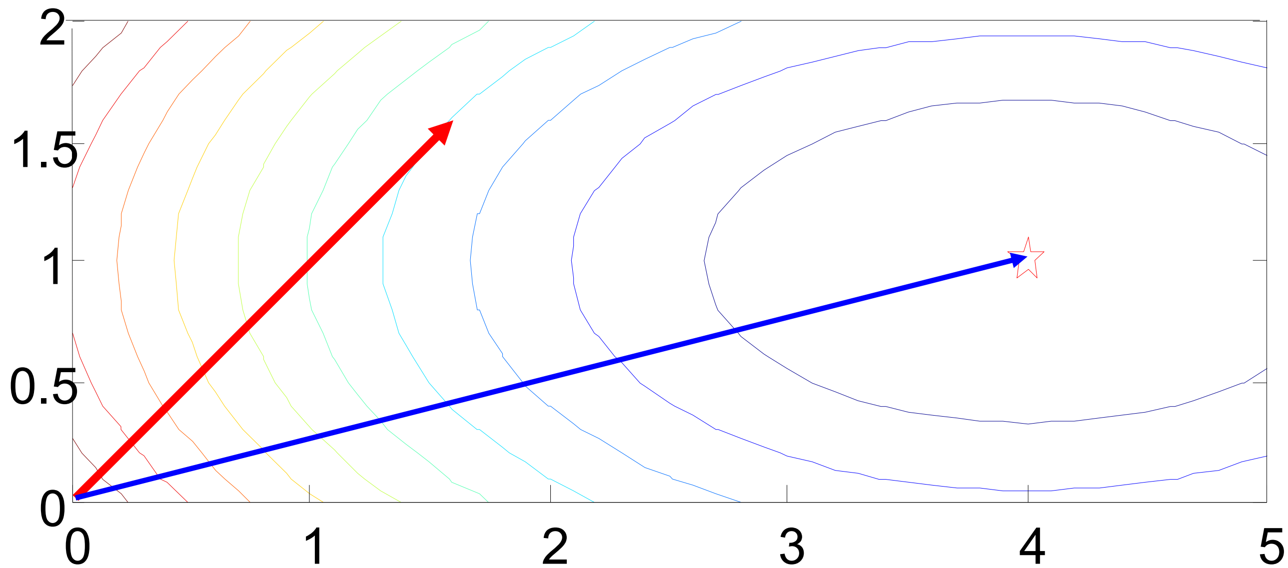
Consider $f(x) = \frac{1}{2}x^T Q x - c^T x$

$$Q = \begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix}, c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$g_0 = \nabla f(x_0) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, p_0^{\mathrm{N}} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \quad x^* = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$



BFGS

- Find the quasi-Newton's direction
  - Give the initial guess of Hessian B_0 = I

$$p_0 = -B_0^{-1} g_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- Use exact line search,

$$f(\alpha) = \tfrac{1}{2}\alpha^2 p_0^T Q p_0 - \alpha c^T p_0$$

$$\alpha_0 = \frac{c^T p_0}{p_0^T Q p_0} = \frac{2}{1.25} = 1.6$$

$$x_1 = x_0 + \alpha_0 p_0 = \begin{pmatrix} 1.6 \\ 1.6 \end{pmatrix}$$

$$g_1 = \nabla f(x_1) = Qx_1 - c = \begin{pmatrix} -0.6 \\ 0.6 \end{pmatrix}$$

$$s_0 = x_1 - x_0 = \begin{pmatrix} 1.6 \\ 1.6 \end{pmatrix} \qquad y_0 = g_1 - g_0 = \begin{pmatrix} 0.4 \\ 1.6 \end{pmatrix}$$

$$B_1 = B_0 - \frac{B_0 s_0 s_0^T B_0}{s_0^T B_0 s_0} + \frac{y_0 y_0^T}{y_0^T s_0} = I - \frac{s_0 s_0^T}{s_0^T s_0} + \frac{y_0 y_0^T}{y_0^T s_0}$$

$$\frac{s_0 s_0^T}{s_0^T s_0} = \begin{pmatrix} .5 & .5 \\ .5 & .5 \end{pmatrix}, \frac{y_0 y_0^T}{y_0^T s_0} = \begin{pmatrix} .05 & .2 \\ .2 & .8 \end{pmatrix}$$

$$B_1 = \begin{pmatrix} .55 & -.30 \\ -.30 & 1.30 \end{pmatrix} \qquad \text{Check } B_1 s_0 = y_0$$

BFGS, the second step

- Find the quasi-Newton's direction p_1

$$p_1 = -B_1^{-1} g_1 = -\begin{pmatrix} .55 & -.30 \\ -.30 & 1.30 \end{pmatrix}^{-1} \begin{pmatrix} -0.6 \\ 0.6 \end{pmatrix} = \begin{pmatrix} 0.96 \\ -0.24 \end{pmatrix}$$

- Use exact line search,

$$f(\alpha) = \frac{1}{2}(x_1 + \alpha p_1)^T Q (x_1 + \alpha p_1) - c^T (x_1 + \alpha p_1)$$

$$\alpha_1 = \frac{c^T p_1 - p_1^T Q x_1}{p_1^T Q p_1} = \frac{0.72}{0.288} = 2.5$$

$$x_2 = x_1 + \alpha_1 p_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

$$g_2 = \nabla f(x_2) = Qx_2 - c = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$s_1 = x_2 - x_1 = \begin{pmatrix} 2.4 \\ -0.6 \end{pmatrix} \quad y_1 = g_2 - g_1 = \begin{pmatrix} 0.6 \\ -0.6 \end{pmatrix}$$

$$B_2 = B_1 - \frac{B_1 s_1 s_1^T B_1}{s_1^T B_1 s_1} + \frac{y_1 y_1^T}{y_1^T s_1}$$

$$\frac{B_1 s_1 s_1^T B_1}{s_1^T s_1} = \begin{pmatrix} .5 & -.5 \\ -.5 & .5 \end{pmatrix}, \frac{y_1 y_1^T}{y_1^T s_1} = \begin{pmatrix} .2 & -.2 \\ -.2 & .2 \end{pmatrix}$$

$$B_2 = \begin{pmatrix} .25 & 0 \\ 0 & 1 \end{pmatrix} = Q$$

SR1

- Use the same initial B_0=I. The same x_1,g_1.

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

$$s_0 = x_1 - x_0 = \begin{pmatrix} 1.6 \\ 1.6 \end{pmatrix} \qquad y_0 = g_1 - g_0 = \begin{pmatrix} 0.4 \\ 1.6 \end{pmatrix}$$

$$B_1 = I + \frac{(y_0 - s_0)(y_0 - s_0)^T}{(y_0 - s_0)^T s_0} = \begin{pmatrix} .25 & 0 \\ 0 & 1 \end{pmatrix} = Q$$

$$p_1 = -B_1^{-1} g_1 = -\begin{pmatrix} .25 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} -0.6 \\ 0.6 \end{pmatrix} = \begin{pmatrix} 2.4 \\ -0.6 \end{pmatrix}$$