

Introdução ao Cálculo Numérico

Método de Newton para resolução de
sistemas de equações não-lineares

- Considere um sistema de n funções não-lineares f_1, \dots, f_n a n variáveis,

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

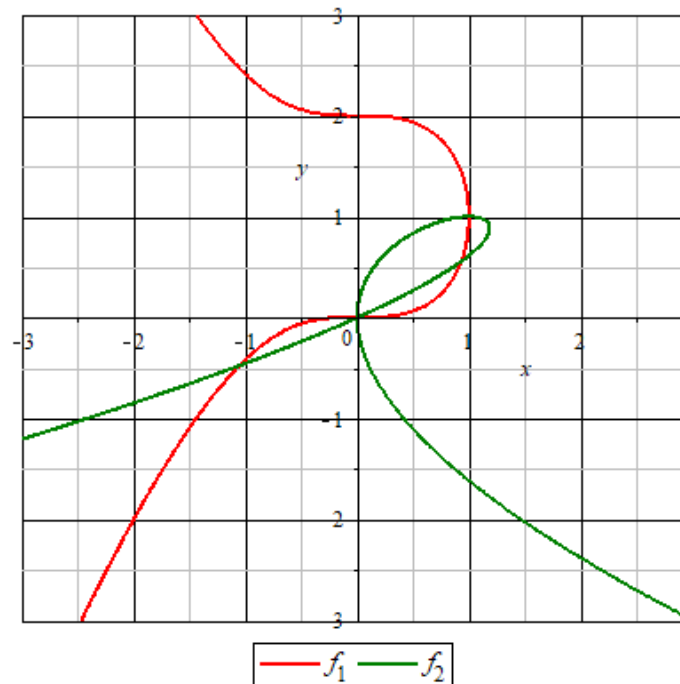
ou, em notação vetorial, $\mathbf{F}(\mathbf{X}) = \mathbf{0}$, onde $\mathbf{X} = (x_1, \dots, x_n)^T$ e $\mathbf{F}: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Método de Newton para sistemas de equações não-lineares

- Por exemplo, para o sistema

$$\begin{cases} x^3 - 2y + y^2 = 0 \\ x^2 - 2xy + y^3 = 0 \end{cases}$$

as curvas correspondentes às expressões em x e y são:



Método de Newton para sistemas de equações não-lineares

- Esse sistema pode ser reescrito na forma vetorial associando as variáveis x e y aos elementos x_1 e x_2 :

$$\mathbf{F}(\mathbf{X}) = \begin{bmatrix} x_1^3 - 2x_2 + x_2^2 \\ x_1^2 - 2x_1x_2 + x_2^3 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Método de Newton para sistemas de equações não-lineares

- Para funções vetoriais, uma aproximação linear para $F(\mathbf{X})$ num ponto \mathbf{X}_0 é dada por uma expansão de Taylor até os termos de 1ª ordem, resultando em

$$F(\mathbf{X}) \cong F(\mathbf{X}_0) + J(\mathbf{X}_0)(\mathbf{X} - \mathbf{X}_0)$$

onde

$$F(\mathbf{X}) = \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{bmatrix}, J(\mathbf{X}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

e $J(\mathbf{X})$ é a matriz jacobiana da F em \mathbf{X} .

- Como desejamos obter as componentes do vetor \mathbf{X} tais que $\mathbf{F}(\mathbf{X}) = \mathbf{0}$, podemos escrever

$$\mathbf{F}(\mathbf{X}_0) + \mathbf{J}(\mathbf{X}_0)\Delta\mathbf{X} = \mathbf{0}$$

onde $\Delta\mathbf{X} = \mathbf{X} - \mathbf{X}_0$.

- Como \mathbf{X} é desconhecido, podemos utilizar a equação acima para, dado um valor inicial \mathbf{X}_0 , determinar qual o valor de $\Delta\mathbf{X}$.

- Para tal, escrevemos

$$J(\mathbf{X}_0)\Delta\mathbf{X} = -\mathbf{F}(\mathbf{X}_0)$$

e observamos que a equação acima é um sistema de equações lineares de n equações a n variáveis, o qual terá solução desde que $J(\mathbf{X}_0)$ admita inversa.

- Supondo que tal solução exista, podemos escrever

$$\mathbf{X} = \mathbf{X}_0 + \Delta\mathbf{X}$$

e caso \mathbf{X} não satisfaça $\mathbf{F}(\mathbf{X}) = \mathbf{0}$, pode-se calcular novamente o sistema anterior, fazendo $\mathbf{X}_0 = \mathbf{X}$.

- Essa ideia leva a estabelecermos o seguinte procedimento iterativo para cálculo da solução de $F(X) = 0$, dada uma estimativa inicial X_0 :

1. A cada iteração $k = 0, 1, \dots$:
2. Calcule a correção ΔX como solução do sistema linear de equações
$$J(X_k)\Delta X = -F(X_k).$$
2. Calcule $X_{k+1} = X_k + \Delta X$.
3. Se $\|F(X_{k+1})\| \cong 0$, então interrompe as iterações.

o qual é o chamado **método de Newton**.

Método de Newton para sistemas de equações não-lineares

- Assim como no caso $f(x) = 0$, aqui também o método de Newton apresenta convergência quadrática quando o vetor \mathbf{X}_k está próximo da solução \mathbf{X} ; da mesma forma, o método pode divergir facilmente caso encontre singularidades ou quase singularidades na matriz $\mathbf{J}(\mathbf{X}_k)$ (o que equivale a $f'(x_k) \cong 0$ naquele caso).

- Observe que calcular a matriz $J(\mathbf{X})$ de forma **explícita**, usando as expressões para as derivadas parciais $\frac{\partial f_i}{\partial x_j}$, não é recomendado, particularmente se n é grande.
- A alternativa é aproximar numericamente $J(\mathbf{X})$, através de diferenças finitas progressivas, conforme descrito em C.T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, 1995:

Método de Newton para sistemas de equações não-lineares

- Seguindo a formulação proposta por aquele autor, calculamos as colunas da matriz $J(\mathbf{X})$ usando as equações abaixo:

$$[J(\mathbf{X})]_j = \begin{cases} \frac{\mathbf{F}(\mathbf{X} + h\|\mathbf{X}\|\mathbf{e}_j) - \mathbf{F}(\mathbf{X})}{h\|\mathbf{X}\|}, & \mathbf{X} \neq \mathbf{0} \\ \frac{\mathbf{F}(h\mathbf{e}_j) - \mathbf{F}(\mathbf{X})}{h}, & \mathbf{X} = \mathbf{0} \end{cases}$$

para $j = 1, \dots, n$, onde $h = \sqrt{\varepsilon_M}$ e \mathbf{e}_j é o j -ésimo vetor canônico de ordem n .

- A norma $\|\mathbf{X}\|$ utilizada é normalmente a norma l_2 ,

$$\|\mathbf{X}\|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$$

- É possível que a matriz $J(\mathbf{X}_k)$ seja singular ou quase singular; nesse caso, a alternativa é perturbar por um valor pequeno algumas das componentes da estimativa \mathbf{X}_k e recalcular $J(\mathbf{X}_k)$.

- A solução do sistema linear de equações

$$J(\mathbf{X}_k)\Delta\mathbf{X} = -\mathbf{F}(\mathbf{X}_k)$$

pode ser feita tanto através de uma fatoração LU como de forma iterativa; como a matriz $J(\mathbf{X}_k)$ não é constante ao longo das iterações e pode ser **esparsa** e **não-simétrica**, em geral se utiliza um método iterativo não-estacionário como o GMRES.

Método de Newton para sistemas de equações não-lineares

- Ainda segundo Kelley, é recomendado que o critério de parada do método de Newton quanto ao valor da função \mathbf{F} avaliada em \mathbf{X}_k seja calculado em termos de duas tolerâncias, uma relativa ($\tau_r \ll 1$) e outra absoluta ($\tau_a \ll 1$), tais que $\tau_a > \tau_r$.
- As iterações deverão proceder até que a condição

$$\|\mathbf{F}(\mathbf{X}_k)\| < \tau_r \|\mathbf{F}(\mathbf{X}_0)\| + \tau_a$$

seja satisfeita.

- O método de Newton pode ser expresso através do seguinte algoritmo, incorporando essas ideias:

```
proc newton(entrada: F,X,tau_r,tau_a,kmax; saída: X)
1. Fx:=F(X)
2. F_0:=||Fx||
3. para k:=1 até kmax
4.   J:=matriz_jacobiana(F,X)
5.   Resolve o sistema  $J^*(\text{delta\_X})=-Fx$ , obtendo delta_X
6.   X:=X+delta_X
7.   Fx:=F(X)
8.   se ||Fx||<tau_r*F_0+tau_a então
9.     para
10.    fim se
11.fim para
```

Método de Newton para sistemas de equações não-lineares

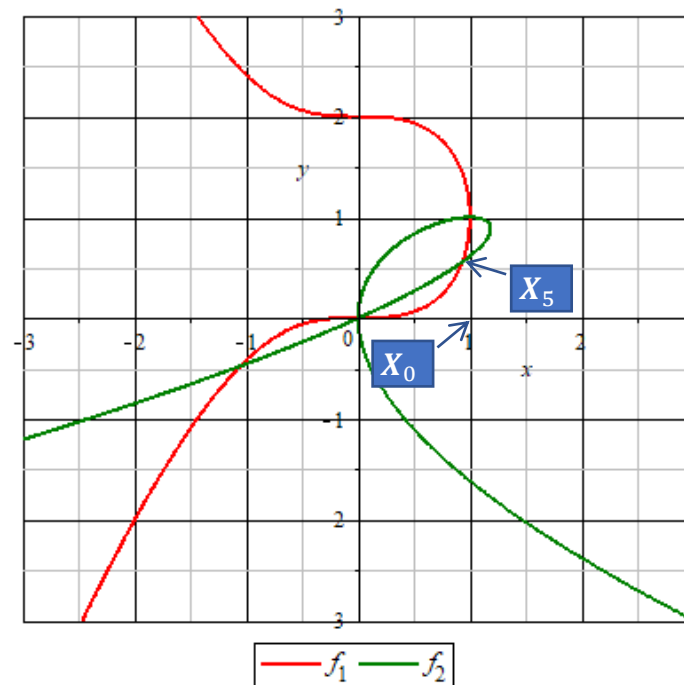
- Retomando o exemplo inicial, vemos que ao se resolver o sistema

$$\mathbf{F}(\mathbf{X}) = \begin{bmatrix} x_1^3 - 2x_2 + x_2^2 \\ x_1^2 - 2x_1x_2 + x_2^3 \end{bmatrix} = \mathbf{0}$$

pelo método de Newton, usando como estimativa inicial o vetor $\mathbf{X}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, sujeito às tolerâncias $\tau_r = 10^{-12}$ e $\tau_a = 10^{-11}$, obtemos os resultados apresentados a seguir:

k	X_k	$\ F(X_k)\ _F$
1	$[1 \quad 0,5]^T$	0,279508
2	$[0,9318182 \quad 0,5454545]^T$	0,0210588
3	$[0,9304621 \quad 0,5588339]^T$	0,000380908
4	$[0,9305332 \quad 0,5592519]^T$	$3,04361 \times 10^{-7}$
5	$[0,9305332 \quad 0,5592522]^T$	$1,31247 \times 10^{-13}$

Observe o comportamento quadrático de convergência do método de Newton!



Método de Newton para sistemas de equações não-lineares

- Dependendo das características de cada equação presente no sistema não-linear, é possível que corrigir \mathbf{X}_k por $\Delta\mathbf{X}$, i.e.

$$\mathbf{X}_{k+1} = \mathbf{X}_k + \Delta\mathbf{X}$$

produza um \mathbf{X}_{k+1} para o qual

$$\|\mathbf{F}(\mathbf{X}_{k+1})\| > \|\mathbf{F}(\mathbf{X}_k)\|.$$

- Existem métodos, chamado de **quase-Newton**, no qual \mathbf{X}_{k+1} é calculado como

$$\mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_k \Delta\mathbf{X},$$

onde α_k é calculado de tal forma a garantir que $\|\mathbf{F}(\mathbf{X}_{k+1})\| < \|\mathbf{F}(\mathbf{X}_k)\|$.

- Em casos mais simples, é possível fazer α_k constante ao longo de todas as iterações.
- Porém, cf. Kelley, existe um critério (chamado de **critério de Armijo**) para se determinar adequadamente α_k , empregando-se uma interpolação parabólica entre três pontos α_a , α_b , α_c da função auxiliar

$$f(\alpha) = \|F(X_k + \alpha\Delta X)\|.$$

- O ponto de mínimo α_k dessa parábola é o valor que minimiza $f(\alpha)$ e, portanto, localmente teremos $\|F(X_{k+1})\| < \|F(X_k)\|$.