April 22, 2021

Last time, we discussed orthonormal bases, considering the concrete question of whether complex exponentials formed an orthonormal basis for $L^2([-\pi, \pi])$. Today, we'll go back to a general discussion of Hilbert spaces, and the rest of the course from here on will be general theory and some concrete applications to particular problems.

Our first topic today will be **length minimizers**: recall that we can describe a norm on V/W for subspaces of a normed vector space, and we did so via an infimum. It makes sense to ask whether this minimal distance is actually achieved:

> **Theorem 178**
>
> Let C be a nonempty closed subset of a Hilbert space H which is **convex**, meaning that for all $v_1, v_2 \in C$, we have $tv_1 + (1 - t)v_2 \in C$ for all $t \in [0, 1]$. Then there exists a unique element $v \in C$ with $||v|| = \inf_{u \in C} ||u||$ (this is a length minimizer).

The convexity condition can alternatively be stated as "the line segment between any two elements of C is contained in C." And to connect this with our discussion earlier, one such example of a set would be $v + W$ for some closed subspace W of C and some $v \in H$.

Remark 179. *The condition that C is closed is required: for example, we can let C be an open disk outside the origin, in which case the minimum norm is not achieved (because it's on the boundary). And convexity is also required – for example, otherwise we could take the complement of an open disk centered at the origin, in which case the minimum norm is achieved on the entire boundary.*

Proof. We should recall that $a = \inf S$ if and only if a is a lower bound for S, and there exists a sequence $\{s_n\}$ in S with $s_n \to a$. If we let $d = \inf_{u \in C} ||u||$, this is some finite number because norms are always bounded from below by 0, and C is nonempty. So there exists some sequence $\{u_n\}$ in C such that $||u_n|| \to d$.

We claim that this sequence is actually Cauchy. To see that, let $\varepsilon > 0$ – because of convergence of $||u_n||$ to d, there exists some N so that for all $n \geq N$, we have

$$2||u_n||^2 < 2d^2 + \frac{\varepsilon^2}{2}.$$

Then for all $n, m \geq N$, the parallelogram law tells us that

$$||u_m - u_n||^2 = 2||u_m||^2 + 2||u_n||^2 - 4\left|\left|\frac{u_n + u_m}{2}\right|\right|^2,$$

and now because $\frac{u_n + u_m}{2}$ lies on the line segment between u_n and u_m (taking $t = \frac{1}{2}$), convexity tells us that it is also in C. Therefore, $\left|\left|\frac{u_n + u_m}{2}\right|\right|^2 \geq d^2$, and thus

$$||u_m - u_n||^2 \leq 2||u_m||^2 - 2d^2 + 2||u_n||^2 - 2d^2 < \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2} = \varepsilon^2$$

by our choice of N, and taking a square root shows that the sequence $\{u_n\}$ is indeed Cauchy. Because our Hilbert space is complete, this means that the sequence also converges, and thus there is some $v \in H$ such that $u_n \to v$, and $v \in C$ as well because our subset C is closed. So now continuity of the norm tells us that

$$||v|| = \lim_{n \to \infty} ||u_n|| = d,$$

and thus we've found our minimizer $v \in C$. To show uniqueness, suppose that $v, \bar{v}$ are both in C and have norm d.

Then the parallelogram law tells us that

$$||v - \overline{v}||^2 = 2||v||^2 + 2||\overline{v}||^2 - 4\left|\left|\frac{v + \overline{v}}{2}\right|\right|^2 \leq 2d^2 + 2d^2 - 4d^2 = 0,$$

again using that $\frac{v+\overline{v}}{2}$ is also in C by convexity, and thus we must have $v - \overline{v} = 0 \implies v = \overline{v}$. $\qquad\square$

We'll obtain some important consequences from this result – the first one is how to decompose our Hilbert space using a closed linear subspace, much like we usually like to do in $\mathbb{R}^n$ and $\mathbb{C}^n$.

Theorem 180

Let H be a Hilbert space, and let $W \subset H$ be a subspace. Then the **orthogonal complement**

$$W^\perp = \{u \in H : \langle u, w \rangle = 0 \quad \forall w \in W\}$$

is a closed linear subspace of H. Furthermore, if W is closed, then $H = W \oplus W^\perp$; in other words, for all $u \in H$, we can write $u = w + w^\perp$ for some unique $w \in W$ and $w^\perp \in W^\perp$.)

A picture to keep in mind is the case where H is $\mathbb{R}^2$ and W is the x-axis – then $W^\perp$ would be the y-axis, and we're saying that all elements can be broken up into a component along the x-axis and a component along the y-axis.

Proof. Showing that $W^\perp$ is a subspace is clear, because if $\langle u_1, w \rangle = 0$ and $\langle u_2, w \rangle = 0$ for all $w \in W$, any linear combination of u_1 and u_2 will also be orthogonal to all $w \in W$ by linearity of the inner product. Furthermore, $W \cap W^\perp = \{0\}$, because any element $w \in W$ that is also in $W^\perp$ must satisfy $\langle w, w \rangle = 0 \implies w = 0$.

To show that $W^\perp$ is closed, let $\{u_n\}$ be a sequence in $W^\perp$ converging to $u \in H$. We wish to show that $\langle u, w \rangle = 0$ for all $w \in W$, so that $u \in W^\perp$ as well. Indeed, by continuity of the inner product, we have

$$\langle u, w \rangle = \lim_{n \to \infty} \langle u_n, w \rangle = \lim_{n \to \infty} 0 = 0,$$

so that our sequential limit is also in our subspace $W^\perp$.

It remains to show that $H = W \oplus W^\perp$ if W is closed. The result is clear for $W = H$, since $W^\perp = \{0\}$ and $H = H \oplus \{0\}$ is a trivial decomposition. Otherwise, if $W \neq H$, then let $u \in H \setminus W$ (that is, u is in H but not W), and define the set

$$C = u + W = \{u + w : w \in W\}.$$

This set C is nonempty because it contains u, and it is convex because for any two elements $u + w_1, u + w_2 \in C$ (for $w_1, w_2 \in W$) and for any $t \in [0, 1]$, we have

$$t(u + w_1) + (1 - t)(u + w_2) = (t + (1 - t))u + tw_1 + (1 - t)w_2 = u + (tw_1 + (1 - t)w_2)$$

and the last term is in W because subspaces are closed under linear combinations. So we now need to show that C is closed: indeed, if $u + w_n$ is a sequence of elements in C that converge to some element $v \in H$, we know that

$$u + w_n \to v \implies w_n \to v - u,$$

and because W is closed, w_n must converge to some element in W. Thus $v - u \in W$, and thus $v = u + w$ for some $w \in W$, which is exactly the definition of being in C. So C is indeed closed.

So returning to the problem, if we want to write an element u of H as a sum of a part in W and a part in $W^\perp$, it makes sense that our component in $W^\perp$ will be the minimizer to C (keeping the $\mathbb{R}^2$ example from above in mind). So

applying Theorem 178, because C is closed and convex, there is some unique $v \in C$ with

$$||v|| = \inf_{c \in C} ||c|| = \inf_{w \in W} ||u + w||.$$

. Since $v \in C$, we know that $u - v \in W$, so we will write $u = (u - v) + v$. Our goal is to show that $v \in W^{\perp}$, and we do this with a variational argument (in physics, this is the Euler-Lagrange equations, and it is another way of phrasing properties of the infimum). If $w \in W$, define the function

$$f(t) = ||v + tw||^2 = ||v||^2 + t^2||w||^2 + 2t\text{Re}\langle v, w \rangle,$$

which is a polynomial in t. We know that $f(t)$ has a minimum at $t = 0$, because all elements of the form $v + tw$ are in C, and we know the minimizer of norm uniquely occurs at v. So $f'(0) = 0$, and thus

$$2\text{Re}\langle v, w \rangle = 0.$$

So the real part of the inner product is zero, and now we can repeat this argument but with $||v + itw||$ instead of $||v + tw||$, which will show us that

$$\text{Re}\langle v, iw \rangle = \text{Im}\langle v, w \rangle = 0.$$

Therefore, $\langle v, w \rangle = 0$, and since this argument was true for all $w \in W$, we must have $v \in W^{\perp}$. It remains to show that this decomposition is unique, and this is true because $W \cap W^{\perp} = \{0\}$: more specifically, if $u = w_1 + w_1^{\perp} = w_2 + w_2^{\perp}$, that means that $w_1 - w_2 = w_2^{\perp} - w_1^{\perp}$ is in both W and $W^{\perp}$, and thus both sides of this equation are 0. So $w_1 = w_2$ and $w_1^{\perp} = w_2^{\perp}$, showing uniqueness. □

The following result is left as an exercise for us:

> **Theorem 181**
>
> If $W \subset H$ is a subspace, then $(W^{\perp})^{\perp}$ is the closure $\overline{W}$ of W. In particular, if W is closed, then $(W^{\perp})^{\perp} = W$.

Now that we have this decomposition $u = w + w^{\perp}$ for our subspace W, we can construct a map which takes in u and outputs w. If we use the $\mathbb{R}^2$ example from above, we can see that this map is a projection onto the x-axis, and more generally we can make the following definition:

> **Definition 182**
>
> Let $P : H \to H$ be a bounded linear operator. Then P is a **projection** if $P^2 = P$.

> **Proposition 183**
>
> Let H be a Hilbert space, and let $W \subset H$ be a closed subspace. Then the map $\Pi_W : H \to H$ sending $v = w + w^{\perp}$ (for $w \in W, w^{\perp} = W^{\perp}$) to w is a projection operator.

Proof. First, we show that Π_W is linear. Indeed, if $v_1 = w_1 + w_1^{\perp}$ and $v_2 = w_2 + w_2^{\perp}$, and we have $\lambda_1, \lambda_2 \in \mathbb{C}$, then

$$\lambda_1 v_1 + \lambda_2 v_2 = (\lambda_1 w_1 + \lambda_2 w_2) + (\lambda_1 w_1^{\perp} + \lambda_2 w_2^{\perp}).$$

The two terms on the right-hand side are in W and $W^{\perp}$, respectively, by closure of subspaces under linear combinations. So $\Pi_W(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 w_1 + \lambda_2 w_2$, which is indeed $\lambda_1 \Pi_W(v_1) + \lambda_2 \Pi_W(v_2)$, as desired. We can also see that Π_W is

bounded, because when $v = w + w^\perp$,

$$||v||^2 = ||w + w^\perp||^2 = ||w||^2 + ||w_\perp||^2 \geq ||w||^2$$

(since the inner product cross term is zero when $\langle w, w^\perp \rangle = 0$). Therefore, $||\Pi_W(v)|| \leq ||v||$, and the operator norm is at most 1. And now we just need to check that $\Pi_W^2 = \Pi_W$: if $v = w + w^\perp$, then $\Pi_W(v) = w$, and then

$$\Pi_W(\Pi_W(v)) = \Pi_W(w) = w = \Pi_W(v),$$

and since this is true for all v, we have $\Pi_W^2 = \Pi_W$, as desired. $\square$

Our next application of length minimizers will be the following important result:

Theorem 184 (Riesz Representation Theorem)

Let H be a Hilbert space. Then for all $f \in H'$, there exists a unique $v \in H$ so that $f(u) = \langle u, v \rangle$ for all $u \in H$.

In other words, every element of the dual can be realized as an inner product with a fixed vector. We've seen something similar before when we proved that the dual of ℓ^p is identified with ℓ^q (for $\frac{1}{p} + \frac{1}{q} = 2$) via a pairing, and the $p = q = 2$ case is the example relevant to Hilbert spaces.

Proof. If such a v exists, it is unique, because $f(u) = \langle u, v \rangle = \langle u, \tilde{v} \rangle = 0$, then $\langle u, v - \overline{v} \rangle = 0$ for all $u \in H$. Setting $u = v - \overline{v}$ tells us that $v - \overline{v} = 0$. So we just need to construct such a v that works.

The easiest case is $f = 0$, because in that case, we take $v = 0$. Otherwise, there exists some $u_1 \in H$ so that $f(u_1) \neq 0$, and we take $u_0 = \frac{u_1}{f(u_1)}$ so that $f(u_0) = 1$. We can then define the nonempty set

$$C = \{u \in H : f(u) = 1\} = f^{-1}(\{1\}),$$

which is closed because f is a continuous function, $\{1\}$ only has one element so is closed, and the preimage of a closed set by a continuous function is a closed set. We claim that C is convex: indeed, if $u_1, u_2 \in C$ and $t \in [0, 1]$, then

$$f(tu_1 + (1 - t)u_2) = tf(u_1) + (1 - t)f(u_2) = t \cdot 1 + (1 - t) \cdot 1 = 1,$$

so that $tu_1 + (1 - t)u_2$ is also in C. So now by Theorem 178, there exists $v_0 \in C$ so that $v_0 = \inf_{u \in C} ||u||$, and we define $v = \frac{v_0}{||v_0||^2}$ (noting that $v_0 \neq 0$ because the infimum is not 0).

We claim that this is the v that we want; in other words, let's check that $f(u) = \langle u, v \rangle$. Indeed, if we let $N = f^{-1}(\{0\}) = \{w \in H : f(w) = 0\}$ be the nullspace of f, then we can check that $C = \{v_0 + w : w \in N\}$ and that $||v_0|| = \inf_{w \in N} ||v_0 + w||$. So by the argument that we made earlier in Theorem 180 using $||v_0 + tw||^2$, $v_0 \in N^\perp$, and now for any $w \in H$,

$$f(u - f(u)v_0) = f(u) - f(u)f(v_0) = 0$$

by linearity of f, and thus $u = (u - f(u)v_0) + f(u)v_0$ is a sum of a component in N and a component in $N^\perp$.

$$\langle u, v \rangle = \frac{1}{||v_0||^2} \langle u, v_0 \rangle = \frac{1}{||v_0||^2} \left[\langle (u - f(u)v_0), v_0 \rangle + f(u)\langle v_0, v_0 \rangle \right],$$

The first term here has $u - f(u)v_0 \in N$ and $v_0 \in N^\perp$, so that inner product is zero, and we're left with

$$= f(u) \frac{\langle v_0, v_0 \rangle}{||v_0||^2} = f(u),$$

92

as desired. So we've found v (a scaled version of the minimizer) so that $f(u) = \langle u, v \rangle$ for all u, concluding the proof. $\qquad\square$

We'll study adjoint operators next time – we defined it as a map from dual spaces to dual spaces, but because we can identify dual spaces of Hilbert spaces with themselves, adjoint operators will be essentially regular operators, and we'll soon see how they relate to solving equations on Hilbert spaces and why they are the analogs of the transpose matrix in finite-dimensional linear algebra as well.

18.102 / 18.1021 Introduction to Functional Analysis
Spring 2021