Convex Analysis for Optimization

Olga Kuryatnikova (Erasmus University Rotterdam)
kuryatnikova@ese.eur.nl

September-October 2024
Lecture 3

Course plan

- Week 1: Introduction to convexity
- Week 2: More on convex sets
- Week 3: Dual view of convex sets + more on convex functions
- Week 4: Dual view of convex functions
- Week 5: Duality and optimization
- Week 6: Introduction to algorithms, descend methods
- Week 7: Proximal methods, projected gradients
- Weeks 8 - 9: Fix point approach, averaged operators

Special cones

Convex cones based on a given set:

- Conic hull (smallest convex cone containing a given set)
- Recession cone (determines directions of unboundedness of a set)
- Polar cone (dual description of a set)
- Dual cone (dual description of a set)
- Normal cone (dual descriptions, optimality conditions)

Polar and dual cones of set $S \subseteq \mathbb{R}^n$

Polar cone: $S^\circ := \{x \in \mathbb{R}^n : x^\top y \leq 0 \ \forall y \in S\}$

Dual cone: $S^* := -S^\circ = \{x \in \mathbb{R}^n : x^\top y \geq 0 \ \forall y \in S\}$

Polar and dual cone properties

For a non-empty set S (so, no convexity or closedness assumed):

$$S^\circ = \mathrm{cl}(S)^\circ = \mathrm{conv}(S)^\circ = \mathrm{cone}(S)^\circ$$

Polar Cone Thm: $(S^\circ)^\circ = \mathrm{cl}(\mathrm{conv}(S))$ for a non-empty cone S

All above also holds for the dual cone (i.e., if we replace $\circ$ with $*$).

Normal cone

Def: the normal cone of $S \subseteq \mathbb{R}^n$ in $x \in S$ is

$$N_S(x) := \{y \in \mathbb{R}^n : y^\top (z - x) \leq 0 \ \forall z \in S\}.$$

That is, the normal cone of S in x is $(S - x)^\circ$.

$N_S(x) = \{0\}$ if $x \in \text{int}(S)$; $N_S(x)$ contains at least one half-line otherwise.

Occurs in duality, optimality conditions.

Hyperplanes

Recall: hyperplane for some $0 \neq a \in \mathbb{R}^n,\ b \in \mathbb{R}$

$H := \{x \in \mathbb{R}^n : a^\top x = b\} = \bar{x} + \{x \in \mathbb{R}^n : a^\top x = 0\}$ for some $\bar{x} \in H$

Def: H separates sets $S, \bar{S} \subseteq \mathbb{R}^n$ if $a^\top y \leq b \leq a^\top x,\ \forall y \in \bar{S}\ \forall x \in S$

Set separation by hyperplane

$S, \bar{S} \subseteq \mathbb{R}^n$ separable: $\exists a \neq 0, b : a^\top y \leq b \leq a^\top x,\ \forall y \in \bar{S}\ \forall x \in S$.
Equiv: $\exists a : \sup_{y \in \bar{S}} a^\top y \leq \inf_{x \in S} a^\top x$ (*)

Properly separable: (*) and $\exists a \neq 0 : \inf_{y \in \bar{S}} a^\top y < \sup_{x \in S} a^\top x$

Strictly separable: $\exists a \neq 0, b : a^\top y < b < a^\top x,\ \forall y \in \bar{S}\ \forall x \in S$

Strongly separable: $\exists a \neq 0 : \sup_{y \in \bar{S}} a^\top y < \inf_{x \in S} a^\top x$
Equiv: $a^\top y \leq a^\top x,\ \forall y \in \bar{S} + B(0, \varepsilon)\ \forall x \in S + B(0, \varepsilon)$, for some $\varepsilon > 0$

Proper Separation Theorem

Thm: Let $S \subseteq \mathbb{R}^n$ be nonempty and convex, and let $\bar{x} \in \mathbb{R}^n$. There is a hyperplane *properly* separating $\bar{x}$ and S if and only if $x \notin \text{ri}(S)$.

Def: hyperplane $H := \bar{x} + \{x \in \mathbb{R}^n : a^\top x = 0\}$ supports set S in point $\bar{x}$ if $\bar{x} \in \text{cl}(S)$ and $\inf_{x \in S} a^\top x = a^\top \bar{x}$ (*i.e.*, $a^\top \bar{x} \leq a^\top x \ \forall x \in S$).

Proof of Proper Separation Theorem

Separating two sets

Let $S, \bar{S} \subseteq \mathbb{R}^n$ be nonempty, convex, and disjoint.

Thm (separation): There is a hyperplane separating S and $\bar{S}$.

Thm (proper s.): There is a hyperplane *properly* separating S and $\bar{S}$ if and only if $\text{ri}(S) \cap \text{ri}(\bar{S}) = \emptyset$.

Thm (strict s.): There is a hyperplane *strictly* separating S and $\bar{S}$ if $S - \bar{S}$ is closed.

Thm (strong s.): There is a hyperplane *strongly* separating S and $\bar{S}$ if and only if $0 \notin \text{cl}(S - \bar{S})$.

Dual description of convex sets

Thm: For $S \subseteq \mathbb{R}^n$, the set cl(conv(S)) is the intersection of the closed halfspaces that contain S.

Corollary: If S is closed and convex, S is the intersection of the closed halfspaces that contain it.

More info on convex functions

- Popular convex functions
- Convexity preserving operations on functions
- Continuity and closedness
- Differentiable convex functions (for next lecture)

Some definitions

Consider $f : \mathbb{R}^n \to (-\infty, \infty]$.

Convexity/continuity/some other property over a set

Def: Let $C \subseteq S \subseteq \mathbb{R}^n$, then f is [property] over S if its restriction to S defined by $[\hat{f} : S \to \overline{\mathbb{R}},\ \hat{f}(x) = f(x)\ \forall x \in S]$ is [property].

Strict convexity

Def: f is strictly convex if $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y)$ for $\alpha \in (0,1)$, $x \neq y$.

Strong convexity

Def: f is strongly convex if $f - \sigma \|x\|_2^2$ is convex for some $\sigma > 0$.

Popular convex functions

- Affine functions: $a^\top x + b$ or $\sum_{i=1}^{n} A_{ij}x_{ij} + b$ if x is a matrix
- Norms: l_p norm for $p \geq 1$: $(\sum_{i=1}^{n} = |x|^{1/p})^p$, ∞-norm: $\max_{i=1}^{n} |x_i|$; spectral norm: $\sigma_{\max}(x) = (\lambda_{\max}(x^\top x))^{1/2}$ if x is a matrix
- Sums of squares of polynomials: $\sum_{j=1}^{m} (p_j(x))^2$
- Max: $\max_{i=1}^{n} x_i$
- Log-sum-exp: $\log(\sum_{i=1}^{n} \exp(x_i))$
- log-determinant: -log(det(x)) is convex on the set of positive definite matrices x

Convexity preserving operations on functions

Consider functions $\mathbb{R}^n \to (-\infty, \infty]$. The following is convex.

- Sum of convex functions, even of infinitely many functions
- For convex f: $f(Ax+b)$, $A \in \mathbb{R}^{n\times m}$, $b \in \mathbb{R}^n$
- Conic combination of convex functions
- Supremum: $\sup_{i\in I} f_i(x)$, $\sup_{y\in Y} f(x,y)$ if f is convex in x for all y
- Partial inf: $\inf_{y\in Y} f(x,y)$ if Y is convex and f is convex in (x,y)
- Integral: $\int_{y\in Y} f(x,y)dy$ if f is convex in x for all y
- Composition: $f(g_1(x),\dots,g_n(x))$ is convex if f is convex and for each $i=1,\dots,n$ at least one of three facts holds: [g_i convex, f non-decreasing in x_i]; [g_i concave, f non-increasing in x_i]; [g_i affine]

Restricting a convex function to a line

Thm: $f : \mathbb{R}^n \to (-\infty, \infty]$ is convex if and only if its restriction to a line $g_{x,v}(t)$ is convex for any fixed x, v, where $g(t) := f(x + tv)$.

Types of continuity

Consider a function $f : S \to \overline{\mathbb{R}}$

Def: f is lower semicontinuous in x if $f(x) \leq \liminf_{y \to x} f(y), \forall (y) \subset S$.

Def: f is continuous in $x \in \mathrm{dom}(f)$ if $f(x) = \lim_{y \to x} f(y), \forall (y) \subset \mathrm{dom}(f)$

Def: f is Lipshitz-continuous with constant $L > 0$ if
$\|f(x) - f(y)\|_2 \leq L\|x - y\|_2$ for all $x, y \in \mathrm{dom}(f)$

Semicontinuity and closedness

Def: $f : S \to \overline{\mathbb{R}}$ is closed if its epigraph $\text{epi}(f)$ is a closed set.

Thm: Function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is closed if and only if

$\iff$ f is lower-semicontinuous

$\iff$ level set $V_\gamma =: \{x \in \mathbb{R}^n : \gamma \geq f(x)\}$ is closed for any $\gamma \in \mathbb{R}$

Continuity and convexity

Thm: $f : S \to \overline{\mathbb{R}}$ proper and convex $\Rightarrow$ f continuous over $\mathrm{ri}(\mathrm{dom}(f))$.

Corollary: A convex function $\mathbb{R}^n \to \mathbb{R}$ is continuous.

If there is time: Farkas' lemma as an example of using polar cones

Farkas' lemma

Let $a_1, \ldots, a_m \in \mathbb{R}^n$. Then $\{x \in \mathbb{R}^n : a_j^\top x \geq 0 \ \forall j = 1, \ldots, m\}$ and $\text{cone}(a_1, \ldots, a_m)$ are closed convex cones dual to each other.

Note: Textbook uses $a_j^\top x \leq 0$, and so the cones become polar.

Interpretations of Farkas' lemma

Let $c, a_1, \ldots, a_m \in \mathbb{R}^n$. Then $c^\top x \geq 0$ for all $x \in S$, where

$$S := \{x \in \mathbb{R}^n : a_j^\top x \geq 0 \ \forall j = 1, \ldots, k, \ a_i^\top x = 0 \ \forall i = k+1, \ldots, m\}$$

if and only if

$$c = \sum_{j=1}^{k} a_j y_j + \sum_{i=k+1}^{m} a_i y_i \text{ for some } y \in \mathbb{R}^m, \ y_1, \ldots, y_k \geq 0.$$

Generalized Farkas' lemma

Let $K \subseteq \mathbb{R}^m$ be a closed convex cone, $c \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times m}$. Let the cone $\{Ay : y \in K^*\}$ be closed. Then $x^\top c \geq 0$ for all $x \in S$, where

$$S := \{x \in \mathbb{R}^n : A^\top x \in K\}$$

if and only if

$$c = Ay \text{ for some } y \in K^*.$$