

Convex Analysis Background

R.D.C. Monteiro¹

¹School of Industrial and Systems Engineering
Georgia Institute of Technology

January 10, 2024

Def: $f : \mathbb{R}^n \rightarrow [-\infty, \infty]$ is differentiable at $\bar{x} \in \mathbb{R}^n$ if $f(\bar{x}) \in \mathbb{R}$ and there exists a linear map $f'(\bar{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$\lim_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x}) - f'(\bar{x})h}{\|h\|} = 0 \quad (1)$$

Facts: If f is differentiable at \bar{x} , then

- 1) $f'(\bar{x})$ satisfying (??) is unique
- 2) $\bar{x} \in \text{int}(\text{dom } f)$ where

$$\text{dom } f := \{x \in \mathbb{R}^n : f(x) < \infty\}$$

Def: Given function $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$, define the epigraph and strict epigraph, respectively, as

$$\begin{aligned}\text{epi } f &:= \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq t\} \\ \text{epi}_s f &:= \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : f(x) < t\}\end{aligned}$$

Def: $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ is said to be **convex** if its epigraph $\text{epi } f$ is a convex set. The set of all such functions is denoted by $\text{E-Conv}(\mathbb{R}^n)$

Proposition

The following conditions about $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ are equivalent:

- a) $\text{epi } f$ is a convex set (i.e., $f \in \text{E-Conv}(\mathbb{R}^n)$)
- b) $\text{epi}_s f$ is a convex set
- c) for every $x_0, x_1 \in \text{dom } f$ and $\alpha \in (0, 1)$, we have

$$f(\alpha x_0 + (1 - \alpha)x_1) \leq \alpha f(x_0) + (1 - \alpha)f(x_1)$$

Notation: Let

$$\Delta_k := \left\{ (\alpha_0, \dots, \alpha_k) \in \mathbb{R}^{k+1} : \sum_{i=0}^k \alpha_i = 1, \alpha_i \geq 0 \right\},$$

Proposition

Let $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be given. Then, $f \in \text{E-Conv}(\mathbb{R}^n)$ iff

$$f(\alpha_0 x_0 + \dots + \alpha_k x_k) \leq \alpha_0 f(x_0) + \dots + \alpha_k f(x_k),$$

for every $x_0, \dots, x_k \in \text{dom } f$ and $(\alpha_0, \dots, \alpha_k) \in \Delta_k$

In particular, for every $x_0, \dots, x_k \in \text{dom } f$ and $x \in \text{co} \{x_0, \dots, x_k\}$, there holds

$$f(x) \leq \max\{f(x_i) : i = 0, \dots, k\}$$

Def: $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ is **proper** if $\text{dom } f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in \mathbb{R}^n$. The set of all proper convex functions is denoted by $\text{Conv}(\mathbb{R}^n)$

Def: $f \in \text{Conv}(\mathbb{R}^n)$ is **closed** if its epigraph is closed. The set of all proper closed convex functions is denoted by $\overline{\text{Conv}}(\mathbb{R}^n)$

Definition

The indicator function $\delta_C : \mathbb{R}^n \rightarrow [0, \infty]$ of a set $C \subset \mathbb{R}^n$ is

$$\delta_C(x) := \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise} \end{cases}$$

Obs: δ_C is convex (resp., proper convex) $\Leftrightarrow C$ is convex (resp., nonempty convex)

Definition

The support function $\sigma_C : \mathbb{R}^n \rightarrow [-\infty, \infty]$ of C is

$$\sigma_C(x) := \sup\{\langle x, c \rangle : c \in C\}$$

If $C = \emptyset$ then $\sigma_C = -\infty$; otherwise, $\sigma_C \in \overline{\text{Conv}}(\mathbb{R}^n)$

Definition

$f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ is **strictly convex** if it is proper and, for every $x_0 \neq x_1 \in \text{dom } f$ and $\alpha \in (0, 1)$,

$$f(\alpha x_0 + (1 - \alpha)x_1) < \alpha f(x_0) + (1 - \alpha)f(x_1)$$

Def: For $\mu \geq 0$, define

$$\text{Conv}_\mu(\mathbb{R}^n) = \{f : \mathbb{R}^n \rightarrow (\infty, \infty] : f - \mu\|\cdot\|^2/2 \in \text{Conv}(\mathbb{R}^n)\}$$

$$\overline{\text{Conv}}_\mu(\mathbb{R}^n) = \{f : \mathbb{R}^n \rightarrow (\infty, \infty] : f - \mu\|\cdot\|^2/2 \in \overline{\text{Conv}}(\mathbb{R}^n)\}$$

Clearly,

$$\text{Conv}_0(\mathbb{R}^n) = \text{Conv}(\mathbb{R}^n) \quad \overline{\text{Conv}}_0(\mathbb{R}^n) = \overline{\text{Conv}}(\mathbb{R}^n)$$

Def: Function $f \in \text{Conv}_\mu(\mathbb{R}^n)$ is called a μ -strongly convex function if $\mu > 0$, or simply, a μ -convex function if $\mu \geq 0$

Def: Let $f : \mathbb{R}^n \rightarrow (-\infty, \infty]$ and nonempty convex set $C \subset \text{dom } f$ be given. f is said to be **strictly convex on C** if $f + \delta_C$ is strictly convex

Proposition

If $f : \mathbb{R}^n \rightarrow (-\infty, \infty]$ is strictly convex on a nonempty convex set C , then

$$\min\{f(x) : x \in C\}$$

has at most one global minimum

Obs: Previous result does not guarantee the existence of a global minimum but simply that there exists at most one (if any).

Def: For nonempty convex set C and $\mu \geq 0$, define

$$\text{Conv}(C) := \{f : C \subset \text{dom } f, f + \delta_C \in \text{Conv}(\mathbb{R}^n)\}$$

$$\overline{\text{Conv}}(C) := \{f : C \subset \text{dom } f, f + \delta_C \in \overline{\text{Conv}}(\mathbb{R}^n)\}$$

$$\text{Conv}_\mu(C) := \{f : C \subset \text{dom } f, f + \delta_C \in \text{Conv}_\mu(\mathbb{R}^n)\}$$

$$\overline{\text{Conv}}_\mu(C) := \{f : C \subset \text{dom } f, f + \delta_C \in \overline{\text{Conv}}_\mu(\mathbb{R}^n)\}$$

$\text{Conv}(C)$ (resp., $\overline{\text{Conv}}(C)$) is the set of convex (resp., closed convex) functions on C .

Proposition

Let $\mu > 0$ and $C \neq \emptyset$ be a convex set. If $f \in \overline{\text{Conv}}_{\mu}(C)$, then the problem

$$f_* := \inf\{f(x) : x \in C\}$$

has a unique optimal solution x_* and

$$f(x) \geq f_* + \frac{\mu}{2} \|x - x_*\|^2 \quad \forall x \in C.$$

Proposition

Assume that f is differentiable on a convex set $\emptyset \neq C \subset \mathbb{R}^n$. Then, the following are equivalent:

- (a) f is convex on C , or equivalently, for every $x, x' \in C$ and $t \in (0, 1)$,

$$f((1-t)x + tx') \leq (1-t)f(x) + tf(x')$$

- (b) for every $x', x \in C$,

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

- (c) for every $x', x \in C$,

$$\langle \nabla f(x') - \nabla f(x), x' - x \rangle \geq 0$$

Proposition

Assume that f is differentiable on a convex set $\emptyset \neq C \subset \mathbb{R}^n$.

Then, the following are equivalent:

- (a) f is strictly convex on C , or equivalently, for every $x, x' \in C$ such that $x \neq x'$ and $t \in (0, 1)$,

$$f((1-t)x + tx') < (1-t)f(x) + tf(x')$$

- (b) for every $x', x \in C$ such that $x' \neq x$,

$$f(x') > f(x) + \langle \nabla f(x), x' - x \rangle;$$

- (c) for every $x', x \in C$ such that $x' \neq x$,

$$\langle \nabla f(x') - \nabla f(x), x' - x \rangle > 0$$

Proposition

Assume that f is differentiable on a convex set $\emptyset \neq C \subset \mathbb{R}^n$. Then, for any constant $\beta \in \mathbb{R}$, the following are equivalent:

- (a) $f - \beta \|\cdot\|^2/2$ is convex on C
- (b) for every $x', x \in C$,

$$f(x') \geq f(x) + \langle \nabla f(x), x' - x \rangle + \frac{\beta}{2} \|x' - x\|^2 \quad (*)$$

- (c) for every $x', x \in C$,

$$\langle \nabla f(x') - \nabla f(x), x' - x \rangle \geq \beta \|x' - x\|^2$$

Letting $\ell_f(x'; x) := f(x) + \langle \nabla f(x), x' - x \rangle$, then $(*)$ becomes

$$f(x') \geq \ell_f(x'; x) + \frac{\beta}{2} \|x' - x\|^2$$

Def: For $m > 0$, f is m -**weakly convex** on C if $f + m\|\cdot\|^2/2$ is convex on C . In such case, we say that m is a **lower curvature** of f on C

Corollary

Assume that f is differentiable on a convex set $\emptyset \neq C \subset \mathbb{R}^n$ and let $m > 0$ be given. Then, the following are equivalent:

- f is m -weakly convex on C
- for all $x, x' \in C$,

$$f(x') \geq \ell_f(x'; x) - \frac{m}{2} \|x' - x\|^2$$

- for all $x, x' \in C$,

$$\langle \nabla f(x') - \nabla f(x), x' - x \rangle \geq -m \|x' - x\|^2$$

Proof: Follows from previous prop with $\beta = -m$

Definition

$M \in \mathbb{R}$ is an **upper curvature** of f on C if

$$\frac{M}{2} \|\cdot\|^2 - f(\cdot)$$

is convex on C

Clearly, $M \in \mathbb{R}$ is an upper curvature of f on C iff

$$f(x') \leq f(x) + \langle \nabla f(x), x' - x \rangle + \frac{M}{2} \|x' - x\|^2 \quad \forall x, x' \in C$$

or equivalently,

$$\langle \nabla f(x') - \nabla f(x), x' - x \rangle \leq M \|x' - x\|^2 \quad \forall x, x' \in C$$

Definition

$f : \mathbb{R}^n \rightarrow (-\infty, \infty]$ is L -smooth on a nonempty convex set $C \subset \mathbb{R}^n$ if f is differentiable on C and ∇f is L -Lipschitz continuous on C , i.e.,

$$\|\nabla f(x') - \nabla f(x)\| \leq L\|x' - x\| \quad x, x' \in C$$

Proposition

Assume that f is L -smooth on a convex set $\emptyset \neq C \subset \mathbb{R}^n$. Then, L is both a lower and an upper curvature of f on C , and hence f is L -weakly convex on C

Assume $f \in \text{Conv}(\mathbb{R}^n)$ and $\bar{x} \in \text{dom } f$

Def: $s \in \mathbb{R}^n$ is called a **subgradient** of f at \bar{x} if

$$f(x) \geq f(\bar{x}) + \langle s, x - \bar{x} \rangle \quad \forall x \in \mathbb{R}^n$$

The set of all subgradients of f at \bar{x} is denoted by $\partial f(\bar{x})$ and the point-to-set map $\partial f(\cdot)$ is called the **subdifferential** of f

Proposition

The following statements hold:

- a) \bar{x} is a global minimizer of

$$\inf\{f(x) : x \in \mathbb{R}^n\}$$

if and only if $0 \in \partial f(\bar{x})$

- b) *for every $x \in \text{dom } f$, the set $\partial f(x)$ is (possibly, empty) closed convex*

Assume that $f : \mathbb{R}^n \rightarrow (-\infty, \infty]$ and $\bar{x} \in \text{dom } f$

Definition

The directional derivative of f at \bar{x} along $d \in \mathbb{R}^n$ is

$$f'(\bar{x}; d) := \lim_{t \downarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t}$$

whenever the above limit exists (possibly, equal to $\pm\infty$).

Proposition

If f is differentiable at \bar{x} then, for every $d \in \mathbb{R}^n$, $f'(\bar{x}; d)$ exists and

$$f'(\bar{x}; d) = f'(\bar{x})d = \langle \nabla f(\bar{x}), d \rangle$$

Proposition

Let $f \in \text{Conv}(\mathbb{R}^n)$, $\bar{x} \in \text{dom } f$ and $d \in \mathbb{R}^n$ be given. Then:

a) the function

$$t > 0 \mapsto \frac{f(\bar{x} + td) - f(\bar{x})}{t}$$

is non-decreasing

b) for every $d \in \mathbb{R}^n$, $f'(\bar{x}; d)$ is well-defined and

$$f'(\bar{x}; d) = \inf_{t>0} \frac{f(\bar{x} + td) - f(\bar{x})}{t}$$

c) $f'(\bar{x}; \cdot)$ is convex

Relationship of subgradient and directional derivative

Proposition

Assume that $f \in \text{Conv}(\mathbb{R}^n)$ and $\bar{x} \in \text{dom } f$. Then,

$$\partial f(\bar{x}) = \{s \in \mathbb{R}^n : \langle s, \cdot \rangle \leq f'(\bar{x}; \cdot)\}$$

and

$$\text{cl}[f'(\bar{x}; \cdot)] = \sup\{\langle s, \cdot \rangle : s \in \partial f(\bar{x})\}$$

Proposition

Assume that $f \in \text{Conv}(\mathbb{R}^n)$ and $\bar{x} \in \text{dom } f$. Then:

- a) $\partial f(\bar{x}) = \emptyset$ if and only if there exists $d_0 \in \mathbb{R}^n$ such that

$$f'(\bar{x}; d_0) = -\infty$$

- b) if $\bar{x} \in \text{ri}(\text{dom } f)$, then $\partial f(\bar{x}) \neq \emptyset$ and

$$f'(\bar{x}; d) = \sup\{\langle d, s \rangle : s \in \partial f(\bar{x})\}$$

- c) $\bar{x} \in \text{int}(\text{dom } f)$ if and only if $\partial f(\bar{x})$ is non-empty and bounded, in which case

$$f'(\bar{x}; d) = \max\{\langle d, s \rangle : s \in \partial f(\bar{x})\}$$

Subgradient versus differentiability

Proposition

Assume that $f \in \text{Conv}(\mathbb{R}^n)$ and $\bar{x} \in \text{dom } f$. Then $\partial f(\bar{x})$ is a singleton if and only if f is differentiable at \bar{x} , in which case

$$\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$$

Def: The **conjugate** of a function $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ is

$$f^*(s) := \sup_{x \in \mathbb{R}^n} \langle s, x \rangle - f(x), \quad \forall s \in \mathbb{R}^n$$

Proposition

Let $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be given. Then, the following statements hold:

- a) $f = +\infty$ if and only if $f^* = -\infty$*
- b) if $f(x^0) = -\infty$ for some $x^0 \in \mathbb{R}^n$, then $f^* = +\infty$*
- c) if $f \neq \infty$ and f is minorized by some affine function (e.g., $f \in \text{Conv}(\mathbb{R}^n)$), then $f^* \in \overline{\text{Conv}}(\mathbb{R}^n)$*

Proposition

Let $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be given. Then:

- a) $-f^*(0) = \inf\{f(x) : x \in \mathbb{R}^n\}$
- b) for every $s \in \mathbb{R}^n$,

$$\begin{aligned} f^*(s) &= -\inf\{f(x) - \langle s, x \rangle : x \in \mathbb{R}^n\} \\ &= \inf\{\beta \in \mathbb{R} : f \geq \langle s, \cdot \rangle - \beta\} \end{aligned}$$

and the infimum is achieved whenever $f^*(s)$ is finite

- c) (Fenchel's inequality) for any $x, s \in \mathbb{R}^n$,

$$f^*(s) \geq \langle x, s \rangle - f(x)$$

Proposition

Let $f \in \text{Conv}(\mathbb{R}^n)$ and $x \in \text{dom } f$ be given. Then,

$$s \in \partial f(x) \iff f^*(s) \leq \langle s, x \rangle - f(x)$$

(or equivalently, $f^*(s) = \langle s, x \rangle - f(x)$)

Proposition

For functions $f, g : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$, vectors $x_0, s_0 \in \mathbb{R}^n$ and scalar $\alpha \in \mathbb{R}$, the following statements hold:

- a) if $g = f + \alpha$, then $g^* = f^* - \alpha$
- b) if $\alpha > 0$ and $g = \alpha f$, then $g^*(s) = \alpha f^*(s/\alpha)$ for every $s \in \mathbb{R}^n$
- c) if $\alpha \neq 0$ and $g(x) = f(\alpha x)$ for every $x \in \mathbb{R}^n$, then $g^*(s) = f^*(s/\alpha)$ for every $s \in \mathbb{R}^n$
- d) if $g(x) = f(x - x_0)$ for every $x \in \mathbb{R}^n$, then $g^*(s) = f^*(s) + \langle s, x_0 \rangle$ for every $s \in \mathbb{R}^n$
- e) if $g(x) = f(x) + \langle x, s_0 \rangle$ for every $x \in \mathbb{R}^n$, then $g^*(s) = f^*(s - s_0)$ for every $s \in \mathbb{R}^n$
- f) if $f \leq g$, then $f^* \geq g^*$

Proposition

If $f \in \overline{\text{Conv}}(\mathbb{R}^n)$ then

$$f^{**} = f$$

where $f^{**} := (f^*)^*$

As a consequence, for every $x \in \mathbb{R}^n$,

$$f(x) = \sup_s \langle s, x \rangle - f^*(s)$$

Proposition

Assume that $f \in \overline{\text{Conv}}(\mathbb{R}^n)$. Then,

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

As a consequence, $\partial f^(0)$ is equal to the set of minimizers of*

$$\inf\{f(x) : x \in \mathbb{R}^n\}$$

Proposition

Assume that $f \in \overline{\text{Conv}}(\mathbb{R}^n)$. Then, the following are equivalent:

- a) f is μ -convex
- b) f^* is $(1/\mu)$ -smooth on \mathbb{R}^n

Application: Assume that $h \in \overline{\text{Conv}}_\mu(\mathbb{R}^n)$ and D is a closed convex set such that $D \cap \text{dom } h \neq \emptyset$. Define

$$g(x) = \max_{y \in D} \langle y, Ax \rangle - h(y) \quad \forall x \in \mathbb{R}^n$$

Then, g is $(\|A\|^2/\mu)$ -smooth

Let $f \in \text{Conv}(\mathbb{R}^n)$, $\bar{x} \in \text{dom } f$ and $\varepsilon \geq 0$ be given

Definition

$s \in \mathbb{R}^n$ is called an ε -**subgradient** of f at \bar{x} if

$$f(x) \geq f(\bar{x}) + \langle s, x - \bar{x} \rangle - \varepsilon \quad \forall x \in \mathbb{R}^n.$$

The set of all ε -subgradients of f at \bar{x} is denoted by $\partial_\varepsilon f(\bar{x})$ and the point-to-set map $\partial_\varepsilon f(\cdot)$ is called the ε -**subdifferential** of f

Proposition

The following statements hold:

- a) \bar{x} is an ε -minimizer of $f_* := \inf\{f(x) : x \in \mathbb{R}^n\}$, i.e.,

$$f(\bar{x}) - f_* \leq \varepsilon$$

or equivalently,

$$f(\bar{x}) \leq f(x) + \varepsilon \quad \forall x \in \mathbb{R}^n$$

if and only if $0 \in \partial_\varepsilon f(\bar{x})$

- b) $\partial_\varepsilon f(\bar{x})$ is a (possibly, empty) closed convex set

Relationship of ε -subgradients and conjugate function

Proposition

Let $f \in \text{Conv}(\mathbb{R}^n)$, $x \in \text{dom } f$ and $s \in \mathbb{R}^n$ be given. Then,

$$s \in \partial_\varepsilon f(x) \iff f^*(s) \leq \langle s, x \rangle - f(x) + \varepsilon$$

Transportation formula for subgradient

Proposition

Let $f \in \text{Conv}(\mathbb{R}^n)$, $x \in \text{dom } f$ and $s \in \mathbb{R}^n$ be given and define

$$\bar{\varepsilon} := f^*(s) + f(x) - \langle s, x \rangle$$

Then, the following statements hold:

(a) if $\bar{\varepsilon} < \infty$, then

$$\bar{\varepsilon} = \min\{\varepsilon : s \in \partial_\varepsilon f(x)\}$$

(b) if $s \in \partial f(\bar{x})$ for some $\bar{x} \in \text{dom } f$, then $s \in \partial_{\bar{\varepsilon}} f(x)$ and

$$\bar{\varepsilon} = f(x) - f(\bar{x}) - \langle s, x - \bar{x} \rangle < \infty$$

Proposition

If $f \in \text{Conv}(\mathbb{R}^n)$ and $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is a linear map, then $f \circ A$ is convex and

$$\partial_\varepsilon(f \circ A)(x) \supset A^* \partial_\varepsilon f(Ax) \quad \forall x \in \mathbb{R}^m$$

and equality holds whenever

$$A(\mathbb{R}^m) \cap \text{ri}(\text{dom } f) \neq \emptyset$$

Proposition

If $f_i \in \overline{\text{Conv}}(\mathbb{R}^n)$ for $i = 1, \dots, m$, then $f := f_1 + \dots + f_m$ is convex and

$$\partial f(x) \supset \partial f_1(x) + \dots + \partial f_m(x) \quad \forall x \in \mathbb{R}^n$$

and equality holds whenever

$$\bigcap_{i=1}^m \text{ri}(\text{dom } f_i) \neq \emptyset$$