P1) $f(x,y) = (x+y-1)^4 + (x-y)^2$

$x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$H_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\alpha = 1$

First update step (k=0)

Gradient at x_0:

$$\nabla f(x_0) = \begin{bmatrix} 4(0+0-1)^3 + 2(0-0) \\ 4(0+0-1)^3 - 2(0-0) \end{bmatrix} = \begin{bmatrix} -4 \\ -4 \end{bmatrix}$$

calculate x_1:

$$x_1 = x_0 - \alpha H_0 \nabla f(x_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -4 \\ -4 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$s_0 = x_1 - x_0 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\nabla f(x_1) = \begin{bmatrix} 4(4+4-1)^3 + 2(4-4) \\ 4(4+4-1)^3 - 2(4-4) \end{bmatrix} = \begin{bmatrix} 1372 \\ 1372 \end{bmatrix}$$

$$y_0 = \nabla f(x_1) - \nabla f(x_0) = \begin{bmatrix} 1372 \\ 1372 \end{bmatrix} - \begin{bmatrix} -4 \\ -4 \end{bmatrix} = \begin{bmatrix} 1376 \\ 1376 \end{bmatrix}$$

H_1 (BFGS & DFP are identical)

$$H_1 = \frac{1}{688} \begin{bmatrix} 345 & -343 \\ -343 & 345 \end{bmatrix}$$

second update (k=1)

calculate x_2:

$$x_2 = x_1 - \alpha H_1 \nabla f(x_1) = \begin{bmatrix} 4 \\ 4 \end{bmatrix} - \frac{1}{688} \begin{bmatrix} 345 & -343 \\ -343 & 345 \end{bmatrix} \begin{bmatrix} 1372 \\ 1372 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$s_1 = x_2 - x_1 = \begin{bmatrix} -4 \\ -4 \end{bmatrix}$$

$$\nabla f(x_2) = \begin{bmatrix} -4 \\ -4 \end{bmatrix}$$

$$y_1 = \nabla f(x_2) - \nabla f(x_1) = \begin{bmatrix} -4 \\ -4 \end{bmatrix} - \begin{bmatrix} 1372 \\ 1372 \end{bmatrix} = \begin{bmatrix} -1376 \\ -1376 \end{bmatrix}$$

i) BFGS Final Inverse Hessian (H_2^{BFGS}):

$$H_2^{BFGS} = \left(I - \frac{s_1 y_1^T}{y_1^T s_1}\right) H_1 \left(I - \frac{y_1 s_1^T}{y_1^T s_1}\right) + \frac{s_1 s_1^T}{y_1^T s_1} = \frac{1}{688}\begin{bmatrix} 689 & -687 \\ -687 & 689 \end{bmatrix}$$

DFP Final Inverse Hessian(H_2^{DFP}):

$$H_2^{DFP} = H_1 + \frac{s_1 s_1^T}{s_1^T y_1} - \frac{H_1 y_1 y_1^T H_1}{y_1^T H_1 y_1} = \frac{1}{688}\begin{bmatrix} 395 & -393 \\ -393 & 395 \end{bmatrix}$$

True Inverse Hessian at x_2:

The Hessian at $x_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is $J(0,0) = \begin{bmatrix} 19 & 10 \\ 10 & 19 \end{bmatrix}$

$$J^{-1}(0,0) = \frac{1}{96}\begin{bmatrix} 14 & -10 \\ -10 & 19 \end{bmatrix} = \begin{bmatrix} 0.1458 & -0.1042 \\ -0.1042 & 0.1458 \end{bmatrix}$$

H_2^{DFP} is significantly closer than th BFGS.

2) $f(x,y)=(1-x)^2+25(y-x^2)^2$, $(x_0,y_0)=(-1,1)$, $H_0=I$

$\alpha=1$

$$\nabla f(x,y)=\begin{bmatrix} -100x(y-x^2)+2x-2 \\ 50(y-x^2) \end{bmatrix}$$

$$\nabla^2 f(x,y)=\begin{bmatrix} 300x^2-100y+2 & -100x \\ -100x & 50 \end{bmatrix}$$

step0 $\rightarrow$ step 1 $(\alpha=1)$

$x_0=(-1,1)$ $\qquad g_0=\nabla f(-1,1)=\begin{bmatrix} -4 \\ 0 \end{bmatrix}$

$x=-1,\ y=1$

$P_0=-H_0 g_0=-I g_0=\begin{bmatrix} 4 \\ 0 \end{bmatrix}$

$\alpha_0=1$

$x_1=x_0+\alpha_0 P_0=(-1,1)+(4,0)=(3,1)$

grad at $x_1=(3,1)$

$g_1=\nabla f(3,1)=\begin{bmatrix} 2404 \\ -400 \end{bmatrix}$

compute s_0, y_0 and

$s_0=(4,0)$

$y_0=g_1-g_0=\begin{bmatrix} 2408 \\ -400 \end{bmatrix}$

$y_0^T s_0=2408(4)=9632$

Step-1

BFGS:

$H_1=(I-\rho s_0 y_0^T)H_0(I-\rho y_0 s_0^T)+\rho s_0 s_0^T$ $\qquad \rho=\frac{1}{y_0^T s_0}=\frac{1}{9632}$

$$H_1^{BFGS}=\begin{bmatrix} 5301/181202 & 50/301 \\ 50/301 & 1 \end{bmatrix}$$

DFP:

$$H_1^{DFP} = H_0 + \frac{s_0 s_0^T}{s_0^T y_0} - \frac{H_0 y_0 y_0^T H_0}{y_0^T H_0 y_0}$$

as $H_0 = I$

$$H_1^{DFP} = \begin{bmatrix} 0.028513 & 0.16165 \\ 0.16165 & 0.97319 \end{bmatrix}$$

Step-2

BFGS

$$P_1^{(B)} = -H_1^{BFGS} g_1$$

$$H_1^{(B)} = \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix}$$

$h_{11} = 0.02925$ $h_2 = 0.1661$, $h_{22} = 1$

$$P_1^{(B)} = \begin{bmatrix} -3.88298 \\ 0.66945 \end{bmatrix}$$

$$x_2^{(B)} = x_1 + P_1^{(B)} = (-0.88298, 1.669951)$$

$$g_2^{(B)} = \begin{bmatrix} 79.359 \\ 49.2397 \end{bmatrix}$$

$$s_1^{(B)} = x_2^{(B)} - x_1 = \begin{bmatrix} -3.882981 \\ 0.669951 \end{bmatrix}$$

$$y_1^{(B)} = g_2^{(B)} - g_1 = \begin{bmatrix} -2329.6401 \\ 499.2397 \end{bmatrix}$$

$$y_1^{(B)T} s_1^{(B)} = 9391.1253$$

$$H_2^{BFGS} = \begin{bmatrix} 0.037682 & 0.18887 \\ 0.18887 & 0.9919 \end{bmatrix}$$

DFP:

$$p_1^{(D)} = -H_1^{(DFP)} g_1$$

$$p_1^{(D)} = \begin{bmatrix} -3.88594 \\ 0.6466 \end{bmatrix}$$

$$x_2^{(D)} = x_1 + p_1^{(D)} = (-0.885945 , 1.6466095)$$

$$g_2^{(0)} = \begin{bmatrix} 72.5709 \\ 43.0855 \end{bmatrix}$$

$$s_1^{(0)} = x_2^{(D)} - x_1 = \begin{bmatrix} -3.8859 \\ 0.6966 \end{bmatrix}$$

$$y_1^{(D)} = g_2^{(D)} - g_1 = \begin{bmatrix} -2331.429 \\ 443.0855 \end{bmatrix}$$

$$y_1^{(D)T} s_1^{(D)} = 9346.3089$$

$$H_2 = H_1 + \frac{s_1 s_1^T}{s_1 y_1} - \frac{H_1 y_1 y_1^T H_1}{y_1^T H_1 y_1}$$

$$H_2^{DFP} = \begin{bmatrix} 0.027931 & 0.1382005 \\ 0.138200 & 0.7286437 \end{bmatrix}$$

At $x_2^{(B)}$

$$\nabla^2 f(x_2^{(B)}) = \begin{bmatrix} 69.951 & 88.2981 \\ 88.2981 & 50 \end{bmatrix}$$

$$(\nabla^2 f)^{-1}(x_2^{(B)}) = \begin{bmatrix} -0.01156 & 0.02092 \\ 0.02092 & -0.01606 \end{bmatrix}$$

$$\Delta^B = H_2^{BFGS} - (\nabla^2 f)^{-1}(x_2^{(B)}) \approx \begin{bmatrix} 0.0492 & 0.16\cancel{84} \\ 0.16895 & 1.0080 \end{bmatrix}$$

At $x_1^{(D)}$

$$(\nabla^2 f)^{-1}(x_2^D) = \begin{bmatrix} -0.011 & 0.02105 \\ 0.021 & -0.0173 \end{bmatrix}$$

∴ with fixed step $\alpha = 1$ DFP produced the approximate inverse Hessian closer to the true Hessiant.

3) $f(x,y) = e^{x+y} + (x-y)^2$ $(x_0, y_0) = (0,0)$, $H_0 = I$.

$$\nabla f(x,y) = \begin{bmatrix} e^{x+y} + 2(x-y) \\ e^{x+y} - 2(x-y) \end{bmatrix}, \quad \nabla^2 f(x,y) = \begin{bmatrix} e^{x+y}+2 & e^{x+y}-2 \\ e^{x+y}-2 & e^{x+y}+2 \end{bmatrix}$$

at $x_0 = (0,0)$

$$g_0 = \nabla f(0,0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$P_0 = -H_0 g_0 = -I \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$\alpha_0 = 1$

$x_1 = x_0 + \alpha_0 P_0 = (-1, -1)$

$$g_1 = \nabla f(-1,-1) = \begin{bmatrix} 0.135335 \\ 0.135335 \end{bmatrix}$$

$$S_0 = x_1 - x_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \qquad y_0 = g_1 - g_0 = \begin{bmatrix} -0.8646 \\ -0.86466 \end{bmatrix}$$

$$y_0^T S_0 = 1.729329$$

$$P_0 = \frac{1}{y_0^T S_0} = 0.57825$$

step ①

BFGS

$$H_1^{BFGS} = (I - P_0 S_0 y_0^T) H_0 (I - P_0 y_0 S_0) + P_0 S_0 S_0^T$$

$$H_1^{BFGS} = \begin{bmatrix} 1.07825 & 0.07825 \\ 0.07825 & 1.07825 \end{bmatrix}$$

DFP:

$H_0 = I$

$$H_1^{DFP} = I + \frac{s_0 s_0^T}{s_0^T y_0} - \frac{y_0 y_0^T}{y_0^T y_0}$$

$$H_1^{DFP} = \begin{bmatrix} 1.07825 & 0.078258 \\ 0.07825 & 1.07825 \end{bmatrix}$$

step-2

$P_1 = -H_1 g_1$

$H_1 = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$

$g_1 = [t, t]$ $t = e^{-2} = 0.135335$

$$H_1 g_1 = (a+b)t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$a = 1.07825$ $b = 0.07825$ $a+b = 1.15651$

$$H_1 g_1 = \cancel{1.15651} \; 0.15651 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$P_1 = \begin{bmatrix} -0.156517693 \\ -0.156517693 \end{bmatrix}$$

$x_2 = x_1 + \alpha_1 P_1 = (-1.1565, -1.1565)$

3)

$$g_2 = \nabla f(x_2) = \begin{bmatrix} 0.098960 \\ 0.098960 \end{bmatrix}$$

$\cancel{s_1 = x_2 - x_1 = \begin{bmatrix} 0.098960 \\ 0.098960 \end{bmatrix}}$ $s_1 = \begin{bmatrix} -0.1565176 \\ -0.1565176 \end{bmatrix}$

$$y_1 = g_2 - g_1 = \begin{bmatrix} -0.036374 \\ -0.036379 \end{bmatrix}$$

$y_1^T s_1 = 0.011385252$

$$\rho_1 = \frac{1}{y_1^T s_1} = 07.82291$$

BFGS

$$H_2^{BFGS} \simeq \begin{bmatrix} 2.65145 & 1.65145 \\ 1.651453 & 2.65145 \end{bmatrix}.$$

DFP

$$H_2^{DFP} \approx \begin{bmatrix} 2.651953 & 1.65195 \\ 1.65195 & 2.65195 \end{bmatrix}$$

True Hessian:-

$$\nabla^2 f(x_2) = \begin{bmatrix} 5+2 & 5-2 \\ 5-2 & 5+2 \end{bmatrix} \approx \begin{bmatrix} 2.098960 & -1.9010395 \\ -1.901039 & 2.09896 \end{bmatrix}$$

$$(\nabla^2 f)^{-1}(x_2) \approx \begin{bmatrix} 2.6512 & 2.40126 \\ 2.4012 & 2.65126 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} 2.6514 & 1.65145 \\ 1.6514 & 2.65145 \end{bmatrix}$$

$$\Delta = H_2 - (\nabla^2 f)^{-1}(x_2) \approx \begin{bmatrix} 0.000191 & -0.7498 \\ -0.74980 & 0.000191 \end{bmatrix}$$

$\therefore$ Both BFGS and DFP are equally close to Hessian

Q-4

$$f(x,y) = \sin(x+y) + (x-y)^2 + x^2 + y^2 \quad (x_0, y_0) = (1,1)$$

$$\nabla f(x,y) = \begin{bmatrix} \cos(x+y) + 2(x-y) + 2x \\ \cos(x+y) - 2(x-y) + 2y \end{bmatrix} = \begin{bmatrix} \cos(x+y) + 4x - 2y \\ \cos(x+y) - 2x + 4y \end{bmatrix}$$

$$\nabla^2 f(x,y) = \begin{bmatrix} -\sin(x+y) + 4 & -\sin(x+y) - 2 \\ -\sin(x+y) - 2 & -\sin(x+y) + 4 \end{bmatrix}$$

step-1 $\alpha_0 = 1$

$$g_0 = \begin{bmatrix} \cos 2 + 4 \cdot 1 - 2 \cdot 1 \\ \cos 2 - 2 \cdot 1 + 4 \cdot 1 \end{bmatrix} = \begin{bmatrix} 1.583 \\ 1.583 \end{bmatrix}$$

$$p_0 = -H_0 g_0 = -g_0 = \begin{bmatrix} -1.583 \\ -1.583 \end{bmatrix}$$

$$x_1 = x_0 + \alpha p_0 = \begin{bmatrix} -0.583 \\ -0.583 \end{bmatrix}$$

$$g_1 = \nabla f(x_1) = \begin{bmatrix} -0.775 \\ -0.775 \end{bmatrix}$$

$$s_0 = x_1 - x_0 = \begin{bmatrix} -1.583 \\ -1.583 \end{bmatrix}, \quad y_0 = g_1 - g_0 = \begin{bmatrix} -2.359 \\ -2.359 \end{bmatrix}$$

$$y_0^T s_0 = 7.4735 \qquad \rho_0 = \frac{1}{y_0^T s_0} = 0.13380$$

BFGS

$$H_1 = (I - \rho_0 s_0 y_0^T) H_0 (I - \rho_0 y_0 s_0^T) + \rho_0 s_0 s_0^T$$

$$H_1^{BFGS} \approx \begin{bmatrix} 0.8356 & -0.1643 \\ -0.1643 & 0.8356 \end{bmatrix}$$

DFP:

$$H_1^{DFP} = I + \frac{s_0 s_0^T}{s_0^T y_0} - \frac{y_0 y_0^T}{y_0^T y_0}$$

$$H_1^{DFP} \approx \begin{bmatrix} 0.8356 & -0.1643 \\ -0.1643 & 0.8356 \end{bmatrix}$$

step -2: (α_1=)

$$p_1 = -H_1 g_1 \approx \begin{bmatrix} 0.5205 \\ 0.5205 \end{bmatrix}$$

$$x_2 = x_1 + \alpha_1 p_1 = \begin{bmatrix} -0.0632 \\ -0.0632 \end{bmatrix}$$

$$g_2 = \nabla f(x_2)$$

$$g_2 = \begin{bmatrix} 0.8654 \\ 0.8654 \end{bmatrix}$$

$$s_1 = x_2 - x_1 = \begin{bmatrix} 0.5205 \\ 0.5205 \end{bmatrix} \qquad y_1 = g_2 - g_1 = \begin{bmatrix} 1.6908 \\ 1.6908 \end{bmatrix}$$

$$y_1^T s_1 = 1.7084 \qquad \rho_1 = \frac{1}{y_1^T s_1} = 0.5853$$

Again s_1 & y_1 are colinear so BFGS and DFP produce sam H_2

$$H_2 = (I - \rho_1 s_1 y_1^T) H_1 (I - \rho_1 y_1 s_1^T) + \rho_1 s_1 s_1^T$$

$$H_2 = \begin{bmatrix} 0.65862 & -0.39137 \\ -0.39137 & 0.65862 \end{bmatrix}$$

~~3~~

Hessian at x_2

$$\nabla^2 f(x_2) = \begin{bmatrix} 9.12 & -1.87 \\ -1.87 & 9.12 \end{bmatrix}$$

$$\left(\nabla^2 f(x_2)\right)^{-1} = \begin{bmatrix} 0.3053 & 0.1386 \\ 0.1386 & 0.3053 \end{bmatrix}$$

$$\Delta = H_2 - (\nabla^2 f)^{-1} = \begin{bmatrix} 0.35 & -0.47 \\ -0.47 & 0.35 \end{bmatrix}$$

$\therefore$ Both BFGS and DFP produce idential inverse Hessian approximations.