

O Método de Newton

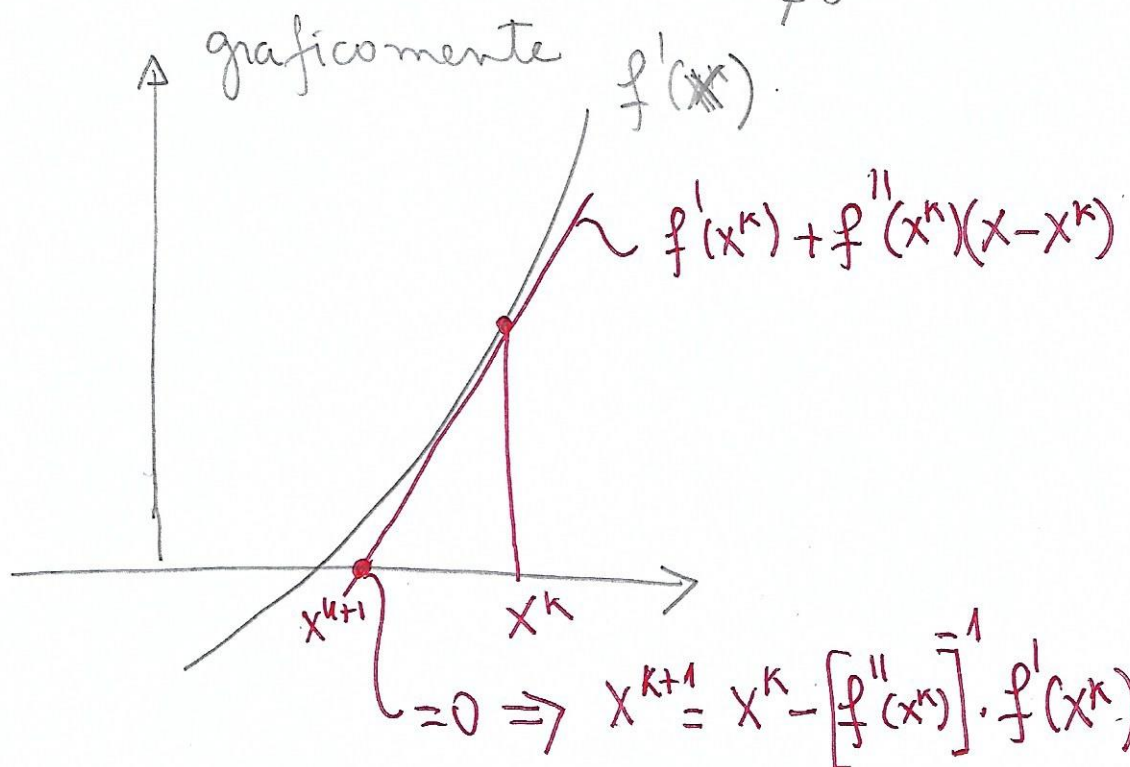
1

Procuramos $\nabla f(x^*) = 0$

Em $\mathbb{R}$ $f'(x^*) = 0 \Rightarrow$

$$f'(x^{k+1}) = f'(x^k) + f''(x^k)(x^{k+1} - x^k) \stackrel{\text{Proposta}}{=} 0$$

$$\Rightarrow \Rightarrow x^{k+1} - x^k = - \frac{f'(x^k)}{f''(x^k)} \quad \text{onde } f''(x^k) \neq 0$$



Em $\mathbb{R}^n$, $f: \mathbb{R}^n \rightarrow \mathbb{R}$

$$\nabla f(x^{k+1}) = \nabla f(x^k) + H(x^k)(x^{k+1} - x^k) \stackrel{\downarrow}{=} 0$$

$$\Rightarrow x^{k+1} = x^k - \underbrace{H(x^k)^{-1} \nabla f(x^k)}_{\text{Passo de Newton}}$$

$\rightarrow$ Se

$H(x_k)$
é não singular

$$d_k = -H(x_k)^{-1} \nabla f(x^k)$$

direção de Newton

① Métodos de Newton (Puro) !!

②

Dado $x_0 \in \mathbb{R}^n$

(*) se $\nabla f(x^k) = 0$ parar se não

$$x^{k+1} = x^k - H(x^k)^{-1} \nabla f(x^k)$$

$k+1 \leftarrow k$ ir a (*)

Problemas

1) se $H(x^k)^{-1}$ não existe está mal definido

2) d_k pode não ser de descida só é de descida se

$H(x^k)$ é definida positiva

$$d_k = \nabla f(x^k)^T, d_k =$$

$$= \nabla f(x^k)^T (-H(x^k) \nabla f(x^k))$$

$$= -\nabla f(x^k)^T H(x^k) \nabla f(x^k) < 0$$

Vamos tentar garantir que seja de descida

Vamos garantir que a direção seja de descida

3

$$\nabla^2 f(x^k) + \mu I$$

sejam λ_i autovalores de $\nabla^2 f(x^u) = H(x^u)$

os autovalores são todos positivos
em cada passo podemos achar μ tal

que

$$-\nabla f(x^k)^T (\nabla^2 f(x^k) + \mu I) \nabla f(x^k) < 0$$

$$\Rightarrow \mu > - \frac{\nabla f(x^k)^T \nabla^2 f(x^k) \nabla f(x^k)}{\|\nabla f(x^k)\|^2}$$

Para garantir a condição do ângulo

$$-\nabla f(x^k)^T (\nabla^2 f(x^k) + \mu I) \nabla f(x^k) < -\delta$$

$$\Rightarrow \mu > - \frac{\nabla f(x^k)^T \nabla^2 f(x^k) \nabla f(x^k)}{\|\nabla f(x^k)\|^2} + \delta$$

Achar
veja que é uma

$$-(\nabla^2 f(x^k) + \mu I) \nabla f(x^k)$$

$$= - \underbrace{\nabla^2 f(x^k) \nabla f(x^k)}_{\text{Newton}} - \underbrace{\mu \nabla f(x^k)^T \nabla f(x^k)}_{\text{direção do gradiente}}$$

Newton

direção
do gradiente

(4)

E uma combinação das duas
quanto maior é o μ mas parece a
direção do Método do gradiente

Propomos o método de Newton
com busca linear

dado $x_0 \in \mathbb{R}^n$, se

se $\nabla f(x^n) \neq 0$ para determinar x^{n+1}

Determinar d_k tal que

$$\nabla^2 f(x^n) d_k = -\nabla f(x^n)$$

Resolver o sistema

(Pode não estar bem definido se $\nabla^2 f(x^n)$ é
singular, Podemos aplicar a técnica para
garantizar a boa definição e a descida)

Passo 2: fazer $x^{n+1} = x^n + \lambda_k d_k$

Usando a busca de armijo
Como no método de direção de descida

Newton para quadrática

5

se f é uma quadrática com Hessiano definido positivo

Newton termina em um passo!

$$f(x) = \frac{1}{2} x^T A x + b^T x + c$$

seja x^0 inicial

$$x^1 = x^0 - \bar{A}^{-1} \nabla f(x^0)$$

$$\nabla f(x^1) = \nabla f(x^0) + A(x^1 - x^0) =$$

$$\nabla f(x^0) - A \bar{A}^{-1} \nabla f(x^0) =$$

$$= 0 \checkmark$$

$x^1 = x^*$ solução do problema.

A direção de Newton pode não ~~ser de descida~~

ser de descida

seja $f(x_1, x_2) = \frac{1}{2}(x^2 - y^2)$ ponto $x^0 = (0, 1)$

$$\nabla f(x^0) = \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}; \quad \nabla^2 f(x^0) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\Rightarrow d_0 = - \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$\nabla f(x^0)^T \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 1 > 0$$

entao toma mos

(6)

- do

OK, mas pode não ter solução

Exemplo:

$$f(x, y) = x^4 + xy + (1+y)^2 \quad X = (0, 0)$$

$$\nabla f(x, y) = \begin{pmatrix} 4x^3 + y \\ x + 2(1+y) \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$\nabla^2 f(x, y) = \begin{pmatrix} 12x & 1 \\ 1 & 2 \end{pmatrix}$$

$$\nabla^2 f(0, 0) = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$$

$$d_N = - \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\Rightarrow \nabla f(0, 0)^T d_N = 0 \quad \checkmark$$

Resultado de convergência local

(7)

Teorema: Seja $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^3$ seja x^* um minimizador local de f em $\mathbb{R}^m$ tal que $\nabla^2 f(x^*)$ é definida positiva, então existe $\rho > 0$ tal que se $x^0 \in B(x^*, \rho)$ e x^k a sequência gerada pelo algoritmo

Newton puro verifica

(i) $\nabla^2 f(x^k)$ é def (+) para todo k

(ii) $\lim_{k \rightarrow \infty} x^k = x^*$

(iii) existe $c > 0$ tal que

$$\|x^{k+1} - x^*\| < c \|x^k - x^*\|^2 \quad \forall k \in \mathbb{N}$$

Prova: Existe $\rho > 0$, $\beta_1 > 0$, $\beta_2 > 0$ tal que para todo x tal que $\|x - x^*\| < \rho$ se verifica que

$$\|H(x)\| < \beta_1$$

$$\text{e} \quad \|\nabla f(x^*)^T - \nabla f(x^T) - H(x)(x^* - x)\| \leq \beta_2 \|x - x^*\|^2$$

seja x^k tal que $\beta_1 \beta_2 \|x_k - x^*\| < 1$ e $\|x^k - x^*\| < \rho$ então

$$\begin{aligned}
 \|x_{k+1} - x^*\| &= \|x^k - x^* - H(x^k)^{-1} \nabla f(x^k)\| \quad (8) \\
 &= \|[H(x^k)]^{-1} [H(x^k)(x^k - x^*) - \nabla f(x^k)]\| = \\
 &= \|[H(x^k)]^{-1} [H(x^k)(x^* - x^k) + \nabla f(x^k) - \nabla f(x^*)]\| \\
 &\leq \|H(x^k)^{-1}\| \cdot \beta_2 \|x^k - x^*\|^2 \leq \theta < 1 \\
 &\leq \beta_1 \cdot \beta_2 \|x^k - x^*\|^2 < \begin{matrix} \uparrow \\ \|x^k - x^*\| \\ \beta_1 \beta_2 \|x^k - x^*\| < 1 \end{matrix}
 \end{aligned}$$

$\Rightarrow$ com $C = \beta_1 \beta_2$ temos (iii)

(i) está garantizado com o ρ

(iii) é

$$\|x_{k+1} - x^*\| < \|x^k - x^*\|$$

$$\Rightarrow x_k \rightarrow x^*$$

Comentários ver que localmente a convergência é quadrática

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} \leq C$$