Math 202B — UCB, Spring 2014 — M. Christ
Topic 3: Introduction to Normed vector spaces and The Hahn-Banach Theorem

(Folland §§5.1 and 5.2.) We will work with vector spaces X over K, where the field K is either the real numbers or the complex numbers. I prefer to write $\mathbb{F}$ instead of K and will do so in class and in these notes; our text prefers the notation K.

Initial definitions, facts, concepts:

- Norm. Associated metric on X. Norm topology.
- A useful variant of the triangle inequality: $|\|x\| - \|y\|| \le \|x - y\|$.
- Convergence of a sequence of elements x_n of X.
- The closed unit ball of X is $\{x \in X : \|x\| \le 1\}$.
- A Banach space is a normed vector space that is complete with respect to the associated metric. Completeness is essential to analysis because it makes it possible to take limits.
- Defn: Equivalent norms on a vector space.
- Absolutely convergent series. Proposition: A normed vector space X is a Banach space if and only if every absolutely convergent series of elements of X converges in X.[1]
- Examples of Banach spaces: There are many, many, many. Some good ones: (i) $\mathbb{F}^d$ for $d \in \mathbb{N}$. (ii) L^1 of any measure space. (iii) $C^0(K)$, the space consisting of all continuous functions from K to $\mathbb{F}$, for any compact topological space K. (iv) ℓ^∞, the set of all bounded sequences $x = (x_1, x_2, \dots, x_n, \dots)_{n\in\mathbb{N}}$. Each of these has a natural norm which I have not specified.
- Defn: Bounded linear map $T : X \to Y$ where X, Y are normed linear spaces. Definition of $\|T\|$. $L(X, Y)$ = set of all such T. This has a natural vector space structure.
- If X, Y are normed vector spaces and $T : X \to Y$ is linear (over $\mathbb{F}$, always), the the following three properties of T are equivalent: (i) T is bounded, (ii) T is continuous, (iii) T is continuous at 0. (Very easy proof.)
- Defn of $\|T\|$ for $T \in L(X, Y)$. This defines a norm for $L(X, Y)$. If Y is complete then $L(X, Y)$ is likewise complete.
- $\|S \circ T\| \le \|S\| \cdot \|T\|$.
- Linear functionals and bounded linear functionals. The *dual space* X^* of X consists of all bounded linear functionals[2] $f : X \to \mathbb{F}$.
- Defn: sublinear functionals. (§5.2) These take values in $\mathbb{R}$ and satisfy $p(x + y) \le p(x) + p(y)$ for all $x, y \in X$, and $p(tx) = tp(x)$ for all *nonnegative* scalars t. The most important example is $p(x) = \|x\|_X$.

In mathematics there are some simple facts whose usefulness seems to be inversely proportional to the degree of difficulty in their proofs. Here is one:

Theorem. All norms on a finite-dimensional vector space are mutually equivalent. □

---

[1]And my calculus students ask what is the use of absolute convergence ...

[2]Most authors prefer other symbols to f. Common choices include T and λ.

I'll prove this for $\mathbb{F} = \mathbb{R}$, just to make it seem more concrete; the proof for $\mathbb{F} = \mathbb{C}$ is identical. In the proof I'll use $\mathbb{R}^d$. Denote its elements by $x = (x_1, \ldots, x_d) = \sum_{i=1}^d x_i e_i$ where $\{e_i\}$ is the standard basis for $\mathbb{R}^d$. $\mathbb{R}^d$ has the standard norm $\|x\|_{\mathbb{R}^d} = (\sum_{i=1}^d x_i^2)^{1/2}$. For any $x \in \mathbb{R}^d$, $\sum_i |x_i| \le d\|x\|_{\mathbb{R}^d}$. Indeed, $|x_k| \le (\sum_i x_i^2)^{1/2}$ for any index k. Then sum over all k.

Proof. Let V be a finite-dimensional vector space. Assume that V has positive dimension d to avoid trivialities. One way to define a particular norm on V is to use the fact that V is isomorphic as a vector space to $\mathbb{R}^d$. Choose some basis $\{v_1, \ldots, v_d\}$ for V over $\mathbb{R}$. Then the correspondence $\mathbb{R}^d \ni x \leftrightarrow \sum_{i=1}^d x_i v_i \in V$ is an isomorphism of vector spaces. Thus we may identify V with $\mathbb{R}^d$. We only have to prove that any two norms on $\mathbb{R}^d$ are equivalent.

Let $\|\cdot\|_X$ be some norm for $\mathbb{R}^d$. The function $F(x) = \|x\|_X$ is continuous from $\mathbb{R}^d$ to $[0, \infty)$. Indeed,

$$|F(x) - F(y)| = |\|x\|_X - \|y\|_X| \le \|x - y\|_X.$$

Let $S = \{x \in \mathbb{R}^d : \|x\|_{\mathbb{R}^d} = 1\}$ be the unit sphere. This is a compact subset of a metric space. Since F is continuous, $F(S)$ is a compact subset of $[0, \infty)$. If $x \in S$ then $x \ne 0$, so $F(x) = \|x\|_X \ne 0$. Thus $0 \notin F(S)$. Since a continuous function on a compact set attains its minimum, there exists $\delta > 0$ such that $F(x) \ge \delta$ whenever $|x| = 1$. Equivalently,

$$\|x\|_X \ge \delta |x|$$

for all $x \in \mathbb{R}^d$.

The converse inequality is even easier. Let $\{e_i\}$ be the standard basis for $\mathbb{R}^d$. Define $A = \max_i \|e_i\|_X < \infty$. Then

$$F(x) = \|\sum_i x_i e_i\|_X \le \sum_i \|x_i e_i\|_X = \sum_i |x_i| \|e_i\|_X \le A \sum_i |x_i| \le Ad|x|.$$

Thus $\|\cdot\|_X$ is equivalent to the standard norm $|\cdot|$. □

Theorem. Every finite-dimensional normed vector space is complete. □

Proof. Let $X = (V, \|\cdot\|_X)$ be a normed vector space. Choose a basis $\{v_i\}$ for V. Any element $x \in V$ can be uniquely expressed as $x = \sum_i x_i v_i$ where $x = (x_i : 1 \le i \le d) \in \mathbb{F}^d$. Define $\|x\|_* = (\sum_i |x_i|^2)^{1/2}$. This is a norm on V. By the above theorem, it is equivalent to the given norm. Any sequence that is Cauchy with respect to the given norm is Cauchy with respect to the equivalent norm $\|\cdot\|_*$, hence converges with respect to $\|\cdot\|_*$, hence converges with respect to the equivalent norm $\|\cdot\|_X$. □

Fact. If X is a finite-dimensional normed vector space then its closed unit ball is compact. Indeed, as above, it suffices to prove when the underlying vector space is $\mathbb{F}^d$. The topology induced on $\mathbb{F}^d$ by the norm $\|\cdot\|_X$ coincides with the standard topology. The unit ball of X is the set of all $x \in \mathbb{R}^d$ for which $\|x\|_X$ is ≤ 1. Since this norm is a continuous function of x, this set is closed. Since this norm is equivalent to the standard norm, this set is bounded. Therefore it is compact. □

It is an important and fundamental fact that the converse is also true.

Theorem. Let X be a normed vector space. If the closed unit ball of X is compact then X is finite-dimensional. □

An awful lot of infinite-dimensional analysis consists largely of attempts to circumvent this general fact in particular circumstances.

Lemma 1. Let $0 \le \gamma < 1$. Let V be a subspace of a normed vector space X. Suppose that for each $x \in X$ there exists $y \in V$ such that $\|x - y\| \le \gamma \|x\|$. Then V is dense in X. □

You will give a proof as part of your solution of problem 5.12(b) of this week's problem set.

Lemma 2. Any finite-dimensional subspace of a normed vector space is closed. □

Assuming the lemmas, here's a proof of the Theorem: Assume that the closed unit ball B of X is compact. We may assume that X has positive dimension. Recursively execute the following greedy algorithm to construct a sequence $v_1, v_2, \dots$ of elements of X. Choose $v_1 \in X$ satisfying $\|v_1\| = 1$. Suppose that $v_1, \dots, v_n$ have been constructed. Let V_n be their span.

If there exists $x \in X$ such that $\|x - y\| > \frac{1}{2}\|x\|$ for every $y \in V_n$, then set $v_{n+1} = x/\|x\|$, and continue the construction. (Any such x is necessarily $\ne 0$.) Note that $\{v_1, \dots, v_{n+1}\}$ is linearly independent, by construction.

If there exists no such x then the construction halts. See below for analysis of this case. If the construction never halts then it produces an infinite sequence v_i of elements of B satisfying $\|v_i - v_j\| \ge \frac{1}{2}$ for every $i \ne j$. Such a sequence has no Cauchy subsequence. This contradicts the assumption that B is compact. So the construction must halt after some finite number of steps.

Suppose that the construction produces $v_1, \dots, v_n$ and then halts. By Lemma 1, the subspace V_n is then dense in X. By Lemma 2, V_n is also closed. We conclude that $V_n = X$, whence X has finite dimension. □

The Hahn-Banach Theorem

The Hahn-Banach theorem is the first substantial theorem of this unit. It is a remarkably versatile tool, widely used throughout mathematics. For instance, it can be invoked to prove existence of solutions of partial differential equations in various contexts. We start with the version for $\mathbb{F} = \mathbb{R}$.

Theorem. Let X be a normed vector space over $\mathbb{R}$. Let V be a subspace of X. Let $f : V \to \mathbb{R}$ be a bounded linear functional. Then there exists $F \in X^*$ whose restriction to V equals f. □

The following more general statement provides a convenient route to a proof, and is sometimes useful in its own right.

Theorem. Let X be a vector space over $\mathbb{R}$. Let p be a sublinear functional on X. Let V be a subspace of X, let $f : V \to \mathbb{R}$ be linear, and suppose that $f(x) \le p(x)$ for all $x \in V$. Then there exists a linear functional $F : X \to \mathbb{R}$ satisfying $F \le p$ and $F|_V = f$. □

Here is an example: Let ℓ^∞ be the normed complex vector space consisting of all sequences $(x_n : n \ge 1)$ with $x_n \in \mathbb{C}$, with norm $\|x\|_\infty = \sup_n |x_n|$. Let V be the subspace consisting of all sequences such that $\lim_{n\to\infty} x_n$ exists. Let $p(x) = \|x\|_\infty$. Define $f(x) = \lim_{n\to\infty} x_n$. The theorem guarantees existence of a generalization of the notion of limit to arbitrary bounded sequences; there exists some $F : \ell^\infty \to \mathbb{C}$ that satisfies

$F(x) = \lim_{n\to\infty} x_n$ whenever this limit exists; $F(x)$ is defined for every bounded sequence x; the generalized limit of a sum equals the sum of the generalized limits; similarly for multiplication by constants; and $|F(x)| \le \sup_n |x_n|$ for all n, so that the generalized limit lies in a reasonable range. Now try to construct an explicit F with all of these properties!

The heart of the matter is this:

Lemma. Let X be a vector space over $\mathbb{R}$. Let p be a sublinear functional on X. Let V be a subspace of X, let $f : V \to \mathbb{R}$ be linear, and suppose that $f(x) \le p(x)$ for all $x \in V$. Let $y \notin V$. There exists a linear functional F from $\mathbb{R}y + V$ to $\mathbb{R}$ satisfying $F \le p$ and $F|_V = f$. □

$\mathbb{R}y + V$ denotes the subset of X consisting of all elements expressible as $ty + v$ with $t \in \mathbb{R}$ and $v \in V$. This is manifestly a subspace, and the representation $x = ty + v$ is uniquely determined by x, since $y \notin V$.

Proof. Let $W = \mathbb{R}y+V$. If F is to be linear, then it must satisfy $F(ty+v) = tF(y)+F(v) = tF(y) + f(v)$ for any $t \in \mathbb{R}$ and $v \in V$. We seek to define F by choosing some $\tau \in \mathbb{R}$ and setting

$$F(ty + v) = t\tau + f(v) \quad \text{for all } t, v.$$

Any choice of τ defines a linear functional F; the only question is whether there exists $\tau \in \mathbb{R}$ satisfying

$$t\tau + f(v) \le p(ty + v) \ \forall t \in \mathbb{R} \quad \text{and} \quad v \in V.$$

For any $t \ne 0$, the right-hand side equals $|t|p(y + |t|^{-1}v)$, while the left-hand side equals $|t| \cdot (\tau + f(|t|^{-1}v))$. By cancelling the common factor of $|t|$, we find tha τ need only satisfy two conditions:

$$\tau + f(v) \le p(y + v) \ \forall v \in V$$
$$-\tau + f(v) \le p(-y + v) \ \forall v \in V.$$

Equivalently,

$$f(v) - p(-y + v) \le \tau \le p(y + w) - f(w) \ \forall v, w \in V.$$

At least one solution τ exists if and only if

$$\sup_{v\in V} f(v) - p(-y + v) \le \inf_{w\in V} p(y + w) - f(w).$$

Thus we only need prove that

$$f(v) - p(-y + v) \le p(y + w) - f(w) \ \forall v, w \in V.$$

This is equivalent to

$$f(v) + f(w) \le p(y + w) + p(-y + v) \ \forall v, w \in V.$$

But by linearity of f, sublinearity of p, and the assumption that $f \le p$ on V,

$$f(v) + f(w) = f(v + w) \le p(v + w) = p((-y + v) + (y + w)) \le p(-y + v) + p(y + w).$$

□

It's easy to complete the proof of the Hahn-Banach Theorem under an extra assumption: Suppose that there exists a sequence of subspaces V_j of X such that $V_0 = V$, each V_{j+1} is equal to $\mathbb{F}x_j + V_j$ for some $x_j \in X$, and $\cup_{j=0}^{\infty} V_j = X$. (Informally speaking, this implies that V has countable codimension.) Define f_1 to be an extension of f to V_1 that satisfies $f_1 \le p$. Define f_{j+1} inductively on j to be an extension of f_j to V_{j+1} that satisfies $f_{j+1} \le p$.

For any $x \in X$, and any indices $i \le j$, if $x \in V_i$ then $x \in V_j$ and $f_i(x) = f_j(x)$. There exists a smallest J, depending on x, such that $x \in X_J$. Define $F(x) = f_J(x)$, which is $\le p(x)$ by construction. It is easy to check that F is linear, and $F|_V = f_0$. □

This simple argument can't be used in general. There are two main alternatives. **(i)** Mimic the argument just given, dropping the extra assumption and using transfinite induction in place of ordinary induction. We don't assume any knowledge of transfinite induction in this course, so that route is unavailable. **(ii)** Use Zorn's Lemma.

Define S to be the set of all partial extensions of f. That is, S is the set of all ordered pairs (W, g) such that W is a subspace of X that contains V, $g : W \to \mathbb{R}$ is linear, $g|_V = f$, and $g \le p$.

Define a partial ordering on S: $(W, g) \le (W', g')$ if $W \subset W'$ and g' is an extension of g to W'.

Any chain in the partial ordered set S is bounded above. Indeed, let $\{(W_\alpha, g_\alpha) : \alpha \in A\}$ be a chain in S. The subspaces W_α are *nested*: If $\alpha, \beta \in A$ then $W_\alpha \subset W_\beta$ or $W_\beta \subset W_\alpha$.

Define $W = \cup_{\alpha \in A} W_\alpha$. A union of subspaces is not in general a subspace, but a union of any nested familhy of subspaces is a subspace (easy). So W is a subspace of x.

If $x \in W$ then there exists $\alpha \in A$ such that $x \in W_\alpha$. If α, β are any two such indices then $g_\alpha(x) = g_\beta(x)$ by definition of the partial ordering on S and because S is a chain. Therefore we may define $g : W \to \mathbb{R}$ by $g(x) = g_\alpha(x)$ for all α such that $x \in W_\alpha$. It is (as in the special case of countable codimension treated above) easy to very that g is linear, and of course $g \le p$. Now $(W_\alpha, g_\alpha) \le (W, g)$ for every $\alpha \in A$. Thus (W, g) is an upper bound for S.

We have shown that the partially ordered set S satisfies the hypotheses of Zorn's Lemma. (Any chain has an upper bound in S.) Therefore it has a maximal element, (W, g). We claim that $W = X$. If so, then $F = g$ is an extension of f to $X = W$ satisfying $F \le p$, as required.

Suppose to the contrary that $W \ne X$. Then there exists $y \in X$ which does not belong to W. Set $W' = \mathbb{F}y + W$, the span of $\{y\} \cup W$; this is a subspace of X. Apply the Lemma to obtain an extension g' of g to W', satisfying $g' \le p$. Now $(W', g') \in S$, and $(W, g) \le (W', g')$ according to the partial order for S. The subspace W' properly contains W since $y \notin W$, so $(W, g) < (W', g')$. Thus (W, g) is not a maximal element of S — a contradiction.

Therefore $W = X$, and the proof is complete. □

Zorn's Lemma says: Let $(S, \le)$ be a partially ordered set. Suppose that any chain in S has an upper bound in S. Then S has a (that is, at least one) maximal element.

A chain is a subset of S that is linearly ordered under $\le$. That is, if $s, s' \in S$ then either $s \le s'$ or $s' \le s$. An upper bound for a subset S' of S is an element $s \in S$ such that $s' \le s$

for every $s' \in S'$.

- The concrete lemma that is really the crux of the matter comes in only at the end, in the proof that $W = X$.
- Zorn's Lemma is equivalent (under the other standard axioms of set theory) to the Axiom of Choice. Applications of Hahn-Banach abound throughout Mathematics. They are all applications of the Axiom of Choice.
- There is a version of Hahn-Banach for $\mathbb{F} = \mathbb{C}$. It is a simple corollary of the case $\mathbb{F} = \mathbb{R}$, using the fact that any vector space over $\mathbb{C}$ can be regarded in a natural way as a vector space over $\mathbb{R}$. See our text for statement and the proof.

Applications of Hahn-Banach

Recall from §5.1 the definition of X^*, the dual space of X. This is the set of all bounded linear mappings from X to $\mathbb{F}$. These are usually called bounded linear *functionals*. One way to think of these: $\ell(x)$ provides one single "bit" of information about x, and does so in a linear manner. Another way: $\ell(x)$ is one single coordinate of x, in some unspecified coordinate system (unless $\ell \equiv 0$).

X^* is a vector space, and is a normed vector space with

$$\|f\|_{X^*} = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)|.$$

A basic consequence of Hahn-Banach is that any normed vector space has a rich dual space X^*.

Theorem. Let X be a normed vector space over $\mathbb{F}$.

(i) If V is a closed subspace of X and $x \notin V$ then there exists $F \in X^*$ satisfying

$$F|_V \equiv 0 \quad \text{and} \quad F(x) \neq 0.$$

Moreover, F can be chosen to satisfy both $F(x) = 1$ and $\|F\|_{X^*} \le \big(\inf_{v \in V} \|x - v\|\big)^{-1}$.

(ii) For any $0 \neq x \in X$ there exists $F \in X^*$ such that $|F(x)| = \|x\|$ and $\|F\| = 1$.

(iii) If $x \neq y$ are distinct elements of X then there exists $F \in X^*$ satisfying $F(x) \neq F(y)$.

□

Proof. (ii) is the special case $V = \{0\}$ of (i). (iii) is proved by applying (ii) to $x - y$. So it suffices to prove (i).

We apply Hahn-Banach with $p(y) = \|y\|$ for all $y \in X$. Let V, x be given. Define W to be the span of $\{x\} \cup V$; $W = \mathbb{F}x + V$.

Let $\delta = \inf_{v \in V} \|x - v\|$. Since x does not belong to the closure of V, $\delta > 0$.

Define $f : W \to \mathbb{F}$ by $f(tx + v) = t\delta$. Note that if $tx + v = t'x + v'$ with $v, v' \in V$ and $t, t' \in \mathbb{F}$, then $(t - t')x = v' - v \in V$. Since $x \notin V$, this forces $t - t' = 0$ and thus $v = v'$. Therefore an element of W has only one representation in the form $tx + v$, so f is well-defined. It is easily checked that f is linear.

Is f bounded? For any $t \neq 0$ and any $v \in V$, $\|tx + v\| = |t|\|x - (-t^{-1}v)\| \geq |t|\delta$.

$$\frac{|f(tx+v)|}{\|tx+v\|} = \frac{|t|\delta}{\|tx+v\|} \leq \frac{|t|\delta}{|t|\delta} = 1.$$

Thus $f \leq p$ on W.

By Hahn-Banach, there exists an extension F of f to all of X satisfying $F(z) \leq p(z) = \|z\|$ for all $z \in X$. Now $-F(z) = F(-z) \leq \|-z\| = \|z\|$ as well. Therefore $|F(z)| \leq \|z\|$ for all $z \in X$. Since F is an extension of f, $F(x) = f(x) = \delta \neq 0$.

Replacing F by $\delta^{-1}F$ yields the quantitative version of the conclusion stated above. □

On Wednesday 2/19 we will discuss the Baire Category Theorem, §5.3 of our text. This will lead to fundamental results about *complete* normed vector spaces.