x_n, y_n)$$
$$k_2 = h \cdot f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1\right) \tag{61}$$

Escreve-se a equação (59) como:

$$y_{n+1} = y_n + k_2 \tag{62}$$

Figura 24: Método de Runge-Kutta de 2ª ordem



O método de Runge-Kutta de 4ª ordem tem as seguintes equações:

$$k_1 = h \cdot f(x_n, y_n)$$
$$k_2 = h \cdot f\left(x_n + \frac{h}{2}, y_n + \frac{1}{2}k_1\right)$$
$$k_3 = h \cdot f\left(x_n + \frac{h}{2}, y_n + \frac{1}{2}k_2\right) \tag{63}$$
$$k_4 = h \cdot f(x_n + h, y_n + k_3)$$
$$y_{n+1} = y_n + \frac{1}{6}\left(k_1 + 2k_2 + 2k_3 + k_4\right)$$

$$x_{n+1} = x_n + h \tag{64}$$

Exemplo

A seguir será apresentado um exemplo usando o método de Runge-Kutta na solução de uma equação diferencial. Será usada a equação diferencial do exemplo anterior, com as mesmas condições iniciais. O passo neste exemplo será reduzido para $0{,}01\,s$. A solução da equação diferencial é encontrada pelo uso iterativo das equações (63) e (64). A seguir é apresentada uma tabela com os valores calculados e com o valor analítico da solução.

Tabela 19: Resultado da solução numérica da equação

n	x_n	y_n	y
0	0,000	1,000000	1,000000
1	0,001	1.050963	1.050963
2	0,002	1.103903	1.103903
3	0,003	1.158903	1.158903
4	0,004	1.216044	1.216044
5	0,005	1.275416	1.275416
...	...		
50	0,500	8.712004	8.712004
100	1,000	64.897798	64.897803
150	1,500	479.259133	479.259192
200	2,000	3540.199525	3540.200110

Comparando-se os resultados da solução numérica usando o método de Euler Tabela 18 com os resultados da solução usando o método de Runge-Kutta Tabela 19, observa-se que neste segundo método a precisão é maior, mesmo com o uso de um passo 10 vezes maior.

15.5 Exercícios

1) Utilizando o método de Runge-Kutta encontre o valor da função dada por $\frac{dy}{dx}=2x-2$ no intervalo de $x=0$ até $x=1$. Utilize $y(0)=1$ e $h=0{,}2$. Mostre todos os passos. Compare graficamente o resultado obtido com o resultado real.

2) Utilizando o método de Runge-Kutta encontre o valor da função dada por $\frac{dy}{dx}=\cos x$ no intervalo de $x=0$ até $x=\frac{\pi}{2}\cdot$ Utilize $y(0)=2$ e $h=\frac{\pi}{10}\cdot$ Mostre todos os passos. Compare graficamente o resultado obtido com o resultado real.

3) Crie um algoritmo para o método estudado e implemente este algoritmo em uma linguagem de programação. Comprove se funcionamento.

AULA 16 - LISTA DE EXERCÍCIOS 2

16.1 Objetivo:

Nesta aula os alunos irão realizar vários exercícios com o objetivo de aprofundar os conhecimentos adquiridos nas aulas anteriores. Esta aula também serve para que as dúvidas que ainda perduram sejam resolvidas bem como para que os alunos se preparem para a avaliação que se segue.

16.2 Exercícios

1. Através do polinômio interpolador de Lagrange, para os pontos da tabela calcule a aproximação de $f(x)$ para $x=1$. Mostre todos os passos do cálculo.

x	-1	0	3
$f\,x$	15	8	-1

2. Seja $f(x)=x^4-5\text{x}$, $x\in[-1,1]$. Aproximar f(x) por um polinômio de 2° grau usando o método dos mínimos quadrados.
3. A tabela a seguir apresenta os pontos que devem ser aproximados por uma parábola passando pela origem.

x	-1,0	-0,75	-0,6	-0,5	-0,3	0	0,2	0,4	0,5	0,7	1,0
$f\,x$	2,05	1,153	0,45	0,4	0,5	0	0,2	0,6	0,512	1,2	1,05

Assim, usando mínimos quadrados, determine α, onde $g(x)=\alpha\,x^{2.}$

4. Calcular utilizando a regra do trapézio: $\int_0^{1,2} e^x \cos x\,dx$ considerando $h=0,2$.
5. Calcular utilizando a regra $\frac{3}{8}$ de Simpson $\int_0^{1,2} e^x \cos x\,dx.$ considerando $h=0,2$.
6. Utilizando o método de Runge-Kutta encontre o valor da função dada por $\frac{dy}{dx}=\frac{y}{1+x^2}$ no intervalo de $x=0$ até $x=3,5$ Utilize $y(0)=1$ e $h=0,7$. Mostre todos os passos. Compare graficamente o resultado obtido com o resultado real.

Anexo I

Símbolo	Nome	lê-se como	Categoria
⇒ →	**equivalência material**	implica; se ... então	lógica proposicional
	A ⇒ B significa: se A for verdadeiro então B é também verdadeiro; se A for falso então nada é dito sobre B. → pode ter o mesmo significado de ⇒.		
	$x = 2 \Rightarrow x^2 = 4$ é verdadeiro, mas $x^2 = 4 \Rightarrow x = 2$ é em geral falso (visto que x pode ser −2)		
⇔ ↔	**equivalência material**	se e somente se	lógica proposicional
	A ⇔ B significa: A é verdadeiro somente se B for verdadeiro e A é falso somente se B é falso		
	$x + 5 = y + 2 \Leftrightarrow x + 3 = y$		
∧	**conjunção lógica**	e	lógica proposicional
	a proposição A ∧ B so é verdadeira ambos forem verdadeiros.		
	Exemplo: 2 = 4 ∧ 1 = 1 é falso		
∨	**disjunção lógica**	ou	lógica proposicional
	a proposição A ∨ B só é falsa se ambos forem falsos.		
	Exemplo: 2 = 4 ∨ 1 = 1 é verdadeiro		
¬ /	**negação lógica**	não	lógica proposicional
	a proposição ¬A é verdadeira se e só se A for falso Uma barra colocada sobre outro operador tem o mesmo significado que "¬" colocado à sua frente		
	Exemplo: ¬(A ∧ B) ⇔ (¬A) ∨ (¬B); $x \notin S \Leftrightarrow \neg(x \in S)$		
∀	**quantificação universal**	para todos; para qualquer; para cada	lógica predicativa
	∀ x: P(x) significa: P(x) é verdadeiro para todos os x		
	Exemplo: $\forall n \in N: n^2 \geq n$		
∃	**quantificação existencial**	existe	lógica predicativa
	∃ x: P(x) significa: existe pelo menos um x tal que P(x) é verdadeiro		
	Exemplo: $\exists n \in N: n + 5 = 2n$		
=	**igualdade**	igual a	todas
	x = y significa: x e y são nomes diferentes para a exacta mesma coisa		
	Exemplo: 1 + 2 = 6 − 3		
:= :⇔	**definição**	é definido como	todas
	x := y significa: x é definido como outro nome para y P :⇔ Q significa: P é difinido como logicamente equivalente a Q		
	Exemplo: cosh x := (1/2)(exp x + exp (−x)); A XOR B :⇔ (A ∨ B) ∧ ¬(A ∧ B)		
{ , }	chavetas de conjunto	o conjunto de ...	teoria de conjuntos
	{a,b,c} significa: o conjunto que consiste de a, b, e c		
	Exemplo: N = {0,1,2,...}		
{ : } { | }	notação de conjuntos	o conjunto de ... tal que ...	teoria de conjuntos
	{x : P(x)} significa: o conjunto de todos os x, para os quais P(x) é verdadeiro. {x | P(x)} é o mesmo que {x : P(x)}.		

	Exemplo: $\{n \in N : n^2 < 20\} = \{0,1,2,3,4\}$		
Ø {}	**conjunto vazio**	conjunto vazio	teoria de conjuntos
	{} significa: o conjunto sem elementos; Ø é a mesma coisa		
	Exemplo: $\{n \in N : 1 < n^2 < 4\} = \{\}$		
∈ ∉	pertença a conjunto	em; está em; é um elemento de; é um membro de; pertence a	teoria de conjuntos
	a ∈ S significa: a é um elemento do conjunto S; a ∉ S significa: a não é um elemento de S		
	Exemplo: $(1/2)^{-1} \in N$; $2^{-1} \notin N$		
⊆ ⊂	**subconjunto**	é um subconjunto [próprio] de	teoria de conjuntos
	Exemplo: A ⊆ B significa: cada elemento de A é também elemento de B (A é um subconjunto de B) A ⊂ B significa: A ⊆ B mas A ≠ B (A é um subconjunto próprio de B)		
	Exemplo: A ∩ B ⊆ A; Q ⊂ R		
∪	**união teórica de conjuntos**	a união de ... com ...; união	teoria de conjuntos
	A ∪ B significa: o conjunto que contém todos os elementos de A e também todos os de B, mas mais nenhuns		
	Exemplo: A ⊆ B ⇔ A ∪ B = B		
∩	intersecção teórica de conjuntos	intersecta com; intersecta	teoria de conjuntos
	A ∩ B significa: o conjunto que contém todos os elementos que A e B têm em comum		
	Exemplo: $\{x \in R : x^2 = 1\} \cap N = \{1\}$		
\	**complemento teórico de conjuntos**	menos; sem	teoria de conjuntos
	A \ B significa: o conjunto que contém todos os elementos de A que não estão em B		
	Exemplo: {1,2,3,4} \ {3,4,5,6} = {1,2}		
( ) [ ] { }	**aplicação de função; agrupamento**	de	teoria de conjuntos
	para a aplicação de função: f(x) significa: o valor da função f no elemento x para o agrupamento: execute primeiro as operações dentro dos parênteses		
	Exemplo: Se $f(x) := x^2$, então $f(3) = 3^2 = 9$; (8/4)/2 = 2/2 = 1, mas 8/(4/2) = 8/2 = 4		
f:X→Y	seta de função	de ... para	funções
	f: X → Y significa: a função f mapeia o conjunto X no conjunto Y		
	Exemplo: Considere a função f: Z → N definida por $f(x) = x^2$		
N	**números naturais**	N	números
	N significa: {0,1,2,3,...}		
	Exemplo: $\{|a| : a \in Z\} = N$		
Z	**números inteiros**	Z	números
	Z significa: {...,−3,−2,−1,0,1,2,3,...}		
	Exemplo: $\{a : |a| \in N\} = Z$		
Q	**números racionais**	Q	números
	Q significa: $\{p/q : p,q \in Z, q \neq 0\}$		
	$3.14 \in Q$; $\pi \notin Q$		

R	**números reais**	R	números
	R significa: {limn→∞ an : ∀ n ∈ N: an ∈ Q, o limite existe}		
	π ∈ R; √(−1) ∉ R		
C	**números complexos**	C	números
	C significa: {a + bi : a,b ∈ R}		
	i = √(−1) ∈ C		
< >	**comparação**	é menor que, é maior que	ordenações parciais
	x < y significa: x é menor que y; x > y significa: x é maior que y		
	Exemplo: x < y ⇔ y > x		
≤ ≥	**comparação**	é menor ou igual a, é maior ou igual a	ordenações parciais
	x ≤ y significa: x é menor que ou igual a y; x ≥ y significa: x é maior que ou igual a y		
	Exemplo: x ≥ 1 ⇒ x² ≥ x		
√	**raiz quadrada**	a raiz quadrada principal de; raiz quadrada	números reais
	√x significa: o número positivo, cujo quadrado é x		
	Exemplo: √(x²) = |x|		
∞	**infinito**	infinito	números
	∞ é um elemento da linha numérica estendida que é maior que qualquer número real; ocorre com frequência em limites		
	Exemplo: limx→0 1/|x| = ∞		
π	**pi**	pi	geometria euclidiana
	π significa: a razão entre a circunferência de um círculo e o seu diâmetro		
	Exemplo: A = πr² é a área de um círculo de raio r		
!	fatorial	fatorial	análise combinatória
	n! é o produto 1×2×...×n		
	Exemplo: 4! = 24		
| |	**valor absoluto**	valor absoluto de; módulo de	números
	|x| significa: a distância no eixo dos reais (ou no plano complexo) entre x e zero		
	Exemplo: |"a" + "bi"| = √(a² + b²)		
|| ||	**norma**	norma de; comprimento de	análise funcional
	||x|| é a norma do elemento x de um espaço vectorial		
	Exemplo: ||"x"+"y"|| ≤ ||"x"|| + ||"y"||		
∑	**soma**	soma em ... de ... até ... de	aritmética
	∑k=1n ak significa: a1 + a2 + ... + an		
	Exemplo: ∑k=14 k² = 1² + 2² + 3² + 4² = 1 + 4 + 9 + 16 = 30		
∏	**produto**	produto em ... de ... até ... de	aritmética
	∏k=1n ak significa: a1a2⋯an		
	Exemplo: ∏k=14 (k + 2) = (1 + 2)(2 + 2)(3 + 2)(4 + 2) = 3 × 4 × 5 × 6 = 360		
∫	**integração**	integral de ... até ... de ... em função de	cálculo
	∫ab f(x) dx significa: a área entre o eixo dos x e o gráfico da função f entre x = a e x = b		

	$\int_0^b x^2\,dx = b^3/3$; $\int x^2\,dx = x^3/3$		
f '	**derivada**	derivada de f; primitiva de f	cálculo
	f '(x) é a derivada da função f no ponto x, i.e. o declive da tangente nesse ponto		
	Exemplo: Se $f(x) = x^2$, então $f'(x) = 2x$		
∇	**gradiente**	del, nabla, gradiente de	cálculo
	∇f (x1, …, xn) é o vector das derivadas parciais (df / dx1, …, df / dxn)		
	Exemplo: Se $f(x,y,z) = 3xy + z^2$ então $\nabla f = (3y, 3x, 2z)$		

Anexo II – Derivadas

1 $\frac{d}{dx}(c) = 0$

2 $\frac{d}{dx}[c\,f(x)] = c\,f'(x)$

3 $\frac{d}{dx}[f(x)+g(x)] = f'(x)+g'(x)$

4 $\frac{d}{dx}[f(x)-g(x)] = f'(x)-g'(x)$

5 $\frac{d}{dx}[f(x)\,g(x)] = f'(x)\,g(x)+f(x)\,g'(x)$

6 $\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)\,g(x)-f(x)\,g'(x)}{[g(x)]^2}$

7 $\frac{d}{dx}f(g(x)) = f'(g(x))\;g'(x)$

8 $\frac{d}{dx}(x^n) = n\,x^{n-1}$

9 $\frac{d}{dx}(e^x) = e^x$

10 $\frac{d}{dx}(a^x) = a^x \ln(a)$

11 $\frac{d}{dx}\ln|x| = \frac{1}{x}$

12 $\frac{d}{dx}\log_a(x) = \frac{1}{x\ln(a)}$

13 $\frac{d}{dx}\text{sen}(x) = \cos(x)$

14 $\frac{d}{dx}\cos(x) = -\text{sen}(x)$

15 $\frac{d}{dx}\text{tg}(x) = \sec^2(x)$

16 $\frac{d}{dx}\text{cossec}(x) = -\text{cossec}(x)\cot(x)$

17 $\frac{d}{dx}\sec(x) = \sec(x)\;\;\text{tg}(x)$

18 $\frac{d}{dx}\text{cotg}(x) = -\text{cossec}^2(x)$

19 $\frac{d}{dx}\text{arc sen}(x) = \frac{1}{\sqrt{1-x^2}}$

20 $\frac{d}{dx}\text{arc cos}(x) = -\frac{1}{\sqrt{1-x^2}}$

21 $\frac{d}{dx}\text{arc tg}(x) = \frac{1}{1+x^2}$

22 $\frac{d}{dx}\text{arc cossec}(x) = -\frac{1}{x\sqrt{x^2-1}}$

23 $\frac{d}{dx}\text{arc sec}(x) = \frac{1}{x\sqrt{x^2-1}}$

24 $\frac{d}{dx}\text{arc cotg}(x) = -\frac{1}{1+x^2}$

25 $\frac{d}{dx}\text{senh}(x) = \cosh(x)$

26 $\frac{d}{dx}\cosh(x) = \text{senh}(x)$

27 $\frac{d}{dx}\text{tgh}(x) = \text{sech}^2(x)$

28 $\frac{d}{dx}\text{cossech}(x) = -\text{cossech}(x)\;\;\text{cotgh}(x)$

29 $\frac{d}{dx}\text{sech}(x) = -\text{sech}(x)\;\text{tgh}(x)$

30 $\frac{d}{dx}\text{cotgh}(x) = -\text{cossech}^2(x)$

31 $\frac{d}{dx}\text{arc senh}(x) = \frac{1}{\sqrt{1+x^2}}$

32 $\frac{d}{dx}\text{arc cosh}(x) = \frac{1}{\sqrt{x^2-1}}$

33 $\frac{d}{dx}\text{arc tgh}(x) = \frac{1}{1-x^2}$

34 $\frac{d}{dx}\text{arc cossech}(x) = -\frac{1}{|x|\sqrt{x^2+1}}$

35 $\frac{d}{dx}\text{arc sech}(x) = -\frac{1}{x\sqrt{1-x^2}}$

36 $\frac{d}{dx}\text{arc cotgh}(x) = \frac{1}{1-x^2}$

Referências Bibliográficas

CLÁUDIO, Dalcidio Moraes; MARTINS, Jussara Maria. ***Cálculo numérico computacional:*** Teoria e prática. São Paulo: Atlas, 1989.

DUKKIPATI, Rao V.. ***Numerical methods.*** New Delhi: New Age, 2010.

BARROSO, Leônidas Conceição. et al. ***Cálculo numérico:*** Com Aplicações. 2. ed. São Paulo: Harbra, 1987.

CHAPRA, Steven C.; CANALE, Raymond P.. ***Métodos numéricos para engenharia.*** 5. ed. São Paulo: Mc Graw Hill, 2008.

BURIAN, Reinaldo; LIMA, Antônio Carlos de; JUNIOR, Annibal Hetem. ***Cálculo numérico:*** Fundamentos de informática. Rio de Janeiro: Ltc, 2011.

FREITAS, Sérgio R. de. **Métodos Numéricos.** UFMS. 2000

SOUZA, Marcone J. F.. **Métodos Numéricos, Notas de aula.** UFOP. 2011

PEDROSA Diogo P. F. **Interpolação, Notas de aula.** UFRN. 2011

BARATTO Giovani. **Solução de Equações Diferenciais Ordinárias Usando Métodos Numéricos.** UFSM. 2007