LECTURE 12

QUASI-NEWTON METHODS

Quadratic model about x_k

$$Q_k(x) = f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2}(x - x_k)^T B_k (x - x_k)$$

Quasi-Newton methods keep an estimate B_k for $\nabla^2 f(x_k)$.

B_k is updated at every iteration.

SECANT EQUATION

How B_k and B_{k+1} are related?

(x_{k+1} is already computed using $Q_k(x)$; now we need to [redacted] form $Q_{k+1}(x)$ and B_{k+1})

Require $Q_{k+1}(x)$ to satisfy

$$\text{(i)} \quad \nabla Q_{k+1}(x_{k+1}) = \nabla f(x_{k+1})$$

$$\text{(ii)} \quad \nabla Q_{k+1}(x_k) = \nabla f(x_k)$$

Note $\nabla Q_{k+1}(x) = \nabla f(x_{k+1}) + B_{k+1}(x - x_{k+1})$

(i) always satisfied regardless of B_{k+1}

(ii) $\nabla Q_{k+1}(x_k) = \nabla f(x_{k+1}) + B_{k+1}(x_k - x_{k+1})$
$= \nabla f(x_k)$

$\Longrightarrow$

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

SECANT EQN

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

where

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

Secant equation is underdetermined. It is a system of n linear equation in $n(n+1)/2$ unknowns. Its solution is not unique.

SR-1 METHOD

Suppose B_{k+1} is a simple rank one update to B_k.

$$(*)\ B_{k+1} = B_k + \sigma u u^T$$

$$(\sigma = 1 \text{ OR } -1 \text{ (SIGN)})$$

subject to

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

Multiply both sides of $(*)$ by s_k from right, and eliminate secant eqn.

$$y_k = B_k s_k + \sigma u (u^T s_k)$$

It follows that (since $u^T s_k$ is a scala

$$(***)\ u = c\,(y_k - B_k s_k)$$

implying

$$y_k = B_k s_k + \sigma c^2 (y_k - B_k s_k)\left((y_k - B_k s_k)^T\right.$$

$$\Longrightarrow$$

$$\sigma c^2 = \left((y_k - B_k s_k)^T s_k\right)^{-1}$$

Plug (***) in (*)

$$B_{k+1} = B_k + \sigma c^2 (y_k - B_k s_k)(y_k - B_k s_k$$

$$= B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)}{\left((y_k - B_k s_k)^T s_k\right)}$$

$$\boxed{\begin{array}{l}\text{SR-1 UPDATE RULE} \\ B_{k+1} = B_k + \dfrac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{\left((y_k - B_k s_k)^T s_k\right)}\end{array}}$$

Letting $H_k = B_k^{-1}$ and $H_{k+1} = B_{k+1}^{-1}$ it can be deduced that

$$\boxed{\begin{array}{l}\text{INVERSE SR-1 UPDATE RULE} \\ H_{k+1} = H_k + \dfrac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(s_k - H_k y_k)^T y_k}\end{array}}$$

EXAMPLE (SR-1)

One iteration of SR-1 on

$$f(x) = x_1^4 + x_1 x_2 + (1+x_2)^2$$

with $x_k = (0, 1)$, $H_k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\alpha_k = 1$.

① Search direction

$$\nabla f(x_k) = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$$p_k^{SR} = -H_k \nabla f(x_k) = \begin{bmatrix} -1 \\ -4 \end{bmatrix}$$

② Next estimate

$$x_{k+1} = x_k + \alpha_k p_k^{SR} = \begin{bmatrix} -1 \\ -3 \end{bmatrix}$$

$$\nabla f(x_{k+1}) = \begin{bmatrix} -7 \\ -5 \end{bmatrix}$$

③ Inverse Hessian update

$$s_k = \underbrace{\begin{bmatrix} -1 \\ -3 \end{bmatrix}}_{x_{k+1}} - \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{x_k} = \begin{bmatrix} -1 \\ -4 \end{bmatrix}$$

$$y_k = \underbrace{\begin{bmatrix} -7 \\ -5 \end{bmatrix}}_{\nabla f(x_{k+1})} - \underbrace{\begin{bmatrix} 1 \\ 4 \end{bmatrix}}_{\nabla f(x_k)} = \begin{bmatrix} -8 \\ -9 \end{bmatrix}$$

$$s_k - H_k y_k = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$$

$$H_{k+1} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{H_k} + \frac{\overbrace{\begin{bmatrix} 7 \\ 5 \end{bmatrix}}^{(s_k - H_k y_k)} \overbrace{\begin{bmatrix} 7 & 5 \end{bmatrix}}^{(s_k - H_k y_k)^T}}{\underbrace{\begin{bmatrix} 7 & 5 \end{bmatrix}}_{(s_k - H_k y_k)^T} \underbrace{\begin{bmatrix} -8 \\ -9 \end{bmatrix}}_{y_k}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{101} \begin{bmatrix} 49 & 35 \\ 35 & 49 \end{bmatrix}$$

$$= \begin{bmatrix} 52/101 & -35/101 \\ -35/101 & 52/101 \end{bmatrix}$$