Convex analysis

Olivier Fercoq

Télécom Paris

Existence of minimizers

Theorem

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

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

Existence of minimizers

Theorem

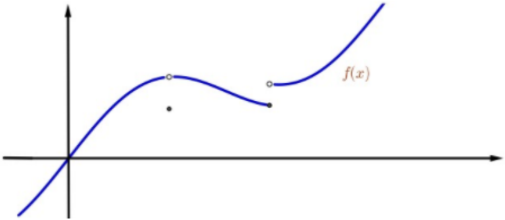
If $f : C \to \mathbb{R}$ is continuous and the set C is compact

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

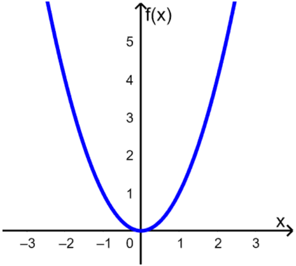
Theorem

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]

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

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

Fermat's rule

Theorem

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

Corollary

Suppose that f is convex and differentiable.

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