

MS211 - Cálculo Numérico
Variações do Método de Newton para Sistemas
Não-Lineares
(Slides Modificados de M. E. Valle)

Introdução

Na aula anterior, vimos o método de Newton para sistemas não-lineares escritos na forma

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0, \\ f_2(x_1, x_2, \dots, x_n) = 0. \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) = 0, \end{cases}$$

em que x_1, x_2, \dots, x_n são as incógnitas e $f_i : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ ou, equivalentemente,

$$\mathbf{F}(\mathbf{x}) = \mathbf{0},$$

em que $\mathbf{F} : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ possui derivadas contínuas num domínio aberto D .

Método de Newton

Dada uma aproximação inicial $\mathbf{x}^{(0)}$, o método de Newton define a sequência $\{\mathbf{x}^{(k)}\}$ da seguinte forma:

■ Enquanto o critério de parada não for satisfeito, faça:

- 1 Resolva: $\mathbf{J}(\mathbf{x}^{(k)})\mathbf{s}^{(k)} = -\mathbf{F}(\mathbf{x}^{(k)})$.
- 2 Defina: $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)}$.

O método de Newton é computacionalmente caro pois, a cada iteração, requer:

- 1 Avaliação da matriz Jacobiana.
- 2 Resolução de um sistema linear ($\mathcal{O}(n^3)$ operações).

Veremos na aula de hoje algumas variações que, embora não apresentem convergência quadrática, efetuam menos operações ou avaliações da matriz Jacobiana por iterações.

Método de Newton Modificado

No método de Newton modificado usamos simplesmente $\mathbf{J}(\mathbf{x}^{(0)})$ no lugar de $\mathbf{J}(\mathbf{x}^{(k)})$.

Formalmente, dada uma aproximação inicial $\mathbf{x}^{(0)}$, o método de Newton modificado define a sequência $\{\mathbf{x}^{(k)}\}$ como segue:

- Calcule a fatoração LU de $\mathbf{J}(\mathbf{x}^{(0)})$, isto é, $\mathbf{LU} = \mathbf{PJ}(\mathbf{x}^{(0)})$
($\mathcal{O}(n^3)$ operações).
- Enquanto o critério de parada não for satisfeito, faça:
 - 1 Resolva o sistema triangular inferior: $\mathbf{Ly} = -\mathbf{PF}(\mathbf{x}^{(k)})$
($\mathcal{O}(n^2)$ operações).
 - 2 Resolva o sistema triangular superior: $\mathbf{Us}^{(k)} = \mathbf{y}$
($\mathcal{O}(n^2)$ operações).
 - 3 Defina: $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)}$.

Observe que a matriz Jacobiana é avaliada uma única vez. A cada iteração, resolvemos apenas sistemas triangulares!

Exemplo 1

Efetue duas iterações do método de Newton modificado, com $\mathbf{x}^{(0)} = [1, 5]^T$, para determinar a solução dos sistema:

$$\mathbf{F}(x, y) = \begin{bmatrix} x + y - 3 \\ x^2 + y^2 - 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Resposta:

A fatoração LU da matriz Jacobiana em $\mathbf{x}^{(0)}$ satisfaz

$$\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}}_{\mathbf{J}(1,5)} = \underbrace{\begin{bmatrix} 1 & 0 \\ 1/2 & 1 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} 2 & 10 \\ 0 & -4 \end{bmatrix}}_{\mathbf{U}}.$$

O passo $\mathbf{s}^{(0)} = [-1.62, -1.38]^T$ é determinado resolvendo

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 1/2 & 1 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} -17 \\ -3 \end{bmatrix}}_{-\mathbf{PF}(\mathbf{x}^{(0)})} \quad \text{e} \quad \underbrace{\begin{bmatrix} 2 & 10 \\ 0 & -4 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} s_1^{(0)} \\ s_2^{(0)} \end{bmatrix}}_{\mathbf{s}^{(0)}} = \underbrace{\begin{bmatrix} -17 \\ 5.5 \end{bmatrix}}_{\mathbf{y}}.$$

Logo, temos $\mathbf{x}^{(1)} = [-0.62, 3.62]^T$ e $\mathbf{F}(\mathbf{x}^{(1)}) = [0, 4.49]^T$.

O passo $\mathbf{s}^{(1)} = [0.56, -0.56]$ da segunda iteração é a solução do sistema linear $J(1, 5)\mathbf{s}^{(1)} = -\mathbf{F}(\mathbf{x}^{(1)})$.
Portanto, temos

$$\mathbf{x}^{(2)} = \begin{bmatrix} -0.62 \\ 3.62 \end{bmatrix} + \begin{bmatrix} 0.56 \\ -0.56 \end{bmatrix} = \begin{bmatrix} -0.07 \\ 3.05 \end{bmatrix}.$$

Observação: O uso da fatoração LU no método de Newton Modificado está geralmente recomendado para $n \gg 0$. Para pequenos n , não é necessário empregá-la.

Exemplo de uma Aplicação do Método de Newton Modificado em Forma Tabelar

Exemplo:

Utilize o método de Newton Modificado para resolver

$$x_1^2 + x_2^2 = 4 \Leftrightarrow x_1^2 + x_2^2 - 4 = 0$$

$$x_1^2 - x_2 = 0$$

com $x^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $\epsilon_1 = \epsilon_2 = 0.15$

Obtemos a tabela

k	$x^{(k)}$	$F(x^{(k)})$	$\ F(x^{(k)})\ _\infty$	$\ s^{(k-1)}\ _\infty$	s^k
0	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -2 \\ 0 \end{pmatrix}$	$2 \geq 0.15$	-	$\begin{pmatrix} 1/2 \\ 3/2 \end{pmatrix}$
1	$\begin{pmatrix} 4/3 \\ 5/3 \end{pmatrix}$	$\begin{pmatrix} +5/9 \\ +1/9 \end{pmatrix}$	$5/9 \geq 0.15$	$2/3 \rightarrow 0.15$	$\begin{pmatrix} -2/3 \\ -1/3 \end{pmatrix}$
2	$\begin{pmatrix} 1.2057 \\ 1.5185 \end{pmatrix}$	$\begin{pmatrix} -0.24526 \\ -0.069606 \end{pmatrix}$	$0.24526 \geq 0.15$	$4/27$	$\begin{pmatrix} 0.14815 \\ \end{pmatrix}$

Resultado é Aprox. de uma raiz $x = \begin{pmatrix} 1.2057 \\ 1.5185 \end{pmatrix}$ com $\epsilon = 0.15$ PARA:

Continuação do Exemplo

Iteração 2: (Newton Modificado)

$$\text{Se } \|F(x^{(1)})\|_{\infty} = \frac{5}{9} \geq \epsilon_1$$

$$\text{e } \|x^{(1)} - x^{(0)}\|_{\infty} = \frac{2}{3} \geq \epsilon_2$$

$$\text{resolva } J(x^{(1)}) \cdot s^{(1)} = -F(x^{(1)}) = -\frac{1}{9} \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 2 & | & -\frac{5}{9} \\ 2 & -1 & | & -\frac{1}{9} \end{pmatrix} \rightarrow \begin{pmatrix} 18 & 18 & | & -5 \\ 18 & -9 & | & -1 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 18 & 18 & | & -5 \\ 0 & -27 & | & 4 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 18 & 18 & | & -5 \\ 0 & -9 & | & \frac{4}{3} \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 18 & 0 & | & -\frac{7}{3} \\ 0 & 1 & | & -\frac{4}{27} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & | & -\frac{7}{54} \\ 0 & 1 & | & -\frac{4}{27} \end{pmatrix}$$

$$: s^{(1)} = \begin{pmatrix} -\frac{7}{54} \\ -\frac{4}{27} \end{pmatrix}, x^{(2)} = x^{(1)} + s^{(1)} \approx \begin{pmatrix} 1.5185 \\ 1.5185 \end{pmatrix}$$

Um Outro Exemplo

Considere o sistema de equações não-lineares

$$\begin{aligned}x_1^2 + x_2^2 &= 4 \\x_1^2 - x_2 &= 0\end{aligned}$$

Utilizando 4 dígitos decimais, a aplicação do método de Newton com precisão $\varepsilon = 0.02$ e chute inicial $x^{(0)} = (1, 1)^t$ resulta na tabela seguinte. Precisa-se verificar se um dos 2 critérios de parada ($\|F(x^{(k)})\|_\infty < \varepsilon?$; $\|s^{(k-1)}\|_\infty < \varepsilon?$) está satisfeito.

k	$x^{(k)}$	$F(x^{(k)})$	$\ F(x^{(k)})\ _\infty$	$\ s^{(k-1)}\ _\infty$	$s^{(k)}$
0	$(1, 1)^t$	$(-2, 0)^t$	2	–	$(0.3333, 0.6667)^t$
1	$(1.3333, 1.6667)^t$	$(0.5556, 0.1111)^t$	0.5556	0.6667	$(-0.0801, -0.1026)^t$
2	$(1.2532, 1.5641)^t$	$(0.0169, 0.0064)^t$	0.0169	0.1026	

Já que $\|F(x^{(2)})\|_\infty \approx 0.0169 < 0.02 = \varepsilon$, obtemos a aproximação $x^{(2)} \approx (1.2532, 1.5641)^t$ como de uma solução do sistema não-linear.

Algoritmo do Método de Newton Modificado

Entrada: Função não-linear \mathbf{F} , aproximação da solução \mathbf{x} e a matriz Jacobiana avaliada em \mathbf{x} , i.e., $\mathbf{J}_x = \mathbf{J}(\mathbf{x})$.

Dados: Número máximo de interações k_{max} ; τ e ϵ .

Inicialize: $k = 0$, $\mathbf{F}_x = \mathbf{F}(\mathbf{x})$ e $Dr = \tau + 1$.

Calcule a fatoração LU: $\mathbf{LU} = \mathbf{PJ}(\mathbf{x})$.

enquanto $k \leq k_{max}$, $\|\mathbf{F}_x\|_\infty > \epsilon$ e $Dr > \tau$ **faça**

- 1 Atualize: $k = k + 1$.
- 2 Resolva o sistema triangular inferior: $\mathbf{Ly} = -\mathbf{PF}_x$.
- 3 Resolva o sistema triangular superior: $\mathbf{Us} = \mathbf{y}$.
- 4 Atualize: $\mathbf{x} = \mathbf{x} + \mathbf{s}$.
- 5 Calcule: $Dr = \|\mathbf{s}\|_\infty$.
- 6 Avalie: $\mathbf{F}_x = \mathbf{F}(\mathbf{x})$.

fim

Saída: Aproximação para a solução é \mathbf{x} .

Métodos Quase-Newton

Nos métodos quase-Newton, também chamados **métodos secantes**, $\mathbf{J}(\mathbf{x}^{(k)})$ é aproximada por $\mathbf{B}^{(k)} \in \mathbb{R}^{n \times n}$. Dessa forma, evita-se avaliar $\mathbf{J}(\mathbf{x}^{(k)})$ a cada iteração.

Com efeito, conhecemos uma aproximação linear de \mathbf{F} em $\mathbf{x}^{(k)}$:

$$\mathbf{L}_k(\mathbf{x}) = \mathbf{F}(\mathbf{x}^{(k)}) + \mathbf{J}(\mathbf{x}^{(k)})(\mathbf{x} - \mathbf{x}^{(k)}).$$

Substituindo $\mathbf{J}(\mathbf{x}^{(k)})$ por uma aproximação $\mathbf{B}^{(k)}$ obtemos

$$\mathbf{S}_k(\mathbf{x}) = \mathbf{F}(\mathbf{x}^{(k)}) + \mathbf{B}^{(k)}(\mathbf{x} - \mathbf{x}^{(k)}).$$

Para obter \mathbf{x}^{k+1} , faça:

- 1 Resolva $\mathbf{B}^{(k)} \mathbf{s}^k = -\mathbf{F}(\mathbf{x}^{(k)})$;
- 2 $\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{s}^k$.

Métodos Quase-Newton: Como Obter $\mathbf{B}^{(k)}$?

No primeiro lugar, pode-se utilizar $\mathbf{B}^{(0)} = \mathbf{J}(\mathbf{x}^{(0)})$.

Note que para $k > 0$ tem-se

$$\mathbf{S}_k(\mathbf{x}^{(k-1)}) = \mathbf{F}(\mathbf{x}^{(k)}) + \mathbf{B}^{(k)}(\mathbf{x}^{(k-1)} - \mathbf{x}^{(k)}).$$

Seja para $k > 0$. Impondo $\mathbf{S}_k(\mathbf{x}^{(k-1)}) = \mathbf{F}(\mathbf{x}^{(k-1)})$, obtemos:

$$\begin{aligned} \mathbf{F}(\mathbf{x}^{(k-1)}) &= \mathbf{F}(\mathbf{x}^{(k)}) + \mathbf{B}^{(k)}(\mathbf{x}^{(k-1)} - \mathbf{x}^{(k)}) \\ \Leftrightarrow \mathbf{F}(\mathbf{x}^{(k-1)}) - \mathbf{F}(\mathbf{x}^{(k)}) &= \mathbf{B}^{(k)}(\mathbf{x}^{(k-1)} - \mathbf{x}^{(k)}) \\ \Leftrightarrow \mathbf{B}^{(k)}(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}) &= \mathbf{F}(\mathbf{x}^{(k)}) - \mathbf{F}(\mathbf{x}^{(k-1)}). \end{aligned}$$

Métodos Quase-Newton: Como Obter $\mathbf{B}^{(k)}$?

Seja para $k > 0$. Considere a equação seguinte:

$$\mathbf{B}^{(k)}(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}) = \mathbf{F}(\mathbf{x}^{(k)}) - \mathbf{F}(\mathbf{x}^{(k-1)}).$$

Tomando

$$\mathbf{s}^{(k-1)} = \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} \quad \text{e} \quad \mathbf{y}^{(k-1)} = \mathbf{F}(\mathbf{x}^{(k)}) - \mathbf{F}(\mathbf{x}^{(k-1)}),$$

concluimos que $\mathbf{B}^{(k)}$ deve satisfazer

$$\mathbf{B}^{(k)}\mathbf{s}^{(k-1)} = \mathbf{y}^{(k-1)}.$$

Esse é um sistema com n equações e n^2 incógnitas. Logo, essa condição não é suficiente para determinar $\mathbf{B}^{(k)}$.

Os métodos quase-Newton diferem entre si pelas condições adicionais impostas sobre $\mathbf{B}^{(k)}$, tais como:

- Obedecer a um princípio de variação mínima com relação a matriz $\mathbf{B}^{(k)}$ da iteração anterior.
- Preservar uma certa estrutura da matriz Jacobiana como simetria e esparsidade.

Método de Broyden

No método de Broyden, define-se

$$\mathbf{B}^{(k+1)} = \mathbf{B}^{(k)} + \mathbf{u}^{(k)} \mathbf{s}^{(k)T},$$

em que

$$\mathbf{u}^{(k)} = \frac{1}{\mathbf{s}^{(k)T} \mathbf{s}^{(k)}} (\mathbf{y}^{(k)} - \mathbf{B}^{(k)} \mathbf{s}^{(k)}).$$

Concluindo, dada uma aproximação inicial $\mathbf{x}^{(0)}$ e uma matriz $\mathbf{B}^{(0)}$, o método de Broyden define a sequência $\{\mathbf{x}^{(k)}\}$ como segue:

■ Enquanto o critério de parada não for satisfeito, faça:

1 Resolva: $\mathbf{B}^{(k)} \mathbf{s}^{(k)} = -\mathbf{F}(\mathbf{x}^{(k)})$.

2 Defina: $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)}$.

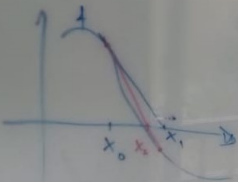
3 Calcule:

$$\mathbf{B}^{(k+1)} = \mathbf{B}^{(k)} + \mathbf{u}^{(k)} \mathbf{s}^{(k)T} = \mathbf{B}^{(k)} + \frac{1}{\mathbf{s}^{(k)T} \mathbf{s}^{(k)}} \mathbf{F}(\mathbf{x}^{(k+1)}) \mathbf{s}^{(k)T}.$$

O Método de Broyden: Outras Informações

$$0 = S_k(x) = F(x^{(k)}) + B^{(k)}(x - x^{(k)})$$

$$\Leftrightarrow -F(x^{(k)}) = B^{(k)}(x - x^{(k)})$$



Seja $k=0$ e $B^{(0)} = J(x^{(0)})$
 Enquanto o critério de parada não é satisfatório, faça:

1. Resolva $B^{(k)} \cdot s^{(k)} = -F(x^{(k)})$
2. $x^{(k+1)} = x^{(k)} + s^{(k)}$
3. $B^{(k+1)}$ = Atualização de $B^{(k)}$

Faça $k=k+1$ e volte ao Passo 1.

Broyden: Para $k=0,1,\dots$

$$B^{(k+1)} = B^{(k)} + u^{(k)} (s^{(k)})^T$$

$$B^{(k+1)} \cdot s^{(k+1)} = B^{(k)} \cdot s^{(k+1)} + u^{(k)} (s^{(k)})^T \cdot s^{(k+1)}$$

$$= B^{(k)} \cdot s^{(k+1)} + u^{(k)} \|s^{(k)}\|^2$$

$$= \gamma^{(k)} \cdot s^{(k)} + (y^{(k)} - B^{(k)} \cdot s^{(k)})$$

$$= y^{(k)}$$

Constante
 Note que $B^{(k+1)} \cdot s^{(k+1)} = y^{(k+1)}$ para $k=0,1,2,\dots$ é equivalente a $B^{(k)} \cdot s^{(k)} = y^{(k)}$ para $k=1,2,\dots$

Exemplo 2

Considere o sistema não-linear

$$\mathbf{F}(x, y) = \begin{bmatrix} x + y - 3 \\ x^2 + y^2 - 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Determine $\mathbf{x}^{(2)}$ produzido pelo método de Broyden com

$$\mathbf{x}^{(0)} = [1, 5]^T \quad \text{e} \quad \mathbf{B}^{(0)} = \mathbf{J}(\mathbf{x}^0) = \begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}.$$

Resposta: Na primeira iteração, encontramos

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}}_{\mathbf{B}^{(0)}} \underbrace{\begin{bmatrix} s_1^{(0)} \\ s_2^{(0)} \end{bmatrix}}_{\mathbf{s}^{(0)}} = \underbrace{\begin{bmatrix} -3 \\ -17 \end{bmatrix}}_{-\mathbf{F}(\mathbf{x}^{(0)})} \implies \mathbf{s}^{(0)} = \begin{bmatrix} -1.62 \\ -1.38 \end{bmatrix}.$$

$$\mathbf{x}^{(1)} = \underbrace{\begin{bmatrix} 1 \\ 5 \end{bmatrix}}_{\mathbf{x}^{(0)}} + \underbrace{\begin{bmatrix} -1.62 \\ -1.38 \end{bmatrix}}_{\mathbf{s}^{(0)}} = \begin{bmatrix} -0.62 \\ 3.62 \end{bmatrix}.$$

$$\begin{aligned} \mathbf{B}^{(1)} &= \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}}_{\mathbf{B}^{(0)}} + \underbrace{\frac{1}{4.49}}_{\mathbf{s}^{(0)T} \mathbf{s}^{(0)}} \underbrace{\begin{bmatrix} 0 \\ 4.49 \end{bmatrix}}_{\mathbf{F}(\mathbf{x}^{(1)})} \underbrace{\begin{bmatrix} -1.62 & -1.38 \end{bmatrix}}_{\mathbf{s}^{(0)T}} \\ &= \begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix} + \frac{1}{4.49} \begin{bmatrix} 0 & 0 \\ -7.27 & -6.20 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0.38 & 8.62 \end{bmatrix}. \end{aligned}$$

Na segunda iteração, temos:

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0.38 & 8.62 \end{bmatrix}}_{\mathbf{B}^{(1)}} \underbrace{\begin{bmatrix} s_1^{(1)} \\ s_2^{(1)} \end{bmatrix}}_{\mathbf{s}^{(1)}} = \underbrace{\begin{bmatrix} 0 \\ -4.49 \end{bmatrix}}_{-\mathbf{F}(\mathbf{x}^{(1)})} \implies \mathbf{s}^{(1)} = \begin{bmatrix} 0.55 \\ -0.55 \end{bmatrix}.$$

Assim,

$$\mathbf{x}^{(2)} = \underbrace{\begin{bmatrix} -0.62 \\ 3.62 \end{bmatrix}}_{\mathbf{x}^{(1)}} + \underbrace{\begin{bmatrix} 0.55 \\ -0.55 \end{bmatrix}}_{\mathbf{s}^{(1)}} = \begin{bmatrix} -0.07 \\ 3.07 \end{bmatrix}$$

Considerações Finais

Na aula de hoje, apresentamos:

- Método de Newton modificado, em que avaliamos e fatoramos a matriz Jacobiana apenas uma vez na inicialização.
- Método de Broyden, que é um método secante. Aqui, a matriz Jacobiana é aproximada por uma matriz $\mathbf{B}^{(k)}$.

Ambos os métodos não possuem a convergência quadrática do método de Newton. Contudo, eles podem ser vantajosos pois requerem menos operações por iteração.

Em particular, o método de Broyden pode ser implementado sem resolver do sistema $\mathbf{B}^{(k)} \mathbf{s}^{(k)} = -\mathbf{F}(\mathbf{x}^{(k)})$ explicitamente a cada iteração usando a chamada fórmula de Sherman-Morrison.