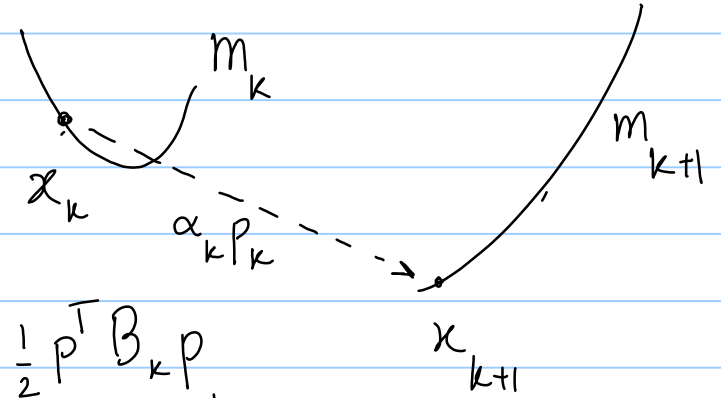
Quasi Newton Method.

NM $\rightarrow$ quadratic

QN $\rightarrow$ Superlinear

SD/CG $\rightarrow$ linear

$$\lim_{k\to\infty} \frac{\| x_{k+1} - x^* \|}{\| x_k - x^* \|} = \begin{cases} 0 & \text{S.L} \\ \mu & 0<\mu<1 \ \text{. Lin} \end{cases}$$



$$m_k(p) = f_k + \nabla f_k^T p + \frac{1}{2} p^T B_k p$$

$$\nabla m_k(p) = \nabla f_k + B_k p$$

quadratic Model of f around x_k.

assume B_k symmetric and P.D.

$m_k(0) = f_k, \quad \nabla m_k(0) = \nabla f_k$, $m_{k+1}(0) = f_{k+1}, \quad \nabla m_{k+1}(0) = \nabla f_{k+1}$

We want $\nabla m_k(p) = 0 \Rightarrow \underline{p_k = B_k^{-1} \nabla f_k}$

to minimize the model m_k

<u>QN motivation</u> $\rightarrow$ DO NOT compute B_{k+1} each time,

Instead, update from B_k.

Consider $m_{k+1} \rightarrow$ (1) does it give me the correct value of ∇f at x_{k+1}? ✓

(2) does it give me the correct value of ∇f at x_k?

$x_k \xrightarrow{\alpha_k p_k} x_{k+1}$

$m_{k+1}(p)$

$p = 0 \rightarrow x_{k+1}$

$p = ? \rightarrow x_k \qquad p = -\alpha_k p_k.$

$$\nabla m_{k+1}(p) = \nabla f_{k+1} + B_{k+1}\, p$$

$$\overbrace{\underbrace{\nabla m_{k+1}(-\alpha_k p_k) = \nabla f_{k+1} - B_{k+1}\alpha_k p_k}_{\text{from defn of } \nabla m_{k+1}} = \nabla f_k}^{\text{we want}}$$

Critical design choice.

$$\nabla f_{k+1} - \nabla f_k = B_{k+1}(\alpha_k p_k) = y_k \quad (\text{defn of } y_k) \qquad (\text{defn of } s_k)$$

$$\boxed{B_{k+1} s_k = y_k}\ , \qquad x_{k+1} - x_k = s_k = \alpha_k p_k$$

$\rightarrow$ Secant eqn

LM by s_k^T $\rightarrow$ $\left[\ \underline{s_k^T B_{k+1} s_k} = \underline{s_k^T y_k} > 0 \quad \text{for } B_{k+1} \text{ P.D.}\ \right]$

B_k symm $\Longrightarrow$ $\frac{(n)(n+1)}{2}$ unknowns.

I have n constraints from Secant eqn.

Remaining constraints up to us.

<u>L-BFGS relation</u> $\qquad s_k = B_{k+1}^{-1} y_k = H_{k+1} y_k.$

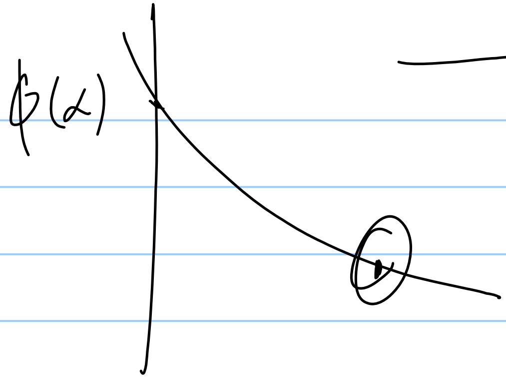
$$H_{k+1} = (I - \rho_k s_k y_k^T) \underbrace{H_k}_{\text{prev step}} (I - \rho_k s_k y_k^T) + \rho_k s_k s_k^T$$

$$\rho_k = (y_k^T s_k)^{-1}$$

Recap:

1. Pick x_0, estimate H_0
2. Inexact L.S. w/ Wolfe Condns $\rightarrow \alpha_k$
3. Direction: $p_k = -H_k \nabla f_k \rightarrow$ go to x_{k+1}
4. Check $\|\nabla f_k\|$ $\rightarrow$ End
   $\rightarrow$ else, use BFGS to get H_{k+1}, then 2, 3.

$\hookrightarrow$ If α_k satisfies the Wolfe condn (curvature Condn) then B_{k+1} is always PD!



$$\longrightarrow \quad \phi'(\alpha) \leqslant c_2\, \phi'(0)$$

$$-\nabla f.(x_k + \alpha p_k)^T p_k$$

$$\nabla f_{k+1}^T p_k \geqslant c_2\, \nabla f_k^T p_k \longrightarrow \text{sub } \nabla f_k^T p_k$$

$$\underbrace{(\nabla f_{k+1} - \nabla f_k)^T p_k}_{y_k^T p_k} \geqslant \underbrace{(1 - c_2)}_{>0}\underbrace{(-1)(\nabla f_k^T p_k)}_{>0} \qquad c_2\,(0,1)$$

$$\text{mult by } \alpha_k \longrightarrow y_k^T(\alpha_k p_k) = \underline{y_k^T s_k > 0}$$

$$B_{k+1}^T s_k = y_k \longrightarrow s_k^T B_{k+1} s_k = s_k^T y_k > 0$$

$$\Rightarrow B_{k+1} \text{ is P.D.}$$

Least Squares Problems

Radioactive decay! Ch 10 of NW

$y(x,t) = x_1 + x_2 \exp(-x_3 t)$ $\underbrace{N}\,L\,\underbrace{L}\,S$

Measure $y_1 \to t_1,\ y_2 \to t_2, \ldots$ (known)

$(x_1, x_2, x_3) \longrightarrow$ unknowns

$\{t_i\}_{i=1}^{m} \xrightarrow{\text{meas}} \{y_i\}_{i=1}^{m}$

Come up with a model.

Error measure:

① $| y_i - y(x, t_i) |$

② $(y_i - y(x, t_i))^2$

$$\hat{x} = \underset{x}{\operatorname{argmin}} \sum_{j=1}^{m} (y_j - y(x, t_j))^2$$

Least Squares

$y(x,t) = x_1 + x_2 t + x_3 t^2 \rightarrow$ linear least squares