DYSFUNCTIONAL ANALYSIS

VIV

1. Preliminaries and Motivation

The study of functional analysis is, at its core, really a (careful) study of infinite dimensional linear algebra. So let's talk about some infinite-dimensional linear algebra! First, let's keep in mind some of the things that we might care about understanding from finite-dimensional linear algebra:

- We have vector spaces over $\mathbb{R}$ with vectors
- that have lengths
- and transposes (dual vectors + dot products)
- and bases
- and there are linear transformations between vector spaces
- that have eigenvalues and eigenvectors.

We'll talk about as many bullet points as we can this week.

2. Length

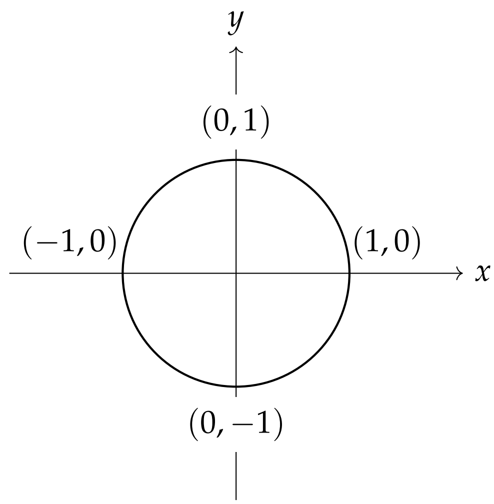
To begin with, when we think of finite dimensional vector spaces, we often think of $\mathbb{R}^n$, something that looks like a space of n-vectors. So, what if we just consider a space of infinite-dimensional vectors, i.e. $\mathbb{R}^\infty$ in the same sense?

Formally this could be OK, but one problem is that this allows things like $(1, 1, 1, 1, \dots)$ or even worse $(1, 2, 3, 4, 5, 6, 7, \dots)$, which don't jibe well with how we think lengths ought to work.

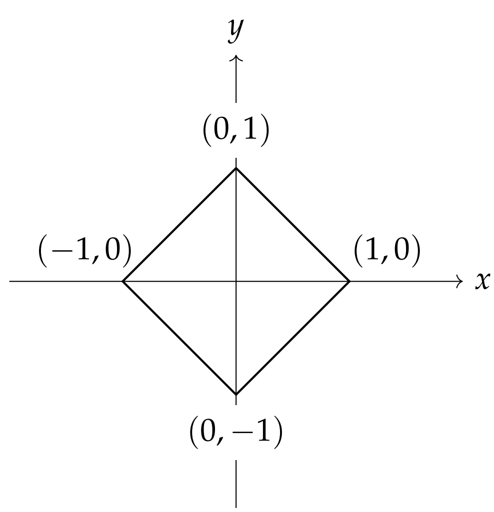
Definition 2.1. A *normed linear space* is a vector space V over $\mathbb{R}$ or $\mathbb{C}$, along with a function, called the *norm* and denoted $||\cdot||$, from V to $\mathbb{R}$ which satisfies:

- $||v|| \geq 0$ for all $v \in V$ (positivity);
- $||v|| = 0 \iff v = 0$ (definiteness);
- $||\alpha v|| = |\alpha|||v||$ for all $v \in V$, for all α in the base field (scaling);
- and $||v + w|| \leq ||v|| + ||w||$ for all v and w in V (triangle inequality).

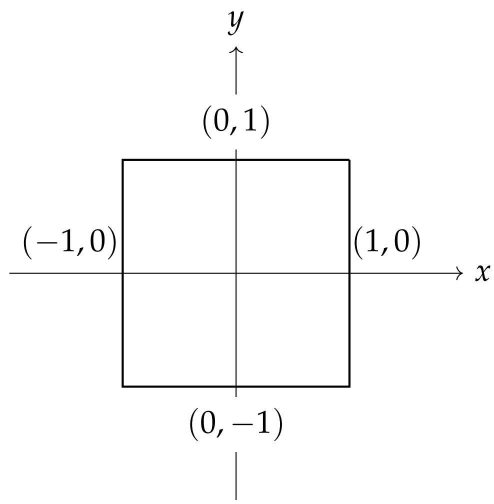
Example 2.2. Let's think about $\mathbb{R}^2$ for the moment. The example of norm that we see most often is the standard length, where $||(a, b)||_2 = \sqrt{a^2 + b^2}$. If we draw the set of vectors with norm 1 (or unit vectors) under this norm, we get the standard unit circle:



However, another norm we could have used is the *taxicab norm*, where $||(a,b)||_t = |a| + |b|$. This norm certainly satisfies positive definiteness and scaling; it's worth thinking about why it also satisfies the triangle inequality. In this case the "unit circle" looks like this:



Yet a third norm we could consider is the *supremum norm*, where $||(a,b)||_s = \max\{a,b\}$. In this case the "unit circle" is a square:



On the other hand, there's in some sense fewer norms than it may seem in the finite dimensional case. If you're counting length by circles or squares or diamonds, no matter what, you'll still be able to tell that a pretty long vector is pretty long and a pretty short vector is pretty short, even if you're fudging about the details of specific lengths in specific cases. This concept is *equivalence* of norms.

Definition 2.3. Let $||\cdot||_a$ and $||\cdot||_b$ be two norms on a finite-dimensional vector space V. We say that $||\cdot||_a$ and $||\cdot||_b$ are *equivalent* if there are positive constants C and D so that, for all $v \in V$,

$$C||v||_a \leq ||v||_b \leq D||v||_a,$$

or equivalently $C||v||_a|| \leq ||v||_b$ and $\frac{1}{D}||v||_b \leq ||v||_a$.

Equivalence (along with the constants C and D) gives a bound on how off the length measurements can be under two different norms. For the finite-dimensional case, we get equivalence everywhere:

Theorem 2.4. *If V is a finite-dimensional vector space, all norms on V are equivalent.*

Lemma 2.5 (Cauchy-Schwarz inequality)**.** *Let $v, w \in V$ be arbitrary vectors, where V has an orthonormal basis $\{e_1, e_2, \dots\}$. Then*

$$|v \cdot w| \leq ||v||_2 ||w||_2,$$

where $||\cdot||_2$ is the standard Euclidean norm, and $v \cdot w$ is the standard dot product.

Proof. Note that $||v||_2 = \sqrt{v \cdot v}$ (and same for w). If $v = 0$, then both sides are 0, and we're done. If not, let $z = u - \frac{u \cdot v}{v \cdot v} v$. Then $z \cdot v = 0$, since

$$z \cdot v = \left(u - \frac{u \cdot v}{v \cdot v} v\right) \cdot v = u \cdot v - \frac{u \cdot v}{v \cdot v} v \cdot v = 0,$$

so z and v are orthogonal. We then apply the Pythagorean theorem to $u = \frac{u\cdot v}{v\cdot v}v + z$, which yields

$$\begin{aligned}||u||_2^2 &= \left|\frac{u\cdot v}{v\cdot v}\right|^2 ||v||_2^2 + ||z||_2^2\\ &= \frac{|u\cdot v|^2}{||v||_2^4}||v||_2^2 + ||z||_2^2\\ \Rightarrow ||v||_2^2||u||_2^2 &= |u\cdot v|^2 + ||z||_2^2||v||_2^2\\ \Rightarrow ||v||_2^2||u||_2^2 &\le |u\cdot v|^2.\end{aligned}$$

□

Remark 2.6. As we will use crucially much later, we never used that V was finite dimensional!!

Proof. We will show that any norm $||\cdot||$ is equivalent to $||\cdot||_2$, the standard Euclidean norm. This is the same as showing that there exists $C > 0$, $D > 0$ so that for all $v \in V$, $C||v|| \le ||v||_2 \le D||v||$. Note that this is automatic for $v = 0$, and that it suffices to show that this holds for v with $||v||_2 = 1$. If we know this result in the case when $||v||_2 = 1$, then for any $v \in V$, $\left|\left|\frac{v}{||v||_2}\right|\right|_2 = 1$, so we know that

$$C\left|\left|\frac{v}{||v||_2}\right|\right| \le 1 \le D\left|\left|\frac{v}{||v||_2}\right|\right|,$$

which by multiplying out by $||v||_2$ implies $C||v|| \le ||v||_2 \le D||v||$ for the same C and D.

Assume V is n-dimensional and let $\{e_1, \dots, e_n\}$ be a basis for V. Let $v \in V$ with $||v||_2 = 1$ and let $v = a_1e_1 + \cdots + a_ne_n$, so that $||v||_2 = \left(\sum_{i=1}^n |a_i|^2\right)^{1/2}$. By the triangle inequality,

$$||v|| = \left|\left|\sum_{i=1}^n a_ie_i\right|\right| \le \sum_{i=1}^n |a_i|||e_i||.$$

But note that this last sum is the same as the dot product of the vector $(|a_1|, \dots, |a_n|)$ and the vector $(||e_1||, \dots, ||e_n||)$, so we have

$$\sum_{i=1}^n |a_i|||e_i|| \le \left(\sum_{i=1}^n |a_i|^2\right)^{1/2}\left(\sum_{i=1}^n ||e_i||^2\right)^{1/2} = \left(\sum_{i=1}^n ||e_i||^2\right)^{1/2},$$

by Cauchy-Schwarz. If we let C be $\left(\sum_{i=1}^n ||e_i||^2\right)^{-1/2}$, which is independent of v, we get that $C||v|| \le 1$, as desired.

Now we want to find D so that $1 \le D||v||$. We've just proven that for all $w \in V$, $C||w|| \le ||w||_2$, which implies that for any sequence of vectors $\{w_i\}_{i=1}^\infty$ with $||w_i||_2 \to 0$, we must also have $||w_i|| \to 0$. In particular, this means that the function $||\cdot|| : V \to \mathbb{R}$ is continuous! Since the set of vectors in V of length 1 is closed and bounded, it's compact, so the continuous function $||\cdot||$ has a minimum on the unit sphere. Let $m > 0$ be the minimum value of $||\cdot||$ on the unit sphere. Then for all v on the unit sphere, $1 \le \frac{1}{m}||v||$, so $||v||_2 \le \frac{1}{m}||v||$. □

But for this argument it was crucial that V be finite-dimensional! The place where we really used this was in saying that "the set of vectors in V of length 1 is closed and bounded." If V is infinite dimensional and has a norm, the set of vectors of length 1 is *not* compact!

Let's see how this works in practice; we'll start by giving a bunch of examples of infinite-dimensional normed vector spaces.

Example 2.7.
- Let
$$\ell_2 = \left\{ \{a_i\}_{i=1}^{\infty} \mid a_i \in \mathbb{R}, \sum_{i=1}^{\infty} |a_i|^2 < \infty \right\},$$
i.e. the set of sequences $\{a_i\}_{i=1}^{\infty}$ where $a_i \in \mathbb{R}$ and whose square sums converge. We can then define the ℓ_2 *norm* on V, denoted $||\cdot||_2$, via $||\{a_i\}||_2 = \left(\sum_{i=1}^{\infty} |a_i|^2\right)^{1/2}$. Note that V is exactly the subset of vectors in $\mathbb{R}^{\infty}$ that have finite length in the ℓ_2 norm.
- The above example embodies our general strategy; pick a norm first, and then take as our space the set of vectors with finite norm. We can instead consider the ℓ_1 *norm* on sequences, given by $||\{a_i\}||_1 = \sum_{i=1}^{\infty} |a_i|$. This gives
$$\ell_1 = \left\{ \{a_i\}_{i=1}^{\infty} \mid a_i \in \mathbb{R}, \sum_{i=1}^{\infty} |a_i| < \infty \right\},$$
under the ℓ_1 norm.
- The *sup norm* or ℓ_∞ *norm* on sequences is given by $||\{a_i\}||_\infty = \sup_i |a_i|$, so we have the space
$$\ell_\infty = \left\{ \{a_i\}_{i=1}^{\infty} \mid a_i \in \mathbb{R}, \sup_i |a_i| < \infty \right\},$$
under the sup norm.
- In this example we can take the sup norm but restrict the vector space we're looking at, to be sequences where only finitely many entries are nonzero. This is denoted c_0.

Note that without considering the norm structure, $c_0 \subseteq \ell_1 \subseteq \ell_2 \subseteq \ell_\infty$. So let's consider as an example c_0, which we'll endow with either the ℓ_1 norm or the ℓ_2 norm. Consider the sequence of vectors $v_n = (\underbrace{1, \dots, 1}_{n}, 0, \dots)$ in ℓ_1. Then $||v_n||_1 = n$, but $||v_n||_2 = \sqrt{n}$. Thus for any $C > 0$, if we pick n so that $\frac{1}{\sqrt{n}} < C$, we have $C||v_n||_1 = Cn > \frac{n}{\sqrt{n}} = \sqrt{n} = ||v_n||_2$, and thus $C||v_n||_1 \not\leq ||v_n||_2$.

Exercise 2.8. Find a sequence of vectors in c_0 that shows that the ℓ_1 norm is not equivalent to the sup norm, and one that shows that the ℓ_2 norm is not equivalent to the sup norm.

This means that we'll generally be talking about infinite dimensional vectors spaces with a specified norm. Since there are notions of length that work so fundamentally differently, it's important to keep track! One last important definition with norms is the notion of a *complete* space: if a vector space has a norm, we have a notion of convergence of sequences and series, and we can ask that if a sequence converges, its limit is in the vector space.

Definition 2.9. A *complete normed vector space* is a vector space V along with a norm $||\cdot||$ satisfying the condition that if $\{v_n\}_{n=1}^{\infty}$ is a Cauchy sequence, i.e. if $\max_{m>n}||v_n - v_m|| \to 0$ as $n \to \infty$, then $\{v_n\}$ has a limit in V, i.e. there exists $v \in V$ with $||v_n - v|| \to 0$.

A complete normed vector space is also called a *Banach space*.

Example 2.10.
- c_0 is *not* complete. Why? Consider the sequence $\{v_n\}_{n=1}^{\infty}$, where $v_n = (1, \frac{1}{2}, \frac{1}{4}, \cdots, \frac{1}{2^n}, 0, \ldots, 0)$, and the vector $v = (1, \frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{2^n}, \ldots)$. The sequence $\{v_n\} \to v$, since the supremum of the entries of $v - v_n$ is $\frac{1}{2^{n+1}}$. But v has infinitely many nonzero entries, so $v \notin c_0$.
- ℓ_1 and ℓ_2 are complete.

3. Bases

The definition of basis that we're used to in the finite case is the following:

Definition 3.1. A *Hamel basis B* of a vector space V is a linearly independent set of vectors $B = \{x_i\}_{i\in I}$ such that every element $v \in V$ can be written as a linear combination of elements of B,

where by "linear combination" we mean a finite linear combination, i.e. $v = \sum_{i=1}^{n} a_i b_i$, and by "linearly independent" we mean that 0 is not a nontrivial finite linear combination of basis elements. Note that we don't even need the index set to be countable; in fact, it most often won't be.

Example 3.2. The set $\{v_n\}_{n=1}^{\infty}$ of vectors in c_0, where v_n has a 1 in the nth entry and 0s elsewhere, is a Hamel basis for c_0.

What's a Hamel basis for ℓ_2? Well, we can start by having the same v_n's that we chose for the Hamel basis of c_0, but since we're only taking finite linear combinations we can't get to vectors like $(1, \frac{1}{2}, \frac{1}{4}, \ldots)$. So maybe we add that one in, but then we can't get to vectors like $(1, 0, \frac{1}{2}, 0, \frac{1}{4}, 0, \ldots)$, and already this starts to get quite unwieldy to work with...and life just doesn't get better.

Fact 3.3. There does exist a Hamel basis for ℓ_2.

Proof. Exercise in logic. (Hint: Zorn's Lemma is necessary!). □

Exercise 3.4. If V is any infinite-dimensional Banach space, then V cannot have a countable Hamel basis.

So we may need a definition that's a bit more robust than the Hamel basis definition; in particular, it sure would be nice to allow infinite sums. Moreover, right now if we start to think about components or projection, it's quite awkward!

Definition 3.5. A countable sequence $\{v_n\}_{n=1}^{\infty}$ in a Banach space V is a *basis* for V if for all $v \in V$, there exists a unique sequence of scalars $\{a_n\}_{n=1}^{\infty}$ such that

$$v = \sum_{n=1}^{\infty} a_n v_n.$$

The series is called the *basis representation* of v with respect to $\{v_n\}$.

Remark 3.6. The convergence in the definition means convergence in the norm of the Banach space.

Remark 3.7. In our definition, *not every space has a basis.* For example, as we'll see later, ℓ_∞ will not have one. This is a big deal!

Example 3.8. Under this definition, the set $\{e_n\}_{n=1}^\infty$ where e_n is the sequence with a 1 in the nth spot and 0s elsewhere is a basis for ℓ_1 and ℓ_2. But, it's *not* a basis for ℓ_∞; why not?

4. Linear Transformations

Now that we have a bit of a handle on what these spaces are, let's talk about linear maps for a bit. In the finite dimensional case, we have a nice representation of a linear transformation $T : V \to W$ as a matrix, given a basis of each, and we can talk about the image and kernel, and so on. But right off the bat in the infinite case, there's a problem: unlike in finite dimensions, a linear function need not even be continuous!

Definition 4.1. A linear transformation $T : V \to W$ between Banach spaces is *continuous at $v \in V$* if for every convergent sequence $\{v_n\}_{n=1}^\infty$ approaching v,

$$\lim_{n\to\infty} T(v_n) = T(v).$$

T is *continuous* if it is continuous at all points.

This is the same definition of continuous that we have for functions on $\mathbb{R}$.

Example 4.2. Consider the linear map $\lambda : \ell_1 \to \mathbb{R}$ given by $\lambda(a_1, a_2, \dots) = \sum_{i=1}^\infty a_i \log i$. Note that for all $(a_1, a_2, \dots) \in \ell_1$, λ is finite.

However, the sequence $\{v_n\}$ where v_n has $\frac{1}{\log n}$ in the nth spot and 0s elsewhere converges to 0, but $\lambda(v_n)$ has a 1 in the nth spot and 0s elsewhere, which is always distance 1 away from 0.

Example 4.3. Let $\{v_i\}_{i\in I}$ be a Hamel basis for an infinite-dimensional Banach space V. By replacing v_i with $\frac{v_i}{||v_i||}$ if necessary, we can assume that $||v_i|| = 1$ for all i. Thus I is uncountable. Let $J = \{j_1, j_2, \dots\}$ be a countable subset of I. Define $\lambda : V \to \mathbb{R}$ on the Hamel basis by the condition that $\lambda(v_{j_n}) = n$ for $n \in \mathbb{N}$ and $\lambda(v_i) = 0$ if $i \notin J$. We can then extend λ linearly to all of V, since each nonzero vector $v \in V$ has a unique representation as a finite linear combination of Hamel basis vectors.

Now consider the sequence $\{w_n\}_{n=1}^\infty$ of vectors where $w_n = \frac{1}{n} v_{j_n}$. Then $\lambda(w_n) = \frac{1}{n}\lambda(v_{j_n}) = \frac{1}{n} \cdot n = 1$ for all n, but $||w_n|| = \frac{1}{n} \to 0$, so $w_n \to 0$ as $n \to \infty$. Thus λ is not continuous.

One thing to note about all of these examples is that the linear transformation at hand is in some sense getting big; there has to be some sequence of vectors that the linear transformation scales more and more drastically. This is encapsulated by the following definition:

Definition 4.4. A linear transformation $T : V \to W$ between Banach spaces V and W is *bounded* if there exists a constant $C \geq 0$ such that for all $v \in V$,

$$||Tv||_W \leq C||v||_V.$$

Proposition 4.5. *For a linear transformation $T : V \to W$ between Banach spaces, the following are equivalent:*

(a) T is continuous at some point $v \in V$.

(b) *T is continuous.*
(c) *T is bounded.*

Proof. (a) $\Rightarrow$ (b): Let $\{u_n\}_{n=1}^{\infty}$ be any convergent sequence in V approaching $u \in V$. Then $\{u_n - u + v\}_{n=1}^{\infty}$ approaches $v \in V$. By continuity at v,

$$\begin{aligned} T(v) &= \lim_{n\to\infty} T(u_n - u + v) \\ \Rightarrow T(v) &= \lim_{n\to\infty} T(u_n) - T(u) + T(v) \\ \Rightarrow T(u) &= \lim_{n\to\infty} T(u_n). \end{aligned}$$

Thus T is continuous at u; since this holds for all $u \in V$, T is continuous.

(b) $\Rightarrow$ (c): Assume not. Let $\{u_n\}_{n=1}^{\infty}$ be a sequence of elements of V such that $||u_n||_V = 1$ and $||T(u_n)||_W \geq n$. Then $\left|\left|\frac{1}{n}u_n\right|\right| = \frac{1}{n}$, so $\left\{\frac{1}{n}u_n\right\} \to 0$ as $n \to \infty$, but

$$\left|\left|T\left(\frac{1}{n}u_n\right)\right|\right|_W = \frac{1}{n}||T(u_n)||_W \geq 1,$$

so $\lim_{n\to\infty} T\left(\frac{1}{n}u_n\right) \neq 0$. Thus T is not continuous, a contradiction.

(c) $\Rightarrow$ (a): We will prove that T is continuous at 0. Let $\{v_n\}_{n=1}^{\infty}$ be any sequence with $v_n \to 0$, i.e. $||v_n||_V \to 0$. Then

$$||T(v_n)||_W \leq C||v_n||_V \to 0,$$

so $\lim_{n\to\infty} T(v_n) = 0$, and thus T is continuous at 0. □

So that's fun! Bounded linear maps are really just the same as continuous maps.

5. Linear Functionals and Separability

Let's now turn our attention to one relatively simple case of linear maps, the case of *linear functionals*, or maps $V \to \mathbb{R}$. Since unbounded maps are so misbehaved, in this case as in general, we're going to restrict our view to bounded linear functionals, which have more hope of being either useful or understandable (or both :P). Linear functionals can add and be scaled, so they themselves form a vector space, called the *dual space* and denoted V^*.

Definition 5.1. Let $T : V \to W$ be a bounded linear map. Let $c \geq 0$ be the infimum of all constants $C \geq 0$ such that for all $v \in V$, $||T(v)||_W \leq C||v||_V$. Then c is called the *operator norm* of T.

Exercise 5.2. Under the operator norm, V^* is a Banach space.

Example 5.3 (Linear functionals: finite-dimensional case)**.** Let V be an n-dimensional vector space with basis $\{e_1, \ldots, e_n\}$. Let $\lambda : V \to \mathbb{R}$ be any linear functional; then λ can be represented by a $1\times n$ matrix as a linear transformation. This is just a row vector, which is also an n-dimensional vector space. Moreover, by taking the transpose, row vectors are isomorphic to column vectors, so $V^* \cong V$. As we'll discuss more later, this isomorphism is *not* canonical! It depends on the choice of basis.

Since transposing twice yields the original vector back again, $(V^*)^* \cong V^* \cong V$ as well. Note that the isomorphism $(V^*)^* \cong V$ can actually be made to be basis independent. One way to think of this is that for all $v \in V$, there is a linear map $e_v : V^* \to \mathbb{R}$ given by evaluation at v, i.e. $\lambda \mapsto \lambda(v)$. This correspondence is the desired isomorphism $V \to (V^*)^*$, and we didn't need to pick a basis.

But as always, life is harder on the infinite side of the fence. We're going to look into two examples in depth, to get a grasp of the picture. The first is the case of ℓ_1 and ℓ_∞.

Proposition 5.4. $\ell_1^* = \ell_\infty$.

Proof. We'll first show that $\ell_\infty \subseteq \ell_1^*$. Let $a = (a_1, a_2, \dots) \in \ell_\infty$, so that $\sup_i |a_i| < \infty$. Now for all $v = (v_1, v_2, \dots) \in \ell_1$, we will say that

$$a(v_1, v_2, \dots) = \sum_{i=1}^{\infty} a_i v_i.$$

Note that this sum converges absolutely since the a_i's are bounded and the v_i's converge absolutely. Also,

$$|a(v_1, v_2, \dots)| = \left|\left|\sum_{i=1}^{\infty} a_i v_i\right|\right| \leq \sum_{i=1}^{\infty} |a_i||v_i| \leq \sup_i |a_i| \sum_{i=1}^{\infty} |v_i| = ||a||_\infty ||v||_1.$$

Thus every element a of ℓ_∞ is a bounded linear transformation from $\ell_1 \to \mathbb{R}$, so $\ell_\infty \subseteq \ell_1^*$.

We'd now like to show that every linear functional in ℓ_1^* arises in this manner. Let $\lambda : \ell_1 \to \mathbb{R}$ be a linear functional. Consider $e_n \in \ell_1$, with a 1 in the nth entry and 0s elsewhere. Let $a_n = \lambda(e_n)$. We'll show that $a = (a_1, a_2, \dots)$ is in ℓ_∞ and corresponds to the same linear functional λ that we started with.

$\underline{a \in \ell_\infty}$: Note that $||e_n||_1 = 1$ for all n. Since λ is bounded, there exists C such that $|\lambda(e_n)| \leq C||e_n||_1$ for all n, which implies that $|a_n| \leq C$. Thus $\sup_n |a_n| < \infty$, so $a \in \ell_\infty$.

$\underline{a \text{ corresponds to } \lambda}$: We'd like to show that for an arbitrary element $v = (v_1, v_2, \dots) \in \ell_1$, $\lambda(v) = a(v)$. By linearity,

$$\lambda(v) = \lambda\left(\sum_{n=1}^{\infty} v_n e_n\right) = \sum_{n=1}^{M-1} v_n \lambda(e_n) + \lambda\left(\sum_{n=M}^{\infty} v_n e_n\right).$$

Since λ is continuous and $|\sum_{n=M}^{\infty} |v_n|| \to 0$ as $M \to \infty$ we can take the limit as $M \to \infty$ on both sides, yielding

$$\lambda(v) = \lim_{M\to\infty} \sum_{n=1}^{M-1} v_n \lambda(e_n) + \lambda\left(\sum_{n=M}^{\infty} v_n e_n\right) = \sum_{n=1}^{\infty} v_n \lambda(e_n) = a(v).$$

□

The question then remains if ℓ_∞ dual is also ℓ_1. The answer (spoiler alert!) is no, but to see why, we'll need separability.

Definition 5.5. A subset S of a Banach space V is *dense* if for all $v \in V$ and for all $\varepsilon > 0$, there exists $s \in S$ such that $||v - s|| < \varepsilon$.

So for example $\mathbb{Q}$ is dense in $\mathbb{R}$, points with rational coordinates are dense in general, and so on.

Definition 5.6. A Banach space V is *separable* if it has a countable dense subset S.

Example 5.7.
- $\mathbb{Q}$ is countable, as is $\mathbb{Q}^n$, so $\mathbb{R}$ is separable, as is $\mathbb{R}^n$.
- Any Banach space with a basis (or countable basis) is separable. If V has basis $\{v_n\}_{n=1}^{\infty}$, then
$$\bigcup_{n=1}^{\infty}\left\{\sum_{i=1}^{n} a_i v_i \mid a_i \in \mathbb{Q}\right\}$$
is a countable dense subset of V.
- Consider the set of sequences which have rational coordinates and only finitely many nonzero coordinates. These are dense in ℓ_1 and in ℓ_2. Also, this is a union of countably many countable sets, so it too is countable; thus ℓ_1 and ℓ_2 are separable.
- ℓ_∞ is *not* separable. The set from the previous example isn't dense anymore, since we can take the sequence of all 1's in ℓ_∞; this distance from this sequence to everything that has finitely many nonzero coordinates is at least 1. Moreover, for every subset $J \subseteq \mathbb{N}$, ℓ_∞ contains an indicator function for J, i.e. an element e_J that is 1 on coordinates in J and 0 on coordinates not in J. The distance between any two of these indicator elements e_J and e_K is $||e_J - e_K||_\infty = 1$. Thus by the triangle inequality, no element $a \in \ell_\infty$ can satisfy $||a - e_J||_\infty < 1/3$ and $||a - e_K||_\infty < 1/3$ for K and J distinct subsets of $\mathbb{N}$. Thus any dense subset of ℓ_∞ must have at least as many elements as the power set of the naturals—which is already uncountable!

Exercise 5.8. If V is a separable Banach space and $S \subseteq V$ is a subset where we consider the same distance function, S is also separable.

Black Box 5.9 (Hahn-Banach)**.** Let V be a Banach space with $Z \subseteq V$ a subspace. Let $v \in V$ be an element such that
$$\inf_{z \in Z} ||v - z||_V = d.$$
Then there exists $\lambda \in V^*$ such that $||\lambda||_{V^*} \leq 1$, $\lambda(y) = d$, and $\lambda(z) = 0$ for all $z \in Z$.

The Hahn-Banach theorem (usually in a different form) is very useful and not hard to prove, but it involves some precise details that we won't go into here. For a proof, see handout.

Theorem 5.10. *Let V be a Banach space. Assume that V^* is separable. Then V is separable.*

Proof. Consider the unit sphere $S^* \subseteq V^*$, which is separable by the exercise. Let $\{\lambda_n\}_{n=1}^{\infty}$ be a countable dense subset of S^*. Thus for each λ_n, the operator norm of λ_n is 1, so there must exist some $z_n \in V$ with $||z_n||_V = 1$ and $|\lambda_n(z_n)| > 1/2$. Let $D \subseteq V$ be the Banach space with basis $\{z_n\}_{n=1}^{\infty}$ and the same norm as V. Since D has a countable basis, it is separable.

Claim: $D = V$.

Proof of claim. Assume not. Then there exists $y \in V \setminus D$; note that since D is a Banach subspace, this means that y has positive distance from D, i.e. $\inf_{z \in D} ||y - z||_V = d > 0$. By our **black box**, there exists $\lambda \in V^*$ with $||\lambda||_{V^*} = 1$, $\lambda(y) \neq 0$, but $\lambda(z) = 0$ for all $z \in D$. Since the λ_n are dense in S^* and $||\lambda||_{V^*} = 1$, there exists a subsequence $\{\lambda_{n_k}\}_k$ that converge to λ in V^*.

However,

$$||\lambda - \lambda_{n_k}||_{V^*} \geq |(\lambda - \lambda_{n_k})(z_{n_k})| = |\lambda_{n_k}(z_{n_k})| \geq \frac{1}{2}.$$

Since $||\lambda - \lambda_{n_k}||_{V^*} \geq \frac{1}{2}$, $||\lambda - \lambda_{n_k}||_{V^*}$ cannot approach 0, contradicting the fact that $\lambda_{n_k} \to \lambda$ as $k \to \infty$. □

□

Now that we have this theorem, we can come back to the dual space question. Since ℓ_1 is separable but ℓ_∞ is not, $\ell_\infty^* \neq \ell_1$; it's something bigger.[1]

6. Hilbert spaces and Riesz Representation

We'll now turn our attention to ℓ_2, which works very differently. Spoiler alert: here "very differently" means "much better"; this class has secretly been a long-form advertisement for the magic of ℓ_2. Why is ℓ_2 so special? It has an extra structure that we can exploit that most Banach spaces don't have.

Definition 6.1. A vector space V is an *inner product space* if there exists a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ satisfying the following for all $u, v, w \in V$ and for all scalars α:

- $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0 \iff x = 0$ (positive definiteness)
- $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$ (linearity)
- $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$ (scaling)
- $\langle x, y \rangle = \langle y, x \rangle$.

Example 6.2. Let V be any finite dimensional vector space with a bases $\{e_1, \ldots, e_n\}$. Then taking standard dot products in this basis is an inner product.

Fact 6.3. If V is an inner product space with inner product $\langle \cdot, \cdot \rangle$, then V is also a normed space, with norm $||v|| = \langle v, v \rangle^{1/2}$.

Example 6.4. Consider ℓ_2, with the inner product that

$$\langle (a_1, a_2, \ldots), (b_1, b_2, \ldots) \rangle = \sum_{i=1}^{\infty} a_i b_i.$$

Note that this is always finite; under this inner product, the single components as vectors are orthogonal, so by Cauchy-Schwarz

$$\left| \sum_{i=1}^{\infty} a_i b_i \right| \leq \left(\sum_{i=1}^{\infty} |a_i|^2 \right)^{1/2} \left(\sum_{i=1}^{\infty} |b_i|^2 \right)^{1/2} < \infty.$$

Also, the norm resulting from this inner product is exactly the ℓ_2 norm back again.

[1]The situation is summarized by the following brief poem:

ℓ_1 dual is ℓ_∞
ℓ_∞ is too big
ℓ_∞ is not separable
ℓ_∞, what a pig!

Credit to Hannah Larson (MC'12) and myself.

So ℓ_2 is special because it's a Banach space with an inner product, which is also called a *Hilbert space*. What does this tell us about the dual space in this case? For each $v \in V$, where V is a Hilbert space, there is a linear functional $\lambda_v : V \to \mathbb{R}$ given by $\lambda_v(w) = \langle v, w \rangle$. However, we ultimately get that that's everything!

Theorem 6.5 (Riesz Representation). *For each $\lambda \in V^*$, there exists $w_\lambda \in V$ such that for all $v \in V$, $\lambda(v) = \langle w_\lambda, v \rangle$. Moreover, $||w_\lambda||_V = ||\lambda||_{V^*}$.*

So, this says that as normed spaces, $V \cong V^*$ canonically when V is a Hilbert space. Note that choosing a Hilbert space structure on V is akin to choosing a basis in the finite-dimensional case, so this is an exact generalization of what's happening in the finite-dimensional universe. To prove this theorem, we first define orthogonal complement and discuss two lemmas.

Definition 6.6. Let V be a Hilbert space and $M \subseteq V$ be a subspace that is complete under the norm. The *orthogonal complement* $M^\perp$ of M is the set

$$M^\perp = \{w \in V \mid \langle u, w \rangle = 0 \forall u \in M\}.$$

Exercise 6.7. $M^\perp$ is also complete under the norm, i.e. it contains its limit points.

Ideally we'd like to say that $V = M + M^\perp = \{u + w \mid u \in M, w \in M^\perp\}$. That is given to us by the following three lemmas.

Lemma 6.8 (Parallelogram Rule). *Let V be a Hilbert space. Then for all $x, y \in V$,*

$$||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2.$$

Proof.

$$\begin{aligned} ||x+y||^2 + ||x-y||^2 &= \langle x+y, x+y \rangle + \langle x-y, x-y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle + \langle x, x \rangle - \langle y, x \rangle - \langle x, y \rangle + \langle y, y \rangle \\ &= 2\langle x, x, \rangle + 2\langle y, y \rangle = 2||x||^2 + 2||y||^2. \end{aligned}$$

□

Lemma 6.9. *Let V be a Hilbert space, $M \subseteq V$ a closed subspace, and let $v \in V$. Then there exists a unique element $u \in M$ that is closest to v.*

Proof. Let $d = \inf_{y \in M} ||v - y||_V$, and choose a sequence $\{y_n\}_{n=1}^\infty$ of points in M so that $||x - y_n||_V \to d$. Then for $n < m$,

$$\begin{aligned} ||y_n - y_m||_V^2 &= ||(v - y_m) - (v - y_n)||_V^2 \\ &= 2||v - y_m||_V^2 + 2||v - y_n||_V^2 - ||(v - y_m) + (v - y_n)||_V^2, \text{by the Parallelogram Rule} \\ &= 2||v - y_m||_V^2 + 2||v - y_n||_V^2 - 4||v - \tfrac{1}{2}(y_m + y_n)||_V^2 \\ &\leq 2||v - y_m||_V^2 + 2||v - y_n||_V^2 - 4d^2, \text{since } \tfrac{1}{2}(y_m + y_n) \in M. \end{aligned}$$

Taking the limit as $n \to \infty$, with the assumption $n < m$, the right hand side becomes $2d^2 + 2d^2 - 4d^2 = 0$, so the sequence $\{y_n\}$ must be a Cauchy sequence, so it converges to an element $u \in M$ by completeness of M. If there were another point u', we could take our sequence to alternate between u and u', showing that $u = u'$. □

Lemma 6.10 (Projection)**.** *Let V be a Hilbert space with closed subspace M. Every $v \in V$ can be written $v = u + w$ with $u \in M$, $w \in M^{\perp}$.*

Proof. Let $v \in V$. By the lemma, there exists a unique element $u \in M$ minimizing distance to v. Define $w = v - u$, so $v = u + w$. Let $d = ||v - w||_V$; then for all $t > 0$, for all $y \in M$,

$$\begin{aligned} d^2 &\leq ||v - (u + ty)||_V^2 = ||w - ty||_V^2 = \langle w - ty, w - ty \rangle \\ &= d^2 - 2t\langle w, y\rangle + t^2 ||y||_V^2. \end{aligned}$$

Thus for all $t > 0$, $-2t\langle w, y\rangle + t^2||y||_V^2 > 0$, or equivalently $t||y||_V^2 > 2\langle w, y\rangle$. We can pick t arbitrarily small, so this must imply that $\langle w, y\rangle = 0$ □

Proof of Riesz. Fix $\lambda \in V^*$. Let $N \subseteq V$ be the set of $v \in V$ with $\lambda(v) = 0$. Since λ is continuous, the preimage of $\{0\}$, which is N, is closed. If $N = V$, then λ is the 0 map, so $w_\lambda = 0$. If not, consider $N^\perp$, which has a nonzero vector w in it by the projection theorem. By scaling, we can assume that $||w||_V^2 = 1$. Then define $w_\lambda = \lambda(w)w$.

To verify that w_λ has the right properties, note first of all that w and N span all of V, since every $v \in V$ can be written

$$v = \left(v - \frac{\lambda(v)}{\lambda(w)}w\right) + \frac{\lambda(v)}{\lambda(w)}w,$$

and

$$\lambda\left(v - \frac{\lambda(v)}{\lambda(w)}w\right) = \lambda(v) - \frac{\lambda(v)}{\lambda(w)}\lambda(w) = 0,$$

so the first component is in N. Moreover, for all v, and noting that w_λ is perpendicular to N,

$$\begin{aligned} \langle v, w_\lambda\rangle &= \left\langle \left(v - \frac{\lambda(v)}{\lambda(w)}w\right) + \frac{\lambda(v)}{\lambda(w)}w, \lambda(w)w \right\rangle \\ &= \lambda(w)\left\langle \left(v - \frac{\lambda(v)}{\lambda(w)}w\right), w\right\rangle + \langle \frac{\lambda(v)}{\lambda(w)}w, \lambda(w)w\rangle \\ &= 0 + \frac{\lambda(v)}{\lambda(w)}\lambda(w) = \lambda(v). \end{aligned}$$

To prove that $||\lambda||_{V^*} = ||w_\lambda||_V$, note that

$$||\lambda||_{V^*} = \sup_{||v||_V = 1} |\lambda(v)| = \sup_{||v||_V=1} |\langle v, w_\lambda\rangle| \leq \sup_{||v||_V=1} ||w_\lambda||_V ||v||_V = ||w_\lambda||_V$$

by Cauchy-Schwarz, and moreover that

$$||\lambda||_{V^*} = \sup_{||v||_V=1} |\lambda(v)| \geq |\lambda(w)| = \langle w, w_\lambda\rangle = ||w_\lambda||_V,$$

so equality holds. □

This completes our proof! The self-dual property of finite dimensional spaces are really a property of Hilbert spaces in all of their shining glory.