

Chapter 1

Convex Analysis

Convex Optimization for Machine Learning & Computer Vision
WS 2018/19

Tao Wu
Yuesong Shen
Zhenzhang Ye

Computer Vision Group
Department of Informatics
TU Munich

Convex Analysis

Tao Wu
Yuesong Shen
Zhenzhang Ye



Convex Set

Convex Function

Existence of Minimizer



Convex Set

Notations

- \mathbb{E} is a *Euclidean space* (i.e., finite dimensional inner product space), equipped with
 - 1 Inner product $\langle \cdot, \cdot \rangle$, e.g., $\langle u, v \rangle = u^T v$ if $\mathbb{E} = \mathbb{R}^n$;
 - 2 Norm $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ satisfying polarization identity:

$$2\|u\|^2 + 2\|v\|^2 = \|u + v\|^2 + \|u - v\|^2.$$

- C is a closed, convex subset of \mathbb{E} .
- J is a convex *objective* function.



Convex Set

Convex Function

Existence of Minimizer



Notations

- \mathbb{E} is a *Euclidean space* (i.e., finite dimensional inner product space), equipped with

- 1 Inner product $\langle \cdot, \cdot \rangle$, e.g., $\langle u, v \rangle = u^\top v$ if $\mathbb{E} = \mathbb{R}^n$;
- 2 Norm $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ satisfying polarization identity:

$$2\|u\|^2 + 2\|v\|^2 = \|u + v\|^2 + \|u - v\|^2.$$

- C is a closed, convex subset of \mathbb{E} .
- J is a convex *objective* function.

Convex Set

Convex Function

Existence of Minimizer

Convex optimization

$$\text{minimize } J(u) \quad \text{over } u \in C.$$

First questions:

- What is a convex set?
- What is a convex function?

Convex set

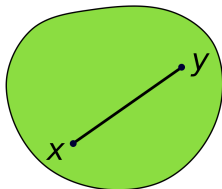
Definition

A set C is said to be **convex** if

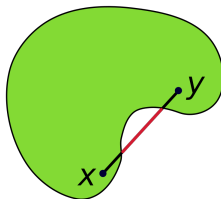
$$\alpha u + (1 - \alpha)v \in C, \quad \forall u, v \in C, \quad \forall \alpha \in [0, 1].$$



convex



non-convex



Recall basic concepts in analysis

Definition

- A set $C \subset \mathbb{E}$ is **open** if $\forall u \in C, \exists \epsilon > 0$ s.t. $B_\epsilon(u) \subset C$, where $B_\epsilon(u) := \{v \in \mathbb{E} : \|v - u\| < \epsilon\}$.
- A set $C \subset \mathbb{E}$ is **closed** if its complement $\mathbb{E} \setminus C$ is open.
- The **closure** of a set $C \subset \mathbb{E}$ is

$$\text{cl } C = \{u \in \mathbb{E} : \exists \{u^k\} \subset C \text{ s.t. } \lim_{k \rightarrow \infty} u^k = u\}.$$

- The **interior** of a set $C \subset \mathbb{E}$ is

$$\text{int } C = \{u \in C : \exists \epsilon > 0 \text{ s.t. } B_\epsilon(u) \subset C\}.$$





Definition

- A set $C \subset \mathbb{E}$ is **open** if $\forall u \in C, \exists \epsilon > 0$ s.t. $B_\epsilon(u) \subset C$, where $B_\epsilon(u) := \{v \in \mathbb{E} : \|v - u\| < \epsilon\}$.
- A set $C \subset \mathbb{E}$ is **closed** if its complement $\mathbb{E} \setminus C$ is open.
- The **closure** of a set $C \subset \mathbb{E}$ is

$$\text{cl } C = \{u \in \mathbb{E} : \exists \{u^k\} \subset C \text{ s.t. } \lim_{k \rightarrow \infty} u^k = u\}.$$

- The **interior** of a set $C \subset \mathbb{E}$ is

$$\text{int } C = \{u \in C : \exists \epsilon > 0 \text{ s.t. } B_\epsilon(u) \subset C\}.$$

- The **relative interior** of a set $C \subset \mathbb{E}$ is

$$\begin{aligned} \text{rint } C &:= \{u \in C : \exists \epsilon > 0 \text{ s.t. } B_\epsilon(u) \cap \text{aff } C \subset C\} \\ &= \{u \in C : \forall v \in C, \exists \alpha > 1 \text{ s.t. } v + \alpha(u - v) \in C\} \end{aligned}$$

if C is convex. Here $\text{aff } C$ stands for the **affine hull** of C .



The following operations preserve the convexity:

- Intersection: $C_1 \cap C_2$.
- Summation: $C_1 + C_2 := \{u^1 + u^2 : u^1 \in C_1, u^2 \in C_2\}$.
- Closure: $\text{cl } C$.
- (Relative) interior: $\text{int } C$, $\text{rint } C$.

– In general, the union of convex sets is not convex.



The following operations preserve the convexity:

- Intersection: $C_1 \cap C_2$.
- Summation: $C_1 + C_2 := \{u^1 + u^2 : u^1 \in C_1, u^2 \in C_2\}$.
- Closure: $\text{cl } C$.
- (Relative) interior: $\text{int } C$, $\text{rint } C$.

– In general, the union of convex sets is not convex.

Convex cone

C is a **cone** if $C = \alpha C$ for any $\alpha > 0$.

C is a **convex cone** if C is a cone and is convex as well.

Separation of convex sets

Theorem (separation of convex sets)

Let C_1, C_2 be nonempty convex subsets in \mathbb{E} .

- ① Assume C_1 is closed and $C_2 = \{w\} \subset \mathbb{E} \setminus C_1$. Then $\exists v \in \mathbb{E}, v \neq 0, \alpha \in \mathbb{R}$ s.t.

$$\langle v, w \rangle > \alpha \geq \langle v, u \rangle, \quad \forall u \in C_1.$$

- ② Assume C_1 is open and $C_2 = \{w\} \subset \mathbb{E} \setminus C_1$. Then $\exists v \in \mathbb{E}, v \neq 0, \alpha \in \mathbb{R}$ s.t.

$$\langle v, w \rangle \geq \alpha \geq \langle v, u \rangle, \quad \forall u \in C_1.$$

- ③ Assume $C_1 \cap C_2 = \emptyset$ and C_1 is open. Then $\exists v \in \mathbb{E}, v \neq 0, \alpha \in \mathbb{R}$ s.t.

$$\langle v, u^1 \rangle \geq \alpha \geq \langle v, u^2 \rangle, \quad \forall u^1 \in C_1, u^2 \in C_2.$$

- ④ Assume $\emptyset \neq \text{int } C_1 \subset \mathbb{E} \setminus C_2$. Then $\exists v \in \mathbb{E}, v \neq 0, \alpha \in \mathbb{R}$ s.t.

$$\langle v, u^1 \rangle \geq \alpha \geq \langle v, u^2 \rangle, \quad \forall u^1 \in C_1, u^2 \in C_2.$$





Convex Function



- An **extended real-valued function** J maps from \mathbb{E} to $\bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$.
- The **domain** of $J : \mathbb{E} \rightarrow \bar{\mathbb{R}}$ is

$$\text{dom } J = \{u \in \mathbb{E} : J(u) < \infty\}.$$

- The function $J : \mathbb{E} \rightarrow \bar{\mathbb{R}}$ is **proper** if $\text{dom } J \neq \emptyset$.

Definition

We say $J : \mathbb{E} \rightarrow \bar{\mathbb{R}}$ is a **convex function** if

- 1 $\text{dom } J$ is a convex set.
- 2 For all $u, v \in \text{dom } J$ and $\alpha \in [0, 1]$ it holds that

$$J(\alpha u + (1 - \alpha)v) \leq \alpha J(u) + (1 - \alpha)J(v).$$

We say J is **strictly convex** if the above inequality is strict for all $\alpha \in (0, 1)$ and $u \neq v$.

Examples

- $J_{data}(u) = \|u - z\|_q^q$, where $q \geq 1$ and $\|\cdot\|_q$ is ℓ^q -norm.
- $J_{regu}(u) = \|Ku\|_{q'}^{q'}$, where K is linear transform and $q' \geq 1$.
- $J(u) = J_{data}(u) + \alpha J_{regu}(u)$.



Examples

- $J_{data}(u) = \|u - z\|_q^q$, where $q \geq 1$ and $\|\cdot\|_q$ is ℓ^q -norm.
- $J_{regu}(u) = \|Ku\|_{q'}^{q'}$, where K is linear transform and $q' \geq 1$.
- $J(u) = J_{data}(u) + \alpha J_{regu}(u)$.
- Negative entropy: $J_\epsilon(u) = \epsilon(u \log(u) + (1 - u) \log(1 - u))$.
- Soft-plus: $J_\epsilon^*(v) = \epsilon \log(1 + \exp(v/\epsilon)) \xrightarrow{\epsilon \rightarrow 0^+} \max(v, 0)$.



Examples

- $J_{data}(u) = \|u - z\|_q^q$, where $q \geq 1$ and $\|\cdot\|_q$ is ℓ^q -norm.
- $J_{regu}(u) = \|Ku\|_{q'}^{q'}$, where K is linear transform and $q' \geq 1$.
- $J(u) = J_{data}(u) + \alpha J_{regu}(u)$.
- Negative entropy: $J_\epsilon(u) = \epsilon(u \log(u) + (1 - u) \log(1 - u))$.
- Soft-plus: $J_\epsilon^*(v) = \epsilon \log(1 + \exp(v/\epsilon)) \xrightarrow{\epsilon \rightarrow 0^+} \max(v, 0)$.
- **Indicator function** ($C \subset \mathbb{E}$ is closed and convex):

$$\delta_C(u) = \begin{cases} 0 & \text{if } u \in C, \\ \infty & \text{otherwise.} \end{cases}$$

- Formulate *constrained optimization* with indicator function:
$$\min J(u) \text{ over } u \in C. \Leftrightarrow \min J(u) + \delta_C(u) \text{ over } u \in \mathbb{E}.$$





(As exercises)

- Any norm (over a normed vector space) is a convex function.
- J is a convex function and K is a linear transform $\Rightarrow J(K\cdot)$ is convex function.
- (Jensen's inequality) $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$ is convex iff

$$J\left(\sum_{i=1}^n \alpha_i u^i\right) \leq \sum_{i=1}^n \alpha_i J(u^i),$$

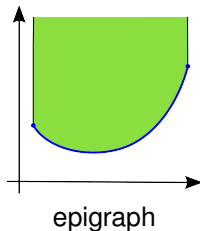
whenever $\{u^i\}_{i=1}^n \subset \mathbb{E}$, $\{\alpha_i\}_{i=1}^n \subset [0, 1]$, $\sum_{i=1}^n \alpha_i = 1$.

Epigraph

Definition

The **epigraph** of a proper function $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$ is

$$\text{epi } J = \{(u, \alpha) \in \mathbb{E} \times \mathbb{R} : J(u) \leq \alpha\}.$$



Theorem

A proper function $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$ is convex (resp. strictly convex) iff $\text{epi } J$ is a convex (resp. strictly convex) set.

Proof: as exercise.



Definition

Assume $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$ with $\text{rint dom } J \neq \emptyset$. We say J is **locally Lipschitz** at $u \in \text{rint dom } J$ with modulus $L_u > 0$ if there exists $\epsilon > 0$ s.t.

$$|J(u^1) - J(u^2)| \leq L_u \|u^1 - u^2\| \quad \forall u^1, u^2 \in B_\epsilon(u) \cap \text{rint dom } J.$$



Convex Set

Convex Function

Existence of Minimizer



Definition

Assume $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$ with $\text{rint dom } J \neq \emptyset$. We say J is **locally Lipschitz** at $u \in \text{rint dom } J$ with modulus $L_u > 0$ if there exists $\epsilon > 0$ s.t.

$$|J(u^1) - J(u^2)| \leq L_u \|u^1 - u^2\| \quad \forall u^1, u^2 \in B_\epsilon(u) \cap \text{rint dom } J.$$

Theorem

A proper convex function $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$ is locally Lipschitz at any $u \in \text{rint dom } J$.

Proof: found in script.

Convex Set

Convex Function

Existence of Minimizer



Existence of Minimizer

Global vs. Local minimizer

Recall the optimization of $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$:

$$\text{minimize } J(u) \quad \text{over } u \in \mathbb{E}.$$

Definition

- 1 $u^* \in \mathbb{E}$ is a **global minimizer** if $J(u^*) \leq J(u)$ for all $u \in \mathbb{E}$.
- 2 u^* is a **local minimizer** if $\exists \epsilon > 0$ s.t. $J(u^*) \leq J(u)$ for all $u \in B_\epsilon(u^*)$.
- 3 In the above definitions, a global/local minimizer is **strict** if $J(u^*) \leq J(u)$ is replaced by $J(u^*) < J(u)$.



Global vs. Local minimizer

Recall the optimization of $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$:

$$\text{minimize } J(u) \quad \text{over } u \in \mathbb{E}.$$



Definition

- 1 $u^* \in \mathbb{E}$ is a **global minimizer** if $J(u^*) \leq J(u)$ for all $u \in \mathbb{E}$.
- 2 u^* is a **local minimizer** if $\exists \epsilon > 0$ s.t. $J(u^*) \leq J(u)$ for all $u \in B_\epsilon(u^*)$.
- 3 In the above definitions, a global/local minimizer is **strict** if $J(u^*) \leq J(u)$ is replaced by $J(u^*) < J(u)$.

Theorem

For any proper convex function $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$, if $u^* \in \text{dom } J$ is a local minimizer of J , then it is also a global minimizer.

Proof: on board.

Does a minimizer always exist?

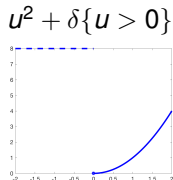
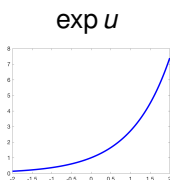
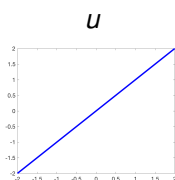


- Consider

$$\text{minimize } J(u) \quad \text{over } u \in \mathbb{E},$$

where $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$ is a proper, convex function.

- Some counterexamples for $J : \mathbb{R} \rightarrow \overline{\mathbb{R}}$:



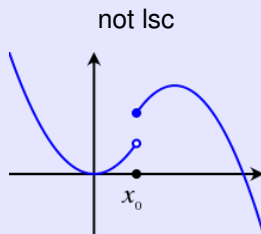
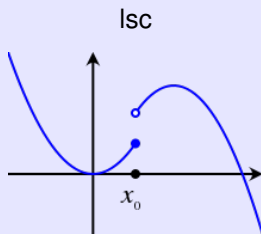
- We shall formalize our observations and derive sufficient conditions for existence.

Sufficient conditions for existence

Definition

- 1 J is **bounded from below** if $J(\cdot) \geq C$ for some $C \in \mathbb{R}$.
- 2 J is **coercive** if $J(u) \rightarrow \infty$ whenever $\|u\| \rightarrow \infty$.
 - Proposition: J is coercive if $\text{dom } J$ is bounded.
- 3 J is **lower semi-continuous** (lsc) at u^* if

$$J(u^*) \leq \liminf_{k \rightarrow \infty} J(u^k), \text{ whenever } u^k \rightarrow u^*.$$



- Proposition: J is lsc iff $\text{epi } J$ is closed.



Sufficient conditions for existence

Definition

- 1 J is **bounded from below** if $J(\cdot) \geq C$ for some $C \in \mathbb{R}$.
- 2 J is **coercive** if $J(u) \rightarrow \infty$ whenever $\|u\| \rightarrow \infty$.
- 3 J is **lower semi-continuous** (lsc) at u^* if

$$J(u^*) \leq \liminf_{k \rightarrow \infty} J(u^k), \text{ whenever } u^k \rightarrow u^*.$$

Theorem

Any proper function $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$, which is bounded from below, coercive, and lsc (everywhere), has a (global) minimizer.

Proof: on board.



Sufficient conditions for existence

Definition

- 1 J is **bounded from below** if $J(\cdot) \geq C$ for some $C \in \mathbb{R}$.
- 2 J is **coercive** if $J(u) \rightarrow \infty$ whenever $\|u\| \rightarrow \infty$.
- 3 J is **lower semi-continuous** (lsc) at u^* if

$$J(u^*) \leq \liminf_{k \rightarrow \infty} J(u^k), \text{ whenever } u^k \rightarrow u^*.$$

Theorem

Any proper function $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$, which is bounded from below, coercive, and lsc (everywhere), has a (global) minimizer.

Proof: on board.

Remarks for infinite dimensions

- 1 Weak compactness in reflexive Banach (e.g. Hilbert) sp.
- 2 J is convex and strongly lsc $\Rightarrow J$ is weakly lsc.





- Recall that a function $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$ is strictly convex if

$$J(\alpha u + (1 - \alpha)v) < \alpha J(u) + (1 - \alpha)J(v),$$

for all $u, v \in \text{dom } J$, $u \neq v$, $\alpha \in (0, 1)$.

Theorem

The minimizer of a strictly convex function $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$ is unique.

Proof: on board.