Equações não lineares

Leonardo F. Guidi

DMPA – IME
UFRGS

Cálculo Numérico

Índice

raizes de polinômios

As equações não lineares constituídas por polinômios de grau $n \in \mathbb{N}$ com coeficientes complexos $a_n, a_{n-1}, \ldots a_0$ (equações algébricas):

$$p(x) \doteq a_n x^n + a_{n-1} x^{n-1} + \ldots a_1 x + a_0 = 0,$$

dispõe de vários métodos para determinar aproximações para suas raízes. Esses métodos foram desenvolvidos a partir da própria estrutura matemática dos polinômios.

raizes de polinômios

As equações não lineares constituídas por polinômios de grau $n \in \mathbb{N}$ com coeficientes complexos $a_n, a_{n-1}, \ldots a_0$ (equações algébricas):

$$p(x) \doteq a_n x^n + a_{n-1} x^{n-1} + \ldots a_1 x + a_0 = 0,$$

dispõe de vários métodos para determinar aproximações para suas raízes. Esses métodos foram desenvolvidos a partir da própria estrutura matemática dos polinômios.

O teorema fundamental da álgebra garante que a equação $p(x) = 0$ possui n soluções (denominadas raízes de $p(x)$) no plano complexo, x_1^*, x_2^*, ..., $x_n^* \in \mathbb{C}$.

raizes de polinômios

As equações não lineares constituídas por polinômios de grau $n \in \mathbb{N}$ com coeficientes complexos $a_n, a_{n-1}, \ldots a_0$ (equações algébricas):

$$p(x) \doteq a_n x^n + a_{n-1} x^{n-1} + \ldots a_1 x + a_0 = 0,$$

dispõe de vários métodos para determinar aproximações para suas raízes. Esses métodos foram desenvolvidos a partir da própria estrutura matemática dos polinômios.

O teorema fundamental da álgebra garante que a equação $p(x) = 0$ possui n soluções (denominadas raízes de $p(x)$) no plano complexo, x_1^*, x_2^*, ..., $x_n^* \in \mathbb{C}$. Portanto, $p(x)$ pode ser reescrito como

$$p(x) \equiv a_n (x - x_1^*)(x - x_2^*) \ldots (x - x_n^*).$$

raizes de polinômios

As equações não lineares constituídas por polinômios de grau $n \in \mathbb{N}$ com coeficientes complexos $a_n, a_{n-1}, \ldots a_0$ (equações algébricas):

$$p(x) \doteq a_n x^n + a_{n-1} x^{n-1} + \ldots a_1 x + a_0 = 0,$$

dispõe de vários métodos para determinar aproximações para suas raízes. Esses métodos foram desenvolvidos a partir da própria estrutura matemática dos polinômios.

O teorema fundamental da álgebra garante que a equação $p(x) = 0$ possui n soluções (denominadas raízes de $p(x)$) no plano complexo, x_1^*, x_2^*, ..., $x_n^* \in \mathbb{C}$. Portanto, $p(x)$ pode ser reescrito como

$$p(x) \equiv a_n (x - x_1^*)(x - x_2^*) \ldots (x - x_n^*).$$

A partir dessa estrutura é possível desenvolver vários métodos específicos para determinar as raízes.

raizes de polinômios

A estrutura dos polinômios também permite desenvolver critérios para a localização das raízes. Em uma equação não linear geral, esse questão pode ser mais complexa.

raizes de polinômios

Teorema

Seja $x*$ qualquer raíz do polinômio

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0.$$

Então

$$\begin{aligned}
|x*| &\leq \max\left\{\left|\frac{a_0}{a_n}\right|, 1 + \left|\frac{a_1}{a_n}\right|, 1 + \left|\frac{a_2}{a_n}\right|, \ldots, 1 + \left|\frac{a_{n-1}}{a_n}\right|\right\}, \\
|x*| &\leq \max\left\{1, \sum_{j=0}^{n-1}\left|\frac{a_j}{a_n}\right|\right\}, \\
|x*| &\leq 2\max\left\{\left|\frac{a_{n-1}}{a_n}\right|, \left|\frac{a_{n-2}}{a_n}\right|^{\frac{1}{2}}, \left|\frac{a_{n-3}}{a_n}\right|^{\frac{1}{3}}, \ldots, \left|\frac{a_0}{a_n}\right|^{\frac{1}{n}}\right\}, \\
|x*| &\leq \sum_{j=0}^{n-1}\left|\frac{a_j}{a_{j+1}}\right|.
\end{aligned}$$

Newton-Raphson modificado

Uma vez definida a região em que as raízes se encontram no plano complexo, escolhemos um ponto nessa região $x^{(0)}$ como aproximação inicial e utilizamos o método de Newton-Raphson usual :

$$x^{(j+1)} = x^{(j)} - \frac{p\left(x^{(j)}\right)}{p'\left(x^{(j)}\right)}.$$

Newton-Raphson modificado

Uma vez definida a região em que as raízes se encontram no plano complexo, escolhemos um ponto nessa região $x^{(0)}$ como aproximação inicial e utilizamos o método de Newton-Raphson usual :

$$x^{(j+1)} = x^{(j)} - \frac{p\left(x^{(j)}\right)}{p'\left(x^{(j)}\right)}.$$

Uma vez determinada a aproximação para a primeira raiz x_1, temos que as demais raízes são também solução do novo polinômio $p_1(x) = \dfrac{p(x)}{x - x_1}$, pois de acordo com o teorema fundamental da álgebra, se x_1, x_2, ..., x_n são raízes de $p(x) = \prod_{j=1}^{n}(x - x_j)$, então, por construção o polinômio $p_1(x)$ possuirá raízes x_2, x_3, ..., x_n.

Newton-Raphson modificado

Uma vez definida a região em que as raízes se encontram no plano complexo, escolhemos um ponto nessa região $x^{(0)}$ como aproximação inicial e utilizamos o método de Newton-Raphson usual :

$$x^{(j+1)} = x^{(j)} - \frac{p\left(x^{(j)}\right)}{p'\left(x^{(j)}\right)}.$$

Uma vez determinada a aproximação para a primeira raiz x_1, temos que as demais raízes são também solução do novo polinômio $p_1(x) = \dfrac{p(x)}{x - x_1}$, pois de acordo com o teorema fundamental da álgebra, se x_1, x_2, ..., x_n são raízes de $p(x) = \prod_{j=1}^{n}(x - x_j)$, então, por construção o polinômio $p_1(x)$ possuirá raízes x_2, x_3, ..., x_n.
Portanto para determinar as raízes seguintes utilizamos a regra de Newton-Raphson para a equação $p_1(x) = 0$:

$$x^{(j+1)} = x^{(j)} - \frac{p_1\left(x^{(j)}\right)}{p_1'\left(x^{(j)}\right)}.$$

Newton-Raphson modificado

Porém dado que

$$\frac{p_1\left(x^{(j)}\right)}{p_1'\left(x^{(j)}\right)} = \frac{p\left(x^{(j)}\right)}{p'\left(x^{(j)}\right) - p\left(x^{(j)}\right)\dfrac{1}{x^{(j)} - x_1}}$$

Newton-Raphson modificado

Porém dado que

$$\frac{p_1\left(x^{(j)}\right)}{p_1'\left(x^{(j)}\right)} = \frac{p\left(x^{(j)}\right)}{p'\left(x^{(j)}\right) - p\left(x^{(j)}\right)\dfrac{1}{x^{(j)} - x_1}}$$

podemos montar a regra a partir do polinômio original $p(x)$ e da raiz conhecida x_1:

$$x^{(j+1)} = x^{(j)} - \frac{p\left(x^{(j)}\right)}{p'\left(x^{(j)}\right) - p\left(x^{(j)}\right)\dfrac{1}{x^{(j)} - x_1}}$$

e de uma outra aproximação inicial para determinar a nova raiz x_2.

Newton-Raphson modificado

Esse procedimento pode ser repetido sucessivamente e uma vez que conheçamos k raízes de $p(x)$, a $k+1$-ésima raíz pode ser determinada através da iteração

$$x^{(j+1)} = x^{(j)} - \frac{p\left(x^{(j)}\right)}{p'\left(x^{(j)}\right) - p\left(x^{(j)}\right)\sum_{m=1}^{k} \dfrac{1}{x^{(j)} - x_m}}.$$

Newton-Raphson modificado

Observação

Naturalmente, as propriedades de convergência desse método são as mesmas do método de Newton-Raphson, ou seja, se a raiz for simples, a convergência será quadrática, se for múltipla a convergência será linear apenas.

Sistemas de equações não lineares

Seja um domínio $U \subseteq \mathbb{R}^n$ e $F : \mathbb{U} \to \mathbb{R}^n$ um função contínua e diferenciável. O sistema de equações

$$F(\mathbf{x}^*) = \mathbf{0} \in \mathbb{R}^n,$$

onde a solução, $\mathbf{x}^* \in \mathbb{R}^n$, é representada por um vetor de componentes $x_1^*, x_2^*, \ldots, x_n^*$ e F possui componentes $F_1(\mathbf{x})$, $F_2(\mathbf{x})$, ..., $F_n(\mathbf{x}) \in \mathbb{R}$, cada uma delas é uma função contínua e diferenciável de n variáveis.

Sistemas de equações não lineares

Seja um domínio $U \subseteq \mathbb{R}^n$ e $F : \mathbb{U} \to \mathbb{R}^n$ um função contínua e diferenciável. O sistema de equações

$$F(\mathbf{x}^*) = \mathbf{0} \in \mathbb{R}^n,$$

onde a solução, $\mathbf{x}^* \in \mathbb{R}^n$, é representada por um vetor de componentes $x_1^*, x_2^*, \ldots, x_n^*$ e F possui componentes $F_1(\mathbf{x})$, $F_2(\mathbf{x})$, ..., $F_n(\mathbf{x}) \in \mathbb{R}$, cada uma delas é uma função contínua e diferenciável de n variáveis.

A natureza do problema é mais complexa, pois conforme aumentamos a dimensão do sistema haverá uma maior possibilidade de “riqueza” no comportamento das soluções.

Sistemas de equações não lineares

Seja um domínio $U \subseteq \mathbb{R}^n$ e $F : \mathbb{U} \to \mathbb{R}^n$ um função contínua e diferenciável. O sistema de equações

$$F(\mathbf{x}^*) = \mathbf{0} \in \mathbb{R}^n,$$

onde a solução, $\mathbf{x}^* \in \mathbb{R}^n$, é representada por um vetor de componentes $x_1^*, x_2^*, \ldots, x_n^*$ e F possui componentes $F_1(\mathbf{x})$, $F_2(\mathbf{x})$, ..., $F_n(\mathbf{x}) \in \mathbb{R}$, cada uma delas é uma função contínua e diferenciável de n variáveis.

A natureza do problema é mais complexa, pois conforme aumentamos a dimensão do sistema haverá uma maior possibilidade de "riqueza" no comportamento das soluções.

Devemos lembrar que cada equação define uma superfície contida em um espaço n-dimensional e as soluções são as intersecções de todas as superfícies.

Sistemas de equações não lineares

Como exemplo, vamos considerar um caso em dimensão dois. Seja $F(x,y) = (F_1(x,y), F_2(x,y))$ onde $F_1(x,y) = \cos(x)\cos(y) - 0.1$ e $F_2(x,y) = \text{sen}(x)\text{sen}(y) - 0.5$. Abaixo estão os gráficos dessas funções:

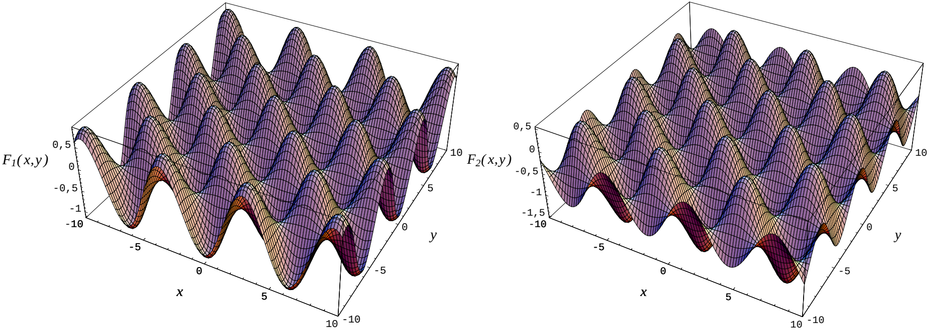


Figura: Gráficos das funções $F_1(x,y)$e $F_2(x,y)$

Sistemas de equações não lineares

Cada equação $F_1(x, y) = 0$ e $F_2(x, y) = 0$ determina um conjunto de curvas (superfícies contidas em um espaço de dimensão dois) no plano $x \times y$. As soluções de $F(\mathbf{x}) = \mathbf{0} \in \mathbb{R}^2$, ou dito de outra forma, do sistema

$$\begin{cases} F_1(x, y) = 0 \\ F_2(x, y) = 0 \end{cases}$$

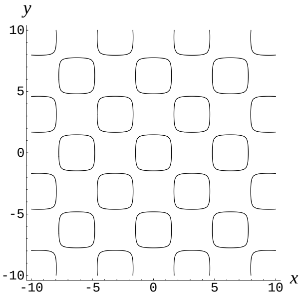
são as intersecções desses dois conjuntos de superfícies:

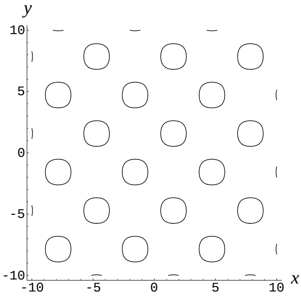
Sistemas de equações não lineares

Cada equação $F_1(x, y) = 0$ e $F_2(x, y) = 0$ determina um conjunto de curvas (superfícies contidas em um espaço de dimensão dois) no plano $x \times y$. As soluções de $F(\mathbf{x}) = \mathbf{0} \in \mathbb{R}^2$, ou dito de outra forma, do sistema

$$\begin{cases} F_1(x, y) = 0 \\ F_2(x, y) = 0 \end{cases}$$

são as intersecções desses dois conjuntos de superfícies:



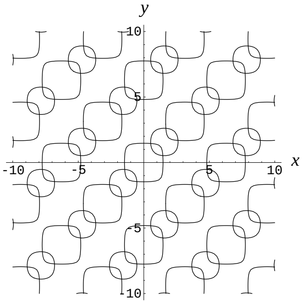


Sistemas de equações não lineares

Cada equação $F_1(x, y) = 0$ e $F_2(x, y) = 0$ determina um conjunto de curvas (superfícies contidas em um espaço de dimensão dois) no plano $x \times y$. As soluções de $F(\mathbf{x}) = \mathbf{0} \in \mathbb{R}^2$, ou dito de outra forma, do sistema

$$\begin{cases} F_1(x, y) = 0 \\ F_2(x, y) = 0 \end{cases}$$

são as intersecções desses dois conjuntos de superfícies:



Sistemas de equações não lineares

Existe um número menor de métodos disponíveis para determinar uma solução aproximada para esse problema.

Vamos estudar apenas a extensão do método Newton-Raphson para sistemas de equações não lineares.

Método Newton-Raphson para sistemas

Se $A(\mathbf{x})$ for uma matriz $n \times n$ não singular (determinante diferente de zero) em alguma vizinhança da solução $\mathbf{x}^*$ da equação , então $\mathbf{x}^*$ também é solução da equação

$$\Phi(\mathbf{x}^*) = \mathbf{x}^*,$$

onde $\Phi(\mathbf{x}) \doteq \mathbf{x} + A(\mathbf{x})F(\mathbf{x})$.

Método Newton-Raphson para sistemas

Se $A(\mathbf{x})$ for uma matriz $n \times n$ não singular (determinante diferente de zero) em alguma vizinhança da solução $\mathbf{x}^*$ da equação , então $\mathbf{x}^*$ também é solução da equação

$$\Phi(\mathbf{x}^*) = \mathbf{x}^*,$$

onde $\Phi(\mathbf{x}) \doteq \mathbf{x} + A(\mathbf{x})F(\mathbf{x})$.
A função Φ permite a construção de uma sequência de aproximações $\left\{\mathbf{x}^{(j)}\right\}_{j=0}^{\infty}$ a partir da regra

$$\mathbf{x}^{(j+1)} = \Phi\left(\mathbf{x}^{(j)}\right)$$

e de uma aproximação inicial $\mathbf{x}^{(0)}$.

Método Newton-Raphson para sistemas

O teorema do ponto fixo garante neste contexto que se $\mathbf{x}^{(0)}$ for suficientemente próximo da solução então existe uma constante $0 \leq K < \infty$ tal que

$$\lim_{j \to \infty} \frac{\left\| \mathbf{x}^{(j+1)} - \mathbf{x}^* \right\|}{\left\| \mathbf{x}^{(j)} - \mathbf{x}^* \right\|} = K,$$

onde a notação $\|\cdot\|$ indica uma norma (medida) para os vetores em dimensão n.

Método Newton-Raphson para sistemas

O teorema do ponto fixo garante neste contexto que se $\mathbf{x}^{(0)}$ for suficientemente próximo da solução então existe uma constante $0 \leq K < \infty$ tal que

$$\lim_{j \to \infty} \frac{\left\| \mathbf{x}^{(j+1)} - \mathbf{x}^* \right\|}{\left\| \mathbf{x}^{(j)} - \mathbf{x}^* \right\|} = K,$$

onde a notação $\|\cdot\|$ indica uma norma (medida) para os vetores em dimensão n. O método de Newton-Raphson consiste em determinar uma função Φ tal que a convergência seja quadrática, i. e., Φ deve ser tal que os elementos da sequência $\left\{\mathbf{x}^{(j)}\right\}_{j=0}^{\infty}$ possuam a seguinte convergência:

$$\lim_{j \to \infty} \frac{\left\| \mathbf{x}^{(j+1)} - \mathbf{x}^* \right\|}{\left\| \mathbf{x}^{(j)} - \mathbf{x}^* \right\|^2} = \tilde{K}$$

para um $0 \leq \tilde{K} < \infty$.

Método Newton-Raphson para sistemas

O teorema do ponto fixo garante neste contexto que se $\mathbf{x}^{(0)}$ for suficientemente próximo da solução então existe uma constante $0 \leq K < \infty$ tal que

$$\lim_{j\to\infty} \frac{\left\|\mathbf{x}^{(j+1)} - \mathbf{x}^*\right\|}{\left\|\mathbf{x}^{(j)} - \mathbf{x}^*\right\|} = K,$$

onde a notação $\|\cdot\|$ indica uma norma (medida) para os vetores em dimensão n.
O método de Newton-Raphson consiste em determinar uma função Φ tal que a convergência seja quadrática, i. e., Φ deve ser tal que os elementos da sequência $\left\{\mathbf{x}^{(j)}\right\}_{j=0}^{\infty}$ possuam a seguinte convergência:

$$\lim_{j\to\infty} \frac{\left\|\mathbf{x}^{(j+1)} - \mathbf{x}^*\right\|}{\left\|\mathbf{x}^{(j)} - \mathbf{x}^*\right\|^2} = \tilde{K}$$

para um $0 \leq \tilde{K} < \infty$.
A função Φ que garante esse comportamento pode ser obtida a partir da série de Taylor para a função F em torno de um dado ponto $x^{(j)}$ (j-ésima aproximação para a solução da equação).

Método Newton-Raphson para sistemas

O desenvolvimento em série de Taylor para F é da forma:

$$F(\mathbf{x}) = F\left(\mathbf{x}^{(j)}\right) + J\left(\mathbf{x}^{(j)}\right)\left(\mathbf{x} - \mathbf{x}^{(j)}\right) + O\left(\left\|\mathbf{x} - \mathbf{x}^{(j)}\right\|^2\right),$$

onde J é a matriz jacobiana associada à função F.

Matriz jacobiana

As componentes i, j da matriz jacobiana são dadas por

$$(J(\mathbf{x}))_{i,j} = \frac{\partial F_i}{\partial x_j}(\mathbf{x})$$

para $F(\mathbf{x}) = (F_1(\mathbf{x}), F_{12}(\mathbf{x}), \ldots, F_n(\mathbf{x}))$, onde cada i-ésima componente é da forma $F_i(\mathbf{x}) = F_i(x_1, x_2, \ldots, x_n)$. As linhas da matriz correspondem aos gradientes das respectivas componentes de F.

Método Newton-Raphson para sistemas

De acordo com a série de Taylor, $\mathbf{x} = \mathbf{x}^*$ temos

$$F\left(\mathbf{x}^{(j)}\right) + J\left(\mathbf{x}^{(j)}\right)\left(\mathbf{x}^* - \mathbf{x}^{(j)}\right) \approx F\left(\mathbf{x}^*\right).$$

O método consiste na construção da sequência de aproximações $\left\{\mathbf{x}^{(j)}\right\}_{j=0}^{\infty}$ a partir de uma aproximação inicial $\mathbf{x}^{(0)}$ e da regra

$$\mathbf{x}^{(j+1)} = \mathbf{x}^{(j)} - J^{-1}(\mathbf{x}^{(j)})F(\mathbf{x}^{(j)})$$

para $j \geq 0$.

Método Newton-Raphson para sistemas

De acordo com a série de Taylor, $\mathbf{x} = \mathbf{x}^*$ temos

$$F\left(\mathbf{x}^{(j)}\right) + J\left(\mathbf{x}^{(j)}\right)\left(\mathbf{x}^* - \mathbf{x}^{(j)}\right) \approx F\left(\mathbf{x}^*\right).$$

O método consiste na construção da sequência de aproximações $\left\{\mathbf{x}^{(j)}\right\}_{j=0}^{\infty}$ a partir de uma aproximação inicial $\mathbf{x}^{(0)}$ e da regra

$$\mathbf{x}^{(j+1)} = \mathbf{x}^{(j)} - J^{-1}(\mathbf{x}^{(j)})F(\mathbf{x}^{(j)})$$

para $j \geq 0$.
Na prática é mais eficiente resolver o sistema linear

$$J\left(\mathbf{x}^{(j)}\right)\delta\mathbf{x}^{(j)} = -F(\mathbf{x}^{(j)})$$

e atualizar a aproximação:

$$\mathbf{x}^{(j+1)} = \mathbf{x}^{(j)} + \delta\mathbf{x}^{(j)}$$

Exemplo

Considere o sistema

$$\begin{cases} x^2 + y^2 - 4x & = & 0 \\ y^2 + 2x - 2 & = & 0 \end{cases}$$

O conjunto de pontos que satisfaz a primeira equação é um círculo de raio 2 e centro $(2, 0)$. No caso da segunda equação, trata-se de um parábola com eixo sobre x e vértice em $(1, 0)$.

Exemplo

Considere o sistema

$$\begin{cases} x^2 + y^2 - 4x &= 0 \\ y^2 + 2x - 2 &= 0 \end{cases}$$

O conjunto de pontos que satisfaz a primeira equação é um círculo de raio 2 e centro $(2,0)$. No caso da segunda equação, trata-se de um parábola com eixo sobre x e vértice em $(1,0)$.

Podemos observar o comportamento das funções no lado esquerdo das equações do sistema através do comando `fplot3d` no Scilab:

```
x=0:0.1:4;y=-3:0.1:3;
deff('y=feq1(x1,x2)','y=x1^2+x2^2-4*x1');
deff('y=feq2(x1,x2)','y=x2^2+2*x1-2');
fplot3d(x,y,feq1);fplot3d(x,y,feq2);xgrid;
```

Exemplo

Já o conjunto de pontos que satisfazem cada equação podem ser desenhados através do comando

```
contour(x,y,feq1,[0 0]);contour(x,y,feq2,[0 0]);xgrid;
```

Através desses comandos é possível verificar que existe uma solução na vizinhança do ponto $(0{,}35; 1{,}23)$. Esse ponto será utilizado como aproximação inicial $x^{(0)}$.