Chapter 1
Convex Analysis

Convex Optimization for Computer Vision
SS 2016

Michael Moeller
Thomas Möllenhoff
Emanuel Laude
Computer Vision Group
Department of Computer Science
TU München

Convexity

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Convex energy minimization problems

This lecture is all about

$$\hat{u} \in \arg\min_{u \in C} E(u),$$

where $C \subset \mathbb{R}^n$ convex set, $E : C \to \mathbb{R}$ convex function.

1. What is a convex set?

Definition

A set $C \subset \mathbb{R}^n$ is called convex, if

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

$\rightarrow$ *Draw a picture.*

$\rightarrow$ *Online TED.*

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Convex set - summary of online TED

Definitions and things to recall from analysis 1:

Definitions

- A set $C \subset \mathbb{R}^n$ is called **open** if for all $x \in C$ there exists a $\epsilon > 0$ such that the ball with radius ϵ around x, $B(x, \epsilon)$, is contained in C: $B(x, \epsilon) \subset C$.
- A set $C \subset \mathbb{R}^n$ is called **closed** if its complement is open.
- A set is closed if and only if it contains all its limit points.
- The **closure** of a set $C \subset \mathbb{R}^n$ is

$$\overline{C} = \{x \mid \text{ there exists a convergent sequence } (x_n)_n \subset C \text{ such that } \lim_{n\to\infty} x_n = x\}$$

- The **interior** of a set $C \subset \mathbb{R}^n$ is

$$\mathring{C} = \{x \in C \mid \text{ there exists } \epsilon > 0 \text{ such that } B(x, \epsilon) \subset C\}$$

Convex set - summary of online TED

The following operations preserve the convexity of a set

- Intersection
- Vector sum
- Closure
- Interior
- Linear Transformation

The union of convex sets is not convex in general.

Polyhedral sets are always convex, cones are not necessarily convex.

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Convex energy minimization problems

Let's get back to what the lecture is all about:

$$\hat{u} \in \arg\min_{u \in C} E(u),$$

where $C \subset \mathbb{R}^n$ convex set, $E : C \to \mathbb{R}$ convex function.

1. What is a convex set? We know this now!

2. What is a convex function?

Definition: Convex Function

We call $E : C \to \mathbb{R}$ a convex function if C is a convex set and for all $u, v \in C$ and all $\theta \in [0, 1]$ it holds that

$$E(\theta u + (1 - \theta)v) \leq \theta E(u) + (1 - \theta)E(v)$$

We call E strictly convex, if the inequality is strict for all $\theta \in]0, 1[$, and $v \neq u$.

$\rightarrow$ *Draw a picture.*

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Convex functions - summary of online TED

The following operations do preserve the convexity of a function

- Summation
- Linear Transformation

The following operations do not preserve the convexity of a function

- Multiplication, Division, Difference
- Composition

The sum of a convex function and a strictly convex function is strictly convex.

Remember to check two conditions to show that a function is convex!

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Convex energy minimization problems

Let's get back to what the lecture is all about:

$$\hat{u} \in \arg\min_{u \in C} E(u), \tag{1}$$

where $C \subset \mathbb{R}^n$ convex set, $E : C \to \mathbb{R}$ convex function.

It is sometimes convenient to "introduce" the constraint $u \in C$ into the energy function E itself. We therefore introduce the notion of **extended real valued functions**.

$$E : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}.$$

The minimization problem (1) can then be written as

$$\hat{u} \in \arg\min_{u \in \mathbb{R}^n} \tilde{E}(u),$$

by defining

$$\tilde{E}(u) = \begin{cases} E(u) & \text{if } u \in C, \\ \infty & \text{else.} \end{cases}$$

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Extended real valued functions

Definition

- For $E : \mathbb{R}^n \to \overline{\mathbb{R}}$, we call

$$\operatorname{dom}(E) := \{u \in \mathbb{R}^n \mid E(u) < \infty\}$$

  the domain of E.
- We call E proper if $\operatorname{dom}(E) \neq \emptyset$.

Revisiting the definition of convex functions

We call $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ a convex function if

1. $\operatorname{dom}(E)$ is a convex set.

2. For all $u, v \in \operatorname{dom}(E)$ and all $\theta \in [0, 1]$ it holds that

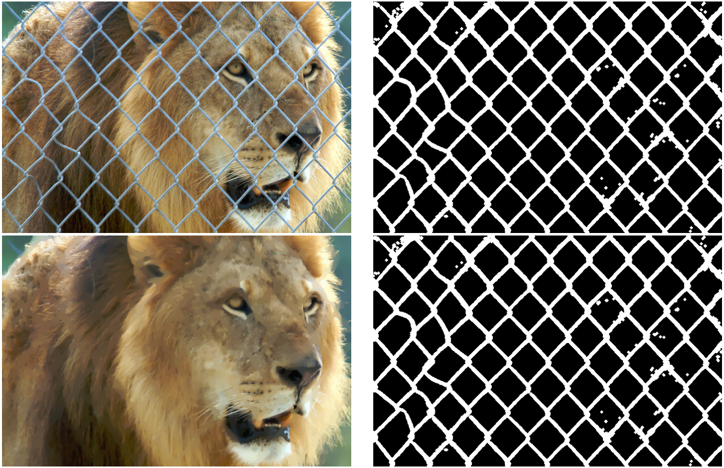
$$E(\theta u + (1-\theta)v) \leq \theta E(u) + (1-\theta)E(v)$$

We call E strictly convex, if the inequality in 2 is strict for all $\theta \in]0, 1[$, and $v \neq u$.

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

First example of an imaging problem: Inpainting

Example: Inpainting



$$\min_{u \in \mathbb{R}^{n \times m}} \sum_{i,j} (u_{i,j} - u_{i-1,j})^2 + (u_{i,j} - u_{i,j-1})^2 \quad \text{s.t. } u_{i,j} = f_{i,j} \; \forall (i,j) \in I$$

with index set I of pixels to keep and suitable boundary

Epigraph of a function

Is there a connection between convex sets and functions?

Defintion: Epigraph

Let $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper function mapping into the extended real line. Then

$$\text{epi}(E) := \{(u, \alpha) \mid E(u) \leq \alpha\}$$

is called the *epigraph* of the function E.

$\rightarrow$ *Draw a picture.*

Theorem

A proper function $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if and only if it's epigraph is convex

Proof: First exercise sheet.

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Properties of convex functions

What is so special about convex functions?

Theorem

Let $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex. Any local minimum of E is global.

Proof: Board.

Theorem: Monotonicity of the gradient

Let $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, convex and differentiable at $u \in \mathrm{dom}(E)$.

$$E(v) - E(u) - \langle \nabla E(u), v - u \rangle \geq 0 \qquad \forall v \in \mathbb{R}^n$$

Proof: Later.

Conclusion

Let $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, convex and differentiable at $u \in \mathrm{dom}(E)$. If $\nabla E(u) = 0$ then u is a global minimum of E.

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Convex energy minimization problems

We said this lecture is all about

$$\hat{u} \in \arg\min_{u \in \mathbb{R}^n} E(u),$$

where $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a convex function.

Does a minimizer of such a function even exist?

Existence of minimizers

Defintion: Lower semi-continuity (l.s.c.)

We call the function $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ lower semi-continuous (l.s.c.), if for all u it holds that

$$\liminf_{v \to u} E(v) \geq E(u).$$

Theorem: Existence of minimizers

Let $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ be l.s.c. and let there exist an α such that the sublevelset

$$\{u \in \mathbb{R}^n \mid E(u) \leq \alpha\}$$

is nonempty and bounded, then

$$\hat{u} \in \arg\min_u E(u)$$

exists.

Proof: Board.

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Closedness and lower semi-continuity

Defintion: Closed function

We call the function $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ closed if it's epigraph is closed.

Theorem: Equivalence of l.s.c. and closedness

For $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ the following two statements are equivalent

- E is lower semi-continuous (l.s.c.).
- E is closed.

Proof: Board.

OnlineTED!

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Continuity of convex functions

Continuity of Convex Functions

If $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex, then E is locally Lipschitz (and hence continuous) on int(dom(E)).

Proof in 1d: Exercise for yourself (solution will be online)

$\to$ *Board: Considering the interior is important!*

Conclusion

If $E : \mathbb{R}^n \to \mathbb{R}$ is convex, then E is continuous.

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Existence of minimizers part II

Definition: Coercivity

A function $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is called *coercive* if $E(v_n) \to \infty$ for all sequences $(v_n)_n$ with $\|v_n\| \to \infty$.

Remark: Coercivity implies that there exists a bounded sublevelset.

Existence of a minimizer for function with full domain

Let $E : \mathbb{R}^n \to \mathbb{R}$ be convex and coercive, then an element $\hat{u} \in \arg\min_u E(u)$ exists.

Proof:

- $\text{dom}(E) = \mathbb{R}^n$, E convex $\Rightarrow$ E is continuous.
- E is coercive, i.e. there exists a non-empty bounded sublevelset.

Uniqueness

When is

$$\hat{u} \in \arg\min_{u \in \mathbb{R}^n} E(u)$$

unique?

Theorem: Uniqueness

If $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ is strictly convex, then there exists at most one local minimum which is the unique global minimum.

Proof: Simple computation.

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Optimality conditions

How can we determine if

$$\hat{u} \in \arg\min_{u \in \mathbb{R}^n} E(u)? \tag{1}$$

In other words,

what is the optimality condition for (1)?

Consider a differentiable E and remember analysis I:

Necessary condition for local extremum is

$$\nabla E(\hat{u}) = 0$$

Sufficient condition? **Convexity!**

Optimality conditions

Example why convex functions are great

(stealing from Thomas' examples)

$$\min_u \frac{1}{2}\|u - f\|_2^2 + \alpha R(\nabla u)$$

with

$$R(d) = \begin{cases} \frac{1}{2}\|d\|_2^2 & \text{if } \|d\|_2 \leq \epsilon \\ \epsilon\|d\| - \frac{1}{2}\epsilon^2 & \text{else.} \end{cases}$$

Examples: Derivatives of convex functions

Getting back: What are the optimality conditions for ...

... $E(u) = \|u - f\|_2^2 = \sum_{i=1}^n (u_i - f_i)^2$?

... $E(u) = \|Au - f\|_2^2$ for a matrix $A \in \mathbb{R}^{m \times n}$?

... $E(u) = \|u\|_1 = \sum_{i=1}^n |u_i|$?

We need a theory for non-differentiable functions!

Illustrate ℓ^1 case for discussion.

The subdifferential

Definition: Subdifferential

Let $E : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex. We call

$$\partial E(u) = \{p \in \mathbb{R}^n \mid E(v) - E(u) - \langle p, v - u \rangle \geq 0,\ \forall v \in \mathbb{R}^n\}$$

the subdifferential of E at u.

- Elements of $\partial E(u)$ are called subgradients.
- If $\partial E(u) \neq \emptyset$, we call E subdifferentiable at u.
- By convention, $\partial E(u) = \emptyset$ for $u \neq \mathrm{dom}(E)$.

Theorem: Optimality condition

Let $0 \in \partial E(\hat{u})$. Then $\hat{u} \in \arg\min_u E(u)$.

The subdifferential

Examples for non-differentiable functions:

- The absolute value function.
- Functional

$$E(u) = \begin{cases} 0 & \text{if } u \geq 0 \\ \infty & \text{else.} \end{cases}$$

- $E(u) = \frac{1}{2}\|u\|^2$

Subdifferential and derivatives

Let the convex function $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be differentiable at $u \in \text{int}(\text{dom}(E))$. Then

$$\partial E(u) = \{\nabla E(u)\}.$$

Proof: Exercise.

This also proves the "Theorem: Monotonicity of the gradient".

The subdifferential

Geometric interpretation of subgradients:

Any subgradient $p \in \partial E(u)$ represents a non-vertical supporting hyperplane to $\text{epi}(E)$ at $(u, E(u))$.

Definition

A supporting hyperplane to a set $S \subset \mathbb{R}^n$ is a hyperplane $\{x \in \mathbb{R}^n \mid \langle a, x \rangle = b\}$, such that

- $S \subset \{x \in \mathbb{R}^n \mid \langle a, x \rangle \leq b\}$ or $S \subset \{x \in \mathbb{R}^n \mid \langle a, x \rangle \geq b\}$
- $\exists y \in \partial S$ (the boundary of S) such that $\langle a, y \rangle = b$.

Let $p \in \partial E(u)$. Then

$$\begin{aligned} & E(v) - E(u) - \langle p, v - u \rangle \geq 0 && \forall v \in \mathbb{R}^n \\ \Rightarrow \quad & \alpha - E(u) - \langle p, v - u \rangle \geq 0 && \forall (v, \alpha) \in \text{epi}(E) \\ \Rightarrow \quad & \left\langle \begin{bmatrix} -p \\ 1 \end{bmatrix}, \begin{bmatrix} v \\ \alpha \end{bmatrix} - \begin{bmatrix} u \\ E(u) \end{bmatrix} \right\rangle \geq 0 && \forall (v, \alpha) \in \text{epi}(E). \end{aligned}$$

$\rightarrow$ *Draw image on the board.*

The subdifferential

Is any convex E subdifferentiable at $x \in \mathrm{dom}(E)$?

$$E(u) = \begin{cases} -\sqrt{u} & \text{if } u \geq 0 \\ \infty & \text{else.} \end{cases}$$

Definition: Relative Interior

The *relative interior* of a convex set M is defined as

$$\mathrm{ri}(M) := \{x \in M \mid \forall y \in M,\ \exists \lambda > 1,\ \text{s.t. } \lambda x + (1-\lambda) y \in M\}$$

Theorem: Subdifferentiability[1]

If E is a proper convex function and $u \in \mathrm{ri}(\mathrm{dom}(E))$, then $\partial E(u)$ is non-empty and bounded.

Partial proof on the board, full proof Rockafellar.

[1] Rockafellar, Convex Analysis, Theorem 23.4

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Subdifferential rules

Theorem: Sum rule[2]

Let E_1, E_2 be convex functions such that

$$\mathrm{ri}(\mathrm{dom}(E_1)) \cap \mathrm{ri}(\mathrm{dom}(E_2)) \neq \emptyset,$$

then it holds that

$$\partial(E_1 + E_2)(u) = \partial E_1(u) + \partial E_2(u).$$

Example: Minimize $(u - f)^2 + \iota_{u \geq 0}(u)$.

Example: Minimize $0.5(u - f)^2 + \alpha|u|$.

[2] Rockafellar, Convex Analysis, Theorem 23.8

Subdifferential rules

Theorem: Chain rule[3]

If $A \in \mathbb{R}^{m \times n}$, $E : \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ is convex, and $\text{ri}(\text{dom}(E)) \cap \text{range}(A) \neq \emptyset$, then

$$\partial(E \circ A)(u) = A^* \partial E(Au)$$

Example: Minimize $\|Au - f\|_2^2$.

[3] Rockafellar, Convex Analysis, Theorem 23.9

Subdifferential rules

We have seen the example of ℓ^1 minimization/denoising

$$\min_u \frac{1}{2}\|u - f\|_2^2 + \alpha\|u\|_1$$

More interesting for imaging: Change of basis, e.g. orthogonal wavelet basis

$$\min_u \frac{1}{2}\|u - f\|_2^2 + \alpha\|Wu\|_1$$

Show example in Matlab!

Michael Moeller
Thomas Möllenhoff
Emanuel Laude

Summary

- **Convex functions**
  - Every local minimum is global
  - First order optimality condition is sufficient
- The **optimality condition** for $\hat{u}$ to minimize E is

$$0 \in \partial E(\hat{u})$$

- **The subdifferential** $\partial E(u)$
  - is set valued.
  - generalizes the derivative.
  - $\partial E(u) = \{\nabla E(u)\}$ is E is differentiable at u.
  - can be identified with supporting hyperplanes to epi(E).
  - Obeys the “usual” sum and chain rules.

We now have all tools that are necessary to discuss a first class of minimization algorithms for determining

$$\hat{u} \in \underset{u}{\operatorname{argmin}}\, E(u)$$

Next lecture: Implementation, convergence, applications!

Michael Moeller
Thomas Möllenhoff
Emanuel Laude