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
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Convex Set

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

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
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

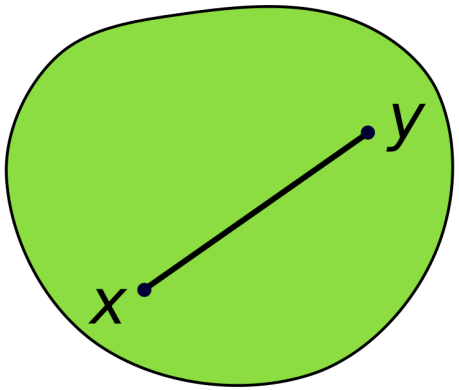
Convex set

Definition

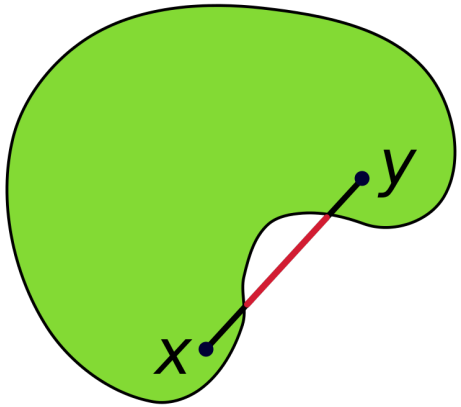
A set C is said to be **convex** if

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

convex



non-convex



Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

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

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

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 convex set $C \subset \mathbb{E}$ is

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

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

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.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

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: cl C
- Interior: int C

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

– *Polyhedral sets* are always convex; *cones* are not necessarily convex.

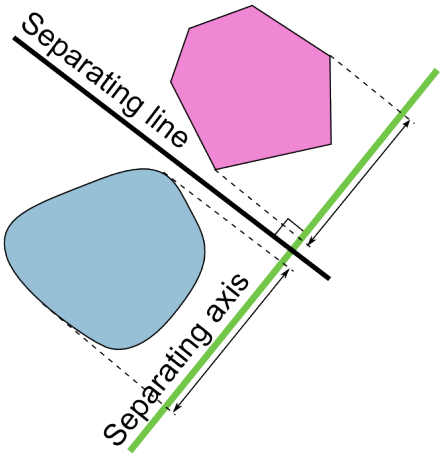
Convex cone

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.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Separation of convex sets



Theorem (separation of convex sets)

Let C_1, C_2 be nonempty convex subsets in $\mathbb{E}$ s.t. $C_1 \cap C_2 = \emptyset$ and C_1 is open. Then there exists a hyperplane separating C_1 and C_2, i.e. $\exists v \in \mathbb{E},\ v \neq 0,\ \alpha \in \mathbb{R}$ s.t.

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

Proof: on board.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Separation of convex sets

Theorem (separation of convex sets)

Let C_1, C_2 be nonempty convex subsets in $\mathbb{E}$ s.t. $C_1 \cap C_2 = \emptyset$ and C_1 is open. Then there exists a hyperplane separating C_1 and C_2, i.e. $\exists v \in \mathbb{E},\ v \neq 0,\ \alpha \in \mathbb{R}$ s.t.

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

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 Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Convex Function

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

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 $\alpha \in [0, 1]$ 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 $u \neq v$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Examples

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

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Examples

- $J_{data}(u) = \|u - f\|_q^q$ where $q \geq 1$ and $\|\cdot\|_q$ is the ℓ^q-norm.
- $J_{regu}(u) = \|Ku\|_q^q$ where 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:

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

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

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

Tao Wu
Thomas Möllenhoff
Emanuel Laude

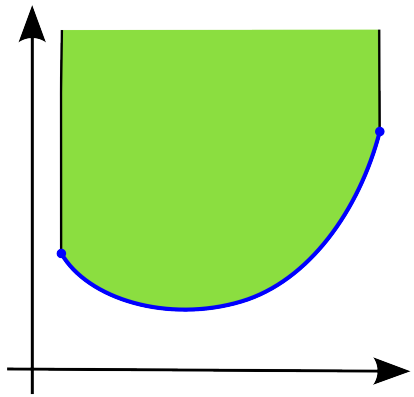
Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

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 $\operatorname{epi} J$ is a convex (resp. strictly convex) set.

Proof: as exercise.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Lipschitz continuity

Definition

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

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

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Lipschitz continuity

Definition

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

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

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.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

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 $\exists \epsilon > 0$ 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 strict inequality $J(u^*) < J(u)$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

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 $\exists \epsilon > 0$ 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 strict inequality $J(u^*) < J(u)$.

Theorem

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

Proof: on board.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Existence of Minimizer

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Does a minimizer always exist?

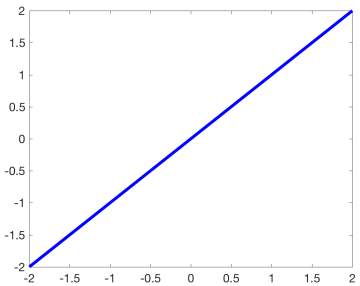
- Consider

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

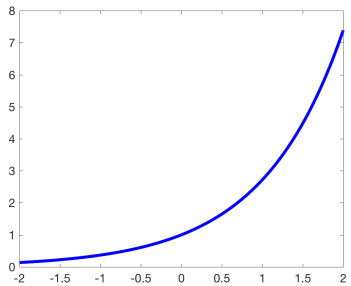
where $J : \mathbb{E} \to \overline{\mathbb{R}}$ is a proper, convex function.

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

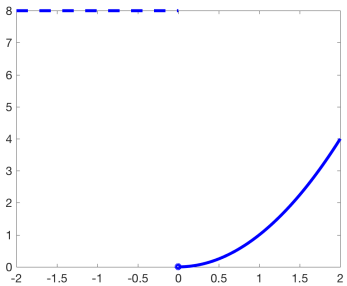
u



$\exp u$



$u^2 + \delta\{u > 0\}$



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

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

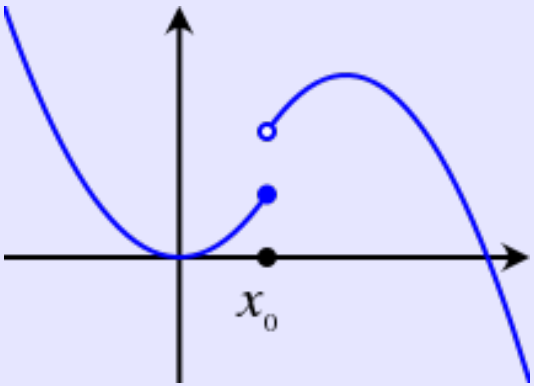
Sufficient conditions for existence

Definition

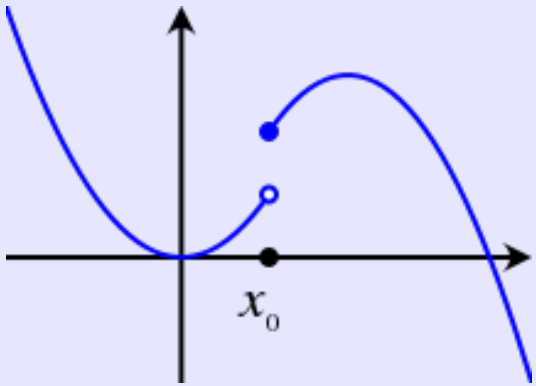
1. J is **bounded from below** if $J(\cdot) \geq C$ for some $C \in \mathbb{R}$.
2. J is **coercive** if $J(u) \to \infty$ whenever $\|u\| \to \infty$.
   - Proposition: J is coercive if dom J is bounded.
3. J is **lower-semicontinuous** (lsc) if

$$J(u^*) \leq \liminf_{u \to u^*} J(u).$$

lsc



not lsc



- Proposition: J is lsc iff epi J is closed.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

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 $J(u) \to \infty$ whenever $\|u\| \to \infty$.
3. J is **lower-semicontinuous** (lsc) if

$$J(u^*) \leq \liminf_{u \to u^*} J(u).$$

Theorem

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

Proof: on board.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

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 $J(u) \to \infty$ whenever $\|u\| \to \infty$.
3. J is **lower-semicontinuous** (lsc) if

$$J(u^*) \leq \liminf_{u \to u^*} J(u).$$

Theorem

Any proper function $J : \mathbb{E} \to \overline{\mathbb{R}}$, which is bounded from below, coercive, and lsc, 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

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Uniqueness

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

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

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

Theorem

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

Proof: on board.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

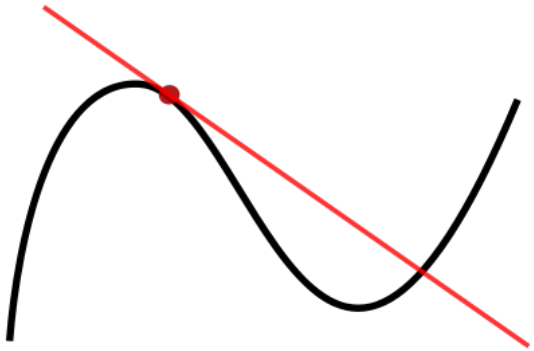
Subdifferential

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Differential



Definition

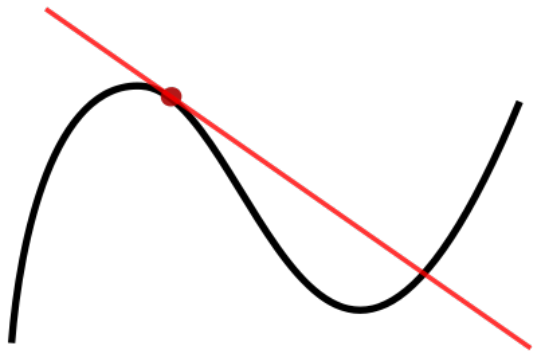
$J : \mathbb{E} \to \overline{\mathbb{R}}$ is called (Fréchet) **differentiable** at $u \in \operatorname{int} \operatorname{dom} J$ and $\nabla J(u) \in \mathbb{E}$ is the (Fréchet) **differential** of J at u if

$$\lim_{h \to 0,\, h \in (\operatorname{dom} J) - u} \frac{J(u+h) - J(u) - \langle \nabla J(u), h \rangle}{\|h\|} = 0.$$

J is said **continuously differentiable** at $u \in \operatorname{int} \operatorname{dom} J$ if $\nabla J(\cdot)$ is continuous on $\operatorname{dom} J \cap B_\epsilon(u)$ for some $\epsilon > 0$.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Differential



Definition

$J : \mathbb{E} \to \overline{\mathbb{R}}$ is called (Fréchet) **differentiable** at $u \in \operatorname{int} \operatorname{dom} J$ and $\nabla J(u) \in \mathbb{E}$ is the (Fréchet) **differential** of J at u if

$$\lim_{h \to 0,\, h \in (\operatorname{dom} J) - u} \frac{J(u+h) - J(u) - \langle \nabla J(u), h \rangle}{\|h\|} = 0.$$

J is said **continuously differentiable** at $u \in \operatorname{int} \operatorname{dom} J$ if $\nabla J(\cdot)$ is continuous on $\operatorname{dom} J \cap B_\epsilon(u)$ for some $\epsilon > 0$.

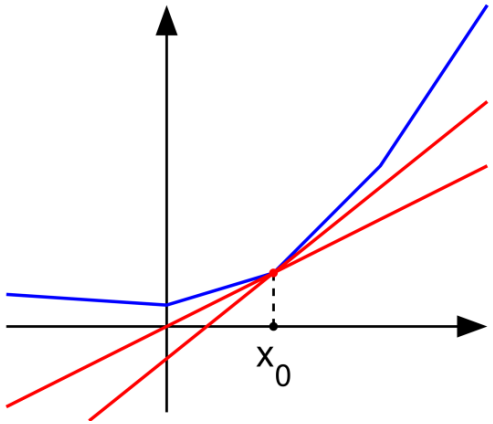
Remark

In functional analysis, $\nabla J(u)$ is treated as a *dual* object in $\mathbb{E}^*$, and $\langle \nabla J(u), h \rangle_{\mathbb{E}^*, \mathbb{E}}$ as *duality pairing*.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Subdifferential

Now we generalize differentiability from differentiable functions to nonsmooth (convex) functions.



Definition

The **subdifferential** of a convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$ at $u \in \operatorname{dom} J$ is defined by

$$\partial J(u) = \{\xi \in \mathbb{E} : J(v) \geq J(u) + \langle \xi, v - u \rangle\}.$$

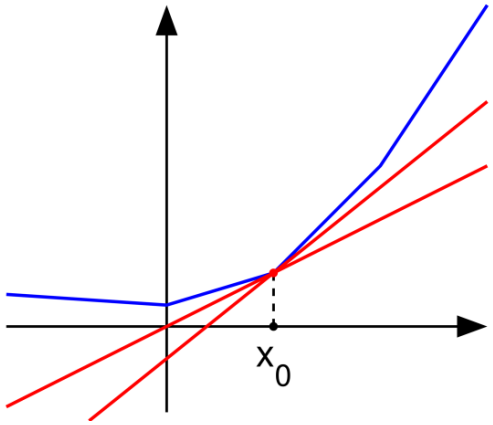
Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Subdifferential

Now we generalize differentiability from differentiable functions to nonsmooth (convex) functions.



Definition

The **subdifferential** of a convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$ at $u \in \operatorname{dom} J$ is defined by

$$\partial J(u) = \{\xi \in \mathbb{E} : J(v) \geq J(u) + \langle \xi, v - u \rangle\}.$$

Geometric interpretation

$\xi \in \partial J(u)$ iff $(\xi, -1)$ is a normal vector for the supporting hyperplane of epi J at $(u, J(u))$.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Examples and basic facts

Basic facts

1. $\partial J(\cdot)$ is a **set-valued map**.
2. If J is cont. differentiable at u, then $\partial J(u) = \{\nabla J(u)\}$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Examples and basic facts

Basic facts

1. $\partial J(\cdot)$ is a **set-valued map**.
2. If J is cont. differentiable at u, then $\partial J(u) = \{\nabla J(u)\}$.

Examples (as exercises)

1. $u \in \mathbb{R}^n \mapsto \|u\|_1$.
2. $u \in \mathbb{R}^n \mapsto \|u\|_\infty$.
3. $X \in \mathbb{R}^{n\times n} \mapsto \|X\|_{1,2} = \sum_i \left(\sum_j |X_{i,j}|^2\right)^{1/2}$.
4. $X \in \mathbb{R}^{n\times n} \mapsto \|X\|_{nuc} = \sum_i \sigma_i(X)$ (sum of singular values).
5. Given a closed convex subset $C \subset \mathbb{E}$ and $u \in C$,

$$\partial \delta_C(u) = N_C(u) = \{\xi \in \mathbb{E} : \langle \xi, v - u\rangle \leq 0 \ \forall v \in C\},$$

known as the **normal cone** of C at u.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Subdifferential calculus

Theorem (chain rule under linear transform)

Let $\widetilde{J}(\cdot) = J(K\cdot)$ with some convex function J and linear transform K. Then

$$\partial\widetilde{J}(u) = K^\top \partial J(Ku)$$

for any $u \in \operatorname{ri}\operatorname{dom} J$.

Example: $J(u) = \|Ku\|_1 \Rightarrow \partial J(u) = K^\top \partial\|\cdot\|_1(Ku)$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Subdifferential calculus

Theorem (chain rule under linear transform)

Let $\widetilde{J}(\cdot) = J(K\cdot)$ with some convex function J and linear transform K. Then

$$\partial\widetilde{J}(u) = K^\top \partial J(Ku)$$

for any $u \in \operatorname{ri}\operatorname{dom} J$.

Example: $J(u) = \|Ku\|_1 \Rightarrow \partial J(u) = K^\top \partial\|\cdot\|_1(Ku)$.

Theorem (summation rule)

If $J(\cdot) = J_1(\cdot) + J_2(\cdot)$ for some convex functions J_1 and J_2, then

$$\partial\widetilde{J}(u) = \partial J_1(u) + \partial J_2(u)$$

for any $u \in \operatorname{ri}\operatorname{dom} J_1 \cap \operatorname{ri}\operatorname{dom} J_2$.

Warning: not true if J_1 or J_2 is non-convex, e.g. $0 = |\cdot| + (-|\cdot|)$.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Properties of subdifferential map

Theorem

Let $J : \mathbb{E} \to \overline{\mathbb{R}}$ be a convex function. Then ∂J is a **monotone operator**, i.e. $\forall u^1, u^2 \in \operatorname{dom} J,\ \xi^1 \in \partial J(u^1),\ \xi^2 \in \partial J(u^2)$:

$$\left\langle \xi^1 - \xi^2, u^1 - u^2 \right\rangle \geq 0.$$

Proof: on board.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Properties of subdifferential map

Theorem

Let $J : \mathbb{E} \to \overline{\mathbb{R}}$ be a convex function. Then ∂J is a **monotone operator**, i.e. $\forall u^1, u^2 \in \operatorname{dom} J,\ \xi^1 \in \partial J(u^1),\ \xi^2 \in \partial J(u^2)$:

$$\left\langle \xi^1 - \xi^2, u^1 - u^2 \right\rangle \geq 0.$$

Proof: on board.

Theorem

Let $J : \mathbb{E} \to \overline{\mathbb{R}}$ be a convex function. Then for any $u \in \operatorname{int} \operatorname{dom} J$, $\partial J(u)$ is a nonempty, compact, and convex subset.

Proof: on board.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Properties of subdifferential map

Theorem

Let $J : \mathbb{E} \to \overline{\mathbb{R}}$ be a convex function. Then ∂J is a **monotone operator**, i.e. $\forall u^1, u^2 \in \operatorname{dom} J,\ \xi^1 \in \partial J(u^1),\ \xi^2 \in \partial J(u^2)$:

$$\left\langle \xi^1 - \xi^2, u^1 - u^2 \right\rangle \geq 0.$$

Proof: on board.

Theorem

Let $J : \mathbb{E} \to \overline{\mathbb{R}}$ be a convex function. Then for any $u \in \operatorname{int} \operatorname{dom} J$, $\partial J(u)$ is a nonempty, compact, and convex subset.

Proof: on board.

Theorem

Let $J : \mathbb{E} \to \overline{\mathbb{R}}$ be a proper, convex, lsc function. Then the set-valued map $\partial J(\cdot)$ is **closed**, i.e. $\xi^* \in \partial J(u^*)$ whenever

$$\exists (u^k, \xi^k) \to (u^*, \xi^*) \in (\operatorname{ri} \operatorname{dom} J) \times \mathbb{E} \ \text{ s.t. } \xi^k \in \partial J(u^k)\ \forall k.$$

Proof: on board.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Optimality condition

Theorem

Given any proper convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$, the sufficient and necessary condition for u^* being a (global) minimizer for J is

$$0 \in \partial J(u^*).$$

Proof: on board.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Optimality condition

Theorem

Given any proper convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$, the sufficient and necessary condition for u^* being a (global) minimizer for J is

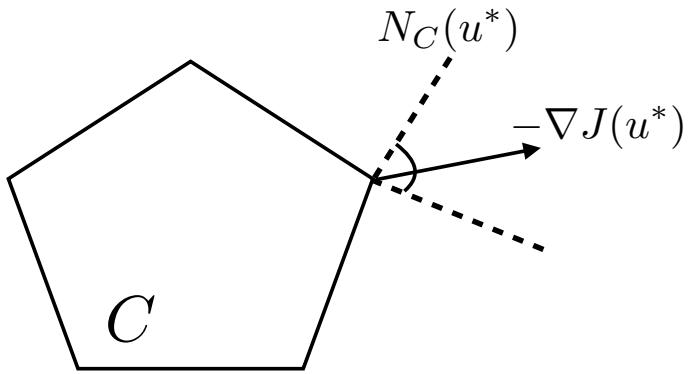
$$0 \in \partial J(u^*).$$

Proof: on board.

Constrained optimization as special case

If $\widetilde{J}(u) = J(u) + \delta_C(u)$ with convex function $J : \mathbb{E} \to \mathbb{R}$ and closed convex subset $C \in \mathbb{E}$, then $0 \in \partial \widetilde{J}(u^*) \rightsquigarrow$

$$0 \in \partial J(u^*) + N_C(u^*).$$



Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Optimality condition

Theorem

Given any proper convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$, the sufficient and necessary condition for u^* being a (global) minimizer for J is

$$0 \in \partial J(u^*).$$

Proof: on board.

Constrained optimization as special case

If $\widetilde{J}(u) = J(u) + \delta_C(u)$ with convex function $J : \mathbb{E} \to \mathbb{R}$ and closed convex subset $C \in \mathbb{E}$, then $0 \in \partial \widetilde{J}(u^*) \rightsquigarrow$

$$0 \in \partial J(u^*) + N_C(u^*).$$

Remark

$0 \in \partial J(u^*) + N_C(u^*)$ is a *geometric* optimality condition. Further characterization replies on the algebraic representation of $N_C(u^*)$, e.g. the Karush-Kuhn-Tucker (KKT) conditions under certain constraint qualifications.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Convex Conjugate

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Legendre-Fenchel transform

Definition

Given a convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$, the **convex conjugate** of J is defined by

$$J^*(p) = \sup_{u \in \mathbb{E}} \langle u, p \rangle - J(u).$$

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Legendre-Fenchel transform

Definition

Given a convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$, the **convex conjugate** of J is defined by

$$J^*(p) = \sup_{u \in \mathbb{E}} \langle u, p \rangle - J(u).$$

Examples (as exercises)

1. $J(\cdot) = \langle w, \cdot \rangle \;\Rightarrow\; J^*(\cdot) = \delta\{\cdot = w\}$. $(w \in \mathbb{E})$
2. $J(\cdot) = \frac{1}{q}\|\cdot\|_q^q \;\Rightarrow\; J^*(\cdot) = \frac{1}{q'}\|\cdot\|_{q'}^{q'}$. $(0 < q < 1, \frac{1}{q} + \frac{1}{q'} = 1)$
3. $J(\cdot) = \|\cdot\| \;\Rightarrow\; J^*(\cdot) = \delta\{\|\cdot\|_* \leq 1\}$. ($\|\cdot\|_*$ is the *dual norm* of $\|\cdot\|$, i.e. $\|v\|_* = \sup_{\|u\| \leq 1} \langle u, v \rangle$)
4. $J(\cdot) = \delta\{\|\cdot\|_* \leq 1\} \;\Rightarrow\; J^*(\cdot) = \|\cdot\|$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Legendre-Fenchel transform

Definition

Given a convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$, the **convex conjugate** of J is defined by

$$J^*(p) = \sup_{u \in \mathbb{E}} \langle u, p \rangle - J(u).$$

Examples (as exercises)

1. $J(\cdot) = \langle w, \cdot \rangle \;\Rightarrow\; J^*(\cdot) = \delta\{\cdot = w\}$. $(w \in \mathbb{E})$
2. $J(\cdot) = \frac{1}{q}\|\cdot\|_q^q \;\Rightarrow\; J^*(\cdot) = \frac{1}{q'}\|\cdot\|_{q'}^{q'}$. $(0 < q < 1, \frac{1}{q} + \frac{1}{q'} = 1)$
3. $J(\cdot) = \|\cdot\| \;\Rightarrow\; J^*(\cdot) = \delta\{\|\cdot\|_* \leq 1\}$. ($\|\cdot\|_*$ is the *dual norm* of $\|\cdot\|$, i.e. $\|v\|_* = \sup_{\|u\| \leq 1} \langle u, v \rangle$)
4. $J(\cdot) = \delta\{\|\cdot\|_* \leq 1\} \;\Rightarrow\; J^*(\cdot) = \|\cdot\|$.

Basic facts

- Scalar multiplication: $\widetilde{J}(\cdot) = \alpha J(\cdot) \;\Rightarrow\; \widetilde{J}^*(\cdot) = \alpha J^*(\cdot/\alpha)$.
- Translation: $\widetilde{J}(\cdot) = J(\cdot - z) \;\Rightarrow\; \widetilde{J}^*(\cdot) = J^*(\cdot) + \langle \cdot, z \rangle$.

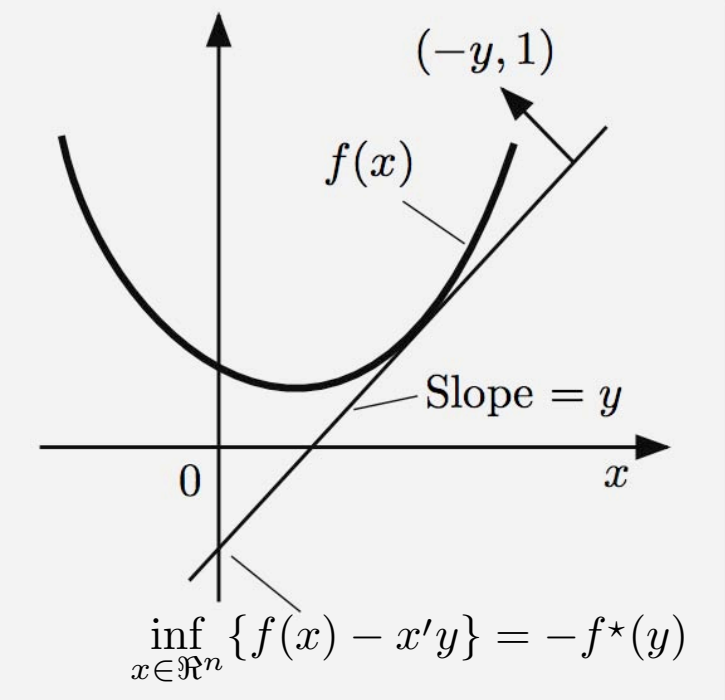
Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Geometric interpretation

Convex conjugation maps:
(non-vertical) supporting hyperplane of the epigraph
to:
intersection with the vertical axis.



Courtesy of Bertsekas

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Fenchel-Young inequality

Theorem (Fenchel-Young inequality)

For all $u \in \operatorname{dom} J,\ p \in \operatorname{dom} J^*$, we have

$$J(u) + J^*(p) \geq \langle u, p \rangle .$$

The equality holds iff $p \in \partial J(u)$.

Proof: $J(u) + J^*(p) \geq \langle u, p \rangle$ follows directly from the definition of convex conjugate; $p \in \partial J(u)$ is the sufficient and necessary condition for: $u = \arg\min_{v \in \mathbb{E}} J(v) - \langle v, p \rangle$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Biconjugate

Theorem

Assume $J : \mathbb{E} \to \overline{\mathbb{R}}$ and $J^{**} = (J^*)^*$ is the **biconjugate** of J.

In general:

1. $J^{**}(\cdot) \leq J(\cdot)$.
2. J^* is convex and lsc.

If J is proper, convex, and lsc, then:

3. $J^{**}(\cdot) = J(\cdot)$.
4. $p \in \partial J(u)$ iff $u \in \partial J^*(p)$.

Proof: on board.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Regularity of J and J^*

Definition

$J : \mathbb{E} \to \mathbb{R}$ is said to be μ-**strongly convex** if $\exists \mu > 0$ s.t. $J(\cdot) - \frac{\mu}{2}\|\cdot\|^2$ is convex.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Regularity of J and J^*

Definition

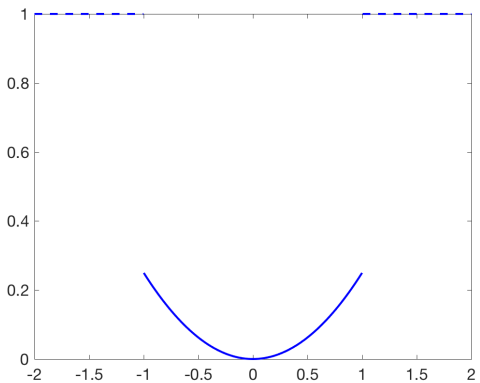
$J : \mathbb{E} \to \mathbb{R}$ is said to be μ-**strongly convex** if $\exists \mu > 0$ s.t. $J(\cdot) - \frac{\mu}{2}\|\cdot\|^2$ is convex.

Theorem

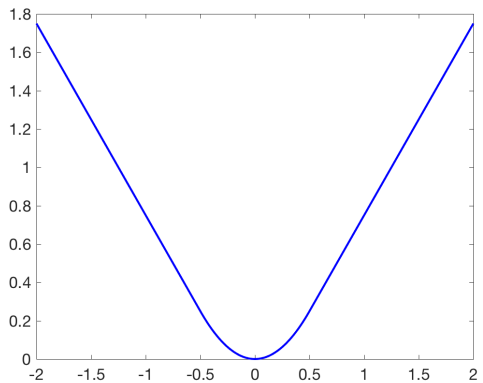
Assume that $J : \mathbb{E} \to \overline{\mathbb{R}}$ is proper, convex, and lsc. Then J is μ-strongly convex iff J^* has $\frac{1}{\mu}$-Lipschitz gradient.

Proof: on board.

$J =$ truncated quadratic



$J^* =$ Huber function



Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Duality Theory

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Fenchel-Rockafellar duality

- Consider

$$\min_{u \in \mathbb{R}^n} F(Ku) + G(u),$$

where $K \in \mathbb{R}^{m \times n}$, and F, G are proper, convex, and lsc.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Fenchel-Rockafellar duality

- Consider

$$\min_{u \in \mathbb{R}^n} F(Ku) + G(u),$$

where $K \in \mathbb{R}^{m \times n}$, and F, G are proper, convex, and lsc.

- The **weak duality** always holds:

$$\begin{aligned}
\mathcal{P}^* &= \min_{u \in \mathbb{R}^n} F(Ku) + G(u) = \min_u \sup_p \langle p, Ku \rangle - F^*(p) + G(u) \\
&\geq \sup_p \inf_u \left\langle K^\top p, u \right\rangle + G(u) - F^*(p) \\
&= \sup_p -G^*(-K^\top p) - F^*(p) = \mathcal{D}^*.
\end{aligned}$$

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Fenchel-Rockafellar duality

- Consider

$$\min_{u \in \mathbb{R}^n} F(Ku) + G(u),$$

where $K \in \mathbb{R}^{m \times n}$, and F, G are proper, convex, and lsc.

- The **weak duality** always holds:

$$\begin{aligned}
\mathcal{P}^* &= \min_{u \in \mathbb{R}^n} F(Ku) + G(u) = \min_u \sup_p \langle p, Ku \rangle - F^*(p) + G(u) \\
&\geq \sup_p \inf_u \left\langle K^\top p, u \right\rangle + G(u) - F^*(p) \\
&= \sup_p -G^*(-K^\top p) - F^*(p) = \mathcal{D}^*.
\end{aligned}$$

- Define the **duality gap**:

$$\mathcal{G}(u, p) = F(Ku) + G(u) + G^*(-K^\top p) + F^*(p).$$

Note that $\mathcal{G}(u, p) = 0$ is an optimality criterion.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Fenchel-Rockafellar duality

- $\mathcal{G}(u^*, p^*) = 0 \Leftrightarrow \mathcal{P}^* = \mathcal{D}^* \Leftrightarrow (u^*, p^*)$ solves the **saddle point problem** with $\mathcal{L}(u, p) := \langle p, Ku \rangle - F^*(p) + G(u)$:

$$\mathcal{L}(u^*, p) \leq \mathcal{L}(u^*, p^*) \leq \mathcal{L}(u, p^*) \quad \forall (u, p).$$

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Fenchel-Rockafellar duality

- $\mathcal{G}(u^*, p^*) = 0 \Leftrightarrow \mathcal{P}^* = \mathcal{D}^* \Leftrightarrow (u^*, p^*)$ solves the **saddle point problem** with $\mathcal{L}(u, p) := \langle p, Ku \rangle - F^*(p) + G(u)$:

$$\mathcal{L}(u^*, p) \leq \mathcal{L}(u^*, p^*) \leq \mathcal{L}(u, p^*) \quad \forall (u, p).$$

Theorem (Fenchel-Rockafellar duality)

Assume $\exists \bar{u} \in \operatorname{dom} G$ s.t. F is continuous at $K\bar{u}$. Then the **strong duality** holds: $\mathcal{P}^* = \mathcal{D}^*$. Moreover, (u^*, p^*) is the optimal solution pair iff

$$\begin{cases} Ku^* \in \partial F^*(p^*), \\ -K^\top p^* \in \partial G(u^*). \end{cases}$$

Proof: on board.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Application: Total-variation image restoration

- Primal problem (let $\alpha > 0,\ q \in [1, \infty],\ \Omega \subset \mathbb{R}^d,\ f \in \mathbb{R}^\Omega$):

$$\min_{u \in \mathbb{R}^\Omega} \alpha \|\nabla u\|_{1,q} + \frac{1}{2}\|u - f\|^2.$$

Here $\|p\|_{1,q} = \sum_{j \in \Omega} |p_j|_{\ell^q}$ for each $p \in (\mathbb{R}^\Omega)^d$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Application: Total-variation image restoration

- Primal problem (let $\alpha > 0,\ q \in [1, \infty],\ \Omega \subset \mathbb{R}^d,\ f \in \mathbb{R}^\Omega$):

$$\min_{u \in \mathbb{R}^\Omega} \alpha \|\nabla u\|_{1,q} + \frac{1}{2}\|u - f\|^2.$$

  Here $\|p\|_{1,q} = \sum_{j \in \Omega} |p_j|_{\ell^q}$ for each $p \in (\mathbb{R}^\Omega)^d$.

- Apply the Fenchel-Rockafellar duality with

$$F(\cdot) = \alpha \|\cdot\|_{1,q}, \quad K = \nabla, \quad G(\cdot) = \frac{1}{2}\|\cdot - f\|^2.$$

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Application: Total-variation image restoration

- Primal problem (let $\alpha > 0,\ q \in [1, \infty],\ \Omega \subset \mathbb{R}^d,\ f \in \mathbb{R}^\Omega$):

$$\min_{u \in \mathbb{R}^\Omega} \alpha \|\nabla u\|_{1,q} + \frac{1}{2}\|u - f\|^2.$$

  Here $\|p\|_{1,q} = \sum_{j \in \Omega} |p_j|_{\ell^q}$ for each $p \in (\mathbb{R}^\Omega)^d$.

- Apply the Fenchel-Rockafellar duality with

$$F(\cdot) = \alpha \|\cdot\|_{1,q}, \quad K = \nabla, \quad G(\cdot) = \frac{1}{2}\|\cdot - f\|^2.$$

- Saddle point problem ($1/q + 1/q' = 1$):

$$\max_p \min_u \langle p, \nabla u \rangle - \delta\{\|p\|_{\infty,q'} \leq \alpha\} + \frac{1}{2}\|u - f\|^2.$$

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Application: Total-variation image restoration

- Primal problem (let $\alpha > 0,\ q \in [1, \infty],\ \Omega \subset \mathbb{R}^d,\ f \in \mathbb{R}^\Omega$):

$$\min_{u \in \mathbb{R}^\Omega} \alpha \|\nabla u\|_{1,q} + \frac{1}{2}\|u - f\|^2.$$

  Here $\|p\|_{1,q} = \sum_{j \in \Omega} |p_j|_{\ell^q}$ for each $p \in (\mathbb{R}^\Omega)^d$.

- Apply the Fenchel-Rockafellar duality with

$$F(\cdot) = \alpha \|\cdot\|_{1,q}, \quad K = \nabla, \quad G(\cdot) = \frac{1}{2}\|\cdot - f\|^2.$$

- Saddle point problem ($1/q + 1/q' = 1$):

$$\max_p \min_u \langle p, \nabla u \rangle - \delta\{\|p\|_{\infty,q'} \le \alpha\} + \frac{1}{2}\|u - f\|^2.$$

- Dual problem:

$$\min_p \frac{1}{2}\|\nabla^\top p - f\|^2 + \delta\{\|p\|_{\infty,q'} \le \alpha\}.$$

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Proximal Operator

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Proximal operator

Definition

Given a proper, convex, lsc function $J : \mathbb{E} \to \overline{\mathbb{R}}$, we define the **proximal operator** of J by

$$\text{prox}_{\tau J}(v) = \arg\min_u J(u) + \frac{1}{2\tau}\|u - v\|^2.$$

Note: The above minimization always has a unique minimizer.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Proximal operator

Definition

Given a proper, convex, lsc function $J : \mathbb{E} \to \overline{\mathbb{R}}$, we define the **proximal operator** of J by

$$\text{prox}_{\tau J}(v) = \arg\min_u J(u) + \frac{1}{2\tau}\|u - v\|^2.$$

Note: The above minimization always has a unique minimizer.

Proximal operator in algorithms

Let us solve the convex optimization:

$$\min_u F(u) + G(u),$$

where G is cont'ly differentiable but F is non-differentiable.

The **proximal gradient** iteration appears as:

$$u^{k+1} = \text{prox}_{\tau F}(u^k - \tau \nabla G(u^k)).$$

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Observations

1. By checking the optimality condition,

$$u = \text{prox}_{\tau J}(v) \Leftrightarrow 0 \in \tau \partial J(u) + u - v \Leftrightarrow u = (I + \tau \partial J)^{-1}(v).$$

Thus, $\text{prox}_{\tau J} = (I + \tau \partial J)^{-1}$, known as the **resolvent** of ∂J.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Observations

1. By checking the optimality condition,

   $$u = \text{prox}_{\tau J}(v) \Leftrightarrow 0 \in \tau\partial J(u) + u - v \Leftrightarrow u = (I + \tau\partial J)^{-1}(v).$$

   Thus, $\text{prox}_{\tau J} = (I + \tau\partial J)^{-1}$, known as the **resolvent** of ∂J.

2. u^* is a **fixed point** of $\text{prox}_{\tau J}$, i.e. $u^* = \text{prox}_{\tau J}(u^*)$,
   $\Leftrightarrow 0 \in \partial J(u^*)$, i.e. u^* is a minimizer of J.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Observations

1. By checking the optimality condition,

   $$u = \text{prox}_{\tau J}(v) \Leftrightarrow 0 \in \tau\partial J(u) + u - v \Leftrightarrow u = (I + \tau\partial J)^{-1}(v).$$

   Thus, $\text{prox}_{\tau J} = (I + \tau\partial J)^{-1}$, known as the **resolvent** of ∂J.

2. u^* is a **fixed point** of $\text{prox}_{\tau J}$, i.e. $u^* = \text{prox}_{\tau J}(u^*)$,
   $\Leftrightarrow 0 \in \partial J(u^*)$, i.e. u^* is a minimizer of J.

3. Derivation of proximal gradient algorithm:

   $$\begin{aligned}
   & u^* \in \arg\min_u F(u) + G(u) \\
   \Leftrightarrow\ & 0 \in \partial F(u^*) + \nabla G(u^*) \\
   \Leftrightarrow\ & u^* + \tau\partial F(u^*) \in u^* - \tau\nabla G(u^*) \\
   \Leftrightarrow\ & u^* = (I + \tau\partial F)^{-1}(u^* - \tau\nabla G(u^*)) \\
   \rightsquigarrow\ & u^{k+1} = \text{prox}_{\tau F}(u^k - \tau\nabla G(u^k)).
   \end{aligned}$$

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Logistic regression (programming exercise)

- MNIST dataset[1] - handwritten digit recognition.
- Train classifier on training set; Evaluate on test set.
- Conv. neural network: 0.23%; **Logistic regression**: 10%.

[1] http://yann.lecun.com/exdb/mnist/

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

updated 23.05.2017

Logistic regression (programming exercise)

- MNIST dataset[1] - handwritten digit recognition.
- Train classifier on training set; Evaluate on test set.
- Conv. neural network: 0.23%; **Logistic regression**: 10%.
- $K = 10$ classes; grayscale images of pixels $M = 28 \times 28$.
- $N = 60000$ training images $X \in \mathbb{R}^{N \times M}$, with ground-truth labels $Y \in \{1, ..., K\}^N$.
- Train **linear classifier** $W \in \mathbb{R}^{K \times M}$ on raw pixel data.

[1] http://yann.lecun.com/exdb/mnist/

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Logistic regression (programming exercise)

- MNIST dataset[1] - handwritten digit recognition.
- Train classifier on training set; Evaluate on test set.
- Conv. neural network: 0.23%; **Logistic regression**: 10%.
- $K = 10$ classes; grayscale images of pixels $M = 28 \times 28$.
- $N = 60000$ training images $X \in \mathbb{R}^{N \times M}$, with ground-truth labels $Y \in \{1, ..., K\}^N$.
- Train **linear classifier** $W \in \mathbb{R}^{K \times M}$ on raw pixel data.
- **Softmax** function:

$$\mathcal{P}(Y_n = k | X_{n,\cdot}, W) = \frac{\exp(\langle W_{k,\cdot}, X_{n,\cdot} \rangle)}{\sum_{k=1}^{K} \exp(\langle W_{k,\cdot}, X_{n,\cdot} \rangle)}.$$

- Minimize negative log-likelihood + regularizer:

$$\min_W \alpha \|W\|_1 - \frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{K} \mathbf{1}\{Y_n = k\} \log \mathcal{P}(Y_n = k | X_{n,\cdot}, W).$$

[1] http://yann.lecun.com/exdb/mnist/

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Examples

1. Indicator function: Let C be a closed, convex subset $\Rightarrow$

$$\text{prox}_{\tau\delta_C}(\cdot) = \text{proj}_C(\cdot).$$

As a result, proximal gradient $\Leftrightarrow$ projected gradient.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Examples

1. Indicator function: Let C be a closed, convex subset $\Rightarrow$

$$\text{prox}_{\tau\delta_C}(\cdot) = \text{proj}_C(\cdot).$$

   As a result, proximal gradient $\Leftrightarrow$ projected gradient.

2. Linear approximation: $\widehat{J}(\cdot) := J(\bar{u}) + \langle \nabla J(\bar{u}), \cdot - \bar{u} \rangle \Rightarrow$

$$\begin{aligned} \text{prox}_{\tau\widehat{J}}(\bar{u}) &= \arg\min_u \frac{1}{2\tau}\|u - \bar{u}\|^2 + \langle \nabla J(\bar{u}), u - \bar{u} \rangle \\ &= \bar{u} - \tau \nabla J(\bar{u}). \quad \text{(gradient descent step)} \end{aligned}$$

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Examples

1. Indicator function: Let C be a closed, convex subset $\Rightarrow$

$$\text{prox}_{\tau\delta_C}(\cdot) = \text{proj}_C(\cdot).$$

As a result, proximal gradient $\Leftrightarrow$ projected gradient.

2. Linear approximation: $\widehat{J}(\cdot) := J(\bar{u}) + \langle \nabla J(\bar{u}), \cdot - \bar{u} \rangle \Rightarrow$

$$\begin{aligned}\text{prox}_{\tau\widehat{J}}(\bar{u}) &= \arg\min_u \frac{1}{2\tau}\|u-\bar{u}\|^2 + \langle \nabla J(\bar{u}), u-\bar{u}\rangle \\ &= \bar{u} - \tau\nabla J(\bar{u}). \quad \text{(gradient descent step)}\end{aligned}$$

3. Quadratic approximation:
$\widetilde{J}(\cdot) := J(\bar{u}) + \langle \nabla J(\bar{u}), \cdot - \bar{u}\rangle + \frac{1}{2}\left\langle \nabla^2 J(\bar{u})(\cdot - \bar{u}), \cdot - \bar{u}\right\rangle \Rightarrow$

$$\begin{aligned}\text{prox}_{\tau\widetilde{J}}(\bar{u}) = \arg\min_u &\frac{1}{2\tau}\|u-\bar{u}\|^2 + \langle \nabla J(\bar{u}), u-\bar{u}\rangle \\ &+ \frac{1}{2}\left\langle \nabla^2 J(\bar{u})(u-\bar{u}), u-\bar{u}\right\rangle \\ = \bar{u} - (\nabla^2 J(\bar{u}) + \frac{1}{\tau}I)^{-1}&\nabla J(\bar{u}). \quad \text{(damped Newton step)}\end{aligned}$$

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Moreau identity

Theorem (Moreau identity)

Let $\tau > 0$ and $J : \mathbb{E} \to \overline{\mathbb{R}}$ be proper, convex, and lsc. Then the following identity holds:

$$\mathrm{id}(\cdot) = \mathrm{prox}_{\tau J}(\cdot) + \tau \, \mathrm{prox}_{\frac{1}{\tau} J^*}(\cdot/\tau).$$

In particular, $\tau = 1 \Rightarrow \mathrm{id}(\cdot) = \mathrm{prox}_{J}(\cdot) + \mathrm{prox}_{J^*}(\cdot)$.

Proof:

$$\begin{aligned}
& v = \tau \, \mathrm{prox}_{\frac{1}{\tau} J^*}(u/\tau) \\
\Leftrightarrow \quad & \left(I + \frac{1}{\tau} \partial J^* \right)^{-1} (u/\tau) \ni v/\tau \\
\Leftrightarrow \quad & \partial J^*(v/\tau) \ni u - v \\
\Leftrightarrow \quad & v/\tau \in \partial J(u - v) \\
\Leftrightarrow \quad & u - v = (I + \tau \partial J)^{-1}(u) = \mathrm{prox}_{\tau J}(u).
\end{aligned}$$

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Moreau identity

Theorem (Moreau identity)

Let $\tau > 0$ and $J : \mathbb{E} \to \overline{\mathbb{R}}$ be proper, convex, and lsc. Then the following identity holds:

$$\mathrm{id}(\cdot) = \mathrm{prox}_{\tau J}(\cdot) + \tau \, \mathrm{prox}_{\frac{1}{\tau} J^*}(\cdot/\tau).$$

In particular, $\tau = 1 \Rightarrow \mathrm{id}(\cdot) = \mathrm{prox}_{J}(\cdot) + \mathrm{prox}_{J^*}(\cdot)$.

Proof:

$$\begin{aligned}
& v = \tau \, \mathrm{prox}_{\frac{1}{\tau} J^*}(u/\tau) \\
\Leftrightarrow \quad & \left(I + \frac{1}{\tau} \partial J^* \right)^{-1} (u/\tau) \ni v/\tau \\
\Leftrightarrow \quad & \partial J^*(v/\tau) \ni u - v \\
\Leftrightarrow \quad & v/\tau \in \partial J(u - v) \\
\Leftrightarrow \quad & u - v = (I + \tau \partial J)^{-1}(u) = \mathrm{prox}_{\tau J}(u).
\end{aligned}$$

Remark

The Moreau identity suggests that if one of $\mathrm{prox}_{J}(\cdot)$ and $\mathrm{prox}_{J^*}(\cdot)$ is computable, so is the other.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Infimal convolution

Definition

Let $F, G : \mathbb{E} \to \overline{\mathbb{R}}$ be proper, convex, and lsc. The **infimal convolution** (or inf convolution) of F and G is defined by

$$(F \square G)(\cdot) = \inf_{u \in \mathbb{E}} \{F(\cdot - u) + G(u)\},$$

with $\operatorname{dom}(F \square G) = \operatorname{dom} F + \operatorname{dom} G$.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Infimal convolution

Definition

Let $F, G : \mathbb{E} \to \overline{\mathbb{R}}$ be proper, convex, and lsc. The **infimal convolution** (or inf convolution) of F and G is defined by

$$(F \square G)(\cdot) = \inf_{u \in \mathbb{E}} \{F(\cdot - u) + G(u)\},$$

with $\operatorname{dom}(F \square G) = \operatorname{dom} F + \operatorname{dom} G$.

Theorem

Let $F, G : \mathbb{E} \to \overline{\mathbb{R}}$ be proper, convex, and lsc. Then

$$(F \square G)^*(\cdot) = F^*(\cdot) + G^*(\cdot).$$

Proof: on board.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Infimal convolution

Definition

Let $F, G : \mathbb{E} \to \overline{\mathbb{R}}$ be proper, convex, and lsc. The **infimal convolution** (or inf convolution) of F and G is defined by

$$(F \square G)(\cdot) = \inf_{u \in \mathbb{E}} \{F(\cdot - u) + G(u)\},$$

with $\operatorname{dom}(F \square G) = \operatorname{dom} F + \operatorname{dom} G$.

Theorem

Let $F, G : \mathbb{E} \to \overline{\mathbb{R}}$ be proper, convex, and lsc. Then

$$(F \square G)^*(\cdot) = F^*(\cdot) + G^*(\cdot).$$

Proof: on board.

Remark

Analogy between inf convolution (under convex conjugation) and convolution (under Fourier transform): $\widehat{F * G}(\cdot) = \widehat{F}(\cdot)\widehat{G}(\cdot)$.

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Moreau envelope

Definition

The **Moreau envelope** of a proper, convex, lsc function $J : \mathbb{E} \to \overline{\mathbb{R}}$ is defined for each $u \in \mathbb{E}$ by

$$\begin{aligned}\mathcal{M}_{\alpha J}(u) &:= \left(J \square \left(\frac{1}{2\alpha}\|\cdot\|^2\right)\right)(u) \\ &= \inf_{v\in\mathbb{E}} \left\{J(v) + \frac{1}{2\alpha}\|v-u\|^2\right\} \\ &= J(\text{prox}_{\alpha J}(u)) + \frac{1}{2\alpha}\|\text{prox}_{\alpha J}(u) - u\|^2.\end{aligned}$$

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Moreau envelope

Definition

The **Moreau envelope** of a proper, convex, lsc function $J : \mathbb{E} \to \overline{\mathbb{R}}$ is defined for each $u \in \mathbb{E}$ by

$$\begin{aligned}\mathcal{M}_{\alpha J}(u) &:= \left(J \square \left(\frac{1}{2\alpha}\|\cdot\|^2\right)\right)(u) \\ &= \inf_{v\in\mathbb{E}} \left\{J(v) + \frac{1}{2\alpha}\|v-u\|^2\right\} \\ &= J(\text{prox}_{\alpha J}(u)) + \frac{1}{2\alpha}\|\text{prox}_{\alpha J}(u) - u\|^2.\end{aligned}$$

Example

$J : u \in \mathbb{R} \mapsto |u| \Rightarrow \mathcal{M}_{\alpha J}$ is Huber function:

$$\mathcal{M}_{\alpha J}(u) = \begin{cases} u^2/(2\alpha) & \text{if } |u| \le \alpha, \\ |u| - \alpha/2 & \text{if } |u| > \alpha. \end{cases}$$

Observation: $\mathcal{M}_{\alpha J}$ does smoothing/regularization on J.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Properties

- Recall the theorem: $(F \square G)^*(\cdot) = F^*(\cdot) + G^*(\cdot) \Rightarrow$

$$(\mathcal{M}_{\alpha J})^* = J^* + (\frac{1}{2\alpha}\|\cdot\|^2)^* = J^* + \frac{\alpha}{2}\|\cdot\|^2.$$

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Properties

- Recall the theorem: $(F \square G)^*(\cdot) = F^*(\cdot) + G^*(\cdot) \Rightarrow$

$$(\mathcal{M}_{\alpha J})^* = J^* + (\frac{1}{2\alpha}\|\cdot\|^2)^* = J^* + \frac{\alpha}{2}\|\cdot\|^2.$$

- Recall the theorem: J is μ-strongly convex iff J^* has $\frac{1}{\mu}$-Lipschitz gradient. $\Rightarrow$

  $\mathcal{M}_{\alpha J}$ has $\frac{1}{\alpha}$-Lipschitz gradient.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator

Properties

- Recall the theorem: $(F \square G)^*(\cdot) = F^*(\cdot) + G^*(\cdot) \Rightarrow$

$$(\mathcal{M}_{\alpha J})^* = J^* + (\frac{1}{2\alpha}\|\cdot\|^2)^* = J^* + \frac{\alpha}{2}\|\cdot\|^2.$$

- Recall the theorem: J is μ-strongly convex iff J^* has $\frac{1}{\mu}$-Lipschitz gradient. $\Rightarrow$

  $\mathcal{M}_{\alpha J}$ has $\frac{1}{\alpha}$-Lipschitz gradient.

- $\nabla \mathcal{M}_{\alpha J}$ can be calculated as:

$$\begin{aligned}
p = \nabla \mathcal{M}_{\alpha J}(u) &\Leftrightarrow u \in \partial(\mathcal{M}_{\alpha J})^*(p) = \partial J^*(p) + \alpha p \\
&\Leftrightarrow u - \alpha p \in \partial J^*(p) \Leftrightarrow \partial J(u - \alpha p) \ni p \\
&\Leftrightarrow \alpha \partial J(u - \alpha p) \ni \alpha p \Leftrightarrow (I + \alpha \partial J)(u - \alpha p) \ni u \\
&\Leftrightarrow u - \alpha p = (I + \alpha \partial J)^{-1} u = \text{prox}_{\alpha J}(u) \\
&\Leftrightarrow \nabla \mathcal{M}_{\alpha J}(u) = \frac{1}{\alpha}(u - \text{prox}_{\alpha J}(u)).
\end{aligned}$$

Convex Set
Convex Function
Existence of Minimizer
Subdifferential
Convex Conjugate
Duality Theory
Proximal Operator