Lakshya Tiwari
1222259194

MAE 598
HW # 2

Problem 1

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

$\Rightarrow$ calculate the gradient of $f(x_1, x_2)$

$$g = \begin{bmatrix} g_{x_1} \\ g_{x_2} \end{bmatrix} = \begin{bmatrix} \frac{df(x_1,x_2)}{dx_1} \\ \frac{df(x_1,x_2)}{dx_2} \end{bmatrix} = \begin{bmatrix} 4x_1 - 4x_2 \\ -4x_1 + 3x_2 + 1 \end{bmatrix}$$

substituting $g = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ to find the stationary points

$\therefore$ $4x_1 - 4x_2 = 0 \Rightarrow x_1 = x_2$

$-4x_1 + 3x_2 + 1 = 0 \Rightarrow \underline{x_1 = 1}$ $\therefore$ $\underline{x_2 = 1}$

$\rightarrow$ calculate hessian

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

$\therefore$ eigen values are
$(4-\lambda)(3-\lambda) - 16 = 0 \Rightarrow \lambda = 7.53$ & -0.53

$|H| = 12 - 16 = -4$ taking det. of H $\leftarrow$ which is < 0

$\therefore$ 1 +ve eigen value & 1 -ve eigen value

$\therefore$ H is indefinite and the stationary point of the function is a saddle. Saddle point $(1, 1)$

→ taylor's expansion at saddle point (1,1)

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

$f(1,1) = 0.5$

$g(1,1) = 0$

$$f(x) = 0.5 + \frac{1}{2}\begin{bmatrix} x_1-1 & x_2-1 \end{bmatrix}\begin{bmatrix} 4 & -4 \\ -4 & 3 \end{bmatrix}\begin{bmatrix} x_1-1 \\ x_2-1 \end{bmatrix}$$

$$= 0.5 + \frac{1}{2}\begin{bmatrix} x_1-1 & x_2-1 \end{bmatrix}\begin{bmatrix} 4x_1-4-4x_2+4 \\ -4x_1+4+3x_2-3 \end{bmatrix}$$

$$= 0.5 + \frac{1}{2}\cancel{\begin{bmatrix} x_1-1 & x_2-1 \end{bmatrix}}\left[(x_1-1)(4x_1-4x_2) + (x_2-1)(-4x_1+3x_2+1)\right]$$

$$f(x) - 0.5 = \frac{1}{2}\left[4x_1^2 + 3x_2^2 - 8x_1x_2 + 2x_2 - 1\right]$$

for finding downslopes RHS < 0

$$\therefore 4x_1^2 + 3x_2^2 - 8x_1x_2 + 2x_2 - 1 < 0$$

$$= (2x_1 - 3x_2 + 1)(2x_1 - x_2 - 1) < 0$$

∴ solution is ← for the downslopes

$$S = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \;\forall\; x_1, x_2 : (2x_1 - 3x_2 + 1)(2x_1 - x_2 - 1) < 0$$

$\{a,b,c,d\} = \{2,1,2,3\}; \{2,3,2,1\}; \{1,2,3,2\}; \{3,2,1,2\}$

Problem 2

a)

$f(x_1, x_2, x_3) = x_1 + 2x_2 + 3x_3 = 1$ in R^3

Pt $\rightarrow (-1, 0, 1)^T$

dist. is given by : $\sqrt{(x_1+1)^2 + x_2^2 + (x_3-1)^2}$

$f(x)$ is min. when $\sqrt{f(x)}$ is min.

min : $(x_1+1)^2 + x_2^2 + (x_3-1)^2$

s.t : $x_1 + 2x_2 + 3x_3 = 1$

$\therefore$ let $x_1 = 1 - 2x_2 - 3x_3$

$$\therefore g = \begin{bmatrix} 2(2-2x_2-3x_3)(-2) + 2x_2 \\ 2(2-2x_2-3x_3)(-3) + 2(x_3-1) \end{bmatrix}$$

$g = 0$ to get the saddle point

$\therefore$ we have $x_2 = -0.143$

$x_3 = 0.786$

calculate $H = \begin{bmatrix} 10 & 12 \\ 12 & 20 \end{bmatrix}$ $|H| = 56 \therefore H > 0$

+ve semi definite

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

$\therefore \lambda \rightarrow 28$
$\quad\;\; \rightarrow 2$

$\lambda_{min} > 0$ $\quad\quad \therefore H \leftarrow$ +ve definite.

$\therefore$ solution $\quad x_1 = -1.072$
$\quad\quad\quad\quad\quad x_2 = -0.143$
$\quad\quad\quad\quad\quad x_3 = 0.786$

Problem 2 Gradient Descent Method

In [88]:
```
#importing libraries
import numpy as np
import matplotlib.pyplot as plt
```

In [89]:
```
#defining objective function
def obj(x):
     return (2 - 2*x[0] - 3*x[1])**2 + (x[0])**2 + (x[1] - 1)**2
```

In [90]:
```
#defining Gradient Function
def grad(x):
     return np.array([ 10*x[0]+12*x[1]-8, 12*x[0]+20*x[1]-14])
```

In [91]:
```
#Hessian function for the problem
def hess(x):
     return np.array([[10,12],[12,20]])
```

In [92]:
```
#termination criteria
epsilon = 1e-3

#intital value
x0 = [12,25]
```

In [93]:
```
t = 0.15
kMax = 100
```

In [94]:
```
xVal = []
objVal = []
gradVal = []
```

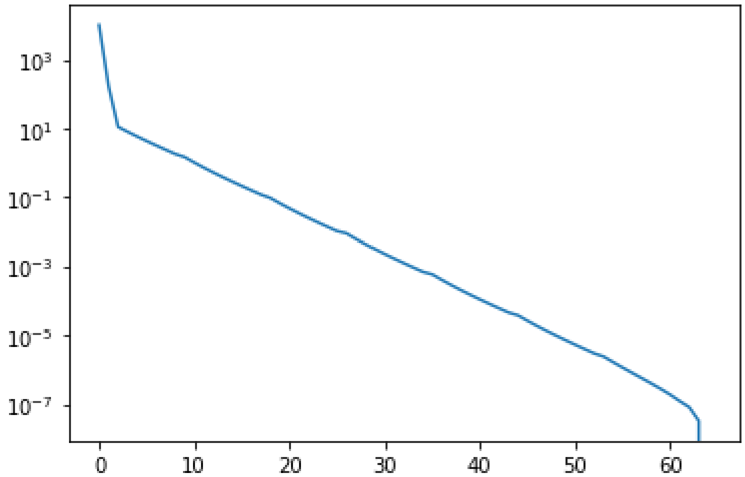
In [95]:
```
#Applying Gradient Descent Algorithm
def alphaDash(x):
    alpha = 1
    k = 0
    while obj(x - alpha*grad(x)) > obj(x) - (t * np.matmul(grad(x).T,grad(x)) * a
        alpha = 0.5 * alpha
        k=k+1
    return alpha


while np.linalg.norm(grad(x0))>epsilon :
    xVal.append(x0)
    gradVal.append(np.linalg.norm(grad(x0)))
    objVal.append(obj(x0))

    alpha=alphaDash(x0)
    x0=x0-alpha*grad(x0)
```

```
Graph for Gradient Descent
```

In [96]:
```
plt.plot(np.arange(len(gradVal)),np.abs(objVal-objVal[-1]))
plt.yscale('log')
```



Newton's Method

In [97]:
```
epsilon = 1e-7 # tolerance
x0 = np.array([10 , 20])
x1=np.array([0 , 0])
k=0
objValue=[]
gradValue=[]
```

In [98]:
```
while np.linalg.norm(x1-x0)>epsilon and k<kMax:
    x1=x0-np.matmul(np.linalg.inv(hess(x0)),grad(x0))
    gradValue.append(np.linalg.norm(grad(x0)))
    objValue.append(obj(x1))
    k+=1
```

In [99]:
```
x1
```

Out[99]:
```
array([-0.14285714,  0.78571429])
```

Problem 3

Let H be a hyperplane

WKT H is in the form $\{x : a^T x = c \text{ for } x \in R^n\}$

taking x_1, x_2 in H lie on a line segment between the two

wkt $x = \lambda x_1 + (1-\lambda)x_2$ for some λ between 0 & 1.

considering the hyperplane condition

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

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

as $a^T x = c$

$= \lambda c + (1-\lambda)c = \lambda c + c - \lambda c$

$= c$

$\therefore \lambda x_1 + (1-\lambda)x_2 \in H$

this shows that x does lie on the hyperplane.

$\therefore$ Hyperplane is convex.

Problem 4

$\min_{p} \max_{k} \{h(a_k^T p, I_t)\}$ subject to $0 \le p_i \le p_{max}$

$p := [p_1, \dots, p_n]^T$ power output n lamps

$k = 1, \dots, m$ fixed parameter for m mirrors.

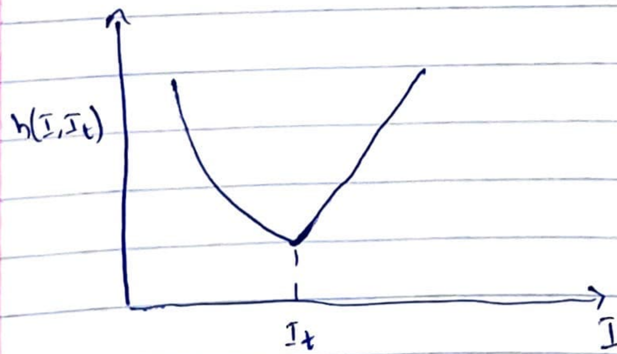
I_t = target intensity where $h(I, I_t) = \begin{cases} I_t / I & \text{if } I \le I_t \\ I / I_t & \text{if } I_t \le I \end{cases}$

a) Problem is convex

$\max\{h_1, h_2 \dots h_n\}$ where h consist $(a_k^T p ; I_t)$ The max of equation is max of h

For h

$$h(I, I_t) = \begin{cases} I_t / I & I \le I_t \\ I / I_t & I_t \le I \end{cases}$$



from the graph we can clearly see that it is a convex function with a psd hessian.
but h also has $a_k^T p$

we find H of the function to prove that h is convex w.r.t to P also.

$$\therefore \frac{\partial h}{\partial p} = \frac{\partial}{\partial p} h(a_k^T p, I_t)$$

$$\frac{\partial h}{\partial p} = \frac{dh}{dI} \frac{\partial a_{k}^{T} p}{\partial p}$$

$\therefore$ gradient

$$g = h' . a_{k}^{T}$$

$\therefore$ Hessian

$$H = \frac{\partial^2 h}{\partial p^2} = \frac{\partial}{\partial p} (h' a^T)$$

$$= h'' . a . a^T$$

as a_k is from $1, \dots, m$

$\therefore$ $a a^T$ should be a $m \times m$ matrix

from graph we know h is convex

$\therefore$ $h'' \geq 0$

for $a . a^T$

for $a \in R^{m \times m}$ $\quad a = a^T$

$\forall$ $w \in R^m$, $w \neq 0$

$w^T a w \geq 0$

we have

$$w^T a_k a_k^T w \geq 0$$

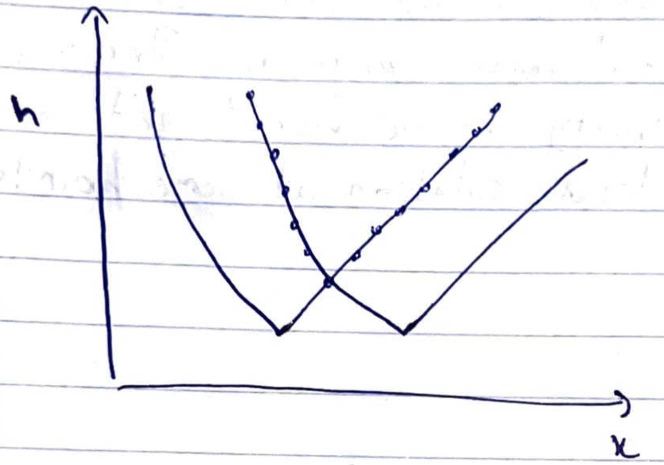
let $w^T a_k = x_i$ the $a_k^T w = x_i$

$\therefore$ $x_i^2 = w^T a_k a_k^T w \geq 0$

$\therefore$ $a . a^T$ is +ve semi definite

$\therefore$ h is a convex function w.r.t p

as we have bunch of h in the main equation each of them are a convex function.



$\therefore$ Max set of a function is also a convex function

$\therefore \quad \min_{p} \max_{h} \{h(a_k^T p, I_t)\}$ is convex.

b) Overall output of any 10 lamps to be less than p^*

we have $\sum_{i=1}^{10} (p_i) < p^*$

This means only 10 lamps are switched on

i.e

$$[1,1,1,1,1,1,1,1,1,1,0 \ldots\ldots 0] \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix} < p^*$$

The above comes out to be a linear equation. It is convex.

$\therefore$ We get an unique solution.

c) if no more than 10 lamps are switched on $(p>0)$

If no more than 10 lamps are switched on the set created is not a feasible one and hence it is not convex anymore. Therefore we might not reach the target intensity so we won't get an unique solution. Instead we might have local solution at some points.

Problem 5

$c(x) \rightarrow$ cost of producing x amount of product (differentiable everywhere)
$y \rightarrow$ price of set product

$$c^*(y) = \max_x \{xy - c(x)\}$$

The eqn given is $xy - c(x)$ It is a linear eqn

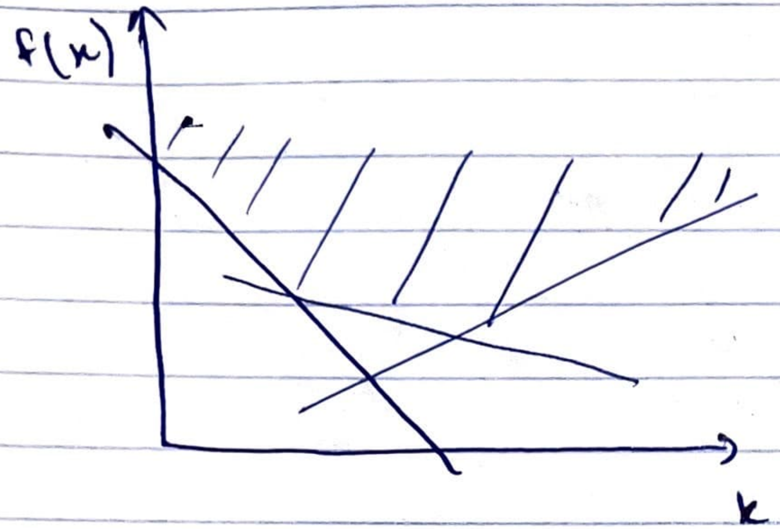
The given function is max of many linear functions. If hessian is found out to check convexity we get $H = 0$

if $H=0$ +ve semi definite.

Thus the linear function is convex

$$g = \frac{df}{dy} = x \qquad H = \frac{\partial^2 f}{\partial y^2} = 0$$

If we have max of such linear functions we can say that



The function is still convex.

∴ we can say that max of the linear function is also convex.

we can say that the profit of the product is a convex function with respect to the price.