

Programa para casa #3

Determinação de raízes de equações *Aplicação do Método de Müller*

April 14, 2025

1 Contextualização

Até esse momento vimos algumas opções de métodos intervalares e abertos, destinados à obtenção de raízes de funções algébricas e transcendentais. Dentre os métodos vistos, um dos mais utilizados em diversas classes de problemas é o método de Newton-Raphson, que possui algumas variantes, como sua versão modificada adaptada para a obtenção de raízes múltiplas. Na verdade, muitos métodos numéricos possuem a característica de serem adaptados para se adequarem à solução de diferentes classes de problemas. Um outro exemplo interessante onde podemos aplicar o método de Newton-Raphson é na solução de sistemas de equações não-lineares. Nesse cenário, o método de Newton-Raphson não chega a resolver o problema, mas nos permite reformulá-lo a fim de transformarmos um sistema de equações não-lineares em múltiplos sistemas lineares que são resolvidos de maneira iterativa por um método próprio para essa finalidade.

1.1 O método de Newton-Raphson aplicado a sistemas não lineares

Existem diversos métodos utilizados para a solução de equações não lineares, que vão desde uso de algoritmos genéticos, método da descida íngreme, métodos de ponto fixo, entre outros. Um método de fácil compreensão e ao mesmo tempo muito eficaz consiste no método de Newton. Geralmente alunos de cursos de cálculo numérico estão familiarizados com o método de Newton para a determinação de zeros de funções e em certo sentido é essa a ideia por trás de seu uso para resolver sistemas de equações não lineares. A ideia do método de Newton é que uma função $f(x)$ qualquer (linear ou não) pode ser aproximada por uma série de Taylor no ponto $x = x_0 + \Delta x$ por seu valor e de sua derivada no ponto x_0 como

$$f(x) \approx f(x_0) + \frac{df}{dx}(x_0)\Delta x \quad (1)$$

aqui $\Delta x = x - x_0$. Caso desejemos encontrar alguma raiz da função $f(x)$, ou seja, determinar o valor (ou valores) de x que fazem com que $f(x) = 0$, igualamos a equação (1) e isolamos x , este procedimento fornece

$$x = x_0 - \frac{f(x_0)}{f'(x_0)} \quad (2)$$

aqui $f'(x_0)$ denota a primeira derivada de f com respeito a x avaliada em x_0 . Como o método se baseia na aproximação dada por (1) o uso da equação (2) não fornecerá a resposta exata com apenas uma única computação. Na realidade o método de Newton é um método iterativo, em que

$$x^{i+1} = x^i - \frac{f(x^i)}{f'(x^i)}, \quad i = 0, 1, 2, \dots \quad (3)$$

A partir do momento em que a diferença entre dois valores consecutivos de x for suficientemente pequena (dentro de uma tolerância pré-estabelecida) dizemos que o método convergiu. Considere por exemplo que desejemos determinar uma raiz da função dada por

$$f(x) = x^2 - 4, \quad (4)$$

aplicando o método descrito em (3) para a equação (4) obtemos

$$x^{i+1} = x^i - \frac{x^{i2} - 4}{2x^i}, \quad i = 0, 1, 2, \dots \quad (5)$$

Sabemos que a equação (4) possui duas raízes reais, dadas por 2 e -2. Considerando $x_0 = 0.5$ podemos montar a seguinte tabela:

i	x^i	x^{i+1}	$ x^{i+1} - x^i $
0	0.5	4.25	3.75
1	4.25	2.5956	1.6544
2	2.5956	2.0683	0.5273
3	2.0683	2.0011	0.0672
4	2.0011	2.0000003183	0.0011284441
5	2.0000003183	2.0	3.18×10^{-7}

Table 1: Determinação de uma das raízes de uma equação não linear algébrica de segundo grau simples via Método de Newton.

Note que a partir da sexta iteração a diferença entre dois valores consecutivos da raiz obtida é da ordem de 10^{-7} . Essa tabela mostra a eficácia do método e sua rápida convergência. O uso do método de Newton para a solução de sistemas não lineares se baseia na generalização deste para encontrar N raízes de N equações não lineares organizadas na forma de um sistema não linear. Um sistema não linear de N equação acopladas pode ser organizado em termos vetoriais como uma função $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. Essa função \mathbf{f} é um vetor com N componentes em que N é o número de equações a ser resolvida. Como exemplo considere o seguinte sistema de duas equações não lineares

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 - 2 \\ x_1^2 + 4x_2^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (6)$$

considere ainda a confecção da seguinte matriz \mathbf{J}

$$\mathbf{J} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2x_1 & 8x_2 \end{pmatrix}. \quad (7)$$

Essa matriz é denominada *matriz Jacobiana do sistema* e é utilizada para escrever a equação de recorrência (5) em forma vetorial como

$$\mathbf{x}^{i+1} = \mathbf{x}^i - \mathbf{J}(\mathbf{x}^i)^{-1} \cdot \mathbf{f}(\mathbf{x}^i). \quad (8)$$

Note que na equação (8) surge a matriz inversa de \mathbf{J} . Isso não significa necessariamente que precisamos inverter a matriz \mathbf{J} para resolver o sistema. Chamando o termo $-\mathbf{J}(\mathbf{x}^i)^{-1} \cdot \mathbf{f}(\mathbf{x}^i)$ de \mathbf{s}^i temos

$$\mathbf{s}^i = -\mathbf{J}(\mathbf{x}^i)^{-1} \cdot \mathbf{f}(\mathbf{x}^i) \quad (9)$$

multiplicando os dois lados da equação (9) por $\mathbf{J}(\mathbf{x}^i)$ obtemos

$$\mathbf{J}(\mathbf{x}^i) \cdot \mathbf{s}^i = -\mathbf{J}(\mathbf{x}^i) \cdot \mathbf{J}(\mathbf{x}^i)^{-1} \cdot \mathbf{f}(\mathbf{x}^i) = -\mathbf{I} \cdot \mathbf{f}(\mathbf{x}^i) = -\mathbf{f}(\mathbf{x}^i), \quad (10)$$

dessa forma temos

$$\mathbf{x}^{i+1} = \mathbf{x}^i + \mathbf{s}^i, \quad i = 0, 1, 2, \dots \quad (11)$$

com \mathbf{s}^i dado pela solução do sistema linear

$$\mathbf{J}(\mathbf{x}^i) \cdot \mathbf{s}^i = -\mathbf{f}(\mathbf{x}^i). \quad (12)$$

Para o exemplo em questão considere um vetor inicial $\mathbf{x}^0 = [1 \ 2]^T$, dessa forma $\mathbf{f}(\mathbf{x}^0) = [3 \ 13]^T$ e

$$\mathbf{J}(\mathbf{x}^0) = \begin{pmatrix} 1 & 2 \\ 2 & 16 \end{pmatrix} \quad (13)$$

Resolvendo o sistema

$$\begin{pmatrix} 1 & 2 \\ 2 & 16 \end{pmatrix} \cdot \mathbf{s}^0 = \begin{pmatrix} -3 \\ 13 \end{pmatrix}, \quad (14)$$

obtemos $\mathbf{s}^0 = [-1.83 \ -0.58]^T$ e $\mathbf{x}^1 = [-0.83 \ 1.42]^T$. Para um próximo ponto temos $\mathbf{f}(\mathbf{x}^1) = [0 \ 4.72]^T$ e

$$\mathbf{J}(\mathbf{x}^1) = \begin{pmatrix} 1 & 2 \\ -1.67 & 11.3 \end{pmatrix} \quad (15)$$

Resolvendo o sistema

$$\begin{pmatrix} 1 & 2 \\ -1.67 & 11.3 \end{pmatrix} \cdot \mathbf{s}^1 = \begin{pmatrix} 0 \\ 4.72 \end{pmatrix}, \quad (16)$$

obtemos $\mathbf{s}^1 = [0.64 \ -0.32]^T$ e $\mathbf{x}^1 = [-0.19 \ 1.1]^T$, repetindo o procedimento diversas vezes a solução convergir para $\mathbf{x} = [0 \ 1]^T$.

Note que o uso do Método de Newton para a solução de sistemas não lineares consiste na transformação de um único sistema não linear em diversos sistemas lineares resolvidos de modo progressivo em um procedimento iterativo até a convergência. A ideia por trás deste método é a de que uma função não linear pode ser aproximada como uma soma de diferentes funções lineares em torno de pequenos intervalos. A próxima seção mostra um exemplo do uso deste método na solução da equação de Navier-Stokes para um problema de camada limite laminar sobre placa plana de um fluido incompressível.

1.2 Exemplo de aplicação: solução de uma EDP não linear

Considere a seguinte equação diferencial parcial

$$u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = \frac{1}{Re} \frac{\partial^2 u}{\partial y^2} \quad (17)$$

a equação (17) é conhecida como equação de Navier-Stokes (sua componente apenas na direção x para um problema de camada limite). Considere a seguinte imagem ilustrativa que representa o domínio de cálculo discretizado no qual desejamos resolver a equação (17). A figura (2) ilustra um modelo esquemático do problema de camada limite a ser abordado.

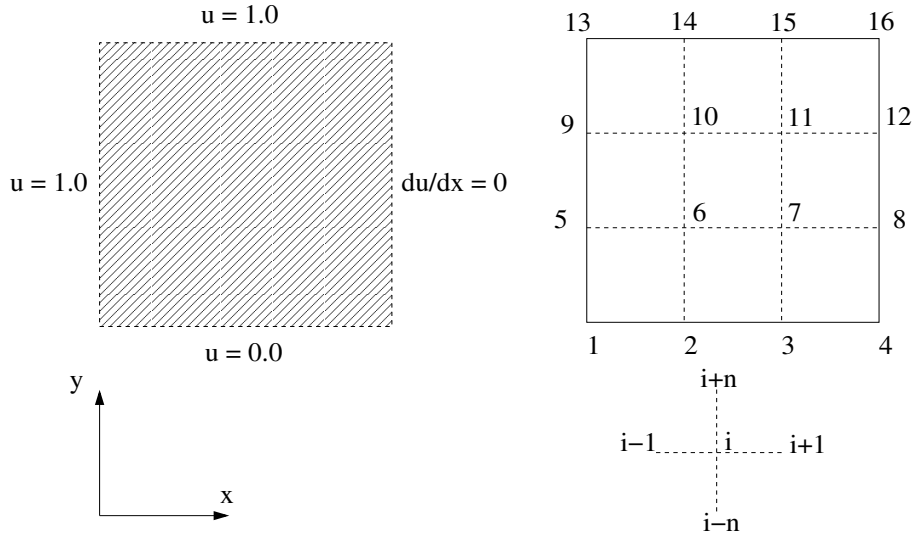


Figure 1: O domínio de cálculo discretizado consiste em uma malha com 16 nós, sendo 4 internos e 12 externos sujeitos às condições de contorno descritas na figura.

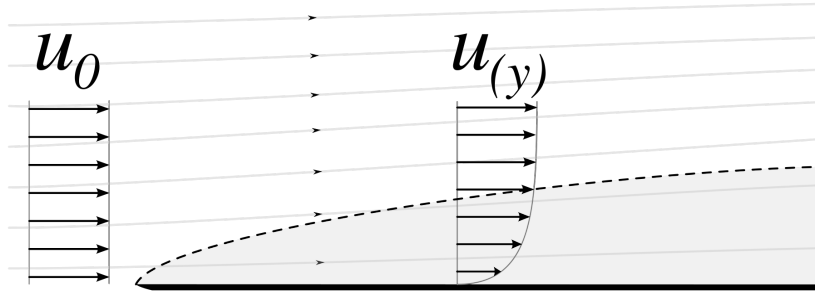


Figure 2: Modelo esquemático do problema de camada limite hidrodinâmica.

As funções escalares u e v são funções de x e y (coordenadas espaciais), de modo que

$$u = u(x, y) \quad \text{e} \quad v = v(x, y) \quad (18)$$

A versão discretizada da equação (17) é dada por

$$u^i \frac{(u^i - u^{i-1})}{\Delta x} + v^i \frac{(u^i - u^{i-n})}{\Delta x} = \frac{1}{Re} \frac{(u^{i+n} + u^{i-n} - 2u^i)}{\Delta y^2} \quad (19)$$

isolando u^i temos

$$u^i \left(\frac{u^i}{\Delta x} + \frac{v^i}{\Delta y} + \frac{2}{Re\Delta y^2} - \frac{u^{i-1}}{\Delta x} \right) - \frac{v^i u^{i-n}}{\Delta y} - \frac{u^{i+n}}{Re\Delta y^2} - \frac{u^{i-n}}{Re\Delta y^2} = 0 \quad (20)$$

Note que a equação acima se refere à um nó i qualquer e a maneira com a qual este se relaciona com seus vizinhos. Para fins de simplificação considere o caso em que $v = 0$, de modo que

$$u^i \left(\frac{u^i}{\Delta x} + \frac{2}{Re\Delta y^2} - \frac{u^{i-1}}{\Delta x} \right) - \frac{u^{i+n}}{Re\Delta y^2} - \frac{u^{i-n}}{Re\Delta y^2} = 0 \quad (21)$$

Podemos escrever esse sistema de equações não lineares para a malha em questão (com 16 nós) da forma

$$\mathbf{f}(\mathbf{u}) = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 - 1 \\ \frac{u_6^2}{\Delta x} - \frac{u_6 u_5}{\Delta x} + \frac{2u_6}{Re\Delta y^2} - \frac{u_{10}}{Re\Delta y^2} - \frac{u_2}{Re\Delta y^2} \\ \frac{u_7^2}{\Delta x} - \frac{u_7 u_6}{\Delta x} + \frac{2u_7}{Re\Delta y^2} - \frac{u_{11}}{Re\Delta y^2} - \frac{u_3}{Re\Delta y^2} \\ u_8 - u_7 \\ u_9 - 1 \\ \frac{u_{10}^2}{\Delta x} - \frac{u_{10} u_9}{\Delta x} + \frac{2u_{10}}{Re\Delta y^2} - \frac{u_{14}}{Re\Delta y^2} - \frac{u_6}{Re\Delta y^2} \\ \frac{u_{11}^2}{\Delta x} - \frac{u_{11} u_{10}}{\Delta x} + \frac{2u_{11}}{Re\Delta y^2} - \frac{u_{15}}{Re\Delta y^2} - \frac{u_7}{Re\Delta y^2} \\ u_{12} - u_{11} \\ u_{13} - 1 \\ u_{14} - 1 \\ u_{15} - 1 \\ u_{16} - 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (22)$$

Considere a matriz Jacobiana \mathbf{J} do sistema definida como

$$\boldsymbol{J} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & [-\frac{1}{Re\Delta y^2}] & 0 & 0 & [-\frac{u_6}{\Delta x}] [\frac{1}{\Delta x}(2u_6 - u_5) + \frac{2}{Re\Delta y^2}] & 0 \\ 0 & 0 & [-\frac{1}{Re\Delta y^2}] & 0 & 0 & [\frac{1}{\Delta x}(2u_7 - u_6) + \frac{2}{Re\Delta y^2}] \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ : & : & : & : & : & : \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (24)$$

Considere o caso em que $\Delta x = \Delta y = 0.1$ e $Re = 100$, essa matriz é dada por

$$\mathbf{J} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -10u_6 & 10(2u_6 - u_5) + 2 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -10u_7 & 10(2u_7 - u_6) + 2 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (25)$$

O método de solução consiste em supor uma condição inicial \mathbf{u}_0 para o vetor \mathbf{u} , substituir essa condição na matriz $\mathbf{J}(\mathbf{u}_0)$ e resolver o seguinte sistema linear

$$\mathbf{J}(\mathbf{u}_0) \cdot \mathbf{u}_0^* = -\mathbf{f}(\mathbf{u}_0) \quad (26)$$

em seguida nós atualizamos o novo valor do vetor \mathbf{u}_1 fazendo

$$\mathbf{u}_1 = \mathbf{u}_0 + \mathbf{u}_0^*, \quad (27)$$

com esse novo valor de \mathbf{u}_1 nós resolvemos o seguinte sistema

$$\mathbf{J}(\mathbf{u}_1) \cdot \mathbf{u}_1^* = -\mathbf{f}(\mathbf{u}_1), \quad (28)$$

em seguida fazemos

$$\mathbf{u}_2 = \mathbf{u}_1 + \mathbf{u}_1^*, \quad (29)$$

e assim sucessivamente até que o método convirja para um valor dentro de uma determinada faixa de tolerância. Podemos resumir o método nos seguintes passos

1. Supor um valor inicial para o vetor \mathbf{u}_0
2. Determinar a matriz Jacobiana $\mathbf{J}(\mathbf{u}_0)$
3. Determinar o vetor $\mathbf{f}(\mathbf{u}_0)$
4. Resolver o sistema linear $\mathbf{J}(\mathbf{u}_0) \cdot \mathbf{u}_0^* = -\mathbf{f}(\mathbf{u}_0)$
5. Corrigir o próximo valor de \mathbf{u}_1 com $\mathbf{u}_1 = \mathbf{u}_0 + \mathbf{u}_0^*$

6. Fazer para j de 1 à N:

- Resolver o sistema linear $\mathbf{J}(\mathbf{u}_j) \cdot \mathbf{u}_j^* = -\mathbf{f}(\mathbf{u}_j)$
- Corrigir o próximo valor de \mathbf{u}_j com $\mathbf{u}_j = \mathbf{u}_j + \mathbf{u}_j^*$

7. Repetir o procedimento até convergir;

Para fins de exemplificação considere o seguinte valor inicial para o vetor \mathbf{u}_0

$$\mathbf{u}_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad (30)$$

Para esses valores iniciais, com $\Delta x = \Delta y = 0.1$ e $Re = 100$, temos que $\mathbf{J}(\mathbf{u}_0)$ é dada por

$$\mathbf{J} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -10 & 12 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -10 & 12 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -10 & 12 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -10 & 12 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (31)$$

e o vetor $\mathbf{f}(\mathbf{u}_0)$ é dado por

$$\mathbf{f}(\mathbf{u}_0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (32)$$

Dessa forma precisamos resolver inicialmente o seguinte sistema

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -10 & 12 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -10 & 12 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -10 & 12 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -10 & 12 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} u_0^{*1} \\ u_0^{*2} \\ u_0^{*3} \\ u_0^{*4} \\ u_0^{*5} \\ u_0^{*6} \\ u_0^{*7} \\ u_0^{*8} \\ u_0^{*9} \\ u_0^{*10} \\ u_0^{*11} \\ u_0^{*12} \\ u_0^{*13} \\ u_0^{*14} \\ u_0^{*15} \\ u_0^{*16} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (33)$$

A solução desse sistema linear pelo método de Jacobi (método iterativo) com 500 iterações fornece

$$\mathbf{u}_0^* = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -0.083916068 \\ -0.1548242 \\ -0.1548242 \\ 0 \\ -0.006993004 \\ -0.018729515 \\ -0.018729515 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (34)$$

Façamos agora a seguinte correção

$$\mathbf{u}_1 = \mathbf{u}_0 + \mathbf{u}_0^*, \quad (35)$$

de modo que

$$\mathbf{u}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0.9160839 \\ 0.8451759 \\ 0.8451759 \\ 1 \\ 0.9930070 \\ 0.9812705 \\ 0.9812705 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad (36)$$

Agora precisamos resolver o seguinte sistema

$$\mathbf{J}(\mathbf{u}_1) \cdot \mathbf{u}_1^* = -\mathbf{f}(\mathbf{u}_1), \quad (37)$$

note que a matriz \mathbf{J} muda a cada nova iteração. Montando o sistema e resolvendo para \mathbf{u}_1^* pelo

mesmo método (Jacobi com 500 iterações) obtemos o novo valor de \mathbf{u}_2 dado por

$$\mathbf{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0.9092013 \\ 0.8277108 \\ 0.8277108 \\ 1 \\ 0.9923854 \\ 0.9790677 \\ 0.9790677 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad (38)$$

Repetindo esse procedimento 10 vezes obtemos \mathbf{u}_{10} , comparando o valor de \mathbf{u}_{10} com o valor de \mathbf{u}_9 temos

$$\mathbf{u}_9 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0.9091543 \\ 0.8274714 \\ 0.8274714 \\ 1 \\ 0.9923811 \\ 0.9790405 \\ 0.9790405 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{u}_{10} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0.9091544 \\ 0.8274716 \\ 0.8274716 \\ 1 \\ 0.9923812 \\ 0.9790406 \\ 0.9790406 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad (39)$$

Para análise da convergência do processo repare no vetor dado por $|\mathbf{u}_{10} - \mathbf{u}_9|$

$$|\mathbf{u}_{10} - \mathbf{u}_9| = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1.1920929E-07 \\ 1.1920929E-07 \\ 1.1920929E-07 \\ 0 \\ 5.9604645E-08 \\ 5.9604645E-08 \\ 5.9604645E-08 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (40)$$

O valor máximo dessa diferença, nesse caso $1.1920929E-07$ pode ser utilizado como o parâmetro que deverá ser menor que determinada tolerância pré-definida pelo usuário para garantir a convergência do método.

1.3 Métodos especiais para polinômios e aplicáveis à raízes complexas

Uma outra classe importante de métodos numéricos voltados à solução de zeros de funções, são os métodos desenvolvidos especificamente para obtenção de raízes de polinômios. Esses métodos são particularmente úteis no sentido em que atacam por uma perspectiva numérica problemas que vem sendo explorados no campo analítico há muitos anos pelos matemáticos. A procura por fórmulas fechadas para obtenção de raízes de polinômios de ordem geral n é uma busca antiga da comunidade matemática, pautada não só pelas diversas aplicações de polinômios na enunciação de problemas práticos, mas também pela elegância e generalidade dos polinômios como estruturas matemáticas.

A fim de explorarmos a relação entre os polinômios e a prática de engenharia, considere um determinado sistema físico, regido por uma equação diferencial ordinária do tipo:

$$a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = F(t), \quad (41)$$

em que a_2, a_1, a_0 representam coeficientes constantes, y representa uma variável cujo comportamento transiente desejamos saber e $F(t)$ é um termo forçante. A equação (41) é um modelo geral para diferentes classes de problemas de engenharia que vão desde a modelagem de um problema de vibrações utilizando um sistema massa-mola-amortecedor até o problema de determinação de correntes transientes é um circuito RLC, também conhecido como um oscilador harmônico. Um caso particular de interesse consiste na solução da equação (41) no contexto em que $F(t) = 0$. Nesse caso, o que estamos afirmando é que o sistema não encontra-se sujeito a uma ação do mundo externo, de tal sorte que a solução dessa equação no limite $F(t) \rightarrow 0$ revelará aspectos fundamentais intrínsecos à estrutura do sistema físico de interesse. Por esse motivo, chamamos essa solução de solução geral e dizemos que a equação diferencial em questão é uma EDO homogênea.

No contexto de uma equação diferencial ordinária de segunda ordem, linear, com coeficientes constantes e homogênea, a solução geral é dada por $y = e^{rt}$, em que r é um coeficiente a ser determinado. Essa forma geral é resultado da cara da nossa equação diferencial que relaciona

a função $y(t)$ que desejamos conhecer com suas derivadas diretamente por meio de coeficientes. A função exponencial é a melhor candidata a solução, uma vez que as derivadas dessa função envolvem ela mesmo a menos de coeficientes decorrentes da aplicação da regra da cadeia. Dessa forma, substituindo a proposta de solução geral $y(t) = e^{rt}$ na equação diferencial de segunda ordem homogênea que desejamos resolver, obtemos:

$$a_2 r^2 e^{rt} + a_1 r e^{rt} + a_0 e^{rt} = 0, \quad (42)$$

o que nos leva à

$$a_2 r^2 + a_1 r + a_0 = 0. \quad (43)$$

A equação (43) é chamada de equação característica ou polinômio característico do sistema. A solução desse polinômio vai gerar valores de r , dependentes dos coeficientes a_0, a_1, a_2 que atendam a física regida por essa equação. Mais ainda, como este é um polinômio de ordem n , em que n é a ordem da equação diferencial a ser resolvida, teremos n raízes para o polinômio característico. Dessa forma, a solução geral da equação homogênea vai ser dada por uma combinação linear de soluções com a cara $y(t) = e^{rt}$ com as diferentes raízes obtidas para r pela solução do polinômio característico. Por exemplo, para o caso em que $n = 2$ e as raízes para r são denotadas por r_1, r_2 , a solução da equação homogênea é dada por:

$$y(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}, \quad (44)$$

em que as constantes c_1 e c_2 são determinadas a partir das condições de contorno do problema.

1.4 O método de Müller

Agora, que a relação entre polinômios e problemas práticos de engenharia foi introduzida, vamos investigar alguns métodos numéricos destinados especialmente à obtenção de raízes de polinômios. O primeiro desses métodos a ser apresentado aqui é o método de Müller, concebido para ser uma extensão do método da secante, em que ao invés de seccionarmos a função $f(x)$, cuja raiz desejamos obter, por uma reta utilizando dois pontos, fazemos isso por meio de uma parábola utilizando três pontos, como ilustrado na figura (3).

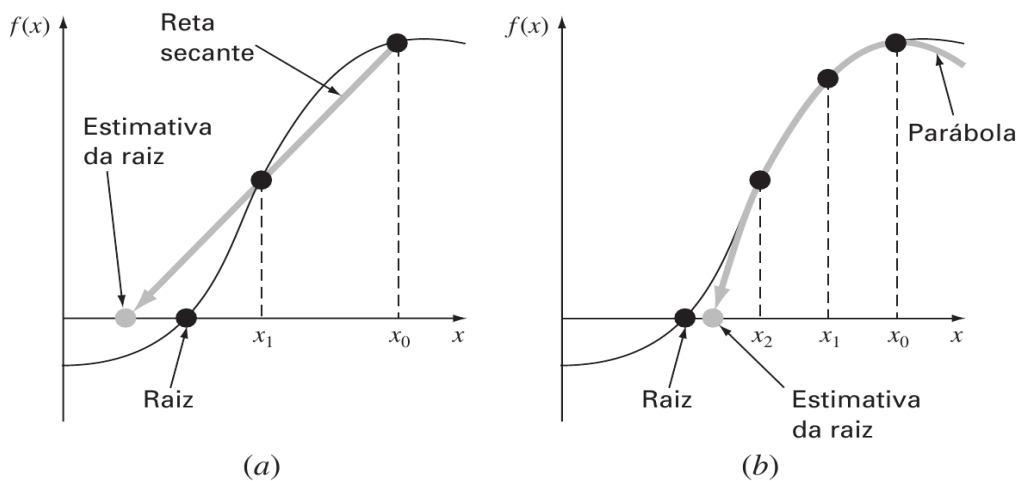


Figure 3: Diferença entre o método da secante (a) e o método de Müller (b).

A essência do método de Müller consiste em descobrir os coeficientes dessa parábola e então projetá-la no eixo x para descobrir onde ela tocará o eixo. Em seguida evoluímos iterativamente

esses passos de construção e projeção da parábola até chegarmos o mais perto possível da raiz. Para compreendermos como o método é construído considere a função $f_2(x) = a(x - x_2)^2 + b(x - x_2) + c$, em que a, b, c são constantes a serem determinadas e x_2 é um ponto ao longo do eixo x que passa sobre a função $f(x)$ no ponto $f(x_2)$. É importante notar aqui que $f(x) \neq f_2(x)$. Quando falamos da função $f(x)$ estamos falando da função que queremos obter a raiz, já $f_2(x)$ é a parábola criada por nós para representar o comportamento de $f(x)$ nas proximidades da raiz que estamos buscando. Dessa forma, o primeiro passo para obtenção dos valores de a, b, c que atendam essa demanda consiste em montar o seguinte sistema de equações:

$$\begin{aligned} f(x_0) &= a(x_0 - x_2)^2 + b(x_0 - x_2) + c \\ f(x_1) &= a(x_1 - x_2)^2 + b(x_1 - x_2) + c \\ f(x_2) &= a(x_2 - x_2)^2 + b(x_2 - x_2) + c, \end{aligned} \quad (45)$$

em que x_0, x_1, x_2 são pontos inicialmente escolhidos pelos quais a parábola que estamos buscando deve passar. Da última equação no sistema (45) é fácil ver que $c = f(x_2)$. Utilizando essa informação e substituindo-a nas outras equações do sistema, após uma série de manipulações algébricas é possível mostrar que a, b, c são dadas por:

$$a = \frac{\delta_1 - \delta_0}{h_1 + h_0}, \quad b = ah_1 + \delta_1, \quad c = f(x_2), \quad (46)$$

em que $h_0, h_1, \delta_0, \delta_1$ são dados por:

$$h_0 = x_1 - x_0, \quad h_1 = x_2 - x_1, \quad \delta_0 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}, \quad \delta_1 = \frac{f(x_2) - f(x_1)}{x_2 - x_1}. \quad (47)$$

Combinando as equações (46) e (47) é possível escrever a, b, c em termos de $x_0, x_1, x_2, f(x_0), f(x_1), f(x_2)$. Isso significa que uma vez que saibamos a cara da função $f(x)$ cujas raízes queremos encontrar, basta utilizarmos as equações (46) e (47) para descobrir os coeficientes a, b, c necessários para montar a parábola que intercepta essa função $f(x)$ nas proximidades da raiz. A finalização do método consiste em retomar a forma de escrevermos a função $f_2(x)$ aplicando a substituição de variáveis $x^* = x - x_2$, de tal sorte que

$$f_2(x) = ax^{*2} + bx^* + c, \quad (48)$$

cujas raízes são

$$x^* = x - x_2 = x_{i+1} - x_i = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \quad (49)$$

Note porém que a aplicação da fórmula quadrática clássica como a conhecemos pode levar a um potencial problema numérico no limite em que $a \rightarrow 0$. Nesse caso, o discriminante $\Delta = b^2 - 4ac \rightarrow b^2$, o que levaria a uma subtração de valores muito próximos para a operação do numerador $-b \pm \sqrt{\Delta}$, o que poderia gerar uma propagação indesejada de erros de arredondamento. Dessa forma, utilizamos para determinação das raízes da parábola $f_2(x^*)$ uma versão alternativa da fórmula quadrática, obtida pela aplicação das fórmulas de Viète, dada por

$$x^* = x_{i+1} - x_i = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}. \quad (50)$$

Essa versão, pouco familiar, da fórmula quadrática é famosa por ser aplicada justamente no método de Müller. Note ainda que a solução que buscamos x_{i+1} evolui de acordo com a seguinte fórmula:

$$x_{i+1} = x_i - \frac{2c}{b \pm \sqrt{b^2 - 4ac}}. \quad (51)$$

Como desejamos caminhar o mais rápido possível rumo à solução, precisamos minimizar as distâncias entre x_{i+1} e x_i . Pelo fato da parábola ter duas raízes, precisamos escolher uma delas para solução que buscamos. Fazemos isso escolhendo a opção de soma no denominador $b + \sqrt{\Delta}$, o que evita a subtração de termos muito próximos no denominador e minimiza a diferença entre x_{i+1} e x_i acelerando a convergência do método na busca da raiz pelo método de projeção da parábola sobre o eixo x . Uma outra vantagem do método é que ele não está restrito apenas a raízes reais, mas vale também para raízes complexas uma vez que não existe nenhuma restrição com relação à positividade do discriminante da parábola construída para representar o comportamento de $f(x)$ nas vizinhanças da raiz.

2 Enunciado da tarefa

Considere um polinômio de ordem 5, dado por

$$f(x) = (x - 1)(x - 3)(x - 5)(x - 7)(x - 9), \quad (52)$$

cujas raízes sabemos ser $x_r = [1, 3, 5, 7, 9]$. Para esse cenário, escreva um programa de computador que busque essas raízes partindo de valores iniciais assumidos pelo usuário que deverão seccionar a função $f(x)$ em dois ou três pontos próximos da raiz e que a partir desses valores encontre a raiz mais próxima tanto pelo método da secante quanto pelo método de Müller. Seu programa deverá comparar a evolução percentual do erro relativo em função do número de iterações para cada um dos métodos e plotar essa informação num gráfico xy .

3 Referências bibliográficas

1. S. C. Chapra, R. P. Canale. “Métodos Numéricos para Engenharia.” McGrawHill, 5a edição (2008): 1-825.