Quasi-Newton Methods

Idea: instead of Newton $d_k = -H_k^{-1} g_k$

let $d_k = -D_k g_k$

$$\left(\begin{array}{l} H_k = \nabla^2 f(x_k) \\ g_k = \nabla f(x_k) \end{array} \right)$$

where $D_{k+1} = D_k + \Delta D_k$

so that

- D_k pos. def.
- $D_{k+1}(g_{i+1} - g_i) = x_{i+1} - x_i$ $\quad i = 1, \dots, k,$ (QN-condition) cf. (Δg)

The reason is that if QN is applied on a quadratic function with pos. def. Hessian H and $d_1, \dots, d_n$ are linearly independant, then

$$D_{n+1}(g_{k+1} - g_k) = x_{k+1} - x_k \quad \overset{(\Delta g)}{\Longleftrightarrow}$$

$$D_{n+1} H (x_{k+1} - x_k) = \underbrace{x_{k+1} - x_k}_{\lambda_k d_k} \quad \Longleftrightarrow$$

$$D_{n+1} H d_k = d_k \quad \Longrightarrow$$

$$D_{n+1} H (d_1 \; d_2 \dots d_n) = \underbrace{(d_1 \dots d_n)}_{\text{invertible}} \quad \Longleftrightarrow$$

$$D_{n+1} H = I$$

The DFP algorithm (30) on p. 82 and the BFGS algorithm p. 89 produce conjugate directions for $f = q$. Hence, equivalent to CG method and termination in at most n line searches.

⊕ Robust, fast convergence, less computations than Newton.

⊖ Store matrices