MA 5037: Optimization Methods and Applications
Chapter 6: Convex Sets

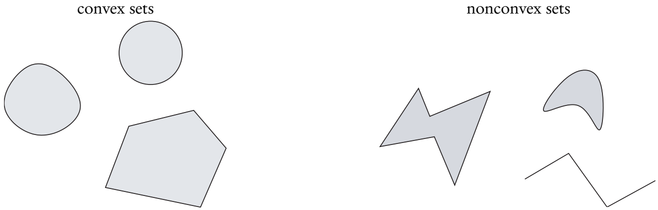
Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 32001, Taiwan

syyang@math.ncu.edu.tw
http://www.math.ncu.edu.tw/∼syyang/

Convex set

- **Definition:** *A set $C \subseteq \mathbb{R}^n$ is called convex if for any $\boldsymbol{x}, \boldsymbol{y} \in C$ and $\lambda \in [0,1]$, we have $\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in C$.*
- **Note 1:** C is convex $\Longleftrightarrow$ for any $\boldsymbol{x}, \boldsymbol{y} \in C$, the line segment $[\boldsymbol{x}, \boldsymbol{y}]$ is in C. i.e., $[\boldsymbol{x}, \boldsymbol{y}] \subseteq C$.
- **Note 2:** The empty set $\varnothing$ is a convex set. (\$ not! then $\exists$ ... $\rightarrow\leftarrow$)
- **Example:** A line in $\mathbb{R}^n$ is a set of the form, $L = \{\boldsymbol{z} + t\boldsymbol{d} : t \in \mathbb{R}\}$, where $\boldsymbol{z}, \boldsymbol{d} \in \mathbb{R}^n$. Let $\boldsymbol{x} = \boldsymbol{z} + t_1\boldsymbol{d} \in L$ and $\boldsymbol{y} = \boldsymbol{z} + t_2\boldsymbol{d} \in L$. Then for any $\lambda \in [0,1]$, $\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} = \boldsymbol{z} + (\lambda t_1 + (1-\lambda)t_2)\boldsymbol{d} \in L$. Therefore, L is a convex set.



Convexity of hyperplanes and half-spaces

- **Note 1:** For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, the closed and open line segments $[\boldsymbol{x}, \boldsymbol{y}]$ and $(\boldsymbol{x}, \boldsymbol{y})$ are convex sets.
- **Note 2:** The entire space $\mathbb{R}^n$ is a convex set.
- **Note 3:** *Let $\boldsymbol{a} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ and $b \in \mathbb{R}$. The following sets are convex:*
  (1) the hyperplane $H = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} = b\}$;
  (2) the half-space $H^- = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} \leq b\}$;
  (3) the open half-space $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} < b\}$.

  Proof of (2): Let $\boldsymbol{x}, \boldsymbol{y} \in H^-$ and $\lambda \in [0, 1]$. We will show that $\boldsymbol{z} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in H^-$. Indeed,

$$\begin{aligned} \boldsymbol{a}^\top \boldsymbol{z} &= \boldsymbol{a}^\top(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) = \lambda(\boldsymbol{a}^\top \boldsymbol{x}) + (1-\lambda)(\boldsymbol{a}^\top \boldsymbol{y}) \\ &\leq \lambda b + (1-\lambda) b = b, \end{aligned}$$

  which implies $\boldsymbol{z} \in H^-$. □

Convexity of balls

Let $\boldsymbol{c} \in \mathbb{R}^n$ and $r > 0$. Let $\|\cdot\|$ be an arbitrary norm defined on $\mathbb{R}^n$. Then the open ball $B(\boldsymbol{c}, r) := \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{c}\| < r\}$ and the closed ball $B[\boldsymbol{c}, r] := \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{c}\| \leq r\}$ are convex.

Proof: We will show the convexity of the closed ball. Let $\boldsymbol{x}, \boldsymbol{y} \in B[\boldsymbol{c}, r]$ and $\lambda \in [0, 1]$. Then $\|\boldsymbol{x} - \boldsymbol{c}\| \leq r$ and $\|\boldsymbol{y} - \boldsymbol{c}\| \leq r$. Let $\boldsymbol{z} = \lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}$. We will show that $\boldsymbol{z} \in B[\boldsymbol{c}, r]$. Indeed,

$$\begin{aligned}
\|\boldsymbol{z} - \boldsymbol{c}\| &= \|\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} - \boldsymbol{c}\| = \|\lambda(\boldsymbol{x} - \boldsymbol{c}) + (1 - \lambda)(\boldsymbol{y} - \boldsymbol{c})\| \\
&\leq \|\lambda(\boldsymbol{x} - \boldsymbol{c})\| + \|(1 - \lambda)(\boldsymbol{y} - \boldsymbol{c})\| \\
&= \lambda\|\boldsymbol{x} - \boldsymbol{c}\| + (1 - \lambda)\|\boldsymbol{y} - \boldsymbol{c}\| \\
&\leq \lambda r + (1 - \lambda) r \\
&= r.
\end{aligned}$$

Therefore $\boldsymbol{z} \in B[\boldsymbol{c}, r]$, establishing the result. □

Note: The above result is true for any norm defined on $\mathbb{R}^n$.

Convexity of ellipsoids

An ellipsoid is a set of the form

$$E = \{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}) := \boldsymbol{x}^\top Q\boldsymbol{x} + 2\boldsymbol{b}^\top \boldsymbol{x} + c \le 0\},$$

where $Q \in \mathbb{R}^{n\times n}$ is positive semidefinite, $\boldsymbol{b} \in \mathbb{R}^n$, and $c \in \mathbb{R}$. Then E is a convex set.

Proof: Let $\boldsymbol{x}, \boldsymbol{y} \in E$, $\lambda \in [0,1]$, and $\boldsymbol{z} := \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}$. Then $f(\boldsymbol{x}) \le 0, f(\boldsymbol{y}) \le 0$ and

$$\begin{aligned} \boldsymbol{z}^\top Q\boldsymbol{z} &= (\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y})^\top Q(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\ &= \lambda^2\boldsymbol{x}^\top Q\boldsymbol{x} + (1-\lambda)^2\boldsymbol{y}^\top Q\boldsymbol{y} + 2\lambda(1-\lambda)\boldsymbol{x}^\top Q\boldsymbol{y}. \end{aligned}$$

Since $\boldsymbol{x}^\top Q\boldsymbol{y} = (Q^{1/2}\boldsymbol{x})^\top(Q^{1/2}\boldsymbol{y})$, by the Cauchy-Schwarz inequality, we have

$$\boldsymbol{x}^\top Q\boldsymbol{y} \le \|Q^{1/2}\boldsymbol{x}\|\|Q^{1/2}\boldsymbol{y}\| = \sqrt{\boldsymbol{x}^\top Q\boldsymbol{x}}\sqrt{\boldsymbol{y}^\top Q\boldsymbol{y}} \le \frac{1}{2}(\boldsymbol{x}^\top Q\boldsymbol{x} + \boldsymbol{y}^\top Q\boldsymbol{y}).$$

Thus, $\boldsymbol{z}^\top Q\boldsymbol{z} \le \lambda\boldsymbol{x}^\top Q\boldsymbol{x} + (1-\lambda)\boldsymbol{y}^\top Q\boldsymbol{y}$. Hence,

$$\begin{aligned} f(\boldsymbol{z}) &\le \lambda\boldsymbol{x}^\top Q\boldsymbol{x} + (1-\lambda)\boldsymbol{y}^\top Q\boldsymbol{y} + 2\lambda\boldsymbol{b}^\top\boldsymbol{x} + 2(1-\lambda)\boldsymbol{b}^\top\boldsymbol{y} + c \\ &= \lambda(\boldsymbol{x}^\top Q\boldsymbol{x} + 2\boldsymbol{b}^\top\boldsymbol{x} + c) + (1-\lambda)(\boldsymbol{y}^\top Q\boldsymbol{y} + 2\boldsymbol{b}^\top\boldsymbol{y} + c) \\ &= \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{y}) \le 0, \end{aligned}$$

establishing the desired result that $\boldsymbol{z} \in E$. $\square$

Convexity is preserved under the intersection

- **Lemma:** *Let $C_i \subseteq \mathbb{R}^n$ be a convex set for any $i \in I$, where I is an arbitrary index set. Then $\cap_{i \in I} C_i$ is convex.*

  Proof: Let $\boldsymbol{x}, \boldsymbol{y} \in \cap_{i \in I} C_i$ and $\lambda \in [0, 1]$. Then $\boldsymbol{x}, \boldsymbol{y} \in C_i, \ \forall\, i \in I$. Since C_i is convex, it follows that $\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \in C_i, \ \forall\, i \in I$. Therefore, $\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \in \cap_{i \in I} C_i$. That is, $\cap_{i \in I} C_i$ is convex. □

- **Example** (convex polytopes): *A set P is called a convex polytope if it has the form $P = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$, where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$.* The convexity of P follows from the fact that it is an intersection of half-spaces and half-spaces are convex:

$$P = \bigcap_{i=1}^{m} \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}_i \boldsymbol{x} \leq \boldsymbol{b}_i\},$$

  where $\boldsymbol{A}_i$ is the ith row of $\boldsymbol{A}$.

Preservation of convexity

1. *Let $C_1, \cdots, C_k \subseteq \mathbb{R}^n$ be convex sets and let $\mu_1, \cdots, \mu_k \in \mathbb{R}$. Then the following set is convex:*

$$\mu_1 C_1 + \mu_2 C_2 + \cdots + \mu_k C_k := \left\{ \sum_{i=1}^{k} \mu_i \boldsymbol{x}_i : \ \boldsymbol{x}_i \in C_i, 1 \leq i \leq k \right\}$$

   Note: if $C \subseteq \mathbb{R}^n$ is a convex set and $\boldsymbol{b} \in \mathbb{R}^n$, then the set $C + \boldsymbol{b} := \{\boldsymbol{x} + \boldsymbol{b} : \boldsymbol{x} \in C\}$ is also convex.

2. *Let $C_i \subseteq \mathbb{R}^{k_i}$ be a convex set for any $i = 1, 2, \cdots, m$. Then the following Cartesian product is convex:*

$$C_1 \times C_2 \times \cdots \times C_m := \{(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m) : \boldsymbol{x}_i \in C_i, 1 \leq i \leq m\}$$

3. *Let $M \subseteq \mathbb{R}^n$ be a convex set and let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. Then the image set $\boldsymbol{A}(M) := \{\boldsymbol{A}\boldsymbol{x} : \boldsymbol{x} \in M\}$ is convex.*

4. *Let $D \subseteq \mathbb{R}^m$ be a convex set and let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. Then the inverse image set, $\boldsymbol{A}^{-1}(D) := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \in D\}$, is convex.*

Convex combinations

- **Definition:** *Given $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in \mathbb{R}^n$, a convex combination of these k vectors is a vector of the form $\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k$, where $\lambda_i \in \mathbb{R}$ and $\lambda_i \geq 0$ for $1 \leq i \leq k$, satisfying $\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$, i.e., $\boldsymbol{\lambda} := (\lambda_1, \lambda_2, \cdots, \lambda_k)^\top \in \Delta_k$.*
- **Note:** A convex set can be defined by the property that *any convex combination of two points from the set is also in the set.*
- **Theorem:** *Let $C \subseteq \mathbb{R}^n$ be a convex set and let $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m \in C$. Then for any $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_m)^\top \in \Delta_m := \{\boldsymbol{\alpha} \in \mathbb{R}^m_+ : \sum_{i=1}^m \alpha_i = 1\}$, we have $\sum_{i=1}^m \lambda_i \boldsymbol{x}_i \in C$. That is, a convex combination of any finite number of points from a convex set is in the set.*

  Proof: We prove the theorem by induction on m. The case $m = 1$ is trivial. Suppose that $m = k$ holds. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{k+1} \in C$ and $\boldsymbol{\lambda} \in \Delta_{k+1}$. If $\lambda_{k+1} = 1$, then $\sum_{i=1}^{k+1} \lambda_i \boldsymbol{x}_i = \boldsymbol{x}_{k+1} \in C$. If $\lambda_{k+1} < 1$, then

$$\boldsymbol{z} := \sum_{i=1}^{k+1} \lambda_i \boldsymbol{x}_i = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i + \lambda_{k+1} \boldsymbol{x}_{k+1} = (1 - \lambda_{k+1}) \overbrace{\sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_{k+1}} \boldsymbol{x}_i}^{\boldsymbol{v}} + \lambda_{k+1} \boldsymbol{x}_{k+1}.$$

  Since $\displaystyle\sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_{k+1}} = \frac{\sum_{i=1}^{k} \lambda_i}{1 - \lambda_{k+1}} = 1$, we have $\boldsymbol{v} \in C$ and hence, $\boldsymbol{z} \in C$. $\square$

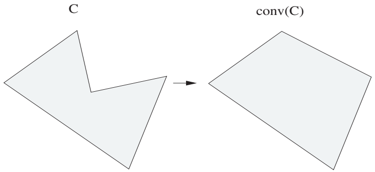
Convex hull

- **Definition:** *(convex hull) Let $S \subseteq \mathbb{R}^n$. Then the convex hull of S is the set comprising all the convex combinations of vectors from S, i.e.,*

$$\mathrm{conv}(S) := \left\{ \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i \,\middle|\, \boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in S, \boldsymbol{\lambda} \in \Delta_k, k \in \mathbb{N} \right\}.$$

- **Note:** *The convex hull* $\mathrm{conv}(S)$ *is a convex set* (**Exercise!**). In fact, $\mathrm{conv}(S)$ is the "smallest" convex set containing S, pls see below.
- **Lemma:** *Let $S \subseteq \mathbb{R}^n$. If $S \subseteq T$ and T is convex, then $\mathrm{conv}(S) \subseteq T$.*

  Proof: Let $\boldsymbol{z} \in \mathrm{conv}(S)$. Then we have $\boldsymbol{z} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$, for some $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k \in S \subseteq T$ and $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_k)^\top \in \Delta_k$. That is, $\boldsymbol{z}$ is a convex combination of elements from T. Since T is convex, by the previous theorem, we obtain $\boldsymbol{z} \in T$. $\square$



A nonconvex set with its convex hull

Carathéodory Theorem

Let $S \subseteq \mathbb{R}^n$ and let $\boldsymbol{x} \in \operatorname{conv}(S)$. Then $\exists\, \boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{n+1} \in S$ such that $\boldsymbol{x} \in \operatorname{conv}(\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{n+1}\})$. That is, $\exists\, \boldsymbol{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_{n+1}) \in \Delta_{n+1}$ such that $\boldsymbol{x} = \sum_{i=1}^{n+1} \lambda_i \boldsymbol{x}_i$.

Proof: Let $\boldsymbol{x} \in \operatorname{conv}(S)$. Then $\exists\, \boldsymbol{x}_1, \cdots, \boldsymbol{x}_k \in S$, $\boldsymbol{\lambda} \in \Delta_k$ s.t. $\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$ with $\lambda_i > 0\ \forall\, i$. If $k \le n+1$, the result is proven. If $k \ge n+2$, then $\boldsymbol{x}_2 - \boldsymbol{x}_1, \cdots, \boldsymbol{x}_k - \boldsymbol{x}_1$ are linearly dependent. Therefore, $\exists\, \mu_2, \cdots, \mu_k$ not all zeros such that $\sum_{i=2}^{k} \mu_i(\boldsymbol{x}_i - \boldsymbol{x}_1) = \boldsymbol{0}$. Let $\mu_1 := -\sum_{i=2}^{k} \mu_i$, we obtain $\sum_{i=1}^{k} \mu_i \boldsymbol{x}_i = \boldsymbol{0}$ and $\sum_{i=1}^{k} \mu_i = 0$, where $\exists\, i$ for which $\mu_i < 0$. Let $\alpha \in \mathbb{R}_+$. Then

$$\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i + \alpha \sum_{i=1}^{k} \mu_i \boldsymbol{x}_i = \sum_{i=1}^{k} (\lambda_i + \alpha\mu_i) \boldsymbol{x}_i \quad \text{and} \quad \sum_{i=1}^{k} (\lambda_i + \alpha\mu_i) = 1.$$

The above representation is a convex combination if and only if

$$\lambda_i + \alpha\mu_i \ge 0, \quad \forall\, i = 1, \cdots, k.$$

Since $\lambda_i > 0\ \forall i$, the above set of inequalities is satisfied for all $\alpha \in [0, \varepsilon]$, where $\varepsilon = \min_{i:\ \mu_i < 0} \left\{ \frac{-\lambda_i}{\mu_i} \right\}$. Taking $\alpha = \varepsilon$, then $\lambda_j + \alpha\mu_j = 0$ for $j = \operatorname{argmin}_{i:\ \mu_i<0} \left\{ \frac{-\lambda_i}{\mu_i} \right\}$. This means that we have found a representation of $\boldsymbol{x}$ as a convex combination of $k-1$ vectors. This process can be carried on until a representation of $\boldsymbol{x}$ as a convex combination of no more than $n+1$ vectors is derived. $\square$

Example: $n = 2$

Let $S = \{x_1, x_2, x_3, x_4\} \subseteq \mathbb{R}^2$, where

$$x_1 = (1,1)^\top, \quad x_2 = (1,2)^\top, \quad x_3 = (2,1)^\top, \quad x_4 = (2,2)^\top.$$

Let $x \in \mathrm{conv}(S)$ be given by

$$x = \frac{1}{8}x_1 + \frac{1}{4}x_2 + \frac{1}{2}x_3 + \frac{1}{8}x_4 = (\frac{13}{8}, \frac{11}{8})^\top$$

By the Carathéodory Theorem, x can be expressed as a convex combination of three of the four vectors x_1, x_2, x_3, x_4. The vectors

$$x_2 - x_1 = (0,1)^\top, \quad x_3 - x_1 = (1,0)^\top, \quad x_4 - x_1 = (1,1)^\top$$

are linearly dependent, and $(x_2 - x_1) + (x_3 - x_1) - (x_4 - x_1) = \mathbf{0}$. i.e., $-x_1 + x_2 + x_3 - x_4 = \mathbf{0}$. Therefore, for any $\alpha \geq 0$ we have

$$x = (\frac{1}{8} - \alpha)x_1 + (\frac{1}{4} + \alpha)x_2 + (\frac{1}{2} + \alpha)x_3 + (\frac{1}{8} - \alpha)x_4.$$

We need guarantee that $\frac{1}{8} - \alpha \geq 0$, $\frac{1}{4} + \alpha \geq 0$, $\frac{1}{2} + \alpha \geq 0$, $\frac{1}{8} - \alpha \geq 0$, which combined with $\alpha \geq 0$ yields that $0 \leq \alpha \leq 1/8$. *Now taking $\alpha = 1/8$, we obtain the convex combination $x = (3/8)x_2 + (5/8)x_3$.*

Convex cones

- **Definition:** *A set S is called a cone if for any $\boldsymbol{x} \in S$ and $\lambda \geq 0$, we have $\lambda \boldsymbol{x} \in S$.*
- **Lemma:** *A set S is a convex cone if and only if the following properties hold: (1) $\boldsymbol{x}, \boldsymbol{y} \in S \Rightarrow \boldsymbol{x} + \boldsymbol{y} \in S$; (2) $\boldsymbol{x} \in S$, $\lambda \geq 0 \Rightarrow \lambda \boldsymbol{x} \in S$.*

  Proof:

  ($\Rightarrow$) Let $\boldsymbol{x}, \boldsymbol{y} \in S$. By the convexity, we have $\frac{1}{2}\boldsymbol{x} + (1 - \frac{1}{2})\boldsymbol{y} \in S$. Since S is a cone, we have $2 \times \frac{1}{2}(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{x} + \boldsymbol{y} \in S$, i.e., property (1) holds. Property (2) is true because S is a cone.

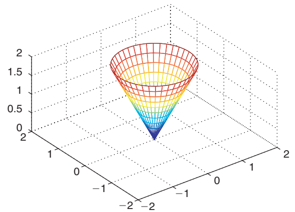
  ($\Leftarrow$) By property (2), S is a cone. Let $\boldsymbol{x}, \boldsymbol{y} \in S$ and $\lambda \in [0, 1]$. Since S is a cone, we have $\lambda \boldsymbol{x} \in S$ and $(1 - \lambda)\boldsymbol{y} \in S$. By property (1), we further have $\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \in S$, establishing the convexity. □
- **Example:** Consider the convex polytope $C = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{0}\}$, where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. The set C is clearly a convex set, see page 6. It is also a cone since

$$\boldsymbol{x} \in C, \lambda \geq 0 \;\Rightarrow\; \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{0}, \lambda \geq 0 \;\Rightarrow\; \boldsymbol{A}(\lambda \boldsymbol{x}) \leq \boldsymbol{0} \;\Rightarrow\; \lambda \boldsymbol{x} \in C.$$

Lorentz cone (ice cream cone)

The Lorentz cone, also called the ice cream cone, is given by

$$L^n := \left\{ (\boldsymbol{x}, t)^\top \in \mathbb{R}^{n+1} : \boldsymbol{x} \in \mathbb{R}^n, t \in \mathbb{R}, \text{ and } \|\boldsymbol{x}\| \leq t \right\}.$$



The boundary of the ice cream cone L^2

The Lorentz cone is in fact a convex cone. Let $(\boldsymbol{x}, t)^\top, (\boldsymbol{y}, s)^\top \in L^n$. Then $\|\boldsymbol{x}\| \leq t$ and $\|\boldsymbol{y}\| \leq s$. The triangle inequality implies that

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \leq t + s.$$

That is, $(\boldsymbol{x}, t)^\top + (\boldsymbol{y}, s)^\top = (\boldsymbol{x} + \boldsymbol{y}, t + s)^\top \in L^n$. We have property (1). To show property (2), take $(\boldsymbol{x}, t)^\top \in L^n$ and $\lambda \geq 0$. Then we obtain $\|\lambda \boldsymbol{x}\| = \lambda \|\boldsymbol{x}\| \leq \lambda t$, so $\lambda (\boldsymbol{x}, t)^\top = (\lambda \boldsymbol{x}, \lambda t)^\top \in L^n$.

Conic combination

- **Definition:** *Given $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in \mathbb{R}^n$, a conic combination of these k vectors is a vector of the form $\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_k \boldsymbol{x}_k$, where $\lambda_i \geq 0$ for all $i = 1, 2, \cdots, k$.*
- **Lemma:** *Let C be a convex cone, and let $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in C$ and $\lambda_1, \lambda_2, \cdots, \lambda_k \geq 0$. Then the conic combination $\sum_{i=1}^k \lambda_i \boldsymbol{x}_i \in C$.*
  Proof: Since C is a convex cone, by property (2), we have $\lambda_i \boldsymbol{x}_i \in C$, $\forall\, i$. By property (1), $\sum_{i=1}^k \lambda_i \boldsymbol{x}_i \in C$. $\square$
- **Definition:** *(conic hull) Let $S \subseteq \mathbb{R}^n$. Then the conic hull of S is the set comprising all the conic combinations of vectors from S, i.e.,*

$$\text{cone}(S) := \left\{ \sum_{i=1}^k \lambda_i \boldsymbol{x}_i \,\middle|\, \boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in S, \boldsymbol{\lambda} \in \mathbb{R}^k_+, k \in \mathbb{N} \right\}.$$

  Note that $\text{cone}(S)$ *is a convex cone.* (**Exercise!**) In fact, we have
- **Lemma:** *Let $S \subseteq \mathbb{R}^n$. If $S \subseteq T$ for some convex cone T, then $\text{cone}(S) \subseteq T$, i.e., the conic hull of S is the smallest convex cone containing S.* (**Exercise!**)

Conic representation theorem

Let $S \subseteq \mathbb{R}^n$ and let $\boldsymbol{x} \in \operatorname{cone}(S)$. Then $\exists\, k$ linearly independent vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in S$ such that $\boldsymbol{x} \in \operatorname{cone}(\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\})$; that is, $\exists\, \boldsymbol{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_k) \in \mathbb{R}^k_+$ such that $\boldsymbol{x} = \sum_{i=1}^k \lambda_i \boldsymbol{x}_i$ and $k \leq n$.

Proof: Let $\boldsymbol{x} \in \operatorname{cone}(S)$. Then $\exists\, \boldsymbol{x}_1, \cdots, \boldsymbol{x}_m \in S$, $\boldsymbol{\lambda} \in \mathbb{R}^m_+$ s.t. $\boldsymbol{x} = \sum_{i=1}^m \lambda_i \boldsymbol{x}_i$ with $\lambda_i > 0\ \forall\, i$. If $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m$ are linearly independent, then $k := m \leq n$ and the result is proven. Otherwise, $\exists\, \mu_1, \cdots, \mu_m \in \mathbb{R}$ not all zeros such that $\sum_{i=1}^m \mu_i \boldsymbol{x}_i = \mathbf{0}$. Let $\alpha \in \mathbb{R}$. Then

$$\boldsymbol{x} = \sum_{i=1}^m \lambda_i \boldsymbol{x}_i = \sum_{i=1}^m \lambda_i \boldsymbol{x}_i + \alpha \sum_{i=1}^m \mu_i \boldsymbol{x}_i = \sum_{i=1}^m (\lambda_i + \alpha \mu_i) \boldsymbol{x}_i.$$

The above representation is a conic combination if and only if

$$\lambda_i + \alpha \mu_i \geq 0, \quad \forall\, i = 1, \cdots, m.$$

Since $\lambda_i > 0$ for all i, we can find $\widetilde{\alpha} \in \mathbb{R}$ s.t. $\lambda_j + \widetilde{\alpha} \mu_j = 0$ for some j and $\lambda_i + \widetilde{\alpha} \mu_i \geq 0$ for the others. Thus we obtain a representation of $\boldsymbol{x}$ as a conic combination of at most $m - 1$ vectors. Continuing this process, we can obtain k linearly independent vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in S$ with $k \leq n$ such that $\boldsymbol{x} \in \operatorname{cone}(\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\})$. $\square$

(Please see textbook page 107 for more details)

Basic feasible solution (BFS)

Linear systems consisting of linear equalities and nonnegativity constraints often appear as constraints in standard formulations of *linear programming problems.*

- **Definition:** *(basic feasible solution)* *Let $P := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$, where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Suppose that the rows of $\boldsymbol{A}$ are linearly independent. Then $\bar{\boldsymbol{x}} \in P$ is a basic feasible solution (BFS) of P if the columns of $\boldsymbol{A}$ corresponding to the indices of the positive values of $\bar{\boldsymbol{x}}$ are linearly independent.*
- **Note:** Since the columns of $\boldsymbol{A}$ reside in $\mathbb{R}^m$, it follows that a BFS has at most m nonzero elements.
- **Example:** Consider the linear system

$$x_1 + x_2 + x_3 = 6, \quad x_2 + x_3 = 3, \quad x_1, x_2, x_3 \geq 0.$$

  A BFS of the system is $(3, 3, 0)$. It satisfies all the constraints and the columns corresponding to the positive elements, $(1, 0)^\top$, $(1, 1)^\top$ are linearly independent.

Existence of a BFS in P

Theorem: *Let $P := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$, where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. If $P \neq \varnothing$, then it contains at least one BFS.*

Proof: Let $\boldsymbol{x} \in P \neq \varnothing$. Then $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$. It follows that $\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$, i.e., $\boldsymbol{b} \in \text{cone}(\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n\})$, where $\boldsymbol{a}_i$ denotes the ith column of $\boldsymbol{A}$. By the conic representation theorem, there exist indices $i_1 < i_2 < \cdots < i_k$ and k numbers $x_{i_1}, x_{i_2}, \cdots, x_{i_k} > 0$ such that $\boldsymbol{b} = \sum_{j=1}^k x_{i_j}\boldsymbol{a}_{i_j}$ and $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \cdots, \boldsymbol{a}_{i_k}$ are linearly independent.

Denote $\bar{\boldsymbol{x}} := \sum_{j=1}^k x_{i_j}\boldsymbol{e}_{i_j}$. Then $\bar{\boldsymbol{x}} \geq \boldsymbol{0}$ and

$$\boldsymbol{A}\bar{\boldsymbol{x}} = \sum_{j=1}^k x_{i_j}\boldsymbol{A}\boldsymbol{e}_{i_j} = \sum_{j=1}^k x_{i_j}\boldsymbol{a}_{i_j} = \boldsymbol{b}.$$

Therefore, $\bar{\boldsymbol{x}} \in P$ and satisfies that the columns of $\boldsymbol{A}$ corresponding to the indices of the positive components of $\bar{\boldsymbol{x}}$ are linearly independent. That is, P contains at least one BFS. $\square$

Closure and interior of a convex set

- **Theorem:** *Let $C \subseteq \mathbb{R}^n$ be a convex set. Then the closure $\text{cl}(C)$ is convex.*

  Proof: Let $\boldsymbol{x}, \boldsymbol{y} \in \text{cl}(C)$ and $\lambda \in [0,1]$. Then $\exists$ sequences $\{\boldsymbol{x}_k\}$, $\{\boldsymbol{y}_k\} \subseteq C$ such that $\boldsymbol{x}_k \to \boldsymbol{x}$ and $\boldsymbol{y}_k \to \boldsymbol{y}$ as $k \to \infty$.
  By the convexity of C, $\lambda \boldsymbol{x}_k + (1-\lambda)\boldsymbol{y}_k \in C$ for any k.
  Since $\lambda \boldsymbol{x}_k + (1-\lambda)\boldsymbol{y}_k \to \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}$, we can conclude that $\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in \text{cl}(C)$, which implies that $\text{cl}(C)$ is convex. $\square$

- **(line segment principle):** *Let $C \subseteq \mathbb{R}^n$ be a convex set, and assume that $\text{int}(C) \neq \varnothing$. Suppose that $\boldsymbol{x} \in \text{int}(C)$, $\boldsymbol{y} \in \text{cl}(C)$. Then $(1-\lambda)\boldsymbol{x} + \lambda \boldsymbol{y} \in \text{int}(C)$ for any $\lambda \in (0,1)$.*

  (Please see textbook page 109 for the proof)

- **Theorem:** *Let $C \subseteq \mathbb{R}^n$ be a convex set. Then the interior $\text{int}(C)$ is convex.*

  Proof: If $\text{int}(C) = \varnothing$, then $\text{int}(C)$ is convex. Let $\boldsymbol{x}, \boldsymbol{y} \in \text{int}(C)$ and $\lambda \in (0,1)$. Then by the line segment principle, $(1-\lambda)\boldsymbol{x} + \lambda \boldsymbol{y} \in \text{int}(C)$. We can conclude that $\text{int}(C)$ is convex. $\square$

Other topological properties

Let $C \subseteq \mathbb{R}^n$ be a convex set and $\mathrm{int}(C) \neq \varnothing$. *Then we have*

- $\mathrm{cl}(\mathrm{int}(C)) = \mathrm{cl}(C)$.

  Proof:

  ($\subseteq$): Since $\mathrm{int}(C) \subseteq C$, we have $\mathrm{cl}(\mathrm{int}(C)) \subseteq \mathrm{cl}(C)$.

  ($\supseteq$): Let $\boldsymbol{x} \in \mathrm{cl}(C)$. We take $\boldsymbol{y} \in \mathrm{int}(C)$. Then by the line segment principle, we have $\boldsymbol{x}_k := \frac{1}{k}\boldsymbol{y} + (1 - \frac{1}{k})\boldsymbol{x} \in \mathrm{int}(C)$ for any $k \geq 1$. Since $\boldsymbol{x}_k \to \boldsymbol{x}$ as $k \to \infty$, we obtain $\boldsymbol{x} \in \mathrm{cl}(\mathrm{int}(C))$. $\square$

- $\mathrm{int}(\mathrm{cl}(C)) = \mathrm{int}(C)$.

  Proof:

  ($\supseteq$): Since $C \subseteq \mathrm{cl}(C)$, we have $\mathrm{int}(\mathrm{cl}(C)) \supseteq \mathrm{int}(C)$.

  ($\subseteq$): Let $\boldsymbol{x} \in \mathrm{int}(\mathrm{cl}(C))$. Then $\exists\, \varepsilon > 0$ s.t. $B(\boldsymbol{x}, \varepsilon) \subseteq \mathrm{cl}(C)$. Let $\boldsymbol{y} \in \mathrm{int}(C)$. If $\boldsymbol{y} = \boldsymbol{x}$, then the result is proved. Otherwise, define $\boldsymbol{z} := \boldsymbol{x} + \alpha(\boldsymbol{x} - \boldsymbol{y})$, where $\alpha = \frac{\varepsilon}{2\|\boldsymbol{x} - \boldsymbol{y}\|}$. Since $\|\boldsymbol{z} - \boldsymbol{x}\| = \frac{\varepsilon}{2}$, we have $\boldsymbol{z} \in \mathrm{cl}(C)$. By the line segment principle, we have $(1 - \lambda)\boldsymbol{y} + \lambda\boldsymbol{z} \in \mathrm{int}(C)$ for $\lambda \in [0, 1)$. Taking $\lambda = \frac{1}{1+\alpha} \in (0, 1)$, we obtain $(1 - \lambda)\boldsymbol{y} + \lambda\boldsymbol{z} = \boldsymbol{x} \in \mathrm{int}(C)$. $\square$

Convex hull of compact set

Theorem: *Let $S \subseteq \mathbb{R}^n$ be a compact set. Then $\operatorname{conv}(S)$ is compact.*

Proof:

- (Boundedness) Since S is bounded, $\exists\, M > 0$ such that $\|\boldsymbol{x}\| \le M$ for any $x \in S$. Let $\boldsymbol{y} \in \operatorname{conv}(S)$. By the Carathéodory theorem it follows that $\exists\, \boldsymbol{x}_1, \cdots, \boldsymbol{x}_{n+1} \in S$ and $\boldsymbol{\lambda} \in \Delta_{n+1}$ s.t. $\boldsymbol{y} = \sum_{i=1}^{n+1} \lambda_i \boldsymbol{x}_i$. Therefore,

$$\|\boldsymbol{y}\| = \|\sum_{i=1}^{n+1} \lambda_i \boldsymbol{x}_i\| \le \sum_{i=1}^{n+1} \lambda_i \|\boldsymbol{x}_i\| \le M \sum_{i=1}^{n+1} \lambda_i = M.$$

- (Closedness) Let $\boldsymbol{y}_k$ be a sequence in $\operatorname{conv}(S)$ and $\boldsymbol{y}_k \to \boldsymbol{y}$ as $k \to \infty$. We wish to show that $\boldsymbol{y} \in \operatorname{conv}(S)$. By the Carathéodory theorem it follows that $\exists\, \boldsymbol{x}_1^k, \cdots, \boldsymbol{x}_{n+1}^k \in S$ and $\boldsymbol{\lambda}^k \in \Delta_{n+1}$ s.t. $\boldsymbol{y}_k = \sum_{i=1}^{n+1} \lambda_i^k \boldsymbol{x}_i^k$. By the compactness of S and Δ_{n+1}, the sequence $\{(\boldsymbol{\lambda}^k, \boldsymbol{x}_1^k, \cdots, \boldsymbol{x}_{n+1}^k)\}$ has a subsequence such that

$$\lim_{j\to\infty} (\boldsymbol{\lambda}^{k_j}, \boldsymbol{x}_1^{k_j}, \cdots, \boldsymbol{x}_{n+1}^{k_j}) = (\boldsymbol{\lambda}, \boldsymbol{x}_1, \cdots, \boldsymbol{x}_{n+1})$$

  with $\boldsymbol{\lambda} \in \Delta_{n+1}$ and $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{n+1} \in S$. Therefore, we have

$$\boldsymbol{y} = \lim_{j\to\infty} \boldsymbol{y}_{k_j} = \lim_{j\to\infty} \sum_{i=1}^{n+1} \lambda_i^{k_j} \boldsymbol{x}_i^{k_j} = \sum_{i=1}^{n+1} \lambda_i \boldsymbol{x}_i,$$

  which means that $\boldsymbol{y} \in \operatorname{conv}(S)$. $\square$

Conic hull of a finite set

Theorem: *Let $S := \{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k\} \subseteq \mathbb{R}^n$. Then $\text{cone}(S)$ is closed.*

Proof:

- By the conic representation theorem, each element of $\text{cone}(S)$ can be represented as a conic combination of a linearly independent subset of $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_k\}$. Let $S_1, \cdots, S_N$ be all the subsets of S comprising linearly independent vectors, then

$$\text{cone}(S) = \cup_{i=1}^{N} \text{cone}(S_i).$$

  It suffices to show that $\text{cone}(S_i)$ is closed for all i. Let $i \in \{1, 2, \cdots, N\}$. Then $S_i = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_m\}$ for some linearly independent vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_m$. We can write $\text{cone}(S_i) = \{\boldsymbol{B}\boldsymbol{y} : \boldsymbol{y} \in \mathbb{R}^m_+\}$, where matrix $\boldsymbol{B} := [\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_m]_{n \times m}$.

- Let $\boldsymbol{x}_k \in \text{cone}(S_i)$ for $k \geq 1$ and $\boldsymbol{x}_k \to \bar{\boldsymbol{x}}$ as $k \to \infty$. We need to show that $\bar{\boldsymbol{x}} \in \text{cone}(S_i)$. Since $\boldsymbol{x}_k \in \text{cone}(S_i)$, $\exists\, \boldsymbol{y}_k \in \mathbb{R}^m_+$ s.t. $\boldsymbol{x}_k = \boldsymbol{B}\boldsymbol{y}_k$. Since the columns of $\boldsymbol{B}$ are linearly independent, we can deduce that

$$\boldsymbol{y}_k = (\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top \boldsymbol{x}_k.$$

  Thus, we have

$$\lim_{k\to\infty} \boldsymbol{y}_k = \lim_{k\to\infty} (\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top \boldsymbol{x}_k = (\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top \bar{\boldsymbol{x}} =: \bar{\boldsymbol{y}}$$

  and $\bar{\boldsymbol{y}} \in \mathbb{R}^m_+$. Therefore,

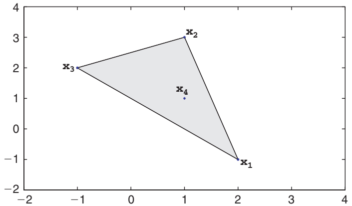
$$\bar{\boldsymbol{x}} = \lim_{k\to\infty} \boldsymbol{x}_k = \lim_{k\to\infty} \boldsymbol{B}\boldsymbol{y}_k = \boldsymbol{B}\bar{\boldsymbol{y}} \in \text{cone}(S_i). \quad \square$$

Extreme points

- **Definition:** *(extreme points) Let $S \subseteq \mathbb{R}^n$ be a convex set. A point $\boldsymbol{x} \in S$ is called an extreme point of S if there do not exist $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$ and $\lambda \in (0,1)$ such that $\boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$. The set of extreme points of S is denoted by $\mathrm{ext}(S)$.*

  That is, an extreme point is a point in the set S that cannot be represented as a nontrivial convex combination of two different points in S.

- **Example:** The set of extreme points of a convex polytope consists of all its vertices.



The convex set $S = \mathrm{conv}\{x_1, x_2, x_3, x_4\}$.
The extreme points set is $\mathrm{ext}(S) = \{x_1, x_2, x_3\}$.

Extreme points and basic feasible solutions

Theorem: *Let $P := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$, where $\boldsymbol{A} \in \mathbb{A}^{m\times n}$ has linearly independent rows and $\boldsymbol{b} \in \mathbb{R}^m$. Then $\bar{\boldsymbol{x}}$ is a basic feasible solution of P if and only if it is an extreme point of P.*

Proof:
($\Rightarrow$): Let $\bar{\boldsymbol{x}} = (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)^\top$ be a basic feasible solution of P. Without loss of generality, assume that $\bar{x}_1, \cdots, \bar{x}_k > 0$ and $\bar{x}_{k+1} = \cdots = \bar{x}_n = 0$, and the first k columns of $\boldsymbol{A}$, denoted by $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$, are linearly independent. Suppose that $\bar{\boldsymbol{x}} \notin \text{ext}(P)$. Then $\exists\, \boldsymbol{y}, \boldsymbol{z} \in P$, $\boldsymbol{y} \neq \boldsymbol{z}$, and $\lambda \in (0,1)$ s.t. $\bar{\boldsymbol{x}} = \lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{z}$. Note that the last $n-k$ components in $\boldsymbol{y}$ and $\boldsymbol{z}$ are zeros. Therefore, we have

$$\sum_{i=1}^{k} y_i\boldsymbol{a}_i = \boldsymbol{b} \text{ and } \sum_{i=1}^{k} z_i\boldsymbol{a}_i = \boldsymbol{b} \Longrightarrow \sum_{i=1}^{k}(y_i - z_i)\boldsymbol{a}_i = \boldsymbol{0},\ y_i - z_i \neq 0 \text{ for some } i \in \{1, 2, \cdots, k\},$$

which implies that $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$ are linearly dependent, a contradiction!
($\Leftarrow$): Suppose that $\widetilde{\boldsymbol{x}} \in P$ is an extreme point, but it is not a basic feasible solution. Thus, the columns corresponding to the positive components of $\widetilde{\boldsymbol{x}}$ are linearly dependent. WLOG, assume that the positive components of $\widetilde{\boldsymbol{x}}$ are exactly the first k components. Then $\exists\, \boldsymbol{y} \in \mathbb{R}^k$ s.t. $\sum_{i=1}^{k} y_i\boldsymbol{a}_i = \boldsymbol{0}$, i.e., $\boldsymbol{A}\widetilde{\boldsymbol{y}} = \boldsymbol{0}$, where $\widetilde{\boldsymbol{y}} = (\boldsymbol{y}, \boldsymbol{0})^\top$. Since the first k components of $\widetilde{\boldsymbol{x}}$ are positive, $\exists\, \varepsilon > 0$ s.t. $\boldsymbol{x}_1 := \widetilde{\boldsymbol{x}} + \varepsilon\widetilde{\boldsymbol{y}} \geq \boldsymbol{0}$ and $\boldsymbol{x}_2 := \widetilde{\boldsymbol{x}} - \varepsilon\widetilde{\boldsymbol{y}} \geq \boldsymbol{0}$. Then we have $\boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{A}\widetilde{\boldsymbol{x}} + \varepsilon\boldsymbol{A}\widetilde{\boldsymbol{y}} = \boldsymbol{b} + \varepsilon\boldsymbol{0} = \boldsymbol{b}$ and $\boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{b}$. Therefore, $\boldsymbol{x}_1, \boldsymbol{x}_2 \in P$. Finally, we have $\widetilde{\boldsymbol{x}} = \frac{1}{2}\boldsymbol{x}_1 + \frac{1}{2}\boldsymbol{x}_2$. This is a contradiction, because $\widetilde{\boldsymbol{x}}$ is an extreme point of P. $\square$

The Krein-Milman theorem

We will state this theorem without a proof.

Krein-Milman theorem: *Let $S \subseteq \mathbb{R}^n$ be a compact convex set. Then*

$$S = \text{conv}(\text{ext}(S)).$$

That is, a compact convex set is the convex hull of its extreme points.