SSY281 Model Predictive Control

Assignment A3

Qun Zhang(961109-6554)

February 19, 2023

Question 1

(a) **convex/strictly convex function**

A function f:$\mathbb{R}^n \rightarrow \mathbb{R}$ is convex if dom $f$is convex, and following the equation shown below:

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \tag{1}$$

For all $x, y \in$ dom f and $0 \leq \theta \leq 1$.

•f is concave if -f is convex

•f is strictly convex if

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y) \tag{2}$$

For all $x, y \in$ dom f and $0 < \theta < 1$.

(b) **convex set**

• A line through x_1 and x_2 are all points x,

$$x = \theta x_1 + (1-\theta)x_2, \quad \theta \in \mathbb{R}. \tag{3}$$

• An affine set contains the line through any two distinct points in the set. All affine sets can be described as solutions to a system of linear equations. For example,$\{x \mid Ax = b\}$.

• A line segment between x_1 and x_2 are all points x, which means:

$$x = \theta x_1 + (1-\theta)x_2, 0 \leq \theta \leq 1. \tag{4}$$

• A convex set contains the line segments between every two points in the set:

$$x_1, x_2 \in \mathcal{S} \Rightarrow \theta x_1 + (1-\theta)x_2 \in \mathcal{S}, \quad 0 \leq \theta \leq 1. \tag{5}$$

• A hyperplane is a set of the form $\left\{x \mid a^\top x = b\right\}$.

• A half-space is a set of the form $\left\{x \mid a^\top x \leq b\right\}$.

Hyperplanes are affine and convex; half-spaces are convex.

• A polyhedron is the intersection of a finite number of half-spaces and hyperplanes or, equivalently, the solution set of a finite number of linear inequalities and equalities:

$$Ax \leq b \tag{6}$$

$$Cx = d. \tag{7}$$

Polyhedra are convex sets.

(c) **Under what conditions, (1) becomes a convex optimization problem**

$$\begin{aligned} &minimize\ f(x) \\ &subject\ to\ g(x) \le 0 \\ &\qquad\qquad h(x) = 0 \end{aligned} \tag{8}$$

where f and gi are convex and h is affine. Assume x^* is regular. Then x^* is globally optimal if and only if the KKT conditions are fulfilled for some $\mu^* \ge 0; \lambda$.

Question 2

(a)

$$S_1 = \{x \in \mathbb{R}^n | \alpha \le \alpha^T x \le \beta\} \tag{9}$$

$$x = \theta x_1 + (1-\theta)x_2, 0 \le \theta \le 1 \tag{10}$$

$$\begin{aligned} \alpha \le \alpha^T x_1 \le \beta \\ \alpha \le \alpha^T x_2 \le \beta \end{aligned} \tag{11}$$

$$\begin{aligned} \alpha^T &= \theta\alpha^T x_1 + (1-\theta)\alpha^T x_2 \\ &\ge \theta\alpha + (1-\theta)\alpha = \alpha \end{aligned} \tag{12}$$

$$\begin{aligned} \alpha^T x &= \theta\alpha^T x_1 + (1-\theta)\alpha^T x_2 \\ &\le \theta\beta + (1-\theta)\beta = \beta \end{aligned} \tag{13}$$

Then, we get that $\alpha \le \alpha^T(\theta x_1 + (1-\theta)x_2) \le \beta$, which satisfy the requirement of convex set. Thus, S_1 is a convex set.

(b)

$$S_2 = \{x|\ \|x-y\| \le f(y), for\, all\, y \in S\} \tag{14}$$

$$x = \theta x_1 + (1-\theta)x_2, 0 \le \theta \le 1 \tag{15}$$

$$\begin{aligned} \|x-y\| &= \|\theta + (1-\theta)x_2 - y\| \\ &= \|\theta(x_1 - y) + (1-\theta)(x_2 - y)\| \\ &\le \|\theta(x_1 - y)\| + \|(1-\theta)(x_2 - y)\| \\ &\le \theta f(y) + (1-\theta) f(y) = f(y) \end{aligned} \tag{16}$$

Thus, we get that $\|(\theta x_1 + (1-\theta)x_2) - y\| \le f(y)$, which satisfy the requirement of convex. Thus, S_2 is a convex set.

(c)

$$S_3 = \{(x, y)|y \leq 2^x, for\, all\, (x, y) \in \mathbb{R}^2\} \tag{17}$$

As is shown in the figure 1, where the shaded part is $y \leq 2^x$. We choose two points inside the set and draw a line between two points. According to the definition, all points on the line should be inside the set. However, from the picture, not every point on the line inside the set. Thus, **the set is not a convex**.

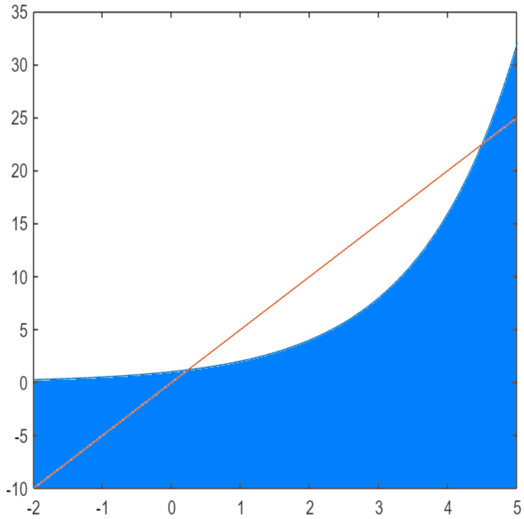

Figure 1: plot of $y \leq 2^x$

Question 3

(a) Considering $y \in R$, $|y|$ can be defined as smallest non negative value ϵ.

$$\begin{aligned} \|y\| = &\min_{\epsilon} \epsilon \\ \text{s.t. } &\epsilon \geq y \\ &\epsilon \geq -y \end{aligned} \tag{18}$$

$$\begin{aligned} &\min_{\epsilon, x} \epsilon \\ \text{s.t. } &\Sigma \geq (Ax - b)_i, \forall i \in (1, 2, ...n) \\ &\Sigma \geq -(Ax - b)_i, \forall i \in (1, 2, ...n) \end{aligned} \tag{19}$$

Where,$\Sigma = ones(size(A,1),1) * \epsilon$.

Then, the equation could be rewritten in the equation shown below:

$$\begin{aligned} &\min_{\epsilon,x} \epsilon \\ \text{s.t.} \quad & -\Sigma \leq (Ax-b)_i \leq \Sigma, \forall i \in (1,2,...n) \end{aligned} \tag{20}$$

(b) Since $Z^T = [x^T \ \epsilon]$, the equation could be rewritten in the form shown below:

$$\begin{aligned} &\min_{\epsilon,x} [0 \ 0 \ 1] Z \\ \text{s.t.} \quad & \underbrace{\begin{bmatrix} A & -\mathbf{1} \\ -A & -\mathbf{1} \end{bmatrix}}_{F} \underbrace{\begin{bmatrix} x \\ \epsilon \end{bmatrix}}_{Z} \leq \underbrace{\begin{bmatrix} b \\ -b \end{bmatrix}}_{g} \end{aligned} \tag{21}$$

Where, **1** is ones(size(A,1),1).

(c) Taking the matrix of A, b and F, g in to the calculation, the optimal value is:

$$Z = \begin{bmatrix} -2.0674 \\ -1.1067 \\ 0.4583 \end{bmatrix} \tag{22}$$

Thus, the minium value of the object function is **0.4583**.

(d) The dual could be achieved from Lagrangian equation:

$$\begin{aligned} \mathbf{L}(x,\mu,\lambda) &= f(Z) + \mu^T g(Z) \\ &= f(Z) + \mu^T (FZ - g) \\ &= -\mu^T g + (C^T + \mu^T F)Z \end{aligned} \tag{23}$$

$$\begin{aligned} q(\mu,\lambda) &= \inf_Z \mathbf{L}(x,\mu,\lambda) \\ &= \inf_Z (-\mu^T + (C^T + \mu^T F)Z) \end{aligned} \tag{24}$$

since $\frac{\partial L}{\partial z} = C^T + \mu^T F = 0$, the equation could be rewritten below:

$$\begin{aligned} &\max_{\mu} q^T = -g^T \mu \\ \text{s.t.} \quad & \mu \geq 0 \\ & F^T \mu + C = 0 \end{aligned} \tag{25}$$

(e) Taking the characters into dual equation we derived above and using matlab function "linprog", then we can get the values of μ and ϵ:

$$\mu = \begin{bmatrix} 0 \\ 0 \\ 0.4095 \\ 0.4284 \\ 0 \\ 0.1621 \\ 0 \\ 0 \end{bmatrix}, \epsilon = 0.4583 \tag{26}$$

(f) According to the complementary slackness condition of dual, we should use the characters where $\mu > 0$ to calculate the Z we want:

$$\mu_i > 0 \rightarrow |AZ|_i = b_i \tag{27}$$

Then, the value of Z should be the same as what we get from the primal

$$Z = \begin{bmatrix} -2.0674 \\ -1.1067 \\ 0.4583 \end{bmatrix} \tag{28}$$

Question 4

(a) According to the equation, we can view that $f(x) = 0.5Z^T HZ$,and the the Z and H should be:

$$Z = \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix}, H = eye(4) \tag{29}$$

And the constraints matrix of inequality and equality should be shown below:

$$Ain = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, bin = \begin{bmatrix} -2.5000 \\ 5.0000 \\ 0.5000 \\ 0.5000 \\ 2.0000 \\ 2.0000 \\ 2.0000 \\ 2.0000 \end{bmatrix} \tag{30}$$

Additionally, the Aeq and beq matrix is shown below:

$$Aeq = \begin{bmatrix} 1.0000 & 0 & -1.0000 & 0 \\ 0.4000 & -1.0000 & 0 & 1.0000 \end{bmatrix}, beq = \begin{bmatrix} 0.6 \\ 0 \end{bmatrix} \tag{31}$$

Taking into the quadprog, the value of Z and fval are:

$$Z = \begin{bmatrix} 2.5000 \\ 0.5000 \\ 1.9000 \\ -0.5000 \end{bmatrix}, fval = \begin{bmatrix} 5.18 \end{bmatrix} \tag{32}$$

(b) The KKT condition is:

$$\begin{gathered} \nabla f\left(x^*\right) + \nabla g\left(x^*\right)\mu^* + \nabla h\left(x^*\right)\lambda^* = 0 \\ \mu^* \geq 0 \\ g\left(x^*\right) \leq 0, \quad h\left(x^*\right) = 0 \\ \mu_i^* g_i\left(x^*\right) = 0, \quad i = 1, \ldots, m \end{gathered} \tag{33}$$

Then the KKT condition of question(a) is shown below:

$$KKT_1 = 10^{-10} \begin{bmatrix} -0.1095 \\ 0.2739 \\ 0.0702 \\ -0.1095 \end{bmatrix}, KKT_2 = \underbrace{\begin{bmatrix} 4.6000 \\ 0 \\ 0 \\ 0.0001 \\ 0 \\ 0.0000 \\ 0.0000 \\ 0 \end{bmatrix}}_{\mu} \tag{34}$$

$$KKT_{31} = \underbrace{\begin{bmatrix} -0.0000 \\ -2.5000 \\ -1.0000 \\ -0.0000 \\ -3.9000 \\ -0.1000 \\ -1.5000 \\ -2.5000 \end{bmatrix}}_{g(x^*)}, KKT_{32} = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{H}, KKT_4 = 10^{-8} \begin{bmatrix} -0.0027 \\ 0 \\ 0 \\ -0.2318 \\ 0 \\ -0.0027 \\ -0.0027 \\ 0 \end{bmatrix} \tag{35}$$

KKT condition holds at solution of question(a). At the optimal solution, μ_1=4.6, $\mu_4 = 0.0001$, which is not equal to zero. So, the constraints that $X_1 \geq 2.5$ and $x_2 \leq 0.5$ are active at optimal solution.

(c) If we remove the lower boundary of x_1, the optimal solution changes;

$$Z = \begin{bmatrix} 0.2885 \\ 0.0577 \\ -0.3115 \\ -0.0577 \end{bmatrix}, fval = \begin{bmatrix} 0.0935 \end{bmatrix} \tag{36}$$

However, if we change the upper boundary of the x_1, the optimal value will be the same:

$$Z = \begin{bmatrix} 2.5000 \\ 0.5000 \\ 1.9000 \\ -0.5000 \end{bmatrix}, fval = \begin{bmatrix} 5.18 \end{bmatrix} \tag{37}$$

Since we change the lower boundary, we changed the optimal active constraints, so the change of the solution is reasonable. If we change the upper boundary, since constraints related to the upper boundary ($\mu = 0$)is inactive, that will not affect our values in question (a).