Lakshya Tiwari
1222259194

ME598/494 Homework 1

1. Solve the following problem using Python SciPy.optimize. Please attach your code and results. Specify your initial guesses of the solution. If you change your initial guess, do you find different solutions? (**30 points**)

$$
\begin{aligned}
\text{minimize:} \quad & (x_1 - x_2)^2 + (x_2 + x_3 - 2)^2 + (x_4 - 1)^2 + (x_5 - 1)^2 \\
\text{subject to:} \quad & x_1 + 3x_2 = 0 \\
& x_3 + x_4 - 2x_5 = 0 \\
& x_2 - x_5 = 0 \\
& -10 \leq x_i \leq 10, \; i = 1, \ldots, 5
\end{aligned}
$$

   Note: Please learn how to use **break points** to debug. You can use Python on Google Colab.

2. Let x and $b \in \mathbb{R}^n$ be vectors and $A \in \mathbb{R}^{n \times n}$ be a square matrix. Define $f : \mathbb{R}^n \to \mathbb{R}$ as $f(x) = b^T x + x^T A x$. (**50 points**)

   (a) What is the gradient and Hessian of $f(x)$ with respect to x?

   (b) Derive the first and second order Taylor's approximations of $f(x)$ at $x = 0$. Are these approximations exact?

   (c) What are the necessary and sufficient conditions for A to be positive definite?

   (d) What are the necessary and sufficient conditions for A to have full rank?

   (e) If there exists $y \in \mathbb{R}^n$ and $y \neq 0$ such that $A^T y = 0$, then what are the conditions on b for $Ax = b$ to have a solution for x?

3. Due to the recent inflation, let's reconsider the Stigler diet problem proposed by Nobel laureate George Stigler after the second World War: Consider that the grocery store offers N types of food of your interest, and each food contains the same M types of nutrition. Denote a_{ij} as the quantity of nutrition type j of food type i for $i = 1, 2, ..., N$ and $j = 1, 2, ..., M$, c_i the unit price of food type i, and b_j the necessary

quantity of nutrition type j for a month. Formulate an optimization problem to determine the minimum grocery cost to satisfy the nutrition needs. (**20 points**)

Problem-1

In [73]:
```
from scipy import optimize
```

In [74]:
```
# Function
fx = lambda x: (x[0] - x[1])**2 + (x[1] + x[2] - 2)**2 + (x[3] - 1)**2 + (x[4] -1)**
```

In [75]:
```
# Constraints
c = ({'type': 'eq', 'fun': lambda x:  x[0] + 3 * x[1]},
{'type': 'eq', 'fun': lambda x:  x[2] + x[3] - 2 * x[4]},
{'type': 'eq', 'fun': lambda x:  x[1] - x[4]})

# Upper and Lower Bounds
b = ((-10, 10),(-10, 10),(-10, 10),(-10, 10),(-10, 10))

# Results
r = optimize.minimize(fx, (0.5,2,1,-0.5,1), method='SLSQP', bounds=b,constraints=c)

# Minimal Value of Function
r.fun
```

Out[75]:
```
4.093023255813957
```

In [76]:
```
# X
r.x
```

Out[76]:
```
array([-0.76744188,  0.25581396,  0.62790694, -0.11627902,  0.25581396])
```

Problem 2

x & $b \in R^n$, $A \in R^{n\times n}$

$f: R^n \to R$ as $f(x) = b^T x + x^T A x$

(a) Gradient & Hessian

Let $f(x) = f(a) + f(b)$ — ①

$f(a) = b^T x$

$f(b) = x^T A x$

for $f(a) = b^T x$

$$f(a) = b^T x = \sum_{j=1}^{n} b_j x_j$$

$$\frac{\partial f(a)}{dx} = \nabla f(a) = \begin{bmatrix} \frac{\partial b_1 x_1}{\partial x} \\ \vdots \\ \frac{\partial b_n x_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} =$$

$\therefore \nabla f(a) = b$ — ②

$$H(a) = \frac{\partial^2 f(a)}{\partial x^2} = \frac{\partial b^T x}{\partial x} = 0 \quad - ③$$

for $f(b) = x^T A x$

$$f(b) = x^T A x = \sum_{j=1}^{n} \sum_{i=1}^{n} x_i a_{ij} x_j$$

$$= \sum_{i=1}^{n} \left(a_{ii} x_i^2 + \sum_{j \neq i} x_i a_{ij} x_j \right)$$

$$\nabla f(b) = \frac{\partial f(b)}{\partial x} = \begin{bmatrix} \sum_{j=1}^{n} x_j a_{j1} \\ \vdots \\ \sum_{j=1}^{n} x_j a_{jn} \end{bmatrix} + \begin{bmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj} x_j \end{bmatrix}$$

$$\therefore \nabla F(b) = \underline{(A^T + A)x}$$

$$H(b) = \frac{\partial^2 f(b)}{\partial x} = A + A^T$$

$\therefore$ we have

gradient $(g(x))$ hessian $(H(x))$

$$\underline{g(x) = b + (A + A^T)x} \qquad \underline{H(x) = A + A^T}$$

if system is symmetric i.e $A = A^T$

$g(x) = b + 2Ax$

$H(x) = 2A$

(b) Taylor approximation at $x = 0$

1st order: $f(x) \approx f(x_0) + \nabla_x f\big|_{x_0}^{n} (x - x_0)$

2nd order: $f(x) \approx f(x_0) + \nabla_x F\big|_{x_0}^{n} (x - x_0) + \frac{1}{2}(x - x_0)^T H\big|_{x_0} (x - x_0)$

$\because f(x) = b^T x + x^T A x$

for $x = 0$

$f(0) = 0$

$g(x) = (A + A^T)x + b$

$g(0) = b$

$H(x) = A + A^T$

$H(0) = A + A^T$

$\therefore$ 1^{st} order

$$f(x) = f(0) + \nabla_x \Big|_{x_0} (x - x_0)$$

$$= 0 + b(x - 0) = bx$$

$$\therefore f(x) = b^T x$$

$\therefore$ 2^{nd} order

$$f(x) = f(0) + \nabla_x \Big|_{x_0} (x - x_0) + \frac{1}{2} (x - x_0)^T H \Big|_{x_0} (x - x_0)$$

$$= 0 + b(x - 0) + \frac{1}{2} (A + A^T)(x - 0)^2 = bx + \frac{(A + A^T)x^2}{2}$$

$$\therefore f(x) = b^T x + \frac{1}{2} x^T (A + A^T) x$$

The approximations are not exact.

(c) A to be +ve definite

$A \in \mathbb{R}^{n \times n}$, $A = A^T$

$\lambda > 0$ means positive definite

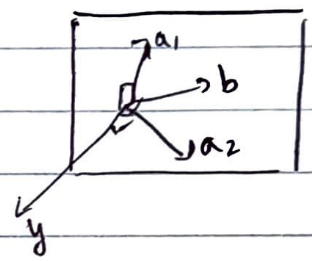
The eigen value has to be greater than 0.

(d) A full rank when

- each of the rows are linearly independent => full row rank
- each of the columns are linearly independent => column full rank
- determinant is nonzero. $|A| \neq 0$

(e) $y \in \mathbb{R}^n$ $y \neq 0$ $A^T y = 0$
conditions on b for $Ax = b$ for x to have a sol^n.

- y is also prependicular to b (b is orthogonal to y)
- b should be in column space of A



Problem 3

Let x_i be the purchased amount of food type $i = 1, 2, 3 \ldots, N$
Let the function be $f(x) = A^T x$
The total nuitrion j in purchased food $= \sum_{i=1}^{N} a_{ij} x_i \quad \forall j = 1, 2, 3 \ldots, M$

$\therefore$ optimization problem : $\min_{x} A' x \quad$ s.t $\sum_{i=1}^{N} a_{ij} x_i \geq b_j \quad \forall j = 1, 2, 3 \ldots, M$

or s.t $Ax \geq b$
where

$A = \begin{bmatrix} a_{11} & \cdots & a_{N1} \\ \vdots & \ddots & \vdots \\ a_{1M} & \cdots & a_{NM} \end{bmatrix} \in \mathbb{R}^{M \times N}$, $b = \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix} \in \mathbb{R}^M$, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N$