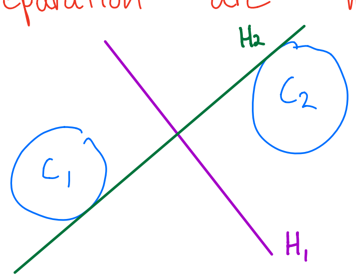
Section II (Separation Theorems)

Part III of this text is concerned w/ duality correspondences, Rockafellar states that separation theorems are the foundations for duality correspondences. Almost everything that follows in this chapter is based on the fact that a hyperplane in $\mathbb{R}^n$ divides $\mathbb{R}^n$ "evenly in two."

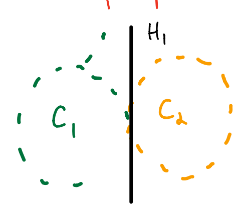
Definition

- A hyperplane H is said to **separate** C_1 & C_2 if C_1 is contained in one of the closed half-spaces assoc. w/ H and C_2 lies in the opposite closed half-space
- H separates C_1 & C_2 **properly** if C_1 & C_2 are not both contained in H itself.
- H separates C_1 & C_2 **strongly** if $\exists$ $\varepsilon > 0$ s.t. $C_1 + \varepsilon B$ is contained in one of the assoc. open half-spaces & $C_2 + \varepsilon B$ is contained in the opposite open half space.
- Other kinds of separation are considered in general, e.g. strict separation (C_1 & C_2 must simply belong to opposing open h.s.), however proper & strong separation are most useful for our purposes.



H_1 properly, strictly, and strongly separates C_1 & C_2

H_2 only properly separates C_1 & C_2



Assume $C_1 \cap C_2 = \emptyset$ but $cl(C_1) \cap cl(C_2) \neq \emptyset$
H_1 properly & strictly separates C_1 & C_2 but not strongly.

Theorem 11.1

Let $C_1, C_2 \in \mathbb{R}^n$ be non-empty sets. $\exists$ a hyperplane separating C_1 & C_2 properly iff $\exists$ a vector b s.t.

(i) (a) $\inf\{\langle x,b\rangle \mid x \in C_1\} \geq \sup\{\langle x,b\rangle \mid x \in C_2\}$,

(b) $\sup\{\langle x,b\rangle \mid x \in C_1\} > \inf\{\langle x,b\rangle \mid x \in C_2\}$.

Additionally, $\exists$ a hyperplane separating C_1 & C_2 strongly iff $\exists$ a vector b s.t.

(ii) (c) $\inf\{\langle x,b\rangle \mid x \in C_1\} > \sup\{\langle x,b\rangle \mid x \in C_2\}$.

Proof

(i) ($\Leftarrow$) Suppose $\exists$ b satisfying (a) & (b) & let β be s.t.

$$\inf_{x \in C_1} \langle x,b\rangle \geq \beta \geq \sup_{x \in C_2} \langle x,b\rangle.$$

$b \neq 0$ & $\beta \in \mathbb{R} \Rightarrow H = \{x \mid \langle x,b\rangle = \beta\}$ is a hyperplane (Thm 1.3).

Then $C_1 \subset \{x \mid \langle x,b\rangle \geq \beta\}$ & $C_2 \subset \{x \mid \langle x,b\rangle \leq \beta\}$.

Condition (b) implies C_1 & C_2 are not both contained in H, i.e., H separates C_1 & C_2 properly. ✓✓

($\Rightarrow$) Assume C_1 & C_2 can be separated properly, then there exists b & β s.t. $H = \{x \mid \langle x,b\rangle = \beta\}$ & $x \in C_1 \Rightarrow \langle x,b\rangle \geq \beta$ & $x \in C_2 \Rightarrow \langle x,b\rangle \leq \beta$ w/ one one of these inequalities strict for at least one $x \in C_1$ or $x \in C_2$, therefore b satisfies (a) & (b). $\square$

(ii) ($\Leftarrow$) If b satisfies (c) we can choose $\beta \in \mathbb{R}$ & $\delta > 0$ s.t. $\langle x,b\rangle \geq \beta + \delta$ $\forall x \in C_1$ & $\langle x,b\rangle \leq \beta - \delta$ $\forall x \in C_2$.

Then $\exists$ $\varepsilon > 0$ such that $|\langle y,b\rangle| < \delta$ $\forall y \in \varepsilon B$. Let $x \in C_1$ & $y \in \varepsilon B$ then

$\langle x+y, b\rangle = \langle x,b\rangle + \langle y,b\rangle > \beta+\delta-\delta = \beta \Rightarrow C_1+\epsilon B \subset \{x \mid \langle x,b\rangle > \beta\}$.

Similarly, we have $C_2+\epsilon B \subset \{x \mid \langle x,b\rangle < \beta\}$. Therefore, $H = \{x \mid \langle x,b\rangle = \beta\}$ separates C_1 & C_2 strongly. ✓✓

($\Rightarrow$) Assume C_1 & C_2 can be separated strongly then $\exists$ b & β s.t

$$\inf_{x\in C_1} \langle x,b\rangle > \inf_{\substack{x\in C_1\\ y\in B}} \langle x,b\rangle + \epsilon\langle y,b\rangle \geq \beta \geq \sup_{\substack{x\in C_2\\ y\in B}} \langle x,b\rangle + \epsilon\langle y,b\rangle > \sup_{x\in C_2} \langle x,b\rangle. \quad \square$$

Now we turn to the existence question which is: given two sets can we find a separating hyperplane?

Theorem 11.2

Let $C\in\mathbb{R}^n$ be a non-empty relatively open cvx set, and let $M\in\mathbb{R}^n$ be a non-empty affine set not meeting C, i.e., $M\cap C=\emptyset$. Then $\exists$ a hyperplane H containing M, s.t. one of the open half-spaces assoc. w/ H contains C.

Proof

If M is a hyperplane (i.e., M is n-1 dimensional) then we are done, M separates C otherwise $M\cap C \neq \emptyset$. Therefore, WLOG assume M is not a hyperplane.

Set up

1) we will construct an affine set M' one dimension higher than M s.t. again $M'\cap C=\emptyset$.
2) In n steps or less this will determine a hyperplane H ($H\cap C=\emptyset$) which will prove the theorem.

WLOG assume $0 \in M$, so that M is a subspace. $0 \notin C$ then $C \subset (C-M)$ but $0 \notin C-M$ $(M \cap C = \emptyset)$. M is not a hyperplane $\Rightarrow$ the subspace $M^\perp$ contains a two-dimensional subspace P. Define $C' = P \cap (C-M)$, C' is a relatively open cvx set in P (Cor 6.5.1: $ri(P \cap (C-M)) = P \cap ri(C-M)$ & Cor 6.6.2: $ri(C-M) = ri(C) - ri(M) = C-M$) & $0 \notin C$. Now we want to find a line L through 0 in P not meeting C' then $M' = M + L$ is one dimension higher than M & $M' \cap C = \emptyset$ (if $(M+L) \cap C \neq \emptyset$ then $L \cap (C-M) \neq \emptyset$, which contradicts $L \cap C' = \emptyset$ by construction). WLOG identify P w/ $\mathbb{R}^2$.

Cases

(i) C' is empty or zero dimensional then finding L is trivial.

(ii) aff C' is a line not containing 0, take L to be the parallel line through 0.

(iii) aff C' is a line containing 0, take L to be the perpendicular line through 0.

(iv) C' is two-dimensional and hence open, then $K = \bigcup_{\lambda > 0} \lambda C'$ is the smallest convex cone containing C' (Cor 2.6.3). Note that K is open and does not contain 0. Therefore K is an open sector of $\mathbb{R}^2$ corresponding to an angle no greater than π. Take L to be the line extending one of the two boundary rays of the sector. $\square$

<u>Theorem 11.3</u> (Main Separation Theorem)

Let $C_1, C_2 \in \mathbb{R}^n$ be non-empty cvx sets. In order that there exists a hyperplane separating C_1 & C_2 ~~properly,~~ it is necessary and sufficient that $\operatorname{ri} C_1$ & $\operatorname{ri} C_2$ have no point in common.

<u>Proof</u>

Consider the cvx set $C = C_1 - C_2$, then $\operatorname{ri} C = \operatorname{ri} C_1 - \operatorname{ri} C_2$ (Thm 6.6.2) so $0 \notin \operatorname{ri} C$ iff $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 = \phi$. Since $0 \notin \operatorname{ri} C$ Thm 11.2 guarantees the existence of a hyperplane containing $M = \{0\}$ s.t. $\operatorname{ri} C$ is contained in one of the assoc. h.s., now $C \subset \operatorname{cl}(\operatorname{ri} C)$ implies C is contained in the closure of the half-space. Therefore if $0 \notin \operatorname{ri}(C)$ $\exists$ b s.t.

$$0 \leq \inf_{x \in C} \langle x, b \rangle = \inf_{x_1 \in C_1} \langle x_1, b \rangle - \sup_{x_2 \in C_2} \langle x_2, b \rangle$$

$$0 < \sup_{x \in C} \langle x, b \rangle = \sup_{x_1 \in C_1} \langle x_1, b \rangle - \inf_{x_2 \in C_2} \langle x_2, b \rangle$$

Now Thm 11.1 implies C_1 & C_2 can be separated properly. Conditions in 11.1 imply $0 \notin \operatorname{ri}(C)$ since they assert the existence of a h.s. $D = \{x \mid \langle x, b \rangle \geq 0\}$ containing C where $\operatorname{ri} D \cap C \neq 0 \Rightarrow \operatorname{ri} C \subset \operatorname{ri} D$ (Cor 6.5.2). $\square$

<u>Example</u> Why does 11.3 only give proper separation?

Consider

$$C_1 = \{(\xi_1, \xi_2) \mid \xi_1 > 0, \xi_2 \geq \xi^{-1}\}, \quad \& \quad C_2 = \{(\xi_1, 0) \mid \xi_1 \geq 0\}.$$

Then $C_1 \cap C_2 = \phi$ but the only separating hyperplane is the ξ_1-axis which contains all of C_2. That is, C_1 & C_2 are properly separated but NOT strongly separated.

Theorem 11.4

Let $C_1, C_2 \in \mathbb{R}$ be non-empty cvx sets. In order that there exists a hyperplane separating C_1 & C_2 strongly, it is necessary & sufficient that

$$\inf\{\|x_1 - x_2\| \mid x_1 \in C_1,\ x_2 \in C_2\} > 0,$$

or in other words that $0 \notin \mathrm{cl}(C_1 - C_2)$.

Proof

By defn of strong separation $\exists$ $\varepsilon > 0$ s.t. $(C_1 + \varepsilon B) \cap (C_2 + \varepsilon B) = \phi$. Considering $(C_1 + \varepsilon B) \cap (C_1 + \varepsilon B) = \phi$ Thm 11.3 implies $C_1 + \varepsilon B$ & $C_1 + \varepsilon B$ can be separated properly, i.e., $C_1 + \frac{\varepsilon}{2}B + \frac{\varepsilon}{2}B$ & $C_2 + \frac{\varepsilon}{2}B + \frac{\varepsilon}{2}B$ belong to opposite (closed) h.s. so that $C_1 + \frac{\varepsilon}{2}B$ & $C_2 + \frac{\varepsilon}{2}B$ belong to opposite open h.s.. That is C_1 & C_2 can be separated strongly iff for some $\varepsilon > 0$ the origin is NOT contained in

$$(C_1 + \varepsilon B) - (C_2 + \varepsilon B) = C_1 - C_2 - 2\varepsilon B$$

i.e. $2\varepsilon B \cap C_1 - C_2 = \emptyset$ for some $\varepsilon > 0$ which says $0 \notin \mathrm{cl}(C_1 - C_2)$. □

Skipping Corollary 11.4.1 b/c it depends on recession.

Corollary 11.4.2

Let $C_1, C_2 \in \mathbb{R}^n$ be non-empty cvx sets whose closures are disjoint. If either set is bounded $\exists$ a hyperplane separating C_1 & C_2 strongly.

Proof

See book, depends on recession & Corollary 11.4.1.

Theorem 11.5 (updated in 18.8)

A closed cvx set C is the intersection of the closed half-spaces which contain it.

Proof

Assume $\emptyset \neq C \neq \mathbb{R}^n$ (otherwise it is trivial). Consider $a \notin C$, then $C_1 = \{a\}$ & $C_2 = C$ satisfy 11.4 implying $\exists$ hyperplane H separating $\{a\}$ & C strongly. One of the closed h.s. assoc. w/ H contains C but not $\{a\}$, therefore the intersection of all closed half spaces containing C contains no points other than those in C. $\square$

Corollary 11.5.1

Let S be any subset of $\mathbb{R}^n$. The $\mathrm{cl}(\mathrm{conv}\, S)$ is the intersection of all the closed half spaces containing S.

Corollary 11.5.2

Let C be a cvx subset of $\mathbb{R}^n$ other than $\mathbb{R}^n$ itself. Then $\exists$ a closed half-space containing C, i.e., there exists some $b \in \mathbb{R}^n$ s.t. the linear fct $\langle \cdot, b \rangle$ is bounded above on C.

Definition

- A <u>supporting half-space</u> to a cvx set $C \in \mathbb{R}^n$ is a closed h.s. containing C & has a point of C in it's boundary
- A <u>supporting hyperplane</u> to a cvx set C is a hyperplane which is the boundary of supporting h.s. to C.

Supporting hyperplanes are $H = \{x \mid \langle x, b\rangle = \beta\}$, $b \neq 0$, w/ $\langle x, b\rangle \leq \beta$ $\forall x \in C$ & $\exists$ $x \in C$ s.t. $\langle x, b\rangle = \beta$, ie, a supporting hyperplane is assoc. w/ a linear fct which achieves it's maximum on C. We consider non-trivial supporting hyperplanes to C, which are supporting hyperplanes to C that do not contain C itself.

Theorem 11.6

Let C be a convex set, & let D be a non-empty cvx subset of C (e.g. a single point). In order that there exists a non-trivial supporting hyperplane to C containing D, it is necessary and sufficient that D be disjoint from $\operatorname{ri} C$.

Proof

$D \subset C \Rightarrow$ the non-trivial supporting hyperplanes (H) to C s.t. $D \subset H$ are hyperplanes separating C & D properly. Such H exist iff $\operatorname{ri} C \cap \operatorname{ri} D = \emptyset$ (Thm 11.3). Assume $D \cap \operatorname{ri} C \neq \emptyset$ then Cor 6.5.2 implies that $\operatorname{ri} D \subset \operatorname{ri} C$ which is a contradiction. Therefore we have $\operatorname{ri} D \cap \operatorname{ri} C = \emptyset$ is equivalent to $D \cap \operatorname{ri} C = \emptyset$. $\square$

Corollary 11.6.1

A cvx set has a non-zero nomal at each of it's bdry pts.

Corollary 11.6.2

Let C be a cvx set. $x \in C$ is a relative bdry pt iff there exists a linear fct h not constant on C s.t. h achieves it's maximum over C at x.