IE 598: Big Data Optimization **Fall 2016**

Lecture 2: Basic Convex Analysis – August 25

Lecturer: Niao He *Scribes: Niao He*

Overview Last lecture, we discussed how optimization problems play an vital role in many engineering and science fields, especially in machine learning. We also discussed why convex optimization is of particular importance: theoretically, we know how to solve convex problems well; practically, they capture numerous interesting applications.

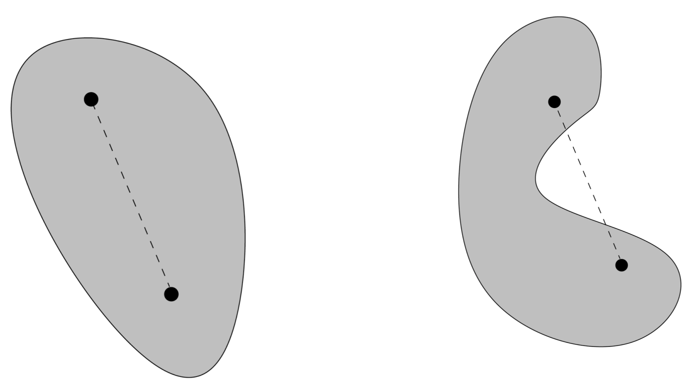
In this lecture, we will introduce the concept of convex set, examples, calculus of convexity, and some geometry results.

2.1 Convex Sets

2.1.1 Definitions

Definition 2.1 (Convex set) *A set $X \subseteq \mathbf{R}^n$ is convex if $\forall x, y \in X, \lambda x + (1-\lambda)y \in X$ for any $\lambda \in [0,1]$.*

In another word, the line segment that connects any two elements lies entirely in the set.



(a) convex set (b) non-convex set

Figure 2.1: Examples of convex sets

Given any elements $x_1, \ldots, x_k$, the combination $\lambda_1 x_1 + \lambda_2 x_1 + \ldots + \lambda_k x_k$ is called

- **Covex**: if $\lambda_i \geq 0, i = 1, \ldots, k$ and $\lambda_1 + \lambda_2 + \ldots + \lambda_k = 1$;
- **Conic**: if $\lambda_i \geq 0, i = 1, \ldots, k$;
- **Linear**: if $\lambda_i \in \mathbf{R}, i = 1, \ldots, k$.

A set is *convex* if all convex combinations of its elements are in the set; a set is a *cone* if all conic combinations of its elements are in the set; a set is a *linear subspace* if all linear combinations of its elements are in the set. Clearly, a linear subspace is always a cone; a cone is always a convex set.

Definition 2.2 (Convex hull) *A convex hull of a set $X \subseteq \mathbf{R}^n$ is the set of all convex combination of its elements, denoted as*

$$conv(X) = \left\{ \sum_{i=1}^{k} \lambda_i x_i : k \in \mathbf{N}, \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1, x_i \in X, \forall i = 1, \ldots, k \right\}.$$

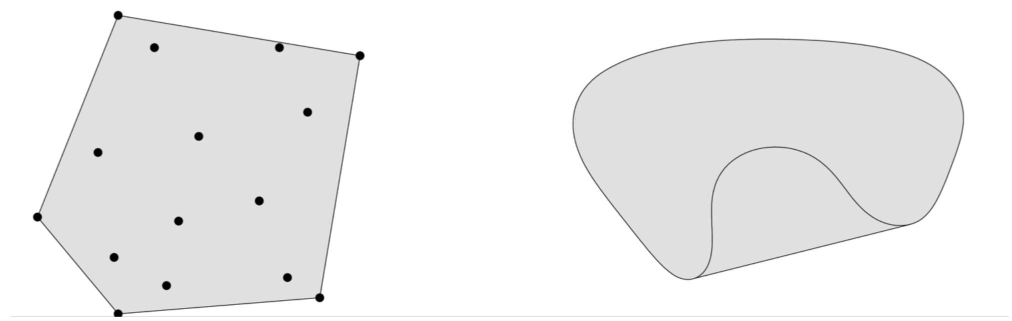


Figure 2.2: Examples of convex hulls

Remark. It follows immediately that

1. A convex hull is always convex.
2. If X is convex, then $conv(X) = X$.
3. For any set X, $conv(X)$ is the smallest convex set that contains X.

2.1.2 Examples of Convex Sets

Example 1. Some simple convex sets:

- *Hyperplane*: $\{x \in \mathbf{R}^n : a^T x = b\}$
- *Halfspace*: $\{x \in \mathbf{R}^n : a^T x \leq b\}$
- *Affine space*: $\{x \in \mathbf{R}^n : Ax = b\}$
- *Polyhedron*: $\{x \in \mathbf{R}^n : Ax \leq b\}$
- *Simplex*: $\{x \in \mathbf{R}^n : x \geq 0, \sum_{i=1}^{n} x_i = 1\} = conv(e_1, \ldots, e_n)$.

Example 2. *Norm balls*:

$$B_{\|\cdot\|} := \{x \in \mathbf{R}^n : \|x - a\| \le r\}$$

for a given norm $\|\cdot\|$ on $\mathbf{R}^n$.

[Note that $\|\cdot\|$ is a norm if

1. (positivity): $\|x\| \ge 0$ and $\|x\| + 0$ if and only if $x = 0$;
2. (homogeneity): $\|\lambda x\| = |\lambda| \cdot \|x\|, \forall \lambda \in \mathbf{R}$;
3. (triangle inequality): $\|x + y\| \le \|x\| + \|y\|, \forall x, y$

A typical example is the ℓ_p-*norm* $(1 \le p \le \infty)$:

$$\|x\|_p = \left(\sum_{i=1}^n x_i^p\right)^{1/p}.$$

- when $p = 2$, $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ is known as the *Euclidean norm*;
- when $p = 1$, $\|x\|_1 = \sum_{i=1}^n |x_i|$ is known as the ℓ_1-*norm*;
- when $p = \infty$, $\|x\|_\infty = \max\{|x_i| : i = 1, \dots, n\}$ is known as the *max norm*.]

It follows immediately from the homogeneity and triangle inequality conditions that a norm ball is always convex. If $x, y \in B_{\|\cdot\|}$ and $\lambda \in [0, 1]$, then

$$\|[\lambda x + (1 - \lambda)y] - a\| \le \lambda\|x - a\| + (1 - \lambda)\|y - a\| \le \lambda r + (1 - \lambda)r = r.$$

Hence, $\lambda x + (1 - \lambda)y \in B_{\|\cdot\|}$.

Example 3. *Ellipsoid*:

$$\{x \in \mathbf{R}^n : (x - a)^T Q(x - a) \le r^2\}$$

where $Q \succ 0$ and is symmetric.

Proof: Define $\|x\|_Q := \|Q^{1/2}x\|_2$, one can show that $\|x\|_Q$ is a valid norm. Since $(x-a)^T Q(x-a) = \|x-a\|_Q^2$, the ellipsoid can be considered as a special norm ball; hence it is convex. ■

Example 4. *ϵ-neighborhood of convex set*, i.e.

$$X^\epsilon := \{x \in \mathbf{R}^n : dist_{\|\cdot\|}(x, X) := \inf_{y \in X} \|x - y\| \le \epsilon\}.$$

Proof:

$$\begin{aligned}
x, y \in X^\epsilon &\Rightarrow \forall \epsilon' > \epsilon, \exists u, v \in X, \text{s. t. } \|x - u\| \le \epsilon, \|y - v\| \le \epsilon \\
&\Rightarrow \forall \lambda \in [0, 1], \exists \lambda u + (1 - \lambda)v \in X, \text{s. t. } \|[\lambda x + (1 - \lambda)y] - [\lambda u + (1 - \lambda)v]\| \le \epsilon' \\
&\Rightarrow \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in X^\epsilon
\end{aligned}$$

■

Example 5. *Convex cones*:

$$X \text{ is a } \underline{\text{convex cone}} \text{ if and only if } \forall x, y \in X, \lambda_1, \lambda_2 \geq 0, \lambda_1 x + \lambda_2 y \in X.$$

- *Nonnegative orthant*: $\mathbf{R}^n_+ = \{x \in \mathbf{R}^n : x \geq 0\}$
- *Lorentz cone*: $L^n = \left\{x \in \mathbf{R}^n : x_n \geq \sqrt{x_1^2 + \ldots + x_{n-1}^2}\right\}$
- *PSD cone*: $S^n_+ = \{X \in \mathbf{S}^{n \times n} : X \succeq 0\}$
- *Polyhedral cone*: $P = \{x \in \mathbf{R}^n : Ax \leq 0\}$

You can verify that all of these are convex cones. Convex cones form an extremely important class of convex sets and enjoy many properties "parallel" to those of general convex sets.

2.1.3 Calculus of Convex Sets

The following operators preserve the convexity of sets, which can be easily verified based on the definition.

1. **Intersection**: If $X_\alpha, \alpha \in \mathcal{A}$ are convex sets, then
$$\cap_{\alpha \in \mathcal{A}} X_\alpha$$
is also a convex set.
2. **Direct product**: If $X_i \subseteq \mathbf{R}^{n_i}, i = 1, \ldots, k$ are convex sets, then
$$X_1 \times \cdots \times X_k := \{(x_1, \ldots, x_k) : x_i \in X_i, i = 1, \ldots, k\}$$
is also a convex set.
3. **Weighted summation**: If $X_i \subseteq \mathbf{R}^n, i = 1, \ldots, k$ are convex sets, then
$$\lambda_1 X_1 + \cdots + \lambda_k X_k := \{\lambda_1 x_1 + \ldots + \lambda_k x_k : x_i \in X_i, i = 1, \ldots, k\}$$
is also a convex set.
4. **Affine image**: If $X \subseteq \mathbf{R}^n$ is a convex set and $\mathcal{A}(x) : x \mapsto Ax + b$ is an affine mapping from $\mathbf{R}^n$ to $\mathbf{R}^k$, then
$$\mathcal{A}(X) := \{Ax + b : x \in X\}$$
is also a convex set.
5. **Inverse affine image**: If $X \subseteq \mathbf{R}^n$ is a convex set and $\mathcal{A}(y) : y \mapsto Ay + b$ is an affine mapping from $\mathbf{R}^k$ to $\mathbf{R}^n$, then
$$\mathcal{A}^{-1}(X) := \{y : Ay + b \in X\}$$
is also a convex set.

Example 7. *Linear matrix inequalities*:

$$C := \{x \in \mathbf{R}^k : x_1 A_1 + \ldots + x_k A_k \preceq B\} = \{x \in \mathbf{R}^k : B - \sum_{i=1}^k x_i A_i \succeq 0\}$$

where $A_1, \ldots, A_k, B \in \mathbf{S}^n$ are symmetric matrices. *Proof:* This is because C can be considered as the inverse affine image of the convex set $\mathbf{S}^n$ under the affine mapping $\mathcal{A} : x \mapsto B - \sum_{i=1}^k x_i A_i$, i.e. $C = \mathcal{A}^{-1}(\mathbf{S}^n)$. ■

2.1.4 Geometry of Convex Sets

Convex sets are special because of their nice geometric properties. We cover some useful results below but without providing the proofs. Details can be found in [BV04,BN01].

Proposition 2.3 *If $X \neq \emptyset$, then*

1. *Both $int(X)$ and $cl(X)$ are convex;*
2. *If $int(X) \neq \emptyset$, then $int(X)$ is dense in $cl(X)$.*

Note that in general, for any set X, $int(X) \subseteq X \subseteq cl(X)$, but $int(X)$ and $cl(X)$ can differ dramatically. For instance, let X be the set of all irrational numbers in $(0,1)$, then $int(X) = \emptyset$, $cl(X) = [0,1]$. The proposition implies that a convex set is perfectly well characterized by its closure or interior if nonempty.

Theorem 2.4 (Carathedory) *Let $X \subseteq \mathbf{R}^n$ be any nonempty set. Then every point in $conv(X)$ is a convex combination of at most $dim(X)+1$ points from X.*

Theorem 2.5 (Seperation) *Two nonempty convex sets X_1, X_2 can be separated by a hyperlane if $int(X_1) \cap \int(X_2) \neq \emptyset$, that is,*

$$\exists a, b, s.\ t.\ X_1 \subseteq \{x : a^T x \leq b\} \text{ and } X_2 \subseteq \{x : a^T x \geq b\}.$$

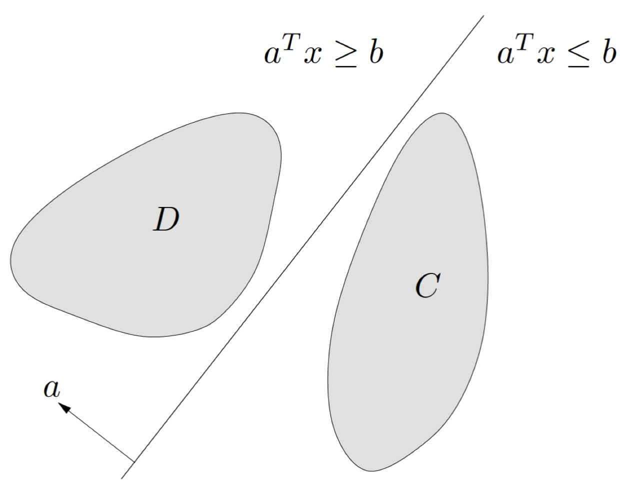


Figure 2.3: Separating hyperplane

Similarly, the supporting hyperplane theorem states that any boundary point of a convex set has a supporting hyperplane passing through it, i.e. if $\bar{x} \in cl(X)/int(X)$, then

$$\exists a, b, \text{s. t. } a^T \bar{x} = b \text{ and } a^T x \geq b, \forall x \in X.$$

References

[BV04] Boyd, S., & Vandenberghe, L. (2004). *Convex optimization.* Cambridge university press.

[BN01] Ben-Tal, A., & Nemirovski, A. (2001). *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, (Vol. 2). SIAM.