Aula 04 - Métodos Quase-Newton

Élton Fontana

Uma das principais dificuldades em aplicar o método de Newton para a resolução de sistemas não-lineares é que o método exige que, a cada iteração, seja determinada a matriz Jacobiana e um sistema linear seja resolvido. Por exemplo, para um sistema com n equações, a determinação da matriz Jacobiana envolve a determinação de n^2 derivadas parciais e o sistema linear gerado em cada iteração será $n \times n$. Para sistemas com muitas equações ou para casos onde as derivadas não podem ser facilmente obtidas, a aplicação do método de Newton é bastante custosa.

Além disso, o método de Newton pode apresentar diversos outros problemas, como por exemplo casos onde a matriz Jacobiana é singular e uma grande dependência da condição inicial utilizada. Diversas modificações foram propostas para melhorar determinados aspectos do método de Newton, em particular para evitar a resolução de um sistema linear a cada iteração. De forma geral, estes métodos são conhecidos como métodos *quase-Newton* e consistem em aproximar a matriz Jacobiana $J(\mathbf{x}^i)$ por alguma outra matriz $\mathbf{B}^i$. A forma como esta matriz é determinada depende do método utilizado. O mais simples destes métodos é o de Newton modificado, que será discutido a seguir.

1 Método de Newton Modificado

O método de Newton possui a vantagem de utilizar o valor da derivada da função em um ponto para estimar o próximo ponto, o que faz com que a convergência do método tenda a ser rápida. Em contrapartida, isto implica em determinar a derivada da função em cada ponto, ou para o caso de sistemas, o valor da matriz Jacobiana em cada ponto. Como consequência, para o caso de sistemas não-lineares, a cada iteração é preciso resolver um sistema linear, o que faz com que o gasto associado à resolução seja alto.

Por exemplo, para uma única equação $f(x) = 0$, o método de Newton pode ser expresso

como:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

Partindo de uma solução inicial x_0 , procede-se com as iterações para refinar a raiz. O método de Newton modificado consiste em considerar $f'(x_i) = f'(x_0)$ para todas as iterações, ou seja, manter o valor da derivada constante e igual ao valor associado à condição inicial. Este método tende a possuir uma convergência muito pior que o método de Newton original, porém não necessita a avaliação da derivada em cada ponto.

Assim, o método de Newton modificado, para encontrar a solução de $f(x) = 0$, é expresso como:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_0)}$$

Exemplo 01: Compare a taxa de convergência do método de Newton e do método de Newton modificado para a busca das raízes da seguinte equação:

$$x^2 - 6 \sin(x) + e^x = 0$$

Considere como condição inicial $x_0 = 2$. A solução exata para o problema é $x = 1.38211$.

1.1 Método de Newton Modificado para Sistemas Não-Lineares

Para o caso da busca de raízes de equações não-lineares, a utilização do método de Newton modificado não se justifica, pois uma vez determinada a derivada de $f(x)$ não existe um gasto grande associado a determinação de $f'(x_i)$. No entanto, para o caso de sistemas não-lineares, cada nova iteração exige a atualização da matriz Jacobiana e a resolução de um novo sistema linear. O método de Newton modificado aplicado para sistemas, de forma semelhante, consiste em considerar a matriz Jacobiana constante e igual ao valor calculado com base na condição inicial.

Com visto anteriormente, o método de Newton para sistemas é expresso como:

$$J(\mathbf{x}_i) \cdot (\mathbf{x}_{i+1} - \mathbf{x}_i) = -\mathbf{F}(\mathbf{x}_i) \quad \rightarrow \quad \mathbf{x}_{i+1} = \mathbf{x}_i - J^{-1}(\mathbf{x}_i)\mathbf{F}(\mathbf{x}_i)$$

Assim, o método de Newton modificado para este caso resulta em:

$$J(\mathbf{x}_0) \cdot (\mathbf{x}_{i+1} - \mathbf{x}_i) = -\mathbf{F}(\mathbf{x}_i) \quad \rightarrow \quad \mathbf{x}_{i+1} = \mathbf{x}_i - J^{-1}(\mathbf{x}_0)\mathbf{F}(\mathbf{x}_i)$$

Para a aplicação do método de Newton, na maioria dos casos é mais conveniente avaliar o sistema linear do que calcular a inversa da matriz Jacobiana. No entanto, para

o método de Newton modificado, como $J^{-1}(\mathbf{x}_0)$ não irá mudar ao longo das iterações, é mais conveniente determinar a matriz inversa para a condição inicial. Com isso, não passa a ser mais necessário resolver um sistema linear a cada iteração.

Exemplo 02: Considere o seguinte sistema não-linear:

$$\begin{aligned}x + y &= 3 \\x^2 + y^2 &= 9\end{aligned}$$

Compare a convergência do método de Newton e do método de Newton modificado, usando como condição inicial $\mathbf{x}^0 = (x_0, y_0) = (1, 5)$. As possíveis soluções para este problema são os pontos $(0, 3)$ e $(3, 0)$.

2 Método de Broyden (Secantes)

O método de Broyden pode ser entendido como uma extensão do método da secante para sistemas não-lineares. No método da secante para a busca das raízes de uma função $f(x) = 0$, o valor da derivada no ponto x_i é avaliado por meio de uma aproximação com diferenças finitas¹

$$f'(x_i) = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$$

Após esta etapa, o ponto x_{i+1} é obtido da mesma forma que através do método de Newton:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

Desta forma, o método da secante não necessita a determinação da derivada da função no ponto x_i . Como desvantagem, este método precisa de dois pontos iniciais para começar o processo iterativo e é bastante sensível com relação às condições adotadas.

Para o caso de um sistema de n equações, a matriz Jacobiana é aproximada de uma forma semelhante:

$$J(\mathbf{x}_i) = \frac{\mathbf{F}(\mathbf{x}_i) - \mathbf{F}(\mathbf{x}_{i-1})}{\mathbf{x}_i - \mathbf{x}_{i-1}}$$

Por simplicidade, pode-se definir:

$$\Delta \mathbf{F}_i = \mathbf{F}(\mathbf{x}_i) - \mathbf{F}(\mathbf{x}_{i-1}) \quad \Delta \mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_{i-1}$$

¹Sugestão: mostre que esta aproximação é um resultado de uma expansão em séries de Taylor truncada no primeiro termo.

Assim, a relação anterior pode ser dada por:

$$J(\mathbf{x}_i) \cdot \Delta \mathbf{x}_i = \Delta \mathbf{F}_i$$

Partindo de uma estimativa inicial para a matriz Jacobiana, pode-se usar a relação anterior para atualizar o valor de $J(\mathbf{x}_i)$ a cada iteração do método de Newton. A princípio, a relação anterior não possui solução única. O método de Broyden consiste em atualizar a matriz Jacobiana de modo que o novo valor obtido $J(\mathbf{x}_{i+1})$ seja o mais próximo possível de $J(\mathbf{x}_i)$ e que satisfaça a relação anterior. Neste contexto, o termo “próximo” significa o valor que minimiza a norma:

$$\|J(\mathbf{x}_i) - J(\mathbf{x}_{i+1})\|$$

Esta condição implica que:

$$J_i = J_{i-1} + \frac{\Delta \mathbf{F}_i - J_{i-1} \Delta \mathbf{x}_i}{\Delta \mathbf{x}_i^T \Delta \mathbf{x}_i} \Delta \mathbf{x}_i^T$$

onde, por conveniência, definiu-se $J_{i-1} = J(\mathbf{x}_{i-1})$ e $J_i = J(\mathbf{x}_i)$.

Para a utilização do método de Newton, a relação anterior pode ser utilizada quando a determinação da matriz Jacobiana não é trivial (quando as funções não possuem derivadas parciais conhecidas, por exemplo). No entanto, não elimina a necessidade da resolução de um sistema linear a cada iteração. Como visto para o método de Newton modificado, uma maneira de evitar a resolução do sistema é através do uso da inversa da matriz Jacobiana:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - J_i^{-1} \mathbf{F}(\mathbf{x}_i)$$

Utilizando a fórmula de Sherman-Morrison em conjunto com a relação anterior para atualizar a matriz Jacobiana, pode-se obter uma relação para atualizar a inversa da matriz Jacobiana:

$$J_i^{-1} = J_{i-1}^{-1} + \frac{\Delta \mathbf{x}_i - J_{i-1}^{-1} \Delta \mathbf{F}_i}{\Delta \mathbf{F}_i^T \Delta \mathbf{F}_i} \Delta \mathbf{F}_i^T J_{i-1}^{-1}$$

As etapas envolvidas na resolução de um sistema linear pelo método de Broyden podem ser resumidas da seguinte forma:

- Considere um sistema linear com n equações da forma:

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f_1(x_1, x_2, x_3, \dots, x_n) \\ f_2(x_1, x_2, x_3, \dots, x_n) \\ \vdots \\ f_n(x_1, x_2, x_3, \dots, x_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- Utilize uma estimativa inicial para a solução $\mathbf{x}_0$ e para a matriz Jacobiana neste ponto J_0 (quando possível, determine as derivadas parciais e avalie os valores em $\mathbf{x}^0$);
- Determine a inversa de J_0 (J_0^{-1});
- Com base em J_0^{-1} , calcule $\mathbf{x}_1$:

$$\mathbf{x}_1 = \mathbf{x}_0 - J_0^{-1}\mathbf{F}(\mathbf{x}_0)$$

- Atualize a matriz Jacobiana para obter J_1^{-1} :

$$J_1^{-1} = J_0^{-1} + \frac{\Delta\mathbf{x}_1 - J_0^{-1}\Delta\mathbf{F}_1}{\Delta\mathbf{F}_1^T \Delta\mathbf{F}_1} \Delta\mathbf{F}_1^T J_0^{-1}$$

Lembrando que $\Delta\mathbf{F}_1 = \mathbf{F}(\mathbf{x}_1) - \mathbf{F}(\mathbf{x}_0)$ e $\Delta\mathbf{x}_1 = \mathbf{x}_1 - \mathbf{x}_0$, portanto, estes valores são conhecidos.

- Com base em J_1^{-1} , determine $\mathbf{x}_2$:

$$\mathbf{x}_2 = \mathbf{x}_1 - J_1^{-1}\mathbf{F}(\mathbf{x}_1)$$

- Atualize novamente a matriz Jacobiana para obter J_2^{-1} e siga com o procedimento até obter convergência.

Exemplo 03: Utilize o método de Broyden para obter uma solução aproximada para o seguinte sistema não-linear:

$$x + y = 3$$

$$x^2 + y^2 = 9$$

Utilize como condição inicial $\mathbf{x}^0 = (x_0, y_0) = (1, 5)$.