Convex analysis

Olivier Fercoq

Télécom Paris

Existence of minimizers

Theorem

If $f : C \to \mathbb{R}$ is continuous *and the set C is* compact

then $\exists x^*$ such that $f(x^*) \leq f(x)$ for all $x \in C$.

Existence of minimizers

Theorem

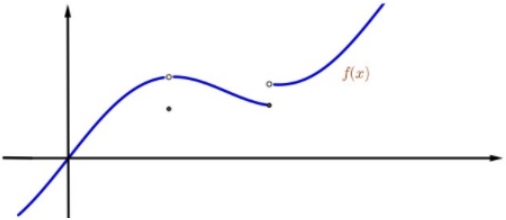
If $f : C \to \mathbb{R}$ is continuous and the set C is compact

then $\exists x^*$ such that $f(x^*) \leq f(x)$ for all $x \in C$.

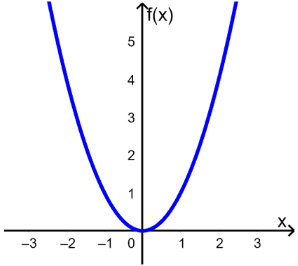
Theorem

If $f : \mathbb{R}^d \to \mathbb{R}$ is lower semi-continuous and coercive
then $\exists x^*$ such that $f(x^*) \leq f(x)$ for all $x \in C$.

Lower semi-continuity



Coercivity



Convex functions

Definition

A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex if $\forall x, y \in \mathbb{R}^d, \forall t \in [0, 1]$,

Convex functions

Definition

A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex if $\forall x, y \in \mathbb{R}^d, \forall t \in [0, 1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

Why do we allow $+\infty$ values?

Let C be a convex set and let the convex indicator of C be $\iota_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{if } x \notin C \end{cases}$

ι_C is convex [proof]

$$\begin{aligned} \min_{x \in \mathbb{R}^d} \ & f(x) \quad = \min_{x \in \mathbb{R}^d} f(x) + \iota_C(x) \\ \text{s.t. } & x \in C \end{aligned}$$

$\rightarrow$ Nonsmooth optimization generalizes constrained optimization

Gradients of convex functions

Proposition

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function, differentiable at x. Then for all $y \in \mathbb{R}^d$,

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$$

- proof last week in lecture

Proposition

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice differentiable convex function. Then for all $x \in \mathbb{R}^d$, $\nabla^2 f(x)$ is a positive semi-definite matrix.

- proof on next slide

Proposition

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function, whose gradient is L-Lipschitz. Then for all $x, y \in \mathbb{R}^d$,

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|x - y\|^2$$

- proof in tutorial session

Hessian matrix of a convex function

Proposition

Let $f : \mathbb{R}^d \to \mathbb{R}$ *be a twice differentiable convex function. Then for all* $x \in \mathbb{R}^d$, $\nabla^2 f(x)$ *is a positive semi-definite matrix.*

Subgradient

Definition

Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$.
A vector $\phi \in \mathbb{R}^d$ is a *subgradient* of f at x if

$$\forall y \in \mathbb{R}^d, \quad f(y) \geq f(x) + \langle \phi, y - x \rangle$$

The set of all subgradients is called the *subdifferential*:

$$\partial f(x) = \{\phi \ : \ \forall y \in \mathbb{R}^d, \ f(y) \geq f(x) + \langle \phi, y - x \rangle\}$$

Examples: $f(x) = |x|$ $\qquad f(x) = \iota_{\mathbb{R}_+}(x)$

Operations on subdifferentials

Theorem

If $f : \mathbb{R}^m \to \mathbb{R}^n$ and $g : \mathbb{R}^d \to \mathbb{R}^m$ are two differentiable function, then

$$J_{f \circ g}(x) = J_f(g(x)) \times J_g(x)$$

Corollary

If f is differentiable and M is a linear operator, then

$$\nabla (f \circ M)(x) = M^\top \nabla f(Mx)$$

Theorem

If f and g are convex and g is differentiable, then

$$\partial (f + g)(x) = \partial f(x) + \{\nabla g(x)\}$$

Proof

Theorem

If f and g are convex and g is differentiable, then $\partial(f+g)(x) = \partial f(x) + \{\nabla g(x)\}$.

Let $\varphi \in \partial f(x)$.

$$f(x) + g(x) + \langle \varphi + \nabla g(x), y - x \rangle = \underbrace{f(x) + \langle \varphi, y-x \rangle}_{\substack{\le f(y) \\ \text{because } \varphi \in \partial f(x)}} + \underbrace{g(x) + \langle \nabla g(x), y - x \rangle}_{\substack{\le g(y) \\ \text{because } \nabla g(x) \in \partial g(x)}}$$

Thus $\varphi + \nabla g(x) \in \partial(f+g)(x)$

Fermat's rule

Theorem

$x \in \arg\min f \Leftrightarrow 0 \in \partial f(x)$

$x \in \arg\min f \Longleftrightarrow \forall y,\ f(y) \geq f(x) \Longleftrightarrow f(y) \geq f(x) + \langle 0, y-x \rangle \Longleftrightarrow 0 \in \partial f(x)$

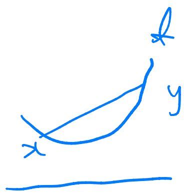
Corollary

Suppose that f is convex and differentiable.

$x \in \arg\min f \Leftrightarrow \nabla f(x) = 0$

Note of Lecture 2.

Convex analysis.



Def: △ function $f: \mathbb{R}^n \to \mathbb{R}$ convex if $\forall x, y \in \mathbb{R}^n$,

$$\forall t \in [0,1], \quad f(tx+(1-t)y) \le t f(x) + (1-t) f(y)$$

△ Set C is convex if $\forall x, y \in \mathbb{R}^n$, $\forall t \in [0,1]$,

$$tx + (1-t)y \in C.$$

Prop. if f is convex and differentiable, then $\forall x, y$,

$$f(y) \ge f(x) + \langle \nabla f(x), y-x \rangle$$

Pf
$$\langle \nabla f(x), y-x \rangle = \lim_{t\to 0} \frac{f(x+t(y-x)) - f(x)}{t}$$
$$= \lim_{t\to 0} \frac{f(ty+(1-t)x) - f(x)}{t}$$
$$\le \lim_{t\to 0} \frac{t f(y) + (1-t) f(x) - f(x)}{t}$$
$$\le f(y) - f(x)$$

Prop: if f is convex and differentiable, $\nabla f(x^*) = 0$,

then $x^* \in \arg\min_x f(x)$

Pf: $f(y) \ge f(x^*) + \langle \nabla f(x^*), y - x^* \rangle = f(x^*) \quad \forall y \in \mathbb{R}^n$.

thus $x^* = \arg\min_x f(x)$

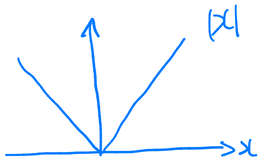
Prop: if f is differentiable and $x^* \in \arg\min_x f(x)$.

then $\nabla f(x^*) = 0$.

Subgradient

g is a subgradient of f at x if

$\forall y,\ f(y) \geq f(x) + \langle g, y-x \rangle$.



the subdifferential of f at x is the set of all subgradients,

we denote it with $\partial f(x)$ A set!

$$\partial |\cdot| (x) = \begin{cases} \{-1\} & \text{if } x<0 \\ \{+1\} & \text{if } x>0 \\ [-1, 1], & \text{if } x=0 \end{cases}$$

Prop: if f is convex and differentiable, $\partial f(x) = \{\nabla f(x)\}\ \forall x$.

other exemples of convex non smooth functions:

- norms
- $z \mapsto \max(0, z)$.
- indicator function of a set: $i_C(x) = \begin{cases} 0, & x \in C \\ +\infty, & x \notin C \end{cases}$

$$\min_{x \in C} f(x) = \min_{x \in \mathbb{R}^n} f(x) + i_C(x)$$

$\partial i_C(x)$ is the normal cone at x

Thm (Fermat's Rule) $x^* \in \arg\min_x f(x) \iff 0 \in \partial f(x^*)$

Pf: $x^* \in \arg\min_x f(x) \iff \forall x,\ f(x) \geq f(x^*)$

$\iff \forall x,\ f(x) \geq f(x^*) + \langle 0, x-x^* \rangle$

$\iff 0 \in \partial f(x^*)$

Rules that preserve convexity:

- if f and g are convex, $f+g$ is convex.
- if f is convex and $a \geq 0$, then $a \cdot f$ is convex.
- if $f: \mathbb{R}^n \to \mathbb{R}$ is convex and $A: \mathbb{R}^n \to \mathbb{R}^n$ is affine, then $f \circ A$ is convex.
- if $f: \mathbb{R}^n \to \mathbb{R}$ is convex and $g: \mathbb{R} \to \mathbb{R}$ is convex and non decreasing, then $g \circ f$ is convex.