Kvasi-Newton metoder

Problem: $\min\limits_{x \in \mathbb{R}^n} f(x)$

Metod: $x^{k+1} = x^k - \alpha_k D_k \nabla f(x^k)$, $k = 0, 1, 2, \ldots$, där D_k är positivt definit och α_k löser $\min f(x^k - \alpha D_k \nabla f(x^k))$, eventuellt bara approximativt.

Idealt: $D_k = \nabla^2 f(x^k)^{-1}$, vilket ger Newtons modifierade metod. Denna är dock ofta för beräkningskrävande.

Åtgärd: använd approximationer av $\nabla^2 f(x^k)^{-1}$ [vilka självklart direkt svarar mot approximationer av $\nabla^2 f(x^k)$].

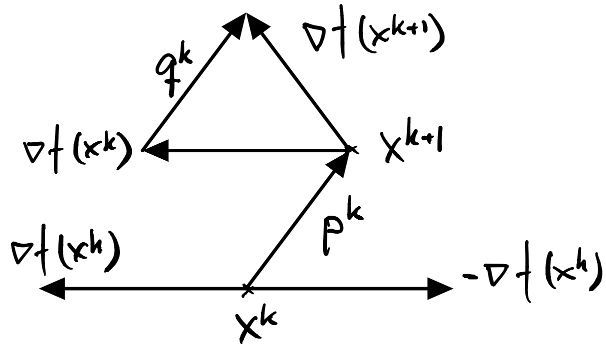
Rimliga krav på approximationer av $\nabla^2 f(x^k)^{-1}$?
Andra ordningens Taylor-utveckling av $f(x)$ i x^{k+1}:

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

$$\Rightarrow \nabla f(x) \approx \nabla f(x^{k+1}) + \nabla^2 f(x^{k+1})(x - x^{k+1}) \underset{x = x^k}{\Longrightarrow}$$

$$\nabla f(x^{k+1}) - \nabla f(x^k) \approx \nabla^2 f(x^{k+1})(x^{k+1} - x^k)$$

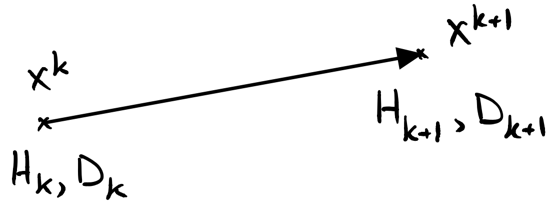
Med $p^k = x^{k+1} - x^k$ och $q^k = \nabla f(x^{k+1}) - \nabla f(x^k)$ fås $\underline{q^k \approx \nabla^2 f(x^k) p^k}$, eller $\underline{\nabla^2 f(x^{k+1})^{-1} q^k = p^k}$.



$p^k = x^{k+1} - x^k$
$q^k = \nabla f(x^{k+1}) - \nabla f(x^k)$

Om $f(x)$ är kvadratisk, säg $f(x) = \frac{1}{2}x^T H x - b^T x$, så fås: $q^k = \nabla f(x^{k+1}) - \nabla f(x^k) = Hx^{k+1} - b - (Hx^k - b) =$ $= H(x^{k+1} - x^k) = Hp^k$. Antag vidare att H är icke-singulär och att $n+1$ punkter $x^0, \dots, x^n$ är givna, sådana att vektorerna $p^0 = x^1 - x^0$, $p^1 = x^2 - x^1, \dots, p^{n-1} = x^n - x^{n-1}$ är linjärt oberoende. Låt $q^k = Hp^k$, $k = 0, \dots, n-1$. Då fås: $[q^0, \dots, \_^{n-1} = [Hp^0, \dots, Hp^{n-1}] =$
$= H[p^0, \dots, p^{n-1}] \Rightarrow H = [q^0, \dots, q^{n-1}][p^0, \dots, p^{n-1}]^{-1}$
och $H^{-1} = [p^0, \dots, p^{n-1}][q^0, \dots, q^{n-1}]^{-1}$.
Alltså: H och H^{-1} kan beräknas mha exakt n stycken differenser p^k och q^k.

Slutsats: en nödvändig egenskap hos approximationer H_{k+1} av $\nabla^2 f(x^{k+1})$ är att $H_{k+1} p^k = q^k$. (Ty annars återskapas inte $\nabla^2 f(x)$ för en kvadratisk funktion.) Analogt måste approximationer D_{k+1} av $\nabla^2 f(x^{k+1})^{-1}$ uppfylla att $D_{k+1} q^k = p^k$.



Uppdatering av D_k (eller H_k) under kravet $D_{k+1} q^k = p^k$ (eller $H_{k+1} p^k = q^k$) karakteriserar *kvasi-Newton metoder*.

Vanligaste uppdateringsstrategin: $D_{k+1} = D_k + \Delta D_k(D_k, p^k, q^k)$ där $\Delta D_k(D_k, p^k, q^k)$ är en enkel uppdatering som väljs så att $D_{k+1} q^k = p^k$ uppfylls. Alternativt: $H_{k+1} = H_k + \Delta H_k(H_k, p^k, q^k)$ där $\Delta H_k(H_k, p^k, q^k)$ är en enkel uppdatering som väljs så att $H_{k+1} p^k = q^k$ uppfylls.