LINEAR ALGEBRA IN INFINITE DIMENSIONS

ARAV BHATTACHARYA

1. Introduction

Many results from linear algebra are exclusive to finite-dimensional vector spaces. As such, one may be inclined to ask when results from linear algebra hold in the case of infinite-dimensional vector spaces; these questions are answered in functional analysis.

Some results, like rank-nullity, simply cannot be translated to an infinite-dimensional setting, as they have no reasonable parallel in such a setting. Other results, like the equivalence of surjective and injective maps (see the example below), simply do not hold in this more general setting.

Example. The derivative map $D : \mathcal{P}(\mathbb{R}) \longrightarrow \mathcal{P}(\mathbb{R})$, where $\mathcal{P}(\mathbb{R})$ is the space of all polynomials with real coefficients, is not injective. Let $x = t$ and $y = t + 1$. Then $Dx = Dy = 1$, but clearly $x \neq y$. However, the derivative is surjective. We prove this using the fact that $\deg(Dx) = \deg x - 1$: due to this property of the derivative map, we choose a set S where there is exactly one polynomial of each degree (barring the zero polynomial). Since there is at least one element of S for each degree, we have that $\operatorname{span} S = \mathcal{P}(\mathbb{R})$. Thus by linearity D is surjective.

Another challenge with working in infinite-dimensional spaces is the challenge of coordinate representations. In linear algebra, many facts can be proved by selecting an arbitrary basis and proving the fact for any given linear combination of basis vectors. However, since only a finite number of basis elements can be summed in the representation of an arbitrary vector, any basis has uncountably many elements (this can be proved using the Baire Category Theorem, which is beyond the scope of this paper; see [Car05]). Thus many proof techniques from linear algebra are no longer feasible in functional analysis.

Despite these multiple challenges, functional analysis is a rich field with results that build on linear algebra. For example, the Riesz representation theorem, which is trivial in finite dimensions, becomes really insightful in infinite dimensions. Also, the fact that all vector spaces of dimension n over either $\mathbb{R}$ or $\mathbb{C}$ are isomorphic to $\mathbb{R}^n$ and $\mathbb{C}^n$ has a somewhat unexpected parallel in infinite dimensions.

2. Prerequisites

Some familiarity with the definitions of linear algebra and real analysis is necessary for reading and fully understanding this paper. Please read [Axl97] if you need to learn linear algebra, and [R$^+$76] if you need to learn real analysis.

Date: July 11, 2022.

3. Banach Spaces

Our first restriction in functional analysis is to normed linear spaces, i.e. vector spaces equipped with a linear norm (a notion of length). We denote the norm of any element by $\|\cdot\|$, and the norm must satisfy the following axioms of a normed linear space for any scalar α and vectors x, y:

$$\begin{aligned}\|x\| &\geq 0\\ \|x\| = 0 &\Leftrightarrow x = 0\\ \|\alpha x\| &= |\alpha|\|x\|\\ \|x+y\| &\leq \|x\| + \|y\|,\end{aligned}$$

the last of which is called the triangle inequality.

Note that we can define a metric on X by $d(x, y) = \|x - y\|$, and this metric is said to be induced by the norm. One can see that the axioms of a metric space are satisfied quite trivially. Thus we note that restricting ourselves to normed linear spaces is useful as it lets us use tools from both linear algebra and real analysis in infinite dimensions. Very loosely, linear algebra helps us examine infinite-dimensional spaces through the lens of finite-dimensional spaces, while analysis helps us translate the theory of finite dimensions to infinite dimensions.

One further restriction on the normed linear spaces we'll look at is *completeness*.

Definition 3.1. Recall that a sequence (x_n) in a metric space (X, d) is Cauchy if for every positive real r there exists some number N such that $d(x_m, x_n) < r$ whenever $m, n > N$. The metric space is *complete* if every Cauchy sequence in X converges to a a point in X.

We call a complete normed linear space a *Banach space*. Essentially, the reason completeness is useful is because it means that any sequence of vectors in a space has the property that pairs of terms in that sequence get closer and closer to each other, then the sequence converges to a limit in the same space. This property might seem very restrictive; however, a result of real analysis states that any metric space (and hence any normed linear space) X is isometric to a dense subspace W of a complete metric space $\hat{X}$ (i.e. elements of W can get arbitrarily close to any element of $\hat{X}$). The space $\hat{X}$, which is unique up to isometry, is called the "completion" of X.

Before moving on, note that any finite-dimensional normed space X is complete, and furthermore every finite-dimensional subspace Y of a normed space X is closed in X.

4. Examples of Banach Spaces

Let's look at a couple of examples (and a non-example) of Banach spaces (see [Kre91] for proofs of completeness).

Example. The spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ are Banach spaces with norm defined by

$$\|x\| = \sqrt{\sum_{j=1}^{n} |x_j|^2}$$

for $x = (x_1, \ldots, x_n)$.

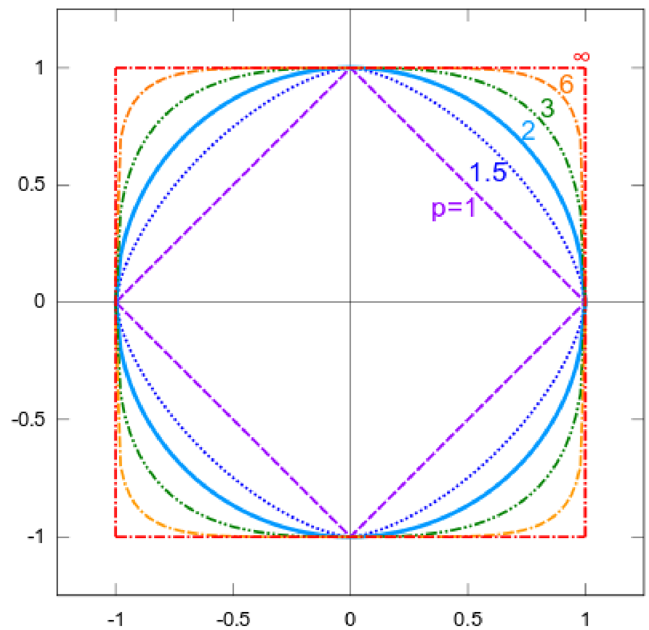

Figure 1. Illustration of the unit circles of some ℓ^p-norms (image credit [Qua11])

Example. Another set of Banach spaces are the sequence spaces ℓ^p, where $p \geq 1$. These spaces contain all sequences $x = (x_j)$ where the sum

$$\sum_{j=1}^{\infty} |x_j|^p$$

converges. If we consider real sequences, we get the *real space* ℓ^p, while if we consider complex sequences we get the *complex space* ℓ^p. For these spaces, the norm is given by

$$\|x\| = \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{\frac{1}{p}}.$$

Example. Another Banach space is the sequence space ℓ^∞, the space of bounded sequences of complex numbers under the norm

$$\|x\| = \sup_{j \in \mathbb{N}} |x_j|$$

for $x = (x_j)$. The nomenclature comes from the fact that the norm on ℓ^∞ is the limit as $p \longrightarrow \infty$ of the norm on ℓ^p. To see this, we can plot the set of $x = (x_1, x_2) \in \mathbb{R}^2$ such that

$$\|x\|_{\ell^p} = x_1|^p + |x_2|^p = 1.$$

See Figure 1.

Example. Another Banach space is the function space $C[a, b]$, the space of continuous functions on the closed interval from a to b, under the norm

$$\|x\| = \max_{t \in [a,b]} |x(t)|.$$

Elements of this space are continuous functions $x : [a, b] \longrightarrow \mathbb{R}$.

Example. One normed linear space that is *not* a Banach space is the space $\mathcal{P}[a, b]$, the space of all polynomials on the interval $[a, b] \subseteq \mathbb{R}$, under the norm

$$\|x\| = \max_{t\in[a,b]} |x(t)|.$$

As an example of a Cauchy sequence in this space that doesn't converge, let $[a, b] = [-1, 1]$ and consider the sequence (x_n) where

$$x_n(t) = \sum_{j=0}^{n} \frac{(-1)^n t^{2n}}{(2n)!} \in \mathcal{P}[-1, 1].$$

However,

$$x(t) = \lim x_n(t) = \lim \sum_{j=0}^{n} \frac{(-1)^n t^{2n}}{(2n)!} = \cos x \notin \mathcal{P}[-1, 1].$$

5. Operators on Banach Spaces

A huge part of linear algebra is the theory of linear operators:

Definition 5.1. A *linear operator* T is an operator such that

(1) the domain $\mathfrak{D}(T)$ is a vector space and the range range T lies in a vector space over the same field,
(2) for all $x, y \in \mathfrak{D}(T)$ and scalars α,

$$T(x + y) = Tx + Ty$$
$$T(\alpha x) = \alpha Tx.$$

The null space, denoted null T, is the set of all $x \in \mathfrak{D}(T)$ such that $Tx = 0$.

In linear algebra, we have tools like the rank-nullity theorem, which says that

$$\dim \operatorname{range} T + \dim \operatorname{null} T = \dim \mathfrak{D}(T);$$

and matrix representations by which any linear operator that takes m-dimensional space to n-dimensional space can be represented by an $(m \times n)$-dimensional matrix. However, we do not have access to these tools in functional analysis simply because of the fact that we are working in infinite-dimensional spaces and so the concepts of dimension and matrix multiplication break down.

Although we don't have a rank-nullity theorem for infinite-dimensional spaces, we have some much looser results about the range and null space:

Theorem 5.2. *Let $T : X \longrightarrow Y$ be a linear operator. Then,*

(1) *The range* range T *is a vector space.*
(2) *The null space* null T *is a vector space.*

Proof. We'll go sequentially in this proof.

(1) We know that range $T \subseteq Y$. Furthermore $T(0) = 0$ so $0 \in$ range T. Lastly, suppose we have $y_1, y_2 \in$ range T. Then there are some $x_1, x_2 \in \mathfrak{D}(T)$ such that $Tx_1 = y_1, Tx_2 = y_2$. Since $\mathfrak{D}(T)$ is a vector space, for scalars a_1, a_2 we have that

$$a_1y_1 + a_2y_2 = a_1Tx_1 + a_2Tx_2 = T(a_1x_1 + a_2x_2) \in \operatorname{range} T.$$

Thus range T is closed under the vector space operations and contains 0, so it is a vector space.

(2) We know that null $T \subseteq \mathfrak{D}(T)$, and clearly $T(0) = 0$ so $0 \in$ null T. Now, suppose we have $x_1, x_2 \in$ null T. Then

$$Tx_1 = Tx_2 = 0.$$

Since T is linear, for scalars a_1, a_2 we have that

$$(a_1x_1 + a_2x_2) = a_1Tx_1 + a_2Tx_2 = 0,$$

so $a_1x_1 + a_2x_2 \in$ null T. Thus null T is closed under the vector space operations and contains 0, so it is a vector space.

■

Many relevant properties of linear operators from linear algebra are not possible to translate over to all linear operators in functional analysis. One restriction that will help us here is to bounded operators:

Definition 5.3. Let X and Y be normed spaces and $T : \mathfrak{D}(T) \longrightarrow Y$ a linear operator, where $\mathfrak{D}(T) \subseteq X$. The operator T is said to be *bounded* if there is a real number c such that for all $x \in \mathfrak{D}(T)$,

$$\|Tx\| \leq c\|x\|.$$

Note that the word bounded here is used differently from its use in calculus. In calculus, a function is bounded if it maps its domain to a bounded set. In functional analysis, an operator is bounded if it maps bounded sets to bounded sets. Furthermore, all linear operators on finite-dimensional normed spaces are bounded (see [Kre91]).

Let's look at some examples (and non-examples) of bounded operators.

Example. The identity operator $I : X \longrightarrow X$ on a nonzero normed space is bounded with norm $\|I\| = 1$.

Example. The zero operator $0 : X \longrightarrow Y$ on a normed space is bounded with norm $\|0\| = 0$.

Example. Consider the normed space $\mathcal{P}[0,1]$ of all polynomials on $[0,1]$ with norm given by $\|x\| = \max|x(t)|$. We define the differentiation operator D on $\mathcal{P}[0,1]$ by

$$Dx(t) = x'(t).$$

This operator is linear but not bounded: let $x_n(t) = t^n$, where $n \in \mathbb{N}$. Then $\|x_n\| = 1$ and

$$Dx_n(t) = x_n'(t) = nt^{n-1},$$

so $\|Dx_n\| = n$ and $\|Dx_n\|/\|x_n\| = n$. Since n is arbitrary, there is no fixed $c \in \mathbb{R}$ such that $\|Dx_n\|/\|x_n\| \leq c$. Thus D is not bounded. Thus there are some unbounded operators that are of practical importance (though their theory is much more complicated than that of bounded operators).

Example. A bounded linear operator on $C[0,1]$ is the integral operator $\int : C[0,1] \longrightarrow \mathbb{R}$ where for $x = x(t) \in C[0,1]$ and some $k(t,\tau) \in C([0,1] \times [0,1])$,

$$\int x = \int_0^1 x(t)\,dt.$$

This operator is clearly linear and we will show that it is also bounded. Note that since by the continuity of k on the closed interval $[0,1] \times [0,1]$, we have that k must be bounded with say $|k(t,\tau)| \leq k_0$ for some $k_0 \in \mathbb{R}$. Then

$$\left\| \int x \right\| = \max_{t \in [0,1]} \left| \int_0^1 k(t,\tau) x(\tau)\, d\tau \right|$$
$$\leq \max_{t \in [0,1]} \int_0^1 |k(t,\tau)| |x(t)|\, d\tau \leq k_0 \|x\|.$$

Thus $\| \int x \| \leq \|x\|$, so $\int$ is bounded with $\| \int \| = k_0$.

6. Bounded Operators

In this section we look at some results concerning bounded linear operators. First of all, restriction to bounded operators is useful because we can define an operator norm for such operators:

Definition 6.1. We define the operator norm by

$$\|T\| = \sup_{x \in \mathfrak{D}(T) \setminus \{0\}} \frac{\|Tx\|}{\|x\|}.$$

If $\mathfrak{D}(T) = \{0\}$, we set $\|T\| = 0$.

We can actually restrict our search for the norm of T to vectors of norm 1. We verify this and the norm axioms below:

Lemma 6.2. *Let $T : X \longrightarrow Y$ be a bounded linear operator. Then:*

(1) *An alternative formula for the norm of T is*

$$\|T\| = \sup_{\|x\|=1} \|Tx\|.$$

(2) *The operator norm satisfies the norm axioms.*

Proof. We once again go sequentially in this proof.

(1) Let $x \in X$ with $\|x\| = a$ and set $y = \frac{1}{a}x \in X$ (where $x \neq 0$). Then $\|y\| = 1$ and by the linearity of T,

$$\|T\| = \sup_{x \neq 0} \frac{1}{a} \|Tx\| = \sup_{x \neq 0} \left\| T\left(\frac{1}{a}x\right) \right\| = \sup_{\|y\|=1} \|Ty\|.$$

(2) Clearly $\|T\| \geq 0$, and $\|0\| = 0$. Conversely if $\|T\| = 0$ then for all $x \in X$, $Tx = 0$ and so clearly $T = 0$. Next, for all $x \in X$,

$$\sup_{\|x\|=1} \|\alpha Tx\| = \sup_{\|x\|=1} |\alpha| \|Tx\| = |\alpha| \sup_{\|x\|=1} \|Tx\|.$$

Lastly, if we have another bounded linear operator S, then

$$\sup_{\|x\|=1} \|(S+T)x\| = \sup_{\|x\|=1} \|Sx + Tx\| \leq \sup_{\|x\|=1} \|Sx\| + \sup_{\|x\|=1} \|Tx\|.$$

Thus all the norm axioms are satisfied by the operator norm.

∎

Another result about bounded linear operators is that these operators are exactly the continuous linear operators:

Theorem 6.3. *A linear operator T on normed spaces is continuous if and only if it is bounded.*

Proof. We first show the if direction. For $T = 0$ the statement is trivial, so take bounded $T \neq 0$. Then $\|T\| \neq 0$. Now, consider any $x_0 \in \mathfrak{D}(T)$ and let $\varepsilon > 0$. Then since T is linear, for every $x \in \mathfrak{D}(T)$ such that

$$\|x - x_0\| < \frac{\varepsilon}{\|T\|},$$

we have that

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\|\|x - x_0\| < \|T\|\frac{\varepsilon}{\|T\|} = \varepsilon.$$

Since we chose x_0 arbitrarily, we have that T is continuous.

Now for the converse. Assume that T is continuous at an arbitrary $x_0 \in \mathfrak{D}(T)$. Then for all $\varepsilon > 0$ there is $\delta > 0$ such that whenever $\|x - x_0\| \leq \delta$, $\|Tx - Tx_0\| \leq \varepsilon$. Now take any $y \neq 0$ in $\mathfrak{D}(T)$ and set

$$x = x_0 + \frac{\delta}{\|y\|}y.$$

Then

$$x - x_0 = \frac{\delta}{\|y\|}y.$$

Hence $\|x - x_0\| = \delta$ so by the linearity of T,

$$\|Tx - Tx_0\| = \|T(x - x_0)\| = \left\|T\left(\frac{\delta}{\|y\|}y\right)\right\| = \frac{\delta}{\|y\|}\|Ty\| \leq \varepsilon.$$

Thus we have that

$$\|Ty\| \leq \frac{\varepsilon}{\delta}\|y\|,$$

so T is bounded. ■

Theorem 6.3 is important because the continuity of bounded linear operators means that for a bounded linear operator T,

$$x_n \longrightarrow x \implies Tx_n \longrightarrow Tx,$$

since as $n \longrightarrow \infty$ we have that

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \leq \|T\|\|x_n - x\| \longrightarrow 0.$$

A last result about bounded linear operators that will be of use to us is below.

Theorem 6.4. *The vector space $B(X, Y)$ of all bounded linear operators from a normed space X into a normed space Y is itself a normed space with norm defined by*

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|.$$

Furthermore, if Y is a Banach space, then so is $B(X, Y)$.

Proof. To see that $\|T\|$ is a valid norm, one can use the definition of the supremum to verify the norm axioms.

Now, suppose that Y is Banach. We consider then an arbitrary Cauchy sequence (T_n) in $B(X, Y)$, and we'll show that (T_n) converges to an operator $T \in B(X, Y)$.

Since (T_n) is Cauchy, for all $\varepsilon > 0$ there is an N such that for all $m, n > N$,

$$\|T_n - T_m\| < \varepsilon.$$

Thus for all $x \in X$, we have that

$$\|(T_n - T_m)x\| \leq \|T_n - T_m\|\|x\| < \varepsilon\|x\|.$$

Thus for any fixed $x \in X$, we clearly see that $(T_n - T_m)x \in Y$ is Cauchy and thus approaches 0. Therefore $T_n x \longrightarrow y \in Y$ by the completeness of Y.

We can thus define an operator

$$T : X \longrightarrow Y : x \longmapsto y.$$

Now, T is linear since

$$\begin{aligned} T(\alpha x + \beta z) &= limT_n(\alpha x + \beta z) \\ &= \lim(\alpha T_n x + \beta T_n z) \\ &= \alpha \lim T_n x + \beta \lim T_n z \\ &= \alpha T x + \beta T z. \end{aligned}$$

Now, we show the boundedness of T. Note that from the above we know that for $m, n > N$ we have that $\|T_n - T_m\| < \varepsilon\|x\|$ for all $x \in X$. Thus

$$\begin{aligned} \|(T_n - T)x\| &= \|T_n x - Tx\| \\ &= \left\|T_n x - \lim_{m\to\infty} T_m x\right\| \\ &= \lim_{m\to\infty} \|T_n x - T_m x\| \\ &\leq \varepsilon\|x\|. \end{aligned}$$

Thus $(T_n - T)$ is bounded. Since T_n is bounded, $T = T_n - (T_n - T)$ is bounded. As a result, $T \in B(X, Y)$. Now, if we take the supremum of all x in the above equation, we get that as $n \longrightarrow \infty$

$$\|(T_n - T)x\| \leq \|T_n - T\|\|x\| = \varepsilon\|x\| \longrightarrow 0.$$

Thus $\|T_n - T\| \longrightarrow 0$, so $T_n \longrightarrow T$. ■

7. Dual Spaces

Now, another interesting aspect of linear algebra is the study of linear functionals and dual spaces. Let's define linear functionals and the dual space:

Definition 7.1. A *linear functional* f is a linear operator with its domain a vector space X and range in the scalar field $\mathbb{F}$ of X; in other words,

$$f : \mathfrak{D}(f) \longrightarrow \mathbb{F},$$

where $\mathbb{F} = \mathbb{R}$ if X is real and $\mathbb{F} = \mathbb{C}$ if X is complex.

Definition 7.2. The set of all bounded linear functionals on a normed space X constitutes a normed space with norm defined by

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)|,$$

and this space is called the dual space of X, denoted X'.

In finite dimensions, all linear functionals are bounded. However, this is not the case in infinite dimensions, so we restrict our dual space to bounded functionals to maintain two important properties from finite dimensions. First of all, $X = X''$ (the proof of this fact requires more tools than we develop in this paper; see [Kre91]). The second property is as follows:

Theorem 7.3. *The dual space X' of any normed space X is a Banach space.*

Proof. Note that the dual space is the space of bounded linear operators from X to $\mathbb{F}$. Thus by Theorem 6.4 and the fact that $\mathbb{R}$ and $\mathbb{C}$ are complete (hence Banach), we have that X' is Banach. ■

Let's look at some examples of functionals:

Example. The *norm* $\|\cdot\| : X \longrightarrow \mathbb{R}$ is a functional on X which is not necessarily linear.

Example. The *dot product* with one factor kept constant is a functional $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ where

$$f(x) = x \cdot a$$

for some fixed $a \in \mathbb{R}^3$. Now, f is linear and bounded with

$$|f(x)| = |x \cdot a| \leq \|x\|\|a\|,$$

so clearly $\|f\| \leq \|a\|$. Taking $x = a$, we get

$$\|f\| \geq \frac{|f(a)|}{\|a\|} = \frac{\|a\|^2}{\|a\|} = \|a\|,$$

so $\|f\| = \|a\|$.

Example. A bounded linear functional on $C[0,1]$ is the *definite integral* $\int : C[0,1] \longrightarrow \mathbb{R}$ where for $x = x(t) \in C[0,1]$,

$$\int x = \int_0^1 x(t)\, dt.$$

This functional is the integral operator from the examples of bounded linear operators with $k(t,\tau)$ being the identity map I, so it is linear and bounded with $\| \int \| = 1$.

8. Hilbert Spaces

What if we wanted an analogue to the familiar dot product from Euclidean space in our infinite dimensional spaces? We can find this analogue by looking at inner product spaces, defined below.

Definition 8.1. An *inner product space* is a vector space X equipped with a scalar product written $\langle \cdot, \cdot \rangle$ that satisfies the following properties:

$$\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle \qquad \langle \alpha x, y\rangle = \alpha\langle x, y\rangle \tag{8.1}$$

$$\langle x, y\rangle = \overline{\langle y, x\rangle} \tag{8.2}$$

$$\langle x, x\rangle \geq 0 \qquad \langle x, x\rangle = 0 \Leftrightarrow x = 0 \tag{8.3}$$

In other words, the inner product is linear in the first argument (1.1), conjugate symmetric (1.2), and positive definite (1.3).

We call an inner product space that is complete in the norm induced by the inner product a *Hilbert space*.

We define a norm on an inner product space by $\|x\| = \sqrt{\langle x, x\rangle}$ and a metric on such a space by $d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y\rangle}$. We will prove that the norm satisfies the normed space axioms, at which point it will be clear that the metric space axioms are also satisfied. Note that the first and second axioms are satisfied by the last two inner product axioms. Also,

$$\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = \alpha\overline{\alpha}\langle x, x\rangle = |\alpha|^2\|x\|^2,$$

and taking square roots on both sides we see that the third norm axiom is satisfied. Lastly, we prove the triangle inequality below.

Lemma 8.2 (Schwarz and triangle inequalities)**.** *The Schwarz inequality states that in an inner product space, the norm induced by the inner product satisfies*

$$|\langle x, y\rangle| \leq \|x\|\|y\|,$$

where equality holds exactly when $\{x, y\}$ is a linearly dependent set.

The inner product norm also satisfies

$$\|x + y\| \leq \|x\| + \|y\|,$$

where equality holds when there is some nonnegative real c such that $x = cy$ or $y = cx$.

Proof. We begin with the Schwarz inequality, and use it to prove the triangle inequality. Notice that the Schwarz inequality holds trivially for $y = 0$, since $\langle x, 0\rangle = 0$. Thus we take $y \neq 0$. For every scalar α we have

$$\begin{aligned} 0 \leq \|x - \alpha y\|^2 &= \langle x - \alpha y, x - \alpha y\rangle \\ &= \langle x, x\rangle - \overline{\alpha}\langle x, y\rangle - \alpha(\langle y, x\rangle - \overline{\alpha}\langle y, y\rangle). \end{aligned}$$

Taking $\overline{\alpha} = \langle y, x\rangle / \langle y, y\rangle$, we simplify our inequality to

$$\begin{aligned} 0 &\leq \langle x, x\rangle - \frac{\langle y, x\rangle}{\langle y, y\rangle}\langle x, y\rangle - \alpha(\langle y, x\rangle - \frac{\langle y, x\rangle}{\langle y, y\rangle}\langle y, y\rangle) \\ &= \langle x, x\rangle - \frac{\langle y, x\rangle}{\langle y, y\rangle}\langle x, y\rangle - \alpha(\langle y, x\rangle - \langle y, x\rangle) \\ &= \langle x, x\rangle - \frac{\langle y, x\rangle}{\langle y, y\rangle}\rangle = \|x\|^2 - \frac{\langle y, x\rangle\,\langle x, y\rangle}{\|y\|^2} \\ &= \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2}. \end{aligned}$$

Rearranging, we get that $|\langle x, y\rangle|^2 \leq \|x\|^2\|y\|^2$, and taking square roots on both sides gives us the Schwarz inequality.

Now, the equality cases require either $y = 0$ or $0 = \|x - \alpha y\|^2$, in which case $x = \alpha y$. In either case, $\{x, y\}$ is linearly dependent.

Now, let's move on to the triangle inequality. By the Schwarz inequality,

$$|\langle x, y\rangle| = |\langle y, x\rangle| \leq \|x\|\|y\|.$$

Thus we have that

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

Taking square roots on both sides, we get the triangle inequality.

We have equality here exactly when

$$\langle x, y\rangle + \langle y, x\rangle = 2\|x\|\|y\|.$$

Note that the left-hand side of this equation is $2\Re\langle x, y\rangle$, where $\Re$ denotes the real part. By the Schwarz inequality, we must then have

$$\Re\langle x, y\rangle = \|x\|\|y\| \geq |\langle x, y\rangle|.$$

Since the real part of any complex number cannot exceed its absolute value, we must have that both sides of the above Schwarz inequality are equal, whereby x, y are linearly dependent; furthermore, the imaginary part of $\langle x, y\rangle$ must be zero, so $\langle x, y\rangle$ is a positive real number.

Without loss of generality, say we have $x = cy$ for nonzero y (the only case this does not cover is $y = 0$, which follows by switching the variables). We will show that c is a real nonnegative number. Since we have that

$$0 \leq \Re\langle x, y\rangle = \Re\langle y, y\rangle = |\langle cy, y\rangle| = c\|y\|^2,$$

we can divide both sides by $\|y\|^2$ to get that $c \geq 0$. ■

Thus by Lemma 8.2, inner product spaces are normed linear spaces, and Hilbert spaces are Banach spaces. There are many equations and inequalities that apply in any inner product space. Let's look at a few.

Lemma 8.3 (parallelogram equality)**.** *In an inner product space, we have that*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Proof. We have that

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

■

9. Examples of Hilbert Spaces

Let's look at some examples (and non-examples) of Hilbert spaces.

Example. The space $\mathbb{R}^n$ is a Hilbert space with inner product

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j y_j,$$

where $x = (x_j)$ and $y = (y_j)$.

Example. The space $\mathbb{C}^n$ is a Hilbert space with inner product

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j \overline{y_j},$$

where $x = (x_j)$ and $y = (y_j)$.

Example. The space ℓ^2 is a Hilbert space with inner product

$$\langle x, y \rangle = \sum_{j=1}^{\infty} x_j y_j$$

for $x = (x_j)$ and $y = (y_j)$. Actually, ℓ^2 is the first Hilbert space, and it was investigated by D. Hilbert in his work on integral equations.

Example. The space ℓ^p with $p \neq 2$ is *not* a Hilbert space, i.e. the ℓ^p norm cannot be obtained from an inner product. Take $x = (1, 1, 0, 0, \cdots, 0, \cdots)$ and $y = (1, -1, 0, 0, \cdots, 0, \cdots)$, where clearly $x, y \in \ell^p$. Thus

$$\|x\| = \|y\| = 2^{1/p} \qquad \|x + y\| = \|x - y\| = 2,$$

so

$$8 = \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) = 2 \cdot 2 \cdot 2^{2/p} = 2^{2/p+2},$$

which does not hold for $p \neq 2$.

Therefore ℓ^p for $p \neq 2$ is a Banach space which is not a Hilbert space.

Example. The space $C[a, b]$ is not an inner product space, i.e. the norm $\|x\| = \max_{t \in [a,b} |x(t)|$ can't be obtained from an inner product. To see this, take $x(t) = 1$ and $y(t) = (t - a)/(b - a)$, so $\|x\| = \|y\| = 1$ and

$$x(t) + y(t) = 1 + \frac{t - a}{b - a},$$
$$x(t) - y(t) = 1 - \frac{t - a}{b - a}.$$

Thus $\|x + y\| = 2$, $\|x - y\| = 1$, and so

$$\|x + y\|^2 + \|x - y\|^2 = 5 \neq 2(\|x\|^2 + \|y\|^2) = 4.$$

10. Inner Products

In this section, we will take a look at some identities and results regarding inner products. First of all, inner products allow us to define orthogonality, which generalizes the concept of perpendicularity to all inner product spaces.

Definition 10.1. An element x of an inner product space X is said to be *orthogonal* to an element $y \in X$ if

$$\langle x, y\rangle = 0.$$

We also say that x and y are orthogonal, and we write $x \perp y$. Similarly, for subsets $A, B \subseteq X$ we write $x \perp A$ if $x \perp a$ for all $a \in A$, and $A \perp B$ if $a \perp b$ for all $a \in A$ and $b \in B$.

We can use this definition of orthogonality and the parallelogram equality to generalize the Pythagorean theorem to inner product spaces:

Theorem 10.2 (Pythagoras)**.** *For elements x, y in an inner product space X, with $x \perp y$, the following equation holds:*

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

Proof. Notice that if $x \perp y$, we have that

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

Thus the parallelogram equality becomes $\|x + y\|^2 + \|x\|^2 + \|y\|^2 = 2(\|x\|^2 + \|y\|^2)$, which simplifies to

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

■

Another equation that we can prove with the parallelogram equality is the Appolonius identity:

Corollary 10.3 (Appolonius)**.** *For any x, y, z in an inner product space, we have that*

$$\|z - x\|^2 + \|z - y\|^2 = \tfrac{1}{2}\|x - y\|^2 + 2\|z - \tfrac{1}{2}\|(x + y)\|^2.$$

Proof. We get this directly by plugging values into the parallelogram identity, namely $z + \frac{1}{2}(x + y)$ and $\frac{1}{2}(x - y)$. Thus we get

$$\begin{aligned} \|z - x\|^2 + \|z - y\|^2 &= \|z + \tfrac{1}{2}(x + y) + \tfrac{1}{2}(x - y)\| + \|z + \tfrac{1}{2}(x + y) + \tfrac{1}{2}(x - y)\| \\ &= 2(\|z + \tfrac{1}{2}(x + y)\|^2 + \|\tfrac{1}{2}(x - y)\|^2) \\ &= 2\|z + \tfrac{1}{2}(x + y)\| + \tfrac{1}{2}\|x - y\|^2. \end{aligned}$$

■

As a last result, we show that the inner product is continuous.

Lemma 10.4. *Any inner product is continuous.*

Proof. Note that the lemma is equivalent to saying that if $(x_n) \longrightarrow x$ and $(y_n) \longrightarrow y$, then $\langle x_n, y_n\rangle \longrightarrow \langle x, y\rangle$. We thus have that as $n \longrightarrow \infty$,

$$\begin{aligned}
|\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle \\
&\leq |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\
&\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\| \longrightarrow 0.
\end{aligned}$$

■

11. Orthogonal Complement

We now build up the concept of an orthogonal complement, extending another key part of finite-dimensional inner product spaces. We first show that any complete convex subset (which doesn't have to be a subspace) of an inner product space contains a unique "closest" vector to any vector in the inner product space.

Theorem 11.1. *Let X be an inner product space and $M \neq \varnothing$ a convex subset which is complete (in the metric induced by the inner product). Then for every given $x \in X$ there exists a unique $y \in M$ whose distance from x is minimized to a value $\delta \in \mathbb{R}$. In other words,*

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| = \|x - y\|.$$

Proof. We start with the existence. By the definition of an infimum there is a sequence (y_n) in M such that we have the decreasing sequence

$$(\delta_n) \longrightarrow \delta, \qquad \delta_n = \|x - y_n\|.$$

We show that (y_n) is Cauchy. Writing $y_n - x = v_n$, we get $\|v_n\| = \delta_n$ and

$$\|v_n + v_m\| = \|y_n + y_m - 2x\| = 2\|\tfrac{1}{2}(y_n + y_m) - x\| \geq 2\delta.$$

Because M is convex, we have that $\frac{1}{2}(y_n + y_m) \in M$. Furthermore, we have $y_n - y_m = v_n - v_m$. By the parallelogram equality,

$$\begin{aligned}
\|y_n - y_m\|^2 = \|v_n - v_m\|^2 &= -\|v_n + v_m\|^2 + 2(\|v_n\|^2 + \|v_m\|^2) \\
&\leq -(2\delta)^2 + 2(\delta_n^2 + \delta_m^2) \longrightarrow 0,
\end{aligned}$$

so (y_n) is Cauchy. Since M is complete, we have a $y \in M$ such that $(y_n) \longrightarrow y$. Since $y \in M$, $\|x - y\| \geq \delta$. Furthermore, $\|x - y\| = \delta$ since

$$\|x - y\| \leq \|x - y_n\| + \|y_n - y\| = \delta_n + \|y_n - y\| \longrightarrow \delta.$$

Next, we show uniqueness. Suppose that $y \in M$ and $y_0 \in M$ both satisfy

$$\|x - y\| = \delta, \qquad \|x - y_0\| = \delta.$$

By the parallelogram equality,

$$\begin{aligned}
\|y - y_0\|^2 &= \|(y - x) - (y_0 - x)\|^2 \\
&= 2\|y - x\|^2 + 2\|y_0 - x\|^2 - \|(y - x) + (y_0 - x)\|^2 \\
&= 4\delta^2 - 4\|\tfrac{1}{2}(y + y_0) - x\|^2.
\end{aligned}$$

Notice that $\frac{1}{2}(y + y_0) \in M$, so

$$\|\tfrac{1}{2}(y + y_0) - x\| \geq \delta.$$

Thus we have that

$$\begin{aligned} 0 \le \|y - y_0\|^2 &= 4\delta^2 - 4\|\tfrac{1}{2}(y + y_0) - x\|^2 \\ &\le 4\delta^2 - 4\delta^2 = 0, \end{aligned}$$

so $\|y - y_0\| = 0$ and clearly $y = y_0$. ■

Now, if we consider complete *subspaces* in Theorem 11.1, we get a much stronger result:

Lemma 11.2. *In Theorem 11.1, let M be a complete subspace Y and $x \in X$ fixed. Then $z = x - y$ is orthogonal to Y.*

Proof. If we did not have that $z \perp Y$, then there would be $y_1 \in Y$ such that

$$\langle z, y_1 \rangle = \beta \neq 0.$$

Evidently, $y_1 \neq 0$. For any scalar α,

$$\begin{aligned} \|z - \alpha y_1\|^2 &= \langle z - \alpha y_1, z - \alpha y_1 \rangle \\ &= \langle z, z \rangle - \overline{\alpha}\langle z, y_1 \rangle - \alpha(\langle y_1, z \rangle - \overline{\alpha}\langle y_1, y_1 \rangle) \\ &= \langle z, z \rangle - \overline{\alpha}\beta - \alpha(\overline{\beta} - \overline{\alpha}\langle y_1, y_1 \rangle). \end{aligned}$$

Note that $\|z\| = \|x - y\| = \delta$, so taking

$$\overline{\alpha} = \frac{\overline{\beta}}{\langle y_1, y_1 \rangle}$$

yields

$$\begin{aligned} \|z - \alpha y_1\|^2 &= \langle z, z \rangle - \frac{\beta\overline{\beta}}{\langle y_1, y_1} - \alpha\left(\overline{\beta} - \frac{\overline{\beta}}{\langle y_1, y_1}\langle y_1, y_1 \rangle\right) \\ &= \langle z, z \rangle - \frac{|\beta|^2}{\langle y_1, y_1 \rangle} < \delta^2, \end{aligned}$$

which is a contradiction since Y is convex and therefore

$$\delta = \inf_{v \in Y} \|x - v\| < \|z - \alpha y_1\| = \|x - (y + \alpha y_1)\|.$$

Thus we must have that $z \perp Y$, so the lemma has been proven. ■

Now, before discussing Lemma 11.2, we make a quick definition.

Definition 11.3. A vector space X is said to be the *direct sum* of two subspaces Y and Z of X, written

$$X = Y \oplus Z,$$

if each $x \in X$ has a unique representation

$$x = y + z \qquad y \in Y, z \in Z.$$

Then Z is called an *algebraic complement* of Y in X and vice versa, and Y, Z is called a *complementary pair* of subspaces in X.

In finite dimensions, we define the *orthogonal complement* $Y^\perp$ of a subspace Y as the subspace of all elements of our inner product space that are perpendicular to Y. The reason we call this space the orthogonal complement is that a finite-dimensional inner product space can be represented as $Y \oplus Y^\perp$ for any subspace Y. Can we generalize this to subspaces of infinite-dimensional spaces? Thanks to Lemma 11.2, the answer (for closed subspaces of Hilbert spaces) is yes.

Theorem 11.4. *Let Y be any closed subspace of a Hilbert space H. Then*

$$H = Y \oplus Y^\perp.$$

Notice that this theorem implies that we can take any vector $x \in H$ and and represent it uniquely as the sum of a $y \in Y$ and $z \in Y^\perp$. We can thus define a projection operator of H onto Y by

$$\begin{aligned} P : H &\longrightarrow Y \\ x &\longmapsto y = Px. \end{aligned}$$

This projection operator is an extension of the finite-dimensional projection operator, except with the restriction that it only applies to closed subspaces. Specifically, this operator maps H onto Y, Y onto itself, and $Y^\perp$ onto $\{0\}$.

Clearly, P is bounded, and since $\|x\| = \|y\| + \|z\|$, we have that $\sup_{\|x\|=1} \|Px\|$ is achieved when $\|z\| = 0$, so

$$\sup_{\|x\|=1} \|Px\| = \|x\| = 1.$$

Thus $\|P\| = 1$.

Furthermore, P is idempotent, i.e. $P^2 = P$ (meaning that if you apply P twice, you get the same end result as applying P once).

Having seen why Thm. 11.4 is important, we now turn to its proof.

Proof of Theorem 11.4. Since H is complete and Y is closed, Y must be complete. Since Y is convex, by Theorem 11.1 and Lemma 11.2 we have that for every $x \in H$ there exists $y \in Y$ such that for some $z \in Y^\perp$,

$$x = y + z.$$

To prove uniqueness, suppose that $x = y + z = y_1 + z_1$, where $y_1 \in Y$, $z_1 \in Y^\perp$. Then $y - y_1 = z - z_1$, but since $y - y_1 \in Y$ while $z - z_1 \in Y^\perp$ we must have that $y - y_1 \in Y \cap Y^\perp = \{0\}$. Thus $y = y_1$, and by the same logic $z = z_1$. ■

12. Orthonormal Sets

In this section, we take a brief look at orthonormal sets:

Definition 12.1. An *orthogonal set* in an inner product space is a subset whose elements are pairwise orthogonal. An *orthonormal set* is an orthogonal set whose elements all have norm 1.

If an orthogonal or orthonormal set is countable, we can arrange it in a sequence and call it an *orthogonal* or *orthonormal sequence*, respectively.

A property of orthonormal sequences is the Bessel inequality:

Theorem 12.2 (Bessel inequality)**.** *Let (e_j) be an orthonormal sequence in an inner product space X. Then for every $x \in X$,*

$$\sum_{j=1}^{\infty} |\langle x, e_j\rangle|^2 \leq \|x\|^2.$$

One great property of orthonormal sequences is that if we know that some x can be represented as an infinite linear combination of elements of an orthonormal sequence (e_j), then the orthonormality makes finding the coefficients quite easy. By the Bessel inequality, we then know that the norm of the infinite series of coefficients multiplied by the elements of orthonormal sequence is bounded. This fact greatly helps us in the search for "bases" of infinite-dimensional inner product spaces.

Proof of 12.2. The proof of the Bessel inequality requires a lot of algebra, but we prove it here since it is a frequently used inequality. This proof is adapted from [Con19].

We will start by showing that a few equations hold for finite orthonormal sets $\{e_1, \ldots, e_n\}$ using inner product axioms and identities defined earlier. Then, we'll derive the Bessel inequality for finite sets. Lastly, we'll use a limiting process to show that the Bessel inequality holds for the infinite case of an orthonormal sequence. The first such equality is that

$$\begin{aligned}
\left\langle x, \sum_{j=1}^{n} \langle x, e_j\rangle\, e_j \right\rangle &= \sum_{j=1}^{n} \langle x, \langle x, e_j\rangle e_j\rangle \\
&= \sum_{j=1}^{n} \overline{\langle \langle x, e_j\rangle\, e_j, x\rangle} \\
&= \sum_{j=1}^{n} \overline{\langle e_j, x\rangle \langle x, e_j\rangle} \\
&= \sum_{j=1}^{n} \langle x, e_j\rangle \overline{\langle x, e_j\rangle} \\
&= \sum_{j=1}^{n} |\langle x, e_j\rangle|^2.
\end{aligned}$$

Thus we have that

$$\begin{aligned}
\left\| x - \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k \right\|^2 &= \left\langle x - \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k, x - \sum_{j=1}^{n} \langle x, e_j\rangle\, e_j \right\rangle \\
&= \left\langle x, x - \sum_{j=1}^{n} \langle x, e_j\rangle\, e_j \right\rangle - \left\langle \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k, x - \sum_{j=1}^{n} \langle x, e_j\rangle\, e_j \right\rangle \\
&= \langle x, x\rangle - \left\langle x, \sum_{j=1}^{n} \langle x, e_j\rangle\, e_j \right\rangle - \left\langle \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k, x \right\rangle + \left\langle \sum_{k=1}^{n} \langle x, e_k\rangle\, e_k, \sum_{j=1}^{n} \langle x, e_j\rangle\, e_j \right\rangle
\end{aligned}$$

$$= \|x\|^2 - \left\langle x, \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\rangle - \overline{\left\langle x, \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\rangle} + \left\| \sum_{k=1}^{n} \langle x, e_k \rangle e_k \right\|^2$$

$$= \|x\|^2 - \left\langle x, \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\rangle - \overline{\left\langle x, \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\rangle} + \sum_{k=1}^{n} \|\langle x, e_k \rangle e_k\|^2$$

$$= \|x\|^2 - \left\langle x, \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\rangle - \overline{\left\langle x, \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\rangle} + \sum_{k=1}^{n} |\langle x, e_k \rangle|^2$$

$$= \|x\|^2 - \sum_{j=1}^{n} |\langle x, e_j \rangle|^2 - \overline{\sum_{j=1}^{n} |\langle x, e_j \rangle|^2} + \sum_{k=1}^{n} |\langle x, e_k \rangle|^2$$

$$= \|x\|^2 - 2\sum_{j=1}^{n} |\langle x, e_j \rangle|^2 + \sum_{k=1}^{n} |\langle x, e_k \rangle|^2$$

$$= \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k \rangle|^2.$$

Now, given that we are looking at a norm in this second equation, we have that

$$0 \leq \left\| x - \sum_{k=1}^{n} \langle x, e_k \rangle e_k \right\|^2 = \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k \rangle|^2,$$

so consequently we have the Bessel inequality for finite sets:

$$\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

Since for each k, we have that

$$|\langle x, e_k \rangle|^2 \geq 0,$$

we have that the sequence (y_n), where

$$y_n = \sum_{k=1}^{n} |\langle x, e_k \rangle|^2,$$

is bounded and increasing. Thus from the monotone convergence theorem, we have that (y_n) converges, say to some $y \in \mathbb{R}$. Since $y_n \leq \|x\|^2$ for each n, we have that $\|h\|^2 \geq \lim y_n = y$. In other words,

$$\|h\|^2 \geq \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2.$$

■

13. Bases of Infinite-Dimensional Spaces

A major proof technique in linear algebra is the use of coordinate representations of a vector under a certain basis. While this is not generalizable to infinite-dimensional spaces, as any such space must have an uncountably large basis (in the sense of linear algebra), we can modify this technique slightly to get it to work in our new setting. In this section, we'll look at a few different types of bases for infinite-dimensional spaces, ultimately leading up to the idea of a complete orthonormal basis. Before discussing these types of bases, we introduce the idea of a separable space:

Definition 13.1. A metric space X is said to be *separable* if it has a countable subset M which is dense in X, i.e. the closure of M is X.

The reason separability is important is that if our Hilbert space is separable, then it can be thought of as small enough to have a countable "basis" of orthonormal elements. Here, we allow for elements of our separable space to be represented as the linear combination of infinitely many elements of the basis. As mentioned earlier, if we only allow for finite linear combinations of basis vectors then our basis must be uncountable. As an aside, note that our new definition of basis is called a *Schauder basis*, while the definition from linear algebra is also known as the definition of a *Hamel basis*. Every nonzero vector space has a Hamel basis (see [Kre91]); however, not every vector space has a Schauder basis. In particular:

Lemma 13.2. *Any nonseparable vector space does not have a Schauder basis.*

Proof. Let V be a vector space, and suppose we have a Schauder basis (e_i) of V. Then for any $x \in V$, we have some (α_i) such that

$$\sum_{i=1}^{\infty} \alpha_i e_i = x.$$

Now, consider the set $Q \subseteq V$ whose elements can be expressed as

$$\sum_{i=0}^{\infty} q_i e_i,$$

where each q_i is rational. Since the number of rational numbers in $\mathbb{R}$, the number of numbers with rational coefficients in $\mathbb{C}$, and the number of elements in our Schauder basis are all countable, we have that Q is countable.

For each α_i, define a sequence of rational approximations (a_{ij}), which must exist by the density of rationals. Then we have a sequence $(x_j) \longrightarrow x$, where

$$x_j = \sum_{i=0}^{\infty} a_{ij} e_j.$$

Note that for all j, $x_j \in Q$, and since Q is countable we must have that the space V is separable. ∎

Our discussion over the last few paragraphs leads to the following defintion:

Definition 13.3. An orthonormal sequence (e_i) in an inner-product space H (which could be finite or infinite) is said to be *maximal* if for $\langle u, e_i \rangle = 0$ for all i means that $u = 0$ for all $u \in H$. We call a maximal orthonormal sequence in a separable Hilbert space a *complete orthonormal basis.*

Essentially, a maximal orthonormal sequence is the largest orthonormal sequence of nonzero elements of an inner-product space.

14. Separable Hilbert Spaces

We now look more closely at separable Hilbert spaces and ultimately generalize one of the major results from finite dimensional linear algebra, which says that that any vector space (or even inner product space) of dimension n over the field $\mathbb{F}$ is isomorphic to $\mathbb{F}^n$:

Theorem 14.1. *A finite-dimensional vector space V over a field $\mathbb{F}$, with $\dim V = n$, is isomorphic to $\mathbb{F}^n$.*

Proof. Choose a basis $v_1, \ldots, v_n$ of V. Then we can write any vector $v \in V$ as

$$v = a_1 v_1 + \cdots + a_n v_n.$$

Now take a basis $e_1, \ldots, e_n$ of $\mathbb{F}^n$. Then there exists a unique element $x \in \mathbb{F}^n$ where

$$x = a_1 e_1 + \cdots + a_n e_n.$$

Thus we define our isomorphism between V and $\mathbb{F}^n$ by mapping every v to its corresponding x. ■

Clearly, the proof above relies on the basis representation of each vector $v \in V$. Unfortunately, we can't utilize such representations for every infinite-dimensional vector space, but in this section we will extend this theorem and its proof technique to the case of separable Hilbert spaces.

We have defined complete orthonormal bases in separable Hilbert spaces; however, before discussing further about complete orthonormal bases, let's verify their existence:

Theorem 14.2. *Every separable inner-product space contains a maximal orthonormal sequence.*

Note that Theorem 14.2 states that any separable *inner-product space* contains a maximal orthonormal set. However, the existence of such sets outside of Hilbert spaces does not matter very much because these sets do not always act like bases outside of these spaces. This is because in incomplete spaces, there is no guarantee that a linear combination of basis elements (which is Cauchy, as we'll prove in the proof of Thm. 14.3) converges.

Now, to prove Theorem 14.2, we will use the *Gram-Schmidt procedure* from linear algebra (see [Axl97] for the finite-dimensional case).

Proof of Theorem 14.2. Take a countable dense subset, i.e. one that can be arranged as a sequence (v_j) and which exists since our space is separable. We will orthonormalize (v_j) to a sequence (e_j) using the aforementioned Gram-Schmidt process while keeping the span of this sequence the same.

Take the first element of the series satisfying $v_j \neq 0$, and then set

$$e_1 = \frac{v_j}{\|v_j\|}.$$

Now, suppose that for the first n elements $v_1, \ldots, v_n$ we have found m nonzero orthonormal elements $e_1, \ldots, e_m$ where $m \leq n$ and

$$\operatorname{span}(e_1, \ldots, e_m) = \operatorname{span}(v_1, \ldots, v_n).$$

If v_{n+1} is in the span of $e_1, \dots, e_m$, then our above equation holds for $v_1, \dots, v_{n+1}$ in place of $v_1, \dots, v_n$. Thus assume that $v_{n+1} \notin \operatorname{span}(e_1, \dots, e_m)$. It follows that

$$w = v_{n+1} - \sum_{j=1}^{\infty} \langle v_{n+1}, e_j \rangle e_j \neq 0,$$

so that we can set

$$e_{m+1} = \frac{w}{\|w\|}.$$

Now notice that $e_{m+1} \perp \{e_1, \dots, e_m\}$.

We can continue this process indefinitely, and ultimately we will either get an orthonormal set which may be finite or infinite and which we can arrange into a sequence (e_j). Now, we claim that any vector $u \in H$ that is perpendicular to each e_j is the zero vector. We prove this using the density of the elements v_n.

Choose u such that $u \perp e_j \, \forall j$. Since $\{v_j\}$ is dense in H, we have that there must be a sequence (w_k) of elements of $\{v_j\}$ with $(w_k) \longrightarrow u$. Each w_k is a finite linear combination of elements e_j by construction, so by the Bessel inequality

$$\|w_k\|^2 = \sum_{j \in \mathbb{N}} |\langle w_k, e_j \rangle|^2 = \sum_{j \in \mathbb{N}} |\langle u - w_k, e_j \rangle|^2 \leq \|u - w_k\|^2,$$

which we get because of the fact that $\langle u, e_j \rangle = 0 \, \forall j$. Thus we see that evidently $\|w_k\| \longrightarrow 0$, so $u = 0$. ■

One fact we will want to verify is that our "basis" indeed has a representation of each element in its Hilbert space. Indeed, we will prove this below.

Theorem 14.3. *If (e_j) is a complete orthonormal basis in a Hilbert space H then for any element $u \in H$ the 'Fourier-Bessel series' converges to u:*

$$u = \sum_{j=1}^{\infty} \langle u, e_j \rangle e_i.$$

Proof. Consider the sequence of partial sums of the Fourier-Bessel series

$$u_N = \sum_{j=1}^{N} \langle u, e_j \rangle e_j.$$

If $m < n$, then

$$\|u_n - u_m\|^2 = \sum_{j=m+1}^{n} |\langle u, e_j \rangle|^2 \leq \sum_{j > m} |\langle u, e_j \rangle|^2,$$

which is small for large m by the Bessel inequality (Thm. 12.2). Since we assume that our space H is complete, note that $u_n \longrightarrow w \in H$. Now, notice that for all $m > j$ we have that $\langle u_m, e_j \rangle = \langle u, e_j \rangle$ and furthermore $|\langle w - u_n, e_j \rangle| \leq \|w - u_n\|$. Thus

$$\langle w, e_j \rangle = \lim_{m \to \infty} \langle u_m, e_j \rangle = \langle u, e_j \rangle$$

for each j. Thus we have that $\langle u - w, e_j \rangle = 0$ for all j and so by the definition of a complete orthonormal basis $u - w = 0$ and thus $u = w$. ■

The existence of a complete orthonormal basis leads us to the following result:

Theorem 14.4. *Any infinite-dimensional separable Hilbert space is isomorphic to ℓ^2.*

Now, in itself this result is remarkable. However, it is even more astonishing when considered as an extension of the aformentioned linear algebra result about spaces of the same dimension over the same field being isomorphic. This extension is especially of note when one considers that the inner product on ℓ^2 is an extension of the finite-dimensional dot product (otherwise known as the inner product on $\mathbb{F}^n$, where $\mathbb{F}$ is either $\mathbb{C}$ or $\mathbb{R}$) to infinite dimensions. Let's prove the isomorphism (the proof is adapted from [Mel14]).

Proof of Theorem 14.4. The theorem is equivalent to saying that there exists a bijective linear map

$$T : H \longrightarrow \ell^2$$

such that for all $u, v \in H$,

$$\langle Tu, Tv\rangle_{\ell^2} = \langle u, v\rangle_H \qquad \|Tu\|_{\ell^2} = \|u\|_H.$$

Thus we choose an orthonormal basis (which must exist by Thm. 14.2) and set

$$Tu = (\langle u, e_j\rangle)_{j=1}^{\infty}.$$

By the Bessel inequality (12.2), T maps H onto ℓ^2. Since all entries in the sequence are linear in u, T is linear.

Furthermore, T is injective since whenever $Tu = 0$, $\langle u, e_i\rangle = 0$ for all i and thus $u = 0$ by the completeness of the basis.

Lastly, to show that T is surjective note that if $(c_j)_{j=1}^{\infty} \in \ell^2$ then

$$u = \sum_{j=1}^{\infty} c_j e_j$$

converges in H. By the same argument as our last proof, the sequence of partial sums is Cauchy since if $n > m$,

$$\left\| \sum_{j=m+1}^{n} c_j e_j \right\|_H^2 = \sum_{j=m+1}^{n} |c_j|^2.$$

Since the inner product is continuous by 10.4, $Tu = (c_j)$ so T is surjective.

The norms are equal because for u as above,

$$\|Tu\|_{\ell^2} = \sqrt{\sum_{j=1}^{\infty} c_j^2} = \sqrt{\sum_{j=1}^{\infty} \langle u, e_j\rangle} = \sqrt{\langle u, u\rangle} = \|u\|_H.$$

■

To sum up the significance of this theorem, we have that any separable complex Hilbert space is isomorphic to either $\mathbb{C}^n$ or ℓ^2, based on the size of its complete orthonormal basis. Now, this makes intuitive sense because the representation of some vector v with respect to an orthonormal basis $e_1, \ldots, e_n$ on $\mathbb{C}^n$ is given by

$$\sum_{j=1}^{n} \langle v, e_j\rangle\, e_j,$$

while the representation on ℓ^2 with respect to an orthonormal basis (e_j) is given by

$$\sum_{j=1}^{\infty} \langle v, e_j \rangle.$$

15. Riesz representation theorem

In finite-dimensional linear algebra, the Riesz representation is relatively trivial, in that it is a direct consequence of the basis representation of vectors in the dual space.

Theorem 15.1 (Riesz representation, finite dimensions)**.** *Suppose V is a finite-dimensional vector space. Then for every linear functional f on V, there exists some $z_f \in V$ such that for all $x \in V$,*

$$f(x) = \langle x, z_f \rangle.$$

Essentially, the theorem states that any inner product in a finite-dimensional space can be represented by a linear functional over that same space. The proof is standard from linear algebra (see [Axl97] for more details about the Riesz representation theorem in finite dimensions).

Proof of Thm 15.1. First we show there exists a vector $z_f \in V$ such that $f(x) = \langle x, z_f \rangle$ for all $x \in V$. Let $e_1, \ldots, e_n$ be an orthonormal basis of V. Then $\forall x \in V$,

$$\begin{aligned} f(x) &= f(\langle x, e_1 \rangle e_1 + \cdots + \langle x, e_n \rangle e_n) \\ &= \langle x, e_1 \rangle f(e_1) + \cdots + \langle x, e_n \rangle f(e_n) \\ &= \langle x, \overline{f(e_1)}\, e_1 + \cdots + \overline{f(e_n)}\, e_n \rangle. \end{aligned}$$

We let

$$z_f = \overline{f(e_1)}\, e_1 + \cdots + \overline{f(e_n)}\, e_n,$$

and now we have $f(x) = \langle x, z_f \rangle$, as desired.

For uniqueness, assume that there exist $w, z \in V$ with $f(x) = \langle x, w \rangle = \langle x, z \rangle$. Then we have that $\langle x, w - z \rangle = 0\, \forall x \in H$, so taking $x = w - z$ we get

$$\langle w - z, w - z \rangle = \|w - z\|^2 = 0,$$

so $w - z = 0$ and thus $w = z$, showing that the representation is unique. ■

The Riesz representation theorem can be extended to arbitrary Hilbert spaces, but we must impose the restriction that f is bounded for the theorem to be true. Ultimately, since in this case the space of bounded linear functionals is the dual space, we again establish the same connection between linear functionals and inner products as in the finite-dimensional case.

Theorem 15.2. *[Riesz representation] Every bounded linear functional f on a Hilbert space H can be represented in terms of the inner product, namely, $f(x) = \langle x, z_f \rangle$ where z_f depends on and is uniquely determined by f and has norm $\|z_f\| = \|f\|$.*

Proof. We start by proving that f has a representation. In the case $f = 0$, then the theorem holds if we take $z = 0$. Thus let $f \neq 0$. Now, let's look at the properties that z_f must have if a representation exists.

Firstly, $z_f \neq 0$ since otherwise $f = 0$. Then also $\langle x, z_f \rangle = 0$ iff $f(x) = 0$, so $\langle x, z_f \rangle = 0$ exactly when x is in the null space null f. Since the inner product of z_f with any element of null f is 0, we can surmise that $z_f \perp$ null f, so $z_f \in (\text{null } f)^\perp$. Now, we know that null f is a closed vector space. Since $f \neq 0$, null $f \neq H$ so that $(\text{null } f)^\perp \neq \{0\}$ by Theorem 11.4. Thus there is a nonzero $z_0 \in (\text{null } f)^\perp$. Let $v = f(x)z_0 - f(z_0)x$ where $x \in H$ is arbitrary. Thus $v \in$ null f since $f(v) = f(x)f(z_0) - f(z_0)f(x) = 0$. Thus since $z_0 \perp$ null f,

$$0 = \langle v, z_0\rangle = \langle f(x)z_0 - f(z_0)x, z_0\rangle = f(x)\langle z_0, z_0\rangle - f(z_0)\langle x, z_0\rangle.$$

Since $\langle z_0, z_0\rangle = \|z_0\|^2 \neq 0$, we can solve for $f(x)$ to get

$$f(x) = \frac{f(z_0)}{\langle z_0, z_0\rangle}\langle x, z_0\rangle.$$

Setting

$$z = \frac{\overline{f(z_0)}}{\langle z_0, z_0\rangle} z_0,$$

we get that $f(x) = \langle x, z\rangle$ as desired, and since $x \in H$ was arbitrary we have proven the existence of a representation.

Now we prove the uniqueness of the representation. Suppose we have two nonzero elements $w, z \in (\text{null } f)^\perp$ with $f(x) = \langle x, w\rangle = \langle x, z\rangle \,\forall x \in H$. Then we have that $\langle x, w - z\rangle = 0 \,\forall x \in H$, so taking $x = w - z$ we get

$$\langle w - z, w - z\rangle = \|w - z\|^2 = 0,$$

so $w - z = 0$ and thus $w = z$, showing that the representation is unique.

Lastly, we prove that $\|z_f\| = \|f\|$. Note that if $f = 0$, then $\|f\| = 0$ and since $z_f = 0$ $\|f\| = \|z_f\|$. Now take $f \neq 0$, and note that $z_f \neq 0$. Then we have that

$$\|z_f\|^2 = \langle z_f, z_f\rangle = f(z) \leq \|f\|\|z_f\|.$$

We can divide by $\|z_f\| > 0$ on both sides to get $\|z_f\| \leq \|f\|$. Now by the Schwarz inequality 8.2, we have that

$$|f(x)| = |\langle x, z_f\rangle| \leq \|x\|\|z\|.$$

Thus

$$\|f\| = \sup_{\|x\|=1} |\langle x, z_f\rangle| \leq \left\langle \frac{z_f}{\|z_f\|}, z_f \right\rangle = \|z_f\|.$$

Thus it is evident that $\|f\| = \|z_f\|$, concluding our proof. ■

We conclude this section with a quick corollary demonstrating the power of Riesz representation that strengthens the connection we saw previously between ℓ^2 and $\mathbb{F}^n$:

Corollary 15.3. *The dual space of the real space ℓ^2 is ℓ^2.*

Proof. An isomorphism of $(\ell^2)'$ onto ℓ^2 is $f \longmapsto z_f$, where z_f is

$$f(x) = \langle x, z_f\rangle.$$

This mapping is a bijection by the Riesz representation theorem 15.2. Note that this mapping is conjugate linear for ℓ^2 by $\alpha f \longmapsto \overline{\alpha} z_f$, so it is not an isomorphism. ■

This is significant because this again parallels the finite-dimensional $\mathbb{F}^n$, where $(\mathbb{F}^n)' = \mathbb{F}^n$.

16. Further Reading

There are a great many applications of functional analysis. In this expository paper, many small results were shown; however, there were not many mentions of applications of this theory. Some applications can be found in solving integral equations, approximation theory, and quantum mechanics. The interested reader may wish to read through some of the books in the references of this paper to learn more functional analysis, and the author specifically recommends [Kre91] if the reader is interested in learning about its applications.

17. Acknowledgements

The author would like to thank Simon Rubenstein-Salzedo and Eric Frankel for helpful conversations and resources.

References

[Axl97] Sheldon Jay Axler. *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. Springer, New York, 1997. URL: http://linear.axler.net/.

[Car05] Neal L Carothers. *A short course on Banach space theory*. Number 64. Cambridge University Press, 2005.

[Con19] John B Conway. *A course in functional analysis*, volume 96. Springer, 2019.

[Kre91] Erwin Kreyszig. *Introductory functional analysis with applications*, volume 17. John Wiley & Sons, 1991.

[Mel14] Richard Melrose. Introduction to functional analysis: 18.102 – spring 2014, 2014. URL: https://math.mit.edu/~rbm/18-102-S14/.

[Qua11] Quartl. *Illustration of the unit circles of some p-norms*. Nov 2011. URL: https://commons.wikimedia.org/w/index.php?curid=17428655.

[R+76] Walter Rudin et al. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1976.

Email address: bhattacharya.arav05@gmail.com