Aopt

Alain, Jonas, Vincent

Exercise 7

Secant Equation:

The secant equation is

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

where $s_k = x_{k+1} - x_k$, $y_k = \nabla f_{k+1} - \nabla f_k$, and B_{k+1} is an approximation of the Hessian $\nabla^2 f_{k+1}$

Prove that the secant equation is valid to find the next hessian approximation B_{k+1}

We need to prove that it satisfies two conditions:

- $\nabla m_{k+1}(0) \overset{!}{=} \nabla f_{k+1}$:

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

$\Rightarrow \nabla m_k(v) = \nabla f_k + B_k v$

$\Rightarrow \nabla m_{k+1}(0) = \nabla f_{k+1} + B_{k+1} \cdot 0 = \nabla f_{k+1}$ ✓

- $\nabla m_{k+1}(-t_k v_k) \overset{!}{=} \nabla f_k$:

$\nabla m_{k+1}(-t_k v_k) = \nabla m_{k+1}(-(x_{k+1} - x_k)) = \nabla m_{k+1}(-s_k)$

$= \nabla f_{k+1} + B_{k+1}(-s_k)$

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

$\Rightarrow -B_{k+1} s_k = \nabla f_k - \nabla f_{k+1}$

$\Rightarrow \nabla f_{k+1} + B_{k+1}(-s_k) = \nabla f_k$

$\Rightarrow \nabla m_{k+1}(-t_k v_k) = \nabla f_k$ ✓ #

Curvature Condition:

- Show that the curvature condition $s_k^T y_k > 0$ is required for the Quasi-Newton method
- Prove that the curvature condition is satisfied when the line search algorithm for the Wolfe conditions is used

---

- To ensure a descent direction, we need $B \in S^n_{++}$, positive definite

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

$\Rightarrow \underbrace{s_k^T B_{k+1} s_k}_{>0 \text{ because } B_{k+1} \text{ positive definite}} = s_k^T y_k$

> 0 because B_{k+1} positive definite $\Rightarrow s_k^T y_k > 0$ #

- Usefull equalities:

$$v_k = -B_k^{-1} \nabla f_k$$

$$s_k = x_{k+1} - x_k = x_k - r_k B_k^{-1} \nabla f_k - x_k = -r_k B_k^{-1} \nabla f_k$$

$$\Rightarrow s_k = r_k v_k \qquad \Rightarrow v_k = \frac{s_k}{r_k}$$

To satisfy the strong Wolfe conditions with $\beta \in (\alpha, 1)$, it must satisfy the curvature condition:

$$\nabla f(x_k + r_k v_k)^T v_k \geq \beta \nabla f(x_k)^T v_k$$

$$\Rightarrow \nabla f_{k+1}^T v_k \geq \beta \nabla f_k^T v_k$$

$$\nabla f_{k+1}^T v_k - \nabla f_k^T v_k \geq \beta \nabla f_k^T v_k - \nabla f_k^T v_k$$

$$(\nabla f_{k+1} - \nabla f_k)^T v_k \geq (\beta - 1) \nabla f_k^T v_k$$

$$y_k^T v_k \geq (\beta - 1) \nabla f_k^T v_k$$

$$y_k^T \cdot \frac{s_k}{r_k} \geq (\beta - 1) \nabla f_k^T v_k$$

$$y_k^T s_k \geq \underbrace{r_k}_{>0} \underbrace{(\beta - 1)}_{<0} \underbrace{\nabla f_k^T v_k}_{<0} > 0$$

$$\Rightarrow y_k^T s_k > 0 \qquad \#$$