ESE 619 Model Predictive Control — **Homework 2: Optimization**

University of Pennsylvania, Spring 2023 — updated February 13, 2023

Student Name: Wei-Cheng (Rory) Huang — *Penn ID: 55353592*

The solutions of this homework are entirely my own. I have discussed these problems with several classmates, they are: Shiming Liang, Yifan Xue and Yifei Shao

1. **Properties of Sets and Functions**

   Let $x_1, x_2 \in \mathbb{R}^n$ be in the feasible set $\mathcal{X}$ of an optimization problrm, i.e.

$$x_1, x_2 \in \mathcal{X} = \{x \in \mathbb{R}^n \,|\, g_i\left(x\right) \leq 0,\, i = 1, \ldots, m,\, h_j\left(x\right) = 0,\, j = 1, \ldots, q\}$$

(1) Can you find simple conditions on functions $g_i\left(x\right),\, i = 1, \ldots, m$, and $h_j\left(x\right),\, j = 1, \ldots, q$ such that the feasible set $\mathcal{X}$ is convex. [2 pts]

Solution: The condition for the feasible set $\mathcal{X}$ to be convex is: $g_i\left(x\right),\, i = 1, \ldots, m$, are all convex and $h_j\left(x\right),\, j = 1, \ldots, q$ are all affine

(2) let $z = \theta x_1 + \left(1 - \theta\right) x_2, \theta \in [0, 1]$ be any point on the line segment between points x_1 and x_2. Can you find conditions on function $f\left(x\right)$ such that 'z is better than the worst of x_1 and x_2', i.e.;

$$f\left(z\right) \leq \max\left\{f\left(x_1\right), f\left(x_2\right)\right\}$$

If yes, can it happen that point z is better than both of them? Try to sketch such a situation. [5 pts]

Solution: if f is quasi-convex, then it satisfies $f\left(z\right) \leq \max\left\{f\left(x_1\right), f\left(x_2\right)\right\}$. A simple case where point z is better than both of x_1 and x_2 is when $f\left(x_1\right) = f\left(x_2\right)$ as the following figure shows

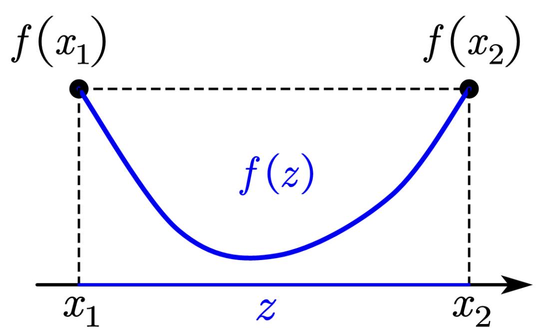

Figure 1: Simple sketch of 'z is better than both of x_1 and x_2'

(3) Which property of the feasible set $\mathcal{X}$ would ensure that point $y = x_1 + x_2$ is feasible, given $x_1, x_2 \in \mathcal{X}$. [3 pts]

Solution: if the feasible set $\mathcal{X}$ is a convex cone, then it would ensure that point $y = x_1 + x_2$ is feasible, given $x_1, x_2 \in \mathcal{X}$

2. **Checking Convexity of Sets**

Which of the following sets are convex? Give reasons for your answers.

(1) A slab, i.e. the set $\{x \in \mathbb{R} \mid \alpha \leq a^{\mathrm{T}} x \leq \beta\}$ where $a \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ [3 pts]
Solution: The slab is a convex set, beacuse it is the Intersection of two halfspaces. Halfspace is a convex set and intersection is convexity-preserved.

(2) Let $s : \mathbb{R}^n \to \mathbb{R}$ be any function with domain $\operatorname{dom}(s) \subseteq \mathbb{R}^n$. Let $\mathcal{S}$ be any subset of $\operatorname{dom}(s)$, though not necessarily a convex one. Is the set $\mathcal{M}$ defined as:

$$\mathcal{M} := \{x \mid \|x - y\| \leq s(y) \text{ for all } y \in \mathcal{S}\}$$

a convex set? Note that $\|\cdot\|$ denotes any norm on $\mathbb{R}^n$ [4 pts]
Solution: the set $\mathcal{M}$ is a convex set. We can show this through convexity-preserved operations, note that:

$$M := \{x \mid \|x - y\| \leq s(y) \text{ for all } y \in \mathcal{S}\} \Leftrightarrow M := \bigcap_{y \in \mathcal{S}} \{x \mid \|x - y\| \leq s(y)\}$$

And for a fixed y, we no that $\{x \mid \|x - y\| \leq s(y)\}$ is a norm ball and the norm ball is convex for any norm on $\mathbb{R}^n$. Therefore, by convexity-preserved property of intersection operation, the set $\mathcal{M}$ is convex.

(3) Consider s and $\mathcal{S}$ as above, again $\|\cdot\|$ denotes any norm on $\mathbb{R}^n$. We consider the set $\tilde{\mathcal{M}}$ defined as:

$$\tilde{\mathcal{M}} := \{x \mid \exists y \in \mathcal{S} \text{ s.t. } \|x - y\| \leq s(y)\}$$

Is $\tilde{\mathcal{M}}$ is a convex set, assuming that s is a convex function? [4 pts]
Solution: the set $\tilde{\mathcal{M}}$ is generally not a convex set, even assuming that s is a convex function, we can consider it from the aspect of convexity-preserved operations, note that:

$$\tilde{\mathcal{M}} := \{x \mid \exists y \in \mathcal{S} \ s.t. \ \|x - y\| \leq s(y)\} \Leftrightarrow \tilde{\mathcal{M}} = \bigcup_{y \in S} \{x \mid \|x - y\| \leq s(y)\}$$

Although for fixed y, $\{x \mid \|x - y\| \leq s(y)\}$ is a norm ball and the norm ball is convex for any norm on $\mathbb{R}^n$, the union of covex sets is not necessarily convex. In fact, we can quickly construct an non-convex example of this. Assume $s(y) = 1$ and $\operatorname{dom}(s) = \mathbb{R}^2$. Take $S = \left\{ \begin{bmatrix} 1 & 0 \end{bmatrix}^{\mathrm{T}}, \begin{bmatrix} -1 & 0 \end{bmatrix}^{\mathrm{T}} \right\}$ which is a subset of $\operatorname{dom}(s)$. Then the sketch of set $\tilde{\mathcal{M}}$ can be shown as the dumbbel-shape grey region below:

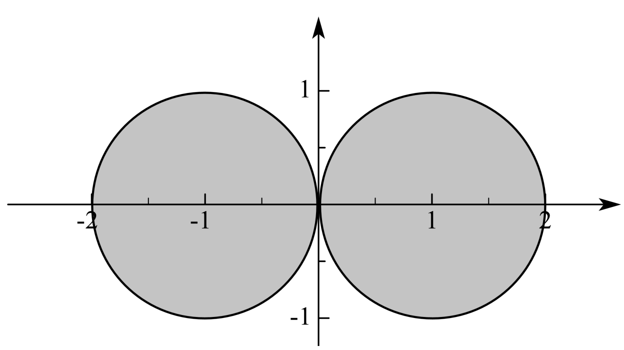

Figure 2: Simple illustration of a certain kind of $\tilde{\mathcal{M}}$: nonconvex

And this set is obviously non-convex, so $\tilde{\mathcal{M}}$ in generaal is not convex.

(4) The set of points closer to a given point x_0 than to given set $\mathcal{Q} \subseteq \mathbb{R}^n$, i.e.

$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in \mathcal{Q}\}$$

where $\|\cdot\|_2$ denotes the standard 2-norm (also called Euclidean norm) [4 pts]
Solution: This set is convex and we can consider it from the aspect of convexity-preserved operations. Note that:

$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in \mathcal{Q}\} \Leftrightarrow \bigcap_{y \in \mathcal{Q}} \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}$$

For a single y, the set $\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}$ is a halfspace and hence is a convex set, which can be proved as below:

$$\begin{aligned}
\|x - x_0\|_2 \leq \|x - y\|_2 &\Leftrightarrow (x - x_0)^{\mathrm{T}} (x - x_0) \leq (x - y)^{\mathrm{T}} (x - y) \\
&\Leftrightarrow 2 (y - x_0)^{\mathrm{T}} x \leq y^{\mathrm{T}} y - x_0^{\mathrm{T}} x_0 \\
&\Leftrightarrow a^{\mathrm{T}} x \leq b, \qquad \text{where } a = 2 (y - x_0)^{\mathrm{T}}, b = y^{\mathrm{T}} y - x_0^{\mathrm{T}} x_0
\end{aligned}$$

Therefore, by convexity-preserved property of intersection operation, the set is convex.

3. **1-Norm, ∞-Norm**

Instead of the standard 2-norm, the 1-norm or the ∞-norm are sometimes used in the MPC cost function. In this exercise you should show that both a minimization problem with a 1-norm objective and an ∞-norm objective

$$\min_x \|Ax\|_p, \quad p \in \{1, \infty\}$$

can be recast as a linear program (LP)

$$\begin{aligned} &\min_y \; b^{\mathrm{T}} y \\ &\text{s.t. } Fy \leq g \end{aligned}$$

with vectors b, g and matrix F defined appropriately. [10 pts]

Solution:

1-norm case:

recall the definition of 1-norm, we can rewrite 1-norm minimization as the following form:

$$\min_{x\in\mathbb{R}^n} \|Ax\|_1 \Leftrightarrow \min_{x\in\mathbb{R}^n} \sum_{i=1}^{N} |(Ax)_i| \Leftrightarrow \min_{x\in\mathbb{R}^n} \sum_{i=1}^{N} \max\{(Ax)_i, -(Ax)_i\}$$

where $(Ax)_i$ represents the i-th component of vector Ax. Using the formulation of piece-wise affine minimization, by introducing the additional variable $t \in \mathbb{R}^{\mathrm{N}}$, we can write down the equivalent optimization problem as below:

$$\min_{x\in\mathbb{R}^n} \sum_{i=1}^{N} \max\{(Ax)_i, -(Ax)_i\} \Leftrightarrow \begin{aligned} &\min_{x\in\mathbb{R}^n, t_i\in\mathbb{R}} \sum_{i=1}^{N} t_i \\ &\text{s.t. } (Ax)_i \leq t_i \\ &\quad\; -(Ax)_i \leq t_i \end{aligned} \Leftrightarrow \begin{aligned} &\min_{x\in\mathbb{R}^n, t\in\mathbb{R}^N} \mathbf{1}_N^{\mathrm{T}} t \\ &\text{s.t. } Ax \leq t \\ &\quad\; -Ax \leq t \end{aligned}$$

Where $\mathbf{1}_N$ is the N dimensional all-1 column vector. Further, we stack x and t up, deonoting as $y = \begin{bmatrix} x^{\mathrm{T}} & t^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{n+m}$, to get the unified augmented form of the problem:

$$\begin{aligned} &\min_{y\in\mathbb{R}^{n+N}} \; b^{\mathrm{T}} y \\ &\text{s.t. } Fy \leq g \end{aligned}$$

And it is easy to get that:

$$b = \begin{bmatrix} \mathbf{0}_n \\ \mathbf{1}_N \end{bmatrix}, \quad g = \mathbf{0}_{2N}, \quad F = \begin{bmatrix} A & -I_N \\ -A & -I_N \end{bmatrix}$$

where I_N denotes that $N \times N$ identity matrix, and $\mathbf{0}_n$,$\mathbf{1}_N$,$\mathbf{0}_{2N}$ denotes the all 0 (or 1) column vector with dimension shown in the corresponding subscript.

∞-norm case:

recall the definition of ∞-norm, we can rewrite ∞-norm minimization as the following form:

$$\min_{x\in\mathbb{R}^n} \|Ax\|_\infty \Leftrightarrow \min_{x\in\mathbb{R}^n} \max\{(Ax)_1, -(Ax)_1, \ldots, (Ax)_{\mathrm{N}}, -(Ax)_{\mathrm{N}}\}$$

where $(Ax)_i$ represents the i-th component of vector Ax. Using the technique of piece-wise affine minimization, by introducing the additional variable $t \in \mathbb{R}$, we can write down the equivalent optimization problem as below:

$$\min_{x\in\mathbb{R}^n} \max\{(Ax)_1, -(Ax)_1, \ldots, (Ax)_{\mathrm{N}}, -(Ax)_{\mathrm{N}}\} \Leftrightarrow \begin{aligned} &\min_{x\in\mathbb{R}^n, t\in\mathbb{R}} t \\ &\text{s.t. } (Ax)_i \leq t \\ &\quad\; -(Ax)_i \leq t \end{aligned} \Leftrightarrow \begin{aligned} &\min_{x\in\mathbb{R}^n, t\in\mathbb{R}} t \\ &\text{s.t. } Ax \leq \mathbf{1}_N t \\ &\quad\; -Ax \leq \mathbf{1}_N t \end{aligned}$$

Where $\mathbf{1}_N$ is the N dimensional all-1 column vector. Further, we stack x and t up, deonoting as $y = \begin{bmatrix} x^{\mathrm{T}} & t \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{n+1}$, to get the unified augmented form of the problem:

$$
\begin{gathered}
\min_{y \in \mathbb{R}^{n+1}} \ b^{\mathrm{T}} y \\
\text{s.t. } Fy \leq g
\end{gathered}
$$

And it is easy to get that:

$$
b = \begin{bmatrix} \mathbf{0}_n \\ 1 \end{bmatrix}, \quad g = \mathbf{0}_{2N}, \quad F = \begin{bmatrix} A & -\mathbf{1}_N \\ -A & -\mathbf{1}_N \end{bmatrix}
$$

where $\mathbf{0}_n$,$\mathbf{1}_N$,$\mathbf{0}_{2N}$ denotes the all 0 (or 1) column vector with dimension shown in the corresponding subscript.

4. **Linear Regression with 1- and ∞-Norms**

We now use the ideas from Exercise 3 to solve a linear regression problem: Use the command `linprog` from MATLAB to solve the LP's stemming from a reformulation of the linear regression problems

$$\min_x \ \|Ax-b\|_p\,, \quad p\in\{1,\infty\}$$

where A is a random matrix and b is a random vector, e.g. created in MATLAB with the commands `A=rand(10,5)` and `b=rand(10,1)`. [10 pts]

Solution: The formulation of the problem can refer to Exercise 3, the only difference is for more general affine objective, g should be modified as $g = \begin{bmatrix} b^{\mathrm{T}} & -b^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ where b here is the affine term in $\|Ax-b\|$

We verify some numerical examples in MATLAB and the code is shown below, we also use `YAlMIP` to directly solve the problem and check the correctness of our method. In the numerical example, we take A be a random 5×3 matrix and $b\in\mathbb{R}^5$ be a radom vector. The following is a result for a certain case (set the random seed to be 0 to reproduce the same result) in this case:

$$A = \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 0.9058 & 0.2785 & 0.9705 \\ 0.1270 & 0.5469 & 0.9572 \\ 0.9134 & 0.9575 & 0.4854 \\ 0.6324 & 0.9649 & 0.8003 \end{bmatrix}, b = \begin{bmatrix} 0.1419 \\ 0.4218 \\ 0.9157 \\ 0.7922 \\ 0.9595 \end{bmatrix}$$

And we have the minimum and minimizer as the following:

$$x_1^* = \begin{bmatrix} -0.1692 \\ 0.8057 \\ 0.3613 \end{bmatrix}, J_1^* = 0.2950$$
$$x_\infty^* = \begin{bmatrix} -0.1165 \\ 0.8325 \\ 0.3998 \end{bmatrix}, J_\infty^* = 0.0926$$

```
%% This is the Matlab Script for ESE 619 HW2 Ex4
clear all
clc
% randomly generated matrices
% use this random seed if you want to reproduce my result
rng(0)
A = rand(5,3);
b = rand(5,1);
[N,n] = size(A);

%% 1-norm minimization
% The formulation refers to Ex3
% optimization variable is to stack x and t up, denote as y
F = [A,-eye(N);-A,-eye(N)];
g = [b;-b];
c = [zeros(n,1);ones(N,1)];
y = linprog(c,F,g);
x = y(1:n);
t = y(n+1:N+n);

% use yalmip to verify the solution
x_sym = sdpvar(n,1);
Objective_1norm = pnorm(A*x_sym-b,1);
Cons = [];
optimize(Cons,Objective_1norm);
solution = value(x_sym);
Flag = (norm(solution-x)<1e-10)
```

```

%% infinity-norm minimization
% The formulation refers to Ex3
% optimization variable is to stack x and t up, denote as y
F_inf = [A,-ones(N,1);-A,-ones(N,1)];
g_inf = [b;-b];
c_inf = [zeros(n,1);1];
y_inf = linprog(c_inf,F_inf,g_inf);
x_inf = y_inf(1:n);
t_inf = y_inf(n+1);

% use yalmip to verify the solution
x_sym_inf = sdpvar(n,1);
Objective_inf = pnorm(A*x_sym_inf-b,inf);
Cons = [];
optimize(Cons,Objective_inf);
solution_inf = value(x_sym_inf);
Flag_inf = (norm(solution_inf - x_inf)<1e-10)
```

5. **Quadratic Program** Consider the optimization problem:

$$\begin{aligned} \min \quad & \frac{1}{2}\left(x_1^2 + x_2^2 + 0.1x_3^2\right) + 0.55x_3 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 1 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \\ & x_3 \geq 0 \end{aligned}$$

1 Show that $x^* = (0.5, 0.5, 0)$ is a local minimum. [5 pts]
Solution:
Solve the Hessian of the function at x^*:

$$\nabla^2 f\left(x^*\right) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix}_{x=x^*} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix} \succ 0$$

According to the Hessian, we can see that the function is a convex function, obviously the feasible region is a convex region. Therefore, this optimization problem is convex.
Then, solve the gradient of the function:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0.1x_3 + 0.55 \end{bmatrix}$$

Consider any feasible y in the feasible region defined by the domain and constraints, we check:

$$\nabla f^{\mathrm{T}}\left(x^*\right)\left(y - x^*\right) = \begin{bmatrix} 0.5 & 0.5 & 0.55 \end{bmatrix} \begin{bmatrix} y_1 - 0.5 \\ y_2 - 0.5 \\ y_3 - 0 \end{bmatrix} = 0.5\left(y_1 + y_2\right) - 0.5 + 0.55y_3 = 0.05y_3 \geq 0$$

According to the first order optimality condition, point x^* is local minimum

2 Is x^* also a global minimum? Explain why, or why not. [5 pts]
Solution: For convex optimization problem, local minimum is global minimum.