Appendix A: Separation theorems in $I\!R^n$

These notes provide a number of separation theorems for convex sets in $I\!R^n$. We start with a basic result, give a proof with the help on an auxiliary result and then continue with a number of more refined results.

Theorem A.1. *Let V be a linear subspace of $I\!R^n$ and $\emptyset \neq K \subseteq I\!R^n$ be convex and compact. If K and V are disjoint, $K \cap V = \emptyset$, then K and V can be strictly separated by a hyperplane: There exists a linear mapping $f : I\!R^n \to I\!R$ with*

$$f(x) > 0, \qquad \forall x \in K,$$

$$f(x) = 0, \qquad \forall x \in V.$$

In particular, $f \not\equiv 0$.

Before we can start proving this, we need an auxiliary result.

Proposition A.2. *Let $\emptyset \neq C \subseteq I\!R^n$ be convex and closed with $0 \notin C$. Then there exist a linear mapping $f : I\!R^n \to I\!R$ and $\alpha > 0$ with*

$$f(x) \geq \alpha, \qquad \forall x \in C,$$

i.e. C and the hyperplane $\{f = 0\} \subseteq I\!R^n$ are disjoint. In particular, $f \not\equiv 0$.

Proof. Choose $r > 0$ with $\overline{U_r(0)} \cap C \neq \emptyset$. Because C is closed, this intersection is compact, and so the continuous function $x \mapsto |x|$ has a minimum over $x \in \overline{U_r(0)} \cap C$ in some point x_0. This x_0 is the projection of the point 0 on C, and it minimises of course $|x|$ also over all of C. (We only take the intersection $\overline{U_r(0)} \cap C$ to have a compact set in order to argue the existence of a minimiser.) For $x \in C$ and $\alpha \in [0, 1]$, we have $\alpha x + (1 - \alpha)x_0 = \alpha(x - x_0) + x_0 \in C$ because C is convex, and therefore

$$|x_0|^2 \leq |x_0 + \alpha(x - x_0)|^2 = |x_0|^2 + 2\alpha(x - x_0) \cdot x_0 + \alpha^2 |x - x_0|^2.$$

Because this holds for all $\alpha \in [0,1]$, we must have

$$0 \leq (x - x_0) \cdot x_0 = x_0 \cdot x - |x_0|^2,$$

and so it is enough to take $f(x) := x_0 \cdot x$ and $\alpha := |x_0|^2$ (which is > 0 since $x_0 \in C$ cannot be 0). **q.e.d.**

Proof of Theorem A.1. Take the algebraic difference

$$\mathcal{C} := K - V := \{c = k - v \mid k \in K, v \in V\}.$$

This is convex like K, and closed because K is compact. Indeed, if $c_n = k_n - v_n \to c$, there exists a subsequence still denoted by (k_n) with $k_n \to k$ for some $k \in K$; so (v_n) also converges to some limit v which is in V because V is closed (this uses that the linear subspace V is finite-dimensional), and so $c = k - v \in C$. Finally, $K \cap V = \emptyset$ implies that $0 \notin C$, and so Proposition A.2 yields the existence of a linear $f : I\!R^n \to I\!R$ and $\alpha > 0$ with $f(x) \geq 2\alpha$ for all $x \in C$. Thus we have, using linearity, that

$$f(k) - f(v) \geq 2\alpha > \alpha \qquad \text{for all } k \in K \text{ and all } v \in V.$$

Taking $k_0 \in K$ fixed and λv instead of v with $\lambda \in I\!R$ yields $f(v) = 0$ for all $v \in V$, and hence also $f(k) > \alpha$ for all $k \in K$. **q.e.d.**

Remark. The same argument still works in almost the same way if we replace the linear subspace V by a convex cone. More precisely, we can use the above argument with all $\lambda \geq 0$, and the conclusion is then only that $f(v) \leq 0$ for all $v \in V$.

A slightly different version of the separation result is as follows.

Theorem A.3. *Let $U \subseteq I\!R^n$ be convex and closed, and $\emptyset \neq K \subseteq I\!R^n$ convex and compact. If $K \cap U = \emptyset$, then there exist $\beta \in I\!R$ and a linear mapping $f : I\!R^n \to I\!R$ with*

$$f(x) > \beta, \qquad \forall x \in K,$$

$$f(x) \leq \beta, \qquad \forall x \in U.$$

In particular, we have again $f \not\equiv 0$.

Proof. As in the proof of Theorem A.1, the set $C := K - U$ is convex and closed and $0 \notin C$; so there exist $\alpha > 0$ and $f : I\!R^n \to I\!R$ linear with $f(x) \geq \alpha$ for all $x \in C$, or

$$f(k) \geq \alpha + f(u) \qquad \text{for all } k \in K \text{ and all } u \in U.$$

Because f is continuous and K is compact, f attains a minimum γ in some point $k_0 \in K$; so we obtain

$$f(k) \geq \gamma = f(k_0) \geq \alpha + f(u) \qquad \text{for all } u \in U,$$

and

$$\delta := \sup_{u \in U} f(u) \leq \gamma - \alpha < \infty.$$

For $\beta := \delta + \frac{\alpha}{2}$, we therefore get

$$f(k) \geq \gamma \geq \alpha + \delta > \beta \qquad \text{for all } k \in K$$

because $\alpha > 0$, and

$$f(u) \leq \delta < \beta \qquad \text{for all } u \in U$$

again because $\alpha > 0$. **q.e.d.**

Before the formulation of the next result, we recall the definition of the *interior* B° of a set $B \subseteq I\!R^n$:

$$B^\circ := \{x \in B \mid U_\varepsilon(x) \subseteq B \text{ for some } \varepsilon > 0\}.$$

Theorem A.4. *Let $A \subseteq I\!R^n$ be convex and $B \subseteq I\!R^n$ convex with $B^\circ \neq \emptyset$. If $A \cap B^\circ = \emptyset$, then there exist $\beta \in I\!R$ and a linear mapping $f : I\!R^n \to I\!R$, $f \not\equiv 0$, with*

$$f(x) \geq \beta, \qquad \forall x \in A,$$

$$f(x) \leq \beta, \qquad \forall x \in B.$$

Proof. 1) Suppose first in addition that A and B are both closed. Take $x_0 \in B^\circ \neq \emptyset$ and set $B_m := \frac{1}{m}x_0 + (1 - \frac{1}{m})B = \{x = \frac{1}{m}x_0 + (1 - \frac{1}{m})b \mid b \in B\}$. Like B, the set B_m is convex and closed, and $x_0 \in B^\circ$ implies that $B_m \subseteq B^\circ$ by Lemma A.5 below. Therefore $K_m := B_m \cap \overline{U_m(x_0)}$ is convex and compact, nonempty because $x_0 \in K_m$, and $K_m \cap A = \emptyset$ due to $A \cap B^\circ = \emptyset$ by assumption. Since A is closed, Theorem A.3 therefore gives the existence of $\beta_m \in I\!R$ and a linear $f_m : I\!R^n \to I\!R$ with

$$f_m(x) > \beta_m, \qquad \forall x \in K_m,$$

$$f_m(x) \leq \beta_m, \qquad \forall x \in A.$$

If we write f_m as $f_m(x) = c_m \cdot x$ with $c_m \in I\!R^n$, then $f_m \not\equiv 0$ yields $c_m \neq 0$ and we can assume by scaling that $|c_m| = 1$. Because $x_0 \in K_m$ for all m, we moreover obtain for fixed $y_0 \in A$ that

$$c_m \cdot y_0 = f_m(y_0) \leq \beta_m < f_m(x_0) = c_m \cdot x_0,$$

and so the sequence (β_m) is bounded.

Now choose a convergent subsequence $(c_{m'}, \beta_{m'})$ with limit (c, β) and set $f(x) := c \cdot x$. Then we obtain for $x \in A$ that

$$f(x) \leq c \cdot x \leq \beta.$$

For $x \in B^\circ$, we have $x \in K_{m'}$ for m' large enough; indeed, if we define b by the requirement that $x = \frac{1}{m'}x_0 + (1 - \frac{1}{m'})b$, solving for b gives $b = (1 + \frac{1}{m'-1})x - \frac{1}{m'-1}x_0$ so that $|b - x| < \delta$ for m' large and hence $b \in U_\delta(x) \subseteq B$. Therefore $c_{m'} \cdot x > \beta_{m'}$ for m' large enough, this yields

$$f(x) = c \cdot x \geq \beta, \qquad \forall x \in B^\circ,$$

and continuity then gives this for $x \in B$ as well. Finally, $|c| = 1$ implies that $f \not\equiv 0$.

2) In general, A and B are convex with $B^\circ \neq \emptyset$. Then the closures $\bar{A}$ and $\bar{B}$ of A and B are also convex, and we have $(\bar{B})^\circ \supseteq B^\circ \neq \emptyset$. If we also have $\bar{A} \cap (\bar{B})^\circ = \emptyset$, the argument in Step 1) shows that one can separate with f and β the sets $\bar{A}$ and $\bar{B}$, and hence of course

also the smaller sets A and B. So it only remains to argue that we do have $\bar{A} \cap (\bar{B})^\circ = \emptyset$, and we now show that this follows from $A \cap B^\circ = \emptyset$.

First, because $B^\circ \neq \emptyset$, we have $(\bar{B})^\circ = B^\circ$. Indeed, the inclusion "$\supseteq$" is obvious. For the converse, if $x \in (\bar{B})^\circ$, there is $\varepsilon > 0$ with $U_\varepsilon(x) \subseteq \bar{B}$ and $x \in \bar{B}$. Lemma B.5 below shows that for any $b_0 \in B^\circ$ and any $y \in \bar{B}$, the point $b_0 + \lambda(y - b_0) = \lambda y + (1-\lambda)b_0$ is still in B° for any $\lambda \in [0,1)$. So if we take $b_0 \in B^\circ \neq \emptyset$ and choose $y := b_0 + (1+\eta)(x - b_0)$, then $|y - x| = \eta|x - b_0| < \varepsilon$ for $\eta > 0$ small enough implies that $y \in U_\varepsilon(x) \subseteq \bar{B}$ and therefore $x \in B^\circ$. This gives the inclusion "$\subseteq$".

Now suppose that $\bar{A} \cap (\bar{B})^\circ \neq \emptyset$. Then $(\bar{B})^\circ = B^\circ$ implies that there exists some $x \in \bar{A} \cap B^\circ$, and so there exist $\varepsilon > 0$ with $U_\varepsilon(x) \subseteq B$ and a sequence $(x_n) \subseteq A$ with $x_n \to x$. So for sufficiently large n and small $\delta > 0$, we have $x_n \in U_\delta(x)$ and $U_\delta(x_n) \subseteq B$, so that $x_n \in A \cap B^\circ$. But this contradicts our assumption that $A \cap B^\circ = \emptyset$, and so we have indeed $\bar{A} \cap (\bar{B})^\circ = \emptyset$. This completes the proof. **q.e.d.**

Finally, we provide the auxiliary result used in the proof of Theorem A.4.

Lemma A.5. *Suppose that $B \subseteq I\!R^n$ is convex with $B^\circ \neq \emptyset$. If $b_0 \in B^\circ$, then for every $x \in \bar{B}$ (the closure of B), the entire "interval"*

$$b_0 + (x - b_0)[0,1) := \{y = \lambda x + (1-\lambda)b_0 \mid 0 \le \lambda < 1\}$$

is still contained in B°.

Proof. If $b_0 \in B^\circ$, then there exists $\varepsilon > 0$ with $U_\varepsilon(b_0) \subseteq B$. Take $x \in \bar{B}$, fix $\lambda \in [0,1)$ and set $y = \lambda x + (1-\lambda)b_0$. Define $\delta := \frac{1-\lambda}{\lambda}\frac{\varepsilon}{2}$ and note that because $x \in \bar{B}$, there exists $z \in U_\delta(x)$ with $z \in B$. If we then set $b' := b_0 + \frac{\lambda}{1-\lambda}(x - z)$, then $b' \in U_{\frac{\varepsilon}{2}}(b_0) \subseteq B$ and $y = \lambda z + (1-\lambda)b'$ with $b' \in B^\circ$, $z \in B$ and $U_{\frac{\varepsilon}{2}}(b') \subseteq U_\varepsilon(b_0) \subseteq B$. But $U_{\frac{\varepsilon}{2}}(b') \subseteq B$ and $z \in B$ together imply that $U_{(1-\lambda)\frac{\varepsilon}{2}}(y) \subseteq B$, because $|v - y| < (1-\lambda)\frac{\varepsilon}{2}$ implies that $|\frac{v}{1-\lambda} - \frac{\lambda}{1-\lambda}z - b'| < \frac{\varepsilon}{2}$, hence $\frac{v}{1-\lambda} - \frac{\lambda}{1-\lambda}z =: w \in U_{\frac{\varepsilon}{2}}(b') \subseteq B$ and therefore $v = \lambda z + (1-\lambda)w \in B$. So we obtain $y \in B^\circ$, and this completes the proof. **q.e.d.**