Aula 11
Variações do Método de Newton para Sistemas Não-Lineares.

MS211 - Cálculo Numérico

Marcos Eduardo Valle

Departamento de Matemática Aplicada
Instituto de Matemática, Estatística e Computação Científica
Universidade Estadual de Campinas

Introdução

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

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

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

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

em que $\mathbf{F} : D \in \mathbb{R}^n \to \mathbb{R}^n$ 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 requer, a cada iteração:

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, são computacionalmente mais baratos.

Método de Newton Modificado

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

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

- Calcule a fatoração LU: $\mathbf{LU} = \mathbf{PJ}(\mathbf{x}^{(0)})$ ($\mathcal{O}(n^3)$ 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 $\mathbf{F}(\mathbf{x}^{(1)}) = [0, 4.53]^T$.
O passo $\mathbf{s}^{(1)} = [0.57, -0.57]$ da segunda iteração é a solução do sistema linear $J(1,5)\mathbf{s}^{(1)} - \mathbf{F}(\mathbf{x}^{(1)})$.
Portanto, temos

$$\mathbf{x}^{(2)} = \begin{bmatrix} -0.62 \\ -3.62 \end{bmatrix} + \begin{bmatrix} 0.57 \\ -0.57 \end{bmatrix} = \begin{bmatrix} -0.06 \\ 3.06 \end{bmatrix}.$$

Algoritmo do Método de Newton Modificado

Entrada: Função não-linear $\mathbf{F}$ e sua matriz Jacobiana $\mathbf{J}$;
Aproximação da solução $\mathbf{x}$.
Dados: Número máximo de interações k_{max}; tolerâncias τ e ϵ.
Inicialize: $k = 0$, $\mathbf{F_x} = \mathbf{F}(\mathbf{x})$ *e* $Er = \tau + 1$.
Calcule a fatoração LU: $\mathbf{LU} = \mathbf{PJ}(\mathbf{x}^{(0)})$.
enquanto $k \leq k_{max}$, $\|\mathbf{F_x}\|_\infty > \epsilon$ *e* $Er > \tau$ **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: $Er = \|\mathbf{s}\|_\infty$.
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**, a matriz Jacobina $\mathbf{J}(\mathbf{x}^{(k)})$ é aproximada por uma certa matriz $\mathbf{B}^{(k)} \in \mathbb{R}^{n \times n}$.

Dessa forma, evita-se avaliar $\mathbf{J}(\mathbf{x}^{(k)})$ a cada iteração.

Com efeito, sabemos que a aproximação linear de $\mathbf{F}$ em $\mathbf{x}^{(k)}$ é

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

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

$$\mathbf{L}(\mathbf{x}^{(k+1)}) = \mathbf{F}(\mathbf{x}^{(k+1)}),$$

encontramos:

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

Equivalentemente, temos

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

Tomando

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

Concluímos que $\mathbf{B}^{(k+1)}$ deve satisfazer

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

Esse é um sistema com n equações e n^2 incógnitas. Logo, essa condição não é suficiente suficiente.

Os métodos quase-Newton diferem entre si pelas condições adicionais impostas sobre $\mathbf{B}^{(k+1)}$, 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: $\mathbf{B}^{(k+1)} = \mathbf{B}^{(k)} + \frac{1}{{\mathbf{s}^{(k)}}^T \mathbf{s}^{(k)}} \mathbf{F}(\mathbf{x}^{(k+1)}) {\mathbf{s}^{(k)}}^T$.

Exemplo 2

Determine $\mathbf{x}^{(2)}$ produzido pelo método de Broyden, com $\mathbf{x}^{(0)} = [1, 5]^T$ e $\mathbf{B}^{(0)} = \begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}$, para determinar a solução do sistema:

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

Resposta: Na primeira iteração, encontramos

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

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

$$\mathbf{B}^{(1)} = \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}}_{\mathbf{B}^{(0)}} + \frac{1}{\underbrace{4.53}_{\mathbf{s}^{(k)^T}\mathbf{s}^{(k)}}} \underbrace{\begin{bmatrix} 0 \\ 4.53 \end{bmatrix}}_{\mathbf{F}(\mathbf{x}^{(1)})} \underbrace{\begin{bmatrix} -1.62 & -1.38 \end{bmatrix}}_{\mathbf{s}^{(0)}}$$

$$= \begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix} + \frac{1}{4.53} \begin{bmatrix} 0 & 0 \\ -7.36 & -6.23 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0.38 & 8.62 \end{bmatrix}.$$

Na segunda iteração, temos:

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0.38 & 8.62 \end{bmatrix}}_{\mathbf{B}^{(1)}} \underbrace{\begin{bmatrix} s_1^{(1)} \\ s_2^{(1)} \end{bmatrix}}_{\mathbf{s}^{(1)}} = \underbrace{\begin{bmatrix} 0 \\ 4.53 \end{bmatrix}}_{\mathbf{F}(\mathbf{x}^{(1)})} \quad \Longrightarrow \quad \mathbf{s}^{(1)} = \begin{bmatrix} 0.55 \\ -0.55 \end{bmatrix}.$$

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 na inicialização.
- Métode de Broyden, que é um metodo secante. Aqui, a matriz Jacobiana é aproximada por uma matriz $\mathbf{B}^{(k)}$.

Em ambos os métodos não possuem a convergência quadrática do método de Newton mas, 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.