Quasi - Newton Methods (Q-N)

- In the *G-M-W* text Q-N means that we iteratively approximate the Hessian.
- Begin with an initial approximation $B_0 \approx G(x_0)$.

- **At each iteration in Algorithm U we replace**

$$B_{k+1} \leftarrow B_k + U_k$$

- **The Hessian describes the change in the gradient**

$$g(x_k + s_k) \approx g(x_k) + G(x_k)s_k \qquad *$$

where $s_k = \alpha p_k$ is the step.

Hessian Updates

- We have B_k, g_k and use

$$B_k p_k = -g_k$$

to find a search direction.

- We find an acceptable step, say

$$s_k = \alpha p_k \longrightarrow x_{k+1} = x_k + s_k$$

- **At the new point we compute a new gradient**

$$g_{k+1} = g(x_{k+1})$$

- **We re-write the *Q-N* condition (*) as**

$$B_{k+1}s_k = g_{k+1} - g_k \equiv y_k$$

- **Here s_k, and y_k are known and**

we want B_{k+1}.

- **Clearly there is not a unique solution**

Hessian Updates

- **One idea is to make the *least* change**

$$B_{k+1} = B_k + U_k$$

where U_k is *small*.

- **First approach is to insist on a**

rank-one update

$$\boldsymbol{U_k = uv^T}$$

- For any $x \in \mathbb{R}^n$ the image point

$$\left[\boldsymbol{uv^T}\right]\boldsymbol{x} = \boldsymbol{u}\langle \boldsymbol{v}, \boldsymbol{x}\rangle$$

is along the $\boldsymbol{u}$ direction. Since the range-space is 1-dimensional the matrix is rank one.

- **Applying the Q-N we get**

$$U_k = \frac{1}{v^T s_k} \left(y_k - B_k s_k \right) v^T$$

for any vector $v \in \mathbb{R}^n$ such that $v^T s_k \neq 0$.

- **If we require symmetry then we**

get

$$\boldsymbol{U_k = \frac{1}{(y_k - B_k s_k)^T s_k} (y_k - B_k s_k)(y_k - B_k s_k)^T}$$

Hessian Updates
Rank Two

➤ **Suppose we choose some $w \in \mathbb{R}^n$ that is orthogonal to s_k. Adding a second term to the**

update

$$\boldsymbol{U_k = uv^T + zw^T}$$

has no effect on the $\boldsymbol{Q}$-$\boldsymbol{N}$ condition since $\left[\boldsymbol{zw^T}\right]\boldsymbol{s_k} = \mathbf{0}$.

- **A variety of rank-two updates have been proposed**
- **One important collection is**

known as the Broyden family (See (4.35), page 119 G-M-W).

- **Two of the more famous are**
  - **Davidon-Fletcher-Powell (DFP) - 1959**
  - **Broyden-Fletcher-Goldfarb-Shanno (BFGS) -1969**

- **The BFGS update can be written as**

$$B_{k+1} = B_k +$$

$$\frac{1}{s_k^T B_k s_k} B_k s_k s_k^T B_k$$

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

Hessian Updates
Properties

- The Broyden family exhibits symmetry and *hereditary* positive definiteness

$$B_k > 0 \rightarrow B_{k+1} > 0$$

- **Finite termination on quadratics**
- **With *exact* 1-D searches members of the family produce the same sequence of iterates (but different matrices $\{B_k\}$)**