March 2, 2021

We'll prove the Hahn-Banach theorem today, which explains how to extend bounded linear functionals on a subspace to the whole normed vector space, answering the question of whether the dual of bounded linear functionals is nontrivial for normed vector spaces.

Last time, we discussed **Zorn's lemma** from set theory (which we can take as an axiom), which tells us that a partially ordered set has a maximal element if every chain has an upper bound. (Remember that this notion involves a generalization $\preceq$ of the usual $\leq$.) As a warmup, today we'll use this axiom to prove a fact about vector spaces. Recall that a **Hamel basis** of a vector space V is a linearly independent set H, where every element of V is a finite linear combination of elements of H. We know that finite-dimensional vector spaces always have a (regular) basis, and this is the analog for infinite-dimensional spaces:

Theorem 47

If V is a vector space, then it has a Hamel basis.

Proof. We'll construct a partially ordered set as follows: let E be the set of linearly independent subsets of V, and we define a partial order $\preceq$ by inclusion of those subsets. We now want to apply Zorn's lemma on E, so first we must check the condition: if C is a chain in E (meaning any two elements can be compared), we can define

$$c = \bigcup_{e \in C} e$$

to be the union of all subsets in the chain. We claim that c is a linearly independent subset: to see that, consider a subset of elements $v_1, v_2, \cdots, v_n \in c$. Pick $e_1, e_2, \cdots, e_n \in C$ such that $v_j \in e_j$ for each j: by induction, because we can compare any two elements in C, we can also order finitely many elements in C as well, and thus there is some J such that $e_j \preceq e_J$ for all $j \in [1, 2, \cdots, n]$. So that means that all of $v_1, \cdots, v_n$ are in e_J, which is a linearly independent set by assumption. So indeed our arbitrary set $v_1, \cdots, v_n \in c$ is linearly independent, meaning c is linearly independent.

And now notice that $e \preceq c$ for all $e \in C$ – that is, c is an upper bound of C. So the hypothesis of Zorn is verified, and we can apply Zorn's lemma to see that E has some maximal element H.

We claim that H spans V – suppose otherwise. Then there is some $v \in V$ such that v is not a finite linear combination of elements in H, meaning that $H \cup \{v\}$ is linearly independent. But then $H \prec H \cup \{v\}$ (meaning $\preceq$ but not equality), so H is not maximal, which is a contradiction. Thus H must have spanned V, and that means H is a Hamel basis of V. □

Now that we've seen Zorn's lemma in action once, we're ready to use it to prove Hahn-Banach:

Theorem 48 (Hahn-Banach)

Let V be a normed vector space, and let $M \subset V$ be a subspace. If $u : M \to \mathbb{C}$ is a linear map such that $|u(t)| \leq C||t||$ for all $t \in M$ (in other words, we have a bounded linear functional), then there exists a **continuous extension** $U : V \to \mathbb{C}$ (which is an element of $\mathcal{B}(V, \mathbb{C}) = V'$) such that $U|_M = u$ and $||U(t)|| \leq C||t||$ for all $t \in V$ (with the same C as above).

This result is very useful – in fact, it can be used to prove that the dual of ℓ^∞ is not ℓ^1, even though the dual of ℓ^1 is ℓ^∞.

To prove it, we'll first prove an intermediate result:

Lemma 49

Let V be a normed space, and let $M \subset V$ be a subspace. Let $u : M \to \mathbb{C}$ be linear with $|u(t)| \le C||t||$ for all $t \in M$. If $x \notin M$, then there exists a function $u' : M' \to \mathbb{C}$ which is linear on the space $M' = M + \mathbb{C}x = \{t + ax : t \in M, a \in C\}$, with $u'|_M = u$ and $|u'(t')| \le C|t'|$ for all $t' \in M'$.

We can think of M as a plane and x as a vector outside of that plane: then we're basically letting ourselves extend u in one more dimension, and the resulting bounded linear functional has the same bound that u did. The reason this is a helpful strategy is that we'll apply Zorn's lemma to the set of all continuous extensions of u, placing a partial order using extension. Then we'll end up with a maximal element, and we want to conclude that this maximal continuous extension is defined on V. So this lemma helps us do that last step of contradiction, much like with the proof of existence for a Hamel basis.

Let's first prove the Hahn-Banach theorem assuming the lemma:

Proof of Theorem 48. Let E be the set of all continuous extensions

$$E = \{(v, N) : N \text{ subspace of } V, M \subset N, v \text{ is a continuous extension of } u \text{ to } N\},$$

meaning that it is a bounded linear functional on N with the same bound C as the original functional u. This is nonempty because it contains (u, M). We now define a partial order on E as follows:

$$(v_1, N_1) \preceq (v_2, N_2) \text{ if } N_1 \subset N_2, v_2|_{N_1} = v_1$$

(in other words, v_2 is a continuous extension of v_1). We can check for ourselves that this is indeed a partial order, and we want to check the hypothesis for Zorn's lemma. To do this, let $C = \{(v_i, N_i) : i \in I\}$ be a chain in E indexed by the set I (so that for all $i_1, i_2 \in I$, we have either $(v_{i_1}, N_{i_1}) \preceq (v_{i_2}, N_{i_2})$ or vice versa).

So then if we let $N = \bigcup_{i \in I} N_i$ be the union of all such subspaces N_i, we can check that this is a subspace of V. This is not too hard to show: let $x_1, x_2 \in N$ and $a_1, a_2 \in \mathbb{C}$. Then we can find indices i_1, i_2 such that $x_1 \in N_{i_1}$ and $x_2 \in N_{i_2}$, and one of these subspaces N_{i_1}, N_{i_2} is contained in the other because C is a chain. So (without loss of generality), we know that x_1, x_2 are both in N_{i_2}, and we can use closure in that subspace to show that $a_1x_1 + a_2x_2 \in N_{i_2} \subset N$.

And now that we have the subspace N, we need to make it into an element of E by defining a linear functional $u : N \to \mathbb{C}$ which satisfies the desired conditions. But the way we do this is not super surprising: we'll define $v : N \to \mathbb{C}$ by saying that for any $t \in N$, we know that $t \in N_i$ for some i, and then we define $v(t) = v_i(t)$. But this is indeed well-defined: if $t \in N_{i_1} \cap N_{i_2}$, it is true that $v_{i_1}(t) = v_{i_2}(t)$, because we're still in a chain and thus one of (v_{i_1}, N_{i_1}) and (v_{i_2}, N_{i_2}) is an extension of the other by definition. Similar arguments (exercise to write out the details) also show that v is linear, and that it's an extension of any v_i (including the bound with the constant C). So $(v_i, N_i) \preceq (v, N)$, and we have an upper bound for our chain.

This means we've verified the Zorn's lemma condition, and now we can say that E has a maximal element (U, N). We want to show that $N = V$ (which would give us the desired conclusion); suppose not. Then there is some $x \in V$ that is not in N, and then Lemma 49 tells us that there is a continuous extension v of U to $N + \mathbb{C}x$, which must then also be a continuous extension of u. So $(v, N + \mathbb{C}x)$ is an element of E, but that means $(U, N) \prec (v, N + \mathbb{C}x)$, contradicting (U, N) being a maximal element. So $N = V$ and we're done. ☐

We'll now return to the (more computational) proof of the lemma:

Proof of Lemma 49. We can check on our own that $M' = M + \mathbb{C}x$ is a subspace (this is not hard to do), but additionally, we can show that the representation of an arbitrary $t' \in M'$ as $t + ax$ (for $t \in M$ and $a \in \mathbb{C}$) is unique.

This is because

$$t + ax = \tilde{t} + \tilde{a}x \implies (a - \tilde{a})x = \tilde{t} - t \in M,$$

which means that $x \in M$ (contradiction) unless $a = \tilde{a}$, which then implies that $t = \tilde{t}$. We need this fact because we want to define our continuous extension in a well-defined way: if we choose an arbitrary $\lambda \in \mathbb{C}$, then the map

$$u'(t + ax) = u(t) + a\lambda$$

is indeed well-defined on M', and then the map $u' : M' \to \mathbb{C}$ is linear. If the bounding constant C is zero, then our map is just zero and we can extend that map by just using the zero function on M'. Otherwise, we can divide by C and thus assume (without loss of generality) that $C = 1$. It remains to choose our λ so that for all $t \in M$ and $a \in \mathbb{C}$, we have $|u(t) + a\lambda| \le ||t + ax||$, which would show the desired bound and give us the continuous extension.

To do this, note that the inequality already holds whenever $a = 0$ (because it holds on M), so we just need to choose λ to make the inequality work for $a \neq 0$. Dividing both sides by $|a|$ yields (for all $a \neq 0$)

$$\left| u\left(\frac{t}{-a}\right) - \lambda \right| \le \left|\left| \frac{t}{-a} - x \right|\right|.$$

We know that $\frac{t}{-a} \in M$ because $t \in M$, so this bound is equivalent to showing that

$$|u(t) - \lambda| \le |t - x| \quad \forall t \in M.$$

To do this, we'll choose the real and imaginary parts of λ. First, we show there is some $\alpha \in \mathbb{R}$ such that

$$|w(t) - \alpha| \le ||t - x||$$

for all $t \in M$, where $w(t) = \frac{u(t)+\overline{u(t)}}{2}$ is the real part of $u(t)$. Notice that $|w(t)| = |\text{Re } u(t)| \le |u(t)| \le ||t||$ by assumption, and because w is real-valued,

$$w(t_1) - w(t_2) = w(t_1 - t_2) \le |w(t_1 - t_2)| \le ||t_1 - t_2||$$

(the middle step here is where we use that w is real-valued). Connecting this back to the expression $||t - x||$, we can add and subtract x from above and use the triangle inequality to get

$$w(t_1) - w(t_2) \le ||t_1 - x|| + ||t_2 - x||.$$

Thus, for all $t_1, t_2 \in M$, we have

$$w(t_1) - ||t_1 - x|| \le w(t_2) + ||t_2 - x||,$$

and thus we can take the supremum of the left-hand side over all t_1s to get

$$\sup_{t \in M} w(t) - ||t - x|| \le w(t_2) + ||t_2 - x||$$

for all $t_2 \in M$, and thus

$$\sup_{t \in M} w(t) - ||t - x|| \le \inf_{t \in M} w(t) + ||t - x||.$$

So **now we choose** α to be a real number between the left-hand side and right-hand side, and we claim this value works. For all $t \in M$, we have

$$w(t) - ||t - x|| \le \alpha \le w(t) + ||t - x||,$$

and now rearranging yields

$$-||t-x|| \leq \alpha - w(t) \leq ||t-x|| \implies |w(t)-\alpha| \leq ||t-x||,$$

and we've shown the desired bound. So now we just need to do something similar for the imaginary part, and we do so by repeating this argument with ix instead of x. This then defines our function u' on all of $M + \mathbb{C}x$, and we're done (we can check that because the desired bound holds on both the real and imaginary "axes" of x, it holds for all complex multiples of x). $\square$

MIT OpenCourseWare
https://ocw.mit.edu

18.102 / 18.1021 Introduction to Functional Analysis
Spring 2021

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.