Hahn-Banach Theorems

©*1995, Paul Garrett, garrett@@math.umn.edu*
version January 27, 1998

- Continuous linear functionals
- Dominated Extension Theorem for *real* spaces
- Separation Theorem for *real* spaces
- Complex scalars
- Corollaries

The first real point here is (roughly) that convex sets can be separated by linear functionals. The second is that continuous linear functionals on subspaces of *locally convex* topological vectorspaces have continuous extensions to the whole space.

Things are proven first, most naturally, for *real* vectorspaces. The complex-linear versions are simply corollaries.

One crucial corollary is that on locally convex topological vectorspaces the continuous linear functionals *separate points*, meaning that for $x \neq y$ there is a continuous linear functional λ so that $\lambda(x) \neq \lambda(y)$. This separation property is essential in a variety of applications, so the hypothesis of local convexity becomes indispensable.

1 Continuous Linear Functionals

Let k be either **R** or **C**, and let V be a k-vectorspace, without any assumptions about topologies for the moment. Then a k-linear k-valued function on V is called a **linear functional**.

If V *has* a topology, and if a linear functional is *continuous*, then reasonably enough we call it a **continuous linear functional**. The space of all continuous linear functionals on V is denoted by V^*, suppressing reference to the field k.

A linear functional λ on V is said to be *bounded* if there is some neighborhood U of 0 in V and constant c so that $|\lambda x| \leq c$ for $x \in U$, where $|\ |$ is the usual absolute value on k. The following proposition is the general topological vectorspace analogue of the corresponding assertion for Banach spaces, in which the 'boundedness' has a different sense.

Proposition: The following three conditions on a linear functional λ on a topological vectorspace V over k are equivalent:

- λ is continuous.
- λ is continuous at 0.
- λ is bounded.

Proof: The first assertion certainly implies the second. Assume the second. Then, given $\varepsilon > 0$, there is a neighborhood U of 0 so that $|\lambda|$ is bounded by ε on U. This proves boundedness. Finally, suppose that $|\lambda(x)| \leq c$ on a neighborhood U of 0. Then given $x \in V$ and given $\varepsilon > 0$, we *claim* that for

$$y \in x + \frac{\varepsilon}{2c} U$$

we have

$$|\lambda(x) - \lambda(y)| < \varepsilon$$

Indeed, letting $x - y = \frac{\varepsilon}{2c} u$ with $u \in U$, we have

$$|\lambda(x) - \lambda(y)| = \frac{\varepsilon}{2c} |\lambda(u)| \leq \frac{\varepsilon}{2c} \cdot c = \frac{\varepsilon}{2} < \varepsilon$$

Done.

2 Dominated Extension Theorem

In this section, all vectorspaces are *real.*

The result here involves only elementary algebra and inequalities (apart from an invocation of transfinite induction) and is the heart of the matter. There is no direct discussion of topological vectorspaces here. The goal here is to *extend* a linear function while maintaining a comparison to another function (denoted p below). Thus, for this section we need *not* suppose that the vectorspaces involved are *topological* vectorspaces.

Our use of the term 'extension' is also standard, that a function f' on a superset X' of a set X is an *extension* of f if f' restricted to the smaller set X is f.

Let V be a *real* vectorspace, without any assumption about topologies. Let

$$p : V \to \mathbf{R}$$

be a *non-negative* real-valued function on V so that

$$p(tv) = t \cdot p(v) \qquad (\text{for } t \geq 0) \qquad \textit{(positive-homogeneity)}$$

$$p(v + w) \leq p(v) + p(w) \qquad \textit{(triangle inequality)}$$

(Thus, p is not quite a *semi-norm*, lacking any description of what happens to $p(tv)$ for $t < 0$).

Theorem: Let λ be a real-linear function on a real vector subspace W of V, so that

$$\lambda(w) \leq p(w)$$

for all $w \in W$. Then there is an extension of λ to a real-linear function Λ on all of V, so that

$$-p(-v) \leq \Lambda(v) \leq p(v)$$

for all $v \in V$.

Proof: The crucial step is to extend the functional 'by one step'. That is, let $v_o \in V$. (Note that it is not necessary to assume that $v_o \notin W$). We attempt to define an extension λ' of λ to $W + \mathbf{R}v_o$ by

$$\lambda'(w + tv_o) = \lambda(w) + ct$$

and see what conditions c must fulfill.

For all $w, w' \in W$ we have

$$\lambda(w) - p(w - v_o) = \lambda(w + w') - \lambda(w') - p(w - v_o)$$

$$\leq p(w + w') - \lambda(w') - p(w - v_o) = p(w - v_o + w' + v_o) - \lambda(w') - p(w - v_o)$$

$$\leq p(w - v_o) + p(w' + v_o) - \lambda(w') - p(w - v_o) = p(w' + v_o) - \lambda(w')$$

That is, we have

$$\lambda(w) - p(w - v_o) \leq p(w' + v_o) - \lambda(w')$$

for all $w, w' \in W$. Let σ be the sup of all the left-hand sides as w ranges over W: since the right-hand side is finite, this sup is finite. And let μ be the inf of the right-hand side as w' ranges over W. We have

$$\lambda(w) - p(w - v_o) \leq \sigma \leq \mu \leq p(w' + v_o) - \lambda(w')$$

Then choose c to be any real number so that

$$\sigma \leq c \leq \mu$$

and define

$$\lambda'(w + tv_o) = \lambda(w) + tc$$

To check the comparison with respect to p is relatively easy: in the inequality

$$\lambda(w) - p(w - v_o) \leq \sigma$$

replace w by w/t with $t > 0$, multiply by t and invoke the positive-homogeneity to obtain

$$\lambda(w) - p(w - tv_o) \leq t\sigma$$

from which follows

$$\lambda'(w - tv_o) = \lambda(w) - tc \leq \lambda(w) - t\sigma \leq p(w - tv_o)$$

Likewise, from

$$\mu \leq p(w + v_o) - \lambda(w)$$

we obtain by a similar trick

$$\lambda'(w + tv_o) = \lambda(w) + tc \leq \lambda(w) + t\mu \leq p(w + tv_o)$$

for $t > 0$, which gives the other half of the desired inequality.

Thus, for all $v \in W + \mathbf{R}v_o$ we have

$$\lambda'(v) \leq p(v)$$

Then, using the linearity of λ', we have

$$\lambda'(v) = -\lambda'(-v) \geq -p(-v)$$

which gives the bottom half of the comparison of λ' and p.

To extend to a functional on the *whole* space still dominated by p is a typical exercise in transfinite induction, which we execute as follows. Let $\mathcal{X}$ be the collection of all pairs (X, μ), where X is a subspace of V (containing W), and where μ is real-linear real-valued function on X so that μ restricted to W is λ and so that

$$-p(-x) \leq \mu(x) \leq p(x)$$

for all $x \in X$. We *order* these by writing $(X, \mu) \leq (Y, \nu)$ if $X \subset Y$ and $\nu|_X = \mu$. By the Hausdorff Maximality Principle, there is a *maximal* totally ordered subset $\mathcal{Y}$ of $\mathcal{X}$. Let

$$V' = \bigcup_{(X,\mu)\in\mathcal{Y}} X$$

be the ascending union of all the subspaces in $\mathcal{Y}$. We define a linear functional λ' on this union as follows: for $v \in V'$, take any X so that $(X, \mu) \in \mathcal{Y}$ and $v \in X$ and define

$$\lambda'(v) = \mu(v)$$

The fact that $\mathcal{Y}$ is totally ordered entails that varying the choice of (X, μ) does not affect the definition of λ'.

Then it just remains to check that V' really is the whole space V. Of course, if it were not, then the first part of the proof of this theorem would create an extension to a properly larger subspace, which would contradict the maximality. *Done.*

---

3 Separation Theorem

Still all vectorspaces are *real*.

Now we suppose that V is a *locally convex* topological vectorspace, meaning that there is a local basis at $0 \in V$ consisting of convex sets. Invoking the dominated extension result just above we see that *open* convex sets can be 'separated' from convex sets:

Theorem: Let X be a non-empty convex open subset of a locally convex topological vectorspace V, and let Y be an arbitrary non-empty convex set in V so that $X \cap Y = \phi$. Then there is a *continuous* real-linear real-valued functional λ on V and a constant c so that

$$\lambda(x) < c \leq \lambda(y)$$

for all $x \in X$ and $y \in Y$.

Proof: Fix $x_o \in X$ and $y_o \in Y$. Since X is open, $X - x_o$ is open, and thus

$$U = (X - x_o) - (Y - y_o) = \{(x - x_o) - (y - y_o) : x \in X,\ \ y \in Y\}$$

is open. Further, since $x_o \in X$ and $y_o \in Y$, U contains 0. Since X, Y are convex, U is convex.

Define the *Minkowski functional* $p = p_U$ attached to U by

$$p(v) = \inf\{t > 0 : v \in tU\}$$

The convexity assures that this function p has the *positive-homogeneity* and *triangle-inequality* properties of the auxiliary functional p mentioned in the dominated extension theorem above.

Let $z_o = -x_o + y_o$. Since $X \cap Y = \phi$, $z_o \notin U$, so $p(z_o) \geq 1$. Define a linear functional λ on $\mathbf{R}z_o$ by

$$\lambda(tz_o) = t$$

We check that λ is 'dominated' by p in the sense of the previous section: for $t \geq 0$,

$$\lambda(tz_o) = t \leq t \cdot p(z_o) = p(tz_o)$$

while for $t < 0$

$$\lambda(tz_o) = t < 0 \leq p(tz_o)$$

Thus, indeed,

$$\lambda(tz_o) \leq p(tz_o)$$

for all real t. Thus, λ extends to a real-linear real-valued functional Λ on V, still so that

$$-p(-v) \leq \Lambda(v) \leq p(v)$$

for all $v \in V$.

From the definition of p, $|\Lambda| \leq 1$ on U. Thus, on $\frac{\varepsilon}{2}U$ we have $|\Lambda| < \varepsilon$. That is, the linear functional Λ is *bounded*, so is *continuous* at 0, so is *continuous* on V.

Then, for arbitrary $x \in X$ and $y \in Y$, we have

$$\Lambda x - \Lambda y + 1 = \Lambda(x - y + z_o) \leq p(x - y + z_o) < 1$$

since $x - y + z_o \in U$. Thus, for all such x, y, we have

$$\Lambda x - \Lambda y < 0$$

Therefore, $\Lambda(X)$ and $\Lambda(Y)$ are *disjoint* convex subsets of the real line. Since Λ is not the zero functional, it is necessarily *surjective*, and so is an *open* map. Thus, $\Lambda(X)$ is open, and

$$\Lambda(X) < \sup \Lambda(X) \leq \Lambda(Y)$$

as desired. *Done.*

4 Complex scalars

The two theorems above, stated and proven there for *real* vectorspaces, have analogues in the complex case, which are really just corollaries of them.

Let V be a complex vectorspace. Given a complex-linear complex-valued functional λ on V, let its real part be

$$u(v) = \text{Re } \lambda(v) = \frac{\lambda(v) + \overline{\lambda(v)}}{2}$$

where the overbar denotes complex conjugation. On the other hand, given a *real*-linear *real*-valued functional u on V, its *'complexification'* is

$$\text{Cx } u(x) = u(x) - iu(ix)$$

where $i = \sqrt{-1}$.

Proposition: For any real-linear functional u on the complex vectorspace V, the complexification Cx u is a complex-linear functional so that

$$\text{Re Cx } u = u$$

And for a complex-linear functional λ

$$\text{Cx Re } \lambda = \lambda$$

(The proof is straightforward computation).

Theorem: Let p be a *seminorm* on the complex vectorspace V. Let λ be a complex-linear function on a complex vector subspace W of V, so that

$$|\lambda(w)| \leq p(w)$$

for all $w \in W$. Then there is an extension of λ to a complex-linear function Λ on all of V, so that

$$|\Lambda(v)| \leq p(v)$$

for all $v \in V$.

Proof: Certainly if $|\lambda| \leq p$ then $|\text{Re } \lambda| \leq p$. Then by the theorem for *real*-linear functionals, there is an extension u of Re λ to a *real*-linear functional so that still $|u| \leq p$. Let

$$\Lambda = \text{Cx } u$$

All that remains to show, in light of the proposition above, is that $|\Lambda| \leq p$.

To this end, for given $v \in V$, let μ be a complex number of absolute value 1 so that

$$|\Lambda(v)| = \mu\Lambda(v)$$

Then

$$|\Lambda(v)| = \mu\Lambda(v) = \Lambda(\mu v) = \text{Re } \Lambda(\mu v) \le p(\mu v) = p(v)$$

where we use the full seminorm property of p. Thus, the complex-linear functional made by 'complexifying' the *real*-linear extension of the real part of λ satisfies the desired bound. *Done.*

Theorem: Let X be a non-empty convex open subset of a locally convex topological vectorspace V, and let Y be an arbitrary non-empty convex set in V so that $X \cap Y = \phi$. Then there is a *continuous* complex-linear complex-valued functional λ on V and a constant c so that

$$\text{Re } \lambda(x) < c \le \text{Re } \lambda(y)$$

for all $x \in X$ and $y \in Y$.

Proof: Invoke the real-linear version of the theorem to make a *real*-linear functional u so that

$$u(x) < c \le u(y)$$

for all $x \in X$ and $y \in Y$. Then, by the proposition, u is the real part of its own complexification $\lambda = \text{Cx } u$. *Done.*

---

5 Corollaries

The corollaries can be stated for real *or* complex scalars.

Corollary: Let V be a locally convex topological vectorspace. Let K and C be *disjoint* sets, where K is a *compact* convex non-empty subset of V, and C is a *closed* convex subset of V. Then there is a continuous linear functional λ on V and there are real constants $c_1 < c_2$ so that for all $x \in K$ and $y \in C$

$$\text{Re } \lambda(x) \le c_1 < c_2 \le \text{Re } \lambda(y)$$

Proof: Let U be a small-enough convex neighborhood of 0 in V so that

$$(K + U) \cap C = \phi$$

Then apply the separation theorem to $X = K + U$ and $Y = C$. The constant c_2 can be taken to be $c_2 = \sup \text{Re}\, \lambda(K + U)$. Since Re $\lambda(K)$ is a compact subset of Re $\lambda(K + U)$, its sup c_1 is strictly less than c_2. *Done.*

Corollary: Let V be a locally convex topological vectorspace, W a subspace, and $v_o \in V$. Let $\overline{W}$ denote the topological closure of W. Then $v_o \notin \overline{W}$ if and only if there is a *continuous* linear functional λ on V so that $\lambda(W) = 0$ while $\lambda(v) = 1$.

Proof: On one hand, if v_o lies in the closure of W, then any continuous function which is 0 on W must be 0 on v_o, as well.

On the other hand, suppose that v_o does *not* lie in the closure of W. Then apply the previous corollary with $K = \{v_o\}$ and $C = \overline{W}$. We find that

$$\mathrm{Re}\ \lambda(\{v_o\}) \cap \mathrm{Re}\ \lambda(\overline{W}) = \phi$$

Since Re $\lambda(\overline{W})$ is a vector subspace of the real line, and is not the whole real line, it must be just $\{0\}$. And then Re $\lambda(v_o) \neq 0$. Then divide λ by the constant Re $\lambda(v_o)$ to obtain a continuous linear functional which is zero on W but is 1 on v_o. *Done.*

Corollary: Let V be a locally convex topological (real) vectorspace. Let λ be a continuous linear functional on a subspace W of V. Then there is a continuous linear functional Λ on V extending λ.

Proof: Without loss of generality, consider $\lambda \neq 0$. Let W_o be the kernel of λ (on W), and pick $w_1 \in W$ so that $\lambda w_1 = 1$. Evidently w_1 is not in the closure of W_o, so there is Λ on the whole space V so that $\Lambda|_{W_o} = 0$ and $\Lambda w_1 = 1$. It is easy to check that this Λ is an extension of λ. *Done.*

Corollary: Let V be a locally convex topological vectorspace. Given two distinct vectors $x \neq y$ in V, there is a continuous linear functional λ on V so that

$$\lambda(x) \neq \lambda(y)$$

Proof: The set $\{x\}$ is compact convex non-empty, and the set $\{y\}$ is closed convex non-empty, so we can apply a corollary just above. *Done.*