553.665 Introduction to Convexity, Fall 2022
Section 10

Hongyu Cheng

November 18, 2022

Definition 1 (Cone of feasible directions)**.** Let $C \subseteq \mathbb{R}^d$ be a convex set, and let $\mathbf{x} \in C$. Define *the cone of feasible directions* as

$$F_C(\mathbf{x}) = \{\mathbf{r} \in \mathbb{R}^d : \exists \varepsilon > 0 \text{ such that } \mathbf{x} + \varepsilon \mathbf{r} \in C\}.$$

Exercise 1. Let $C \subseteq \mathbb{R}^d$ and $\mathbf{x} \in C$, show that $F_C(\mathbf{x})$ is a convex cone.

Proof. $\forall \mathbf{r}^1, \mathbf{r}^2 \in F_C(\mathbf{x})$, by definition there exist $\varepsilon_1, \varepsilon_2 > 0$ such that $\mathbf{x} + \varepsilon_1 \mathbf{r}^1, \mathbf{x} + \varepsilon_2 \mathbf{r}^2 \in C$. For any $\lambda, \gamma \geq 0$, notice that

$$\mathbf{x} + \frac{\varepsilon_1}{2\lambda}(2\lambda \mathbf{r}^1),\ \mathbf{x} + \frac{\varepsilon_2}{2\gamma}(2\gamma \mathbf{r}^2) \in C,$$

so we have that $2\lambda\mathbf{r}^1, 2\gamma\mathbf{r}^2 \in F_C(\mathbf{x})$. Let $\varepsilon_3 = \min\{\frac{\varepsilon_1}{2\lambda}, \frac{\varepsilon_2}{2\gamma}\}$, by convexity of C one can obtain that $\mathbf{x} + \varepsilon_3(2\lambda\mathbf{r}^1), \mathbf{x} + \varepsilon_3(2\gamma\mathbf{r}^2) \in C$. Again by convexity, we have

$$\mathbf{x} + \varepsilon_3(\lambda\mathbf{r}^1 + \gamma\mathbf{r}^2) = \frac{1}{2}(\mathbf{x} + \varepsilon_3(2\lambda\mathbf{r}^1)) + \frac{1}{2}(\mathbf{x} + \varepsilon_3(2\gamma\mathbf{r}^2)) \in C,$$

so $\lambda\mathbf{r}^1 + \gamma\mathbf{r}^2 \in F_C(\mathbf{x})$, which proves $F_C(\mathbf{x})$ is a convex cone. □

Remark 1. $F_C(\mathbf{x})$ may not be closed: consider $C = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\}$, and let $\mathbf{x} = (-1, 0)$. Then $F_C(\mathbf{x}) = \{\mathbf{r} \in \mathbb{R}^2 : \mathbf{r}_1 > 0\} \cup \{\mathbf{0}\}$.

Exercise 2. Let $P \subseteq \mathbb{R}^d$ be a polyhedron given by $P = \{\mathbf{x} \in \mathbb{R}^d : A\mathbf{x} \leq \mathbf{b}\}$. Let $\mathbf{a}^i,\ i = 1, \dots, m$ be the rows of A. For any $\bar{\mathbf{x}} \in P$, define $\text{tight}(\bar{\mathbf{x}}) = \{i : \langle \mathbf{a}^i, \mathbf{x} \rangle = \mathbf{b}_i\}$. Show that

$$F_P(\bar{\mathbf{x}}) = \{\mathbf{r} \in \mathbb{R}^d : \langle \mathbf{a}^i, \mathbf{r} \rangle \leq 0 \text{ for all } i \in \text{tight}(\bar{\mathbf{x}})\}.$$

Proof. The set $\{\mathbf{r} \in \mathbb{R}^d : \langle \mathbf{a}^i, \mathbf{r} \rangle \leq 0 \text{ for all } i \in \text{tight}(\bar{\mathbf{x}})\} \subseteq F_P(\bar{\mathbf{x}})$, since
Case 1: $i \in \text{tight}(\bar{\mathbf{x}})$. For all $\varepsilon > 0$, $\langle \mathbf{a}^i, \mathbf{x} + \varepsilon\mathbf{r} \rangle = \langle \mathbf{a}^i, \mathbf{x} \rangle + \varepsilon\langle \mathbf{a}^i, \mathbf{r} \rangle \leq \langle \mathbf{a}^i, \mathbf{x} \rangle = \mathbf{b}_i$.
Case 2: $i \notin \text{tight}(\bar{\mathbf{x}})$. $\langle \mathbf{a}^i, \mathbf{x} \rangle < \mathbf{b}_i$, once $\varepsilon > 0$ is small enough, the inequality $\langle \mathbf{a}^i, \mathbf{x} + \varepsilon\mathbf{r} \rangle \leq \mathbf{b}_i$ will hold.

To show the reverse direction, consider any $\mathbf{r} \in \mathbb{R}^d$ has $\langle \mathbf{a}^i, \mathbf{r} \rangle > 0$ for some $i \in \text{tight}(\bar{\mathbf{x}})$, then for any $\varepsilon > 0$ we have

$$\langle \mathbf{a}^i, \bar{\mathbf{x}} + \varepsilon\mathbf{r} \rangle = \langle \mathbf{a}^i, \bar{\mathbf{x}} \rangle + \varepsilon\langle \mathbf{a}^i, \mathbf{r} \rangle > \mathbf{b}_i,$$

so $\mathbf{r} \notin F_C(\bar{\mathbf{x}})$. □

Exercise 3. Let $C \subseteq \mathbb{R}^d$ be a nonempty, closed, convex set. Then the following are all true:

1. Let $\mathbf{y} \in \mathbb{R}^d \backslash C$, then $\text{Proj}_C(\mathbf{y}) = \mathbf{x}$ if and only if $\mathbf{y} - \mathbf{x} \in N_C(\mathbf{x})$.
2. If $\mathbf{x} \in \text{int}\,(C)$, then $N_C(\mathbf{x}) = \{0\}$.
3. If $\mathbf{x} \in \text{bd}\,(C)$, then $\{0\} \subsetneq N_C(\mathbf{x})$.

Proof.

1. ($\Longrightarrow$): Notice that

$$\begin{array}{rcl} & \langle \mathbf{y}-\mathbf{x}, \mathbf{z}\rangle \le \langle \mathbf{y}-\mathbf{x}, \mathbf{x}\rangle \le \langle \mathbf{y}-\mathbf{x}, \mathbf{y}\rangle, & \forall \mathbf{z} \in C, \\ \Longrightarrow & \langle \mathbf{y}-\mathbf{x}, \mathbf{z}-\mathbf{x}\rangle \le 0, & \forall \mathbf{z} \in C, \\ \Longrightarrow & \mathbf{y}-\mathbf{x} \in N_C(\mathbf{x}). & \end{array}$$

   ($\Longleftarrow$): Since $\mathbf{y} - \mathbf{x} \in N_C(\mathbf{x})$, by definition of normal cone,

$$\begin{array}{rcl} & \langle \mathbf{y}-\mathbf{x}, \mathbf{z}-\mathbf{x}\rangle \le 0, & \forall \mathbf{z} \in C, \\ \Longrightarrow & \langle \mathbf{y}-\mathbf{x}, \mathbf{y}-\mathbf{x}\rangle + \langle \mathbf{y}-\mathbf{x}, \mathbf{z}-\mathbf{y}\rangle \le 0, & \forall \mathbf{z} \in C, \\ \Longrightarrow & \|\mathbf{y}-\mathbf{x}\|^2 \le (\mathbf{y}-\mathbf{x})^T(\mathbf{y}-\mathbf{z}) \le \|\mathbf{y}-\mathbf{x}\|\|\mathbf{y}-\mathbf{z}\|, & \forall \mathbf{z} \in C, \\ \Longrightarrow & \|\mathbf{y}-\mathbf{x}\| \le \|\mathbf{y}-\mathbf{z}\|, & \forall \mathbf{z} \in C, \\ \Longrightarrow & \mathbf{x} = \text{Proj}_C(\mathbf{y}). & \end{array}$$

2. $\forall \mathbf{r} \in N_C(\mathbf{x})$, there exists $\varepsilon > 0$ such that $\mathbf{x} + \varepsilon\mathbf{r} \in C$ since $\mathbf{x} \in \text{int}\,(C)$. Then

$$\begin{array}{rc} & \langle \mathbf{r}, (\mathbf{x}+\varepsilon\mathbf{r}) - \mathbf{x}\rangle \le 0, \\ \Longrightarrow & \varepsilon\langle \mathbf{r}, \mathbf{r}\rangle \le 0, \\ \Longrightarrow & \mathbf{r} = 0, \end{array}$$

   which implies $N_C(\mathbf{x}) = \{0\}$.

3. By Lemma 2.3.4 in lecture notes, there exists $\mathbf{y} \in \mathbb{R}^d \backslash C$ such that $\text{Proj}_C(\mathbf{y}) = \mathbf{x}$. Then part 1 implies $\mathbf{y} - \mathbf{x} \in N_C(\mathbf{x})$. Therefore, $\{0\} \subsetneq \{0, \mathbf{y} - \mathbf{x}\} \subseteq N_C(\mathbf{x})$.

□

Exercise 4. Consider the following standard form polyhedron in $\mathbb{R}^d$, defined by some $A \in \mathbb{R}^{m \times d}$, $\mathbf{b} \in \mathbb{R}^m$:

$$P = \{\mathbf{x} \in \mathbb{R}^d : A\mathbf{x} = \mathbf{b}, \mathbf{x} \ge 0\}.$$

1. Prove that every $\bar{\mathbf{x}} \in P$ has $N_P(\bar{\mathbf{x}}) = \{-(\mathbf{s} + A^T\mathbf{y}) : (\mathbf{s}, \mathbf{y}) \in \mathbb{R}^d \times \mathbb{R}^m, \mathbf{s} \ge 0, \mathbf{s}_i = 0 \text{ for } i \in I(\bar{\mathbf{x}})\}$, where $I(\bar{\mathbf{x}}) = \{i : \bar{\mathbf{x}}_i > 0\}$.
2. Prove that if $-\mathbf{c} \in \text{int}\,(N_P(\bar{\mathbf{x}}))$ for some $\bar{\mathbf{x}} \in P$, then $\bar{\mathbf{x}}$ is the unique minimizer of $\langle \mathbf{c}, \cdot \rangle$ over P.

Proof. 1. For any $\mathbf{c} \in \mathbb{R}^d$, the function $f(\mathbf{x}) = \langle \mathbf{c}, \mathbf{x} \rangle$ is convex and differentiable with $\nabla f(\mathbf{x}) = \mathbf{c}$, then

$$\begin{aligned}
-\mathbf{c} \in N_P(\bar{\mathbf{x}}) &\iff \bar{\mathbf{x}} \text{ minimizes } \langle \mathbf{c}, \mathbf{x} \rangle \text{ over } P, \\
&\iff \exists \mathbf{y} \in \mathbb{R}^m \text{ s.t. } A^T\mathbf{y} \leq \mathbf{c},\ \langle \mathbf{b}, \mathbf{y} \rangle = \langle \mathbf{c}, \bar{\mathbf{x}} \rangle, \\
&\iff \exists \mathbf{y} \in \mathbb{R}^m \text{ s.t. } A^T\mathbf{y} \leq \mathbf{c},\ \langle A\bar{\mathbf{x}}, \mathbf{y} \rangle = \langle \mathbf{c}, \bar{\mathbf{x}} \rangle, \\
&\iff \exists \mathbf{y} \in \mathbb{R}^m \text{ s.t. } A^T\mathbf{y} \leq \mathbf{c},\ \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y}) = 0, \\
&\iff \exists \mathbf{y} \in \mathbb{R}^m, \mathbf{s} \geq 0 \text{ s.t. } \mathbf{s} = \mathbf{c} - A^T\mathbf{y},\ \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y}) = 0, \\
&\iff \exists \mathbf{y} \in \mathbb{R}^m, \mathbf{s} \geq 0 \text{ s.t. } \mathbf{c} = \mathbf{s} + A^T\mathbf{y},\ \mathbf{x}^T\mathbf{s} = 0, \\
&\iff -\mathbf{c} \in \{-(\mathbf{s} + A^T\mathbf{y}) : \mathbf{y} \in \mathbb{R}^m, \mathbf{s} \geq 0, \langle \bar{\mathbf{x}}, \mathbf{s} \rangle = 0\}.
\end{aligned}$$

2. Suppose to the contrary both $\bar{\mathbf{x}}$ and $\bar{\mathbf{x}}'$ minimize the objective $\langle \mathbf{c}, \mathbf{x} \rangle$ but $\bar{\mathbf{x}} \neq \bar{\mathbf{x}}'$. Since $c \in \text{int}(N_P(\bar{\mathbf{x}}))$, for small enough $\varepsilon > 0$, $-c - \varepsilon(\bar{\mathbf{x}} - \bar{\mathbf{x}}') \in N_P(\bar{\mathbf{x}})$. This implies $\bar{\mathbf{x}}$ minimizes $\langle \mathbf{c} + \varepsilon(\bar{\mathbf{x}} - \bar{\mathbf{x}}'), \mathbf{x} \rangle$ over P. However, this is contradicted by $\bar{\mathbf{x}}' \in P$ as it has

$$\begin{aligned}
\langle \mathbf{c} + \varepsilon(\bar{\mathbf{x}} - \bar{\mathbf{x}}'), \bar{\mathbf{x}}' \rangle &= \langle \mathbf{c} + \varepsilon(\bar{\mathbf{x}} - \bar{\mathbf{x}}'), \bar{\mathbf{x}} \rangle + \langle \mathbf{c} + \varepsilon(\bar{\mathbf{x}} - \bar{\mathbf{x}}'), \bar{\mathbf{x}}' - \bar{\mathbf{x}} \rangle \\
&= \langle \mathbf{c} + \varepsilon(\bar{\mathbf{x}} - \bar{\mathbf{x}}'), \bar{\mathbf{x}} \rangle + \underbrace{\langle \mathbf{c}, \bar{\mathbf{x}}' \rangle - \langle \mathbf{c}, \bar{\mathbf{x}} \rangle}_{=0} - \varepsilon \underbrace{\left\| \bar{\mathbf{x}} - \bar{\mathbf{x}}' \right\|_2^2}_{>0} \\
&< \langle \mathbf{c} + \varepsilon(\bar{\mathbf{x}} - \bar{\mathbf{x}}'), \bar{\mathbf{x}} \rangle.
\end{aligned}$$

□