IE 521 Convex Optimization

Lecture 3: Separation Theorems

Niao He

28th January 2019
Outline

Warm-up
  Quick Review
  Questions

Separation Theorems
  Definitions of Separation, Strict and Strong Separation
  Separation Hyperplane Theorem
  Strong Separation Hyperplane Theorem

Theorems of Alternatives
  Farkas’ Lemma
  LP Duality
Quick Review

▶ Radon’ Theorem
  ▶ Any set of $d + 2$ points in $\mathbb{R}^d$ can be partitioned into two disjoint sets whose convex hulls intersect.

▶ Helley’s Theorem
  ▶ If every $(d + 1)$ of the sets from a collection of $n$ sets in $\mathbb{R}^d$ intersect ($n > d$), then the whole collection of sets intersect.
What’s in common and what’s different?

Figure: Separation of sets
Separation of Sets

**Definition.** Let $S$ and $T$ be two nonempty convex sets in $\mathbb{R}^n$. A hyperplane $H = \{ x \in \mathbb{R}^n : a^T x = b \}$ with $a \neq 0$ is said to separate $S$ and $T$ if $S \cup T \not\subset H$ and

\[ S \subset H^- = \{ x \in \mathbb{R}^n : a^T x \leq b \} \]

\[ T \subset H^+ = \{ x \in \mathbb{R}^n : a^T x \geq b \} \]

Figure: Separation of two sets

- Separation is equivalent to say

\[ \sup_{x \in S} a^T x \leq \inf_{x \in T} a^T x \text{ and } \inf_{x \in S} a^T x < \sup_{x \in T} a^T x. \]
Strict Separation of Sets

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

$$S \subset H^- = \{x \in \mathbb{R}^n : a^T x < b\}$$

$$T \subset H^+ = \{x \in \mathbb{R}^n : a^T x > b\}$$

Figure: Strict Separation of two sets
Strong Separation of Sets

**Definition.** Let $S$ and $T$ be two nonempty convex sets in $\mathbb{R}^n$. A hyperplane $H = \{ x \in \mathbb{R}^n : a^T x = b \}$ with $a \neq 0$ is said to strongly separate $S$ and $T$ if there exits $b' < b < b''$ such that

$$S \subset \{ x \in \mathbb{R}^n : a^T x \leq b' \}$$
$$T \subset \{ x \in \mathbb{R}^n : a^T x \geq b'' \}$$

**Remark.**

- Strong separation $\implies$ strict separation.
- Strict separation $\nRightarrow$ strong separation.
- Strong separation is equivalent to say
  $$\sup_{x \in S} a^T x < \inf_{x \in T} a^T x.$$
Example

In all examples, $S$ and $T$ are two closed and convex sets

(A) $S$ and $T$ are separated, but not strictly;
(B) $S$ and $T$ are strictly separated, but not strongly;
(C) $S$ and $T$ are strongly separated.
Separation Hyperplane Theorem

**Theorem.** Let $S$ and $T$ be two nonempty convex sets. Then $S$ and $T$ can be separated if and only if

$$\text{rint}(S) \cap \text{rint}(T) = \emptyset.$$ 

**Corollary.** Let $S$ be a nonempty convex set and $x_0 \in \partial S$. There exists a supporting hyperplane $H = \{x : a^T x = a^T x_0\}$ with $a \neq 0$ such that

$$S \subset \{x : a^T x \leq a^T x_0\}, \text{ and } x_0 \in H.$$
Proof of Separation Theorem

S and T can be separated iff \( \text{rint}(S) \cap \text{rint}(T) = \emptyset \).

- **Necessity.**
  - \( S \) and \( T \) are separated implies that for some \( a \neq 0 \),
    \[
    \sup_{x \in S} a^T x \leq \inf_{x \in T} a^T x.
    \]
  - If \( z \in \text{rint}(S) \cap \text{rint}(T) \), then
    \[
    z = \arg\max_{x \in S} \{ a^T x \} = \arg\min_{x \in T} \{ a^T x \}.
    \]
  - The linear function \( f(x) = a^T x \) has to be constant on both \( S \) and \( T \), i.e., \( S \cap T \subset H \). (why?)
Proof of Separation Theorem

\( S \) and \( T \) can be separated iff \( \text{rint}(S) \cap \text{rint}(T) = \emptyset \).

▶ **Sufficiency.** Based on constructive steps:

1. Separation of a convex set \( S \) and \( x_0 \notin \text{cl}(S) \) (key step);
2. Separation of a convex set \( S \) and \( x_0 \notin \text{rint}(S) \);
3. Separation of 0 and \( \text{rint}(S) - \text{rint}(T) \);
4. Separation of \( S \) and \( T \).
**Proposition.** Let $S$ be convex and closed, $x_0 \notin S$. Then $x_0$ and $S$ can be separated.

**Proof.** Define

$$d(\{x_0\}, S) := \inf\{\|x_0 - x\|_2 : x \in S\}$$

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

Then $d(\{x_0\}, S) > 0$ and proj($x_0$) exists and is unique (why?). The hyperplane

$$H := \{x : a^T x = b\}, a = x_0 - \text{proj}(x_0), b = a^T x_0 - \frac{||a||_2}{2}$$

separates $x_0$ and $S$, i.e. $a^T x < b, \forall x \in S, a^T x_0 > b.$ (why?)

**Figure:** Separation of a convex set and a point
**Strong Separation Hyperplane Theorem**

**Theorem.** Let $S$ and $T$ be two nonempty convex sets. Then $S$ and $T$ can be **strongly separated** if and only if

$$\text{dist}(S, T) := \inf\{\|s - t\|_2 : s \in S, t \in T\} > 0.$$ 

In particular, if $S - T$ is closed, then $S$ and $T$ can be strongly separated.
Remarks

- “$S - T$ is closed” is only a sufficient condition for strong (strict) separation, not a necessary condition.

- Even if both $S$ and $T$ are closed convex, $S - T$ might not be closed, and they might not even be strictly separated.

- 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 strongly separated.
Proof of Strong Separation Theorem

S and T can be strongly separated iff dist(S, T) > 0.

- **Necessity.** If S and T are strongly separated, then 
  \( \exists a \neq 0 : \alpha := \sup_{x \in S} a^T x < \inf_{y \in T} a^T y := \beta. \)
  Hence, \( \forall x \in S, y \in T: \)
  \[
  \|a\|_2 \|y - x\|_2 \geq a^T (y - x) \geq \beta - \alpha \Rightarrow \|y - x\|_2 \geq \frac{\beta - \alpha}{\|a\|_2}.
  \]

- **Sufficiency.** Suppose \( r := \text{dist}(S, T) > 0 \), then \((S - T)\) and \(B(0, r)\) are two disjoint convex sets. By Separation Theorem, \( \exists a \neq 0, \)
  \[
  \sup_{z \in S - T} a^T z = \sup_{x \in S, y \in T} a^T (x - y) \leq \inf_{z \in B(0,r)} a^T z < 0.
  \]
Feasibility of Linear System

Example. Show that the following system have no solution.

\[
\begin{align*}
  x_1 - x_2 + 2x_3 & \leq 0 & \ldots \times 3 \\
  -x_1 + x_2 - x_3 & \leq 0 & \ldots \times 5 \\
  2x_1 - x_2 + 3x_3 & \leq 0 & \ldots \times 3 \\
  4x_1 - x_2 + 10x_3 & > 0 \\
\end{align*}
\]

\[3 \times \text{Eq.}(1) + 5 \times \text{Eq.}(2) + 3 \times \text{Eq.}(3) \Rightarrow 4x_1 - x_2 + 10x_3 \leq 0.\]

Note the system

\[
\begin{align*}
y_1 - y_2 + 2y_3 & = 4 \\
y_1 - y_2 - y_3 & = -1 \\
2y_1 - y_2 + 3y_2 & = 10 \\
y_1, y_2, y_3 & \geq 0 \\
\end{align*}
\]

has a solution \( y = (3, 5, 3). \)
The Celebrated Farkas’ Lemma

Theorem. (Farkas’ Lemma) Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^{m}$. Exactly one of the following sets must be empty:

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

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

▶ System (i) and (ii) are often called **strong alternatives**, i.e. exactly one of them must be feasible.

▶ This is an example of “theorem on alternatives”.

**Figure:** Gyula Farkas (1847–1930)
Geometric View of Farkas’ Lemma

Let \( A = [a_1 \mid a_2 \mid \ldots \mid a_n] \), define the cone
\[
\text{Cone} \{a_1, \ldots, a_n\} = \left\{ \sum_{i=1}^{n} x_i a_i : x_i \geq 0, i = 1, \ldots, n \right\}
\]

\[
\{Ax = b, x \geq 0\} \text{ is infeasible} \iff b \not\in \text{Cone} \{a_1, \ldots, a_n\}
\]

\[
\implies \exists y, y^T a_i \leq 0, \forall i, y^T b > 0
\]

Farkas’ lemma can be regarded as a special case of the separation theorem.
Proof of Farkas’ Lemma

- System (i) feasible ⇒ system (ii) infeasible. Otherwise, $0 < b^T y = (Ax)^T y = x^T (A^T y) \leq 0$.

- System (i) infeasible ⇒ system (ii) feasible. Let $C = \text{Cone} \{a_1, \ldots, a_n\}$. Then $C$ is convex and closed. And $b \notin C$.
  - By the separation theorem, $b$ and $C$ can be strongly separated, i.e. $\exists y \neq 0 \in \mathbb{R}^m, \gamma \in \mathbb{R}$, s.t. $y^T z \leq \gamma, \forall z \in C, y^T b > \gamma$.
  - Since $0 \in C$, we have $\gamma \geq 0$.
  - Show that $\gamma = 0$. Suppose $\gamma > 0$, and $\exists z_0 \in C$ such that $y^T z_0 > 0$, then $y^T (\alpha z_0) > \gamma$ for $\alpha$ large enough.
  - Since $a_1, \ldots, a_n \in C$, we have $y^T a_i \leq 0, \forall i = 1, \ldots, m$, i.e., $A^T y \leq 0$. 

Farkas’ Lemma
Remarks

► The closedness of the cone $\text{Cone}\{a_1, \ldots, a_n\}$ is crucial here. Note that in general, when $S$ is not a finite set, $\text{Cone}(S)$ is not always closed.

► Farkas’ Lemma can also be proved by Fourier–Motzkin elimination.

► Result can be generalized to convex inequalities.
Variant of Farkas’ Lemma

Theorem. Exactly one of the following two sets must be empty:

(i) \( \{ x \in \mathbb{R}^n : Ax \leq b \} \)
(ii) \( \{ y \geq 0 : A^T y = 0, b^T y < 0 \} \)

Theorem. Exactly one of the following two sets must be empty:

(i) \( \{ x \in \mathbb{R}^n : Ax = b \} \)
(ii) \( \{ y \in \mathbb{R}^m : A^T y = 0, b^T y \neq 0 \} \)
Duality of Linear Program

Consider the primal and dual pair of linear programs

\[
\begin{align*}
\text{(P)} : \quad & \min \ c^T x \\
\text{s.t.} \quad & Ax = b \\
\quad & x \geq 0
\end{align*}
\]

\[
\begin{align*}
\text{(D)} : \quad & \max \ b^T y \\
\text{s.t.} \quad & A^T y \leq c
\end{align*}
\]

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

**Proof:** Homework Exercise.
Who introduced LP duality?

**Figure:** Leonid Kantorovich (1912–1986)  
**Figure:** George Dantzig (1914–2005)  
**Figure:** John von Neumann (1903–1957)

Who introduced LP duality?

Who introduced LP duality?

Who introduced LP duality?

Who introduced LP duality?
References

- Boyd & Vandenberghe, Chapter 2.5
- Ben-Tal & Nemirovski, Chapter 1.2.5-1.2.6