# April 8, 2021

Last time, we introduced the concept of a **pre-Hilbert space** (a vector space that comes equipped with a Hermitian inner product). This inner product is positive definite, linear in the first argument, and becomes complex conjugated when we swap the two arguments, and we can use this quantity to define

$$||v|| = \langle v, v \rangle^{1/2}$$

for any $v$ in the pre-Hilbert space. And we want to show that this is actually a norm – towards that goal, recall that last time, we showed the Cauchy-Schwarz inequality

$$|\langle u, v \rangle| \leq ||u|| \, ||v||$$

for all $u, v$ in the pre-Hilbert space. We'll now put that to use:

---

**Theorem 143**

If $H$ is a pre-Hilbert space, then $|| \cdot ||$ is a norm on $H$.

---

*Proof.* We need to prove the three properties of the norm. For positive definiteness, note that we do have $||v|| \geq 0$ for all $v$, and

$$||v|| = 0 \iff \langle v, v \rangle = 0 \iff v = 0$$

because the Hermitian inner product is (defined to be) positive definite. Furthermore, for any $v \in H$, we have

$$\langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle \implies ||\lambda v|| = |\lambda| \, ||v||$$

by taking square roots of both sides, which shows homogeneity. So we just need to show the triangle inequality: indeed, if we have $u, v \in H$, then

$$\boxed{||u + v||^2} = \langle u + v, u + v \rangle = ||u||^2 + ||v||^2 + 2\text{Re}(\langle u, v \rangle).$$

Because $\text{Re}(z) \leq ||z||$ and using the Cauchy-Schwarz inequality, this can be bounded by

$$\leq ||u||^2 + ||v||^2 + 2|\langle u, v \rangle| \leq ||u||^2 + ||v||^2 + 2||u|| \, ||v|| = \boxed{(||u|| + ||v||)^2},$$

and now taking square roots of both sides yields the desired inequality. $\square$

The Cauchy-Schwarz inequality can also help us in other ways:

---

**Theorem 144** (Continuity of the inner product)

If $u_n \to u$ and $v_n \to v$ in a pre-Hilbert space equipped with the norm $|| \cdot ||$, then $\langle u_n, v_n \rangle \to \langle u, v \rangle$.

---

*Proof.* Notice that if $u_n \to u$ and $v_n \to v$, that means that $||u_n - u|| \to 0$ and $||v_n - v|| \to 0$ as $n \to \infty$. Therefore, we can bound

$$|\langle u_n, v_n \rangle - \langle u, v \rangle| = |\langle u_n, v_n \rangle - \langle u, v_n \rangle + \langle u, v_n \rangle - \langle u, v \rangle|$$

and factoring and using the triangle inequality for $\mathbb{C}$ gives us

$$= |\langle u_n - u, v_n \rangle + \langle u, v_n - v \rangle| \leq |\langle u_n - u, v_n \rangle| + |\langle u, v_n - v \rangle|.$$

The Cauchy-Schwarz inequality then allows us to bound this by

$$\leq ||u_n - u|| \cdot ||v_n|| + ||u|| \cdot ||v_n - v||,$$

and now because $v_n \to v$ we know that $||v_n|| \to ||v||$, and this convergent sequence of real numbers must be bounded. Thus, our new bound is

$$\leq ||u_n - u|| \cdot \sup_n ||v_n|| + ||u|| \cdot ||v_n - v||,$$

and now because $||u_n - u||, ||v_n - v|| \to 0$, the linear combination of them given above also converges to 0, and we're done. Thus $\langle u_n, v_n \rangle$ indeed converges to $\langle u, v \rangle$ (by the squeeze theorem). □

---

**Definition 145**

A **Hilbert space** is a pre-Hilbert space that is complete with respect to the norm $|| \cdot || = \langle \cdot, \cdot \rangle^{1/2}$.

---

**Example 146**

The space of $n$-tuples of complex numbers $\mathbb{C}^n$ with inner product $\langle \underline{z}, \underline{w} \rangle = \sum_{j=1}^n z_j \overline{w_j}$ is a (finite-dimensional) Hilbert space.

---

**Example 147**

The space $\ell^2 = \{\underline{a} : \sum_n |a_n|^2 < \infty\}$ is a Hilbert space, where we define

$$\langle \underline{a}, \underline{b} \rangle = \sum_{k=1}^{\infty} a_k \overline{b_k}.$$

---

In this latter example, we can check that $\langle a, a \rangle^{1/2}$ coincides with the $\ell^2$ norm $||a||_2$. And it turns out that every **separable** Hilbert space (which are the ones that we'll primarily care about) can be mapped in an isometric way to one of these two examples, so the examples above are basically the two main types of Hilbert spaces we'll often be seeing! But here's another one that we'll see often:

---

**Example 148**

Let $E \subset \mathbb{R}$ be measurable. Then $L^2(E)$, the space of measurable functions $f : E \to \mathbb{C}$ with $\int_E |f|^2 < \infty$, is a Hilbert space with inner product

$$\langle f, g \rangle = \int_E f\overline{g}.$$

---

We might notice that we focused on $\ell^2$ and $L^2$, and that's because the inner product only induces the $\ell_2$ norm in the way that it's written right now. But we might want to ask whether there's an inner product that we could put on the other $\ell^p$ or $L^p$ so that they are also Hilbert spaces (so that we get out the appropriate norm), and the answer turns out to be **no**. We'll see that through the following result:

---

**Proposition 149** (Parallelogram law)

Let $H$ be a pre-Hilbert space. Then for any $u, v \in H$, we have

$$||u + v||^2 + ||u - v||^2 = 2 \left( ||u||^2 + ||v||^2 \right).$$

In addition, if $H$ is a normed vector space satisfying this equality, then $H$ is a pre-Hilbert space.

We can use this result (which can be verified by computation) to see that there are always $u, v$ which make this inequality not satisfied if $p \neq 2$ for the $\ell^p$ and $L^p$ spaces. And now that we have this inner product, we can start doing more work in the "linear algebra" flavor:

---

**Definition 150**

Let $H$ be a pre-Hilbert space. Two elements $u, v \in H$ are **orthogonal** if $\langle u, v \rangle = 0$ (also denoted $u \perp v$), and a subset $\{e_\lambda\}_{\lambda \in \Lambda} \subset H$ is **orthonormal** if $||e_\lambda|| = 1$ for all $\lambda \in \Lambda$ and for all $\lambda_1 \neq \lambda_2$, $\langle e_{\lambda_1}, e_{\lambda_2} \rangle = 0$.

---

**Remark 151.** *We may notice that the index set we're using is some arbitrary set $\Lambda$ instead of $\mathbb{N}$: we'll mainly be interested in the case where we have a finite or countably infinite orthonormal set, but the definition makes sense more generally.*

We'll see some examples corresponding to each of the examples of Hilbert spaces $\mathbb{C}^n, \ell^2, L^2$ that we described above:

---

**Example 152**

The set $\{(0, 1), (1, 0)\}$ is an orthonormal set in $\mathbb{C}^2$, and $\{(0, 0, 1), (0, 1, 0)\}$ is an orthonormal set in $\mathbb{C}^3$.

---

**Example 153**

Let $\underline{e}_n$ be the sequence which is 1 in the $n$th entry and 0 everywhere else, we find that $\{e_n\}_{n \geq 1}$ is an orthonormal subset of $\ell^2$.

---

**Example 154**

The functions $f_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$ (as elements of $L^2([-\pi, \pi])$) form an orthonormal subset of $L^2([-\pi, \pi])$. (This is because the integral $\int_{-\pi}^{\pi} e^{imx} \overline{e^{inx}} dx = \int_{-\pi}^{\pi} e^{i(m-n)x} dx$ is zero unless $m = n$; if we're uncomfortable integrating a complex exponential, we can break it up into its real and imaginary parts by Euler's formula.)

---

Notice that we haven't talked about whether the spaces $\ell^2$ and $L^2$ are separable, but it was an exercise for us to show that the continuous functions are dense in $L^p$ (for all $p < \infty$) and the Weierstrass approximation tells us that continuous functions on a closed and bounded interval can be uniformly approximated by a polynomial. So the polynomials are dense in $L^p$, and to get to a countable dense subset, we only consider the polynomials with rational coefficients, and there are indeed countably many of those. So for all $L^p$ with $p$ finite, the polynomials with coefficients of the form $q_1 + iq_2$ for rational $q_1 + q_2$ form a countable dense subset of $L^p([a, b])$, and thus those $L^p$ spaces are separable. And the set of sequences which terminate after some point form a dense subset in $\ell^p$ for any finite $p$ as well, so we can get our countable dense subset of $\ell^p$ by looking at the set of sequences of rationals that terminate eventually!

---

**Theorem 155** (Bessel)

Let $\{e_n\}$ be a countable (finite or countably infinite) orthonormal subset of a pre-Hilbert space $H$. Then for all $u \in H$, we have

$$\sum_n |\langle u, e_n \rangle|^2 \leq ||u||^2.$$

---

*Proof.* First, we do the finite case. If $\{e_n\}_{n=1}^N$ is a finite collection of orthonormal vectors in $H$, we can verify that

$$\left\|\sum_{n=1}^N \langle u, e_n\rangle e_n\right\|^2 = \left\langle \sum_n \langle u, e_n\rangle e_n, \sum_m \langle u, e_m\rangle e_m\right\rangle,$$

and we can pull out some numbers to write this as

$$= \sum_{n,m} \langle u, e_n\rangle \overline{\langle u, e_m\rangle}\langle e_n, e_m\rangle,$$

By orthonormality, the inner product $\langle e_n, e_m\rangle$ is only nonzero when $n = m$ (in which case it's equal to 1), so that we end up with $\sum_{n=1}^N \langle u, e_n\rangle^2$. We can also say by linearity that

$$\left\langle u, \sum_{n=1}^N \langle u, e_n\rangle.e_n\right\rangle = \sum_{n=1}^N \overline{\langle u, e_n\rangle}\langle u, e_n\rangle = \sum_{n=1}^N |\langle u, e_n\rangle|^2.$$

From here, note that

$$0 \le \left\|u - \sum_{n=1}^N \langle u, e_n\rangle e_n\right\|^2,$$

where the term inside the parentheses can be thought of as **the projection of u onto the orthogonal space to the** $e_i$**s**. We can then rewrite this by expanding in the same way we previously did for $||u + v||^2$, and we get

$$0 \le ||u||^2 + \left\|\sum_{n=1}^N \langle u, e_n\rangle e_n\right\|^2 - 2\,\mathrm{Re}\left\langle u, \sum_{n=1}^N \langle u, e_n\rangle e_n\right\rangle.$$

Both of the last two terms now just give us multiples of $\sum_{n=1}^N |\langle u, e_n\rangle|^2$ by our work above, and we end up with

$$\boxed{0 \le} ||u||^2 + \sum_{n=1}^N |\langle u, e_n\rangle|^2 - 2\sum_{n=1}^N |\langle u, e_n\rangle|^2 = \boxed{||u||^2 - \sum_{n=1}^N |\langle u, e_n\rangle|^2},$$

and this is exactly what we want to show for the finite case. And the infinite case follows by taking $N \to \infty$: more formally, if $\{e_n\}$ is an orthonormal subset of $H$, then

$$\sum_{n=1}^N |\langle u, e_n\rangle|^2 \le ||u||^2 \implies \lim_{N\to\infty} \sum_{n=1}^N |\langle u, e_n\rangle|^2 \le ||u||^2,$$

and this proves the result that we want for all countable orthonormal subsets of $H$. $\qquad\square$

Orthonormal sets are not the most useful thing on their own for studying a pre-Hilbert space $H$, since we might leave out some vectors in our span. That motivates this next definition:

**Definition 156**

An orthonormal subset $\{e_\lambda\}_{\lambda\in\Lambda}$ of a pre-Hilbert space $H$ is **maximal** if the only vector $u \in H$ satisfying $\langle u, e_\lambda\rangle = 0$ for all $\lambda \in \Lambda$ is $u = 0$.

**Example 157**

The $n$ standard basis vectors in $\mathbb{C}^n$ form a maximal orthonormal subset. (A non-example would be any proper subset of that set.)

We'll soon see that a countably infinite maximal orthonormal subset basically serves the same purpose as an orthonormal basis does in linear algebra, but not every element will be able to be written as a **finite** linear combination of the orthonormal subset elements (like was possible with a Hamel basis).

> **Theorem 159**
>
> Every nontrivial pre-Hilbert space has a maximal orthonormal subset.

We can prove this result by using Zorn's lemma and thinking of subsets as being ordered by inclusion. But if that scares us (because of the use of the Axiom of Choice), we can do a slightly less strong proof by hand:

> **Theorem 160**
>
> Every nontrivial **separable** pre-Hilbert space $H$ has a **countable** maximal orthonormal subset.

(Recall that a space is **separable** if it has a countable dense subset.)

*Proof.* We'll use the **Gram–Schmidt process** from linear algebra as follows. Because $H$ is separable, we can let $\{v_j\}_{j=1}^{\infty}$ be a countable dense subset of $H$ such that $||v_1|| \neq 0$.

We claim that for all $n \in \mathbb{N}$, there exists a natural number $m(n)$ and an orthonormal subset $\{e_1, \cdots, e_{m(n)}\}$ so that the span of this subset is the span of $\{v_1, \cdots, v_n\}$, and $\{e_1, \cdots, e_{m(n+1)}\}$ is the union of $\{e_1, \cdots, e_{m(n)}\}$ and either the empty set (if $v_{n+1}$ is already in the span) or some vector $e_{m(n+1)}$ (otherwise). In other words, we can come up with a finite orthonormal subset that has the same span as the first $n$ vectors of our countable dense subsets, and we can keep constructing this iteratively by adding at most one element.

We'll prove this claim by induction. For the base case $n = 1$, we can take $e_1 = \frac{v_1}{||v_1||}$, which indeed satisfies all of the properties we want. Now for the inductive step, suppose that our claim holds for $n = k$, and now we want to span $v_1$ through $v_{k+1}$ instead of just $v_1$ through $v_k$. If $v_{k+1}$ is already in the span of $\{v_1, \cdots, v_k\}$, then the span of $\{e_1, \cdots, e_{m(k)}\}$ is the same as the span of $\{v_1, \cdots, v_k\}$, which is the same as the span of $\{v_1, \cdots, v_{k+1}\}$. So in this case, we don't need to add anything, and all of our conditions are still satisfied. Otherwise, $v_{k+1}$ is not in the span of $\{v_1, \cdots, v_k\}$, and we'll define

$$w_{k+1} - \sum_{j=1}^{m(k)} \langle v_{k+1}, e_j \rangle e_j$$

to be $v_{k+1}$ with components along the other $v_i$s subtracted off. This vector is not zero, or else $v_{k+1}$ would be in the span of the existing $e_i$s and thus in the span of the existing $v_i$s. We then define the normalized version $e_{m(k+1)} = \frac{w_{k+1}}{||w_{k+1}||}$ to add to our orthonormal subset: this is a unit vector by construction, and for any $1 \leq \ell \leq k$ we can indeed check orthogonality:

$$\langle e_{m(k+1)}, e_\ell \rangle = \frac{1}{||m_{k+1}||} \left\langle v_{k+1} - \sum_{j=1}^{m(k)} \langle v_{k+1}, e_j \rangle e_j, e_\ell \right\rangle$$

now simplifies because the first $m(k)$ $e$'s are already orthonormal: we just pick out $j = \ell$ from the sum and we have

$$= \frac{1}{||m_{k+1}||} \left( \langle v_{k+1}, v_\ell \rangle - \langle v_{k+1}, e_\ell \rangle \right) = 0.$$

Therefore, $\{e_1, \cdots, e_{m(k)}, e_{m(k+1)}\}$ is indeed an orthonormal subset, and that proves the desired claim.

It now remains to show that the collection of all $e_\ell$s forms a maximal orthonormal subset. We define the set

$$S = \bigcup_{n=1}^{\infty} \{e_1, \cdots, e_{m(n)}\};$$

this is an orthonormal subset of $H$ which can be finite or countably infinite, and we want to show that $S$ is maximal. And here is where we use the fact that the $v_i$s are dense in $H$: suppose that we have some $u \in H$ so that $\langle u, e_\ell \rangle = 0$ for all $\ell$. Then we can find a sequence of elements $\{v_{j(k)}\}_k$ such that

$$\lim_{k \to \infty} v_{j(k)} \to u.$$

Because the span of the $v_i$s and the $e_i$s are the same, we know that each $v_{j(k)}$ is in the span of $\{e_1, \cdots, e_{m(j(k))}\}$, so now

$$\boxed{||v_{j(k)}||^2} = \sum_{\ell=1}^{m(j(k))} |\langle v_{j(k)}, e_\ell \rangle|^2$$

(we have equality for a finite set of such orthonormal elements), and now we can rewrite this as

$$= \sum_{\ell=1}^{m(j(k))} |\langle v_{j(k)} - u, e_\ell \rangle|^2 \leq \boxed{||v_{j(k)} - u||^2}$$

by Bessel's inequality. But because $v_{j(k)} \to u$ by construction, this means that $||v_{j(k)}|| \to 0$, and thus the limit of the $v_{j(k)}$s, which is $u$, must be zero. That proves that our orthonormal basis is indeed maximal. $\qquad \square$

Next time, we'll understand more specifically what it means for these maximal orthonormal subsets to serve as replacements for bases from linear algebra!

18.102 / 18.1021 Introduction to Functional Analysis
Spring 2021