HW 2 | MAE 598: Design Optimization | Prof. Max Yi Ren

By: Daniel Rivera

ME598/494 Homework 2

1. (20 points) Show that the stationary point (zero gradient) of the function

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

is a saddle (with indefinite Hessian).

Find the directions of downslopes away from the saddle. To do this, use Taylor's expansion at the saddle point to show that

$$f(x_1, x_2) = f(1, 1) + (a\partial x_1 - b\partial x_2)(c\partial x_1 - d\partial x_2),$$

with some constants a, b, c, d and $\partial x_i = x_i - 1$ for $i = 1, 2$. Then the directions of downslopes are such $(\partial x_1, \partial x_2)$ that

$$f(x_1, x_2) - f(1, 1) = (a\partial x_1 - b\partial x_2)(c\partial x_1 - d\partial x_2) < 0.$$

2. (a) (10 points) Find the point in the plane $x_1 + 2x_2 + 3x_3 = 1$ in $\mathbb{R}^3$ that is nearest to the point $(-1, 0, 1)^T$. Is this a convex problem? Hint: Convert the problem into an unconstrained problem using $x_1 + 2x_2 + 3x_3 = 1$.

   (b) (40 points) Implement the gradient descent and Newton's algorithm for solving the problem. Attach your codes in the report, along with a short summary of your findings. The summary should include: (1) The initial points tested; (2) corresponding solutions; (3) A log-linear convergence plot.

3. (5 points) Prove that a hyperplane is a convex set. Hint: A hyperplane in $\mathbb{R}^n$ can be expressed as: $\mathbf{a}^T\mathbf{x} = c$ for $\mathbf{x} \in \mathbb{R}^n$, where $\mathbf{a}$ is the normal direction of the hyperplane and c is some constant.

4. (15 points) Consider the following illumination problem:

$$\begin{aligned} \min_{\mathbf{p}} \quad & \max_k \{h(\mathbf{a}_k^T\mathbf{p}, I_t)\} \\ \text{subject to:} \quad & 0 \le p_i \le p_{\max}, \end{aligned}$$

where $\mathbf{p} := [p_1, ..., p_n]^T$ are the power output of the n lamps, $\mathbf{a}_k$ for $k = 1, ..., , m$ are fixed parameters for the m mirrors, I_t the target intensity level. $h(I, I_t)$ is defined as follows:

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

(a) (5 points) Show that the problem is convex.

(b) (5 points) If we require the overall power output of any of the 10 lamps to be less than p^*, will the problem have a unique solution?

(c) (5 points) If we require no more than 10 lamps to be switched on ($p > 0$), will the problem have a unique solution?

5. (10 points) Let $c(x)$ be the cost of producing x amount of product A and assume that $c(x)$ is differentiable everywhere. Let y be the price set for the product. Assuming that the product is sold out. The total profit is defined as

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

Show that $c^*(y)$ is a convex function with respect to y.

ME598/494 Homework 2

1. (20 points) Show that the stationary point (zero gradient) of the function

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

is a saddle (with indefinite Hessian).

Find the directions of downslopes away from the saddle. To do this, use Taylor's expansion at the saddle point to show that

$$f(x_1, x_2) = f(1, 1) + (a\partial x_1 - b\partial x_2)(c\partial x_1 - d\partial x_2),$$

with some constants a, b, c, d and $\partial x_i = x_i - 1$ for $i = 1, 2$. Then the directions of downslopes are such $(\partial x_1, \partial x_2)$ that

$$f(x_1, x_2) - f(1, 1) = (a\partial x_1 - b\partial x_2)(c\partial x_1 - d\partial x_2) < 0.$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{\partial x_2} \end{bmatrix}$$

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

2 equations, 2 unknowns:

$$4x_1 - 4x_2 = 0$$

$$\Rightarrow x_1 - x_2 = 0$$

$$x_1 = x_2$$

$$-4x_1 + 3x_2 + 1 = 0$$

$$-4x_1 + 3x_1 + 1 = 0$$

$$-x_1 = -1$$

$$\Rightarrow x_1 = x_2 = 1$$

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

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

$\Rightarrow$ Thus saddle point is $x^* = [1, 1]^T$

$$H(1,1) = \begin{bmatrix} 4 & -4 \\ -4 & 3 \end{bmatrix}$$

The Hessian H is indefinite (it has one eigenvalue that is positive and another that is negative). This shows that $x^* = [1,1]^T$ is a saddle point

Using Taylor's expansion at saddle point $(1,1)$

$$f(x_1, x_2) = f(1,1) + \cancelto{0}{\nabla f(x - x^*)} \cdot \partial x + \frac{1}{2} \partial x^T H \partial x$$
$@(1,1)$

Let $H = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} a & -b \\ -c & d \end{bmatrix}$ and $\begin{bmatrix} x_1 - x_1^* \\ x_2 - x_1^* \end{bmatrix} = \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix} = \begin{bmatrix} \partial x_1 \\ \partial x_2 \end{bmatrix}$

Thus

$$f(x_1, x_2) - f(1,1) = \frac{1}{2} \begin{bmatrix} \partial x_1 & \partial x_2 \end{bmatrix} \begin{bmatrix} a & -b \\ -c & d \end{bmatrix} \begin{bmatrix} \partial x_1 \\ \partial x_2 \end{bmatrix}$$

$$\underset{@(1,1)}{f(x_1,x_2) - f(1,1)} = \frac{1}{2}\begin{bmatrix} a\partial x_1 - c\partial x_2 & -b\partial x_1 + d\partial x_2 \end{bmatrix}\begin{bmatrix} \partial x_1 \\ \partial x_2 \end{bmatrix}$$

$$\underset{@(1,1)}{f(x_1,x_2) - f(1,1)} = \frac{1}{2}\left[\partial x_1(a\partial x_1 - c\partial x_2) + \partial x_2(d\partial x_2 - b\partial x_1) \right]$$

$$f(x_1,x_2) - f(1,1) = \frac{1}{2}\left(a\partial x_1^2 - c\partial x_1\partial x_2 + d\partial x_2^2 - b\partial x_1\partial x_2 \right)$$

$$f(x_1,x_2) - f(1,1) = \frac{1}{2}\left(a\partial x_1^2 + (d-c)\partial x_1\partial x_2 + d\,\partial x_2^2 \right)$$

factor into a,b,c,d

$$\Rightarrow \quad f(x_1,x_2) - f(1,1) = (a\partial x_1 - b\partial x_2)(c\partial x_1 - d\partial x_2) < 0$$

$a\partial x_1 - b\partial x_2 < 0$ AND $c\partial x_1 - d\partial x_2 > 0$

OR

$a\partial x_1 - b\partial x_2 > 0$ AND $c\partial x_1 - d\partial x_2 < 0$

Thus the direction of the downslope is:

$$f(x_1,x_2) - f(1,1) = (a\partial x_1 - b\partial x_2)(c\partial x_1 - d\partial x_2) < 0$$

where a,b,c,d are some constants

and

$\partial x_i = x_i - 1$ for $i = 1,2$

2. (a) (10 points) Find the point in the plane $x_1 + 2x_2 + 3x_3 = 1$ in $\mathbb{R}^3$ that is nearest to the point $(-1, 0, 1)^T$. Is this a convex problem? Hint: Convert the problem into an unconstrained problem using $x_1 + 2x_2 + 3x_3 = 1$.

(b) (40 points) Implement the gradient descent and Newton's algorithm for solving the problem. Attach your codes in the report, along with a short summary of your findings. The summary should include: (1) The initial points tested; (2) corresponding solutions; (3) A log-linear convergence plot.

The normal vector of the plane is needed:

$n = \langle 1, 2, 3 \rangle$

thus $\langle -1+t, 0+2t, 1+3t \rangle$ is the line of the normal vector to the plane where t is some arbitrary variable. Plug in this line into the plane equation:

$$-1+t + 2(2t) + 3(1+3t) = 1$$

$$-1 + t + 4t + 3 + 9t = 1$$

$$14t = -1$$

$$\Rightarrow t = -1/14$$

Thus the point $(-15/14, -1/7, 11/14)$ is closest to the point $(-1, 0, 1)^T$

To determine if this is a convex problem, we must first determine the Hessian (H):

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

$$\text{s.t.} \quad x_1 + 2x_2 + 3x_3 = 1$$

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

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

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

$$=4x_2^2+6x_2x_3-4x_2+6x_2x_3+9x_3^2-6x_3-4x_2-6x_3+4+x_2^2+x_3^2-2x_3+1$$

$$=5x_2^2+12x_2x_3-8x_2+10x_3^2-14x_3+5$$

$$\nabla f=\begin{vmatrix}\frac{df}{dx_2}\\\frac{df}{dx_3}\end{vmatrix}=\begin{vmatrix}10x_2+12x_3-8\\12x_2+20x_3-14\end{vmatrix}$$

$$\nabla f(x_i) = \begin{bmatrix} 2(x_1+1) \\ 2x_2 \\ 2(x_3-1) \end{bmatrix} = \begin{bmatrix} 2x_1+2 \\ 2x_2 \\ 2x_3-2 \end{bmatrix}$$

$i=1,2,3$

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

The Hessian has all positive eigen values, making it positive-definite. Thus, this is a convex problem

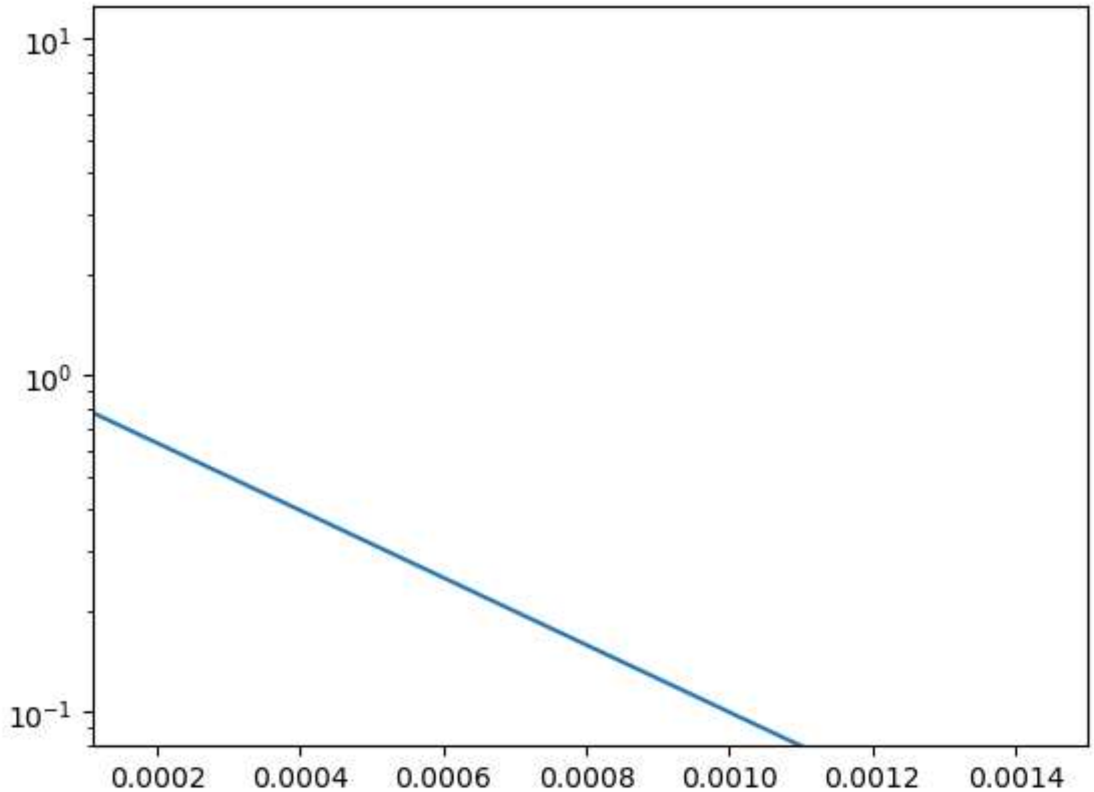
Summary of Findings: the initial guess can be anything and the algorithm will continue to converge to the values of x_2 and x_3 (since we can solve easily for x_1).

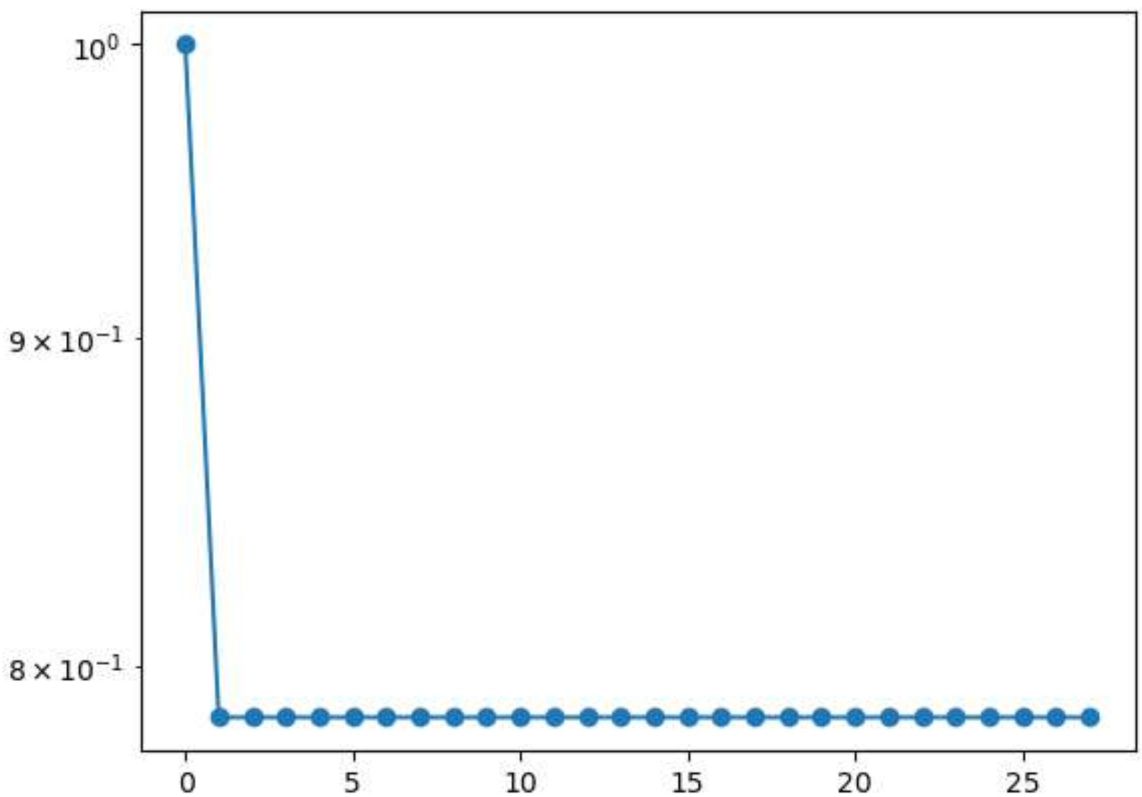
In [1]:

```
#Submission by: Daniel Rivera | MAE 598: Design Optimization | Prof. Max Yi R
#Problem 2 Answer using gradient and Newton's Method
import numpy as np
import matplotlib.pyplot as plt
#Original Function
obj = lambda x: 5 * x[0] ** 2 + 12 * x[0] * x[1] - 8 * x[0] + 10 * x[1] ** 2


def grad(x):  # gradient found for 2 dimensional function
    return [10 * x[0] + 12 * x[1] - 8, 12 * x[0] + 20 * x[1] - 14]

#Eps is used to minimize error
eps = 1e-3
#Specify Input
xin = int(input('Input initial guess:'))  # different guesses were tested wit
#Let n be your counter
n = 0
#Create a solution that initializes with your input
soln = [[xin, xin]]

x = [xin, xin]

error = np.linalg.norm(grad(x))

#start Newton's method that uses lambda and phi
def line_search(x):
    a = 1.
    phi = lambda a, x: [obj(x) - a * .8 * grad(x)[0] ** 2, obj(x) - a * .8 *

    while phi(a, x)[0] < obj([x[0] - a * grad(x)[0], x[1] - a * grad(x)[1]])
            [x[0] - a * grad(x)[0], x[1] - a * grad(x)[1]]):
        # outputted value only returning one unchanged number error in this w
        a = .5 * a

    return a

#use a while loop to minimize error
while error >= eps:
    a = line_search(x)
    x[0] = x[0] - a * grad(x)[0]
    x[1] = x[1] - a * grad(x)[1]
    soln.append(x)
    error = np.linalg.norm(grad(x))

print(soln)
#zip your solution to convert your tupled solution into a plottable list
y1, y2= zip(*soln)
print(y1)
print(y2)
#plot x2 and x3 and call them y1 and y2 respectively (x1 can be solved mathem
plt.yscale("log")
plt.xlim(0.00011,0.0015)
plt.scatter(range(len(y1)),y1)
plt.plot(range(len(y1)),y1)
plt.show()
```

```
plt.yscale("log")
plt.scatter(range(len(y2)),y2)
plt.plot(range(len(y2)),y2)
plt.show()
```

```
input initial guess:1
[[1, 1], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461], [-0.14256492718762875, 0.7855466747113461]]
(1, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875, -0.14256492718762875)
(1, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461, 0.7855466747113461)
```


10^1
10^0
10^{-1}
0.0002
0.0004
0.0006
0.0008
0.0010
0.0012
0.0014


10^0
9×10^{-1}
8×10^{-1}
0
5
10
15
20
25

In [ ]:

3. (5 points) Prove that a hyperplane is a convex set. Hint: A hyperplane in $\mathbb{R}^n$ can be expressed as: $\mathbf{a}^T\mathbf{x} = c$ for $\mathbf{x} \in \mathbb{R}^n$, where $\mathbf{a}$ is the normal direction of the hyperplane and c is some constant.

First we should define ovr hyperplane "$\mathcal{H}$" :

$$\mathcal{H} = \{ x / x \in \mathbb{R}^n , a^T x = c \}$$

let x_1 and x_2 be points in $\mathcal{H}$, thvs

$a^T x_1 = c$ and $a^T x_2 = c$

Using eigenvalves we know that :

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

$\boxed{\text{Thvs } \lambda x_1 + (1-\lambda)x_2 \text{ belongs to } \mathcal{H}\text{, which means that } \mathcal{H} \text{ is convex}}$

Pg.20

4. (15 points) Consider the following illumination problem:

$$\min_{\mathbf{p}} \quad \max_k \{h(\mathbf{a}_k^T \mathbf{p}, I_t)\}$$

$$\text{subject to:} \quad 0 \le p_i \le p_{\max},$$

where $\mathbf{p} := [p_1, ..., p_n]^T$ are the power output of the n lamps, $\mathbf{a}_k$ for $k = 1, ..., m$ are fixed parameters for the m mirrors, I_t the target intensity level. $h(I, I_t)$ is defined as follows:

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

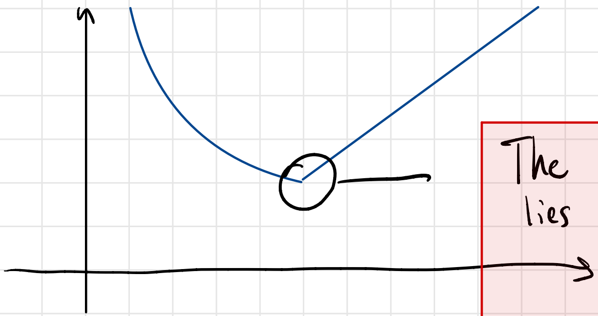
(a) (5 points) Show that the problem is convex.

(b) (5 points) If we require the overall power output of any of the 10 lamps to be less than p^*, will the problem have a unique solution?

(c) (5 points) If we require no more than 10 lamps to be switched on ($p > 0$), will the problem have a unique solution?

a.) We can graphically depict the max $\{h_1, h_2\}$

where $h = \begin{cases} I_t/I & \text{if } I \le I_t \\ I/I_t & \text{if } I_t \le I \end{cases}$



The solution of this function lies at the intersection of I_t/I and I/I_t, making this a convex problem

Furthermore, we can prove this mathematically:

$$\min \left(I_t - \sum_{k,n=1}^{n,m} a_k^T p_n \right)^2$$

$$\text{s.t.} \quad 0 \le p_i \le p_{max}$$

$$f(p_1 \cdots p_n) = 2\left(I_t - \sum_{k,n=1}^{n,m} a_k^T p_n\right)^2$$

$$f(p) = (I_t - a_i^T p)^2$$

$$\nabla(f(p)) = 2(I_t - a_i^T p)\cdot(-a_i)$$

$$\nabla(f(p)) = -2a_i I_t + 2a_i^T a_i p$$

$$H(f(p)) = 0 + 2a_i a_i^T$$

$$H(p) = \sum_{k=1}^{n} a_k a_k^T$$

Thus, the Hessian is positive Semi-definite which means that the problem is convex as long as the number of mirrors is greater than or equal to the number of lamps ($m \geq n$). Thus $\max\{f(p)\}$ is <u>convex</u> as well as the $\min\max\{f(p)\}$.

b.) $p = [p_1 + \dots + p_{10}]^T < p^*$

Applying a contraint on p allows there to be a unique solution as there will be only one way to maximize.

c.) $n \leq 10$ thus

$p = [p_1 + \dots + p_{10}]^T$

Applying this constraint establishes a set that is not convex. We simply are unable to determine the amount of local solutions we have. In other words, the summation of the total power could be dominated by any of the ten lamps.

5. (10 points) Let $c(x)$ be the cost of producing x amount of product A and assume that $c(x)$ is differentiable everywhere. Let y be the price set for the product. Assuming that the product is sold out. The total profit is defined as

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

Show that $c^*(y)$ is a convex function with respect to y.

let $xy = a$ where a is the product of A's selling price and the quantity of product A. Assuming that this business has a "successful business model," we can assume that

$$a \gg c(x)$$

Since a successful business should be selling their product at a higher price than it costs to make, thus:

$$c^*(y) = \max \left\{ \left(a(x,y) - c(x) \right) \right\}$$

$$\nabla(c^*(y)) = \begin{bmatrix} \frac{dc^*}{dx} \\ \frac{dc^*}{dy} \end{bmatrix}$$

$$\nabla(c^*(y)) = \begin{bmatrix} y - \frac{\partial c(x)}{\partial x} \\ x - \frac{\partial c(x)}{\partial y} \end{bmatrix}$$

$$H(c^*(y)) = \begin{bmatrix} -\dfrac{\partial^2 c(x)}{\partial x^2} & 1 - \dfrac{\partial^2 c(x)}{\partial x \partial y} \\ 1 - \dfrac{\partial^2 c(x)}{\partial y \partial x} & -\dfrac{\partial^2 c(x)}{\partial y^2} \end{bmatrix}$$

$\boxed{\text{Since } \det|H(c^*(y))| > 0, \text{ then } c^*(y) \text{ is convex with respect to } y}$