

BFGS Quasi-Newton

(Broyden, Fletcher, Goldfarb, Shanno)

$$GD: x_{k+1} = x_k - \eta \nabla f(x_k)$$

$$N: x_{k+1} = x_k - \eta \nabla^2 f(x_k)^{-1} \nabla f(x_k)$$

- Qualquer aproximação quadrática define uma direção

$$f(x_k + d) \approx m_k(d) = f(x_k) + \nabla f(x_k)^T d + \frac{1}{2} d^T B_k d$$

$$d_x = -B_k^{-1} \nabla f(x_k)$$

B_k
definida
positiva

$$QN: x_{k+1} = x_k - \eta_k \underbrace{B_k^{-1}}_{\text{Line search}} \nabla f(x_k)$$

Atualizada iterativamente

* Propriedades desejadas

① $B_k^{-1} \nabla f(x_k)$ deve ser fácil de computar

② B_k deve aproximar a Hessiana

③ $m_k(d)$ de capturar informação sobre a curvatura de f na ~~na~~ direção d .

BFGS :

$$f(x_k + d) \approx m_k(d) \quad , \quad f(x_{k+1} + d) \approx m_{k+1}(d)$$

Requerimento:

m_{k+1} deve ter o mesmo gradiente de f nas últimas duas iterações.

$$(i) \quad \left. \nabla m_{k+1} \right|_{d=0} = \nabla f(x_{k+1})$$

$$\begin{aligned} [x_{k+1} &= x_k + \eta_k d_k \\ x_k &= x_{k+1} - \eta_k d_k] \end{aligned}$$

$$(ii) \quad \left. \nabla m_{k+1} \right|_{d=-\eta_k d_k} = \nabla f(x_k)$$

$$m_{k+1}(d) = f(x_{k+1}) + \nabla f(x_{k+1})^T d + \frac{1}{2} d^T B_k d$$

$$\nabla m_{k+1}(d) = \nabla f(x_{k+1}) + B_{k+1} d$$

$$(i) \quad \nabla m_{k+1}(0) = \nabla f(x_{k+1}) \quad - \text{satisfeito automaticamente}$$

$$(ii) \quad \nabla m_{k+1}(-\eta_k d_k) = \nabla f(x_{k+1}) - \eta_k B_{k+1} d_k = \nabla f(x_k)$$

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

$$\begin{aligned} x_{k+1} &= x_k + \eta_k d_k \Rightarrow \eta_k d_k = x_{k+1} - x_k \\ B_{k+1} \underbrace{(x_{k+1} - x_k)}_{s_k} &= \underbrace{\nabla f(x_{k+1}) - \nabla f(x_k)}_{y_k} \end{aligned}$$

$$B_{k+1} s_k = y_k \quad (2)$$

* Definindo a aproximação de Hessiana
 $B_{k+1} = \arg \min \|B - B_k\|$ (B_{k+1} é próxima de B_k)
 $B = B^T$ (simétrica)

$$B_{s_{k+1}} = y_k$$

Each choice of norm gives a different B_{k+1}

BFGS

$$H_k = B_k^{-1}$$

① H_{k+1} é próxima de H_k

② H_{k+1} é simétrica

③ $B_{s_{k+1}} = y_k \rightarrow s_k = (B_{k+1})^{-1} y_k$

$$\rightarrow s_k = H_k y_k$$

$$H_{k+1} = \arg \min \|H - H_k\|$$

s.a.

$$H = H^T$$

$$H y_k = s_k$$

Norma de Frobenius

$$\|A\|_F^2 = \sum_{i,j} \|A_{ij}\|^2$$

Norma de Frobenius ponderada

$$\|A\|_w^2 = \|w^{\frac{1}{2}} A w^{\frac{1}{2}}\|_F^2$$

BFGS usa

$$W = \int_0^1 \nabla^2 f(x_k + t h_k d_k) dt$$

↳ definida positiva

Então;

$$H_{k+1} = (I - P_k s_k y_k^T) H_k (I - P_k y_k s_k^T) + P_k s_k s_k^T$$

$$P_k = \frac{1}{y_k^T s_k}$$

Algoritmo BFGS

Dado x_0

$k = 0$

Repetir enquanto $\|\nabla f(x_k)\| > \epsilon$

$$d_k = -H_k \nabla f(x_k)$$

$$x_{k+1} = x_k + h_k d_k = x_k - h_k H_k \nabla f(x_k)$$

$$s_k = x_{k+1} - x_k$$

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

$$H_{k+1} = (I - P_k s_k y_k^T) H_k (I - P_k y_k s_k^T) + P_k s_k s_k^T$$

$$P_k = \frac{1}{s_k^T s_k}$$

Busca em Linha

- Decremento suficiente

$$f(x_k + \eta_k d_k) \leq f(x_k) + \alpha_1 \eta_k \nabla f(x_k)^T d_k$$

- condição de declive

$$\nabla f(x_k + \eta_k d_k)^T d_k \leq \alpha_2 \cdot \nabla f(x_k)^T d_k$$

$$0 < \alpha_1 < 1$$

$$0 < \alpha_2 < 1$$

Handwritten text in a cursive script, likely from a 19th-century document. The text is written in dark ink on aged paper. The first line appears to be a date or a reference number, followed by several lines of descriptive text. The handwriting is somewhat faded and difficult to decipher.

Continuation of the handwritten text. The script remains consistent with the first section. There are some small, illegible markings that could be initials or abbreviations. The overall tone of the document suggests a formal or official record.

Further handwritten text, showing a continuation of the narrative or record. The ink is slightly lighter here, and the paper shows more signs of age and wear. The text is still written in the same cursive style.

More lines of handwritten text. The script is consistent throughout the document. There are some small, illegible markings that could be initials or abbreviations. The overall tone of the document suggests a formal or official record.

Final section of handwritten text. The script is consistent throughout the document. There are some small, illegible markings that could be initials or abbreviations. The overall tone of the document suggests a formal or official record.