Background on Convex Analysis

John Duchi

Outline

Why convex analysis and optimization?

Consider problem

$$\underset{x}{\text{minimize}} \;\; f(x) \quad \text{subject to} \;\; x \in X.$$

When is this (efficiently) solvable?

- When things are convex
- *If* we can formulate a numerical problem as minimization of a convex function f over a convex set X, then (roughly) it is solvable

Convex sets

Definition

A set $C \subset \mathbb{R}^n$ is *convex* if for any $x, y \in C$

$$tx + (1-t)y \in C \quad \text{for all } t \in C$$

Examples

Hyperplane: Let $a \in \mathbb{R}^n$, $b \in \mathbb{R}$,

$$C := \{x \in \mathbb{R}^n : \langle a, x \rangle = b\}.$$

Polyhedron: Let $a_1, a_2, \ldots, a_m \in \mathbb{R}^n$, $b \in \mathbb{R}^m$,

$$C := \{x : Ax \leq b\} = \{x \in \mathbb{R}^n : \langle a_i, x \rangle \leq b_i\}$$

Examples

Norm balls: let $\|\cdot\|$ be any norm,

$$C := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

Basic properties

Intersections: If C_α, $\alpha \in \mathcal{A}$, are all convex sets, then

$$C := \bigcap_{\alpha \in \mathcal{A}} C_\alpha \text{ is convex.}$$

Minkowski addition: If C, D are convex sets, then $C + D := \{x + y : x \in C, y \in D\}$ is convex.

Basic properties

Convex hulls: If $x_1, \ldots, x_m$ are points,

$$\mathrm{Conv}\{x_1, \ldots, x_m\} := \left\{ \sum_{i=1}^m t_i x_i : t_i \geq 0, \sum_{i=1}^m t_i = 1 \right\}.$$

Projections

Let C be a closed convex set. Define

$$\pi_C(x) := \operatorname*{argmin}_{y \in C} \{\|x - y\|_2^2\}.$$

Existence: Assume $x = 0$ and that $0 \notin C$. Show that if x_n is such that $\|x_n\|_2 \to \inf_{y \in C} \|y\|_2$ then x_n is Cauchy.

Projections

Characterization of projection

$$\pi_C(x) := \operatorname*{argmin}_{y \in C} \left\{ \|x - y\|_2^2 \right\}$$

We have $\pi_C(x)$ is the projection of x onto C if and only if

$$\langle \pi_C(x) - x, y - \pi_C(x) \rangle \geq 0 \;\text{ for all } y \in C.$$

Separation Properties

Separation of a point and convex set: Let C be closed convex and $x \notin C$. Then there exists a *separating hyperplane*: there is $v \neq 0$ and a constant b such that

$$\langle v, x \rangle > b \;\text{ and }\; \langle v, y \rangle \leq b \text{ for all } y \in C.$$

Proof: Show the stronger result

$$\langle v, x \rangle \geq \langle v, y \rangle + \|x - \pi_C(x)\|_2^2 \, .$$

Separation Properties

Separation of two convex sets: Let C be convex and compact and D be closed convex. Then there is a non-zero separating hyperplane v

$$\inf_{x \in C} \langle v, x \rangle > \sup_{y \in D} \langle v, y \rangle .$$

Proof: $D - C$ is closed convex and $0 \notin D - C$

Supporting hyperplanes

Supporting hyperplane: A hyperplane $\{x : \langle v, x \rangle = b\}$ *supports* the set C if C is contained in the halfspace $\{x : \langle v, x \rangle \leq b\}$ and for some $y \in \operatorname{bd} C$ we have $\langle v, y \rangle = b$

Supporting hyperplanes

Let C be a convex set with $x \in \operatorname{bd} C$. Then there exists a non-zero supporting hyperplane H passing through x. That is, a $v \neq 0$ such that

$$C \subset \{y : \langle v, y \rangle \leq b\} \quad \text{and} \quad \langle v, x \rangle = b.$$

Proof:

Convex functions

A function f is *convex* if its domain $\mathrm{dom}\, f$ is a convex set and for all $x, y \in \mathrm{dom}\, f$ we have

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y) \ \text{ for all } t \in [0,1].$$

(Define $f(z) = +\infty$ for $z \notin \mathrm{dom}\, f$)

Convex functions and epigraphs

Duality between convex sets and functions. A function f is convex if and only if its epigraph

$$\operatorname{epi} f := \{(x,t) : f(x) \le t, t \in \mathbb{R}\}$$

is convex

Minima of convex functions

Why convex? Let x be a local minimizer of f on the convex set C. Then *global* minmization:

$$f(x) \le f(y) \text{ for all } y \in C.$$

Proof: Note that for $y \in C$,

$$f(x + t(y - x)) = f((1 - t)x + ty) \le (1 - t)f(x) + tf(y).$$

Subgradients

A vector g is a *subgradient* of f at x if

$$f(y) \geq f(x) + \langle g, y - x \rangle \quad \text{for all } y.$$

Subdifferential

The *subdifferential* (subgradient set) of f at x is

$$\partial f(x) := \{g : f(y) \geq f(x) + \langle g, y - x\rangle \text{ for all } y\}.$$

Subdifferential examples

Let $f(x) = |x| = \max\{x, -x\}$. Then

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

Existence of subgradients

Theorem: Let $x \in \operatorname{int} \operatorname{dom} f$. Then $\partial f(x) \neq \emptyset$

Proof: Supporting hyperplane to $\operatorname{epi} f$ at the point $(x, f(x))$

Optimality properties

Subgradient optimality A point x minimizes f if and only if $0 \in \partial f(x)$. Immediate:

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

Advanced optimality properties

Subgradient optimality Consider problem

$$\underset{x}{\text{minimize}} \;\; f(x) \quad \text{subject to} \;\; x \in C.$$

Then x solves problem if and only if there exists $g \in \partial f(x)$ such that

$$\langle g, y - x \rangle \geq 0 \;\; \text{for all } y \in C.$$

Subgradient calculus

Addition Let $f_1, \ldots, f_m$ be convex and $f = \sum_{i=1}^m f_i$. Then

$$\partial f(x) = \sum_{i=1}^m \partial f_i(x) = \left\{ \sum_{i=1}^m g_i : g_i \in \partial f_i(x) \right\}$$

(Extends to infinite sums, integrals, etc.)

Subgradient calculus

Composition Let $A \in \mathbb{R}^{n \times m}$ and $f : \mathbb{R}^n \to \mathbb{R}$ be convex, with $h(x) = f(Ax)$. Then

$$\partial h(x) = A^T \partial f(Ax) = \{A^T g : g \in \partial f(Ax)\}.$$

Subgradient calculus: maxima

Let f_i, $i = 1, \ldots, m$, be convex

$$f(x) = \max_i f_i(x)$$

Then with $I(x) = \{i : f_i(x) = f(x) = \max_j f_j(x)\}$,

$$\partial f(x) = \operatorname{Conv}\{g_i : g_i \in \partial f_i(x), i \in I(x)\}.$$

Subgradient calculus: suprema

Let $\{f_\alpha : \alpha \in \mathcal{A}\}$ be a collection of convex functions,

$$f(x) = \sup_{\alpha \in \mathcal{A}} f_\alpha(x).$$

Then if the supremum is attained and $\mathcal{A}(x) = \{\alpha : f_\alpha(x) = f(x)\}$,

$$\partial f(x) \subset \mathrm{Conv}\,\{g_\alpha : g_\alpha \in \partial f_\alpha(x), \alpha \in \mathcal{A}(x)\}$$

Examples of subgradients

Norms: Recall ℓ_∞-norm, $\|\cdot\|_\infty$,

$$\|x\|_\infty = \max_j |x_j|$$

Then

$$\partial \|x\|_\infty = \mathrm{Conv}\left\{\{e_i : \langle e_i, x\rangle = \|x\|_\infty\} \cup \{-e_i : \langle -e_i, x\rangle = \|x\|_\infty\}\right\}$$

Examples of subgradients

General norms: Recall dual norm $\|\cdot\|_*$ of norm $\|\cdot\|$

$$\|y\|_* = \sup_{x:\|x\|\leq 1} \langle x, y \rangle \quad \text{and} \quad \|x\| = \sup_{y:\|y\|_*\leq 1} \langle x, y \rangle \,.$$

Then

$$\partial \|x\| = \{y : \|y\|_* \leq 1, \langle y, x \rangle = \|x\|\}.$$