Chapter 1
Convex Analysis

Convex Optimization for Machine Learning & Computer Vision
SS 2017

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Computer Vision Group
Department of Informatics
TU München

Convex Set

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Optimization

Assume

- $\mathbb{E}$ is a Euclidean space (finite dimensional vector space), equipped with the inner product $\langle\cdot,\cdot\rangle$, e.g. $\langle u, v\rangle = u \cdot v$.
- C is a closed convex subset in $\mathbb{E}$.
- J is a convex objective function.

Convex optimization

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

First questions:

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

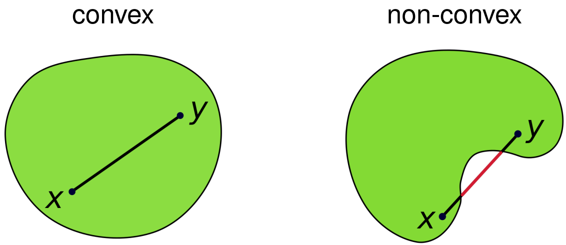
Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex set

Definition

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

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



Tao Wu
Thomas Möllenhoff
Emanuel Laude

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 $\mathbb{E} \backslash C$ is open.
- The **closure** of a set $C \subset \mathbb{E}$ is

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

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

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

Tao Wu
Thomas Möllenhoff
Emanuel Laude

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 $\mathbb{E} \backslash C$ is open.
- The **closure** of a set $C \subset \mathbb{E}$ is

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

- 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 <u>convex</u> set $C \subset \mathbb{E}$ is

$$\text{ri}\, C = \{u \in C : \forall v \in C, \exists \alpha > 1 \text{ s.t. } v + \alpha(u - v) \in C\}.$$

Basic properties

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: $\mathrm{cl}\, C$
- Interior: $\mathrm{int}\, C$

– The union of convex sets is not convex in general.

Basic properties

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: $\operatorname{cl} C$
- Interior: $\operatorname{int} C$

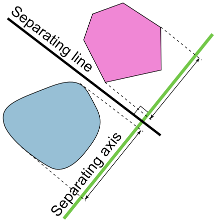
– The union of convex sets is not convex in general.

– *Polyhedral sets* are always convex; *cones* are not necessarily 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 convex subsets in $\mathbb{E}$ such that $C_1 \cap C_2 = \emptyset$ and C_1 is open. Then there exists a hyperplane separating C_1 and C_2, i.e. $\exists v \in \mathbb{E},\ \alpha \in \mathbb{R}$ such that

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

Proof: on board.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Separation of convex sets

Theorem (separation of convex sets)

Let C_1, C_2 be convex subsets in $\mathbb{E}$ such that $C_1 \cap C_2 = \emptyset$ and C_1 is open. Then there exists a hyperplane separating C_1 and C_2, i.e. $\exists v \in \mathbb{E},\ \alpha \in \mathbb{R}$ such that

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

Proof: on board.

Remarks

1. The proof works in any real Hilbert space.
2. Corollary: In a real Hilbert space, any (strongly) closed convex subset C is weakly closed.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Function

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex functions

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

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

- The function $J : \mathbb{E} \to \overline{\mathbb{R}}$ is **proper** if $\operatorname{dom} J \neq \emptyset$.

Definition

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

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

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

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

Examples

- $J_{data}(u) = \|u - z\|_p^p$ where $p \geq 1$ and $\|\cdot\|_p$ is the ℓ^p-norm.
- $J_{regu}(u) = \|Ku\|_q^q$ where $q \geq 1$ and K is linear transform.
- $J(u) = J_{data}(u) + \alpha J_{regu}(u)$ where $\alpha > 0$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Examples

- $J_{data}(u) = \|u - z\|_p^p$ where $p \geq 1$ and $\| \cdot \|_p$ is the ℓ^p-norm.
- $J_{regu}(u) = \|Ku\|_q^q$ where $q \geq 1$ and K is linear transform.
- $J(u) = J_{data}(u) + \alpha J_{regu}(u)$ where $\alpha > 0$.
- **Indicator function**:

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

  where C is a convex subset of $\mathbb{E}$.

- Alternative formulation of constrained optimization:

$$\min J(u) \text{ over } u \in C. \quad \Leftrightarrow \quad \min J(u) + \delta_C(u) \text{ over } u \in \mathbb{E}.$$

Basic facts

(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) $J : \mathbb{E} \to \overline{\mathbb{R}}$ is convex iff

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

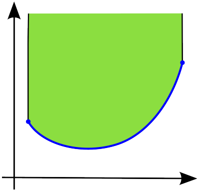
whenever $\{u^i\}_{i=1}^n \subset \mathbb{E}$, $\{\alpha_i\}_{i=1}^n \subset [0,1]$, $\sum_{i=1}^n \alpha_i = 1$.

Epigraph

Definition

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

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



epigraph

Theorem

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

Proof: as exercise.

Lipschitz continuity

Definition

Assume $J : U \to \mathbb{R}$, and U is a nonempty open subset of $\mathbb{E}$.

1. J is **(globally) Lipschitz** with modulus $L > 0$ if

$$|J(u^1) - J(u^2)| \leq L\|u^1 - u^2\| \quad \forall u^1, u^2 \in U.$$

2. J is **locally Lipschitz** at $u \in U$ with modulus $L_u > 0$ if there exists $\epsilon > 0$ s.t.

$$|J(u^1) - J(u^2)| \leq L_u\|u^1 - u^2\| \quad \forall u^1, u^2 \in B_\epsilon(u) \cap U.$$

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Lipschitz continuity

Definition

Assume $J : U \to \mathbb{R}$, and U is a nonempty open subset of $\mathbb{E}$.

1. J is **(globally) Lipschitz** with modulus $L > 0$ if

$$|J(u^1) - J(u^2)| \leq L\|u^1 - u^2\| \quad \forall u^1, u^2 \in U.$$

2. J is **locally Lipschitz** at $u \in U$ with modulus $L_u > 0$ if there exists $\epsilon > 0$ s.t.

$$|J(u^1) - J(u^2)| \leq L_u\|u^1 - u^2\| \quad \forall u^1, u^2 \in B_\epsilon(u) \cap U.$$

Theorem

A proper convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$ is locally Lipschitz at any $u \in \operatorname{ri} \operatorname{dom} J$.

Proof: on board.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Global vs. Local minimizers

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

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

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 $u^* \in \operatorname{dom} J$ and $\exists \epsilon > 0$ s.t. $J(u^*) \leq J(u)$ for all $u \in B_\epsilon(u^*)$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Global vs. Local minimizers

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

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

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 $u^* \in \operatorname{dom} J$ and $\exists \epsilon > 0$ s.t. $J(u^*) \leq J(u)$ for all $u \in B_\epsilon(u^*)$.

Theorem

Assume that $J : \mathbb{E} \to \overline{\mathbb{R}}$ is a proper convex function. Then any local minimizer of J is also global.

Proof: on board.