SSY281 Assignment 3

Abdulrahman Hameshli

February 22, 2024

Contents

1 Question 1

1.1 A

In the realm of mathematics, a convex function possesses a unique trait: any straight line segment connecting two points on its graph lies entirely above or on the graph itself. Put simply, for a function to be convex, it must satisfy the condition:

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta)f(x_2)$$

Here, θ ranges from 0 to 1. If this holds true for any θ, the function is convex. In strict convexity, the inequality becomes strict.

To visualize, imagine two functions depicted in Figure 1: one convex and the other not. The rightmost function illustrates a violation of convexity, with a segment where some points fall beneath its graph.

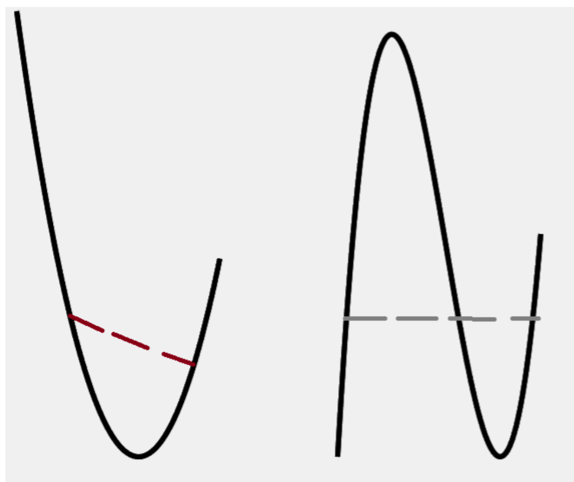


Figure 1: convex and non convex function

1.2 B

A set is deemed convex if it possesses the property of containing all the points along any line segment between two points within the set. This implies that any point lying between two others in the set must also be a part of the set. Formally, this can be expressed as follows:

For any $x_1, x_2 \in S$, where S represents the set, it follows that $\theta x_1 + (1 - \theta)x_2 \in S$ for $0 \leq \theta \leq 1$.

To visually understand this concept, Figure 2 illustrate convex sets to the left and non-convex set to the right.

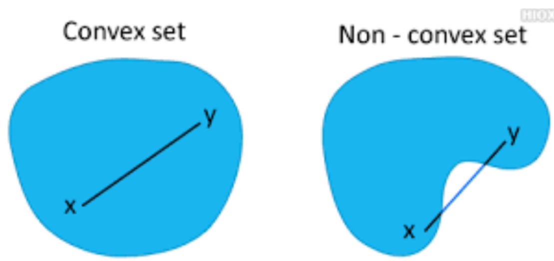


Figure 2: convex and non convex set

1.3 c

In the realm of optimization, an optimization problem transitions to convexity under the following conditions:

- The objective function $f(x)$ must exhibit convexity.
- The inequality constraints, as denoted by $g(x)$, must demonstrate convexity.
- The equality constraints, represented by $h(x)$, must adhere to affine properties.

2 Question 2

2.1 A

By partitioning the set into two distinct subsets, we obtain the following:

Subset S_1 is defined as the set of all x in the real numbers such that $\alpha \leq a^T x \leq \beta$, which can be expressed as the intersection of subsets S_{11} and S_{12} (Equation 5).

This division results in two new subsets:

Subset S_{11} consists of all x in the real numbers for which $a^T x \leq \beta$.

Subset S_{12} consists of all x in the real numbers for which $a^T x \geq \alpha$

Given that the intersection of convex sets is itself a convex set, it suffices to demonstrate the convexity of both S_{11} and S_{12}. It is noteworthy that these subsets are half-spaces, represented as $\{x \mid a^T x \leq b\}$. Using the definition of a convex set, it can be proven that half-spaces are indeed convex sets.

Consider two points x_1 and x_2 belonging to S_{11}. According to the definition:

$$\theta x_1 + (1-\theta)x_2 \leq \beta$$

which implies:

$$a^T(\theta x_1) + (1-\theta)a^T x_2 \leq \theta\beta + (1-\theta)\beta$$

This simplifies to:

$$a^T(\theta x_1) + (1-\theta)a^T x_2 \leq \beta$$

indicating that $a^T(\theta x_1) + (1-\theta)a^T x_2$ belongs to S_{11}. The same reasoning can be applied to S_{12}, confirming its convexity. Thus, both S_{11} and S_{12} are convex sets.

2.2 B

Absolutely, this assortment also possesses convex properties. The set $S_2 = \{x \mid ||x - y|| \leq f(y), \forall y \in S\}, S \in \mathbb{R}^n$ can be reformulated as:

$$S_2 = \{x \mid ||x - y|| - f(y) \leq 0, \forall y \in S\}, S \in \mathbb{R}^n$$

Using this representation, we can once again verify convexity by inspecting the function governing set S_2, represented as $S_2 = \{x \mid g(x) \leq 0\}$ where $g(x) = ||x - y|| - f(y)$. The set S_2 exhibits convexity if, for any pair of points x_1 and x_2, every intermediary point between $\theta x_1 + (1 - \theta)x_2$ also falls within the set. Assuming x_1 and x_2 belong to S_2, we find:

$$g(\theta x_1 + (1 - \theta)x_2) = ||(\theta x_1 + (1 - \theta)x_2) - y|| - f(y)$$

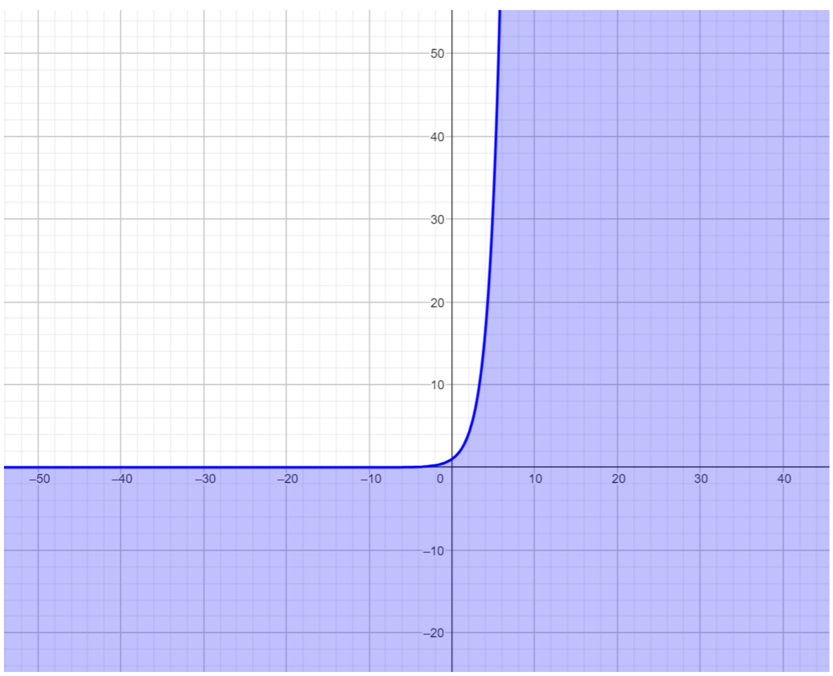
Here, we employ the triangle inequality, $||a + b|| \leq ||a|| + ||b||$, and expand the expression to:

$$g(\theta x_1 + (1 - \theta)x_2) \leq \theta(||x1 - y|| - f(y)) + (1 - \theta)(||x2 - y|| - f(y))$$

Thus, we've established that the function governing the points within the set is convex. Consequently, the set itself adheres to convexity principles.

2.3 C

set $S_3 = \{(x, y) \mid y \leq 2x, \forall (x, y) \in \mathbb{R}^2\}$ is not convex. The following pictures show the curve of the set :



We can clearly see that the curve is not convex for all (x,y) an example that not satisfy the convexity for example linear segment between two points (x1,y1) = (-20,0) and (x2,y2) = (3,7)

$$x = \theta x_1 + (1-\theta)x_2$$
$$y = \theta y_1 + (1-\theta)y_2$$

$$x = 2^{0.5\cdot -20+0.5\cdot 3} = 0.0028$$
$$y = 0.5\cdot 0 + 0.5\cdot 7 = 3.5$$

since the points are outside it means the set is concavew

3 Question 3

3.1 A

The reason the problems yield the same result is because both optimization tasks essentially aim to minimize the maximum absolute deviation between the vector Ax and b.

Problem (3), which seeks to minimize $\|Ax - b\|_\infty$, aims to find x that minimizes the largest deviation between the i-th component of Ax and b (infinity norm).

On the other hand, problem (5), which minimizes ϵ subject to $-\epsilon \leq (Ax-b)_i \leq \epsilon$, focuses on minimizing the deviation ϵ of the i-th component of $(Ax - b)_i$, ensuring that the deviation does not exceed the value of ϵ.

The optimal solution in problem (5) aligns with the optimal solution of the infinity norm of $Ax - b$, which essentially constitutes optimization problem (3). Hence, the two problems are equivalent.

3.2 B

Let's consider a scenario where the vector $x^\top$ has a length of $n \times 1$. In this scenario, the combined vector $z^\top$ will have a length of $n + 1$. Our objective is to minimize ϵ in the optimization problem. We can achieve this by setting the vector $c^\top$ to consist of n zeros and a 1 at the last position (position $n + 1$), like so: $c^\top = [0, 0, \ldots, 0, 1]$. Consequently, we obtain:

$$\min[x^\top \epsilon] \quad c^\top [x^\top \epsilon]^\top = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ \epsilon \end{bmatrix} = \epsilon$$

With the given constrain F and g are:

$$Axi - \epsilon \leq b \quad \text{and} \quad -Axi - \epsilon \leq -b$$

Expressing them together in a matrix format yields:

$$\underbrace{\begin{bmatrix} A & -1 \\ -A & -1 \end{bmatrix}}_{F_i} \underbrace{\begin{bmatrix} x_i \\ \epsilon \end{bmatrix}}_{z_i} \leq \underbrace{\begin{bmatrix} b \\ -b \end{bmatrix}}_{g_i}$$

Given that we have n number of x, the matrix needs to be enlarged considerably:

$$
\begin{bmatrix}
A & 0 & 0 & \cdots & 0 & -1 \\
-A & 0 & 0 & \cdots & 0 & -1 \\
0 & A & 0 & \cdots & 0 & -1 \\
0 & -A & 0 & \cdots & 0 & -1 \\
0 & 0 & A & \cdots & 0 & -1 \\
0 & 0 & -A & \cdots & 0 & -1 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & A & -1 \\
0 & 0 & 0 & \cdots & -A & -1
\end{bmatrix}
\begin{Bmatrix}
x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \\ \epsilon
\end{Bmatrix}
\leq
\begin{Bmatrix}
b \\ -b \\ b \\ \vdots \\ b \\ -b
\end{Bmatrix}
$$

In short, $Fz \leq g$ and it's important to note that the 0s and -1s in the above expression must match the dimensions of A.

3.3 C

Solving the linear programs by using MATLAB given A and B matrices:

$$\min_{x,\epsilon} \ \epsilon \qquad \underbrace{\begin{bmatrix} A & -1 \\ -A & -1 \end{bmatrix}}_{F_i} \underbrace{\begin{bmatrix} x_i \\ \epsilon \end{bmatrix}}_{z_i} \leq \underbrace{\begin{bmatrix} b \\ -b \end{bmatrix}}_{g_i}$$

where F becomes:

$$\begin{bmatrix} \begin{bmatrix} 0.4889 & 0.2939 \\ 1.0347 & -0.7873 \\ 0.7269 & 0.8884 \\ -0.3034 & -1.1471 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \\ -\begin{bmatrix} 0.4889 & 0.2939 \\ 1.0347 & -0.7873 \\ 0.7269 & 0.8884 \\ -0.3034 & -1.1471 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \end{bmatrix}$$

And g becomes :

$$\begin{bmatrix} \begin{bmatrix} -1.0689 \\ -0.8095 \\ -2.9443 \\ 1.4384 \end{bmatrix} \\ -\begin{bmatrix} -1.0689 \\ -0.8095 \\ -2.9443 \\ 1.4384 \end{bmatrix} \end{bmatrix}$$

we obtaint:

$$\begin{bmatrix} \begin{bmatrix} 0.4889 & 0.2939 \\ 1.0347 & -0.7873 \\ 0.7269 & 0.8884 \\ -0.3034 & -1.1471 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \\ -\begin{bmatrix} 0.4889 & 0.2939 \\ 1.0347 & -0.7873 \\ 0.7269 & 0.8884 \\ -0.3034 & -1.1471 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \epsilon \end{bmatrix} \leq \begin{bmatrix} \begin{bmatrix} -1.0689 \\ -0.8095 \\ -2.9443 \\ 1.4384 \end{bmatrix} \\ -\begin{bmatrix} -1.0689 \\ -0.8095 \\ -2.9443 \\ 1.4384 \end{bmatrix} \end{bmatrix}$$

By solving the result becomes:

$$z^* = \begin{bmatrix} x_1^* \\ x_2^* \\ \epsilon^* \end{bmatrix} = \begin{bmatrix} -2.0674 \\ -1.1067 \\ 0.4583 \end{bmatrix}$$

3.4 D

We encounter the problem:

$$\min_{x,\epsilon} c^\top z \quad \text{s.t.} \quad Fz - h \leq 0$$

This problem can be reformulated. By incorporating the constraint into the objective function, we arrive at the Lagrangian:

$$\mathcal{L}(z, \mu, \lambda) = c^\top z + \mu^\top (Fz - h) + \lambda^\top 0$$

$$\mathcal{L}(z, \mu, \lambda) = c^\top z + \mu^\top (Fz - h)$$

$$q(\mu, \lambda) = \inf_z \mathcal{L}(z, \mu, \lambda)$$

where μ and λ are Lagrange multipliers.

We define the Dual function using this Lagrangian:

$$\max_{\mu} -h^\top \mu \quad \text{s.t.} \quad \mu \geq 0, \quad F^\top \mu + c^\top = 0$$

3.5 E

Using MATLAB to solve this linear program gives us the solution for μ as shown below :

$$\mu^* = \begin{bmatrix} 0.0000 \\ 0.0000 \\ 0.4095 \\ 0.4284 \\ 0.0000 \\ 0.1621 \\ 0.0000 \\ 0.0000 \end{bmatrix}$$

3.6 F

With calculated μ in 3.E we get the primal solution:

$$z^* = \begin{bmatrix} -2.0674 \\ -1.1067 \\ 0.4583 \end{bmatrix}$$

4 Question 4

4.1 A

The problem can be written as follow:

$$\min_{x,u} \; \frac{1}{2} \begin{bmatrix} x_1 & x_2 & u_0 & u_1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{Q} \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}}_{p^\top} \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix}$$

$$\text{s.t} \quad \underbrace{\begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{G} \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix} \leq \underbrace{\begin{bmatrix} -2.5 \\ 5 \\ 0.5 \\ 0.5 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}}_{h}$$

$$\underbrace{\begin{bmatrix} 1 & 0 & -b & 0 \\ -A & 1 & 0 & -b \end{bmatrix}}_{A_{eq}} \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix} = \underbrace{\begin{bmatrix} Ax_0 \\ 0 \end{bmatrix}}_{b_{eq}}$$

where initial condition is $X_0 = 1.5$ we get the following result:

$$x^* = \begin{bmatrix} x_1^* \\ x_2^* \\ u_0^* \\ u_1^* \end{bmatrix} = \begin{bmatrix} 2.5000 \\ 0.5000 \\ 1.9000 \\ -0.5000 \end{bmatrix}$$

4.2 B

The Karush-Kuhn-Tucker (KKT) conditions are as follows:

$$
\begin{aligned}
\nabla f(x^*) + \nabla g(x^*)\mu^* + \nabla h(x^*)\lambda^* &= 0 \\
\mu^* &\geq 0 \\
g(x^*) &\leq 0 \\
h(x^*) &= 0 \\
\mu_i^* g(x^*) &= 0
\end{aligned}
$$

According to the provided MATLAB script:

$$
\mu^* = \begin{bmatrix} 4.6000 \\ 0 \\ 0 \\ 0.0001 \\ 0 \\ 0.0000 \\ 0.0000 \\ 0 \end{bmatrix}
$$

$$
\lambda^* = \begin{bmatrix} 1.9000 \\ -0.5000 \end{bmatrix}
$$

This implies that the second condition is satisfied. To verify the third and fourth conditions, we need to ensure that the inequalities and equalities hold at x^*. Upon analyzing the solution, it's evident that the inequalities hold for every state, thus $g(x^*) \leq 0$, meaning the third condition holds. Now, let's check the equality:

$$
x^* = \begin{bmatrix} 2.5000 \\ 0.5000 \\ 1.9000 \\ -0.5000 \end{bmatrix}
$$

$$
\begin{bmatrix} 1 & 0 & -B & 0 \\ -A & 1 & 0 & -B \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ u_1^* \\ u_2^* \end{bmatrix} = \begin{bmatrix} Ax^0 \\ 0 \end{bmatrix}
$$

$$
\begin{bmatrix} 2.5 - 1.9 \\ -0.4 \cdot 2.5 + 0.5 + 0.5 \end{bmatrix} = \begin{bmatrix} 0.4 \cdot 1.5 \\ 0 \end{bmatrix}
$$

Since the equality constraints hold true as well, $h(x^*) = 0$, and the fourth condition is also satisfied.

It can be easily verified that $\mu_i^* g(x^*) = 0$, meaning the fifth condition holds true as well.

The first condition can be confirmed by:

$$
\nabla f(x^*) = x^* \nabla g(x^*)\mu^* = A^T \text{ inequality } \mu^* \nabla h(x^*)\lambda^* = A^T \text{ equality } \lambda^*
$$

$$\nabla f(x^*) = x^*$$
$$\nabla g(x^*)\mu^* = A_{inequality}^T \mu^*$$
$$\nabla h(x^*)\lambda^* = A_{equality}^T \lambda^*$$

Which yields:

$$\begin{bmatrix} 2.5000 \\ 0.5000 \\ 1.9000 \\ -0.5000 \end{bmatrix} + \begin{bmatrix} -4.6000 \\ 0.0001 \\ 0.0000 \\ -0.0000 \end{bmatrix} + \begin{bmatrix} 2.1000 \\ -0.5000 \\ -1.9000 \\ 0.5000 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Hence, the fifth condition is true.

Analyzing μ^*, it's evident that the first and fourth elements are not 0, indicating that these constraints are active, i.e., x_1 is at the lower bound and x_2 is at the upper bound.

4.3 C

If the lower boundary on x_1 is lifted, meaning x_1 is no longer restricted to 5, the outcome of the `quadprog` algorithm changes:

$$x^* = \begin{bmatrix} 0.2885 \\ 0.0577 \\ -0.3115 \\ -0.0577 \end{bmatrix}$$

However, if the upper boundary is lifted, i.e., x_1 is allowed to exceed 2.5, the `quadprog` result becomes:

$$x^* = \begin{bmatrix} 2.5000 \\ 0.5000 \\ 1.9000 \\ -0.5000 \end{bmatrix}$$

Removing the restriction on x_1 alters the outcome significantly. Without the upper bound, x_1 achieves its maximum possible value, which in turn influences other constraints differently. Consequently, the solution improves compared to the scenario where x_1 is capped at 2.5.