SEPARATION OF CONVEX SETS

We know from finite dimensional geometry that disjoint convex sets can be separated by planes. In what follows, I follow closely the exposition in the book of H. Brezis, 'Analyse fonctionelle'.

Definition 0.1. *Let X be a real normed vector space and $f : X \to \mathbb{R}$ be a linear functional, not necessarily continuous. The set*

$$H = \{x \in X : f(x) = \alpha\}$$

is called a hyperplane in X.

We have the simple

Proposition 0.2. *The hyperplane H is closed if an only if f is continuous.*

Proof. Suppose that H is closed. Then the complement of H in X, H^c is open. Pick any $x_0 \in H^c$. We may assume that $f(x_0) > \alpha$. There exists an open ball $B(x_0, r) \subset H^c$. Suppose that there exists $x_1 \in B(x_0, r)$ with $f(x_1) < \alpha$. Then for all $0 \le t \le 1$, $x_t = (1-t)x_0 + tx_1 \in B(x_0, r)$ and hence $f(x_t) \ne \alpha$. Pick

$$s = \frac{f(x_0) - \alpha}{f(x_0) - f(x_1)}$$

which, by assumption, is in $(0, 1)$. But $f(x_s) = \alpha$ which contradicts the fact that $x_s \in H^c$. Hence we have that $f(x) > \alpha$ for all $x \in B(x_0, r)$. Hence for all $z \in B(0, 1)$ $f(x_0 - rz) > \alpha$ or

$$f(z) < \frac{f(x_0) - \alpha}{r} \ .$$

Hence

$$\|f\| \le \frac{f(x_0) - \alpha}{r} \ .$$

The converse is evident. □

Definition 0.3. *let $A, B \subset X$ be two sets. We say that the hyperplane H separates the set A and B if*

$$f(x) \le \alpha \ , x \in A \ , f(x) \ge \alpha \ , \ x \in B \ .$$

We say that H separates A, B strictly if there exists $\varepsilon > 0$ such that

$$f(x) \le \alpha - \varepsilon \ , x \in A \ , f(x) \ge \alpha + \varepsilon \ , \ x \in B \ .$$

The following lemma is useful an is usually attributed to Hermann Minkowski.

Lemma 0.4. *Let C be a non-empty, open and convex set. Assume further that $0 \in C$. For $x \in X$ set*

$$p(x) = \inf\{t > 0 : \frac{x}{t} \in C\} \ .$$

Then

$$a) \ p(\lambda x) = \lambda p(x) \ , \lambda > 0 \ , x \in X \ ,$$
$$b) \ p(x + y) \le p(x) + p(y) \ , x, y \in X \ ,$$

$$c)\ C = \{x \in X : p(x) < 1\}\ .$$

Moreover, there exists a positive constant K such that

$$p(x) \leq K\|x\|\ .$$

Proof. The statement a) is obvious. To prove c), denote the set on the right side in c) by C'. Obviously

$$p(x) \leq 1\ ,\ x \in C\ .$$

If $x \in C$ then $(1+\varepsilon)x \in C$ for some $\varepsilon > 0$ since C is open and convex. Hence $p(x) \leq \frac{1}{1+\varepsilon} < 1$. Thus, $C \subset C'$. Conversely, if $x \in C'$, then $p(x) < 1$. There exists $0 < \alpha < 1$ so that $\frac{x}{\alpha} \in C$. But

$$x = \alpha\frac{x}{\alpha} + (1-\alpha)0 \in C$$

because C is convex. Hence, $C = C'$. To prove b) pick any$x, y \in X$. Then

$$\frac{x}{p(x)+\varepsilon}, \frac{y}{p(y)+\varepsilon} \in C$$

because of c). Set

$$s = \frac{p(y)+\varepsilon}{p(x)+p(y)+2\varepsilon}$$

and note that $0 \leq s \leq 1$. Since C is convex

$$\frac{x+y}{p(x)+p(y)+2\varepsilon} = (1-s)\frac{x}{p(x)+\varepsilon} + s\frac{y}{p(y)+\varepsilon} \in C\ .$$

Hence, using c), $p(\frac{x+y}{p(x)+p(y)+2\varepsilon}) < 1$ and hence

$$p(x+y) < p(x) + p(y) + 2\varepsilon$$

for any $\varepsilon > 0$. For the last point, since C is open, there exists $r >$ so that the open ball $B(0,r) \subset C$. Hence, by c), $p(x) < 1$ for all $x \in B(0,r)$. Hence for all $x \in X$, $p(x) \leq \frac{\|x\|}{r}$. ☐

Remark 0.5. *The function $p(x)$ defines in the previous lemma is often called the* **Minkowski gauge for** *C.*

We first prove a separation theorem for a convex set and a point.

Lemma 0.6. *Let C be a non-empty open convex set and let $x_0 \in X$ with $x_0 \notin C$. There exists a bounded linear functional such that $f(x) < f(x_0)$ for all $x \in C$, i.e., the hyperplane H defined by $f = f(x_0)$ separates x_0 from C.*

Proof. By shifting the set we may assume that $0 \in C$. Define the subspace

$$E = \{tx_0 : t \in \mathbb{R}\}\ .$$

On this subspace we have the linear functional

$$g(tx_0) = tp(x_0)\ ,$$

where p is the Minkowski gauge for C. Note that $p(x_0) \geq 1$ since $x_0 \notin C$. We have that $g(x) \leq p(x)$ for all $x \in E$. If $t > 0$ then $g(tx_0) = tp(x_0) = p(tx_0)$. If $t \leq 0$ then $g(tx_0) = -p(|t|x_0) \leq 0 \leq p(tx_0)$. By the Hahn-Banach theorem we may extend g to a linear functional f on X with the property that $f(x) \leq p(x)$. Because $p(x) \leq K\|x\|$, we have that

$$|f(x)| \leq K\|x\|\ , x \in X\ .$$

Hence f is bounded. Further, $f(x_0) = g(x_0) = p(x_0) \geq 1$ and $f(x) \leq p(x) < 1$ for all $x \in C$. □

Theorem 0.7. *Let $A, B \subset X$ be two non-empty disjoint convex sets. Assume that A is open. Then there exists a closed hyperplane H that separates A and B.*

Proof. Set $C = \{x \in X : x = y - z, y \in A, z \in B\}$. It is easily seen that the set C is convex. Since

$$C = \cup_{z \in B}(A - z)$$

and A is open so is C. C is not empty and moreover, the origin $0 \notin C$. By the previous lemma there exists a bounded linear functional f on X which separates C and the origin,i.e., $f(x) < 0$ all $x \in C$, or $f(y - z) < 0$ for all $y \in A, z \in B$. Hence we have that $f(y) < f(z)$ for all $y \in A, z \in B$. From this it follows that

$$\sup_{y \in A} f(y) \leq \inf_{z \in B} f(z)$$

and choosing α between these two numbers yields the desired hyperplane. □

Theorem 0.8. *Let A, B be two non-empty convex and disjoint sets. Assume that A is compact and B is closed. Then there exists a closed hyperplane that separates the sets strictly.*

Proof. Consider the set

$$A_\varepsilon = A + B(0, \varepsilon)\ , B_\varepsilon = B + B(0, \varepsilon)\ .$$

Recall that A_ε is the set of all vectors x that can be written as $x = y + z$ where $y \in A$ and $z \in B(0, \varepsilon)$. Both sets A_ε and B_ε are non-empty, open and convex. This is trivial to verify. For ε small enough $A_\varepsilon \cap B_\varepsilon = \emptyset$. If not there would be a sequence $\varepsilon_n \to 0$ and a sequence of points x_n such that $x_n \in A_{\varepsilon_n} \cap B_{\varepsilon_n}$. This means that $x_n = y_n + z_n$ where $y_n \in A$ and $z_n \in B(0, \varepsilon_n)$ and since A is compact there is a convergent subsequence y_{n_k} which converges to some point in $y \in A$. Hence $x_{n_k} \to y \in A$ and since B is closed $y \in B$, a contradiction. By the previous theorem, there exists a closed hyperplane that separates A_ε and B_ε. Hence, there exists a bounded linear functional f and a number α such that

$$f(x + z_1) \leq \alpha \leq f(y + z_2)\ , x \in A, y \in B, z_1, z_2 \in B(0, \varepsilon)\ .$$

Hence

$$f(x) + \varepsilon\|f\| \leq \alpha \leq f(y) - \varepsilon\|f\|\ , x \in A, y \in B\ .$$

□

Remark 0.9. *If the underlying space is finite dimensional, then a stronger statement holds. Assume that A and B are disjoint, and convex, then they can be separated by a hyperplane. No additional assumptions on A and B are needed. This is false if the underlying space is infinite dimensional as the following example again taken from Brezis's book, shows.*

Take X to be ℓ_1, i.e., the space of all summable sequences. Let X be the set of all sequences of the form

$$x_{2n} = 0\ ,$$

i.e.,

$$x = (x_1, 0, x_3, 0, x_5, 0, \dots)$$

where we, of course, assume that

$$\sum_{j=1}^{\infty} |x_{2j-1}| < \infty$$

Consider the set Y given by all sequence that satisfy

$$y_{2n} = \frac{1}{2^n} y_{2n-1} \ ,$$

i.e., the sequences of the form

$$y = (y_1, \frac{y_1}{2}, y_3, \frac{y_3}{2^2}, y_5, \frac{y_5}{2^3}, \dots)$$

where

$$\sum_{j=1}^{\infty} |y_{2j-1}| < \infty \ .$$

It is easy to see that X as well as Y are closed subspaces of ℓ_1. Consider the point in ℓ_1 given by $c_{2n-1} = 0$ and $c_{2n} = \frac{1}{2^n}$, i.e.,

$$c = (0, \frac{1}{2}, 0, \frac{1}{2^2}, 0, \frac{1}{2^3}, 0, \frac{1}{2^4}, \dots) \ .$$

This point is neither in X nor in Y. However, it is in the closure of the sum, $\overline{X+Y}$, in fact we have

$$\overline{X+Y} = \ell_1 \ .$$

To prove this consider the set S of all sequences $z \in \ell_1$ that have only finitely many non-zero elements. We know that $\overline{S} = \ell_1$. However, every element in S can be written as a sum of points in X and in Y. We simply have to consider the system of equations

$$z_j = x_j + y_j \ , j = 1, 2, 3, \dots \ ,$$

i.e.,

$$z_{2n-1} = x_{2n-1} + y_{2n-1} \ , z_{2n} = \frac{y_{2n-1}}{2^n} \ , n = 1, 2, \dots \ ,$$

which is solved by setting

$$y_{2n-1} = 2^n z_{2n} \ , \ x_{2n-1} = z_{2n-1} - 2^n z_{2n} \ .$$

Because the sequence z consists of zeros for all but finitely many terms,

$$\sum_j |y_{2j-1}| < \infty \ , \sum_j |x_{2j-1}| < \infty \ .$$

Hence, $\overline{X+Y} = \ell_1$. Note that $c \notin X + Y$ since this means that

$$y_{2n-1} = 1 \ , y_{2n} = \frac{1}{2^n} \ , x = 0 \ .$$

The sequence y_j is not summable. This example shows that, in general, it is not true that whenever X, Y are closed subspaces then $X + Y$ is closed. Note that this carries over to ℓ_p $1 < p < \infty$. Now define the set

$$A = X - c \ , B = Y \ .$$

Both sets are closed and disjoint, for if $d \in A \cap B$ then

$$x - c = d = y$$

for some $x \in X$ and $y \in Y$ and hence $c = x - y$ which is not true. Hence, we have two disjoint closed sets A, B. The set B is linear and the set A is affine. Moreover, $\overline{A + B} = \ell_1$. That the two sets cannot be separated by a closed hyperplane hinges now on the following lemma.

Lemma 0.10. *Let B be a linear space and f a linear functional. Suppose that there exists $\alpha \in \mathbb{R}$ such that $f(y) \geq \alpha$. Then $f(y) = 0$ for all $y \in B$.*

Proof. Clearly $0 = f(0) \geq \alpha$ implies that $\alpha \leq 0$. Further for any $\lambda > 0$ and $y \in B$, $f(\lambda y) \geq \alpha$ and hence $f(y) \geq \alpha/\lambda$. Since λ may be any positive number we have that $f(y) \geq 0$ for all $y \in B$. Hence, we may choose $\alpha = 0$. The relation $-f(y) = f(-y) \geq 0$ implies that $f(y) \leq 0$ for all $y \in B$ and hence $f(y) = 0$ for all $y \in B$.

□

Suppose that there exists a bounded linear functional f and $\alpha \in \mathbb{R}$ with

$$f(x) \leq \alpha, x \in A \ , f(y) \geq \alpha, y \in B \ .$$

From the previous lemma we may choose $\alpha = 0$ and we also have that $f(y) = 0, y \in B$ and $f(x) \leq 0, x \in A$. Pick any $z \in \ell_1$. There exists a sequence of points $x_n \in A$ and $y_n \in B$ so that $z = \lim_{n\to\infty}(x_n + y_n)$. Hence, since f is bounded we have that

$$f(z) = \lim_{n\to\infty} f(x_n + y_n) = \lim_{n\to\infty}[f(x_n) + f(y_n)] \leq 0 \ .$$

By replacing z by $-z$ we find that $f(z) = 0$ and since z is arbitrary, f must be the zero functional. Hence A and B cannot be separated.