Minkowski's separating hyperplane theorem

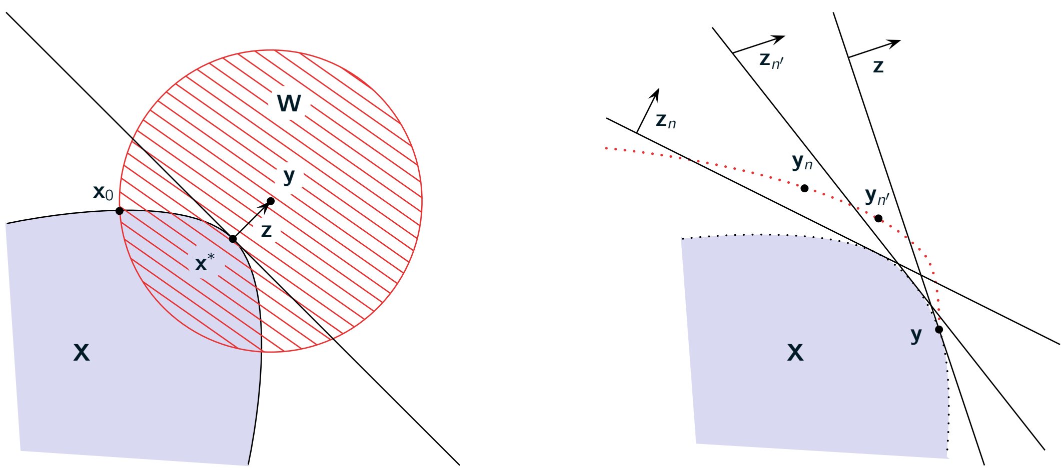
Bruno Salcedo[1]

Supplementary material for Knowledge, belief and rationality
Pennsylvania State University · New Economic School · Higher School of Economics
Summer 2013

Theorem 1 *For every pair of convex non-empty sets $\mathbf{X}, \mathbf{Y} \subseteq \mathbb{R}^n$ with disjoint interiors, there exists a non-null vector $\mathbf{z} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ and a scalar $\mu \in \mathbb{R}$ such that:*

$$\sup\{\mathbf{z} \cdot \mathbf{y} \mid \mathbf{y} \in \mathbf{Y}\} \leq \mu \leq \inf\{\mathbf{z} \cdot \mathbf{x} \mid \mathbf{x} \in \mathbf{X}\}$$

The theorem follows from the two following lemmas.



Lemma 1 Lemma 2

Figure (1) Proof of Minkowski's Separating Hyperplane Theorem

Lemma 1 *For every non-empty, closed and convex set $\mathbf{X} \subseteq \mathbb{R}^n$ and every point $\mathbf{y} \in \mathbb{R}^L \backslash \mathbf{X}$, there exists a non-null vector $\mathbf{z} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ such that $\mathbf{z} \cdot \mathbf{y} > \sup\{\mathbf{z} \cdot \mathbf{x} \mid \mathbf{x} \in \mathbf{X}\}$.*

Proof. Let $\mathbf{X} \subseteq \mathbb{R}^n$ be non-empty and convex and fix a point $\mathbf{y} \in \mathbb{R}^n \backslash \mathbf{X}$. Since $\mathbf{X} \neq \emptyset$ we can fix some arbitrary point $\mathbf{x}_0 \in \mathbf{X}$. Consider the compact non-empty set $\mathbf{W} = \{\mathbf{w} \in \mathbb{R}^n \mid \|\mathbf{y} - \mathbf{w}\|_2 \leq \|\mathbf{x}_0 - \mathbf{y}\|_2\}$ and let $\mathbf{X}' = \mathbf{X} \cap \mathbf{W}$. In words, $\mathbf{X}'$ is the set of points in $\mathbf{X}$ that are as close to $\mathbf{y}$ as $\mathbf{x}_0$ according to the Euclidean metric. Clearly we have $\mathbf{x}_0 \in \mathbf{X}'$ and thus $\mathbf{X}'$ is a compact non-empty set. Hence, by continuity of the Euclidean norm and Weierstrass' theorem, we know that there exists some point $\mathbf{x}^* \in \arg\min\{\|\mathbf{y} - \mathbf{x}\|_2 \mid \mathbf{x} \in \mathbf{X}'\}$. By construction, $\mathbf{x}^*$ is the point of $\mathbf{X}$ which is closest to $\mathbf{y}$ according to the Euclidean metric. Let $\mathbf{z} = \mathbf{y} - \mathbf{x}^* \in \mathbb{R}^n$, since $\mathbf{y} \notin \mathbf{X}$ we know that $\mathbf{z} \neq \mathbf{0}$.

[1]bruno@psu.edu, brunosalcedo.com

Let $\mathbf{x} \in \mathbf{X}$ be an arbitrary point and let $\mathbf{v} : (0,1) \to \mathbb{R}^n$ be the function given by $\mathbf{v}_\mu = \mu\mathbf{x} + (1-\mu)\mathbf{x}^*$. By convexity of $\mathbf{X}$ we know that $\mathbf{v} \in \mathbf{X}^{(0,1)}$. Thus, by definition of $\mathbf{x}^*$, $0 < \|\mathbf{y} - \mathbf{x}^*\|_2 \leq \|\mathbf{y} - \mathbf{v}_\mu\|_2$ for all $\mu \in [0,1]$ and:

$$\begin{aligned}
0 &\geq \|\mathbf{y} - \mathbf{x}^*\|_2^2 - \|\mathbf{y} - \mathbf{v}_\mu\|_2^2 \\
&= \mathbf{z} \cdot \mathbf{z} - \big(\mathbf{y} - (\mu\mathbf{x} + (1-\mu)\mathbf{x}^*)\big) \cdot \big(\mathbf{y} - (\mu\mathbf{x} + (1-\mu)\mathbf{x}^*)\big) \\
&= \mathbf{z} \cdot \mathbf{z} - \big(\mathbf{z} + \mu(\mathbf{x}^* - \mathbf{x})\big) \cdot \big(\mathbf{z} + \mu(\mathbf{x}^* - \mathbf{x})\big) \\
&= \mathbf{z} \cdot \mathbf{z} - \mathbf{z} \cdot \mathbf{z} - 2\mu\mathbf{z} \cdot (\mathbf{x}^* - \mathbf{x}) - \mu^2(\mathbf{x}^* - \mathbf{x}) \cdot (\mathbf{x}^* - \mathbf{x}) \\
&= -2\mu\mathbf{z} \cdot (\mathbf{x}^* - \mathbf{x}) - \mu^2\|\mathbf{x}^* - \mathbf{x}\|_2
\end{aligned}$$

This implies that:

$$\mathbf{z} \cdot (\mathbf{x}^* - \mathbf{x}) \geq -\frac{1}{2}\mu\|\mathbf{x}^* - \mathbf{x}\|_2 \xrightarrow[\mu \to 0]{} 0$$

Hence $\mathbf{z} \cdot \mathbf{x}^* \geq \mathbf{z} \cdot \mathbf{x}$. Since $\mathbf{x}$ was arbitrary, this implies $\mathbf{z} \cdot \mathbf{x}^* = \max\{\mathbf{z} \cdot \mathbf{x} \mid \mathbf{x} \in \mathbf{X}\}$. Finally, notice that $0 < \mathbf{z} \cdot \mathbf{z} = \mathbf{z} \cdot (\mathbf{y} - \mathbf{x}^*)$ and thus $\mathbf{z} \cdot \mathbf{y} > \mathbf{z} \cdot \mathbf{x}^*$. ■

Lemma 2 *For every non-empty and convex set $\mathbf{X} \subseteq \mathbb{R}^n$ and every point $\mathbf{y}^*$ in the boundary of $\mathbf{X}$, there exists a non-null vector $\mathbf{z} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ such that $\mathbf{z} \cdot \mathbf{y} \geq \sup\{\mathbf{z} \cdot \mathbf{x} \mid \mathbf{x} \in \mathbf{X}\}$.*

Proof. Let $\mathbf{X} \subseteq \mathbb{R}^n$ be a non-empty convex set and consider a boundary point $\mathbf{y} \in \mathrm{fro}(\mathbf{X})$. By definition of boundary, there exists a sequence $(\mathbf{y}_n)_{n \in \mathbb{N}} \in (\mathrm{int}(\mathbb{R}^n \backslash \mathrm{cl}(\mathbf{X})))^{\mathbb{N}}$ such that $\lim \mathbf{y}_n = \mathbf{y}$. By Lemma 1, there exists a sequence $(\mathbf{w}_n)_{n \in \mathbb{N}} \in (\mathbb{R}^n \backslash \{\mathbf{0}\})^{\mathbb{N}}$, such that $\mathbf{w}_n \cdot \mathbf{y}_n > \sup\{\mathbf{w}_n \cdot \mathbf{x} \mid \mathbf{x} \in \mathrm{cl}(\mathbf{X})\}$ for all $n \in \mathbb{N}$. For every $n \in \mathbb{N}$ we have $\mathbf{w}_n \neq \mathbf{0}$ and, consequently $\|\mathbf{w}_n\|_2 > 0$. We can thus define the sequence $(\mathbf{z}_n)_{n \in \mathbb{N}} \in \mathbf{B}^{\mathbb{N}}$ given by $\mathbf{z}_n = \mathbf{w}_n / \|\mathbf{w}_n\|_2$, where $\mathbf{B} = \{\mathbf{v} \in \mathbb{R}^n \mid \|\mathbf{v}\|_2 = 1\}$ is the unit circle in $\mathbb{R}^n$. This transformation preserves the inequalities $\mathbf{z}_n \cdot \mathbf{y}_n > \mathbf{z}_n \cdot \mathbf{x}$ for all $\mathbf{x} \in \mathrm{cl}(\mathbf{X})$ and all $n \in \mathbb{N}$. Since $\mathbf{B}$ is compact, we know that $(\mathbf{z}_n)_{n \in \mathbb{N}}$ has a convergent subsequence converging to some limit $\mathbf{z} \in \mathbf{B}$. Since weak inequalities are preserve under limits of linear functions, we have $\mathbf{z} \cdot \mathbf{y} \geq \mathbf{z} \cdot \mathbf{x}$ for all $\mathbf{x} \in \mathrm{cl}(\mathbf{X})$. Consequently, since $\mathbf{X} \subseteq \mathrm{cl}(\mathbf{X})$, we have $\mathbf{z} \cdot \mathbf{y} \geq \sup\{\mathbf{z} \cdot \mathbf{x} \mid \mathbf{x} \in \mathbf{X}\}$. ■

Let $\mathbf{X}, \mathbf{Y} \subseteq \mathbb{R}^n$ be convex and have disjoint interiors and let $\mathbf{W} = \mathrm{int}(\mathbf{X}) - \mathrm{int}(\mathbf{Y}) \subseteq \mathbb{R}^n$. Since $\mathrm{int}(\mathbf{Y}) \cap \mathrm{int}(\mathbf{X}) = \emptyset$, we know that $\mathbf{0} \notin \mathbf{W}$. Simple algebra shows that $\mathbf{W}$ is convex. From the previous lemmas it follows that there exists some $\mathbf{z} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$ such that $0 = \mathbf{z} \cdot \mathbf{0} \geq \mathbf{z} \cdot (\mathbf{x} - \mathbf{y})$ for all $\mathbf{x} \in \mathbf{X}$ and all $\mathbf{y} \in \mathbf{Y}$ (if $0 \in \mathrm{fro}(\mathbf{W})$ use Lemma 1, otherwise use Lemma 2). Which implies that $\mathbf{z} \cdot \mathbf{x} \geq \mathbf{z} \cdot \mathbf{y}$ for all $\mathbf{x} \in \mathbf{X}$ and all $\mathbf{y} \in \mathbf{Y}$, and hence $\sup\{\mathbf{z} \cdot \mathbf{y} \mid \mathbf{y} \in \mathbf{Y}\} \leq \inf\{\mathbf{z} \cdot \mathbf{x} \mid \mathbf{x} \in \mathbf{X}\}$.

QED

Ü///