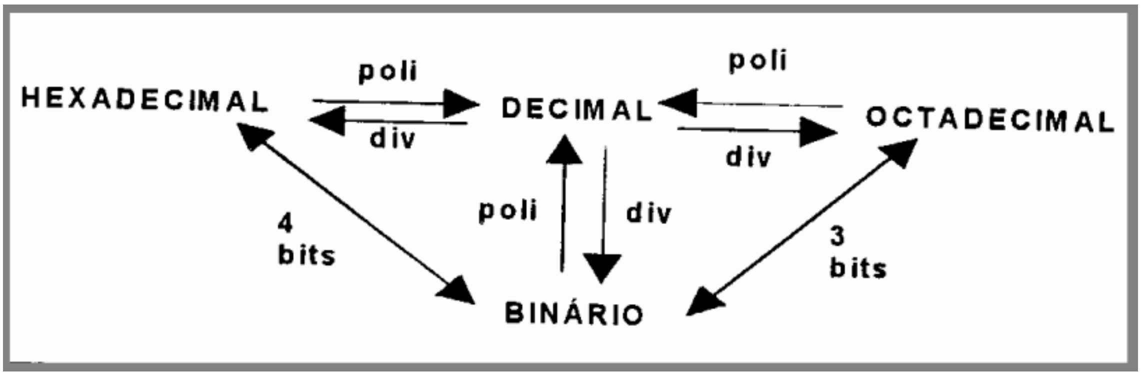
Cálculo Numérico

Sistema de Numeração/Teoria dos Erros/Zeros de Funções/Sistemas Lineares

1. Sistema de Numeração

1.1. Conversão de Base Numérica

É o nome dado à passagem da representação de um número de uma base numérica para outra, alterando a simbologia para se adequar à nova base. A base que normalmente usamos é a decimal ou base dez, pois contém dez algarismos (0, 1, 2, 3, 4, 5, 6, 7, 8, 9).



1.2. Aritméticas Binárias

Como o computador manipula os dados (números) através de uma representação binária, iremos estudar agora a aritmética do sistema binário, a mesma usada pela ULA (Unidade Lógica e Aritmética) dos processadores.

1.2.1. Soma

0 + 0 = 0
0 + 1 = 1
1 + 0 = 1
1 + 1 = 0 (e "vai um" para o dígito de ordem superior)

1.2.2. Subtração

0 - 0 = 0
0 - 1 = 1 ("vem um do próximo")
1 - 0 = 1
1 - 1 = 0

1.2.3. Multiplicação

0 x 0 = 0
0 x 1 = 0
1 x 0 = 0
1 x 1 = 1

1.2.4. Divisão

A divisão é análoga à uma divisão de decimais, trabalhando com multiplicação e subtração.

1.3. Complemento de 1

Outra maneira de representar números binários negativos, consiste em inverter todos os bits, ou seja onde tem 0 ele é substituído por 1 e onde existe 1 é substituído por 0.

Ex: $(0101)_2 = (1010)_2$; $(0111)_2 = (1000)_2$;

1.4. Complemento de 2

Os números negativos nessa abordagem são representados primeiramente aplicando-lhes a regra do complemento de 1 e em seguida soma 1 ao resultado. O principal objetivo do complemento de 2 é trazer uma representação única ao número zero e possibilitar a soma de números positivos e negativos, sem se preocupar com os sinais, pois nessa abordagem a soma de números positivos e negativos pode ser feita normalmente como a soma de dois números positivos. A CPU pode realizar uma subtração indiretamente pela adição do complemento de dois do Subtraendo com Minuendo.

Ex: $(101)_2 = (1011)_2$; $(0111)_2 = (1001)_2$

2. Teoria dos Erros

Conjunto de operações que têm por objetivo determinar o valor de uma grandeza.

2.1. Aritmética de Ponto Flutuante

2.1.1. Representação Numérica nas Máquinas Computacionais

dt = Número de dígitos

Número $= \underbrace{(0,d_1d_2...d_t)}_{\text{Mantissa}} \times \underbrace{B}_{B\,=\,\text{Base}}^{e}$ e = Expoente

Aonde:

$d_1 \neq 0$;

exp ϵ [m, M] ; m = limitante inferior do expoente; M = limitante superior do expoente

Nota: Máquina Computacional tem memória finita

2.2. Arredondamento e Truncamento de Ponto Flutuante

Tipos de Arredondamentos: Matemático; Estatístico e ABNT.

Critério ABNT (Este será o critério adotado no curso de Cálculo Numérico)

Verifica o dígito posterior ao dígito a ser arredondado, se for > 5, então (soma +1);

Se for < 5, então (mantém o dígito)

Se for =5, então (Se o dígito anterior for ímpar, soma + 1, ao dígito anterior. Se o dígito anterior for par, mantém o dígito anterior).

2.3. Overflow e Underflow – SPF (Sistema de Ponto Flutiante)

t = Número de dígitos

SPF(B, t, m, M) exp ϵ [m;M], máquina opera por arredondamento ABNT

B = Base

2.4. Palavra de 16 Bits

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Expoente (Posição 2 a 5)

Sinal do expoente (Posição 1)

Mantissa (Posição 6 a 15)

Sinal do número (Posição 0)

2.5. Erro Absoluto e Relativo

"Erro" é a diferença entre um valor medido e um valor verdadeiro de uma grandeza. "Incerteza" é a quantificação da dúvida sobre o resultado da medição.

Erro Absoluto (EA)

EA = |Vv – Va|

Vv = Valor verdadeiro, valor exato ou valor original; Va = Valor aproximado

Erro Relativo (ER)

ER = $\frac{Ea}{Vv}$

Erro Relativo (ER) em percentual

ER = $\frac{Ea}{Vv}$*100

2.6. Máximo Erro Relativo de Arredondamento (Propagação de Erro)

Análise de Erros nas Operações Aritméticas de Ponto Flutuante.

Onde: RA = $\frac{1}{2}x10^{-t+1}$, porque o erro sempre aumenta.

Nota: Os números são considerados exatamente representados, quando ERx=0; ERy=0;

Os cálculos são efetuados em pares.

Adição $ER(x+y) < ERx\left|\frac{x}{x+y}\right| + ERy\left|\frac{y}{x+y}\right| + RA$

Subtração $ER(x-y) < ERx\left|\frac{x}{x-y}\right| - ERy\left|\frac{y}{x-y}\right| + RA$

Divisão $ER(x/y) < ERx - ERy + RA$

Multiplicação $ER(x*y) < ERx + ERy + RA$

3. Zeros de Funções

Métodos numéricos aplicados na medida em que a complexidade aumenta nas soluções analíticas para obter as raízes de funções polinomiais.

Condições: f(x)=0 e se f(a) * f(b) < 0, então vai existir um zero da função no intervalo [a ; b].

3.1. Teorema de Bolzano

Se f(a) * f(b) < 0, então vai existir um zero da função no intervalo **[a ; b]**. Zero de funções não lineares e reais.

X	-∞	-99	-13	-7	-5	-1	0	2	4	6	8	10
f(x)	-	-	-	-	+	+	+	-	-	+	+	+

4.

f(a) * f(b) < 0 => x_o ϵ [a ; b]

4.1. Método da Bisseção

- +

\- +

Este método numérico da bissecção pode ser adotado para calcular o zero de funções reais não lineares quando a raiz **x = xo** está no **intervalo de [a ; b].**

A raiz da função não linear pode ser estimada pela média aritmética do intervalo $x_k = \frac{a+b}{2}$, em que a tolerância (**restrição)** de **$|f(x_k)| < \varepsilon$, ou $|b - a| < \varepsilon$**.

Onde $\mathbf{X_k}$, é a estimativa do zero da função não linear.

4.2. Método da Falsa Posição / Posição Falsa ou Regula Falsi

Este método numérico da falsa posição pode ser adotado para calcular o zero de funções reais não lineares.

\- +

Então, dado o **intervalo de [a ; b],** a raiz aproximada da função não linear pode ser estimada por $X_k = \frac{a*f(b)-b*f(a)}{f(b)-f(a)}$, em que a tolerância (**restrição)** de **$|f(x_k)| < \varepsilon$, ou $|b - a| < \varepsilon$**.

Onde $\mathbf{X_k}$, é a estimativa do zero da função não linear. Lembrando a regra: **f(a) * f(b) < 0**.

4.3. Método do Ponto Fixo

A partir da expressão f(x) = 0, podemos determinar a expressão **x = g(x)**, que é a função de iteração que será utilizada para calcular o valor de x. Para tanto, temos: **$X_{k+1} = g(X_k)$.** Sendo X_{k+1}, a estimativa da raiz da função não linear atual e $g(X_k)$, a estimativa da raiz da função não linear anterior. Na qual, a estimativa da raiz atual depende da estimativa da raiz anterior.

4.4. Método Newtin Rapshon

Dada à expressão **f(x) = 0** e a estimativa inicial X_0, podemos calcular a raiz real estimada da função não linear, de acordo com a **estimativa atual (X_{k+1})** e a **estimativa anterior (X_k)**, a partir da expressão

$$X_{k+1} = X_k - \frac{f(X_k)}{f'(X_k)}.$$

Este é o método no qual a estimativa atual depende da estimativa anterior, além disso, o método depende do **valor da função f no ponto X_k da estimativa anterior** e também do **valor da derivada da função f no ponto X_k da estimativa anterior**.

4.5. Método Secante

Dada duas raízes reais com estimativas iniciais: X_0 e X_1 e à expressão f(x) = 0, podemos calcular a raiz real estimada da função não linear, de acordo com a **estimativa atual (X_{k+1})** e as **duas estimativas anteriores (X_k e X_{k-1})**, a partir da expressão $X_{k+1} = \frac{X_{k-1} * (f(X_k) - X_k * f(X_k - 1)}{f(X_k) - f(X_{k-1})}$, sendo que, a estimativa atual depende das duas estimativas anteriores.

5. Sistemas Lineares – Métodos Diretos

É um conjunto de equações lineares, com m equações e n incógnitas. Para reduzir este sistema, serão utilizados métodos numéricos diretos e iterativos, transformando em sistemas triangulares, através de escalonamento.

5.1. Método Gauss

É o sistema diagonal em que os elementos a_{ij} da matriz coeficiente [A] são iguais à zero, para i < j.

O método de Gauss consiste em, por meio de um número de (n-1) passos, transformar o sistema linear A.x=B em um sistema triangular equivalente, U.x=C. Este método é mais usado em sistemas lineares de pequeno e médio portes (n=30 e n=50 respectivamente).

As operações elementares para transformar o sistema inicial em outro equivalente são as seguintes:

· Multiplicar uma linha por um número real diferente de zero;
· Somar ou subtrair a uma linha outra linha;
· Somar a uma linha outra linha multiplicada por um número diferente de zero.
· Trocar a ordem de duas linhas;
· Trocar a ordem das colunas que correspondem as incógnitas do sistema, levando em consideração as trocas realizadas na hora de escolher o novo sistema equivalente;
· Eliminar linhas proporcionais ou que sejam combinações lineares de outras linhas;
· Eliminar linhas nulas (0 0 0 ... 0);

Para reduzir os sistemas lineares de equações será necessário utilizar escalonamento.

Passos:

1. Identificar os Pivôs (elementos da diagonal principal);

2. Zerar os elementos das linhas abaixo do Pivô;

3. Identificar o multiplicador da linha onde está o Pivô, definido pela razão entre o elemento da linha a ser zerado e o pivô;

4. Multiplicar a linha aonde se encontra o pivô pelos elementos da linha a ser zerada;

primeiro temos o sistema:

$$\begin{cases} 2x_1 + 8x_2 + 3x_3 = 2 \\ x_1 + 3x_2 + 2x_3 = 5 \\ 2x_1 + 7x_2 + 4x_3 = 8 \end{cases}$$

para começamos escrevemos sua matriz aumentada:

$$\begin{bmatrix} 2 & 8 & 3 & 2 \\ 1 & 3 & 2 & 5 \\ 2 & 7 & 4 & 8 \end{bmatrix}$$

Solução: X1=3; X2 = -2; X3 =4.

5.2. Método Pivoteamento Parcial

Semelhante ao método de Gauss. Minimiza a amplificação de erros de arredondamento durante as eliminações. Consiste em escolher o elemento de maior módulo em cada coluna para ser o Pivô.

Portanto, faz-se necessário identificar o **maior elemento em módulo** da coluna em que se encontra o primeiro Pivô. Após identificar trocar a linha que se encontra o Pivô pela linha onde está o maior elemento em módulo. Assim sucessivamente, para cada Pivô da matriz a ser escalonada.

Depois aplicar o procedimento de Gauss para zerar os elementos da diagonal principal abaixo do Pivô, através da técnica de escalonamento.

5.3. Método Jordan

É o sistema diagonal em que os elementos a_{ij} da matriz coeficiente [A] são iguais à zero, para i<>j.

Será aplicado o procedimento de Gauss para zerar os elementos da diagonal principal abaixo do Pivô, através da técnica de escalonamento, e também zerar os elementos da diagonal principal acima do Pivô.

5.4. Método Fatoração LU ou Decomposição LU

O objetivo é fatorar a matriz dos coeficientes A em um produto de duas matrizes L e U.

Ax = b; Ax = LU

LUx = b = > L*(Ux) = b; Fazendo Ux = y, temos: Ly = b

A **matriz U** é a matriz escalonada pelo método de Gauss, somente dos coeficientes da matriz A.

A **matriz L** é a matriz definida com os elementos da diagonal principal iguais a 1, acima destes elementos iguais a ZERO e abaixo dos elementos da diagonal principal os elementos multiplicadores utilizados pelo método de sistema de Gauss, utilizados para escalonar a matriz A.

Para resolver a fatoração LUx = b, será necessário resolver dois sistemas:

1. Resolver o sistema Ly = b
2. Depois resolver o sistema Ux = y.

Nota: b (**vetor constante b** da matriz A); y (**vetor constante y**, resultado do sistema Ly = b).