Rumitkumar Patel ASU ID: 1219783279

Homework-2

1) $f(x_1, x_2) = 2x_1^2 - 4x_1x_2 + 1.5x_2^2 + x_2$

Now, gradient is given by,

$$g(x_1, x_2) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4x_1 - 4x_2 \\ -4x_1 + 3x_2 + 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Here, it is given that at stationary point gradient is zero.

So, to find stationary point, we have two equations and two variables.

$$4x_1 - 4x_2 = 0 \quad \text{—①}$$
$$-4x_1 + 3x_2 + 1 = 0 \quad \text{—②}$$

$$-x_2 = -1$$
$$\boxed{x_2 = 1}$$

$$4x_1 - 4(1) = 0$$
$$4x_1 = 4$$
$$\boxed{x_1 = 1}$$

$\therefore$ stationary point $x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Now, Hessian is given by,

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 4 & -4 \\ -4 & 3 \end{bmatrix}$$

Now, let's have λ as eigen values of Hessian,

to find eigen values

$$|H - \lambda I| = 0$$

$$\begin{vmatrix} 4-\lambda & -4 \\ -4 & 3-\lambda \end{vmatrix} = 0 \Rightarrow (4-\lambda)(3-\lambda) - 16 = 0.$$

$$\lambda^2 - 7\lambda - 4 = 0.$$

$$\Delta = b^2 - 4ac$$
$$= 49 - 4(1)(-4)$$
$$= 49 + 16$$
$$= 65$$

$\therefore$ Roots are $\lambda = \dfrac{-(-7) \pm \sqrt{65}}{2}$

So, $\lambda = \dfrac{7+\sqrt{65}}{2}, \dfrac{7-\sqrt{65}}{2}$

$\qquad\qquad \lambda_1 \qquad\qquad \lambda_2$

Here, eigenvalue λ_1 is positive and eigenvalue λ_2 is negative, so, we can conclude that the stationary point $(1,1)$ is a saddle point.

The taylor's expansion is given by,

$$f(x) = f(x_0) + \nabla x f\Big|_{x_0}^{T} (x - x_0) + \frac{1}{2} \frac{d^2 f}{dx^2}\Big|_{x_0} (x - x_0)^2$$

$$f(x,y) = f(1,1) + \nabla f\Big|_{(1,1)}^{T} (x - x^*) + \frac{1}{2} (x - x^*)^T H\Big|_{(1,1)} (x - x^*)$$

($\because$ Taylor expansion in matrix form when saddle point is taken as reference point of $(1,1)$)

$$f(x,y) = f(1,1) + 0 + \frac{1}{2} \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix}^T \begin{bmatrix} 4 & -4 \\ -4 & 3 \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix}$$

(Zero gradient from question)

$$\begin{array}{c|c} \cancel{2}\partial x_1 & 2\partial x_1 \\ -\cancel{6} & -\cancel{2} \\ -3 & -1 \end{array}$$

Let's consider $\partial x_i = x_i - 1$ for $i = 1, 2$.

$$f(x,y) = f(1,1) + \frac{1}{2}\begin{bmatrix}\partial x_1 & \partial x_2\end{bmatrix}\begin{bmatrix}4 & -4 \\ -4 & 3\end{bmatrix}\begin{bmatrix}\partial x_1 \\ \partial x_2\end{bmatrix}$$

$$f(x,y) = f(1,1) + \frac{1}{2}\left[(4\partial x_1 - 4\partial x_2) \quad (-4\partial x_1 + 3\partial x_2)\right]\begin{bmatrix}\partial x_1 \\ \partial x_2\end{bmatrix}$$

$$f(x,y) = f(1,1) + \frac{1}{2}\left[4\partial x_1^2 - 4\partial x_1\partial x_2 - 4\partial x_1\partial x_2 + 3\partial x_2^2\right]$$

$$f(x,y) = f(1,1) + \frac{1}{2}\left[4\partial x_1^2 + 3\partial x_2^2 - 8\partial x_1\partial x_2\right]$$

$$f(x,y) = f(1,1) + \frac{1}{2}(2\partial x_1 - 3\partial x_2)(2\partial x_1 - \partial x_2)$$

$$f(x,y) - f(1,1) = \left(\partial x_1 - \frac{3}{2}\partial x_2\right)\left(\partial x_1 - \frac{1}{2}\partial x_2\right) < 0$$

Now, to get the downslope one of the both should be negative and other should be positive.

So, $\left(\partial x_1 - \frac{3}{2}\partial x_2\right) < 0$ and $\left(\partial x_1 - \frac{1}{2}\partial x_2\right) > 0$

OR

$\left(\partial x_1 - \frac{3}{2}\partial x_2\right) > 0$ and $\left(\partial x_1 - \frac{1}{2}\partial x_2\right) < 0$

2.
(a) Let's say (x_1, x_2, x_3) is a point in $\mathbb{R}^3$.

Distance between two point is given by,

$$\sqrt{(x_1-a)^2+(x_2-b)^2+(x_3-c)^2}$$

Here, for ~~sym~~ Simplicity, we will consider the distance as

$$(x_1-a)^2+(x_2-b)^2+(x_3-c)^2$$

Consider the point (a,b,c) as $(-1,0,1)^T$

therefore,
We need to

$$\min_{x_1, x_2, x_3} (x_1+1)^2+x_2^2+(x_3-1)^2 \quad \text{—— (1)}$$

subjected to $x_1+2x_2+3x_3=1$

Now, making the problem unconstrained, Substitute

$x_1 = 1-(2x_2+3x_3)$ in the equation (1)

function

$$f = (1-2x_2-3x_3+1)^2 + x_2^2 + (x_3-1)^2$$

Taking gradient

$$g(x_2,x_3) = \begin{bmatrix} \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 2(2-2x_2-3x_3)(-2)+2x_2 \\ 2(2-2x_2-3x_3)(-3)+2(x_3-1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So, $10x_2+12x_3-8=0$ — (1) $\times 6$

$12x_2+20x_3-14=0$ — (2) $\times 5$

$$60x_2 + \cancel{36}^{72}x_3 - 48 = 0$$

$$60x_2 + 100x_3 - 70 = 0$$

$$- \qquad - \qquad +$$

$$-28x_3 + 22 = 0.$$

$$x_3 = \frac{-22}{-28}$$

$$\boxed{x_3 = \frac{11}{14}}$$

$$10x_2 + \cancel{12}^{6} \times \frac{11}{\cancel{14}_7} - 8 = 0.$$

$$\cancel{140x_2} \quad 70x_2 + 66 - 56 = 0.$$

$$x_2 = \frac{-10}{70}$$

$$\boxed{x_2 = \frac{-1}{7}}$$

$$x_1 = 1 - 2x_2 - 3x_3$$

$$= 1 - 2\left(-\frac{1}{7}\right) - 3\left(\frac{11}{14}\right)$$

$$= 1 + \frac{2}{7} - \frac{33}{14}$$

$$= \frac{14 + 4 - 33}{14}$$

$$\boxed{x_1 = \frac{-15}{14}}$$

Hessian is given by,

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_3} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 10 & 12 \\ 12 & 20 \end{bmatrix}$$

eigen values. $|H - \lambda I| = 0$

$$\begin{vmatrix} 10-\lambda & 12 \\ 12 & 20-\lambda \end{vmatrix} = 0.$$

$$(10-\lambda)(20-\lambda) - 144 = 0.$$

$$200 - 30\lambda + \lambda^2 - 144 = 0.$$

$$\lambda^2 - 30\lambda + 56 = 0.$$

$$(\lambda - 28)(\lambda - 2) = 0.$$

$$\boxed{\lambda = 28, 2}$$

Here, Hessian H is definite, So, we can say that the given problem is convex problem.

Problem-2(b) ~~st~~ is in Github-repo

3) Hyperplane is expressed by

$S = \{ a^T x = c \mid x \in \mathbb{R}^n \}$

The definition of convex set states that $\forall\, x_1, x_2 \in S$ and $\forall\, \lambda \in [0,1]$ $\lambda x_1 + (1-\lambda) x_2 \in S$, then S can be convex set.

So, according to the definition,

Let's select x_1 & $x_2 \in S$

So, for $\lambda x_1 + (1-\lambda) x_2$ should belong to S

to prove that Let's consider the hyper plane $a^T x = c$.

$$a^T(\lambda x_1 + (1-\lambda) x_2) = c$$

which must be true.
So, taking left hand side.

$$\text{L.H.S.} = a^T(\lambda x_1 + (1-\lambda) x_2)$$

$$= \lambda a^T x_1 + (1-\lambda) a^T x_2$$

$$= \lambda \cdot c + (1-\lambda) \cdot c \qquad (\because \text{from } a^T x = c)$$

$$= \cancel{\lambda c} + c - \cancel{\lambda c}$$

$$= c$$

$$= \text{R.H.S.} \quad (\text{Right Hand side})$$

~~So, we can conclude that~~

$$a^T(\lambda x_1 + (1-\lambda)x_2) = c$$

which holds true.

So, we can conclude that a hyper plane is a ~~x~~ convex set.

4>. $\min_{p} \max_{k} \{h(a_k^T p, I_t)\}$

S.t : $0 \le p_i < p_{max}$.

$p = [p_1, \dots, p_n]^T \rightarrow$ n lamps

$a_k \quad k = 1, \dots, m \xrightarrow{m}$ parameters of mirrors.

I_t - target Intensity.

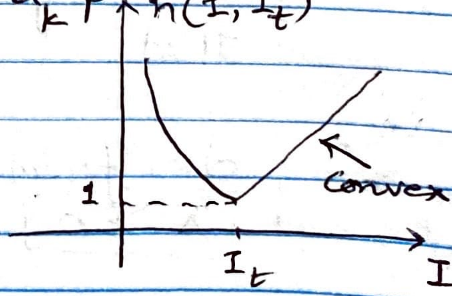
$$h(I, I_t) = \begin{cases} \dfrac{I_t}{I}, & \text{if } I \le I_t \\ \dfrac{I}{I_t}, & \text{if } I_t < I \end{cases}$$

To show problem is convex, ~~we n~~ Hessian must be atleast positive Semi-definite.

here, let's consider $I_k = a_k^T p$

So, for any $I = a^T p$.

$h(I, I_t) = (a^T p, I_t)$



$g(I, It) = \dfrac{\partial h}{\partial p} = \dfrac{\partial h}{\partial I} \dfrac{\partial I}{\partial p}$ ($\because$ chain Rule)

$= h' a$

$H(I, I_t) = \dfrac{\partial^2 h}{\partial p^2} = \dfrac{\partial h'}{\partial I} \dfrac{\partial I}{\partial p} = \dfrac{\partial (h'a)}{\partial I} \dfrac{\partial a^T p}{\partial p}$

$= h'' a a^T \qquad h$ convex $\Rightarrow h'' \ge 0$

we need to find eigenvalues of Hessian matrix. we will be neglecting h'' because it is a constant.

$$(aa^T)X = \lambda X$$

λ = eigenvalues
X = eigen vectors.

Let's multiply both sides by X^T

$$X^T(aa^T)X = \lambda X^T X$$

$$\lambda = \frac{X^T(aa^T)X}{X^T X}$$

Let's have $w \neq 0$ and $w = a^T X$
So, $w^T = (a^T X)^T = X^T a$.

So, $\lambda = \frac{w^T w}{X^T X}$

So, from above results we can say that $w^T w \geq 0$ and $X^T X \geq 0$.

So, from this result eigen value ~~is also~~ is also $\boxed{\lambda \geq 0}$

So, we can say that Hessian is positive semi-definite. So we can say that the function $h(a^T p, I_t)$ is convex.

To find $h(a^T p, I_t)$ is strictly convex we need to look at another condition.

$$h(I, I_t) = \frac{I}{I_t}, \quad I \leq I_t$$

and here $I = a_k^T p \Rightarrow \cancel{a_k} \sum_{i=1}^{n} a_k^T p_i$

So, $h(I, I_t) = \dfrac{\sum_{i=1}^{n} a_k^T p_i}{I_t}$

and $h(I, I_t)$ ~~= I~~ to be greater than or equal to zero, I must be, $I \geq 0$.

another condition:

$$h(I, I_t) = \frac{I_t}{I}, \quad \cancel{\Rightarrow} \; I_t \leq I$$

Now, $I = a_k^T p = \sum_{i=1}^{n} a_k^T p_i$

So, $h(I, I_t) = \dfrac{I_t}{\sum_{i=1}^{n} a_k^T p_i}$

here, I is at the denominator so, if $I = 0$ then $h(I, I_t)$ would be undefined. So, $I \neq 0$. So, I must be $\boxed{I > 0.}$

which means, $a_k^T p > 0$. This is valid for all k terms. So, we can say that $h(a_k^T p, I_t)$ is $\boxed{\text{strictly convex}}$ So,

maximum of convex function would also be

So, $\min_{p} \max_{k} \{h(a_k^T p, I_t)\}$ would be strictly convex.

$\boxed{\text{a convex function.}}$

→ Also $p_i \geq 0$ and $p_i < \max$. So, p_i is positive. So we can consider the $\boxed{\text{problem is a convex problem.}}$

(b) According to the combination definition.

from n lamps, if we choose any 10 lamps, we would have,

$\binom{n}{10}$ combinations of 10 lamps that has compined power less than p^*.

Let's have.

So, eg. $p_1 + p_2 + \cdots + p_{10} \leq p^*$

So, we can say that ~~the~~ this above constraint is linear combination of power of all the lamps. ~~So~~
So, due to Linear combination, the convexity remains the same. So, It has $\boxed{\text{a unique Solution}}$

(c) Now, if we consider maximum of 10 lamps on at the same time then, we would have 2^{10} possibilidgties of lamps to be on or off. In that possibilities, for m mirrors, and other constraints, problem would become non-convex, and ~~we wt~~ there might be infinitely many Solution. So, we can say that it will $\boxed{\text{not have a unique Solution}}$

5>. $c^*(y) = \max_{x} \{xy - c(x)\}$

Here, y is the cost of the product so taking that as a variable.

$\cancel{f(y)}$ $f(y) = xy - c(x)$

gradient $g(y) = \frac{\partial f}{\partial y} = x$

Hessian $H(y) = \frac{\partial^2 f}{\partial y^2} = 0.$

Here, as we can see Hessian is zero. So, the problem is linear ~~prob~~ function. and linear functions are also convex functions.

So, because the function is a convex function we can definitely say that $\max_{x} \{xy - c(x)\}$ is a convex problem.

So, from that conclusion, we can say that $c^*(y)$ is a convex function with respect to y.