7. Separating Hyperplane Theorems I

Daisuke Oyama

Mathematics II

May 3, 2024

Separating Hyperplane Theorem

Proposition 7.1 (Separating Hyperplane Theorem)

Suppose that $C \subset \mathbb{R}^N$, $C \neq \emptyset$, is convex and closed, and that $b \notin C$.
Then there exist $p \in \mathbb{R}^N$ with $p \neq 0$ and $c \in \mathbb{R}$ such that

$$p \cdot y \leq c < p \cdot b \text{ for all } y \in C.$$

Lemma 7.2

Suppose that $C \subset \mathbb{R}^N$, $C \neq \emptyset$, *is closed, and that* $b \notin C$.
Let $\delta = \inf\{\|z - b\| \mid z \in C\}$.
Then $\delta > 0$, *and there exists* $y^* \in C$ *such that* $\delta = \|y^* - b\|$.

Lemma 7.3

Suppose that $C \subset \mathbb{R}^N$, $C \neq \emptyset$, *is closed and convex, and that* $b \notin C$.
Let $y^* \in C$ *be such that* $\|y^* - b\| = \min\{\|z - b\| \mid z \in C\}$. *Then*

$$(b - y^*) \cdot (z - y^*) \leq 0 \text{ for all } z \in C.$$

Lemma 7.4

Suppose that $C \subset \mathbb{R}^N$, $C \neq \emptyset$, *is closed and convex, and that* $b \notin C$.
Then there exists a unique $y^* \in C$ *such that* $\|y^* - b\| = \min\{\|z - b\| \mid z \in C\}$.

Proof of Lemma 7.3

- Let $y^* \in C$ be such that $\|b - y^*\| = \min\{\|b - z\| \mid z \in C\}$.
- Take any $z \in C$ and any $\alpha \in (0, 1)$.
- Since $(1 - \alpha)y^* + \alpha z \in C$, we have

$$\begin{aligned} \|b - y^*\|^2 &\leq \|b - [(1 - \alpha)y^* + \alpha z]\|^2 \\ &= \|(b - y^*) - \alpha(z - y^*)\|^2 \\ &= \|b - y^*\|^2 - 2\alpha(b - y^*) \cdot (z - y^*) + \alpha^2\|z - y^*\|^2, \end{aligned}$$

  and therefore,

$$(b - y^*) \cdot (z - y^*) \leq \frac{\alpha}{2}\|z - y^*\|^2.$$

- Then let $\alpha \to 0$.

Proof of Proposition 7.1

- Let $y^* \in C$ be as in Lemma 7.4.
- By Lemma 7.2, $(y^* - b) \cdot (y^* - b) > 0$.
- By Lemma 7.3, $(b - y^*) \cdot (z - y^*) \leq 0$ for all $z \in C$.
- Therefore,

  $$(b - y^*) \cdot z \leq (b - y^*) \cdot y^* < (b - y^*) \cdot b$$

  for all $z \in C$.
- Let $p = b - y^*$ and $c = (b - y^*) \cdot y^*$.

Dual Representation of a Convex Set

For $K \subset \mathbb{R}^N$, $K \neq \emptyset$, define the function $\phi_K \colon \mathbb{R}^N \to (-\infty, \infty]$ by

$$\phi_K(p) = \sup_{x \in K} p \cdot x,$$

which is called the *support function* of K.

Proposition 7.5

Let $K \subset \mathbb{R}^N$, $K \neq \emptyset$, be a closed convex set. Then

$$K = \{x \in \mathbb{R}^N \mid p \cdot x \leq \phi_K(p) \text{ for all } p \in \mathbb{R}^N\}.$$

More generally, for any nonempty set K,

$$\mathrm{Cl}(\mathrm{Co}\, K) = \{x \in \mathbb{R}^N \mid p \cdot x \leq \phi_K(p) \text{ for all } p \in \mathbb{R}^N\}.$$

Proof

- $K \subset$ (RHS): By definition.
- $K \supset$ (RHS):

  Let $b \notin K$.
- Since K is closed and convex, by the Separating Hyperplane Theorem, there exist $\bar{p} \neq 0$ and $c \in \mathbb{R}$ such that

  $$\bar{p} \cdot z \leq c < \bar{p} \cdot b \text{ for all } z \in K,$$

  and hence

  $$\phi_K(\bar{p}) = \sup_{z \in K} \bar{p} \cdot z < \bar{p} \cdot b.$$
- This means that $b \notin$ (RHS).

Supporting Hyperplane Theorem

Proposition 7.6 (Supporting Hyperplane Theorem)

Suppose that $C \subset \mathbb{R}^N$, $C \neq \emptyset$, is convex, and that $b \notin \operatorname{Int} C$. Then there exists $p \in \mathbb{R}^N$ with $p \neq 0$ such that

$$p \cdot y \leq p \cdot b \text{ for all } y \in C.$$

For proof, we will use the following fact:

Fact 1

For any convex set $C \subset \mathbb{R}^N$, $\operatorname{Int} C = \operatorname{Int}(\operatorname{Cl} C)$.

The equality does not hold in general for nonconvex sets;
for example, $[0, 1/2) \cup (1/2, 1]$.

Proof

- Let $b \notin \operatorname{Int} C$.

  Since C is convex, $b \notin \operatorname{Int}(\operatorname{Cl} C)$ by Fact 1.

- Therefore, there is a sequence $\{b^m\}$ with $b^m \notin \operatorname{Cl} C$ such that $b^m \to b$.

- Since C is convex, $\operatorname{Cl} C$ is also convex (Proposition 4.12).

- Then by the Separating Hyperplane Theorem, for each m there exists $p^m \in \mathbb{R}^N$ with $p^m \neq 0$ such that

  $$p^m \cdot y < p^m \cdot b^m \text{ for all } y \in C.$$

- Without loss of generality we assume that $\|p^m\| = 1$ for all m.

- $\{p^m\}$ has a convergent subsequence $\{p^{m_k}\}$ with a limit p, where $p \neq 0$ since $\|p\| = 1$.

- Letting $k \to \infty$ we have $p \cdot y \leq p \cdot b$ for all $y \in C$.

Separating Hyperplane Theorem

Proposition 7.7 (Separating Hyperplane Theorem)

Suppose that $A, B \subset \mathbb{R}^N$, $A, B \neq \emptyset$, *are convex, and that* $A \cap B = \emptyset$.

Then there exists $p \in \mathbb{R}^N$ *with* $p \neq 0$ *such that*

$$p \cdot x \leq p \cdot y \text{ for all } x \in A \text{ and } y \in B.$$

Proof

- Since A and B are convex, $A - B = \{x - y \in \mathbb{R}^N \mid x \in A,\ y \in B\}$ is also convex (Proposition 4.5).
- Since $A \cap B = \emptyset$, $0 \notin A - B$.
- Thus by the Supporting Hyperplane Theorem, there exists $p \in \mathbb{R}^N$ with $p \neq 0$ such that

  $$p \cdot z \leq p \cdot 0 \text{ for all } z \in A - B,$$

  or

  $$p \cdot x \leq p \cdot y \text{ for all } x \in A \text{ and } y \in B.$$

Separating Hyperplane Theorem

Proposition 7.8 (Strong Separating Hyperplane Theorem)

Suppose that $A, B \subset \mathbb{R}^N$, $A, B \neq \emptyset$, are convex and closed, and that $A \cap B = \emptyset$.
If A or B is bounded, then there exist $p \in \mathbb{R}^N$ with $p \neq 0$ and $c_1, c_2 \in \mathbb{R}$ such that

$$p \cdot x \leq c_1 < c_2 \leq p \cdot y \text{ for all } x \in A \text{ and } y \in B.$$

Proof

- Since A and B are convex, $A - B$ is also convex.
- Since A and B are closed and A or B is bounded, $A - B$ is closed. ($\rightarrow$ Homework)
- Since $A \cap B = \emptyset$, $0 \notin A - B$.
- Thus by the Separating Hyperplane Theorem, there exist $p \in \mathbb{R}^N$ with $p \neq 0$ and $c \in \mathbb{R}$ such that

  $$p \cdot z \leq c < p \cdot 0 \text{ for all } z \in A - B,$$

  or

  $$p \cdot (x - y) \leq c < 0 \text{ for all } x \in A \text{ and } y \in B.$$

- Thus we have

  $$\sup_{x \in A} p \cdot x - \inf_{y \in B} p \cdot y \leq c < 0.$$

  Let $c_1 = \sup_{x \in A} p \cdot x$ and $c_2 = \inf_{y \in B} p \cdot y$, where $c_1 < c_2$.

Separation with Nonnegative/Positive Vectors

Lemma 7.9

For $A \subset \mathbb{R}^N$, $A \neq \emptyset$, suppose that $A - \mathbb{R}^N_{++} \subset A$.
For $p \in \mathbb{R}^N$, if there exists $c \in \mathbb{R}$ such that $p \cdot x \leq c$ for all $x \in A$, then $p \geq 0$.

Proof

- Assume that $p_n < 0$.
- Fix any $x^0 \in A$ and any $\varepsilon > 0$.

  We have $x^0 - (te_n + \varepsilon\mathbf{1}) \in A - \mathbb{R}^N_{++} \subset A$ for all $t > 0$, while $p \cdot [x^0 - (te_n + \varepsilon\mathbf{1})] = p \cdot x^0 - tp_n - \varepsilon p \cdot \mathbf{1} \to \infty$ as $t \to \infty$, contradicting the assumption that $p \cdot x \leq c$ for all $x \in A$.

Separation with Nonnegative/Positive Vectors

Proposition 7.10

Suppose that $C \subset \mathbb{R}^N$, $C \neq \emptyset$, *is convex.*
If $C \cap \mathbb{R}^N_{++} = \emptyset$, *then there exists* $p \geq 0$ *with* $p \neq 0$ *such that*

$$p \cdot x \leq 0 \text{ for all } x \in C.$$

Proof

- Let $A = C - \mathbb{R}^N_{++}$.
- Since C and $\mathbb{R}^N_{++}$ are convex, A is also convex.
- Since $C \cap \mathbb{R}^N_{++} = \emptyset$, $0 \notin A$.
- Thus by the Supporting Hyperplane Theorem, there exists $p \in \mathbb{R}^N$ with $p \neq 0$ such that

  $$p \cdot z \leq p \cdot 0 \text{ for all } z \in A.$$

- Since $A - \mathbb{R}^N_{++} \subset A$, we have $p \geq 0$ by Lemma 7.9.
- We have

  $$p \cdot x \leq p \cdot y \text{ for all } x \in C \text{ and } y \in \mathbb{R}^N_{++}.$$

  Letting $y \to 0$, we have $p \cdot x \leq 0$ for all $x \in C$.

Separation with Nonnegative/Positive Vectors

Proposition 7.11

Suppose that $C \subset \mathbb{R}^N$, $C \neq \emptyset$, is convex and closed.
If $C \cap \mathbb{R}^N_+ = \{0\}$, then there exist $p \gg 0$ and $c \geq 0$ such that

$p \cdot x \leq c$ *for all* $x \in C$.

Proof

- Let $\Delta = \{x \in \mathbb{R}^N_+ \mid x_1 + \cdots + x_N = 1\}$.
- C is convex and closed and Δ is convex and compact.
- Since $C \cap \mathbb{R}^N_+ = \{0\}$, $C \cap \Delta = \emptyset$.
- Thus by Proposition 7.8, there exist $p \in \mathbb{R}^N$ with $p \neq 0$ and $c \in \mathbb{R}$ such that

  $$p \cdot x \leq c < p \cdot y \text{ for all } x \in C \text{ and } y \in \Delta,$$

  where $c \geq 0$ since $0 \in C$.
- For each n, since $e_n \in \Delta$, we have $0 \leq c < p \cdot e_n = p_n$.

Efficient Production

Let $Y \subset \mathbb{R}^N$ be the production set of a firm.

Definition 7.1

- A production vector $y \in Y$ is *efficient* if there is no $y' \in Y$ such that $y' \geq y$ and $y' \neq y$.
- $y \in Y$ is *weakly efficient* if there is no $y' \in Y$ such that $y' \gg y$.

- y: efficient $\Rightarrow$ y: weakly efficient

Proposition 7.12

Suppose that Y is convex.
Then for any weakly efficient production vector $\bar{y} \in Y$, there exists $p \geq 0$ with $p \neq 0$ such that

$$p \cdot \bar{y} \geq p \cdot y \text{ for all } y \in Y.$$

Proof

- Let $\bar{y} \in Y$ be weakly efficient.
- Then $(Y - \{\bar{y}\}) \cap \mathbb{R}^N_{++} = \emptyset$, where $Y - \{\bar{y}\}$ is convex.
- Thus by Proposition 7.10, there exists $p \geq 0$ with $p \neq 0$ such that $p \cdot z \leq 0$ for all $z \in Y - \{\bar{y}\}$, or $p \cdot y \leq p \cdot \bar{y}$ for all $y \in Y$.

From Profit Function to Production Set

- Let $Y \subset \mathbb{R}^N$, $Y \neq \emptyset$, be the production set of a firm, and let $\phi_Y \colon \mathbb{R}^N \to (-\infty, \infty]$ be the support function of Y:

$$\phi_Y(p) = \sup_{y \in Y} p \cdot y.$$

- Suppose that Y is convex and closed.

  Then, as we have seen,

$$Y = \{y \in \mathbb{R}^N \mid p \cdot y \leq \phi_Y(p) \text{ for all } p \in \mathbb{R}^N\}.$$

- What additional assumptions are needed to recover Y from the profit function, which is defined only for nonnegative, or positive, price vectors (where we allow the profit function to take values in $(-\infty, \infty]$)?

- *Free disposal*: $Y - \mathbb{R}^N_+ \subset Y$.
- *No free production*: $Y \cap \mathbb{R}^N_+ \subset \{0\}$.
- *The ability to shut down*: $0 \in Y$.

Proposition 7.13

1. *If Y is nonempty, convex, and closed and satisfies free disposal, then*

$$Y = \{y \in \mathbb{R}^N \mid p \cdot y \leq \phi_Y(p) \text{ for all } p \in \mathbb{R}^N_+\}.$$

2. *If Y is nonempty, convex, and closed and satisfies free disposal, no free production, and the ability to shut down, then*

$$Y = \{y \in \mathbb{R}^N \mid p \cdot y \leq \phi_Y(p) \text{ for all } p \in \mathbb{R}^N_{++}\}.$$

Proof

1

- $Y \subset$ (RHS): Immediate.
- $Y^{\mathrm{c}} \subset$ (RHS)$^{\mathrm{c}}$: Suppose that $\bar{y} \notin Y$.
- Since Y is nonempty, convex, and closed, there exist $\bar{p} \neq 0$ and c such that

  $$\bar{p} \cdot y \leq c < \bar{p} \cdot \bar{y} \text{ for all } y \in Y,$$

  and hence $\phi_Y(\bar{p}) < \bar{p} \cdot \bar{y}$, by the Separating Hyperplane Theorem.
- Since Y satisfies free disposal, i.e., $Y - \mathbb{R}^N_+ \subset Y$ (which implies $Y - \mathbb{R}^N_{++} \subset Y$), we have $\bar{p} \geq 0$ by Lemma 7.9.
- Hence, $\bar{y} \notin$ (RHS).

Proof

2

- $Y \subset$ (RHS): Immediate.
- $Y^c \subset (\text{RHS})^c$: Suppose that $\bar{y} \notin Y$.
- Since Y is nonempty, convex, and closed and satisfies free disposal, there exist $p^1 \neq 0$ with $p^1 \geq 0$ and c_1 such that

$$p^1 \cdot y \leq c_1 < p^1 \cdot \bar{y} \text{ for all } y \in Y.$$

- Since $Y \cap \mathbb{R}^N_+ = \{0\}$ by no free production and the ability to shut down, by Proposition 7.11 there exist $p^2 \gg 0$ and c_2 such that

$$p^2 \cdot y \leq c_2 \text{ for all } y \in Y.$$

- Let $\varepsilon > 0$ be small enough that $c_1 + \varepsilon c_2 < p^1 \cdot \bar{y} + \varepsilon p^2 \cdot \bar{y}$.

- Then we have

$$(p^1 + \varepsilon p^2) \cdot y \leq c_1 + \varepsilon c_2 < (p^1 + \varepsilon p^2) \cdot \bar{y} \text{ for all } y \in Y,$$

  and hence, $\phi_Y(p^1 + \varepsilon p^2) < (p^1 + \varepsilon p^2) \cdot \bar{y}$. where $p^1 + \varepsilon p^2 \gg 0$.
- Hence, $\bar{y} \notin$ (RHS).

Subgradients and Subdifferentials

Let $X \subset \mathbb{R}^N$ be a non-empty convex set.

Definition 7.2

For a function $f\colon X \to \mathbb{R}$ and $\bar{x} \in X$, if

$$f(x) \leq f(\bar{x}) + p \cdot (x - \bar{x})$$

holds for all $x \in X$, then

- $p \in \mathbb{R}^N$ is called a *subgradient* of f at $\bar{x}$,
- the set of all subgradients of f at $\bar{x}$, denoted by $\partial f(\bar{x})$, is called the *subdifferential* of f at $\bar{x}$, and
- the correspondence $x \mapsto \partial f(x)$ is called the subdifferential of f.

(Usually a subgradient is defined to be p that satisfies the converse inequality, and sometimes p that satisfies the above inequality is called a *supergradient*.)

Subgradients and Subdifferentials

Let $X \subset \mathbb{R}^N$ be a non-empty convex set.

Proposition 7.14

Suppose that $f\colon X \to \mathbb{R}$ is concave.
If $\bar{x} \in \operatorname{Int} X$ and f is differentiable at $\bar{x}$, then $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$.

Proposition 7.15

Suppose that $f\colon X \to \mathbb{R}$ is concave.
Then $\partial f(\bar{x}) \neq \emptyset$ for all $\bar{x} \in \operatorname{Int} X$.

Fact 2

Suppose that $f\colon X \to \mathbb{R}$ is concave.
If $\partial f(\bar{x}) = \{p\}$, then f is differentiable at $\bar{x}$ (and $p = \nabla f(\bar{x})$).

Proof of Proposition 7.15

- Let $f\colon X \to \mathbb{R}$ be a concave function, and let $\bar{x} \in \operatorname{Int} X$.
- $\operatorname{hyp} f$ is convex by the concavity of f.
- We also have $(\bar{x}, f(\bar{x})) \notin \operatorname{Int}(\operatorname{hyp} f)$.
- Thus by the Supporting Hyperplane Theorem, there exists $(p, q) \in \mathbb{R}^N \times \mathbb{R}$ with $(p, q) \neq (0, 0)$ such that

$$p \cdot x + qy \geq p \cdot \bar{x} + q(f(\bar{x})) \text{ for all } (x, y) \in \operatorname{hyp} f.$$

- We must have $q < 0$:
  - If $q > 0$, as $y \to -\infty$ the inequality would be violated.
  - If $q = 0$, we would have $p \neq 0$ and $p \cdot x \geq p \cdot \bar{x}$ for all $x \in X$, where $\bar{x} \in \operatorname{Int} X$.

    Letting $x = \bar{x} - \varepsilon p$ for sufficiently small $\varepsilon > 0$ leads to a contradiction.

  So that we may let $q = -1$.

- Therefore, we in particular have

$$p \cdot x - f(x) \geq p \cdot \bar{x} - f(\bar{x}) \text{ for all } x \in X,$$

or

$$f(x) \leq f(\bar{x}) + p \cdot (x - \bar{x}) \text{ for all } x \in X,$$

which means that $p \in \partial f(\bar{x})$.