Convex Sets and Functions

Instructor: Jin Zhang

Department of Mathematics
Southern University of Science and Technology

Fall 2023

Contents

1. Convex Sets
2. Extended Real-valued Functions
3. Convex Functions
4. Support Functions

Definitions

- (Convex Sets) A set Ω of $\mathbb{R}^n$ is convex if for any $x, y \in \Omega$ and $0 \leq \lambda \leq 1$ it holds that $\lambda x + (1-\lambda)y \in \Omega$.
- (convex combination) Given $\omega_1, \cdots, \omega_m \in \mathbb{R}^n$, the element $x = \sum_{i=1}^m \lambda_i \omega_i$, where $\lambda_1, \cdots, \lambda_m \in \mathbb{R}$ satisfy $\sum_{i=1}^m \lambda_i = 1$, is a convex combination of $\omega_1, \cdots, \omega_m$.
- (Affine Mappings) A mapping $\mathcal{B} : \mathbb{R}^p \to \mathbb{R}^q$ is affine if there exist a linear mapping $\mathcal{A} : \mathbb{R}^p \to \mathbb{R}^q$ and an element $b \in \mathbb{R}^q$ such that $\mathcal{B}(x) = \mathcal{A}(x) + b$ for all $x \in \mathbb{R}^p$.

Proposition

$\Omega \subseteq \mathbb{R}^n$ is convex if and only if it contains all convex combinations of its elements.

(continued).

Convexity-preserving

1. Let $\Omega_1 \subseteq \mathbb{R}^p$ and $\Omega_2 \subseteq \mathbb{R}^q$ be convex sets. Then the Cartesian product $\Omega_1 \times \Omega_2$ is a convex set of $\mathbb{R}^p \times \mathbb{R}^q$.
2. Let $\Omega_1, \Omega_2 \subseteq \mathbb{R}^n$ be convex, $\alpha, \beta \in \mathbb{R}$. Then $\alpha\Omega_1 + \beta\Omega_2$ is convex, where

$$\alpha\Omega_1 + \beta\Omega_2 := \left\{ w \in \mathbb{R}^n \middle| \exists\, w_1 \in \Omega_1, w_2 \in \Omega_2 \text{ s.t. } w = \alpha w_1 + \beta w_2 \right\}.$$

3. Let $\{\Omega_\alpha\}_{\alpha \in \mathcal{I}}$ be collection of convex subsets of $\mathbb{R}^n$. Then the intersection $\cap_{\alpha \in \mathcal{I}} \Omega_\alpha$ is convex.

(continued).

Definition (Convex Hull) Let $\Omega \subseteq \mathbb{R}^n$. The convex hull of Ω is defined by

$$\text{co}\Omega = \bigcap \left\{ C \,\middle|\, C \text{ is convex and } \Omega \subseteq C \right\}.$$

Properties

- The convex hull of Ω is the smallest convex set containing Ω.
- For any $\Omega \subseteq \mathbb{R}^n$, its convex hull admits the representation

$$\text{co}\Omega = \left\{ \sum_{i=1}^{m} \lambda_i a_i \,\middle|\, \sum_{i=1}^{m} \lambda_i = 1,\ \lambda_i \geq 0,\ a_i \in \Omega,\ m \in \mathbb{N} \right\}.$$

Definitions

- (Lines) Given two elements a and b in $\mathbb{R}^n$, the line connecting them is

$$\mathcal{L}[a,b] := \{\lambda a + (1-\lambda)b \,|\, \lambda \in \mathbb{R}\}\,.$$

- (Affine Sets) $\Omega \subseteq \mathbb{R}^n$ is affine if $\mathcal{L}[a,b] \subset \Omega$ for any $a, b \in \Omega$.
- (Affine Hulls) The affine hull of a set $\Omega \subseteq \mathbb{R}^n$ is

$$\text{aff}\Omega := \bigcap \left\{ C \,\middle|\, C \text{ is affine and } \Omega \subset C \right\}.$$

- (Relative Interiors of Convex Sets) Let $\Omega \subset \mathbb{R}^n$ be convex. Then $v \in \Omega$ belongs to the relative interior of Ω denoted by riΩ if there exists some neighborhood V of v such that $V \cap \text{aff}\Omega \subset \Omega$.

Notations

- $\mathbb{R} \cup \{-\infty, \infty\} = [-\infty, \infty]$, $\mathbb{R} \cup \{\infty\} = (-\infty, \infty] = \overline{\mathbb{R}}$.
- We focus on $f : \mathbb{R}^n \to [-\infty, \infty]$ (an extended real-valued function). The (effective) domain of f is the set

$$\mathrm{dom} f = \{x \in \mathbb{R}^n : f(x) < \infty\}$$

and the epigraph of f is the set

$$\mathrm{epi} f = \{(x, y) : f(x) \leq y, x \in \mathbb{R}^n, y \in \mathbb{R}\}.$$

- We say $f : \mathbb{R}^n \to [-\infty, \infty]$ is proper if $\mathrm{dom} f \neq \varnothing$ and $f(x) > -\infty$ for any $x \in \mathbb{R}^n$.

Example: The indicator function associated with a set $C \subseteq \mathbb{R}^n$, given by $\delta_C(x) = 0$ if $x \in C$ and $\delta_C(x) = \infty$ otherwise, is extended real-valued.

(continued).

Definitions

- $f : \mathbb{R}^n \to [-\infty, \infty]$ is closed if its epigraph is closed.
- $f : \mathbb{R}^n \to [-\infty, \infty]$ is lower semicontinuous at $x \in \mathbb{R}^n$ if for any sequence $\{x_n\}_{n\geq 1}$ for which $x_n \to x$ as $n \to \infty$, we have

$$f(x) \leq \liminf_{n\to\infty} f(x_n).$$

  Moreover, $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is lower semicontinuous if it is lower semicontinuous at each point in $\mathbb{R}^n$.
- For any $\alpha \in \mathbb{R}$, the α-level set of $f : \mathbb{R}^n \to [-\infty, \infty]$ is the set

$$\text{Lev}(f, \alpha) = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}.$$

(continued).

Theorem: equivalence of closedness, lower semicontinuity, and closedness of level sets

Let $f : \mathbb{R}^n \to [-\infty, \infty]$. Then the following claims are equivalent:

1. f is lower semicontinuous.
2. f is closed.
3. For any $\alpha \in \mathbb{R}$, the α-level set $\text{Lev}(f, \alpha)$ is closed.

(continued).

Theorem: Preserving Closedness

1. Let $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^p$ be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^p$ and $b \in \mathbb{R}^p$ and let $f : \mathbb{R}^p \to [-\infty, \infty]$ be closed. Then $g : \mathbb{R}^n \to [-\infty, \infty]$ given by $g(x) = f(\mathcal{A}(x) + b)$ is closed.
2. Let $f_1, \cdots, f_m : \mathbb{R}^n \to \overline{\mathbb{R}}$ be closed and let $\alpha_1, \cdots, \alpha_m \in \mathbb{R}_+$. Then $f = \sum_{i=1}^m \alpha_i f_i$ is closed.
3. Let $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}, i \in \mathcal{I}$ be closed, where $\mathcal{I}$ is a given index set. Then the supremum function

$$f(x) = \sup_{i \in \mathcal{I}} f_i(x)$$

is closed.

Examples on Closedness:

- (closedness of indicators of closed sets). Let $C \subseteq \mathbb{R}^n$. Then the indicator function δ_C is closed if and only if C is a closed set.
- The function $f : \mathbb{R} \to \overline{\mathbb{R}}$, given by $f(x) = \frac{1}{x}$ for $x > 0$ and ∞ otherwise, is a closed function.
- (support functions). Let $\varnothing \neq C \subseteq \mathbb{R}^n$. Then the support function of C, given by

  $$\sigma_C(y) = \sup_{x \in C} \langle y, x \rangle \text{ for any } y \in \mathbb{R}^n,$$

  is a closed function.

Theorem

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be continuous over its domain and suppose $\mathrm{dom} f$ is closed. Then f is closed.

Example:

- Consider the function $f_\alpha : \mathbb{R} \to \overline{\mathbb{R}}$ associated with $\alpha \in \mathbb{R}$ given by

$$f_\alpha(x) = \begin{cases} \alpha, & \text{if } x = 0, \\ x, & \text{if } 0 < x \leq 1, \\ \infty, & \text{else.} \end{cases}$$

  Then f_α is closed if and only if $\alpha \leq 0$, and it is continuous over its domain if and only if $\alpha = 0$.
- (l_0-norm). The l_0-norm function $f : \mathbb{R}^n \to \mathbb{R}$, given by

$$f(x) = \|x\|_0 = \#\{i : x_i \neq 0\},$$

  is a closed function.

(continued).

Weierstrass theorem for closed functions

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper closed and assume that $C \subseteq \mathbb{R}^n$ is compact satisfying $C \cap \mathrm{dom} f \neq \varnothing$. Then

- f is bounded below over C.
- f attains its minimal value over C.

(continued)

Definition: Coerciveness

A proper function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is called coercive if

$$\lim_{\|x\|\to\infty} f(x) = \infty.$$

Theorem: attainment under coerciveness

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper closed and coercive and let $\Omega \subseteq \mathbb{R}^n$ be a closed set satisfying $\Omega \cap \operatorname{dom} f \neq \varnothing$. Then f attains its minimal value over Ω.

Definition: convex functions

$f : \mathbb{R}^n \to [-\infty, \infty]$ is convex if $\text{epi} f$ is a convex set.

Recall that $\alpha + \infty = \infty (\forall \alpha \in \mathbb{R})$, $\infty + \infty = \infty$, $\alpha \cdot \infty = \infty (\alpha > 0)$.

Properties

- If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex, then $\text{dom} f$ is a convex set.
- $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if and only if for any for any $x, y \in \mathbb{R}^n$ and $0 \leq \lambda \leq 1$ it holds that $f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$.
- **Jensen's Inequality:** Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex. Then for any $x^1, \cdots, x^k \in \mathbb{R}^n$ and $\lambda_1, \cdots, \lambda_k \in \mathbb{R}_+$ satisfying $\sum_{i=1}^k \lambda_i = 1$ it holds that
$$f\left(\sum_{i=1}^k \lambda_i x^i\right) \leq \sum_{i=1}^k \lambda_i f(x^i).$$

(continued).

operations preserving convexity:

1. Let $\mathcal{A}$ be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^p$ and $b \in \mathbb{R}^p$, and let $f : \mathbb{R}^p \to \overline{\mathbb{R}}$ be a convex function. Then $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ given by $g(x) = f(\mathcal{A}(x) + b)$ is convex.
2. Let $f_1, \cdots, f_m : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex and let $\alpha_1, \cdots, \alpha_m \in \mathbb{R}_+$. Then $f = \sum_{i=1}^m \alpha_i f_i$ is convex.
3. Let $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}, i \in \mathcal{I}$ be convex. Then the supremum function

$$f(x) = \sup_{i \in \mathcal{I}} f_i(x)$$

   is convex.
4. **partial minimization:** Let $f : \mathbb{R}^p \times \mathbb{R}^q \to \overline{\mathbb{R}}$ be convex satisfying the following property:

   for any $x \in \mathbb{R}^p$ there exists $y \in \mathbb{R}^q$ for which $f(x, y) < \infty$.

   Then $g : \mathbb{R}^p \to [-\infty, \infty)$ given by $g(x) = \inf_{y \in \mathbb{R}^q} f(x, y)$ is convex.

Definitions

- **Distance functions.** Given $\varnothing \neq C \subseteq \mathbb{R}^n$, the distance function associated with C is defined by

$$d_C(x) = d(x; C) = \inf \{\|x - \omega\| \mid \omega \in C\}.$$

- **Infimal Convolution.** Let $h_1, h_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper. The infimal convolution of h_1, h_2 is defined by

$$(h_1 \square h_2)(x) = \inf_{u \in \mathbb{R}^n} \{h_1(u) + h_2(x - u)\}.$$

Example 1: Let $C \subset \mathbb{R}^n$ be nonempty. If the norm on $\mathbb{R}^n$ is given by $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$. Then the function

$$\varphi_C(x) = \frac{1}{2}\left(\|x\|^2 - d_C^2(x)\right)$$

is convex.

(continued).

Theorem: (Convexity of the infimal convolution).

Let $h_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper convex and let $h_2 : \mathbb{R}^n \to \mathbb{R}$ be real-valued convex. Then $h_1 \Box h_2$ is convex.

Example 2: (Distance Function)

Let $\varnothing \neq C \subset \mathbb{R}^n$ be convex. The distance function associated with C is convex.

Theorem: Local Lipschitz Continuity of Convex Functions

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex. Let $x_0 \in \operatorname{int}(\operatorname{dom} f)$. Then there exists $\varepsilon > 0$ and $L > 0$ such that $\mathbb{B}(x_0, \varepsilon) \subseteq C$ and

$$|f(x) - f(x_0)| \leq L \|x - x_0\|$$

for all $x \in \mathbb{B}(x_0, \varepsilon)$

Theorem: continuity of closed convex univariate functions

Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be proper closed and convex. Then f is continuous over $\operatorname{dom} f$.

Recall that the support function of $\varnothing \neq C \subset \mathbb{R}^n$ is defined as

$$\sigma_C(y) = \sup_{x \in C} \langle y, x \rangle \text{ for any } y \in \mathbb{R}^n.$$

Lemma: (convexity of support functions) Let $\varnothing \neq C \subset \mathbb{R}^n$. Then σ_C is a convex function.

(continued).

Lemmas

- **positive homogeneity:** For any $\varnothing \neq C \subseteq \mathbb{R}^n$ and $\alpha > 0$,
$$\sigma_C(\alpha y) = \alpha \sigma_C(y).$$
- **subadditivity:** For any $\varnothing \neq C \subseteq \mathbb{R}^n$,
$$\sigma_C(y_1 + y_2) \leq \sigma_C(y_1) + \sigma_C(y_2).$$
- For any $\varnothing \neq C \subseteq \mathbb{R}^n$ and $\alpha > 0$,
$$\sigma_{\alpha C}(y) = \alpha \sigma_C(y).$$
- For any For any $\varnothing \neq A, B \subseteq \mathbb{R}^n$,
$$\sigma_{A+B}(y) = \sigma_A(y) + \sigma_B(y).$$

(continued).

Example 1 (support functions of finite sets)

Suppose that

$$C = \{b_1, \cdots, b_m\},$$

where $b_1, \cdots, b_m \in \mathbb{R}^n$. Then

$$\sigma_C = \max\{\langle b_1, y\rangle, \cdots, \langle b_m, y\rangle\}.$$

Example 2 (support functions of cones)

Let $K \subset \mathbb{R}^n$ be a cone, i.e., for any $x \in K$, we have $\alpha x \in K$ for all $\alpha \geq 0$. Define the polar cone of K as

$$K^\circ = \{y \in \mathbb{R}^n \mid \langle y, x\rangle \leq 0 \text{ for all } x \in K\}.$$

Then we have $\sigma_K = \delta_{K^\circ}$.

(continued).

Example 3: (support function of the nonnegative orthant)

Notice that $\left(\mathbb{R}^n_+\right)^\circ = \mathbb{R}^n_-$, it follows that

$$\sigma_{\mathbb{R}^n_+} = \delta_{\mathbb{R}^n_-}.$$

Example 4: (support functions of unit balls)

Consider $\mathbb{R}^n$ endowed with a norm $\|\cdot\|$ and the unit ball given by

$$\mathbb{B}_{\|\cdot\|}(0,1) = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}.$$

We have for any $y \in \mathbb{R}^n$

$$\sigma_{\mathbb{B}_{\|\cdot\|}(0,1)}(y) = \max_{\|x\| \leq 1} \langle y, x \rangle = \|y\|_*.$$

(continued).

Some theorems of the alternative

(Farkas's lemma-first formulation) Let $c \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$. Then exactly one of the following is true:

A. There exists $x \in \mathbb{R}^n$ such that $Ax \leq 0$ and $c^T x > 0$.

B. There exists $y \in \mathbb{R}^m_+$ such that $A^T y = c$.

(Farkas's lemma-second formulation) Let $c \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$. Then the following two claims are equivalent:

A. The implication $Ax \leq 0 \Rightarrow c^T x \leq 0$ holds true.

B. There exists $y \in \mathbb{R}^m_+$ such that $A^T y = c$.

(Gordan Theorem) Let $B \in \mathbb{R}^{m \times n}$. Then exactly one of the following is true:

A. There exists $x \in \mathbb{R}^n$ such that $Bx < 0$.

B. There exists $0 \neq y \geq 0$ such that $B^T y = 0$.

Proof of Farkas's lemma-first formulation

Consider the following linear programming (LP) problem and its dual problem

$$\min_{y} \left\{ 0 \mid A^T y = c, y \geq 0 \right\} (P).$$
$$\max_{x} \left\{ c^T x \mid Ax \leq 0 \right\} (D).$$

If Claim B holds, then by weak duality theorem, we have $0 \geq c^T x$ if x satisfies $Ax \leq 0$, which is contradict with Claim A.
If Claim A holds, then the optimal value of (D) is ∞. For any feasible point in (P), it is contradict with weak duality theorem $c^T x \leq 0$.

(continued).

Example 5 (support functions of convex polyhedral cones)

Let $A \in \mathbb{R}^{m\times n}$. Define the set

$$S = \{x \in \mathbb{R}^n \,|\, Ax \leq 0\}.$$

Then we have $\sigma_S(y) = \delta_{S^\circ}(y)$, where $S^\circ = \left\{A^T\lambda \,|\, \lambda \in \mathbb{R}^m_+\right\}$.

Example 6 (support functions of affine sets).

Let $B \in \mathbb{B}^{m\times n}$, $b \in \mathbb{R}^m$. Define the affine set

$$C = \{x \in \mathbb{R}^n \,|\, Bx = b\}.$$

We assume that C is nonempty. The support function is given by

$$\sigma_C(y) = \langle y, x_0\rangle + \sigma_{\mathsf{Range}(B^T)}(y).$$

Example 7 (a closed, convex and noncontinuous function). Consider the following set in $\mathbb{R}^2$:

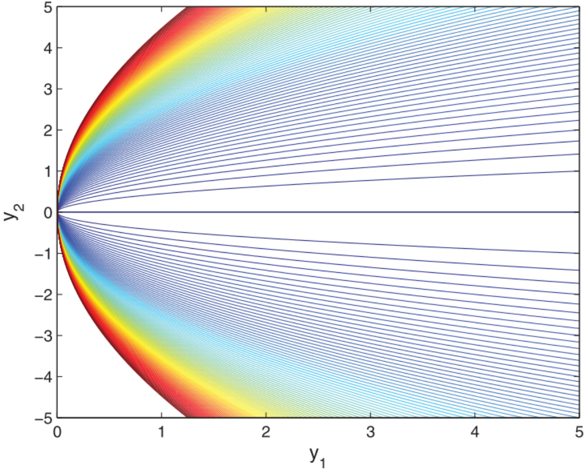
$$C = \left\{ (x_1, x_2)^T \,\middle|\, x_1 + \frac{x_2^2}{2} \leq 0 \right\}.$$

Then the support function is given by

$$\sigma_C(y) = \begin{cases} \frac{y_2^2}{2y_1}, & y_1 > 0, \\ 0, & y_1 = y_2 = 0, \\ \infty & \text{else.} \end{cases}$$

(continued).

σ_C is not continuous at $(y_1, y_2) = (0, 0)$. The contour lines of σ_C are plotted in the following figure.



(continued).

Theorem (strict separation theorem).

Let $C \subseteq \mathbb{R}^n$ be nonempty closed and convex and let $y \notin C$. Then there exist $p \in \mathbb{R}^n \backslash \{0\}$ and $\alpha \in \mathbb{R}^n$ such that

$$\langle p, y \rangle > \alpha$$

and

$$\langle p, x \rangle \leq \alpha \text{ for all } x \in C.$$

Lemma

Let $A, B \subseteq \mathbb{R}^n$ be nonempty closed and convex. Then $A = B$ if and only if $\sigma_A = \sigma_B$.

Lemma

Let $A \subseteq \mathbb{R}^n$ be nonempty. Then

(a). $\sigma_A = \sigma_{\text{cl}A}$;

(b). $\sigma_A = \sigma_{\text{co}A}$.

Example 8 (support of the unit simplex)

Consider the unit simplex set $\Delta_n = \left\{x \in \mathbb{R}^n \,|\, e^T x = 1, x \geq 0\right\}$, where e is the column vector of all ones. Then

$$\sigma_{\Delta_n}(y) = \max\{y_1, y_2, \cdots, y_n\}.$$