April 29, 2021

We'll continue discussing compactness today – recall that a subset K of a metric space X is **compact** if every sequence $\{x_n\}_n$ in K has a subsequence that is convergent in K. While being closed and bounded is equivalent to being compact in $\mathbb{R}^n$, this is not true in general Hilbert spaces (for example, take the orthonormal basis vectors in ℓ^2). So we need an additional condition – last time, we proved that if we have a convergent sequence $\{v_n\}_n$ in H converging to v, then the subset $K = \{v_n : n \in \mathbb{N}\} \cup \{v\}$ is compact, and it has **equi-small tails** with respect to any orthonormal subset. Here, the definition is that if $\{e_k\}_k$ is a countable orthonormal subset of H, then for all $\varepsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $\tilde{v} \in K$ (either an element of the sequence or v), we have

$$\sum_{k>N} |\langle \tilde{v}, e_k \rangle|^2 < \varepsilon^2.$$

(We know that this sum over all k is bounded, and thus convergent, by Bessel's inequality for any individual $\tilde{v}$, so we can always find an N that makes this work for a fixed $\tilde{v}$, but the condition requires it simultaneously for all $\tilde{v} \in K$. So we can think of "equi-small tails" as really meaning "uniformly small tails.") It turns out that this condition suffices (and is necessary) for compactness:

> **Theorem 197**
>
> Let H be a separable Hilbert space, and let $\{e_k\}_k$ be an orthonormal basis of H. Then a subset $K \subset H$ is compact if and only if K is closed, bounded, and has equi-small tails with respect to $\{e_k\}$.

Proof. For the forward direction, first suppose that K is compact. We know by general metric space theory that K is then closed and bounded, and we'll show that K has equi-small tails with respect to $\{e_k\}$ by contradiction. Suppose otherwise: then there exists some ε_0 such that for each natural N, there is some $u_N \in K$ such that

$$\sum_{k>N} |\langle u_N, e_k \rangle|^2 \geq \varepsilon_0^2.$$

This then gives us a sequence $\{u_n\}$ (by picking such a u_N for every natural number N), and thus by the assumption of compactness, there is some subsequence $\{v_m\} = \{u_{n_m}\}$ and some $v \in K$ such that $v_m \to v$. But we also know that for all $n \in \mathbb{N}$, $\sum_{k>n} |\langle v_n, e_k \rangle| \geq \varepsilon_0^2$, because $v_m = u_{n_m}$ is the nth or later term of the original sequence (so summing over $k > n_m$ is at most the value we get summing over $k > n$). That means that the subset $\{v_n : n \in \mathbb{N}\} \cup \{v\}$ does not have equi-small tails, which is a contradiction of our previous theorem. So if K is compact, then it must have equi-small tails, as desired.

On the other hand, suppose K is closed, bounded, and has equi-small tails. We wish to show that any sequence $\{u_n\}$ has a convergent subsequence in K. Because K is closed, any sequence that converges will converge in K, so we just need to show that there is some convergent subsequence. We know that any bounded sequence of complex numbers has a convergent subsequence (showing convergence of the real and imaginary parts by Bolzano-Weierstrass), so the plan is to expand $\{u_n\}$ in terms of the orthonormal basis of H and think about the coefficients along each basis vector. Since K is bounded, there is some $C \geq 0$ (only depending on K) so that for all n, $||u_n|| \leq C$. Therefore, for all k and for all n, the "Fourier coefficient"

$$|\langle u_n, e_k \rangle| \leq ||u_n|| \cdot ||e_k|| \leq C,$$

and thus for each fixed k, we get a bounded sequence of coefficients along the kth basis vector: specifically, we have

the bounded sequence of numbers

$$\{\langle u_n, e_k \rangle\}_n$$

in $\mathbb{C}$. Thus, by Bolzano-Weierstrass (fixing $k = 1$), there is some subsequence $\{\langle u_{n_1(j)}, e_1 \rangle$ of $\{\langle u_n, e_1 \rangle\}_n$ which converges in $\mathbb{C}$ (in other words, we have a subset of the original $\{u_n\}$s in which the **first entry converges**). And now $\{\langle u_{n_1(j)}, e_2 \rangle$ is still a bounded sequence, and thus by Bolzano-Weierstrass again we have a **further** subsequence $\{\langle u_{n_2(j)}, e_2 \rangle$ which converges. So we now have a subset of the original $\{u_n\}$s in which the first and second entries both converge (since the first subsequence converges in the first entry, and thus any subsequence of it will also converge in the first entry).

We can repeat this argument arbitrarily many times: further subsequences of the $u_{n_2(j)}$s gives us a subsequence $u_{n_\ell(j)}$ such that $\{\langle u_{n_\ell(j)}, e_\ell \rangle$ converges, meaning that we have convergence along our sequence in the first ℓ entries. If we now define

$$v_\ell = u_{n_\ell}(\ell) \quad \forall \ell \in \mathbb{N},$$

then the $\{v_\ell\}_\ell$ form a subsequence of the $\{u_n\}_n$s with convergence in the kth entry (for any fixed k) as $\ell \to \infty$. This on its own doesn't mean that the sequence converges, but here is where we will use the fact that K has equi-small tails. It suffices to show that $\{v_\ell\}_\ell$ is Cauchy (because H is complete and K is closed). For any $\varepsilon > 0$, having equi-small tails tells us that there is some N such that

$$\sum_{k>n} |\langle v_\ell, e_k \rangle|^2 < \frac{\varepsilon^2}{16}$$

for all $\ell \in \mathbb{N}$. Now because the N sequences $\{\langle v_\ell, e_1 \rangle\}_\ell$ through $\{\langle v_\ell, e_N \rangle\}_\ell$ each converge, we can then find an M such that for all $\ell, m \geq M$,

$$\sum_{k=1}^{N} |\langle v_\ell, e_k \rangle - \langle v_m, e_k \rangle|^2 < \frac{\varepsilon^2}{4}.$$

We claim that this M is the one that we want for our sequence $\{v_\ell\}$: indeed,

$$||v_\ell - v_m|| = \sum_{k=1}^{N} \left[|\langle v_\ell - v_m, e_k \rangle|^2 + \sum_{k>N} |\langle v_\ell - v_m, e_k \rangle|^2 \right]^{1/2}$$

(because we have an orthonormal **basis**, the norm squared is the sum of the Fourier coefficients). Now using the fact that $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, we can bound this as

$$\leq \sum_{k=1}^{N} \left[|\langle v_\ell - v_m, e_k \rangle|^2 \right]^{1/2} + \left[\sum_{k>N} |\langle v_\ell - v_m, e_k \rangle|^2 \right]^{1/2}.$$

By our choice of M, the first term is at most $\frac{\varepsilon}{2}$, and we can use the ℓ^2 triangle inequality for the second term, thinking of that second term as the difference of the sequences $\{\langle v_\ell, e_k \rangle\}_k$ and $\{\langle v_m, e_k \rangle\}_k$. Thus we have the bound

$$< \frac{\varepsilon}{2} + \left[\sum_{k>N} |\langle v_\ell, e_k \rangle|^2 \right]^{1/2} + \left[\sum_{k>N} |\langle v_m, e_k \rangle|^2 \right]^{1/2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon,$$

where the last inequality comes from how we chose N. So our subsequence is Cauchy, thus convergent, and thus K is compact. $\qquad \square$

Example 198

Let K be the set (not subspace) of sequences $\{a_k\}_k$ in ℓ^2 satisfying $|a_k| \leq 2^{-k}$ – this set is known as the **Hilbert cube**, and it is compact.

It may seem unwieldy that we make this definition with respect to an orthonormal basis, but we can characterize compact sets in another way as well:

> **Theorem 199**
>
> A subset $K \subset H$ is compact if and only if K is closed, bounded, and for all $\varepsilon > 0$, there exists a finite-dimensional subspace $W \subset H$ so that for all $u \in K$, $\inf_{w \in W} \|u - w\| < \varepsilon$.

In other words, our additional condition is that we can approximate the points in K by a finite-dimensional subspace. This proof also involves a similar "diagonal argument," and notably it works for non-separable Hilbert spaces as well, but we can read about the proof on our own. This should be a believable result, because the equi-small tail condition we worked with in our previous proof was basically saying that we can approximate points in K by the first N vectors in our orthonormal basis (since the contribution from the other basis vectors is small).

We'll now start to talk about various **classes of operators**, and we'll start with the simplest ones. From linear algebra, we know that matrices are operators in finite-dimensional vector spaces, and we can represent them with an array of numbers. We can now generalize that definition to our current setting.

> **Fact 200**
>
> From here on, H will be a Hilbert space, and we'll denote $\mathcal{B}(H, H)$ by $\mathcal{B}(H)$.

> **Definition 201**
>
> A bounded linear operator $T \in \mathcal{B}(H)$ is a **finite rank operator** if the range of T (a subspace of H) is finite-dimensional. We denote this as $T \in \mathcal{R}(H)$.

> **Example 202**
>
> If H is a finite-dimensional Hilbert space, then every linear operator is of finite rank. For a more interesting example, for any positive integer n, the operator
>
> $$T a = \left\{ \frac{a_1}{1}, \frac{a_2}{2}, \cdot, \frac{a_n}{n}, 0, \cdots \right\}$$
>
> is a finite rank operator (because the image is spanned by the first n standard basis vectors).

> **Proposition 203**
>
> The set $\mathcal{R}(H)$ is a subspace of $\mathcal{B}(H)$.

Proof. The range of a scalar multiple of an operator is the same as the original range, and the sum of two finite rank operators has range contained in the direct sum of the individual ranges (which is also finite-dimensional). □

We'll now prove that these finite rank operators are really like matrices:

> **Theorem 204**
>
> An operator $T \in \mathcal{B}(H)$ is in $\mathcal{R}(H)$ if and only if there exists an orthonormal set $\{e_k\}_{k=1}^{L}$ and an array of constants $\{c_{ij}\}_{i,j=1}^{L} \subset \mathbb{C}$, such that
>
> $$T u = \sum_{i,j=1}^{L} c_{ij} \langle u, e_j \rangle e_i.$$

Proof. The backwards direction is clear: if T has such a representation, then the range of T is contained in the span of the L vectors $\{e_1, \cdots, e_L\}$ and is thus finite-dimensional. Now suppose that T is a finite rank operator. Then we can find an orthonormal basis $\{\bar{e}_k\}_{k=1}^N$ of the range of T, such that

$$Tu = \sum_{k=1}^{N} \langle Tu, \bar{e}_k \rangle \bar{e}_k$$

(since Tu is in the range, it must be this particular combination of the orthonormal basis vectors). Now by the definition of the adjoint operator, we can rewrite this sum as

$$= \sum_{k=1}^{n} \langle u, T^* \bar{e}_k \rangle \bar{e}_k = \sum_{k=1}^{N} \langle u, v_k \rangle \bar{e}_k,$$

where we've define $v_k = T^* \bar{e}_k$. If we now apply the Gram-Schmidt process to the vectors $\{\bar{e}_1, \cdots, \bar{e}_N, v_1, \cdots, v_N\}$, we get an orthonormal subset $\{e_1, \cdots, e_L\}$ with the same span as our original $\bar{e}_i$s and v_is. Thus, there exist constants a_{ki}, b_{kj} so that (expanding in terms of the new orthonormal subset)

$$\bar{e}_k = \sum_{i=1}^{L} a_{ki} e_i, \quad \bar{v}_k = \sum_{j=1}^{L} b_{kj} e_j.$$

Thus, substituting back in,

$$Tu = \sum_{i,j=1}^{L} \left(\sum_{k=1}^{N} a_{ki} \overline{b_{kj}} \right) \langle u, e_j \rangle e_i,$$

and now the term in the inner parentheses is our desired c_{ij}. $\qquad\square$

And with this characterization, we can now describe our finite rank linear operators more explicitly: for example, the nullspace of T contains the set of vectors orthogonal to all of the e_ks.

> **Theorem 205**
> If $T \in \mathcal{R}(H)$, then $T^* \in \mathcal{R}(H)$, and for any $A, B \in \mathcal{B}(H)$, $ATB \in \mathcal{R}(H)$.

In other words, $\mathcal{R}(H)$ is a "star-closed, two-sided ideal in the space of bounded linear operators" — it's closed under two-sided multiplication and adjoints.

Proof. We'll leave the closure under multiplication as an exercise: the main point is that if T has a finite-dimensional range, the range of AT is also finite-dimensional, and whatever happens with B doesn't really matter. For closure under adjoints, if T is a finite rank operator, then we can write

$$Tu = \sum_{i,j=1}^{L} c_{ij} \langle u, e_j \rangle e_i,$$

and thus

$$\langle u, T^* v \rangle = \langle Tu, v \rangle = \left\langle \sum_{i,j=1}^{L} c_{ij} \langle u, e_j \rangle e_i, v \right\rangle.$$

By linearity in the first entry, we can rewrite this inner product as

$$= \sum_{i,j} c_{ij} \langle u, e_j \rangle \langle e_i, v \rangle,$$

and we can now use linearity to pull things into the second component instead:

$$= \left\langle u, \sum_{i,j} \overline{c_{ij}} \overline{\langle e_i, v \rangle} e_j \right\rangle = \left\langle u, \sum_{i,j} \overline{c_{ij}} \langle v, e_i \rangle e_j \right\rangle.$$

But since this is true for all $u \in H$, we've shown that

$$\left\langle u, T^* v - \sum_{i,j} \overline{c_{ij}} \langle v, e_i \rangle e_j \right\rangle = 0$$

for all $u, v \in H$, and thus we must have $T^* v = \sum_{i,j=1}^{L} \overline{c_{ij}} \langle v, e_i \rangle e_j$ for all $v \in H$, and thus $T^* \in \mathcal{R}(H)$ as well: in fact, we can recover the coefficients in terms of the coefficients for T by reindexing as

$$= \sum_{i,j=1}^{L} \overline{c_{ji}} \langle v, e_j \rangle e_i.$$

Thus, the coefficients of the matrix governing T^* are obtained by taking the conjugate transpose of the ones for T. □

Since the set of finite rank linear operators is a subspace of the Banach space of bounded linear operators, which come with a norm, it makes sense to ask if the subspace of finite rank operators is closed (under the norm). In other words, if $T_n \in \mathcal{R}(H)$, and $||T_n - T|| \to 0$ as $n \to \infty$, we want to know whether $T \in \mathcal{R}(H)$. It turns out the answer is **no**:

Example 206

Let $T_n : \ell^2 \to \ell^2$ be a sequence of operators defined as

$$T_n a = \left\{ \frac{a_1}{1}, \cdots, \frac{a_n}{n}, 0, \cdots \right\}.$$

We can imagine that the limit T of these operators is the infinite-dimensional "diagonal matrix" with entries $(1, \frac{1}{2}, \frac{1}{3}, \cdots)$: specifically, defining

$$T a = \left\{ \frac{a_1}{1}, \frac{a_2}{2}, \frac{a_3}{3}, \cdots \right\},$$

we can check that $||T - T_n|| \leq \frac{1}{n+1}$, but T is not of finite rank (since $T(k e_k) = e_k$ is in the range for each standard basis vector e_k). In other words, the space of finite rank linear operators (which are nice because we can solve linear equations involving them using matrices) is not closed. But we still want to know about the closure of $\mathcal{R}(H)$, and the hope is that we still have a useful characterization:

Definition 207

An operator $K \in \mathcal{B}(H)$ is a **compact operator** if $\overline{K(\{u \in H : ||u|| \leq 1\})}$, the closure of the image of the unit ball under K, is compact.

We'll show next time that K is a compact operator if and only if it is in the closure of $\mathcal{R}(H)$, meaning that there is a sequence of finite rank operators converging to K. These compact operators will come up in useful problems – for example, T in our example above is compact, and the inverse of many differential operators will turn out to be compact as well. And as a sanity check before we do the proof next time, finite rank operators are indeed compact operators, because the image of the unit ball will be a bounded subset of a finite-dimensional subspace, and thus the closure of that image is a closed and bounded subset of a finite-dimensional subspace, which is compact by Heine-Borel.

18.102 / 18.1021 Introduction to Functional Analysis
Spring 2021