Formulário - Cálculo Numérico

Wellington José Corrêa
Universidade Tecnológica Federal do Paraná
Campus Campo Mourão
DAMAT, 2017.

Nome:

1 Conversão de Base e Erros

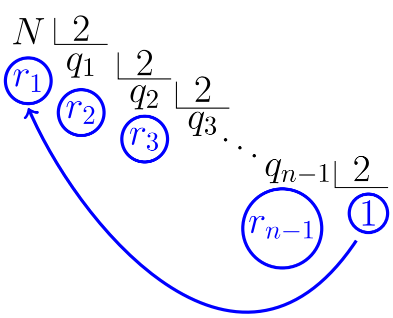
Base Binária

Um número na base 2 pode ser escrito como:

$$\boxed{\sum_{i=n}^{m} a_i\, 2^i;\;\; a_i = \{0, 1\},\; n, m \in \mathbb{Z};\; n \leq 0 \text{ e } m \geq 0.}$$

- Para mudar da base 2 para base 10, basta multiplicar o dígito binário por uma potência de 2 adequada.
- Para converter um número da base 10 para a base 2, tem-se que aplicar um processo para a parte inteira e outro para a parte fracionária.

1. Parte inteira:



Portanto,

$$\boxed{(N)_{10} = (1\, r_{n-1}\, r_{n-2}\, \dots\, r_3\, r_2\, r_1)_2\ .}$$

2. Parte Fracionária.
   (a) Multiplica-se o número fracionário por 2;
   (b) Do resultado do passo (a), a parte inteira é o primeiro dígito binário;
   (c) Do resultado do passo (b), a parte fracionária é novamente multiplicada por 2;
   (d) O processo continua até que a parte fracionária seja nula.

Aritmética de Ponto Flutuante

$$\boxed{\underbrace{\pm}_{\text{sinal}}\; 0,\; \underbrace{d_1 d_2 d_3 \dots d_p}_{\text{mantissa}} \times B^{\overbrace{e}^{\text{expoente}}}\ .}$$

Erros

Denotando por a_{ex} o valor exato e a_{aprox} o valor aproximado, temos:

Erro absoluto:

$$\boxed{E_{\text{abs}} = |a_{\text{ex}} - a_{\text{aprox}}|}$$

Erro relativo:

$$\boxed{E_{rel} = \frac{E_{\text{abs}}}{a_{\text{aprox}}}}$$

2 Raízes de Equações

Em todos os métodos, escolha a, b de modo que

$$\boxed{f(a)\, \cdot\, f(b) < 0.}$$

Usaremos o seguinte critério de parada:

$$\boxed{|x_{k+1} - x_k| \leq \varepsilon\, .}$$

Método da Bissecção

$$\boxed{x_{k+1} = \frac{a_k + b_k}{2},\; k = 0, 1, \dots}$$

Número mínimo de iterações k :

$$\boxed{k \geq \log_2\left(\frac{b-a}{\varepsilon}\right) - 1\, .}$$

A seguir, encontra-se um esquema prático para resolver este método:

k	a_k	b_k	$f(a_k)$	$f(b_k)$	x_k	$f(x_k)$	$\lvert x_k - x_{k-1}\rvert$
0	a	b	$f(a)$	$f(b)$	$\frac{a+b}{2}$	$f\left(\frac{a+b}{2}\right)$	$-$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Método do Ponto Fixo

Teorema 2.1 *Seja $\xi \in [a, b]$ uma raiz da equação $f(x) = 0$ e $F(x)$ contínua, diferenciável em $[a, b]$ e $F(x) \in [a, b]$. Se $|F'(x)| \leq k < 1$ para todos os pontos em $[a, b]$ e $x_0 \in [a, b]$, então, os valores dados pela equação*

$$\boxed{x_{k+1} = F(x_k),\; k = 0, 1, 2, \dots}$$

converge para ξ.

Método de Newton

Teorema 2.2 *Uma condição suficiente para a convergência do método de Newton é, se $f(a) \cdot f(b) < 0$, $f'(x), f''(x)$ forem não-nulas e preservem o sinal em (a, b) de modo que*

$$f(x_0) \cdot f''(x_0) > 0.$$

$$\boxed{x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, k = 0, 1, 2, ...}$$

Método da Secante

Considere $x_0 = a$ e $x_1 = b$.

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

Método Regula Falsi

Considere $x_0 = a$ e $x_1 = b$.

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

O método *Regula Falsi* retém o ponto no qual o valor da função tem sinal oposto no ponto mais recente, assegurando desta forma, que a raiz continue isolada entre dois pontos.

3 Sistemas Lineares

Decomposição LU

Se $\det(A_k) \neq 0$, $k = 1, 2, \ldots, n-1$.
Então,

$$\boxed{A = L \cdot U},$$

onde

$$\boxed{L = \begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & 0 & \\ l_{31} & l_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{pmatrix}, U = \begin{pmatrix} u_{11} & u_{21} & u_{31} & \cdots & u_{n1} \\ & u_{22} & u_{23} & \cdots & u_{n2} \\ & & u_{33} & \cdots & u_{3n} \\ & 0 & & \ddots & \vdots \\ & & & & u_{nn} \end{pmatrix}}$$

E ainda,

$$\boxed{\det(A) = u_{11}\, u_{22}\, \ldots\, u_{nn}.}$$

Procedimento:

- Passo 1: Primeira linha de U.

$$\boxed{u_{1j} = a_{1j},\ j = 1, 2, \ldots, n\,.}$$

- Passo 2: Primeira coluna de L.

$$\boxed{l_{i1} = \frac{a_{i1}}{u_{11}},\ i = 2, \ldots, n\,.}$$

- Passo 3: Segunda linha de U.

$$\boxed{u_{2j} = a_{2j} - l_{21}\, u_{1j},\ j = 2, \ldots, n\,.}$$

- Passo 4: Segunda coluna de L.

$$\boxed{l_{i2} = \frac{a_{i2} - l_{i1}\, u_{12}}{u_{22}},\ i = 3, \ldots, n\,.}$$

- Passo 5: Terceira linha de U, 3ª coluna de L, 4ª linha de U, 4ª de L, etc.

$$\boxed{u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik}\, u_{kj},\ i \leq j}$$

$$\boxed{l_{ij} = \frac{a_{ij} - \displaystyle\sum_{k=1}^{j-1} l_{ik}\, u_{kj}}{u_{jj}},\ i > j\,.}$$

Resolução de sistema linear $A\, X = B$:

Resolva primeiramente

$$\boxed{L \cdot Y = B}$$

e depois

$$\boxed{U \cdot X = Y\,.}$$

Método de Cholesky

Se $A = A^t$ e $\det(A) > 0$, $k = 1, 2, \ldots, n$, então

$$\boxed{A = G \cdot G^t}$$

onde

$$\boxed{G = \begin{pmatrix} g_{11} & & & & \\ g_{21} & g_{22} & & 0 & \\ g_{31} & g_{32} & g_{33} & & \\ \vdots & \vdots & \vdots & \ddots & \\ g_{n1} & g_{n2} & g_{n3} & \cdots & g_{nn} \end{pmatrix}.}$$

Além disso,

$$\boxed{\det(A) = (g_{11}\, g_{22}\, \ldots\, g_{nn})^2\,.}$$

Procedimento:

- Elementos Diagonais

$$\boxed{g_{11} = \sqrt{a_{11}}}$$

$$\boxed{g_{ii} = \left(a_{ii} - \sum_{k=1}^{i-1} g_{ik}^2\right)^{\frac{1}{2}},\ i = 2, 3, \cdots, n\,.}$$

- Elementos não diagonais de G
  1. Primeira coluna:

$$\boxed{g_{i1} = \frac{a_{i1}}{g_{11}},\ i = 2, 3, \cdots, n\,.}$$

  2. Segunda coluna:

$$\boxed{g_{i2} = \frac{a_{i2} - g_{i1}\, g_{21}}{g_{22}},\ i = 3, 4, \cdots, n\,.}$$

  3. Demais colunas:

$$\boxed{g_{ij} = \frac{a_{ij} - \displaystyle\sum_{k=1}^{j-1} g_{ik} g_{jk}}{gjj},\ 2 \leq j < i\,.}$$

Resolução de sistema linear $A\,X = B$:

$$\boxed{G \cdot Y = B} \qquad \boxed{G^t \cdot X = Y\,.}$$

Método de Eliminação de Gauss

Considere o sistema linear $A\,X = B$ caracterizado matricialmente pela matriz aumentada:

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

onde $i, j = 1, 2, \dots, n$, $a_{ij}^{(1)} = a_{ij}$ e $b_i^{(1)} = b_i$.

De modo geral, o k-ésimo passo do método da eliminação de Gauss é obtido por

$$\boxed{\begin{aligned} a_{ij}^{(k+1)} &= a_{ij}^{(k)} - a_{kj}^{(k)}\,\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}},\ k = 1, 2, \dots, n-1; \\ &\qquad\qquad i = k+1, \dots, n; \\ b_i^{(k+1)} &= b_i^{(k)} - b_k^{(k)}\,\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}},\ j = k, k+1, \dots, n\,. \end{aligned}}$$

Norma de Matriz

- Norma do Máximo (norma linha):

$$\boxed{\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|;}$$

- Norma da Soma (norma coluna):

$$\boxed{\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|;}$$

- Norma Euclidiana:

$$\boxed{\|A\|_E = \sqrt{\sum_{i,j=1}^{n} a_{ij}^2}}$$

Método de Gauss-Jacobi

Convergência: São duas situações:

1. Se A for estritamente diagonalmente dominante (e.d.d.), isto é,

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

2. Ou se $\|F\|_\infty < 1$ ou $\|F\|_1 < 1$, onde $F = (f_{ij})$ é dada por

$$f_{ij} = \begin{cases} 0, & i = j \\ -\frac{a_{ij}}{a_{ii}}, & i \neq j \end{cases} \tag{1}$$

Algoritmo:

$$\boxed{\begin{aligned} x_1^{(k+1)} &= \frac{1}{a_{11}}\left(b_1 - a_{12}\,x_2^{(k)} - a_{13}\,x_3^{(k)} - \dots - a_{1n}\,x_n^{(k)}\right) \\ x_2^{(k+1)} &= \frac{1}{a_{22}}\left(b_2 - a_{21}\,x_1^{(k)} - a_{23}\,x_3^{(k)} - \dots - a_{2n}\,x_n^{(k)}\right) \\ \vdots\quad & \quad \vdots \\ x_n^{(k+1)} &= \frac{1}{a_{nn}}\left(b_n - a_{n1}\,x_1^{(k)} - a_{n2}\,x_2^{(k)} - \dots - a_{nn-1}\,x_{n-1}^{(k)}\right) \end{aligned}}$$

Critério da Parada:

$$\boxed{\frac{\|X^{(k+1)} - X^{(k)}\|_\infty}{\|X^{(k+1)}\|_\infty} < \varepsilon}$$

onde ε é a precisão pré-fixada.

Método de Gauss-Seidel

Convergência: Ocorre nas seguintes situações:

1. O critério de Sanssenfeld for satisfeito, isto é, se

$$\boxed{\max_{1 \le i \le n} \beta_i < 1,}$$

onde

$$\boxed{\beta_i = \sum_{j=1}^{i-1} \left|\frac{a_{ij}}{a_{ii}}\right| \beta_j + \sum_{j=i+1}^{n} \left|\frac{a_{ij}}{a_{ii}}\right|,\ i = 1, 2, \dots, n.}$$

2. Se a matriz A do sistema $AX = B$ for estritamente diagonalmente dominante.

3. O critério das linhas é satisfeito, isto é, $\|F\|_\infty < 1$, onde F é dado em (1).

Algoritmo:

$$\boxed{\begin{aligned} x_1^{(k+1)} &= \frac{1}{a_{11}}\left(b_1 - a_{12}\,x_2^{(k)} - a_{13}\,x_3^{(k)} - \dots - a_{1n}\,x_n^{(k)}\right) \\ x_2^{(k+1)} &= \frac{1}{a_{22}}\left(b_2 - a_{21}\,x_1^{(k+1)} - a_{23}\,x_3^{(k)} - \dots - a_{2n}\,x_n^{(k)}\right) \\ x_3^{(k+1)} &= \frac{1}{a_{33}}\left(b_3 - a_{31}\,x_1^{(k+1)} - a_{32}\,x_2^{(k+1)} - a_{34}\,x_4^{(k)} \dots - a_{3n}\,x_n^{(k)}\right) \\ \vdots\quad & \quad \vdots \\ x_n^{(k+1)} &= \frac{1}{a_{nn}}\left(b_n - a_{n1}\,x_1^{(k+1)} - a_{n2}\,x_2^{(k+1)} - \dots - a_{nn-1}\,x_{n-1}^{(k+1)}\right) \end{aligned}}$$

Critério da Parada:

$$\boxed{\frac{\|X^{(k+1)} - X^{(k)}\|_\infty}{\|X^{(k+1)}\|_\infty} < \varepsilon}$$

onde ε é a precisão pré-fixada.