



# Chapter 1

## Convex Analysis

*Convex Optimization for Machine Learning & Computer Vision*  
WS 2019/20

Convex Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator

Tao Wu  
Zhenzhang Ye

Computer Vision Group  
Department of Informatics  
TU Munich



# Convex Set

Convex Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator



## 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

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator



## 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

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator

## Convex optimization

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

First questions:

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

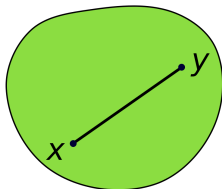
# Convex set

## Definition

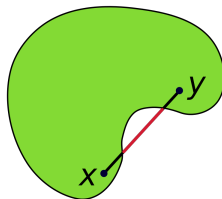
A set  $C$  is said to be **convex** if

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

convex



non-convex



### 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

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



### Convex Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator



### 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

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

with  $\text{aff } C$  the **affine hull** of  $C$ . If  $C$  is a *convex* set, then

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

### 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:  $\text{cl } C$ .
- Interior and relative interior:  $\text{int } C$ ,  $\text{rint } C$ .

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





## Basic properties

The following operations preserve the convexity:

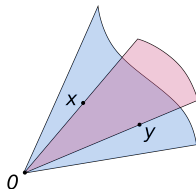
- Intersection:  $C_1 \cap C_2$ .
- Summation:  $C_1 + C_2 := \{u^1 + u^2 : u^1 \in C_1, u^2 \in C_2\}$ .
- Closure:  $\text{cl } C$ .
- Interior and relative interior:  $\text{int } C, \text{rint } C$ .

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.



Convex cone.





## Theorem (separation of convex sets)

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

- 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 Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator



# Convex Function

# Convex functions

- 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

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



## Examples

- $J_{data}(u) = \|u - z\|_p^p$ , where  $p \geq 1$  and  $\|\cdot\|_p$  is  $\ell^p$ -norm.
- $J_{regu}(u) = \|Ku\|_q^q$ , where  $K$  is linear transform and  $q \geq 1$ .
- $J(u) = J_{data}(u) + \alpha J_{regu}(u)$ , where  $\alpha > 0$ .
- Negative binary entropy ( $\epsilon > 0$ ):  
 $J_\epsilon(u) = \epsilon(u \log(u) + (1 - u) \log(1 - u))$ .
- Soft plus:  $J_\epsilon(v) = \epsilon \log(1 + \exp(v/\epsilon))$ .



## Examples

- $J_{data}(u) = \|u - z\|_p^p$ , where  $p \geq 1$  and  $\|\cdot\|_p$  is  $\ell^p$ -norm.
- $J_{regu}(u) = \|Ku\|_q^q$ , where  $K$  is linear transform and  $q \geq 1$ .
- $J(u) = J_{data}(u) + \alpha J_{regu}(u)$ , where  $\alpha > 0$ .
- Negative binary entropy ( $\epsilon > 0$ ):  
 $J_\epsilon(u) = \epsilon(u \log(u) + (1 - u) \log(1 - u))$ .
- Soft plus:  $J_\epsilon(v) = \epsilon \log(1 + \exp(v/\epsilon))$ .
- **Indicator function** of a convex set  $C \subset \mathbb{E}$ :

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

Formulate *constrained optimization* with indicator function:

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





(As exercises)

- Any norm (over a normed vector space) is a convex function.
- $J$  is a convex function and  $A$  is an affine transform  
 $\Rightarrow u \mapsto J(A(u))$  is a 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  $\{u^i\}_{i=1}^n \subset \mathbb{E}$ ,  $\{\alpha_i\}_{i=1}^n \subset [0, 1]$ ,  $\sum_{i=1}^n \alpha_i = 1$ .

(Hence it is an equivalent definition of a convex function.)

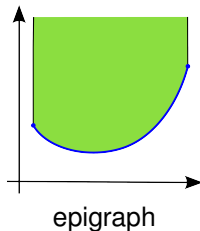


# Epigraph

## 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

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator



## 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

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator



Convex Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator

# 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?



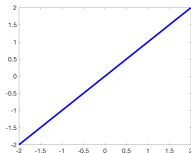
- Consider

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

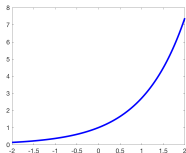
where  $J : \mathbb{E} \rightarrow \overline{\mathbb{R}}$  is a proper, convex function.

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

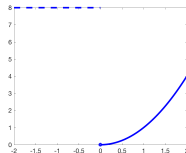
$u$



$\exp u$



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



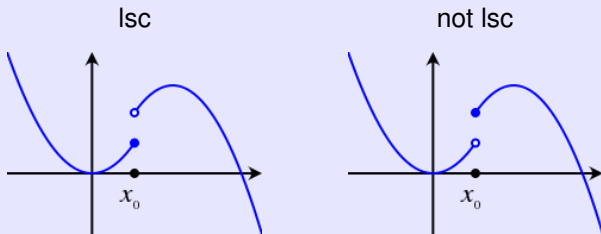
- Next we formalize our observations and derive sufficient conditions for existence.

# 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.





## 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.

Convex Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator



Convex Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator

- 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.



# Subdifferential

Convex Set

Convex Function

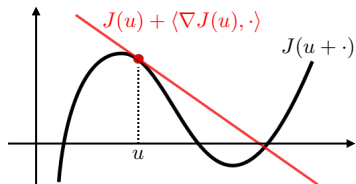
Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator



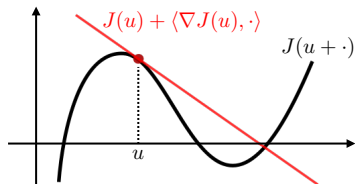
## Definition

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

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

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





## Definition

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

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

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

## Remark

If  $J$  is a convex function, the definition of differentiability can be extended to  $J : \text{aff dom } J \rightarrow \overline{\mathbb{R}}$  for  $u \in \text{rint dom } J$ .

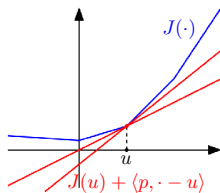


# Subdifferential

## Definition

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

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

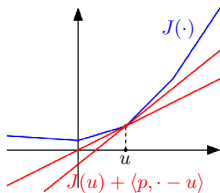


# Subdifferential

## Definition

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

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



## Geometric interpretation

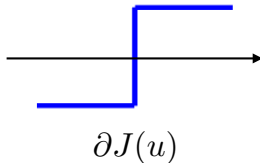
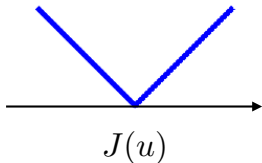
$p \in \partial J(u)$  iff  $(p, -1)$  is a normal vector for the supporting hyperplane of  $\text{epi } J$  at  $(u, J(u))$ , i.e.,

$$\left\langle \begin{bmatrix} p \\ -1 \end{bmatrix}, \begin{bmatrix} u \\ J(u) \end{bmatrix} \right\rangle \geq \left\langle \begin{bmatrix} p \\ -1 \end{bmatrix}, \begin{bmatrix} v \\ \alpha \end{bmatrix} \right\rangle, \quad \forall (v, \alpha) \in \text{epi } J.$$



# Subdifferential: Examples

①  $J(u) = |u|$ .

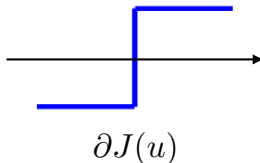
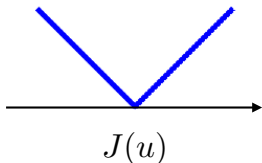




# Subdifferential: Examples



①  $J(u) = |u|$ .



② Given a closed, convex subset  $C \subset \mathbb{E}$  and  $u \in C$ ,

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

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

Convex Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

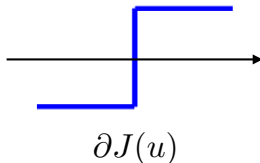
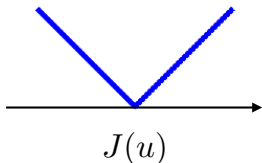
Duality Theory

Proximal Operator

## Subdifferential: Examples



①  $J(u) = |u|.$



② Given a closed, convex subset  $C \subset \mathbb{E}$  and  $u \in C$ ,

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

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

③  $J(u) = \|\|u\|\| \Rightarrow \partial J(0) = \{p : \|\|p\|\|_* \leq 1\}$ .  $\|\|\cdot\|\|_*$  is the dual norm of  $\|\|\cdot\|\|$ , i.e.,  $\|\|p\|\|_* = \sup\{\langle p, u \rangle : \|\|u\|\| \leq 1\}$ .

## Theorem (chain rule under linear transform)

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

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

whenever  $Ku \in \text{rint dom } J$ .

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



Convex Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator

# Subdifferential calculus



## Theorem (chain rule under linear transform)

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

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

whenever  $Ku \in \text{rint dom } J$ .

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

## Theorem (summation rule)

Let  $\tilde{J}(\cdot) = J_1(\cdot) + J_2(\cdot)$ , where  $J_1, J_2$  are convex functions s.t.

$$\text{rint dom } J_1 \cap \text{rint dom } J_2 \neq \emptyset.$$

Then for any  $u \in \text{dom } J_1 \cap \text{dom } J_2$ , we have

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

# Properties of subdifferential map

## Theorem

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

Proof: on board.

Convex Analysis

Tao Wu  
Zhenzhang Ye



Convex Set

Convex Function

Existence of Minimizer

Subdifferential

Convex Conjugate

Duality Theory

Proximal Operator

## Properties of subdifferential map

### Theorem

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

Proof: on board.

### Theorem

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

$$\langle p^1 - p^2, u^1 - u^2 \rangle \geq 0.$$

Proof: on board.



## Properties of subdifferential map

### Theorem

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

Proof: on board.

### Theorem

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

$$\langle p^1 - p^2, u^1 - u^2 \rangle \geq 0.$$

Proof: on board.

### Theorem

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

$$\exists (u^k, p^k) \rightarrow (u^*, p^*) \in (\text{dom } J) \times \mathbb{E} \text{ s.t. } p^k \in \partial J(u^k) \quad \forall k.$$

Proof: on board.



# Optimality condition

## Theorem

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

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

Proof: on board.





## Optimality condition

### Theorem

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

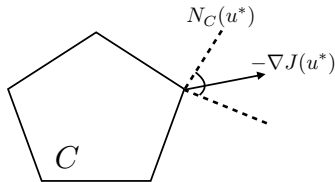
$$0 \in \partial J(u^*).$$

Proof: on board.

### Constrained optimization as special case

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

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

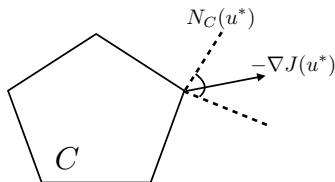


## Optimality condition

### Constrained optimization as special case

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

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



### Remark

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



# Optimality condition



## Constrained optimization as special case

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

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

## Remark

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

## Example: Normal cone of linear-inequality constraints

Let  $C = \{u \in \mathbb{R}^n : Au \leq b\}$  with given  $b \in \mathbb{R}^m, A \in \mathbb{R}^{m \times n}$ . Then

$$N_C(u) = \{A^\top \lambda : \lambda \geq 0, \lambda_i = 0 \text{ whenever } (Au - b)_i < 0\}.$$