Sistemas não lineares. Método de Newton

MS211 – Cálculo Numérico

Giuseppe Romanazzi

3 Novembro 2020

Conteúdo

Sistemas não lineares, $F(x) = 0$

Suponha de ter n incógnitas que indicamos com $x_1, x_2, \ldots, x_n$ e de ter n equações não lineares entre as variáveis, estaremos no caso de ter o seguinte sistema não linear para resolver:

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

A função f_i descreve a equação não linear i.
Por exemplo se for $n = 2$ e $f_1(x_1, x_2) = x_1 x_2 - x_1 + 1$, $f_2(x_1, x_2) = x_1^2 - 2$ é como se quiséssemos encontrar x_1, x_2 tais que $\begin{cases} x_1 x_2 - x_1 + 1 = 0 \\ x_1^2 - 2 = 0 \end{cases}$ que tem duas soluções $(x_1, x_2) = (\sqrt{2}, 1 - \frac{1}{\sqrt{2}})$ e $(-\sqrt{2}, 1 + \frac{1}{\sqrt{2}})$.

Sistemas não lineares, $F(x) = 0$

Note que o sistema

$$\begin{cases} f_1(x_1, x_2, \ldots, x_n) = 0 \\ f_2(x_1, x_2, \ldots, x_n) = 0 \\ \ldots \\ f_n(x_1, x_2, \ldots, x_n) = 0 \end{cases} \iff F(x) = 0$$

pode ser escrito simplesmente como $F(x) = 0$, onde $x \in \mathbb{R}^n$ é um vetor de dimensão n que tem como elementos os x_i incógnitos: $x = (x_1, x_2, \ldots, x_n)^t$.
A função F transforma os vetores em $\mathbb{R}^n$ em outros vetores em $\mathbb{R}^n$, ousseja $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$,
satisfaz

$$F(x) = F(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), f_2(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n))^t$$

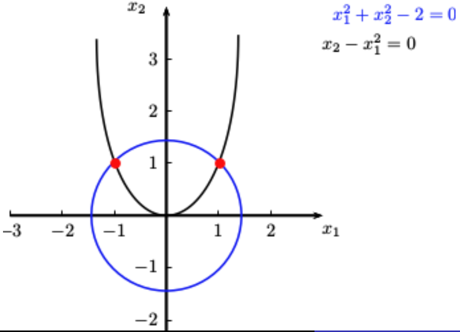
Exemplos de sistemas não lineares, interpretação geométrica

Diferentemente dos sistemas lineares, não é conhecido um critério para determinar a-priori se um sistema não linear de n equações do tipo $f_i(x) = 0$ admite nenhuma, uma ou mais soluções.
Por exemplo um sistema de duas equações não lineares e de duas incógnitas x_1, x_2 pode ter zero, uma, duas, três ,...,k soluções ou até infinitas soluções.
É possível achar quantas soluções o sistema tem, fazendo o gráfico dos pontos (x_1, x_2) que satisfazem as equações $f_i(x_1, x_2) = 0$ com $i = 1, 2$ e ver quantas interseções existem.

Exemplos de sistemas com $n = 2$ equações não lineares

$$\begin{cases} f_1(x_1, x_2) = x_1^2 + x_2^2 - 2 = 0 \\ f_2(x_1, x_2) = x_2 - x_1^2 = 0 \end{cases}$$

Notamos que a primeira equação descreve o conjuntos de pontos que distam $\sqrt{2}$ da origem $(0, 0)$. Ou seja o gráfico de $f_1(x_1, x_2) = 0$ descreve a circunferência de centro $(0, 0)$ e raio $\sqrt{2}$. Em vez a segunda equação descreve os pontos que se encontram na parábola de equação $x_2 = x_1^2$ que tem vértex na origem. Observando os gráficos notamos que temos duas soluções do sistema, porque os gráficos intersecam-se em dois pontos.



Matriz Jacobiana de F

Seja $F : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ a função que descreve o sistema não linear e $x \in D$ um ponto no domínio da F.

Definição Matriz Jacobiana

Chama se matriz jacobiana de F no ponto x o simplesmente Jacobiana de F em x a seguinte matriz de dimensão n

$$J_F(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1}(x) & \frac{\partial f_n}{\partial x_2}(x) & \dots & \frac{\partial f_n}{\partial x_n}(x) \end{pmatrix} = \begin{pmatrix} \nabla f_1^t(x) \\ \nabla f_2^t(x) \\ \vdots \\ \nabla f_n^t(x) \end{pmatrix}$$

Denotado com $\nabla f_i(x) \in \mathbb{R}^n$ o gradiente de f em x ou seja, $\nabla f_i(x) = (\frac{\partial f_i}{\partial x_1}(x) \quad \frac{\partial f_i}{\partial x_2}(x) \quad \dots \quad \frac{\partial f_i}{\partial x_n}(x))^t$. Temos também

$$J_F(x) = (\nabla f_1(x)^t \quad \nabla f_2(x)^t \quad \dots \quad \nabla f_n(x)^t)^t.$$

O Jacobiano estende o conceito de derivada para funções $n-$dimensionais

Exemplo de Jacobiano associado a um sistema não linear

$$\begin{cases} -2x_1^3 + 2x_2 + x_3 = 2 \\ 4x_1 - 2x_2 = -1 \\ -x_3 + x_1 = 0 \end{cases}$$

Observamos que o sistema pode ser escrito assim

$$\begin{cases} f_1(x) = -2x_1^3 + 2x_2 + x_3 - 2 = 0 \\ f_2(x) = 4x_1 - 2x_2 + 1 = 0 \\ f_3(x) = x_1 - x_3 = 0 \end{cases}$$

Portanto o Jacobiano de $F = (f_1 \quad f_2 \quad f_3)^t$ num ponto $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ é

$$J_F(x) = \begin{pmatrix} -6x_1^2 & 2 & 1 \\ 4 & -2 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

Normas de Matrizes e vetores

A **norma de matriz** induzida de uma norma vetorial é
$\|A\| := \sup_{x\in\mathbb{R}^n, x\neq 0} \dfrac{\|Ax\|}{\|x\|}$.
Exemplo de normas de matrizes: sendo que $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$, obtemos

$$\|A\|_\infty = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{i=1,\ldots,n} \sum_{j=1}^{n} |a_{ij}|.$$

Usando que $\|x\|_1 = \sum_{i=1}^{n} |x_i|$, obtemos

$$\|A\|_1 = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{j=1,\ldots,n} \sum_{i=1}^{n} |a_{ij}|.$$

Temos que $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$ é a norma euclideana e a norma associada é $\|A\|_2 = \sqrt{\rho(AA^t)}$ que é a raiz quadrada do máximo autovalor em modulo de AA^t.

Métodos iterativos. Critérios de precisão ou de parada

Queremos construir métodos iterativos que dada uma aproximação inicial $x^{(0)}$ da solução x do sistema $F(x) = 0$ geram uma sequência $\{x^{(k)}\}$ que converge a solução x do sistema, ou seja tal que $\lim_{k\to\infty} x^{(k)} = x$.

Lembre que este significa que por um dado $\epsilon > 0$ existe um $k_0 > 0$ tal que para cada $k > k_0$ $\|x^{(k)} - x\| < \varepsilon$.

Sendo que não se conhece a solução x usaremos três possíveis critérios para parar o processo iterativo com sucesso:

- $\|F(x^{(k)})\| < \varepsilon$ (controle ou critério do resíduo)
- $\|x^{(k+1)} - x^{(k)}\| < \varepsilon$ (controle sobre a distância entre duas iterações consecutivas)
- $\dfrac{\|x^{(k+1)} - x^{(k)}\|}{\|x^{(k+1)}\|} < \varepsilon$ (controle sobre a distância relativa)

Usaremos nesta aula a norma do máximo: $\|x\|_\infty = \max_i |x_i|$

Critérios de insucesso

Há dois possíveis controles/critérios para parar o processo iterativo com um insucesso

- $k > k_{max}$
- $\|F(x^{(k)})\| > F_{max}$

onde k_{max} é o número máximo de iterações aceites do método e F_{max} é o máximo valor de F aceite. Estes são controles feitos para evitar que o ciclo iterativo continue ao infinito no caso que não haja convergência.

Método de Newton

Generalizamos o método de Newton $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$, usado para resolver uma só equação não linear,
e resolvemos agora um sistema de n equações não lineares.

Consideramos a aproximação $x^{(k)}$ da solução do sistema não linear $F(x) = 0$, o método de Newton acha a próxima solução através a formula

$$x^{(k+1)} = x^{(k)} - (J_F(x^{(k)}))^{-1} F(x^{(k)})$$

Interpretação geométrica

Lembramos que o método de Newton no caso escalar de uma equação procurava o zero de f determinando em cada passo k o zero da reta tangente a f no ponto $(x_k, f(x_k))$, então:
x_{k+1} era o zero da reta $r(x) = f(x_k) + f'(x_k)(x - x_k)$. Esta reta $r(x)$ pode ser vista como a linearização de f(x) ao longo do ponto $(x_k, f(x_k))$.

No caso vetorial $L(x) = F(x^{(k)}) + J_F(x^{(k)})(x - x^{(k)})$, é a linearização de F passante por $(x^{(k)}, F(x^{(k)}))$, porque a matriz Jacobiana generaliza no caso n dimensional o conceito de derivada.
O zero da função linear $L(x)$ **é o vetor** y **tal que** $L(y) = 0$.
$L(y) = 0 \iff J_F(x^{(k)})(y - x^{(k)}) = -F(x^{(k)})$

$$\iff y = x^{(k)} - J_F(x^{(k)})^{-1}F(x^{(k)})$$

Por isso $x^{(k+1)} := x^{(k)} - J_F(x^{(k)})^{-1}F(x^{(k)})$ é o zero de $L(x)$.

Determinar $x^{(k+1)}$

Em vez de determinar a inversa de $J_F(x^{(k)})$ podemos obter $x^{(k+1)}$ observando que

$$x^{(k+1)} = x^{(k)} - (J_F(x^{(k)}))^{-1}F(x^{(k)}) \iff J_F(x^{(k)})(x^{(k+1)} - x^{(k)}) = -F(x^{(k)}).$$

Então obtemos $x^{(k+1)}$ em dois passos

1. Resolvemos o sistema $J_F(x^{(k)})s^{(k)} = -F(x^{(k)})$
2. Sendo que $s^{(k)} = x^{(k+1)} - x^{(k)}$, obtemos $x^{(k+1)}$ somando o vetor $s^{(k)}$ com a iteração anterior $x^{(k)}$, ou seja $x^{(k+1)} = x^{(k)} + s^{(k)}$.

O sistema linear $J_F(x^{(k)})s^{(k)} = -F(x^{(k)})$ pode ser resolvido usando métodos diretos ou iterativos. Claramente se for usado um método iterativo iremos achar uma aproximação de $s^{(k)}$ e não o seu valor exato.

Algoritmo

Input: F, $x^{(0)}$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$

Passo 1: Computa o vetor $F(x^{(k)})$ e matriz Jacobiana $J_F(x^{(k)})$

Passo 2: Se $\|F(x^{(k)})\|_\infty < \varepsilon_1$ então dá em output $x^{(k)}$. Sai do ciclo iterativo.

Passo 3: Senão resolve o sistema linear $J_F(x^{(k)})s^{(k)} = -F(x^{(k)})$, determinando $s^{(k)}$

Passo 4: $x^{(k+1)} = x^{(k)} + s^{(k)}$

Passo 5: Se $\|s^{(k)}\| < \varepsilon_2$ então dá em output $x^{(k+1)}$; Sai do ciclo.

Passo 6: $k \leftarrow k + 1$, vai ao passo 1

Note que $s^{(k)} = x^{(k+1)} - x^{(k)}$

Teorema de Convergência do método de Newton

Seja $F : D \subset \mathbb{R}^n \to \mathbb{R}^n$ *a função que descreve o sistema não linear* $F(x) = 0$ *e seja* $\Omega_0 \subset D$ *tal que*

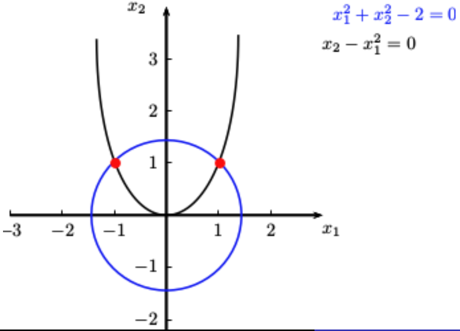
- Ω_0 *contem um único zero* x *de* F*, ou seja contem uma única solução do sistema.*
- F *é diferenciável em* Ω_0*, ou seja existem todas as derivadas de* F *e são continuas em* Ω*.*
- $det(J_F(x)) \neq 0$*, ou analogamente existe* $\beta > 0$ *tal que* $\|J_F(x)^{-1}\| < \beta$ *onde* x *é o zero procurado*
- *A matriz jacobiana* $J_F(y)$ *satisfaz a condição de Lipschitz em* Ω_0*: existe* $\gamma > 0$ *tal que por cada* $y, z \in \Omega_0$ *temos* $\|J_F(y) - J_F(z)\| \leq \gamma \|y - z\|$

então se $x^{(0)} \in \Omega_0$ *temos que cada iteração* $x^{(k)}$ *do método de Newton estará em* Ω_0 *e temos a convergência do método para o zero* x *procurado:* $\lim\limits_{k \to \infty} x^{(k)} = x$*. Mais, temos que a convergência é quadrática e vale que existe* $C > 0$ *tal que* $\|x^{(k+1)} - x\| \leq C \|x^{(k)} - x\|^2$*.*

Exemplo de resolução numérica de um sistema não linear com o método de Newton

$$\begin{cases} x_1^2 + x_2^2 - 2 = 0 \\ x_2 - x_1^2 = 0 \end{cases}$$

Trata se de encontrar as interseções de uma circunferência de centro na origem e raio $\sqrt{2}$ com uma parábola com vertex na origem. São esperadas duas interseções, ou seja duas soluções do sistema



$$\begin{cases} x_1^2 + x_2^2 - 2 = 0 \\ x_2 - x_1^2 = 0 \end{cases}$$

Note que $F(x) = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 - 2 \\ x_2 - x_1^2 \end{pmatrix}$.

Procuramos a solução do sistema que satisfaz um dos dois critérios de aproximação $\|F(x^{(k)})\| < \varepsilon_1$ ou $\|x^{(k)} - x^{(k-1)}\| < \varepsilon_2$, com $\varepsilon_1 = 10^{-1}$, $\varepsilon_2 = 10^{-1}$. Então pararemos o método quando um dos dois critério for satisfeito.

Decidimos de encontrar o zero de $F(x)$ que tem $x_1 > 0$.

A matriz jacobiana associada ao sistema é $J_F(x) = \begin{pmatrix} 2x_1 & 2x_2 \\ -2x_1 & 1 \end{pmatrix}$.

Notamos que o seu determinante é $2x_1 + 4x_1x_2 = 2x_1(1 + 2x_2)$ que resulta ser nulo somente se $x_1 = 0$ ou se $x_2 = -\frac{1}{2}$.

Então a região Ω_0 onde pegar o $x^{(0)}$ tem de estar distante do eixo das x_2 (de eq. $x_1 = 0$) e da reta $x_2 = -\frac{1}{2}$. Observando que o zero procurado tem um $x_2 > 0$ e um $x_1 > 0$,

então podemos escolher um $\Omega_0 :=]\delta, +\infty[\times [0, +\infty[$, com $\delta > 0$ ou seja Ω_0 é conteúdo no primeiro quadrante.

Analisamos se a condição de Lipschitz (existe $\gamma > 0$ tal que $\|J_F(y) - J_F(z)\| \leq \gamma\|y - z\| \quad \forall y, z \in \Omega_0$) for satisfeita em Ω_0.
Senão temos de refinar Ω_0.
Usando a norma do máximo

$$\|J_F(y) - J_F(z)\|_\infty = \left\| \begin{pmatrix} 2(y_1 - z_1) & 2(y_2 - z_2) \\ -2(y_1 - z_1) & 0 \end{pmatrix} \right\|_\infty =$$

$$= \max\{|2(y_1 - z_1)| + |2(y_2 - z_2)|, |2(y_1 - z_1)|\} = 2|y_1 - z_1| + 2|y_2 - z_2|.$$

Agora sendo que $\|y - z\|_\infty = \max\limits_{i=1,2} |y_i - z_i|$ e
$\|J_F(y) - J_F(z)\|_\infty = 2|y_1 - z_1| + 2|y_2 - z_2| \leq$
$2 \max\limits_{i=1,2} |y_i - z_i| + 2 \max\limits_{i=1,2} |y_i - z_i| \leq 4\|y - z\|_\infty$,
então vale a condição de Lipschitz em todo $\mathbb{R}^2$ com $\gamma = 4$. Então sendo também que F é diferenciável e o zero procurado está em Ω_0 podemos pegar um qualquer $x^{(0)}$ com $x_1^{(0)} > 0$ e $x_2^{(0)} > 0$ para ter a convergência do método de Newton ao zero (com $x_1 > 0$) de F.

Escolhemos a aproximação inicial $x^{(0)} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}$.

Observamos que com $x^{(0)} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}$, obtemos

$F(x^{(0)}) = \begin{pmatrix} \sqrt{2}^2 + 0 - 2 \\ 0 - \sqrt{2}^2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$ e portanto

$\|F(x^{(0)})\|_\infty = 2 > \epsilon_1$.

Determinamos então $x^{(1)}$. Resolvendo: o sistema $J_F(x^{(0)})s = -F(x^{(0)})$ e depois obtemos $x^{(1)}$ como $x^{(1)} = x^{(0)} + s$.

O sistema $J_F(x^{(0)})s = -F(x^{(0)})$ é

$$\begin{pmatrix} 2\sqrt{2} & 0 \\ -2\sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

A solução do sistema é $s_1 = 0$, $s_2 = 2$ Portanto

$$x^{(1)} = x^{(0)} + s = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 2 \end{pmatrix}.$$

Observamos que $\|x^{(1)} - x^{(0)}\|_\infty = \|s\|_\infty = 2 > \varepsilon_2$ e

$\|F(x^{(1)})\|_\infty = \left\| \begin{pmatrix} \sqrt{2}^2 + 2^2 - 2 \\ 2 - \sqrt{2}^2 \end{pmatrix} \right\|_\infty = \left\| \begin{pmatrix} 4 \\ 0 \end{pmatrix} \right\|_\infty = 4 > \varepsilon_1.$

Precisamo então computar $x^{(2)} = x^{(1)} + s$ onde $s = (s_1 \quad s_2)^t$ é o vetor solução do sistema $J_F(x^{(1)})s = -F(x^{(1)})$ que é

$$\begin{pmatrix} 2\sqrt{2} & 4 \\ -2\sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = - \begin{pmatrix} 4 \\ 0 \end{pmatrix}$$

A solução é $s = \begin{pmatrix} -\frac{\sqrt{2}}{5} \\ -0.8 \end{pmatrix}$ e portanto

$$x^{(2)} = x^{(1)} + s = \begin{pmatrix} \sqrt{2} \\ 2 \end{pmatrix} + \begin{pmatrix} -\frac{\sqrt{2}}{5} \\ -0.8 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{24}}{5} \approx 1.1314 \\ 1.2 \end{pmatrix}.$$

Observamos que

$\|x^{(2)} - x^{(1)}\|_\infty = \max\{|s_1|, |s_2|\} = \max\{0.2828, 0.8\} = 0.8 > \varepsilon_2$ e

$\|F(x^{(2)})\| = \left\| \begin{pmatrix} 0.72 \\ -0.080 \end{pmatrix} \right\|_\infty = 0.72 > \varepsilon_1.$

Prosseguindo obteremos $x^{(3)} = \begin{pmatrix} 1.0128 \\ 1.0118 \end{pmatrix}$, que tem

$F(x^{(3)}) = \begin{pmatrix} 0.0495 \\ -0.0141 \end{pmatrix}$, então $\|F(x^{(3)})\|_\infty = 0.0495 < \varepsilon_1$. Portanto $x^{(3)}$ é a solução achada do método de Newton associada ao $\epsilon_1 = 0.1$.