Linear and Convex Optimization Homework 02

Qiu Yihang, 2021/10/03-10/05

1. ***Proof:***

By the definition, $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in f^{-1}(C), \forall \theta \in [0,1],\ \exists \boldsymbol{y}_1, \boldsymbol{y}_2 \in C$ **s.t.** $\boldsymbol{y}_1 = f(\boldsymbol{x}_1), \boldsymbol{y}_2 = f(\boldsymbol{x}_2)$.

Let $\bar{\theta} = 1 - \theta$. Since C is convex, $\theta \boldsymbol{y}_1 + \bar{\theta} \boldsymbol{y}_2 \in C$.

Since $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$ is an affine function,

$$f(\theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2) = \boldsymbol{A}(\theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2) + \boldsymbol{b} = \theta \boldsymbol{A}\boldsymbol{x}_1 + \theta \boldsymbol{b} + \bar{\theta} \boldsymbol{A}\boldsymbol{x}_2 + \bar{\theta} \boldsymbol{b} = \theta \boldsymbol{y}_1 + \bar{\theta} \boldsymbol{y}_2 \in C,$$

$$\textbf{i.e. } \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2 \in f^{-1}(C).$$

Thus, $f^{-1}(C)$ is also convex. *Qed.* ■

2. ***Proof:***

First we prove $\boldsymbol{0} \notin C$ by contradiction. If $\boldsymbol{0} \in C$, by definition we know $\exists \boldsymbol{x}_1 \in C_1, \boldsymbol{x}_2 \in C_2$ **s.t.** $\boldsymbol{x}_1 - \boldsymbol{x}_2 = \boldsymbol{0}$, **i.e.** $C_1 \cap C_2 = \{\boldsymbol{x}_1\} \neq \emptyset$. Contradiction.

Thus, $\boldsymbol{0} \notin C$.

Then we prove C is a nonempty set.

Since C_1 and C_2 are both nonempty sets, there exist at least one $\boldsymbol{x}_1 \in C_1$ and one $\boldsymbol{x}_2 \in C_2$. By definition, we have $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in C$. Thus, C is a nonempty set.

Now we prove C is a convex set.

$\forall \boldsymbol{x}, \boldsymbol{y} \in C, \forall \theta \in [0,1]$, by definition, $\exists \boldsymbol{x}_1, \boldsymbol{y}_1 \in C_1,\ \boldsymbol{x}_2, \boldsymbol{y}_2 \in C_2$ **s.t.** $\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2, \boldsymbol{y} = \boldsymbol{y}_1 - \boldsymbol{y}_2$.

Let $\bar{\theta} = 1 - \theta$.

Since C_1 and C_2 are both convex sets, $\boldsymbol{z_1} \triangleq \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{y}_1 \in C_1,\ \boldsymbol{z_2} \triangleq \theta \boldsymbol{x}_2 + \bar{\theta} \boldsymbol{y}_2 \in C_2$.

Thus, $\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y} = \theta \boldsymbol{x}_1 - \theta \boldsymbol{x}_2 + \bar{\theta} \boldsymbol{y}_1 - \bar{\theta} \boldsymbol{y}_2 = (\theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{y}_1) - (\theta \boldsymbol{x}_2 + \bar{\theta} \boldsymbol{y}_2) = \boldsymbol{z_1} - \boldsymbol{z_2} \in C$.

Therefore, C is a convex set.

In conclusion, C is a nonempty convex set and $\boldsymbol{0} \notin C$.

Qed. ■

3. (a) ***Proof:***

$\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \textbf{int}\, C,\ \forall \theta \in [0,1]$, Since $\textbf{int}\, C \subset C,\ \boldsymbol{x}_1, \boldsymbol{x}_2 \in C$.

Since C is convex, $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in C, \forall \theta \in [0,1]$, let $\bar{\theta} = 1 - \theta,\ \boldsymbol{y} \triangleq \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2 \in C$.

Since $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \textbf{int}\, C,\ \exists\, \varepsilon > 0$ **s.t.** $B(\boldsymbol{x}_1, \varepsilon) \subset C, B(\boldsymbol{x}_2, \varepsilon) \subset C$.

$\forall \boldsymbol{z} \in B(\boldsymbol{y}, \varepsilon),\ \|\boldsymbol{z} - \boldsymbol{y}\| < \varepsilon$,

$$\textbf{i.e. } \boldsymbol{z} - \boldsymbol{y} + \boldsymbol{x}_1 \in B(\boldsymbol{x}_1, \varepsilon) \subset C,\ \boldsymbol{z} - \boldsymbol{y} + \boldsymbol{x}_2 \in B(\boldsymbol{x}_2, \varepsilon) \subset C,$$

$$\boldsymbol{z} = \boldsymbol{y} + (\boldsymbol{z} - \boldsymbol{y}) = \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2 + (\theta + \bar{\theta})(\boldsymbol{z} - \boldsymbol{y}) = \theta(\boldsymbol{z} - \boldsymbol{y} + \boldsymbol{x}_1) + \bar{\theta}(\boldsymbol{z} - \boldsymbol{y} + \boldsymbol{x}_2) \in C.$$

Thus, $B(\boldsymbol{y}, \varepsilon) \subset C$, **i.e.** $\boldsymbol{y} \in \textbf{int}\, C$.

In other words, $\textbf{int}\, C$ is convex.

Qed. ■

(b) ***Proof:***

$\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \bar{C}, \forall \theta \in [0,1]$, let $\bar{\theta} = 1 - \theta, \boldsymbol{y} = \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2, \partial C = \bar{C} \backslash \textbf{int}\, C \triangleq \{\boldsymbol{x}: \boldsymbol{x} \in \bar{C}, \boldsymbol{x} \notin \textbf{int}\, C\}$.

There are three cases:

CASE 1. $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \textbf{int}\, C$. Since C is convex, from **(a)** we know $\boldsymbol{y} \in \textbf{int}\, C \subset \bar{C}$.

CASE 2. $\boldsymbol{x}_i \in \textbf{int}\, C, \boldsymbol{x}_j \in \partial C, \{i, j\} = \{1,2\}$. Let $\boldsymbol{x}_1 \in C, \boldsymbol{x}_2 \in \partial C$.

(1) When $\theta = 0$, $\boldsymbol{y} = \boldsymbol{x}_2 \in \bar{C}$.

(2) When $\theta \in [0,1)$, since $\boldsymbol{x}_2 \in \partial C$, for given $\varepsilon > 0$, we can always find $\widetilde{\boldsymbol{x}}_2 \in B(\boldsymbol{x}_2, \varepsilon)$ **s.t.** $\widetilde{\boldsymbol{x}}_2 \in \textbf{int}\, C$. We have $\|\widetilde{\boldsymbol{x}}_2 - \boldsymbol{x}_2\| < \varepsilon$.

From **(a)** we know $\theta \boldsymbol{x}_1 + \bar{\theta} \widetilde{\boldsymbol{x}}_2 \in \textbf{int}\, C$ and $B(\theta \boldsymbol{x}_1 + \bar{\theta} \widetilde{\boldsymbol{x}}_2, \varepsilon) \subset C \subset \bar{C}$.

Meanwhile, $\boldsymbol{y} = \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2 = (\theta \boldsymbol{x}_1 + \bar{\theta} \widetilde{\boldsymbol{x}}_2) + \bar{\theta}(\boldsymbol{x}_2 - \widetilde{\boldsymbol{x}}_2) \in B(\theta \boldsymbol{x}_1 + \bar{\theta} \widetilde{\boldsymbol{x}}_2, \theta \varepsilon)$

$\subset B(\theta \boldsymbol{x}_1 + \bar{\theta} \widetilde{\boldsymbol{x}}_2, \varepsilon) \subset \bar{C}$.

CASE 3. $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \partial C$.

(1) When $\theta = 0$, $\boldsymbol{y} = \boldsymbol{x}_2 \in \bar{C}$.

(2) When $\theta = 1$, $\boldsymbol{y} = \boldsymbol{x}_1 \in \bar{C}$.

(3) When $\theta \in (0,1)$, since $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \partial C$, for given $\varepsilon > 0$,

we can always find $\widetilde{\boldsymbol{x}}_1 \in B(\boldsymbol{x}_1, \varepsilon), \widetilde{\boldsymbol{x}}_2 \in B(\boldsymbol{x}_2, \varepsilon)$ **s.t.** $\widetilde{\boldsymbol{x}}_1, \widetilde{\boldsymbol{x}}_2 \in \textbf{int}\, C$.

We have $\|\widetilde{\boldsymbol{x}}_1 - \boldsymbol{x}_1\| < \varepsilon, \|\widetilde{\boldsymbol{x}}_2 - \boldsymbol{x}_2\| < \varepsilon$.

From **(a)** we know $\theta \widetilde{\boldsymbol{x}}_1 + \bar{\theta} \widetilde{\boldsymbol{x}}_2 \in \textbf{int}\, C$ and $B(\theta \widetilde{\boldsymbol{x}}_1 + \bar{\theta} \widetilde{\boldsymbol{x}}_2, \varepsilon) \subset C \subset \bar{C}$.

Meanwhile, $\boldsymbol{y} = \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2 = (\theta \widetilde{\boldsymbol{x}}_1 + \bar{\theta} \widetilde{\boldsymbol{x}}_2) + \bar{\theta}(\boldsymbol{x}_2 - \widetilde{\boldsymbol{x}}_2) + \theta(\boldsymbol{x}_1 - \widetilde{\boldsymbol{x}}_1)$.

Since $\left\|\bar{\theta}(\boldsymbol{x}_2 - \widetilde{\boldsymbol{x}}_2) + \theta(\boldsymbol{x}_1 - \widetilde{\boldsymbol{x}}_1)\right\| \leq \theta \|\widetilde{\boldsymbol{x}}_1 - x_1\| + \bar{\theta}\|\widetilde{\boldsymbol{x}}_2 - x_2\| \leq \theta \varepsilon + \bar{\theta} \varepsilon \leq \varepsilon$,

$\boldsymbol{y} \in B(\theta \widetilde{\boldsymbol{x}}_1 + \bar{\theta} \widetilde{\boldsymbol{x}}_2, \varepsilon) \subset \bar{C}$.

Thus, $\boldsymbol{y} \in \bar{C}$.

In other words, $\bar{C}$ is convex.

Qed. ■

4. (a) ***Proof:***

$\forall \boldsymbol{y}_1, \boldsymbol{y}_2 \in C$, by definition we know $\exists \varphi_1, \ldots, \varphi_m, \mu_1, \ldots, \mu_m$ **s.t.**

$$\sum_{i=1}^{m} \varphi_i \boldsymbol{x}_i = \boldsymbol{y}_1, \sum_{i=1}^{m} \mu_i \boldsymbol{x}_i = \boldsymbol{y}_2, \sum_{i=1}^{m} \varphi_i = 1, \sum_{i=1}^{m} \mu_i = 1, \varphi_i \geq 0, \mu_i \geq 0 \; (i = 1,2, \ldots, m).$$

$\forall \theta \in [0,1]$, let $\bar{\theta} = 1 - \theta$,

$$\theta \boldsymbol{y}_1 + \bar{\theta} \boldsymbol{y}_2 = \sum_{i=1}^{m} \theta \varphi_i \boldsymbol{x}_i + \sum_{i=1}^{m} \bar{\theta} \mu_i \boldsymbol{x}_i = \sum_{i=1}^{m} (\theta \varphi_i + \bar{\theta} \mu_i) \boldsymbol{x}_i \,,$$

$$\sum_{i=1}^{m} \theta \varphi_i + \bar{\theta} \mu_i = \theta \sum_{i=1}^{m} \varphi_i + \bar{\theta} \sum_{i=1}^{m} \mu_i = \theta + \bar{\theta} = 1 \,,$$

$$\theta \varphi_i + \bar{\theta} \mu_i \geq 0 \; (i = 1,2, \ldots, m).$$

Thus, $\theta \boldsymbol{y}_1 + \bar{\theta} \boldsymbol{y}_2 \in C$.

In other words, C is convex.

Qed. ■

(b) ***Proof:***

First we prove $C \subset \textbf{conv}\, S$ by contradiction.

If there exists $\boldsymbol{x} \in C$ **s.t.** $\boldsymbol{x} \notin \textbf{conv}$ S, by definition we know $\boldsymbol{x} = \sum_{i=1}^{m} \theta_i \boldsymbol{x}_i$.

Meanwhile, since $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in \textbf{conv}\, S$ and $\textbf{conv}\, S$ is convex, by theorem, convex combination $\boldsymbol{x} \in \textbf{conv}\, S$. (The theorem will be proved below.)

Contradiction.

Thus, $\forall \boldsymbol{x} \in \boldsymbol{C}, \boldsymbol{x} \in \textbf{conv}\, S$, **i.e.** $C \subset \textbf{conv}\, S$.

By definition we know $S \subset C$ (let $\theta_i = \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases}$ $(i \in \{1, \dots, m\})$, we get $\boldsymbol{x}_k \in C$).

Since $\textbf{conv}\, S$ is the smallest convex set containing S (by the definition of convex hull) and C is a convex set containing S, we know $\textbf{conv}\, S \subset C$.

Therefore, $C = \textbf{conv}\, S$.

Qed. ■

(The theorem given in the ppt and used above can be proved as follows.)

Thm. If C is convex and $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in C$, any convex combination $\boldsymbol{x} = \sum_{i=1}^{m} \theta_i \boldsymbol{x}_i \in C$.

Proof. Prove the theorem by induction.

We prove

$$\boldsymbol{y}_n = \frac{\sum_{i=1}^{n} \theta_i \boldsymbol{x}_i}{\sum_{i=1}^{n} \theta_i} \in C \ (n \in \{1, \dots, m\}).$$

BASE STEP. When $n = 1$, obviously $\boldsymbol{y}_1 = \boldsymbol{x}_1 \in C$.

When $n = 2$, obviously $\boldsymbol{y}_2 \in C$ (by the definition of convex sets).

INDUCTIVE STEP.

Suppose when $n = k < m$, $\boldsymbol{y}_n \in C$.

Let $\bar{\theta}_{k+1} = 1 - \theta_{k+1}$, then $\bar{\theta}_{k+1} \geq 0, \theta_{k+1} \geq 0$

(by the definition of convex combination).

Since $\frac{\sum_{i=1}^{k} \theta_i}{\sum_{i=1}^{k+1} \theta_i} + \frac{\theta_{k+1}}{\sum_{i=1}^{k+1} \theta_i} = 1$, by the definition of convex sets, we have

$$\frac{\sum_{i=1}^{k} \theta_i}{\sum_{i=1}^{k+1} \theta_i} \boldsymbol{y}_k + \frac{\theta_{k+1}}{\sum_{i=1}^{k+1} \theta_i} \boldsymbol{x}_{k+1} \in C.$$

i.e.

$$\boldsymbol{y}_{k+1} = \frac{\sum_{i=1}^{k+1} \theta_i \boldsymbol{x}_i}{\sum_{i=1}^{k+1} \theta_i} = \frac{\sum_{i=1}^{k} \theta_i \boldsymbol{x}_i + \theta_{k+1} \boldsymbol{x}_{k+1}}{\sum_{i=1}^{k+1} \theta_i} = \frac{\sum_{i=1}^{k} \theta_i}{\sum_{i=1}^{k+1} \theta_i} \frac{\sum_{i=1}^{k} \theta_i \boldsymbol{x}_i}{\sum_{i=1}^{k} \theta_i} + \frac{\theta_{k+1}}{\sum_{i=1}^{k+1} \theta_i} \boldsymbol{x}_{k+1}$$

$$= \frac{\sum_{i=1}^{k} \theta_i}{\sum_{i=1}^{k+1} \theta_i} \boldsymbol{y}_k + \frac{\theta_{k+1}}{\sum_{i=1}^{k+1} \theta_i} \boldsymbol{x}_{k+1} \in C$$

Thus, $\boldsymbol{y}_n \in C$ still holds when $n = k + 1$ $(n \leq m)$.

Therefore, $\boldsymbol{y}_m \in C$.

Since $\sum_{i=1}^{m} \theta_i = 1$ (by the definition of convex combination), we have

$$\boldsymbol{x} = \sum_{i=1}^{m} \theta_i \boldsymbol{x}_i = \frac{\sum_{i=1}^{m} \theta_i \boldsymbol{x}_i}{\sum_{i=1}^{m} \theta_i} = \boldsymbol{y}_m \in C.$$

□

5. ***Proof:***

Consider the case of $\|\boldsymbol{x}-\boldsymbol{x_0}\|_2 \le \|\boldsymbol{x}-\boldsymbol{x}_i\|_2$.

Let $\boldsymbol{x}=(y_1,\dots,y_n), \boldsymbol{x}_i=\left(x_{i_1},\dots,x_{i_n}\right)$ $(i=0,1,\dots,K)$.

$$\|\boldsymbol{x}-\boldsymbol{x_0}\|_2 \le \|\boldsymbol{x}-\boldsymbol{x}_i\|_2 \Leftrightarrow \sqrt{\sum_{j=1}^{n}\left(y_j-x_{0_j}\right)^2} \le \sqrt{\sum_{j=1}^{n}\left(y_j-x_{i_j}\right)^2}$$

可以尝试 $\|x-x_0\|_2^2 = (x-x_0)^T(x-x_0)$.

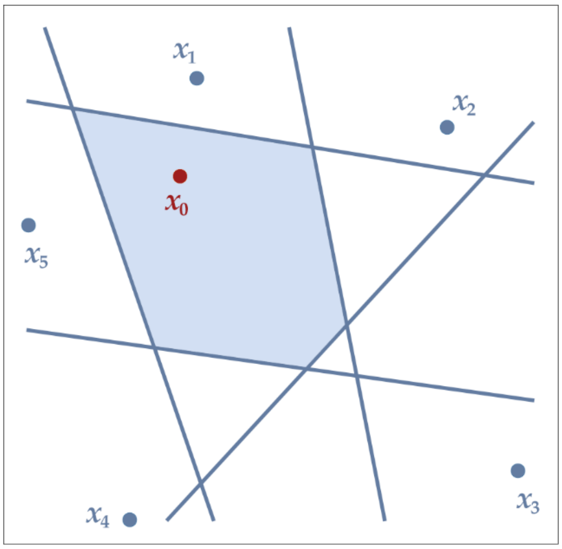
$$\Leftrightarrow 2\sum_{j=1}^{n}\left(x_{i_j}-x_{0_j}\right)y_j \le 2\sum_{j=1}^{n}\left(x_{i_j}^2-x_{0_j}^2\right) \Leftrightarrow 2(\boldsymbol{x}_i-\boldsymbol{x_0})^T\boldsymbol{x} \le \boldsymbol{x}_i^T\boldsymbol{x}_i-\boldsymbol{x}_0^T\boldsymbol{x}_0$$

Thus, we can find

$$\boldsymbol{A}=2\cdot\begin{pmatrix}(\boldsymbol{x}_1-\boldsymbol{x_0})^T\\ \vdots \\ (\boldsymbol{x}_K-\boldsymbol{x_0})^T\end{pmatrix}, \boldsymbol{b}=\begin{pmatrix}\boldsymbol{x}_1^T\boldsymbol{x}_1-\boldsymbol{x}_0^T\boldsymbol{x}_0\\ \vdots \\ \boldsymbol{x}_K^T\boldsymbol{x}_K-\boldsymbol{x}_0^T\boldsymbol{x}_0\end{pmatrix}$$

s.t. $V=\{\boldsymbol{x}: \boldsymbol{A}\boldsymbol{x}\le\boldsymbol{b}\}$, **i.e.** V is a polyhedron.

A visualization of V when $n=2$ is as follows.



Qed. ■