SOLUTIONS OF HOMEWORK - 3

1) a)

$$A = \begin{bmatrix} -1 & 0.45 & -0.4 \\ 0.45 & -1 & 0.45 \\ -0.4 & 0.45 & -1 \end{bmatrix}$$

$R_1 = [a_{11} - r_1 \, , \; a_{11} + r_1]$ $\qquad r_1 = 0.45 + 0.4 = 0.85$

$R_2 = [a_{22} - r_2 \, , \; a_{22} + r_2]$ $\qquad r_2 = 0.45 + 0.45 = 0.9$

$R_3 = [a_{33} - r_3 \, , \; a_{33} + r_3]$ $\qquad r_3 = 0.4 + 0.45 = 0.85$

$\Rightarrow \quad R_1 = [-1.85 \, , \; -0.15]$

$R_2 = [-1.9 \, , \; -0.1]$

$R_3 = [-1.85 \, , \; -0.15] \qquad \Rightarrow \quad -1.9 \leq \lambda_i \leq -0.1$

b) By part a), we know that all of the eigenvalues of this matrix are between -1.9 and -0.1, so we can be able to say that all eigenvalues are negative which means X is negative definite. i.e $\eta = -1.9$

c) the smallest eigenvalue of this matrix is $\lambda_{min} \cong 1.87$. again by part a) the lower bound for this matrix holds.

d) $f := \mathbb{R}^3 \to \mathbb{R}$

Newton Direction: $\nabla^2 f(x_k) \, p_n = -\nabla f(x_k)$

$$A p_n = [1 \; 1 \; 1]^T \quad \Rightarrow \quad A \begin{bmatrix} p_n^1 \\ p_n^2 \\ p_n^3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\Rightarrow \quad p_n = \begin{bmatrix} -1.457 \\ -2.312 \\ -1.457 \end{bmatrix} \qquad \nabla f(x_k)^T . p_n = [-1 \, -1 \, -1] \, p_n = 5.226 > 0$$

Therefore, p_n is not descent direction.

Modified Search direction: $\nabla^2 f(x_k) + (-\eta + 0.1) I = \nabla^2 \hat{f}(x_k)$ ②

$$\nabla^2 f(x_k) + 2I = \hat{\nabla} f(x_k)$$

Then $\nabla^2 \hat{f}(x_k) = \begin{bmatrix} 1 & 0.45 & -0.4 \\ 0.45 & 1 & 0.45 \\ -0.4 & 0.45 & 1 \end{bmatrix} \cdot P_m = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$$P_m = \begin{bmatrix} 2.821 \\ -1.538 \\ 2.821 \end{bmatrix}$$

So $\nabla f(x_k)^T P_m = [-1\ -1\ -1]\ P_m = -4.104 < 0$

Hence P_m is a descent direction.

2) a) $A = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$

$p(\lambda) = \det(A - \lambda I)$

$0 = \lambda^2 - 4 \Rightarrow \lambda_1 = 2 \quad \lambda_2 = -2 \Rightarrow V_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

$V_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

$$\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = A$$

b) $\hat{\lambda}_2 = |\lambda_2| = 2$

$$\hat{A} = V \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} V^T = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

c) $\nabla f(\bar{x}) = [1\ -1]^T$

$$\underbrace{-\hat{A}^{-1} \cdot \nabla f(\bar{x}) = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}}_{\text{Newton Modified}}$$

Search Direction

$\nabla f(\bar{x})^T P_m = [1\ -1] \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix} = -1 < 0$ so P_m is the descent direction.

3) a) Apply DFP to $f(x_1,x_2) = x_1^2 + x_2^2$ starting with ③

$x_k = (1,-1)$ and $H_k = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

We use exact line search.

Search Direction

$$\nabla f(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \Rightarrow \nabla f(x_k) = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$$p_k = -H_k \nabla f(x_k)$$

$$= \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

Step Lengths

If f is a quadratic fnc., then exact step length is given by

$$\alpha_k = \frac{-\phi'(0)}{p_k^T \nabla^2 f(x_k) p_k}$$

$$= \frac{-\nabla f(x_k)^T p_k}{p_k^T \nabla^2 f(x_k) p_k}, \quad \nabla^2 f(x_k) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\Rightarrow \alpha_k = \frac{-\begin{bmatrix} 2 & -2 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix}}{\begin{bmatrix} -4 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix}} = \frac{3}{10}$$

Consequently, $x_{k+1} = x_k + \alpha_k p_k = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{3}{10} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -0.2 \\ -0.4 \end{bmatrix}$

$$\nabla f(x_{k+1}) = \nabla f\left(\begin{bmatrix} -0.2 \\ -0.4 \end{bmatrix} \right) = \begin{bmatrix} -0.4 \\ -0.8 \end{bmatrix}$$

4

$$y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$

$$= \begin{bmatrix} -0.4 \\ -0.8 \end{bmatrix} - \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix}$$

$$s_k = x_{k+1} - x_k = \alpha_k p_k = \frac{3}{10} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -1.2 \\ 0.6 \end{bmatrix}$$

$$\Rightarrow s_k^T = [-1.2 \quad 0.6]$$

$$H_k = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\boxed{\lambda_1 = 1, \; \lambda_2 = 2 > 0}$$

so $H_k \succ 0$

$$s_k^T y_k = [-1.2 \quad 0.6] \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix} = 3.6 > 0$$

Since $H_k \succ 0$ and $s_k^T y_k > 0$ then $H_{k+1} \succ 0$.

$$H_{k+1} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} + \frac{s_k s_k^T}{s_k^T y_k} - \frac{(H_k y_k)(H_k y_k)^T}{y_k^T H_k y_k}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\begin{bmatrix} -1.2 \\ 0.6 \end{bmatrix} [-1.2 \quad 0.6]}{\underbrace{[-1.2 \quad 0.6] \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix}}_{3.6}} - \frac{\overbrace{\left(\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix}\right)}^{\begin{bmatrix} -4.8 \\ 1.2 \end{bmatrix}} \overbrace{\left(\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix}\right)^T}^{[-4.8 \quad 1.2]}}{\underbrace{[-2.4 \quad 1.2] \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix}}_{12.96}}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.4 & -0.2 \\ -0.2 & 0.1 \end{bmatrix} - \begin{bmatrix} 23.04/12.96 & -5.76/12.96 \\ -5.76/12.96 & 1.44/12.96 \end{bmatrix}$$

$$= \begin{bmatrix} 2.4 & -0.2 \\ -0.2 & 1.1 \end{bmatrix} - \begin{bmatrix} 1.\overline{7} & -0.\overline{4} \\ -0.\overline{4} & 0.\overline{1} \end{bmatrix}$$

$$\simeq \begin{bmatrix} 0.63 & 0.24 \\ 0.24 & 0.99 \end{bmatrix}$$

b) Similar to part a

$$H_{k+1} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} + \left(1 + \frac{y_k^T H_k y_k}{s_k^T y_k}\right) \frac{s_k s_k^T}{s_k^T y_k} - \frac{s_k y_k^T H_k + H_k y_k s_k^T}{s_k^T y_k}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} + \left(1 + \frac{\begin{bmatrix} -2.4 & 1.2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix}}{\begin{bmatrix} -1.2 & 0.6 \end{bmatrix} \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix}}\right) \frac{\begin{bmatrix} -1.2 \\ 0.6 \end{bmatrix} \begin{bmatrix} -1.2 & 0.6 \end{bmatrix}}{\begin{bmatrix} -1.2 & 0.6 \end{bmatrix} \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix}}$$

$$- \frac{\begin{bmatrix} 1.2 \\ 0.6 \end{bmatrix} \begin{bmatrix} -2.4 & 1.2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix} \begin{bmatrix} 1.2 & 0.6 \end{bmatrix}}{\begin{bmatrix} -1.2 & 0.6 \end{bmatrix} \begin{bmatrix} -2.4 \\ 1.2 \end{bmatrix}}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1.84 & -0.92 \\ -0.92 & 0.46 \end{bmatrix} - \begin{bmatrix} 3.2 & -1.2 \\ -1.2 & 0.4 \end{bmatrix} = \begin{bmatrix} 0.64 & 0.28 \\ 0.28 & 1.06 \end{bmatrix}$$

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

$$= \left(B_k - \frac{y_k s_k^T B_k}{y_k^T s_k}\right)\left(I - \frac{s_k y_k^T}{y_k^T s_k}\right) + \frac{y_k y_k^T}{y_k^T s_k}$$

If $B_{k+1} \succ 0$, then for any nonzero z $\quad z^T B_{k+1} z > 0$

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

$$= z^T B_k z - z^T \frac{y_k s_k^T B_k z}{y_k^T s_k} - z^T \frac{B_k s_k y_k^T}{y_k^T s_k} z + z^T \frac{y_k s_k^T B_k s_k y_k^T}{y_k^T s_k y_k^T s_k} z$$

$$+ z^T \frac{y_k y_k^T}{y_k^T s_k} z.$$

if $\rho = \dfrac{1}{y_k^T s_k}$ $\qquad w = \dfrac{s_k y_k^T z}{y_k^T s_k}$

$$= z^T B_k z - w^T B_k z - z^T B_k w + w^T B_k w + \rho (z^T y_k)(y_k^T z)$$

$+ \rho \| z^T y_k \|^2$ since

if $z^T y_k \neq 0$ then $B_{k+1} \succ 0$ when $\rho \|z^T y_k\|^2 > 0$

if $z^T y_k = 0$ then by positive definiteness of B_k, $B_{k+1} \succ 0$

all eigenvalues of B_{k+1} is positive. All of the eigen values of

$H_{k+1} = B_{k+1}^{-1}$ are also positive. so $H_{k+1} \succ 0$.

5) Assume that α_k is min. of $\phi(\alpha)$.

$$\phi(\alpha) = f(x_k + \alpha p_k)$$

$$\phi'(\alpha_k) = f'(x_k + \alpha_k p_k) . p_k$$

$$= \nabla f(x_k + \alpha_k p_k)^T p_k = 0$$

if p_k is a descent direction, then $p_k^T \nabla f(x_k) < 0$.

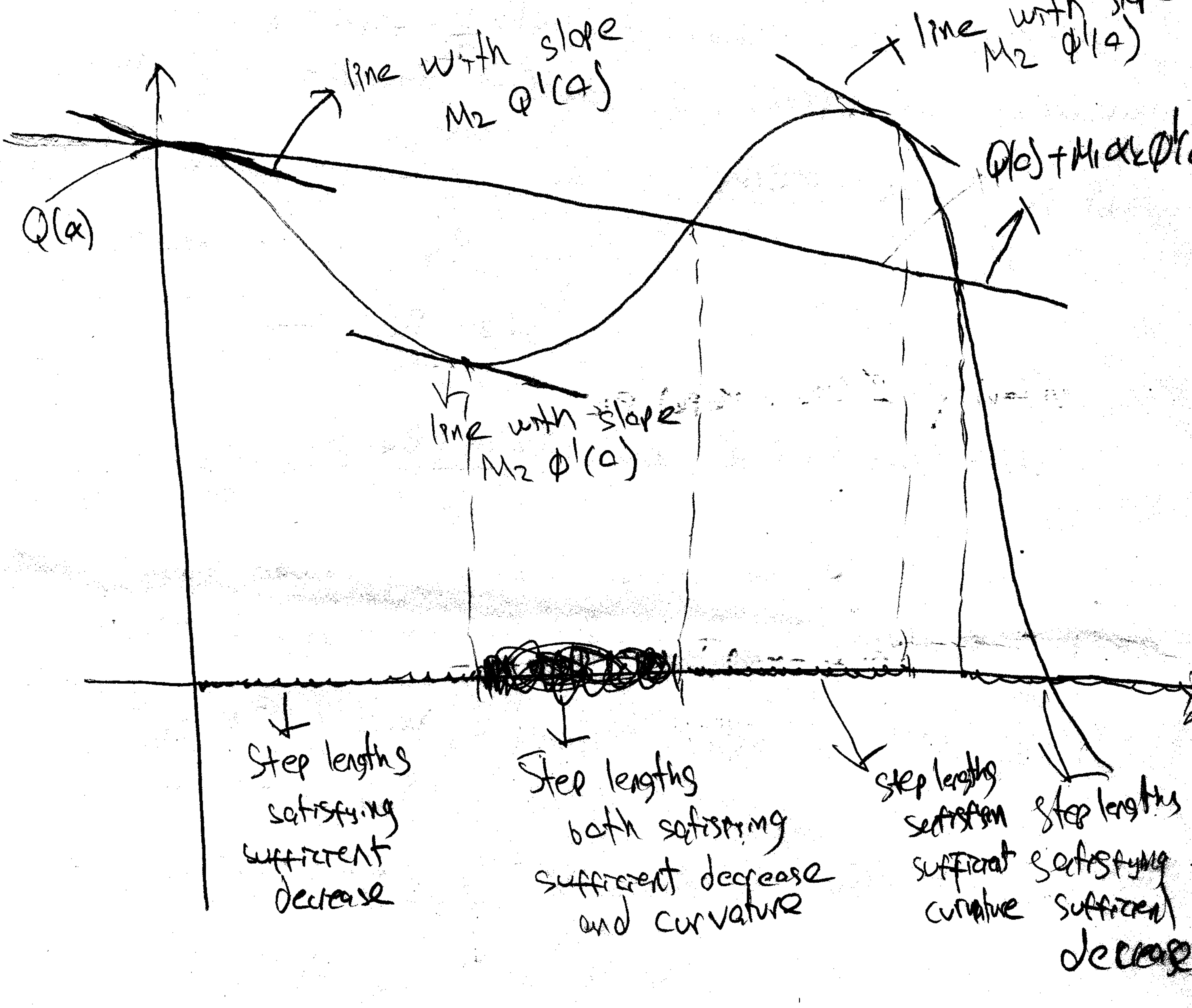
$$s_k^T y_k = (x_{k+1} - x_k)^T (\nabla f(x_{k+1}) - \nabla f(x_k))$$

$$= (x_k + \alpha_k p_k - x_k)^T (\nabla f(x_{k+1}) - \nabla f(x_k))$$

$$= \underbrace{\alpha_k p_k^T \nabla f(x_{k+1})}_{\text{This is equal to zero since } \nabla f(x_{k+1})^T p_k = 0} - \alpha \underbrace{p_k^T \nabla f(x_k)}_{\text{this is smaller than zero}} > 0$$

so $s_k^T y_k > 0$.

7)


line with slope
$M_2\ \phi'(0)$
$\phi(\alpha)$
line with slope
$M_2\ \phi'(0)$
line with slope
$M_2\ \phi'(0)$
$\phi(0)+M_1\alpha\phi'(0)$
Step lengths satisfying sufficient decrease
Step lengths both satisfying sufficient decrease and curvature
step lengths satisfying sufficient curvature
Step lengths satisfying sufficient decrease

8) a) $p_k = -H_k \nabla f(x_k)$

then $|\cos\theta_k| = \dfrac{|\nabla f(x_k)^T p_k|}{\|\nabla f(x_k)\| \|p_k\|} = \dfrac{|-\nabla f(x_k)^T H_k \nabla f(x_k)|}{\|\nabla f(x_k)\| \|H_k \nabla f(x_k)\|} \geq \dfrac{\lambda_{min} H}{\lambda_{max} H}$

$= \frac{1}{2} + 2^{-k}$

we find a lower bound $|\cos\theta_k| \geq \frac{1}{2}$ then by Zoutendijk's theorem $\lim \nabla f(x_k) = 0$, therefore convergence is garanteed.

b) $p_k = -B_k^{-1} \nabla f(x_k)$ so as $k \to \infty$ $B_k \to \begin{bmatrix} 3 & \\ & 0 \end{bmatrix}$

$\Rightarrow B_k \to$ singularity

and $B_k^{-1} \to \begin{bmatrix} \infty & \\ & 1/3 \end{bmatrix}$

$$\lim |\cos\theta_k| \geq \lim \frac{\lambda_{min}(B_k^{-1})}{\lambda_{max}(B_k^{-1})} = \frac{1/3}{\infty} = 0$$

so we can not be able to find a lower bound for $|\cos\theta_k|$.

c) by (b) we have $|\cos\theta_k| \geq \dfrac{\lambda_{min}(\nabla^2 f(x_k)^{-1})}{\lambda_{max}(\nabla^2 \tilde{f}(x_k)^{-1})}$

Now $\nabla^2 f(x_k) = A = V^T \Lambda V$, $V^T V = I$

$\Rightarrow V^{-1} = V^T$, $(V^T)^{-1} = V \Rightarrow \nabla^2 f(x_k)^{-1} = A^{-1}$

$= V^{-1} \Lambda^{-1} V^{T^{-1}} = V^T \tilde{\Lambda}^{-1} V = V^T \begin{bmatrix} \frac{1}{|\lambda_1| + 0.1} & \\ & \frac{1}{|\lambda_2| + 0.1} \end{bmatrix}$

then $|\cos\theta_k| \geq \dfrac{\lambda_{min}(\nabla^2 f(x_k)^{-1})}{\lambda_{max}(\nabla^2 \tilde{f}(x_k)^{-1})} = \dfrac{1/|\lambda_1| + 0.1}{1/|\lambda_2| + 0.1} = c$

Hence By Zoutendijk' thm $\lim\limits_{k\to\infty} \nabla f(x_k) = 0$.

Thus convergence is garanteed.