MAE : 598 Design Optimization

Vedant Rajesh Manadik
ASU ID - 1222505793

Homework 02

Problem 01

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

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

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

i) Gradient $\nabla f(x_1, x_2) = \dfrac{\partial f(x_1, x_2)}{\partial x_1} \quad \dfrac{\partial f(x_1, x_2)}{\partial x_2}$

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \frac{\partial (2x_1^2 - 4x_1x_2 + 1.5x_2^2 + x_2)}{\partial x_1}$$

$$\frac{\partial f(x_1, x_2)}{\partial x_2} = \frac{\partial (2x_1^2 - 4x_1x_2 + 1.5x_2^2 + x_2)}{\partial x_2}$$

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

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

ii) Hessian

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

$$H(x_1, x_2) = \begin{bmatrix} 4 & -4 \\ -4 & 3 \end{bmatrix}$$

iii) Eigenvalues

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

$$\left| \begin{bmatrix} 4 & -4 \\ -4 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0$$

$$\begin{vmatrix} 4-\lambda & -4 \\ -4 & 3-\lambda \end{vmatrix} = 0$$

$$(4-\lambda)(3-\lambda) - (-4 \times -4) = 0$$

$$12 - 4\lambda - 3\lambda + \lambda^2 - 16 = 0$$

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

$$\lambda = (7.5311, -0.5311)$$

$$\boxed{\text{Eigenvalue}(\lambda) = \lambda_1 = 7.5311, \ \lambda_2 = -0.5311}$$

Its a indefinite as one eigenvalue is positive and the other is negative. ∴ The stationary point is a saddle point.

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

Using Taylors expansion (second order approximation)

$$f(x_1, x_2) = (2x_1^2 - 4x_1x_2 + 1.5x_2^2 + x_2)_{(x_1, x_2)} + \begin{bmatrix} 4x_1 - 4x_2 \\ -4x_1 + 3x_2 + 1 \end{bmatrix}^T_{(x_1, x_2)}$$

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

$$= (2 \cdot 1^2 - 4 \times 1 \times 1 + 1.5 \times 1^2 + 1) + \begin{bmatrix} 4 \cdot 1 - 4 \cdot 1 \\ -4 \cdot 1 + 3 \cdot 1 + 1 \end{bmatrix}^T \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix}$$

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

Substituting $x_1 - 1 = \partial x_1$ and $x_2 - 1 = \partial x_2$

$$= \frac{1}{2} + \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} \partial x_1 \\ \partial x_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \partial x_1 & \partial x_2 \end{bmatrix} \begin{bmatrix} 4 & -4 \\ -4 & 3 \end{bmatrix} \begin{bmatrix} \partial x_1 \\ \partial x_2 \end{bmatrix}$$

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

$$= \frac{1}{2} + \frac{1}{2} \left((4\partial x_1 - 4\partial x_2)\partial x_1 + (-4\partial x_1 + 3\partial x_2)\partial x_2 \right)$$

$$= \frac{1}{2} + \frac{1}{2} \left(4\partial x_1^2 - 4\partial x_1 \partial x_2 - 4\partial x_1 \partial x_2 + 3\partial x_2^2 \right)$$

$$= \frac{1}{2} + \frac{1}{2}\left(4\partial x_1^2 - 8\partial x_1 \partial x_2 + 3\partial x_2^2\right)$$

$$= \frac{1}{2} + \frac{1}{2}\left(4\partial x_1^2 - 2\partial x_1 \partial x_2 - 6\partial x_1 x_2 + 3\partial x_2^2\right)$$

$$= \frac{1}{2} + \frac{1}{2}\left(2\partial x_1(2\partial x_1 - \partial x_2) - 3\partial x_2(2\partial x_1 - \partial x_2)\right)$$

$$= \frac{1}{2} + \frac{1}{2} \times (2\partial x_1 - 3\partial x_2)(2\partial x_1 - \partial x_2)$$

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

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

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

Problem 02

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.

→ Distance between 2 points $= \sqrt{(x_1+1)^2 + x_2^2 + (x_3-1)^2}$

we know $\min f(x) = \min(\sqrt{f(x)})$

$\therefore f(x) = (x_1+1)^2 + (x_2^2) + (x_3-1)^2$ — a

$x_1 + 2x_2 + 3x_3 = 1$ — given

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

Substituting value of x_1 in eqn a

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

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

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

$(4 + 4x_2^2 + 9x_3^2 - 8x_2 + 12x_2x_3 - 12x_3) + x_2^2 + x_3^2 - 2x_3 + 1$

$4x_2^2 + x_2^2 + 9x_3^2 + x_3^2 - 8x_2 - 12x_3 - 2x_3 + 12x_2x_3 + 5$

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

$\therefore$ The object function $f(x_2, x_3) = 5x_2^2 + 10x_3^2 - 8x_2 - 14x_3 + 12x_2x_3 + 5$

Gradient $\nabla f(x_2 x_3) = \begin{bmatrix} \dfrac{\partial f(x_2\, x_3)}{\partial x_2} \\ \dfrac{\partial f(x_2\, x_3)}{\partial x_3} \end{bmatrix}$

$$= \begin{bmatrix} 10x_2 + 0 - 8 - 0 + 12x_3 + 0 \\ 0 + 20x_3 - 14 + 12x_2 + 0 \end{bmatrix}$$

$$\nabla f(x_2 x_3) = \begin{bmatrix} 10x_2 + 12x_3 - 8 \\ 20x_3 + 12x_2 - 14 \end{bmatrix}$$

Now, Hessian

$$H = \begin{bmatrix} \dfrac{\partial f(x_2 x_3)}{\partial x_2^2} & \dfrac{\partial f(x_2\, x_3)}{\partial x_2 x_3} \\ \dfrac{\partial f(x_2 x_3)}{\partial x_3 x_2} & \dfrac{\partial f(x_2 x_3)}{\partial x_3^2} \end{bmatrix}$$

$$H = \begin{bmatrix} 10 & 12 \\ 12 & 20 \end{bmatrix}$$

For Eigenvalue,

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

$$\left| \begin{bmatrix} 10 & 12 \\ 12 & 20 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0$$

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

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

$$200 - 10\lambda - 20\lambda + \lambda^2 - 144 = 0$$

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

$$\lambda = (28, 2)$$

$$\lambda_1 = 28, \quad \lambda_2 = 2$$

$\therefore$ Since the eigenvalues for the matrix are positive

The function is strictly convex

$\boxed{\because \text{ This is a convex problem.}}$

Now

$$\nabla f(x_2, x_3) = 0$$

$$\begin{bmatrix} 10x_2 + 12x_3 - 8 \\ 12x_2 + 20x_3 - 14 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$10x_2 + 12x_3 - 8 = 0 \longrightarrow 10x_2 + 12x_3 = 8 \quad -①$$

$$12x_2 + 20x_3 - 14 = 0 \longrightarrow 12x_2 + 20x_3 = 14 \quad -②$$

Multiply eqⁿ ① by 5 and eqⁿ ② by 3

then substracting eqⁿ ② from eqⁿ ①

$$50x_2 + 60x_3 = 40$$

$$\underline{-\;36x_2 + 60x_3 = 42}$$

$$14x_2 = -2$$

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

Substituting value of x_2 in eqn (i)

$$10x_2 + 12x_3 = 8$$

$$10 \times \left(-\frac{1}{7}\right) + 12x_3 = 8$$

$$12x_3 = 8 + \frac{10}{7}$$

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

Substituting value of x_2 and x_3 in eqn (i)

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

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

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

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

The point nearest to point $(-1,0,1)^T$.
is $(-15/14, -1/7, 11/14)$ which lie in plane
$x_1 + 2x_2 + 3x_3 = 1$ in $\mathbb{R}^3$

b) Attached after problem 05.

Problem 03

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.

$\rightarrow$ A hyperplane is expressed as

$$a^T x = c$$

for $x \in \mathbb{R}^n$

let x_1 and x_2 points on hyperplane

$$x_1, x_2 \in \text{hyperplane.}$$

$$\therefore \quad a^T x_1 = c$$
$$a^T x_2 = c$$

for any $\lambda \in [0, 1]$

$$\lambda x_1 + (1-\lambda) x_2$$

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

$$\cancel{\lambda a^T x_1} + a^T x_2 - \cancel{\lambda a^T x_2}$$

$$a^T x_2$$

$$= c$$

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

$\boxed{\therefore \text{ This hyperplane is a Convex set}}$

Problem 04.

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?

$\Rightarrow$

a) $\min_{P} \max_{K} \{h(a_K^T P, I_t)\}$

I_t = target intensity

For $I < I_t$

$$h = \frac{I_t}{I} \implies y = \frac{c}{x}$$

which is a hyperbola till $I = I_t$

i) Hessian

$$H(h) = \frac{\partial^2 (I_t/I)}{\partial I^2}$$

$$= \frac{2 I_t}{I^3}$$

$\therefore$ for $P_i > 0$ the $I \geq 0$

So as $I \geq 0$

$$\rightarrow \frac{2 I_t}{I^3} > 0$$

$\therefore$ Its convex for $I \leq I_t$

Now when $I > I_t$,

$$h = \frac{I}{I_t} \Rightarrow y = mx$$

So we can say its linear and $H(h) = 0$
$\therefore$ it is also convex

$$\boxed{\text{And therefore } h(I, I_t) \text{ is a convex problem}}$$

b)

As the function is linear, its not strictly convex for the part of the function when $h = I/I_t$

Now
$$h = \begin{cases} I_t/I & I \leq I_t \\ I/I_t & I \geq I_t \end{cases}$$

$$I = ap$$

The solution for $\max\{h_1, h_2\}$ must lie on the intersection of both functions. It will stay the same as the intersection point would be last. as the function stops decreasing and we get best function value

Now

$$\frac{a_1^T p}{I_t} = \frac{a_2^T p}{I_t} = \dots\dots \frac{a_n^T p}{I_t}$$

Because of these existing linear function are equal. This intersection is a unique solution.

Taking the functions representing intersection conditions

$$\min_{p} \frac{I_t}{a^T p} = \frac{I_t}{I}$$

$$\text{s.t} \quad (a_1 - a_2)^T p = 0$$

$$\vdots$$

$$(a_n - a_{n+1})^T p = 0$$

Hessian of this functions

$$H = \frac{2 I_t^*}{I^{*3}}$$

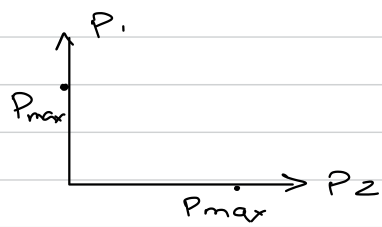
when $p_i > 0$, $I^* \geq 0$

$$\therefore \quad \frac{2 I_t}{I^3} > 0$$

And the function is convex as Hessian is positive definite.

$\therefore$ This will have a unique solution for the overall power output of any of the 10 lamps to be less than p^*

c) If we use only 2 lamps (P_1 and P_2). Any of these lamps can give power output between 0 and P_{max}. So its possible that the power output wouldn't lie in between.



$\therefore$ Its not certain weather it will have a unique soln

Problem Q5

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.

$\longrightarrow$

$c^*(y) = \max\limits_x \{xy - c(x)\}$ = total profit

$c(x)$ = cost of producing x amount of product A.

i) Gradient

$$= \frac{\partial c^*(y)}{\partial y}$$

$$= \frac{\partial (xy - c(x))}{\partial y}$$

$$= x - 0$$

$$= x$$

ii) Hessian

$$H = \frac{\partial^2 c^*(y)}{\partial y^2}$$

$$= \frac{\partial (x)}{\partial y}$$

$$H = 0$$

iii) Eigenvalues

The hessian is 0, as a result the eigenvalues (λ_1, λ_2) are also equal to zero.

now

$\lambda \geq 0$ $\therefore$ the hessian is positive semi definite

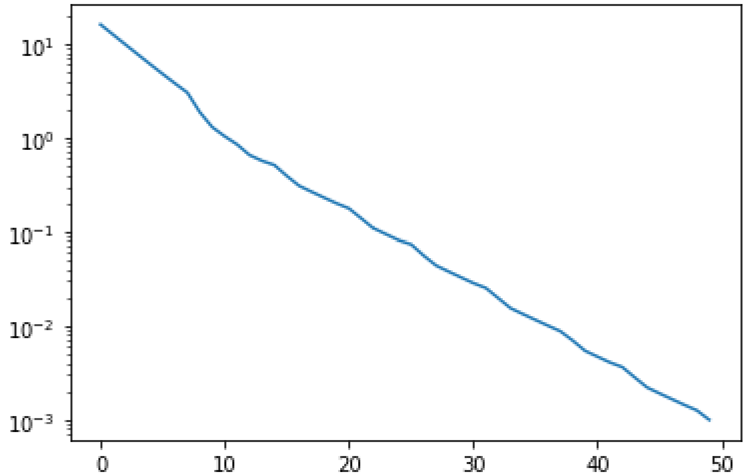
and $\therefore$ $C^*(y)$ is a convex function. w.r.t y

Maximum of multiple convex function is also convex

$\therefore$ $C^*(y) = \max_x \{xy - C(x)\}$ is a convex function w.r.t y

```
import numpy as np
import matplotlib.pyplot as plt
obj_lambda = lambda x: (2 - 2*x[0] - 3*x[1])**2 + x[0]**2 + (x[1]-1)**2
def grad_var(x):
     return np.array([10*x[0] + 12*x[1] - 8, 20*x[1] + 12*x[0] -14])
eps = 1e-3
initial_guess= np.array([0,0])
print(initial_guess)
counter_1 = 0
solution_array = [initial_guess]
initial_x = solution_array[counter_1]
error = np.linalg.norm(grad_var(initial_x))
step_size_a = 0.01
def lineSerachFunction(x, d):
    a = 1.  # initialize step size

    def phi(a,x,d):
        return obj_lambda(x)+a*0.8*np.dot(grad_var(x),d)

    while phi(a,x,d)<obj_lambda(x+a*d):
        a = 0.5*a

    return a
counter_2 = 0
err_array = [np.linalg.norm(grad_var(initial_x))]
# print(grad_var(initial_x))
i = [0]
hess_matrix = ([10,12],[12,20])
while error >= eps:
    d = -grad_var(initial_x)

    a = lineSerachFunction(initial_x, d)
    # print('this is a', a)
    initial_x = initial_x + a * d
    #print(initial_x)
    solution_array.append(initial_x)
    error = np.linalg.norm(grad_var(initial_x))
    err_array.append(error)
    counter_2 = counter_2 + 1
    i.append(counter_2)

x1 = 1 - 2*initial_x[0] - 3*initial_x[1]
print(f"x1:{x1} , x2:{initial_x[0]}, x3:{initial_x[1]}")
print("Minimum Distance:",np.sqrt((2 - 2*initial_x[0] - 3*initial_x[1])**2 + initial_x[0]**2
plt.plot(i,err_array)
plt.yscale('log')
```

```
[0 0]
x1:-1.071439502165902 , x2:-0.14244266863401012, x3:0.7854416131446407
Minimum Distance: 0.26726170261985605
```



Colab paid products - Cancel contracts here

✓ 0s completed at 10:12 PM

```
import numpy as np
import matplotlib.pyplot as plt

obj_lambda = lambda x: (2 - 2*x[0] - 3*x[1])**2 + x[0]**2 + (x[1]-1)**2
def grad_var(x):
  return np.array([10*x[0] + 12*x[1] - 8, 20*x[1] + 12*x[0] -14])

eps = 1e-3
initial_guess= np.array([0,0])
print(initial_guess)
counter_1 = 0
solution_array = [initial_guess]
initial_x = solution_array[counter_1]
error = np.linalg.norm(grad_var(initial_x))
step_size_a = 0.01

def lineSerachFunction(x, d):
  a = 1. # initialize step size

  def phi(a,x,d):
    return obj_lambda(x)+a*0.8*np.dot(grad_var(x),d)

  while phi(a,x,d)<obj_lambda(x+a*d):
    a = 0.5*a

  return a

counter_2 = 0
err_array = [np.linalg.norm(grad_var(initial_x))]
#print(grad_var(initial_x))
i = [0]
hess_matrix = ([10,12],[12,20])
while error >= eps:
  d = -np.matmul(np.linalg.inv(hess_matrix),grad_var(initial_x))
  #print('this is D',d)
  a = lineSerachFunction(initial_x, d)
  #print('this is a', a)
  initial_x = initial_x + d
  #print(initial_x)
  solution_array.append(initial_x)
  error = np.linalg.norm(grad_var(initial_x))
  err_array.append(error)
  counter_2 = counter_2 + 1
  i.append(counter_2)

x1 = 1 - 2*initial_x[0] - 3*initial_x[1]
print(f"x1:{x1} , x2:{initial_x[0]}, x3:{initial_x[1]}")
print("Minimum Distance:",np.sqrt((2 - 2*initial_x[0] - 3*initial_x[1])**2 + initial_x[0]**2
plt.plot(i,err_array)
plt.yscale('log')
```

```
[0 0]
x1:-1.071428571428572 , x2:-0.1428571428571428, x3:0.7857142857142858
Minimum Distance: 0.26726124191242445
```