DFP Method: Quasi Newton method

Newton method
$X_{R+1} = X_R - \underbrace{H_f^{-1}(X_R)\nabla f(X_R)^t}_{\text{direction } d_i}$
$\alpha_i = 1$

1. start with the initial point x_1 and positive definite symmetric matrix B_1 (If matrix B_1 is not given, take Identity matrix.) of order n, when function has n-variables.
   Set i = 1
2. Compute $\nabla f(x_1)$ and set $d_i = -[B_i]\nabla f(x_i)$.
3. find α_i by minimizing function $f(x_i + \alpha_i d_i)$
4. set $x_{i+1} = x_i + \alpha_i d_i$
5. Check x_{i+1} for optimality condition. If it satisfies then stop, otherwise go to the next step.
6. Update matrix B_i to B_{i+1} using $B_{i+1} = B_i + M_i + N_i$
   1. where $M_i = \alpha_i \frac{d_i d_i^t}{d_i^t g_i}$
   2. And $N_i = \frac{-(B_i g_i)(B_i g_i)^t}{g_i^t B_i g_i}$, where $g_i = \nabla f(x_i + 1) - \nabla f(x_i)$
   3. Set i = i + 1 and go to step 2.

$B_2 = B_1 + M_1 + N_1$

$M_1 = \alpha_1 \frac{d_1 d_1^t}{d_1^t g_1}$

$g_i = \nabla f(x_{i+1}) - \nabla f(x_i)$

$g_1 = \nabla f(x_2) - \nabla f(x_1)$

Note: d_i *is column vector and hence* d_i^t *is row vector, similiar for other vectors*

question: Minimize $f(x,y) = 3x^2 - 4xy + 2y^2 + 4x + 6$ starting from the point (0,0) using DFP Method. (Quasi Newton method)

$\|\nabla f(x_R)\| = (0,0)$ stopping condition

Solution :-

$$f(x,y) = 3x^2 - 4xy + 2y^2 + 4x + 6$$

$$\nabla f(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 6x - 4y + 4 \\ -4x + 4y \end{pmatrix}$$

$-12 + 8 + 4$

$-4 \times -2 + 8$

$$\nabla f(0,0) = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \qquad d_1 = -B_1 \nabla f(0,0)$$

$$d_1 = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 0 \end{bmatrix} = -\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 0 \end{bmatrix}$$

Find the optimal length α_1 by minimising

Find the optimal length α_1 by

$$g_1 = f\left((0,0) + \alpha_1(-4,0)\right)$$

$$g_1 = f(-4\alpha_1, 0)$$

$$f(x,y) = 3x^2 - 4xy + 2y^2 + 4x + 6$$

$$g_1 = 3(-4\alpha_1)^2 + 4(-4\alpha_1) + 6$$

$$g_1 = (16\times 3)\alpha_1^2 - 16\alpha_1 + 6$$

$$\frac{dg_1}{d\alpha_1} = (32\times 3)\alpha_1 - 16 \quad , \quad \frac{dg_1}{d\alpha_1} = 0$$

$$(32\times 3)\alpha_1 - 16 = 0$$

$$\alpha_1 = \frac{16}{32\times 3} = \frac{1}{6}$$

$$\frac{d^2 g_1}{d\alpha_1^2} = 32\times 3 > 0 \rightarrow \text{minima}$$

Set
$$X_2 = X_1 + \alpha_1 d_1$$
$$= (0,0) + \frac{1}{6}(-4,0) = \left(-\frac{2}{3}, 0\right)$$

Stopping condition

$$\left\|\nabla f\left(-\frac{2}{3}, 0\right)\right\| = \left(0, \frac{8}{3}\right) \neq (0,0)$$

Second Iteration

$$B_2 = B_1 + M_1 + N_1$$

$$B_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + M_1 + N_1$$

$$M_1 = \alpha_1 \frac{d_1 d_1^t}{d_1^t g_1}$$

$$g_1 = \nabla f(x_2) - \nabla f(x_1)$$

$$g_1 = \nabla f(-\tfrac{2}{3}, 0) - \nabla f(0,0)$$

$$g_1 = (0, \tfrac{8}{3}) - (4, 0)$$

$$g_1 = \begin{pmatrix} -4 \\ \frac{8}{3} \end{pmatrix} \qquad d_1 = \begin{pmatrix} -4 \\ 0 \end{pmatrix}$$

$$\boxed{A_{m\times n}\, B_{n\times l} = (AB)_{m\times l}}$$

$$M_1 = \frac{1}{6} \frac{\begin{pmatrix} -4 \\ 0 \end{pmatrix}_{2\times 1} (-4, 0)_{1\times 2}}{(-4, 0)_{1\times 2} \begin{pmatrix} -4 \\ 8/3 \end{pmatrix}_{2\times 1}} = \frac{1}{6} \begin{bmatrix} 16 & 0 \\ 0 & 0 \end{bmatrix} \cdot \frac{1}{16}$$

$$\overrightarrow{(-4,0)} \begin{pmatrix} -4 \\ 8/3 \end{pmatrix}\downarrow = 16 + 0 = 16$$

$$\overrightarrow{\begin{pmatrix} -4 \\ 0 \end{pmatrix}} (-4, 0)\downarrow = \begin{bmatrix} 16 & 0 \\ 0 & 0 \end{bmatrix}$$

$$M_1 = \begin{bmatrix} 1/6 & 0 \\ 0 & 0 \end{bmatrix}$$

$$g_1 = \begin{pmatrix} -4 \\ 8/3 \end{pmatrix}$$

$$N_1 = - \frac{(B_1 g_1)(B_1 g_1)^t}{g_1^t B_1 g_1}$$

$$B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 8/3 \end{bmatrix} = g_1$$

$$N_1 = - \frac{g_1 g_1^t}{g_1^t g_1} = - \frac{\begin{pmatrix} -4 \\ 8/3 \end{pmatrix}_{2\times 1} (-4, 8/3)_{1\times 2}}{[-4, 8/3]_{1\times 2} \begin{bmatrix} -4 \\ 8/3 \end{bmatrix}_{2\times 1}}$$

$$= 16 + \frac{64}{9}$$

$$N_1 = \begin{bmatrix} -9/13 & 6/13 \\ 6/13 & -4/13 \end{bmatrix}$$

$$B_2 = B_1 + M_1 + N_1$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1/6 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -\frac{9}{13} & \frac{6}{13} \\ \cdots & -\frac{4}{\cdots} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1/6 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -\frac{7}{13} & \frac{0}{13} \\ 6/13 & -4/13 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 37/78 & 6/13 \\ 6/13 & 9/13 \end{bmatrix}$$

$$d_2 = -B_2 \nabla f(x_2) = -\begin{bmatrix} \frac{37}{78} & 6/13 \\ 6/13 & 9/13 \end{bmatrix} \begin{bmatrix} 0 \\ 8/3 \end{bmatrix}$$

$$d_2 = \begin{bmatrix} \frac{-16}{13} \\ -24/13 \end{bmatrix}$$

Find α_2 by minimizing the function

$$f(x_2 + \alpha_2 d_2) = f\left(\left(-\frac{2}{3}, 0\right) + \alpha_2\left(\frac{-16}{13}, \frac{-24}{13}\right)\right)$$

Exercise

$$= f\left(\frac{-2}{3} - \frac{16}{13}\alpha_2, \frac{-24}{13}\alpha_2\right)$$

$$\alpha_2 = \frac{13}{12}$$

$$x_3 = x_2 + \alpha_2 d_2$$

$$x_3 = \left(-\frac{2}{3}, 0\right) + \frac{13}{12}\left(\frac{-16}{13}, \frac{-24}{13}\right)$$

$$x_3 = (-2, -2)$$

Stopping condition:- $\|\nabla f(x_3)\| = (0,0)$

$x_3 = (-2,-2)$ is minima.