SAARTHAK MARATHE
ME17B162

classmate
Date 17/9/20
Page 1

ME7223 : ASSIGNMENT 3

1) $x = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

$$x^T M x = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & -4 & 5 \\ 3 & 5 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 & -6 & 13 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$\therefore x^T M x = 5 - 12 = -7 < 0$

$\therefore$ given matrix is not positive definite.

We can define $x = A_j$ where $A_j[1][j] = 1$ and rest terms 0.
$\therefore$ with $x^T M x$ we get the $M[j][j]$ term as the answer and thus can extract any diagonal term of M matrix.
In this scenario,
If any of the $x^T M x < 0$ then ~~MQ~~ M is not positive definite.
$\therefore$ For M to be positive definite, none of $M[j][j]$ can be negative.

2) a) If $f(x)$ is continuous and doubly differentiable then H is symmetric as $\frac{\partial^2 f}{\partial x_i \cdot \partial x_j} = \frac{\partial^2 f}{\partial x_j \, \partial x_i}$

i) As 'H' has doubly differentiation terms, all $f(x)$ with degree 2 will have H independent of x. The terms should not have -ve integer powers.

eg. $f(x) = x_1^2 + x_2^2$

$\Rightarrow H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

ii) The functions with ~~xs~~ x terms having -ve integer powers or with functions having degree more than two and non- ~~int~~ integer powers.

eg. $f(x) = \sqrt{x_1 x_2}$

$$H = \begin{bmatrix} -\frac{1}{4}\sqrt{\frac{x_2}{x_1^3}} & \frac{1}{4\sqrt{x_1 x_2}} \\ \frac{1}{4\sqrt{x_1 x_2}} & -\frac{1}{4}\sqrt{\frac{x_1}{x_2^3}} \end{bmatrix}$$

$f(x) = x_1/x_2$

$$H = \begin{bmatrix} 0 & -1/x_2^2 \\ -1/x_2^2 & 2x_1/x_2^3 \end{bmatrix}$$

d) $\bar{y} = d\bar{x}$

$$H d\bar{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i \Rightarrow H d\bar{x} = \begin{bmatrix} f_{x_1^2} & f_{x_1 x_2} \\ f_{x_1 x_2} & f_{x_2^2} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix}$$

$$\Rightarrow H d\bar{x} = \begin{bmatrix} f_{x_1^2} \cdot dx_1 + f_{x_1 x_2} \cdot dx_2 \\ f_{x_1 x_2} \cdot dx_1 + f_{x_2^2} \cdot dx_2 \end{bmatrix} \qquad f_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

This denotes the small change in the gradient of f ($\nabla^2 f$).
Physically it is the displacement in direction of rate of change of the function.
It is the error of quadratic approx. of f (linear error).

e) $\bar{y} = d\bar{x}$

$$d\bar{x}^T H d\bar{x} = \begin{bmatrix} dx_1 & dx_2 \end{bmatrix} \begin{bmatrix} f_{x_1^2} & f_{x_1 x_2} \\ f_{x_1 x_2} & f_{x_2^2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} f_{x_1^2} dx_1 + f_{x_1 x_2} dx_2 & f_{x_1 x_2} dx_1 + f_{x_2^2} dx_2 \end{bmatrix} \begin{bmatrix} dx_1 & dx_2 \end{bmatrix}^T$$

$$\therefore d\bar{x}^T H d\bar{x} = f_{x_1^2} dx_1^2 + 2 f_{x_1 x_2} dx_1 dx_2 + f_{x_2^2} dx_2^2$$

This is essentially: $\sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i dx_j$

This is quadratic error in quadratic approx of f. Physically, it can ~~be~~ be used to find infinitesimal ~~change in~~ change ($d^2 f$) of the function.

3) $\min f = (x_1-2)^2 + (x_2-1)^2$

$g_1 : x_1 + x_2 - 2 \le 0$

$g_2 : x_1^2 - x_2 \le 0$

Using Kuhn-Tucker conditions :

$\nabla f + \lambda \nabla g = 0$

$\lambda \ge 0$

$g \le 0$

$\therefore \nabla f = \begin{bmatrix} 2(x_1-2) \\ 2(x_2-1) \end{bmatrix}, \nabla g_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \nabla g_2 = \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix}$

$\therefore \begin{bmatrix} 2(x_1-2) \\ 2(x_2-1) \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix} = \nabla f + \lambda \nabla g$

a) for $x_1 = \begin{Bmatrix} 1.5 \\ 0.5 \end{Bmatrix}$

$\nabla f + \lambda \nabla g = \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 0$

$\Rightarrow -1 + \lambda_1 + 3\lambda_2 = 0, \ -1 + \lambda_1 - \lambda_2 = 0 \Rightarrow \lambda_2 = \lambda_1 - 1$

$\Rightarrow -1 + \lambda_1 + 3\lambda_1 - 3 = 0 \Rightarrow \lambda_1 = 1 \Rightarrow \lambda_2 = 0$

As $\lambda_1 \ge 0, \lambda_2 = 0$: this gives a ~~local minimum~~ regular point as $g_1 = 0, \ g_2 \le 0$ not a local minima

b) $x_2 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$

$\nabla f + \lambda \nabla g = \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 0$

$\Rightarrow -2 + \lambda_1 + 2\lambda_2 = 0, \ \lambda_1 - \lambda_2 = 0 \Rightarrow \lambda_2 = \lambda_1$

$\Rightarrow 3\lambda_1 = 2 \Rightarrow \lambda_1 = \frac{2}{3}, \lambda_2 = \frac{2}{3}$

As $\lambda_1, \lambda_2 \ge 0$: $\lambda \ge 0$ is satisfied

$g_1(x_2) = 1 + 1 - 2 \le 0$ ✓

$g_2(x_2) = 1 - 1 \le 0$ ✓

$\therefore x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a VALID local minima.

Scanned with CamScanner

c) $x_3 = \begin{Bmatrix} 2 \\ 0 \end{Bmatrix}$

$$\nabla f + \lambda \nabla g = \begin{bmatrix} 0 \\ -2 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 4 \\ -1 \end{bmatrix} = 0$$

$\Rightarrow \lambda_1 + 4\lambda_2 = 0, \quad -2 + \lambda_1 - \lambda_2 = 0$

$\Rightarrow \lambda_1 = -4\lambda_2 \Rightarrow -2 - 5\lambda_2 = 0 \Rightarrow \lambda_2 = \frac{-2}{5} \Rightarrow \lambda_1 = \frac{8}{5}$

As $\lambda_2 \neq 0$, the point $x_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ is not a local minima

4) max $f(x_1, x_2) = 2x_1 + \beta x_2$

$g_1 : x_1^2 + x_2^2 - 5 \leq 0$

$g_2 : x_1 - x_2 - 2 \leq 0$

GRAPHS ON THE NEXT PAGE

KKT for maximization :

$\nabla f + \lambda \nabla g = 0$

$\lambda \leq 0$

$g \leq 0$.

$$\nabla f = \begin{bmatrix} 2 \\ \beta \end{bmatrix}, \quad \nabla g_1 = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \nabla g_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$\Rightarrow \begin{bmatrix} 2 \\ \beta \end{bmatrix} + \lambda_1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0$ for $x_1^* = 1, x_2^* = 2$.

$\Rightarrow 2 + 2\lambda_1 + \lambda_2 = 0, \quad \beta + 4\lambda_1 - \lambda_2 = 0$

$\Rightarrow \lambda_2 = \beta + 4\lambda_1$

$\Rightarrow 2 + 2\lambda_1 + \beta + 4\lambda_1 = 0 \Rightarrow \lambda_1 = -\frac{(2+\beta)}{6}$

$\Rightarrow \lambda_2 = \frac{6\beta - 8 - 4\beta}{6} = \frac{2\beta - 8}{6} = \frac{\beta - 4}{3}$

$g_1(1,2) = 1 + 4 - 5 = 0 \leftarrow$ active $\Rightarrow \lambda_1 < 0$

$g_2(1,2) = 1 - 2 - 2 < 0 \leftarrow$ inactive $\Rightarrow \lambda_2 = 0$

As $\lambda_2 = 0 \Rightarrow \beta = 4$, $\lambda_1 < 0 \Rightarrow \beta \geq -2 \Rightarrow \boxed{\beta = 4}$

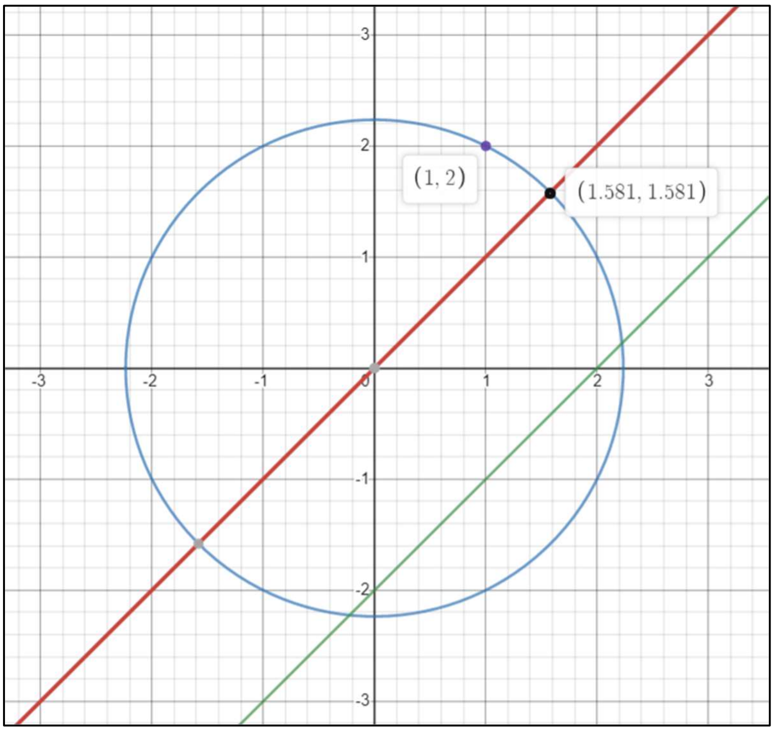
$\Rightarrow \lambda_1 = \frac{-6}{6} = -1 < 0$ satisfied

Ans. $\beta = 4$

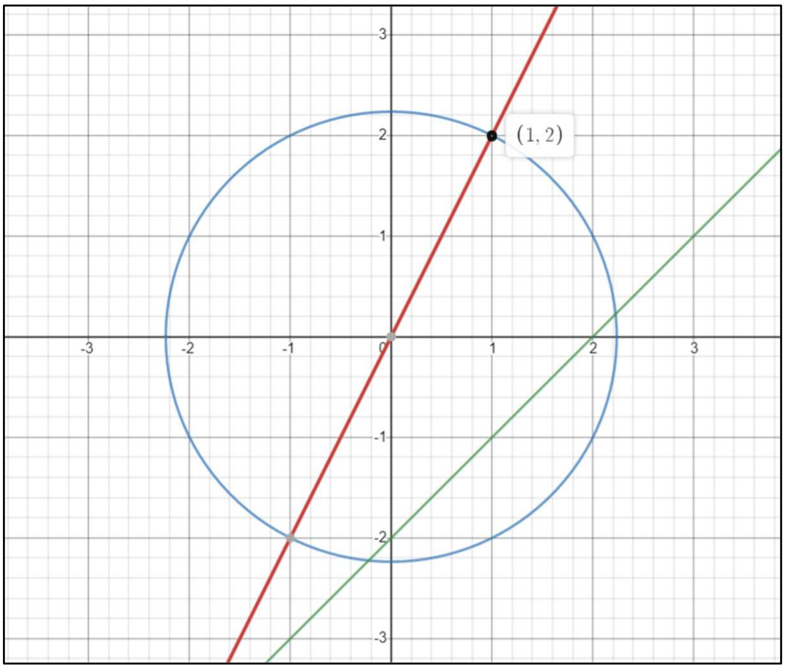
SAARTHAK SANDIP MARATHE | ME17B162

For each of the graphs:

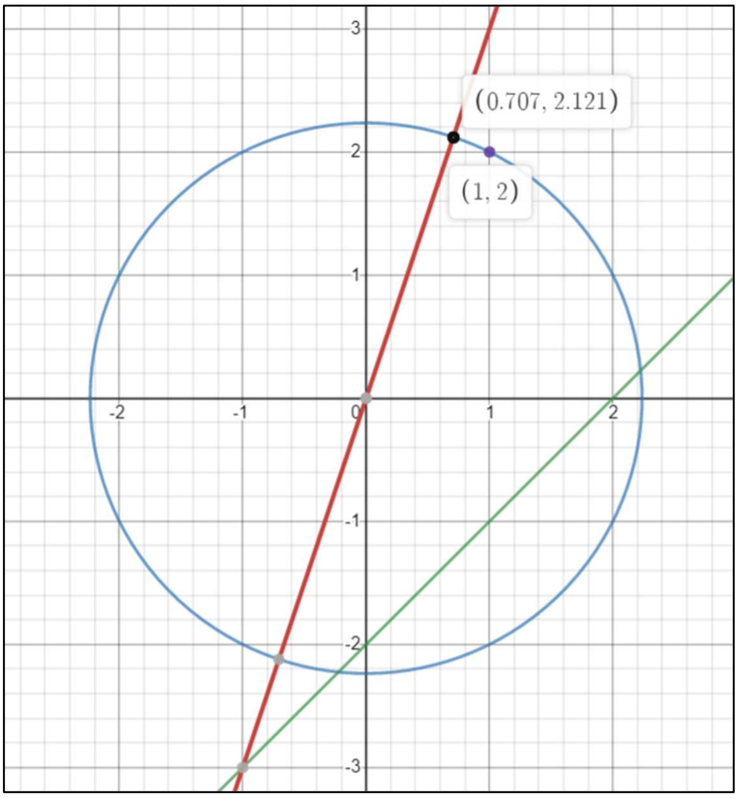
Red line = f(X)
Blue circle = g1(X)
Green line = g2(X)



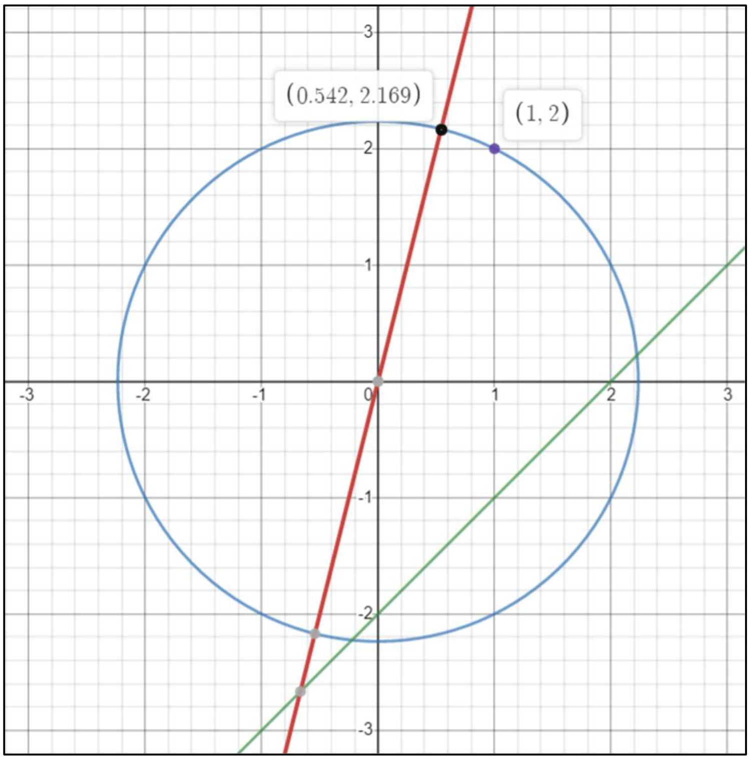
Beta = 2



Beta = 4



Beta = 6



Beta = 8

As we can see, only for beta = 4 the maximum point for f(X) coincide with the given (1,2) point. For beta<4, the value shifts to the right to that of beta=4 and it shifts to the left for beta>4 values.

Thus, beta = 4 is our optimal solution

5) min $f(x) = (x_1-1)^2 + (x_2-5)^2$

$g_1: -x_1^2 + x_2 - 4 \leq 0$

$g_2: -(x_1-2)^2 + x_2 - 3 \leq 0$

$$\nabla f = \begin{bmatrix} 2(x_1-1) \\ 2(x_2-5) \end{bmatrix}, \quad \nabla g_1 = \begin{bmatrix} -2x_1 \\ 1 \end{bmatrix}, \quad \nabla g_2 = \begin{bmatrix} -2(x_1-2) \\ 1 \end{bmatrix}, \quad S = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix}$$

a) $x_1 = [-1 \;\; 5]^T$

$g_1(x_1) = -1 + 5 - 4 = 0$ ← active

$g_2(x_1) = -9 + 5 - 3 < 0$ ← inactive

Usability ~~region~~ condition :

$S^T \nabla f < 0$

$$\Rightarrow [S_1 \;\; S_2] \begin{bmatrix} -4 \\ 0 \end{bmatrix} < 0 \Rightarrow -4S_1 < 0 \Rightarrow S_1 > 0$$

Usable direction : $\boxed{S_1 > 0, \; S_2 \in \mathbb{R}}$

Feasible condition :

$S^T \nabla g_1(x_1) < 0, \quad S^T \nabla g_2(x_1) < 0$

ACTIVE

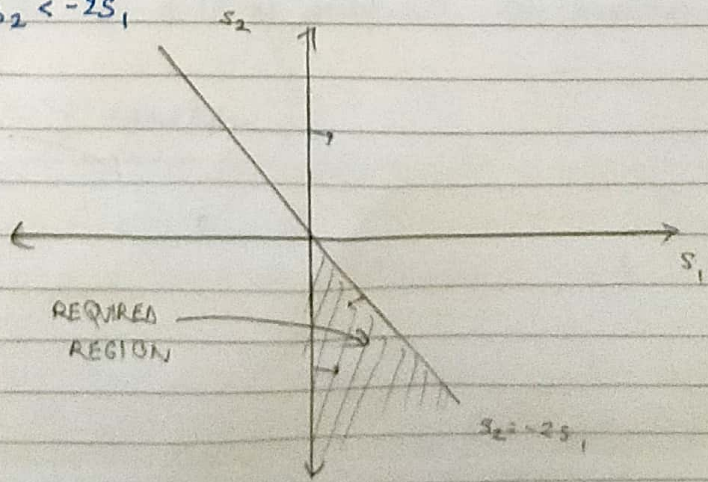
$$\Rightarrow [S_1 \;\; S_2] \begin{bmatrix} 2 \\ 1 \end{bmatrix} < 0 \Rightarrow 2S_1 + S_2 < 0 \Rightarrow S_2 < -2S_1$$

$$\Rightarrow [S_1 \;\; S_2] \begin{bmatrix} 6 \\ 1 \end{bmatrix} < 0 \Rightarrow 6S_1 + S_2 < 0 \Rightarrow S_2 < -6S_1$$

INACTIVE

~~So~~ the usable and feasible regions are ~~an~~ intersection

$S_1 > 0, \; S_2 < -2S_1$



b) $x_2 = [2\ 3]^T$

Usability condition :

$s^T \nabla f(x_2) < 0$

$\Rightarrow \begin{bmatrix} s_1 & s_2 \end{bmatrix} \begin{bmatrix} 2 \\ -4 \end{bmatrix} < 0 \Rightarrow 2s_1 - 4s_2 < 0 \Rightarrow s_1 - 2s_2 < 0$

$g_1(x_2) = -4 + 3 - 4 > 0$ ← Inactive

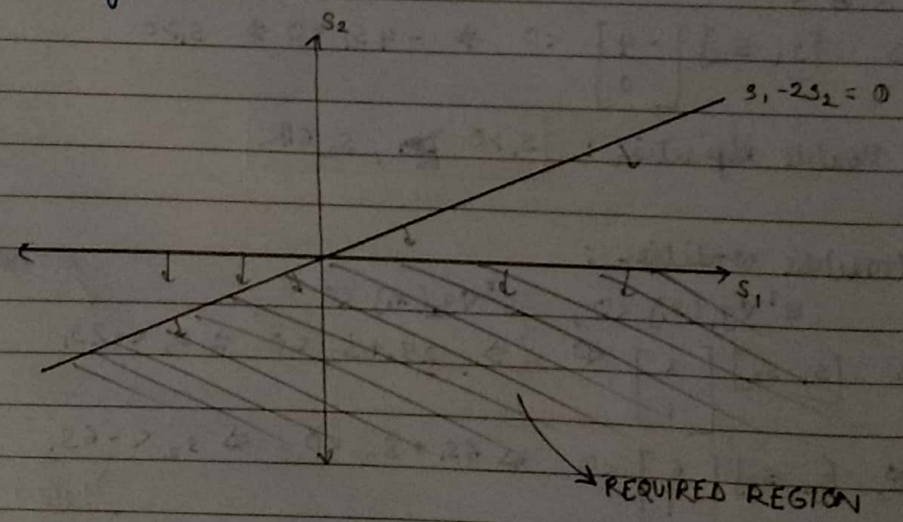
$g_2(x_2) = 3 - 3 = 0$ ← active.

Feasibility condition :

$\therefore s^T \nabla g_2(x_2) < 0$

$\Rightarrow \begin{bmatrix} s_1 & s_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} < 0 \Rightarrow s_2 < 0$

Required region :



6) a) $f(x_1, x_2) = -4x_1 + x_1^2 - 2x_1x_2 + 2x_2^2$

$$\nabla f = \begin{bmatrix} -4 - 2x_2 + 2x_1 \\ -2x_1 + 4x_2 \end{bmatrix} \Rightarrow H = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$$

Finding eigenvalue:

$$\begin{vmatrix} 2-\lambda & -2 \\ -2 & 4-\lambda \end{vmatrix} = 0 \Rightarrow (2-\lambda)(4-\lambda) - 4 = 0$$

$\therefore \quad 8 - 4\lambda - 2\lambda + \lambda^2 - 4 \Rightarrow \lambda^2 - 6\lambda - 4 = 0$

$\therefore \quad \lambda = \dfrac{6 \pm \sqrt{36-16}}{2} = \dfrac{6 \pm 2\sqrt{5}}{2} \Rightarrow \lambda = 3 \pm \sqrt{5}$

As $\lambda_1, \lambda_2 > 0 \Rightarrow f(x_1, x_2)$ is convex.

b) $f(x_1, x_2) = 2x_1 + 3x_2 - x_1^3 - 2x_2^2$

$$H = \begin{bmatrix} -6x_1 & 0 \\ 0 & -4 \end{bmatrix}$$

As its a diagonal matrix, one of eigenvalue $= -4$.

As $-4 < 0 \Rightarrow f(x_1, x_2)$ is not convex.

7) min $f(x_1, x_2) = 9x_1^2 - 18x_1x_2 + 13x_1 - 4$.

$g: \; x_1^2 + x_2^2 + 2x_1 \geq 16$

$$H_f = \begin{bmatrix} 18 & -18 \\ -18 & 0 \end{bmatrix} = 18 \begin{bmatrix} +1 & -1 \\ -1 & 0 \end{bmatrix}$$

Eigenvalues: $18 \begin{vmatrix} 1-\lambda & -1 \\ -1 & -\lambda \end{vmatrix} = 0 \Rightarrow -\lambda + \lambda^2 - 1 = 0 \Rightarrow \lambda = \dfrac{1 \pm \sqrt{5}}{2}$

As $\dfrac{1+\sqrt{5}}{2} > 0$, $\dfrac{1-\sqrt{5}}{2} < 0 \Rightarrow H$ is indefinite. The problem is not convex.

We need not check 'g' function.

8) $f(x) = \frac{1}{2}x^T A x + bx + c$

Let $x_0 = [x_1\ x_2\ x_3\ \dots\ x_n]^T$

We can write $\frac{1}{2}x^T A x = \frac{1}{2}[c_{11}x_1^2 + c_{22}x_2^2 + c_{33}x_3^2 \dots + c_{nn}x_n^2] + \frac{1}{2}[\sum_{i \neq j}\sum c_{ij}x_i x_j]$

where $c_1, c_2, \dots, c_n$ are columns of A. $A = \begin{bmatrix} | & | & | & & | \\ c_1 & c_2 & c_3 & \dots & c_n \\ | & | & | & & | \end{bmatrix}$

$\therefore f'(x) = (c_{11}x_1 + c_{22}x_2 \dots + c_{nn}x_n) + b + \frac{1}{2}(\sum_{i \neq j}\sum c_{ij}x_i)$

$\therefore f'(x) = Ax + b$

$\Rightarrow f''(x) = \sum c_{ij} = A$

As $A = \begin{bmatrix} 18\kappa & 2-2\kappa \\ 0 & 2 \end{bmatrix}$, $b = [\kappa^2\ 1]$, $c = \alpha$,

all being real values, we can claim that $f(x)$ is doubly differentiable.

a) Given above conditions of doubly differentiable f, we can have $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$

Thus, H of $f(x)$ is symmetric.

b) No, we could have the function with saddle points making it neither concave or convex.

c) Convexity is checked using the Hessian which is 2nd order differential so we can drop '$bx + c$' in $f(x)$ and concentrate on the quadratic terms.

$x^T A x = [x_1\ x_2]\begin{bmatrix} 18\kappa & 2-2\kappa \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 18\kappa x_1 \\ 2x_1 - 2\kappa x_1 + 2x_2 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$x^T A x = 18\kappa x_1^2 + 2x_1x_2 - 2\kappa x_1 x_2 + 2x_2^2$

i) Direction: $x_2 = (1-\kappa)x_1$

$\Rightarrow x^T A x = 18\kappa x_1^2 + 2(1-\kappa)^2 x_1^2 + 2(1-\kappa)^2 x_1^2$

$\therefore x^T A x = (18\kappa + 4 - 8\kappa + 4\kappa^2)x_1^2 = (4\kappa^2 + 10\kappa + 4)x_1^2$

For it to be convex :

$4\alpha^2 + 10\alpha + 4 > 0$ $\Rightarrow$ as H will be $= \begin{bmatrix} 4\alpha^2+10\alpha+4 & 0 \\ 0 & 0 \end{bmatrix}$

$\Rightarrow 2(2\alpha+1)(\alpha+2) > 0$

$\Rightarrow \alpha \in (-\infty, -2) \cup \left(\frac{-1}{2}, \infty\right)$

ii) Direction $x_1 = x_2$:

$x^T A x = 18\alpha x_1^2 + 2(1-\alpha)x_1^2 + 2x_1^2 = (18\alpha + 2 - 2\alpha + 2)x_1^2$

$\Rightarrow x^T A x = (16\alpha + 4)x_1^2$

similar to previous :

$16\alpha + 4 \leq 0$ ← condition for ~~convexity~~ concavity

$\Rightarrow \alpha \leq \frac{-1}{4}$ $\Rightarrow$ ~~$\alpha \in \left(-\frac{1}{4}, \infty\right)$~~

$\Rightarrow \alpha \in \left(-\infty, -\frac{1}{4}\right)$

$\Rightarrow$ The required region : $\underline{\alpha \in (-\infty, -2) \cup \left(-\frac{1}{2}, -\frac{1}{4}\right)}$

d) $x_2 = (1-\alpha)x_1$, along this the global ~~max~~ min occurs for $\alpha \in (-\infty, -2) \cup \left(\frac{-1}{2}, \infty\right)$

As $\alpha = -8$ is present in this range,

the local ~~max~~ min will be global minima .