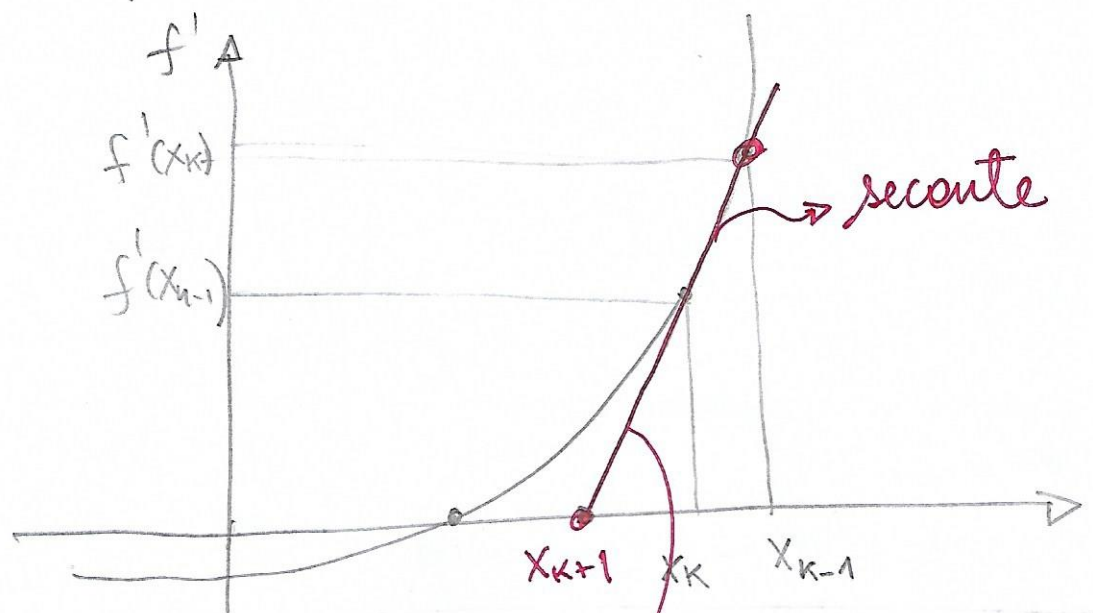


Metodos quase-Newton

1

Em $\mathbb{R}^1$ método da secante para
Achar $f'(x) = 0$



$$\Rightarrow \underbrace{\tilde{f}'(x) - f'(x_k)}_{\text{Aproximação}} = \frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}} \cdot (x - x_k)$$

$$\Rightarrow \tilde{f}'(x_{k+1}) \approx 0$$

$$x_{k+1} = x_k - \left[\frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}} \right]^{-1} f'(x_k)$$

↳ Aproximação de $f''(x_k)$
coef angular da
secante

Método em $\mathbb{R}^m$

(2)

$$x_{k+1} = x_k - \boxed{B_k} (\nabla f(x_k))$$

Aproximação da inversa
de $H(x^k)$

Com a condição Secante

$$\boxed{H_k [\nabla f(x_k) - \nabla f(x_{k-1})] = x_k - x_{k-1}}$$

ou H_k

Algoritmo para encontrar o mínimo

x_0 arbitrário, Seja B_0 = uma matriz simétrica def +

Se $\nabla f(x^k) \neq 0$, os passos para obter x^{k+1} são

Passo 1: calcular $d_k = -B_k \nabla f(x^k)$

Passo 2: Determinar λ_k através de uma busca linear e definir $x^{k+1} = x^k + \lambda_k d_k$

Passo 3: Definir $p_k = x^{k+1} - x^k$, $q_k = \nabla f(x^{k+1}) - \nabla f(x^k)$
e calcular B_{k+1} tal que

$$\boxed{H_{k+1} q_k = p_k}$$

condição secante

A questão é como construir H_{k+1} de maneira que cumpra a condição secante e seja econômica computacionalmente (4)

Família de Broyden (unegão de posto 1)

$$H_{k+1} = H_k + \underbrace{\mu_k \mu_k^T}_{\text{matriz de posto 1}}$$

condição secante

$$H_{k+1} q_k = H_k q_k + \mu_k \boxed{\mu_k^T q_k} = p_k$$

↗ número

se $\mu_k^T p_k \neq 0$ então a direção de μ_k é

$$\mu_k = \alpha (p_k - H_k q_k)$$

vamos determinar α

$$q_k^T p_k = q_k^T H_{k+1} q_k = q_k^T H_k q_k + \alpha^2 [q_k^T (p_k - H_k q_k)]^2$$

$$\Rightarrow \alpha^2 = \frac{q_k^T (p_k - H_k q_k)}{[q_k^T (p_k - H_k q_k)]^2} \Rightarrow$$

desde que $q_k^T (p_k - H_k q_k) \neq 0 \Rightarrow$

$$H_{k+1} = H_k + \alpha^2 (p_k - H_k q_k) (p_k - H_k q_k)^T$$

$$H_{k+1} = H_k + \frac{(p_k - H_k q_k) (p_k - H_k q_k)^T}{(q_k^T (p_k - H_k q_k))}$$

verjam que se $q_k^T (p_k - H_k q_k) > 0$

entao se H_k e' def (+) entao H_{k+1} e'
def positiva (que e garantia de direcao
de descida

Mas queremos uma conexao de ordem 2

Método de Davidson - Fletcher - Powell

$$H_{k+1} = H_k + u_k u_k^T - N_k N_k^T$$

$$p_k = H_{k+1} q_k = H_k q_k + u_k u_k^T q_k + N_k N_k^T q_k$$

$$\underbrace{p_k}_{u_k} - \underbrace{H_k q_k}_{N_k} = \underbrace{(u_k^T q_k) u_k}_{u_k} - \underbrace{(N_k^T q_k) N_k}_{N_k}$$

se $u_k = p_k / (p_k^T q_k)^{1/2}$

entao $N_k = \frac{H_k q_k}{(q_k^T H_k q_k)^{1/2}}$

se verifica a condicao e

$$H_{k+1} = H_k + \frac{p_k p_k^T}{p_k^T q_k} - \frac{H_k q_k q_k^T H_k}{q_k^T H_k q_k}$$

o Método:

(6)

dado x_0 arbitrário e H_0 def (+), $K=0$

Passo 1 Se $\nabla f(x_K) = 0$ parar

Passo 2 fazer $d_K = -H_K g_K$

Passo 3 Minimizar $f(x_K + \alpha d_K)$ para obter x_{K+1} isto é DPF (busca exata)
ou Ache x_{K+1} com armijo
isto é uma opção

Passo 4: $q_K = \nabla f(x_{K+1}) - \nabla f(x_K)$ e $p_K = x_{K+1} - x_K$

$$H_{K+1} = H_K + \frac{p_K p_K^T}{p_K^T q_K} - \frac{H_K q_K q_K^T H_K}{q_K^T H_K q_K}$$

obviamente necessitamos que $p_K^T q_K \neq 0$

Primer Resultado (com a busca exata)

Se H_K é def + então H_{K+1} também é def +

$$d^T H_{K+1} d = d^T H_K d + \frac{(d^T p_K)^2}{p_K^T q_K} - \frac{(d^T H_K q_K)^2}{q_K^T H_K q_K}$$

(7)

dados que H_k é def (+) podemos definir

$$W = H_k^{1/2} d \quad \text{e} \quad z = H_k^{1/2} q_k$$

com isto temos que

$$d^T H_{k+1} d = \boxed{\overset{(1)}{W^T W - \frac{(W^T z)^2}{z^T z}}} + \boxed{\overset{(2)}{\frac{(d^T p_k)^2}{p_k^T q_k}}}$$

dados que a busca é exata

$$\nabla f(x_{k+1}) \perp (x_{k+1} - x_k)$$

$$\Rightarrow \nabla f(x_{k+1})^T (p_k) = 0$$

$$\begin{aligned} \Rightarrow p_k^T q_k &= p_k^T \nabla f(x_{k+1}) - p_k^T \nabla f(x_k) = \\ &= - p_k^T \nabla f(x_k) \end{aligned}$$

$$\Rightarrow p_k^T q_k = \alpha_k \nabla f(x_k)^T H_k \nabla f(x_k) > 0$$

com isto garantizamos $p_k^T q_k \neq 0$

Analizando

$$W^T W - \frac{(W^T z)^2}{(z^T z)^2} = \frac{(W^T W)(z^T z) - (W^T z)^2}{(z^T z)^2} \geq 0$$

veja-se que não pode ser zero (1) e (2) simultaneamente

(8)

① é zero se w e z são colineares

$$w = \beta z$$

$$\Rightarrow H_K^{1/2} d = \beta H_K^{1/2} q_K$$

$$\Rightarrow d = \beta q_K$$

$$\Rightarrow d^T p_K = \beta q_K^T p_K = \beta d_K^T \nabla f^T(x_K) H_K \nabla f(x_K) \neq 0$$

$\Rightarrow$ que não podem ser zero simultaneamente

Terminação finita para quadráticas

Se f é uma função quadrática com Hessiana definida positiva DFP termina em n passo no máximo.

Teorema: Se f é quadrática com Hessiano G def +, então o Método DFP

$$p_i^T G p_j = 0 \quad 1 \leq i < j \leq K \quad (1)$$

$$H_{K+1} G p_i = p_i \quad 0 \leq i \leq K \quad (2)$$

Prova: observamos que

$$q_K = \nabla f(x_{K+1}) - \nabla f(x_K) = G(x_{K+1} - x_K) = G p_K$$

além disso

e

(9)

$$H_{k+1} G p_k = H_{k+1} q_k = p_k \quad (3)$$

vamos $k=0$

(1) é verdade (hipótese vazia)

$$(2) H_1 G p_0 = p_0 \text{ por (3)}$$

vamos a supor que para $l-1$ é verdade ou seja

$$HI \begin{cases} p_i^T G p_j = 0 & 0 \leq i \leq j \leq l-1 \quad (4) \\ H_l G p_i = p_i & 0 \leq i \leq l-1 \quad (5) \end{cases}$$

vamos a provar a Tese indutiva para l

$$\nabla f(x_l) = \nabla f(x_{l+1}) + G(p_l - p_{l+1}) \quad 0 \leq i \leq l-1$$

$$p_i^T \nabla f(x_l) = p_i^T \nabla f(x_{l+1}) + \underbrace{p_i^T G(p_l - p_{l+1})}_{=0}$$

$$\text{mas } p_{l-1}^T \nabla f(x_l) = 0 \Rightarrow$$

$$\Rightarrow \boxed{p_i^T \nabla f(x_l) = 0} \quad 0 \leq i \leq l-1$$

Agora de 5

$$H_l G p_i = p_i \Rightarrow$$

$$p_i^T \nabla f(x_l) = p_i^T G H_l \nabla f(x_l) = p_i(x_l) G p_l = 0$$

 $\Rightarrow$ vale para $i < l$ $\checkmark$ está provado (2)

use mos

$$q_k^T H_k G p_i = q_k^T p_i = p_k^T G p_i = 0$$

(10)

$$\Rightarrow H_{k+1} G p_i = H_k G p_i = p_i \quad 0 \leq i < k$$

prova (3) para k

Exercício 2.14: Sejam $f(x) = \frac{1}{2} x^T G x - b^T x$ (quadrática)

A simétrica def (+) sejam

$x^0, x^1, \dots, x^m \in \mathbb{R}^n$ e definimos $\delta^j = x^j - x^0$

$r^j = \nabla f(x^j) - \nabla f(x^0)$, $j=0, 1, \dots, m$

se δ^j são LI então

$$\tilde{x} = x^m - [\delta^1 \dots \delta^m] [\delta^1 \dots \delta^m]^{-1} \nabla f(x^m)$$

é minimizador global de f

Prova:

1) temo que provar que $\delta^1, \dots, \delta^m$ são LI

$$r^j = \nabla f(x^j) - \nabla f(x^0) = G(x^j - x^0) = G \delta^j$$

$\Rightarrow G$ é simétrica def (+) $\Rightarrow$ cm δ^j são LI
então r^j são LI

$$[r^1 \dots r^m] = G [\delta^1 \dots \delta^m]$$

$$\Rightarrow [r^1 \dots r^m]^{-1} = [\delta^1 \dots \delta^m]^{-1} G^{-1}$$

$$\Rightarrow \tilde{x} = x^m - G^{-1} \nabla f(x^m)$$

$$\Rightarrow \nabla f(\tilde{x}) = G \tilde{x} - b = G x^m - \nabla f(x^m) - b = 0$$

Vamos a provar que DFP termina em n iterações (11)

Teorema: Se f é quadrática com Hessiana G definida positiva o DFP termina em n iterações

Prova

$$x^m = x^{m-1} - H_{n-1} \nabla f(x^{m-1})$$

$$H_{n-1} G p_i = p_i \quad 0 \leq i < n$$

$$p_j^T p_i = 0 \Rightarrow p_0, p_1, \dots, p_{n-1} \text{ são LI}$$

$$\sum_{i=0}^{n-1} \alpha_i p_i = 0 \Rightarrow p_j \sum_{i=0}^{n-1} \alpha_i G p_i = 0$$

$$\Rightarrow \alpha_j p_j^T G p_j = 0$$

$$\Rightarrow \alpha_j = 0$$

como

$$H_{n-1} G p_i = p_i \quad 0 \leq i \leq n-1 \Rightarrow$$

$$p_i \text{ são LI} \Rightarrow$$

$$H_{n-1} G = I \Rightarrow G^{-1} = H_{n-1}$$

$$G(x^n) = b = G(x^{n-1}) - G G^{-1} \nabla f(x^{n-1}) - b = \nabla f(x^{n-1}) - \nabla f(x^{n-1}) = 0$$

$$\Rightarrow \nabla f(x^n) = 0$$