April 5, 2021

Newton: $\min_{x \in \mathbb{R}^n} f(x)$

for $k = 1, 2, 3, \ldots$

$$x_{k+1} = x_k - H_k^{-1} \nabla f_k$$

H_k SPD $\Rightarrow$ descent direction

How expensive?

1. compute H_k, ∇f_k
2. $H_k p_k = \nabla f_k$ $O(n^3)$

Problems:

1. Difficult to compute H_k
2. $O(n^3)$ expensive
3. $O(n^2)$ storage

Quasi-Newton: cheap Hessian approx

$$H(x_k) \approx B_k$$

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

>0

$+ O(\|\alpha_k p_k\|^2)$

if G_k SPD, $\exists\ \alpha_k > 0$ s.t.

$f(x_{k+1}) < f(x_k)$

Claim: G_{k+1} is SPD if G_k SPD and α_k chosen "appropriately"

Limited memory BFGS (L-BFGS)

only use last l (< 20) steps of $y_k, \alpha_k p_k$, when computing $G_k \nabla f_k$