IE 521: Convex Optimization | **Spring 2017, UIUC**

Lecture 3: Separation Theorems – January 30

Instructor: Niao He | *Scribe: Shuanglong Wang*

Courtesy warning: These notes do not necessarily cover everything discussed in the class. Please email TA (swang157@illinois.edu) if you find any typos or mistakes.

In this lecture, we cover the following topics

- Separation Theorems
- The Farkas Lemma
- Duality of Linear Programs

Reference: Boyd & Vandenberghe, Chapter 2.5; Ben-Tal & Nemirovski, Chapter 1.2

3.1 Separation of Convex Sets

Definition 3.1 *Let S and T be two nonempty convex sets in $\mathbf{R}^n$, A hyperplane $H = \{x \in \mathbf{R}^n : a^T x = b\}$ with $a \neq 0$ is said to separate S and T if*

a) $S \subset H^- = \{x \in \mathbf{R}^n : a^T x \leq b\}$ *and* $T \subset H^+ = \{x \in \mathbf{R}^n : a^T x \geq b\}$

b) $S \cup T \not\subset H$

Note that a) implies that

$$\sup_{x \in S} a^T x \leq \inf_{x \in T} a^T x$$

and b) implies that

$$\inf_{x \in S} a^T x < \sup_{x \in T} a^T x$$

The separation is strict if $S \subset \{x \in \mathbf{R}^n : a^T x \leq b'\}$ and $T \subset \{x \in \mathbf{R}^n : a^T x \geq b''\}$, with $b' < b''$. Note that strict separation is equivalent to

$$\sup_{x \in S} a^T x < \inf_{x \in T} a^T x$$

Question: When can S and T be separated? strictly separated? Necessary conditions?

Theorem 3.2 *Let S and T be two nonempty convex sets. Then S and T can be separated if and only if $rint(S) \cap rint(T) = \emptyset$*

Corollary 3.3 *Let S be a nonempty convex set and $x_0 \in \partial S$. Then there exists a supporting hyperplane $H = \{x : a^T x = a^T x_0\}$ such that $S \subset \{x : a^T x \le a^T x_0\}$ and $x_0 \in H$.*

We will prove a special case of the theorem and corollary.

Theorem 3.4 *Let S be closed and convex and $x_0 \notin S$, Then there exists a hyperplane that strictly separated x_0 and S.*

Proof: Define the projection of x_0, denoted as $\text{proj}(x_0)$ to be the point in S that is closest to x_0:

$$\text{proj}(x_0) = \arg\min_{x \in S} \| x - x_0 \|_2^2$$

Note that $\text{proj}(x_0)$ exists and is unique.

- Existence: due to closedness of S
- Uniqueness: If x_1, x_2 are both closest to x_0 in S, then $\| x_1 - x_0 \|_2 = \| x_2 - x_0 \|_2 = d$. Consider $z = \frac{x_1 + x_2}{2} \in S$, then $\| z - x_0 \| \ge d$. Since $\| (x_0 - x_1) + (x_0 - x_2) \|_2^2 + \| (x_0 - x_1) - (x_0 - x_2) \|_2^2 = 2 \| x_0 - x_1 \|_2^2 + 2 \| x_0 - x_2 \|_2^2$. We have $4 \| x_0 - z \|_2^2 + \| x_1 - x_2 \|_2^2 = 4d^2$. Hence $\| x_1 - x_2 \|_2^2 = 0$, i.e. $x_1 = x_2$.

Next we show that strict separation is given by $H := \{x : a^T x = b\}$ with $a = x_0 - \text{proj}(x_0)$. $b = a^T x_0 - \frac{\|a\|_2^2}{2}$, i.e. $a^T x < b$, $\forall x \in S$, $a^T x_0 > b$.
By definition of projection and convexity, $\forall \lambda \in [0, 1], x \in S$,

$$\lambda x + (1 - \lambda)\text{proj}(x_0) \in S$$

Let $\phi(\lambda) = \| \lambda x + (1 - \lambda)\text{proj}(x_0) - x_0 \|_2^2 = \| \text{proj}(x_0) - x_0 + \lambda(x - \text{proj}(x_0)) \|_2^2$
Then

$$\phi(\lambda) \ge \phi(0), \forall \lambda \in [0, 1]$$

Hence, $\phi'(0) \ge 0$, i.e. $-2a^T(x - \text{proj}(x_0)) \ge 0$. This implies

$$a^T x \le a^T \text{proj}(x_0) = a^T(x_0 - a) = a^T x_0 - \| a \|^2 < a^T x_0 - \frac{\| a \|^2}{2} = b$$

∎

Corollary 3.5 *Let S and T be two nonempty convex sets and $S \cap T = \emptyset$. Assume $S - T$ is closed, then S and T can be strictly separated.*

Proof: Let $Y = S - T$. Since Y is a weighted sum of two convex sets, Y is nonempty and convex. Since $S \cap T$, $0 \notin Y$, from the precious theorem, $\exists a, b$ such that $a^T y < b < 0$. This implies that

$$a^T x < b + a^T z, \quad \forall x \in S, z \in T$$

Hence, $\sup_{x \in S} a^T x < \inf_{z \in T} a^T z$, i.e. S and T can be strictly separated. ∎

Remark

1. Even if both S and T are closed convex, $S - T$ might not be closed, and they might not be strictly separated.
2. When both S and T are closed convex, $S \cap T = \emptyset$ and at least one of them is bounded, then $S - T$ is closed, and S and T can be strictly separated

3.2 Theorems of alternatives

Theorem 3.6 *(**Farkas' Lemma**) Exactly one of the following sets must be empty:*

(i) $\{x \in \mathbf{R}^n : Ax = b, x \geq 0\}$

(ii) $\left\{y \in \mathbf{R}^m : A^T y \leq 0, b^T y > 0\right\}$

where $A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$.

Remark

- System (i) and (ii) are often called strong alternative, i.e. exactly one of them must be feasible.
- Farka's Lemma is particularly useful to prove infeasibility of an linear program
- Geometric interpretation: let $A = [a_1|a_2|...|a_n]$,

$$\text{Cone}\{a_1, ..., a_n\} = \left\{\sum_{i=1}^{n} x_i a_i : x_i \geq 0, i = 1, ..., n\right\}$$

$$(ii)\text{empty} \Longleftrightarrow b \notin \text{Cone}\{a_1, ..., a_n\} \Longrightarrow \exists y, y^T a_i \leq 0, \forall i = 1, ..., n, y^T b > 0$$

  Farkas' lemma can be regarded as a special case of the separation theorem.

Proof: First, we show that if system (ii) feasible, then system (i) infeasible. Otherwise, $0 < b^T y = (Ax)^T y = x^T(A^T y) \leq 0$, contradiction!
Second, we show that if system (i) infeasible, then system (ii) feasible. Let $C = \text{Cone}\{a_1, ..., a_n\}$, then C is convex and closed. Now that $b \notin C$, by the separation theorem, b and C can be (strictly) separated, i.e.

$$\exists y \in \mathbf{R}^m, \gamma \in \mathbf{R}, y \neq 0, \text{ such that }, y^T z \leq \gamma, \forall z \in C, y^T b > \gamma$$

Since $0 \in C$, we have $\gamma \geq 0$. Suppose $\gamma > 0$, and $\exists z_0 \in C$ such that $y^T z_0 > 0$, then we have $y^T(\alpha z_0) > \gamma$ when α is large enough. Hence, it suffices to set $\gamma = 0$. Since $a_1, .., a_n \in C$, we have $y^T a_i \leq 0$, $\forall i = 1, ..., m$, i.e. $A^T y \leq 0$. ■

Remark: The fact that $\text{Cone}\{a_1, ...a_m\}$ is closed is crucial. Note that in general, when S is not a finite set, $\text{Cone}(S)$ is not always closed. e.g. the conic hull of a solid circle $S = \{(x_1, x_2) : x_1^2 + (x_2 - 1)^2 < 1\}$ is the open halfspace $\{(x_1, x_2) : x_2 > 0\}$.

Variant of Farkas' Lemma Exactly one of the following two sets must be empty:

1. $\{x \in \mathbf{R}^n : Ax \leq b\}$
2. $\left\{y \geq 0 : A^T y = 0, b^T y < 0\right\}$

Proof: Exercise in HW1. ∎

3.3 LP strong duality

Consider the primal and dual pair of linear programs

$$(P) \quad \begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned} \qquad\qquad (D) \quad \begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & A^T y \leq c \end{aligned}$$

Theorem 3.7 *If (P) has a finite optimal value, then so does (D) and the two values equal each other.*

Proof: Exercise in HW 1. ∎

Remark The theorem of alternatives can be generalized to systems with convex constraints, and the strong duality of linear program can be extended to general convex programs.