April 7, 2021

$$f(x+p) \approx f(x) + \nabla f(x)^T p + \frac{1}{2} p^T H(x) p$$

Newton: exact minimization of quadratic approx (H SPD)

$$p = -H(x)^{-1} \nabla f(x)$$

$$f(x+p) \approx f(x) - \nabla f(x)^T H^{-1}(x) \nabla f(x)$$ ≥ 0 (H SPD)

$x_{k+1} = x_k + p_k$ $\quad x_{k+1} = x_k + \alpha_k p_k$ $\quad \alpha_k > 0$

Modified Hessian $\quad \tilde{H} = H + E \quad \tilde{H}$ SPD

Quasi-Newton (e.g. BFGS) $\quad$ approx H with B $\quad p = -B^{-1} \nabla f(x)$

$x_{k+1} = x_k + \alpha_k p_k$ search direction

step size

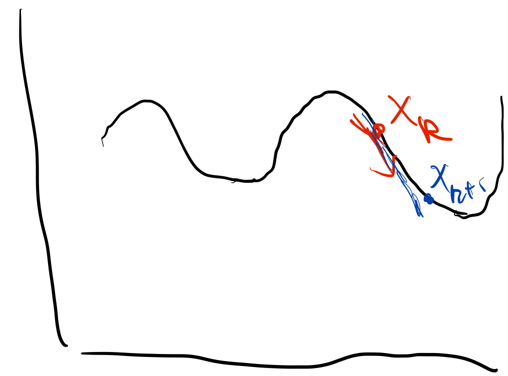
Gradient descent aka steepest descent

$$f(x+p) \approx f(x) + \nabla f(x)^T p$$

$$p = -\nabla f(x) \quad \text{(descent direction)}$$

$$f(x+p) \approx f(x) - \|\nabla f(x)\|_2^2 \quad > 0$$

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



, SGD)

$$f(x) = \frac{1}{2} x^T A x + x^T b + c$$

$$\alpha_k \quad \underset{\alpha}{\operatorname{argmin}} \; f(x_k - \alpha g_k)$$

$$= \underset{\alpha}{\operatorname{argmin}} \; \cancel{\frac{1}{2} x_k^T A x_k} + \alpha^2 \frac{1}{2} g_k^T A g_k - \alpha g_k^T A x_k$$

$$+ (\cancel{x_k} - \alpha g_k)^T b$$

$$\frac{\partial}{\partial \alpha} = 0 \Rightarrow \alpha g_k^T A g_k \underbrace{- g_k^T A x_k - g_k^T b}_{-g_k^T (A x_k + b) = -g_k^T g_k}$$

$$\alpha_k = \frac{g_k^T g_k}{g_k^T A g_k}$$