Quasi-Newton Methods: Conjugate Gradients

Given the first order Taylor series expansion, we know that

$$\nabla_x F(x_k + s_k) = \nabla_x F(x_k) + \nabla^2_{xx} F(x_k) s_k$$

for $s_k \in N(x_k, \delta)$. This expression yields

$$\nabla_x F(x_k + s_k) - \nabla_x F(x_k) = \nabla^2_{xx} F(x_k) s_k$$

which yields

$$s_k(\nabla_x F(x_k + s_k) - \nabla_x F(x_k)) = s_k \nabla^2_{xx} F(x_k) s_k.$$

This indicates that $s_k(\nabla_x F(x_k + s_k) - \nabla_x F(x_k))$ contains the same curvature information as the original Hessian. This forms the Quasi-Newton curvature restriction that the approximated curvature along the step equal the actual curvature. Setting

$$y_k = \nabla_x F(x_k + s_k) - \nabla_x F(x_k)$$

this amount to

$$\begin{aligned} y_k &= \nabla^2_{xx} F(x_{k+1}) s_k \\ y_k &= B_{k+1} s_k. \end{aligned}$$

If

$$\begin{aligned} B_{k+1} &= B_k + uv' \\ y_k &= (B_k + uv') s_k \end{aligned}$$

Thus,

$$\begin{aligned} u(v's_k) &= y_k - B_k s_k \\ u &= \frac{1}{v's_k}(y_k - B_k s_k) \\ B_{k+1} &= B_k + \frac{1}{v's_k}(y_k - B_k s_k) v' \end{aligned} .$$

The simpliest construction which maintains symmetry is

$$B_{k+1} = B_k + \frac{1}{(y_k - B_k s_k)' s_k}(y_k - B_k s_k)(y_k - B_k s_k)'.$$

Other than the simple update, no vector v will satisfy the restriction of symmetry. One alternative is start with a nonsymmetric matrix $B^{(1)}$ and generate $B^{(2)}$ by

$$B^{(2)} = \frac{1}{2}(B^{(1)} + B^{(1)\prime})$$

The sequence of this replacement taken to the limit yields

$$B_{k+1} = B_k + \frac{1}{v's_k}\left((y_k - B_k s_k)v' + v(y_k - B_k s_k)'\right) - \frac{(y_k - B_k s_k)' s_k}{(v's_k)^2} vv'.$$

To prominent conjugate gradient formulations involve setting $v = y_k$ yielding the Davidon-Fletcher-Powell (DFP) update:

$$B_{k+1} = B_k - \frac{1}{s_k' B_k s_k} B_k s_k s_k' B_k + \frac{1}{y_k s_k} y_k y_k' + (s_k' B_k s_k) w_k w_k',$$

where

$$w_k = \frac{1}{y_k' s_k} y_k - \frac{1}{s_k' B_k s_k} B_k s_k.$$

And the Broyden-Fletcher-Golfarb-Shanno (BFGS) update

$$B_{k+1} = B_k - \frac{1}{s_k' B_k s_k} B_k s_k s_k' B_k + \frac{1}{y_k' s_k} y_k y_k'.$$