HAHN-BANACH THEOREM

1. The Hahn-Banach theorem and extensions

We assume that the real vector space X has a function $p : X \mapsto \mathbb{R}$ such that

$$p(ax) = ap(x), \ a \geq 0, \quad p(x+y) \leq p(x) + p(y). \tag{1.1}$$

Theorem 1.1 (Hahn-Banach)**.** *Let Y be a subspace of X on which there is ℓ such that*

$$\ell y \leq p(y) \qquad \forall y \in Y, \tag{1.2}$$

where p is given by (1.1)*. Then ℓ can be extended to all X in such a way that the above relation holds, i.e. there is a $\tilde{\ell} : X \mapsto \mathbb{R}$ such that*

$$\tilde{\ell}|_Y = \ell, \quad \ell x \leq p(x) \quad \forall x \in X. \tag{1.3}$$

Proof. The proof is done in two steps: we first show how to extend ℓ on a direction x not contained in Y, and then use Zorn's lemma to prove that there is a maximal extension.

If $z \notin Y$, then to extend ℓ to $Y \oplus \{\mathbb{R}x\}$ we have to choose ℓz such that

$$\ell(\alpha z + y) = \alpha \ell x + \ell y \leq p(\alpha z + y) \quad \forall \alpha \in \mathbb{R}, y \in Y.$$

Due to the fact that Y is a subspace, by dividing for α if $\alpha > 0$ and for $-\alpha$ for negative α, we have that ℓx must satisfy

$$\ell y - p(y-x) \leq \ell x \leq p(y+x) - \ell y, \qquad y \in Y.$$

Thus we should choose ℓx by

$$\sup_{y \in Y} \big\{ \ell y - p(y-x) \big\} \leq \ell x \leq \inf_{y \in Y} \big\{ p(y+x) - \ell y \big\}.$$

This is possible since by subadditivity (1.1) for $y, y' \in \mathbb{R}$

$$\ell(y+y') \leq p(y+y') \leq p(y+x) + p(y'-x).$$

Thus we can extend ℓ to $Y \oplus \{\mathbb{R}x\}$ such that $\ell(y+\alpha x) \leq p(y + \alpha x)$.

To apply Zorn's lemma, consider all extensions of ℓ to linear spaces Z containing Y as a subspace and such that $\ell z \leq p(z)$. We give a partial order this set by inclusion, i.e. $(Z, \ell) \leq (Z', \ell')$ if $Z \subset Z'$ and $\ell'|_Z = \ell$.

Clearly every totally ordered subset $\{(Z_\alpha, \ell_\alpha)\}_{\alpha \in A}$ has an upper bound, given by taking the element

$$Z = \bigcup_{\alpha \in A} Z_\alpha, \quad \ell\big|_{Z_\alpha} = \ell_\alpha.$$

By Zorn's lemma there exists a maximal element $\{\tilde{Z}, \tilde{\ell}\}$.

If $\tilde{Z} \subsetneq X$, then there is an element $x \in X \setminus Z$, so that by repeating the analysis of the first part of the proof it follows that $(\tilde{Z}, \tilde{\ell})$ can be extended. This is a contradiction, hence $\tilde{Z} = X$, and $\tilde{\ell} : X \mapsto \mathbb{R}$ satisfies

$$\ell x \leq p(x) \quad \forall x \in X.$$

□

Note that

$$-\ell x \leq p(-x) \quad \Longrightarrow \quad -p(-x) \leq \ell x \leq p(x). \tag{1.4}$$

We now consider two extension of the above theorem. The important part of the theorem is that we can extend a linear functional (this is trivial by using again Zorn's lemma) in such a way that some requirements are respected, namely

$$\ell x \leq p(x).$$

One can thus ask the question if in some cases we can extend a linear functional respecting more conditions.

Theorem 1.2 (Agnew-Morse)**.** *Let X be a real vector space and let $\mathcal{A} = \{\mathbf{A}_\nu\}_\nu$ a collection of commuting linear operators:*

$$\mathbf{A}_\nu : X \mapsto X, \quad \mathbf{A}_\nu \mathbf{A}_\mu = \mathbf{A}_\mu \mathbf{A}_\nu. \tag{1.5}$$

Assume that $p : X \mapsto \mathbb{R}$ is positively homogeneous, subadditive as in (1.1) *and*

$$p(\mathbf{A}_\nu x) = p(x). \tag{1.6}$$

Let Y be a subspace of X, invariant for every $\mathbf{A}_\nu$, and assume that on Y there is linear functional $\ell : Y \mapsto \mathbb{R}$ such that

$$\ell(y) \leq p(y), \quad \ell(\mathbf{A}_\nu y) = \ell y. \tag{1.7}$$

Then ℓ can be extended to the whole X in such a way ℓ is invariant w.r.t. $\mathcal{A}$ and dominated by p:

$$\ell(\mathbf{A}_\nu x) = \ell x, \quad \ell x \leq p(x) \quad \forall \in X. \tag{1.8}$$

Proof. The first step is to extend the family $\mathcal{A}$ is such a way the identity $\mathbf{I}$ belongs to $\mathcal{A}$ and

$$\mathbf{A}, \mathbf{B} \in \mathcal{A} \quad \Longrightarrow \quad \mathbf{AB} \in \mathcal{A}. \tag{1.9}$$

Clearly this extension do not change the assumptions on Y, ℓ, p.

We thus define the new positive homogeneous function $g : X \mapsto \mathbb{R}$ by

$$g(x) = \inf \bigg\{ p(\mathbf{C}x), \quad \mathbf{C} = \sum_i a_i \mathbf{A}_i, \ \sum_i a_i = 1, \ a_i > 0, \mathbf{A}_i \in \mathcal{A} \bigg\}. \tag{1.10}$$

Using basic computations, one can show that g satisfies

$$g(x) \leq p(x), \quad g(\alpha x) = \alpha g(x) \ \alpha > 0, \quad g(x+y) \leq g(x) + g(y). \tag{1.11}$$

The advantage of g w.r.t. p is that g is invariant w.r.t. convex combinations.

If $\mathbf{C}$ is a convex combination of $\mathbf{A}_i \in \mathcal{A}$, then

$$\mathbf{C}Y \subset Y, \quad \ell(\mathbf{C}y) = \ell y \leq p(\mathbf{C}y).$$

Thus on Y we have also that $\ell(y) \leq g(y)$.

We now apply Hahn-Banach theorem to obtain that there is some extension $\tilde{\ell} : X \mapsto \mathbb{R}$ such that

$$\tilde{\ell} x \leq g(x).$$

To prove that from the above inequality it follows that $\tilde{\ell}(\mathbf{A}x) = \tilde{\ell}x$, $\mathbf{A} \in \mathcal{A}$, we consider the particular linear combination

$$\mathbf{C}_n = \frac{1}{n} \sum_{i=0}^n \mathbf{A}^i, \quad \mathbf{C}_n(\mathbf{I} - \mathbf{A}) = \frac{1}{n}(\mathbf{I} - \mathbf{A}^n).$$

By the definition of g it follows

$$g((\mathbf{I} - \mathbf{A})x) \leq p(\mathbf{C}_n(\mathbf{I} - \mathbf{A})x) = \frac{1}{n} p((\mathbf{I} - \mathbf{A}^n)x).$$

We finally have

$$\ell x - \ell(\mathbf{A}x) = \ell((\mathbf{I} - \mathbf{A})x) \leq g((\mathbf{I} - \mathbf{A})x) \leq \frac{1}{n} p(x - \mathbf{A}^n x) \leq \frac{p(x) + p(-x)}{n} \to 0.$$

Computing the same expression for $-x$, we conclude the proof. □

We now extend the theorem to complex valued spaces.

Theorem 1.3. *If X is a complex vector space, and p is subadditive and*

$$p(ax) = |a| p(x), \quad a \in \mathbb{C},$$

and $\ell : Y \mapsto \mathbb{C}$ linear such that $|\ell(y)| \leq p(y)$ for $y \in Y$. Then there is an extension $\tilde{l} : X \mapsto \mathbb{C}$ such that

$$|\ell x| \leq p(x) \quad x \in X. \tag{1.12}$$

Proof. The idea in this proof is to split the functional in two parts, which are real linear functionals.

Let

$$\ell y = \ell_1 y + i\ell_2 y, \quad \ell_1, \ell_2 : Y \mapsto \mathbb{R}.$$

Clearly ℓ_1, ℓ_2 are real, and they satisfy $\ell_1(iy) = \ell_2(y)$.

Conversely, if $\ell_1 : X \mapsto \mathbb{R}$ is a linear functional, then

$$\ell x = \ell_1 x - i\ell_1(ix) \tag{1.13}$$

is a complex valued linear functional.

We can extend ℓ_1 by means of the real Hahn-Banach theorem, and define $\ell : X \mapsto \mathbb{C}$ by (1.13), and we have for some $\alpha = \alpha(x)$ on the unitary disk

$$|\ell(x)| = \ell(\alpha x) = \ell_1(\alpha x) \le p(\alpha x) = p(x).$$

□

2. An example: Banach limits

Let B be the space of all bounded real sequences over $\mathbb{N}$. Define the function

$$p(x) = \limsup_{n\to\infty} x_n. \tag{2.1}$$

One can easily check that p is positively homogeneous and subadditive. Let Y be the subspace of the sequence having limit.

If we consider the left translation

$$\mathbf{A}(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots), \tag{2.2}$$

it is clear that p, Y are invariant w.r.t. $\mathbf{A}$.

We thus conclude that there is an extension $\mathrm{LIM} : B \mapsto \mathbb{R}$ of the operator $\lim : Y \mapsto \mathbb{R}$ such that

(1) if the sequence x converges, then
$$\mathrm{LIM} x = \lim x;$$
(2) LIM is invariant w.r.t. left translation;
(3) it holds
$$\liminf_{n\to\infty} x_n \le \mathrm{LIM} x \le \limsup_{n\to\infty} x_n.$$

The last part follows from (1.4).

3. Applications and geometrical interpretations

3.1. **Extension of linear functionals.** A particular class of subadditive and homogeneous functions p are the seminorms. We recall that if X is a locally convex topological vector space, then there exists a separating family of continuous seminorms $\mathcal{P}$, and conversely if we have a separating family of seminorms $\mathcal{P}$ then we can generate a topology such that every $p \in \mathcal{P}$ is continuous.

The continuity of $p \in \mathcal{P}$ implies in particular that the set $V = \{p(x) < 1\}$ is open, and p coincides with the Minkowski functional μ_V. Conversely, if V is balanced, convex and open, then μ_V is a continuous seminorm and $V = \{\mu_V < 1\}$.

We can restate Theorem 1.3 in terms of seminorms:

Theorem 3.1. *Assume $\ell : M \mapsto \mathbb{C}$ is a linear functional over a subspace of X such that for a seminorm p*

$$|\ell(x)| \le p(x). \tag{3.1}$$

Then ℓ can be extended to the whole X in such a way that (3.1) *holds.*

Clearly the linear functional ℓ satisfying (3.1) is continuous, and conversely if f is continuous then (3.1) holds. We can thus say that ℓ *can be extended in such a way it remains continuous.*

If the space X is normed, in particular, we have the implication (in the sense that there is an extension of ℓ)

$$|\ell x| \le \alpha\|x\| \ \forall x \in M \quad \Longrightarrow \quad |\ell x| \le \alpha\|x\| \ \forall x \in X. \tag{3.2}$$

Moreover, by taking the linear subspace $\mathbb{C}x_0$, with $x_0 \in X$ fixed, and

$$\ell(\alpha x_0) = \alpha \|x_0\|,$$

we have

Corollary 3.2. *For every x_0 fixed in the normed space X, there is a continuous functional ℓ_0 such that*

$$\ell x_0 = \|x_0\|, \quad |\ell x| \leq \|x\| \ \forall x \in X.$$

Given a linear functional $\ell \neq 0$, we define *hyperplane* in X the set

$$H_\alpha = \Big\{ x \in X : \ell x = \alpha \Big\}, \tag{3.3}$$

with α fixed in $\mathbb{C}$. For real vector spaces, the two sets $\{\ell x < c\}$ and $\{\ell x < c\}$, $c \in \mathbb{R}$, are the two half spaces.

We recall that

Proposition 3.3. *The functional ℓ is continuous if and only if H_α is closed.*

Proof. If ℓ is continuous, then it is closed.

Conversely, if H is closed and not equal to X, then there is an open $x_0 + V$, with V balanced convex neighborhood of 0, such that

$$H \cap (x_0 + V) = \emptyset.$$

By translation an replacing ℓx with $\ell(x - x_0)$, we can assume $x_0 = 0$, and thus we have that for all $x \in V$

$$\ell x \neq \alpha.$$

Since V is balanced, an ℓ is linear then we have also for all $\gamma \leq 1$

$$\gamma(\ell x) \neq \alpha.$$

Then we obtain $|\ell x| \leq \alpha$ for all $x \in V$. Thus

$$\frac{|\beta|}{|\alpha|} V \subset \ell^{-1}\{|\beta| < \epsilon\}.$$

□

Remark 3.4. It is important to observe here that the seminorm p does not need to be one of the seminorms generating the topology. For example, we can use the following procedure.

A point $x_0 \in S \subset X$ is called *interior* if for all $y \in X$ there is $t > 0$ such that $x_0 + ty \in S$. If K is a convex set with an interior point let us assume 0, then the *gauge* of K is

$$p_K(x) = \inf\big\{ a : x/a \in K \big\}.$$

Proposition 3.5. *The function $p_K(x)$ is well defined, positively homogeneous and subadditive. Moreover for all $x \in K$, $p_K(x) \leq 1$, and $p_K(x) < 1$ only if x is in the interior of K. The converse is also true, given the function p.*

As an application,

Theorem 3.6. *Let $K \subset X$ be a convex subset such that all points in K are interior. Then for all $y \in X \setminus K$ there is ℓ, c such that*

$$\ell y = c, \quad \ell x < c, \ x \in K.$$

Proof. Assume that 0 is an interior point of K, and take the gauge p_K. Consider the one dimensional space $\mathbb{R}y$, and set $\ell y = 1$. Since $y \notin K$, then $p_K(y) \geq 1$, so that we can extend ℓ to X in such a way that $\ell x \leq p_K(x)$. Thus $\ell|_K < 1$, and by construction $\ell y = 1$. □

In this case, however, one cannot say that the extension of a continuous linear functional is continuous.

3.2. **Geometric versions.** We consider in this section *real* topological vector spaces. As in the extension of Hahn-Banach theorem to complex spaces, if the vector space is complex, in the statement of the next results one has to replace the value of the functional with its real part.

Theorem 3.7. *Let A, B disjoint nonempty convex subset of the topological vector space X, and assume A has not empty interior and B closed. Then there is a continuous functional ℓ such that*

$$\sup_{x\in A}\{\ell x\} \leq \gamma \leq \inf_{x\in B}\{\ell x\}. \tag{3.4}$$

Geometrically we can say that the hyperplane $\{\ell x = \gamma\}$ *separates* the sets A, B. Since we have only the large inequality, we say that ℓ *separates in broad sense.*

Proof. If we fix $a_0 \in \overset{\circ}{A}$, $b_0 \in B$, then the set

$$C = \overset{\circ}{A} - B + (b_0 - a_0)$$

contains an open convex neighborhood of 0. Let μ_C be the Minkowski functional of C, so that

$$C = \{\mu_C < 1\}.$$

Since $A \cap B = \emptyset$, then $x_0 = b_0 - a_0 \notin C$, so that $\mu_C(x_0) \geq 1$.

Define on the one dimensional subspace $\mathbb{R}x_0$ the linear functional

$$\ell(\alpha x) = \alpha \leq \mu_V(\alpha x_0).$$

Hence, by the Hahn-Banach theorem, there is a continuous functional $\ell : X \mapsto \mathbb{R}$ such that

$$|\ell x| \leq \mu_C(x), \quad \ell x_0 = 1.$$

If $a \in \overset{\circ}{A}$, $b \in B$, then we have

$$\ell(a - b) + 1 = \ell(a - b + x_0) < 1,$$

so that

$$\ell a < \ell b, \quad \sup_{a\in \overset{\circ}{A}} \ell a \leq \inf_{b\in B} \ell b.$$

Thus we obtain the theorem, because ℓ is continuous. □

If A is instead compact, and B closed, then we can separate A, B *in the strict sense.*

Theorem 3.8. *Let A, B closed convex sets in X locally convex, and A compact. Then there is a continuous functional ℓ and two constants $\gamma_1, \gamma_2 \in \mathbb{R}$ such that*

$$\ell a \leq \gamma_1 < \gamma_2 < \ell b, \quad a \in A,\ b \in B. \tag{3.5}$$

Proof. We recall that since A is compact and B closed there is a convex balanced neighborhood of 0 such that

$$(A + V) \cap (B + V) = \emptyset.$$

We then use Theorem 3.7 with $A + V$, obtaining that there is a continuous functional $\ell : X \mapsto \mathbb{R}$ such that

$$\sup_{a\in A+V} \ell a = \sup_{a\in A} \ell a + \sup_{x\in V} \ell x \leq \inf_{b\in B} \ell b.$$

If $\ell V = \{0\}$, then since $X = \cup_n nV$ we have $\ell = 0$, which is a contradiction. Thus there is some $\kappa > 0$ such that

$$\sup_{a\in A} \ell a + \frac{k}{2} \leq \inf_{b\in B} \ell b - \frac{\kappa}{2}.$$

The conclusion follows. □

An important and useful application is the following

Corollary 3.9. *If $M \subset X$ is a linear subspace and $x_0 \notin \bar{M}$, then there is a continuous functional ℓ such that*

$$\ell M = \{0\}, \quad \ell x_0 = 1. \tag{3.6}$$

Proof. We use Theorem 3.8 with $A = \{x_0\}$, $B = \bar{M}$. We thus have that there is a continuous functional $\ell : X \mapsto \mathbb{R}$ and a constant $\kappa > 0$ such that

$$\sup_{x \in \bar{M}} \ell x \leq \ell x_0 - \kappa.$$

Since $\bar{M}$ is a linear subspace, then if $\ell M \neq \{0\}$ we have

$$\sup_{x \in \bar{M}} \ell x = +\infty,$$

so that we have

$$\sup_{x \in \bar{M}} \ell x = 0, \quad \ell x_0 \geq \kappa > 0.$$

The functional $(1/\kappa)\ell : X \mapsto \mathbb{R}$ satisfies (3.6). □

4. Dual space of a locally convex topological vector space

We define the dual space X^* of a topological vector space X as

$$X^* = \Big\{ \ell : X \mapsto F, \ell \text{ continuous} \Big\}, \quad F = \mathbb{R} \text{ or } \mathbb{C}. \tag{4.1}$$

The function

$$p_\ell(x) = |\ell x| \tag{4.2}$$

with ℓ continuous is a continuous seminorm on X. Conversely, on X^* the function

$$p_x(\ell) = |\ell x| \tag{4.3}$$

is a seminorm on X^*.

By construction, if $\ell \neq 0$, then there is some $x \in X$ such that $\ell x \neq 0$, so that the family of seminorms

$$\mathcal{P}^* = \Big\{ p_x : x \in X \Big\} \tag{4.4}$$

is separating, hence generating a topology which makes X^* a locally convex topological vector space.

Conversely, in general the family

$$\mathcal{P} = \Big\{ p_\ell : \ell \in X^* \Big\} \tag{4.5}$$

is not separating, as for example in L^p, $0 < p < 1$.

For locally convex topological vector spaces X, we have the following result

Theorem 4.1. *If X is locally convex, then $\mathcal{P}$ is separating.*

Proof. We need only to apply Theorem 3.8 to the sets $\{0\}$ and $\{x\}$ for all $x \in X$ fixed. □

Thus the family of continuous seminorm $\mathcal{P}$ generates a topology, which in general is weaker than the original topology of X.

5. Exercises

(1) Prove that the Banach limit can be extended in such a way it is invariant w.r.t. the Cesaro means operator, i.e.
$$\mathbf{C}(x_1, x_2, \dots) = \bigg(x_1, \frac{x_1 + x_2}{2}, \frac{x_1 + x_2 + x_3}{3}, \dots \bigg).$$
(2) In the space $X = \{f : [0,1] \mapsto \mathbb{C}, f \text{ uniformly bounded}\}$ with the product topology, consider $K = \{|f| \leq 1\}$.
   - Show that K is convex and absorbing, but the gauge of K is not continuous.
   - Deduce that the functionals constructed in Theorem 3.6 are not continuous.

(3) Prove that if E is a Banach space, $H = \{\ell x = \alpha\}$ an hyperplane, either H is closed or dense in E.

(4) Using the fact that the dual of

$$\ell_2 = \Big\{ a = (a_n)_{n\in\mathbb{N}} : \|a\|_2^2 = \sum_{n=1}^{\infty} |a_n|^2 < \infty \Big\}$$

is again ℓ_2, consider the two subspaces

$$L_1 = \mathbb{C}\begin{pmatrix} 1 & 0 & 0 & \dots \end{pmatrix},$$

$$L_2 = \operatorname{span}\Big\{ \begin{pmatrix} 1 & 2 & 0 & 0 & \dots \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 & 0 & \dots \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 2 & \dots \end{pmatrix}, \dots \Big\}.$$

Show that $L_1 \oplus L_2$ is dense in ℓ_2.

(5) Consider the operator $T : \ell_2 \mapsto \ell_2$ defined by

$$T\Big(\{a_n\}_{n\in\mathbb{N}}\Big) = \Big\{ \frac{a_n}{n} \Big\}_{n\in\mathbb{N}}.$$

Show that $T\ell_2$ is dense in ℓ_2 but it is not all ℓ_2, and it has empty interior.

(6) Consider the Banach space

$$\ell_1 = \Big\{ a = \{a_n\}_{n\in\mathbb{N}} : \|a\|_1 = \sum_{n=1}^{\infty} |a_n| < \infty \Big\},$$

and the two subspaces

$$X = \Big\{ a \in \ell_1 : a_{2n} = 0 \ \forall n \geq 1 \Big\}, \quad Y = \Big\{ a \in \ell_1 : a_{2n} = \frac{a_{2n-1}}{2^n} \ \forall n \geq 1 \Big\}.$$

- Show that X, Y are closed, and that $\overline{X \oplus Y} = \ell_1$.
- Find $c \in \ell_1$ such that $c \notin X \oplus Y$.
- Let now $Z = X - c$, so that $Z \cap Y = \emptyset$. Deduce that there are not continuous hyperplane separating Z and Y.

(7) If X is a normed space, define the *duality map* as

$$x \mapsto F(x) = \Big\{ \ell \in X^* \colon \|\ell\|_{X^*} = \|x\|_X, \ell x = \|x\|_X^2 \Big\}. \tag{5.1}$$

Use Hahn-Banach theorem to show that $F(x) \neq \emptyset$.

(8) A normed space is strictly convex if

$$x \neq y \quad \Longrightarrow \quad \Big\| \frac{x+y}{2} \Big\|_X < \frac{1}{2}\big(\|x\| + \|y\|\big). \tag{5.2}$$

Show that is this case $F(x) \subset X^*$ reduces to a single point.

(9) Consider the Banach space

$$X = \Big\{ u \in C([0,1], \mathbb{R}) \colon u(0) = 0 \Big\}, \quad \|u\|_X = \max_{t\in[0,1]} |u(t)|.$$

and the linear functional

$$\ell u = \int_0^1 u(t)dt.$$

Show that $\|\ell\|_{X^*} = 1$, bur there are no $u \in X$ such that $\|u\|_X = 1$ and $\ell u = 1$.