MATH31001/MATH41001/MATH61001: LINEAR ANALYSIS

Chapter 2: Normed Vector Spaces

Vector Spaces and Dimension

Definition. A *vector space* (or linear space) over $\mathbb{R}$ (or $\mathbb{C}$) is a non-empty set V and binary operations $V \times V \to V : (x, y) \mapsto x + y$ (addition) and $\mathbb{R} \times V \to V : (\lambda, x) \mapsto \lambda x$ (or $\mathbb{C} \times V \to V : (\lambda, x) \mapsto \lambda x$) (scalar multiplication) such that

(1) $(x + y) + z = x + (y + z)$, for all $x, y, z \in V$;
(2) $x + y = y + x$, for all $x, y \in V$;
(3) there exists $0 \in V$ such that $x + 0 = x$, for all $x \in V$;
(4) for any $x \in V$, there exists $-x \in V$ such that $x + (-x) = 0$;
(5) $\lambda(x + y) = \lambda x + \lambda y$, for all $\lambda \in \mathbb{R}$ (or $\mathbb{C}$), $x, y \in V$;
(6) $(\lambda + \mu)x = \lambda x + \mu x$, for all $\lambda, \mu \in \mathbb{R}$ (or $\mathbb{C}$), $x \in V$;
(7) $\lambda(\mu x) = (\lambda\mu)x$, for all $\lambda, \mu \in \mathbb{R}$ (or $\mathbb{C}$), $x \in V$; and
(8) $1x = x$, for all $x \in V$.

Definition. We say that a set $\{x_\alpha\}_{\alpha \in I} \subset V$ is *linearly independent* if, for any finite collection $\{\alpha_1, \dots, \alpha_k\} \subset I$,

$$\sum_{i=1}^{k} \lambda_{\alpha_i} x_{\alpha_i} = 0 \ (\lambda_{\alpha_i} \text{ scalars}) \qquad \Longrightarrow \qquad \lambda_{\alpha_1} = \cdots = \lambda_{\alpha_k} = 0.$$

If $\{x_\alpha\}_{\alpha \in I}$ is not linearly independent then we say that it is linearly dependent.

Definition. We say that a set $\{x_\alpha\}_{\alpha \in I} \subset V$ *spans* V if every $x \in V$ can be written in the form

$$x = \sum_{i=1}^{k} \lambda_{\alpha_i} x_{\alpha_i},$$

for some indices $\{\alpha_1, \dots, \alpha_k\}$ and scalars λ_{α_i}.

Definition. If $\{x_\alpha\}_{\alpha \in I} \subset V$ is linearly independent and spans V then we say that it is a *basis* for V. (Strictly speaking, we should call such a set a Hamel basis.)

We say that V is *finite dimensional* if it has a finite basis $\{x_1, \dots, x_n\}$, say. Part (i) of the next lemma shows that any other basis also has n elements and we say that V has dimension n ($\dim V = n$).

If V is not finite dimensional then we say that V is *infinite dimensional*.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Lemma 2.1 (Linear Algebra). *Let V be a vector space.*
(i) Suppose that $\{x_1, \dots, x_n\}$ is a basis for V. If $\{y_1, \dots, y_m\}$ is another basis for V then $n = m$ and we say that V has dimension n ($\dim V = n$).
(ii) Suppose that V has dimension n. Then any set in V containing at least $n+1$ elements is linearly dependent.

Example 1. $\mathbb{R}^n$ and $\mathbb{C}^n$ are finite dimensional and have dimension n(!). The standard basis is given by $\{e_i\}_{i=1}^n$, where $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ (with the 1 in the ith place). (Here, of course, we are considering $\mathbb{C}^n$ as a vector space over $\mathbb{C}$. $\mathbb{C}^n$ is also a vector space over $\mathbb{R}$, in which case the dimension is $2n$.)

However, from the point of view of analysis, the most interesting spaces are infinite dimensional.

Example 2. Let $V = C([0,1], \mathbb{R})$. Addition and scalar multiplication are defined pointwise: for $f, g \in C([0,1], \mathbb{R})$ and $\lambda \in \mathbb{R}$,

$$(f+g)(x) = f(x) + g(x),$$

$$(\lambda f)(x) = \lambda f(x).$$

We claim that $C([0,1], \mathbb{R})$ is infinite dimensional. To see this, we shall finid an infinite subset of $C([0,1], \mathbb{R})$ which is linearly independent. Consider $\{p_n\}_{n=0}^\infty$, where $p_n(x) = x^n$. If, for all $x \in [0,1]$,

$$\sum_{i=1}^k \lambda_{n_i} x^{n_i} = 0$$

then $\lambda_{n_1} = \cdots = \lambda_{n_k} = 0$, so $\{p_n\}_{n=0}^\infty$ is linearly independent. Thus, $C([0,1], \mathbb{R})$ cannot be finite dimensional: if it had dimension n the any set containing more that n elements would be linearly dependent (Lemma 2.1(ii)).

Example 3. $R[0,1]$ = the space of Riemann integrable functions $f : [0,1] \to \mathbb{R}$. This is an infinite dimensional space since it contains $C([0,1], \mathbb{R})$.

Example 4. Let

$$l^1 = \left\{ (x_i)_{i=1}^\infty : \sum_{i=1}^\infty |x_i| < +\infty, \ x_i \in \mathbb{C} \right\}.$$

(Check that l^1 is a vector space). Then l^1 is infinite dimensional. To see this consider $\{e_n\}_{n=1}^\infty$, where

$$e_n = (0, \dots, 0, 1, 0, \dots) \text{ (with the 1 in the } n\text{th place)}.$$

Clearly, for each $n \geq 1$, $e_n \in l^1$. Then

$$\sum_{i=1}^k \lambda_{n_i} e_{n_i}$$

is the vector in l^1 with λ_{n_i} in the n_ith place, $i = 1, \dots, k$, and zeros elsewhere. Thus if

$$\sum_{i=1}^{k} \lambda_{n_i} e_{n_i} = 0$$

then we must have $\lambda_{n_1} = \cdots = \lambda_{n_k} = 0$. Therefore $\{e_n\}_{n=1}^{\infty}$ is an infinite linearly independent set in l^1, so l^1 is infinite dimensional.

Example 5. Let

$$l^{\infty} = \left\{ (x_i)_{i=1}^{\infty} : \sup_{1 \le i < \infty} |x_i| < +\infty,\ x_i \in \mathbb{C} \right\}.$$

(Check that l^{∞} is a vector space). Then l^{∞} is infinite dimensional: as for l^1 the set $\{e_n\}_{n=1}^{\infty}$ provides an infinite linearly independent set.

Norms on Vector Spaces

Definition. A norm on a vector space V over $\mathbb{R}$ (or $\mathbb{C}$) is a function $\|\cdot\| : V \to \mathbb{R}$ such that

(1) $\|x\| \ge 0$ and $\|x\| = 0$ if and only if $x = 0$;
(2) $\|\lambda x\| = |\lambda| \|x\|$, for all $x \in V$ and $\lambda \in \mathbb{R}$ (or $\mathbb{C}$);
(3) $\|x + y\| \le \|x\| + \|y\|$, for all $x, y \in V$ (the *triangle inequality.*

Example 1. Let us first consider norms on $\mathbb{R}^n$ or $\mathbb{C}^n$. The most obvious norm is the Euclidean norm or 2-norm:

$$\|(x_1, x_2, \dots, x_n)\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

That this satisfies the fist two properties of a norm is obvious. The third follows from the Cauchy-Schwarz Inequality below.

Lemma 2.2 (Cauchy-Schwarz Inequality). *Suppose $a_i, b_i \in \mathbb{R}$, $i = 1, \dots, k$. Then*

$$\sum_{i=1}^{k} a_i b_i \le \left(\sum_{i=1}^{k} a_i^2 \right)^{1/2} \left(\sum_{i=1}^{k} b_i^2 \right)^{1/2}.$$

Proof. We have

$$\sum_{i=1}^{k} (a_i t + b_i)^2 \ge 0$$

for all $t \in \mathbb{R}$. By multiplying out each bracket, we may rewrite this inequality as

$$At^2 + 2Bt + C \ge 0,$$

where

$$A = \sum_{i=1}^{k} a_i^2, \quad B = \sum_{i=1}^{k} a_i b_i, \quad C = \sum_{i=1}^{k} b_i^2.$$

If $A > 0$ then the desired inequality follows by taking $t = -B/A$, so that $B^2 - AC \leq 0$. If $A = 0$ then $a_i = 0$, $i = 1, \ldots, k$, and the inequality is clearly satisfied. $\square$

We can use the Cauchy-Schwarz Inequality to check that $\|\cdot\|_2$ satsfies the triangle inequality on $\mathbb{R}^n$:

$$\begin{aligned}
\|(x_1, \ldots, x_n) + (y_1, \ldots, y_n)\|_2^2 &= \sum_{i=1}^{n} |x_i + y_i|^2 \\
&= \sum_{i=1}^{n} |x_i^2 + 2x_i y_i + y_i^2| \\
&\leq \sum_{i=1}^{n} \left(|x_i|^2 + 2|x_i||y_i| + |y_i|^2\right) \\
&\leq \sum_{i=1}^{n} |x_i|^2 + 2\left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2} + \sum_{i=1}^{n} |y_i|^2 \\
&= \left(\left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} + \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2}\right)^2 \\
&= \left(\|(x_1, \ldots, x_n)\|_2 + \|(y_1, \ldots, y_n)\|_2\right)^2,
\end{aligned}$$

where we used the Cauchy-Schwarz Inequality with $a_i = |x_i|$ and $b_i = |y_i|$.

There are other norms which can be defined on $\mathbb{R}^n$. In each case parts (1) and (2) of the definition of norm are clearly satisfied.

(1) the "1-norm"

$$\|(x_1, \ldots, x_n)\|_1 = \sum_{i=1}^{n} |x_i|.$$

For the triangle inequality we just use the usual triangle inequality:

$$\begin{aligned}
\|(x_1, \ldots, x_n) + (y_1, \ldots, y_n)\|_1 &= \sum_{i=1}^{n} |x_i + y_i| \\
&\leq \sum_{i=1}^{n} (|x_i| + |y_i|) \\
&= \|(x_1, \ldots, x_n)\|_1 + \|(y_1, \ldots, y_n)\|_1.
\end{aligned}$$

(2) the "supremum norm" or "∞-norm"

$$\|(x_1, \ldots, x_n)\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

The triangle inequality is easily verified.

(3) the "p-norm"

$$\|(x_1,\dots,x_n)\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}.$$

Of course, $p=2$ and $p=1$ are above. In this case, the triangle inequality follows from Minkowski's Inequality below. Note that one has

$$\|(x_1,\dots,x_n)\|_\infty = \lim_{p\to+\infty} \|(x_1,\dots,x_n)\|_p,$$

which motivates the notation $\|\cdot\|_\infty$.

Lemma 2.3 (Minkowski's Inequality). *For $p \geq 1$,*

$$\left(\sum_{i=1}^n |a_i+b_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^n |b_i|^p\right)^{1/p}.$$

Proof. Omitted. **(Extra reading for MATH41001/MATH61001.)** □

Remark. The proof uses Hölder's Inequality: if $p,q>1$ satisfy $1/p+1/q=1$ then, for $a_i, b_i \in \mathbb{C}$, $i=1,\dots,n$,

$$\sum_{i=1}^n |a_i b_i| \leq \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} \left(\sum_{i=1}^n |b_i|^q\right)^{1/q},$$

with equality if and only if $|a_i|^p/|b_i|^q$ is constant.

There are natural infinite dimensional analogues for all of these – but notice how we have to use the correct space each time ($l^2, l^1, l^\infty, \dots$) to have the norm defined.

Example 2. $l^2 = \left\{(x_i)_{i=1}^\infty : \sum_{i=1}^\infty |x_i|^2 < +\infty,\ x_i \in \mathbb{C}\right\}$ (which is an infinite dimensional space for the same reason that l^1 is) with the norm

$$\|(x_i)_{i=1}^\infty\|_2 = \left(\sum_{i=1}^\infty |x_i|^2\right)^{1/2}.$$

To prove that this satisfies the triangle inequality, notice that the Cauchy-Schwarz inequality continues to hold for infinite sums, with the same proof, provided that everything converges.

Example 3. $l^1 = \{(x_i)_{i=1}^\infty : \sum_{i=1}^\infty |x_i| < +\infty,\ x_i \in \mathbb{C}\}$ with the norm

$$\|(x_i)_{i=1}^\infty\|_1 = \sum_{i=1}^\infty |x_i|.$$

Remark. We have $l^1 \subset l^2$ but $l^1 \neq l^2$.

To see this, first take $(x_i)_{i=1}^\infty \in l^1$. Then $\sum_{i=1}^\infty |x_i| < +\infty$ so, in particular, $\lim_{i\to+\infty} x_i = 0$. Thus there exists $N \geq 1$ such that

$$i \geq N \quad \Longrightarrow \quad |x_i| < 1.$$

If $|x_i| < 1$ then $|x_i|^2 < |x_i|$, so this holds for all $i \geq N$. Applying the Comparison Test, we then have that $\sum_{i=1}^\infty |x_i|^2 < +\infty$, so that $(x_i)_{i=1}^\infty \in l^2$. Thus $l^1 \subset l^2$.

On the other hand, consider $(1, \frac{1}{2}, \frac{1}{3}, \dots)$. We know

$$\sum_{i=1}^\infty \frac{1}{i} = +\infty \quad \text{so} \quad \left(1, \frac{1}{2}, \frac{1}{3}, \dots\right) \notin l^1.$$

However,

$$\sum_{i=1}^\infty \frac{1}{i^2} < +\infty \quad \text{so} \quad \left(1, \frac{1}{2}, \frac{1}{3}, \dots\right) \in l^2.$$

Thus $l^1 \neq l^2$.

Example 4. $l^\infty = \left\{(x_i)_{i=1}^\infty : \sup_{1\leq i<\infty} |x_i| < +\infty, \ x_i \in \mathbb{C}\right\}$ with the norm

$$\|(x_i)_{i=1}^\infty\|_\infty = \sup_{1\leq i<\infty} |x_i|.$$

Example 5. For $p \geq 1$, $l^p = \{(x_i)_{i=1}^\infty : \sum_{i=1}^\infty |x_i|^p < +\infty, \ x_i \in \mathbb{C}\}$ (again this is an infinite dimensional space) with the norm

$$\|(x_i)_{i=1}^\infty\|_p = \left(\sum_{i=1}^\infty |x_i|^p\right)^{1/p}.$$

That the triangle inequality holds in this case follows from the fact that Minkowski's Inequality continues to hold for infinite sums, with the same proof, provided that everything converges.

Exercise. For $1 \leq p < q < +\infty$ show that $l^p \subset l^q$ but $l^p \neq l^q$. Also show that $l^p \subset l^\infty$ but $l^p \neq l^\infty$.

Example 6. Now we consider the infinite dimensional space $C([0,1], \mathbb{R})$. There are several norms we can define here:

(a) "uniform norm" or "supremum norm" or "∞-norm"

$$\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|.$$

(b)

$$\|f\|_1 = \int_0^1 |f(x)| dx.$$

(c)

$$\|f\|_2 = \left(\int_0^1 |f(x)|^2 dx \right)^{1/2}.$$

For the triangle inequality, use the Cauchy-Schwarz inequality for integrals:

$$\left| \int_0^1 f(x)g(x)dx \right| \leq \left(\int_0^1 |f(x)|^2 dx \right)^{1/2} \left(\int_0^1 |g(x)|^2 dx \right)^{1/2}.$$

(d) for $0 < p < +\infty$,

$$\|f\|_p = \left(\int_0^1 |f(x)|^p dx \right)^{1/p}.$$

For the triangle inequality, use Minkowski's inequality for integrals:

$$\int_0^1 |f(x) + g(x)|^p dx \leq \left(\int_0^1 |f(x)|^p dx \right)^{1/p} + \left(\int_0^1 |g(x)|^p dx \right)^{1/p}.$$

We have put $\|\cdot\|_\infty$ first because it is the most important one or, more accurately, the one best suited to the space $C([0,1],\mathbb{R})$. We shall see in what respect the other norms are lacking later.

Metrics and Topology

Let $\|\cdot\|$ be a norm on the vector space V. This defines a metric on V by $d(x,y) = \|x-y\|$. The axioms for a metric:

(1) $d(x,y) \geq 0$ and $d(x,y) = 0$ if and only if $x = y$;
(2) $d(x,y) = d(y,x)$; and
(3) $d(x,y) \leq d(x,z) + d(z,y)$

follow from the corresponding properties of the norm $\|\cdot\|$.

Recall that we write

$$B(x,r) = \{y \in V : d(x,y) < r\}$$

(the open ball of radius r around $x \in V$). This is an open set.

Recall that, in general, a set U is open, if, for every $x \in U$, there exists $r > 0$ such thaT $B(x,r) \subset U$. **Note that this definition depends on the choice of norm/metric.**

Let $\mathcal{U}$ denote the collection of all open sets in V (with the norm $\|\cdot\|$). They form a *topology*:

(i) $\varnothing, V \in \mathcal{U}$;
(ii) (closure under finite intersections)

$$U_1, \ldots, U_k \in \mathcal{U} \quad \Longrightarrow \quad \bigcap_{i=1}^k U_i \in \mathcal{U};$$

(iii) (closure under arbitrary unions: the index set $\mathcal{A}$ does not even have to be countable)

$$\{U_\alpha\}_{\alpha \in \mathcal{A}} \subset \mathcal{U} \quad \Longrightarrow \quad \bigcup_{\alpha \in \mathcal{A}} U_\alpha \in \mathcal{U}.$$

Suppose we have two norms, $\|\cdot\|$ and $\|\cdot\|'$, on V, each giving rise to a topology on V. When are these topologies the same (i.e. that a set is open for one norm if and only if it is open for the other)? A necessary and sufficient condition is given by the following.

Definition. We say that two norms $\|\cdot\|$ and $\|\cdot\|'$ are equivalent ($\|\cdot\| \sim \|\cdot\|'$) if there exist $C_1, C_2 > 0$ such that

$$C_1\|x\|' \leq \|x\| \leq C_2\|x\|',$$

for all $x \in V$.

(This is obviously an equivalence relation. In particular, if $\|\cdot\| \sim \|\cdot\|'$ and $\|\cdot\|' \sim \|\cdot\|''$ are equivalent then $\|\cdot\| \sim \|\cdot\|''$.)

Lemma 2.4. *Two norms on V give the same topology (i.e. the same open sets) if and only if they are equivalent.*

Proof. Omitted. □

Example. Let's check that $(\mathbb{R}^n, \|\cdot\|_1)$ and $(\mathbb{R}^n, \|\cdot\|_2)$ are equivalent.

Write $x = (x_1, \dots, x_n)$. Then

$$\|x\|_2^2 = |x_1|^2 + \cdots + |x_n|^2 \leq (|x_1| + \cdots + |x_n|)^2 = \|x\|_1^2,$$

so

$$\|x\|_2 \leq \|x\|_1.$$

On the other hand,

$$\begin{aligned}\|x\|_1^2 &= (|x_1| + \cdots + |x_n|)^2 \\ &= |x_1|^2 + \cdots |x_n|^2 + 2|x_1x_2| + 2|x_1x_3| + \cdots + 2|x_{n-1}x_n|.\end{aligned}$$

Now, for $i \neq j$,

$$|x_i|^2 - 2|x_ix_j| + |x_j|^2 = (|x_i| - |x_j|)^2 \geq 0,$$

so

$$2|x_ix_j| \leq |x_i|^2 + |x_j|^2 \leq \|x\|_2^2.$$

Thus

$$\|x\|_1^2 \leq \|x\|_2^2 + \frac{n(n-1)}{2}\|x\|_2^2 < n^2\|x\|_2^2,$$

so

$$\|x\|_1 \leq n\|x\|_2.$$

Lemma 2.5. *On $\mathbb{R}^n$ (and $\mathbb{C}^n$), all norms are equivalent.*

Proof. Let $\|\cdot\|_1$ be the 1-norm on $\mathbb{R}^n$ and let $\|\cdot\|$ be an arbitrary norm. We shall show that $\|\cdot\|_1$ and $\|\cdot\|$ are equivalent.

As usual e_i is the basis vector with 1 in the ith place and 0 elsewhere. Write $x = (x_1, \dots, x_n) = \sum_{i=1}^n x_ie_i$ and let $M = \max_{1\leq i\leq n} ||e_i||$. Then

$$\|x\| \leq \sum_{i=1}^n |x_i|\|e_i\| \leq M\sum_{i=1}^n |x_i| = M\|x\|_1.$$

Now we shall show that $\inf_{0\neq x\in\mathbb{R}^n}\|x\|/\|x\|_1$ is positive. If it isn't, then we can find a sequence x_i such that

$$\lim_{i\to+\infty}\|x_i\|/\|x_i\|_1=0.$$

Set $y_i=x_i/\|x_i\|_1$, so

$$y_i\in\{y\in\mathbb{R}^n:\ \|y\|_1\le 1\}.$$

This set is closed and bounded (hence compact), so y_i has a convergent subsequence y_{i_j} with limit y. In other words $\lim_{j\to+\infty}\|y_{i_j}-y\|_1=0$ and (since $\|y_{i_j}\|_1=1$) $\|y\|_1\neq 0$. However, we also have

$$\big|\|y_{i_j}\|-\|y\|\big|\le\|y_{i_j}-y\|\le M\|y_{i_j}-y\|_1\to 0,\ \text{as } j\to+\infty,$$

so $\|y\|=\lim_{j\to+\infty}\|y_{i_j}\|=0$. But $\|y\|=0$ if and only if $y=0$, giving a contadiction to $\|y\|_1\neq 0$. Therefore, we can define

$$0<m=\inf_{0\neq x\in\mathbb{R}^n}\frac{\|x\|}{\|x\|_1}.$$

Clearly, $\|x\|\ge m\|x\|_1$, as required. □

However, this result is not true for infinite dimensional spaces.

Example. Consider the space $C([0,1],\mathbb{R})$. Then the two norms $\|f\|_\infty=\sup_{x\in[0,1]}|f(x)|$ and $\|f\|_1=\int_0^1|f(x)|dx$ are *not* equivalent. For example, consider the sequence

$$f_n(x)=\begin{cases}1-nx & \text{if } 0\le x\le 1/n\\ 0 & \text{if } 1/n<x\le 1\end{cases}.$$

Then $\|f_n\|_\infty=1$, for all n, but

$$\|f_n\|_1=\int_0^{1/n}(1-nx)dx=\left[x-\frac{nx^2}{2}\right]_0^{1/n}=\frac{1}{n}-\frac{1}{2n}=\frac{1}{2n}\to 0,$$

as $n\to+\infty$.

Banach Spaces

Recall that a sequence x_n in the normed space $(V,\|\cdot\|)$ is called a *Cauchy sequence* if, for every $\epsilon>0$, there exists $N\ge 1$ such that

$$n,m\ge N\quad\Longrightarrow\quad\|x_n-x_m\|<\epsilon.$$

Definition. A normed space $(V,\|\cdot\|)$ is said to be *complete* if every Cauchy sequence in V converges to a point in V (i.e. if x_n is a Cauchy sequence in V then there exists $x\in V$ such that $x_n\to x$, as $n\to+\infty$. (Of course, $x_n\to x$ means $\|x_n-x\|\to 0$.)

Definition. A normed vector space $(V,\|\cdot\|)$ which is also complete is called a *Banach space.*

Remark. Suppose that $\|\cdot\|\sim\|\cdot\|'$. Then $(V,\|\cdot\|)$ is complete if and only if $(V,\|\cdot\|')$ is complete.

Remark. Banach spaces are named after Stefan Banach (1892-1945). He introduced them and developed their theory in the 1920s and 1930s.

Example 1. $\mathbb{R}^n$ and $\mathbb{C}^n$ (with any norm) are Banach spaces. (Why?)

Theorem 2.6. *$C([0,1],\mathbb{R})$ with the norm $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$ is a Banach space. (So is $C([0,1],\mathbb{C})$ with the same proof.)*

Proof. Suppose that $\{f_n\}_{n=1}^\infty$ is a Cauchy sequence for $\|\cdot\|_\infty$. This means that, given $\epsilon > 0$, there exists $N \geq 1$ such that, for $n, m \geq N$ and for any $x \in [0,1]$,

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty < \epsilon. \tag{$*$}$$

In other words, $\{f_n(x)\}_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{R}$ and so, since $\mathbb{R}$ is complete (Example 1), it has a limit $f(x)$, say. We may now let $m \to +\infty$ in $(*)$ to obtain, for $n \geq N$,

$$|f_n(x) - f(x)| = \lim_{m\to+\infty} |f_n(x) - f_m(x)| \leq \limsup_{m\to+\infty} \|f_n - f_m\|_\infty \leq \epsilon.$$

Now the above inequality tells us that $f_n(x)$ converges to $f(x)$ *uniformly* on $[0,1]$ and so, by a theorem in Metric Spaces, f is continuous (i.e., it's an element of $C([0,1],\mathbb{R})$). Furthermore, taking the supremum over $x \in [0,1]$ in the inequality gives that, for $n \geq N$,

$$\|f_n - f\|_\infty = \sup_{x\in[0,1]} |f_n(x) - f(x)| \leq \epsilon,$$

so $\lim_{n\to+\infty} \|f_n - f\|_\infty = 0$, as required. □

Remark. Note the three steps in the above proof:

(1) identify a potential limit f for the Cauchy sequence f_n;
(2) show that f is in the desired space (here $C([0,1],\mathbb{R})$);
(3) show that f_n converges to f in the appropriate norm (here $\|\cdot\|_\infty$).

We will use this scheme again in the next theorem.

Theorem 2.7. *l^1 with the norm $\|\cdot\|_1$ is a Banach space.*

Proof. Suppose that $\{(x_i^{(n)})_{i=1}^\infty\}_{n=1}^\infty$ is a Cauchy sequence for $\|\cdot\|_1$. Then, given $\epsilon > 0$, there exists $N \geq 1$ such that, for $n, m \geq N$,

$$\sum_{i=1}^\infty |x_i^{(n)} - x_i^{(m)}| = \|(x_i^{(n)})_{i=1}^\infty - (x_i^{(m)})_{i=1}^\infty\|_1 < \epsilon \tag{$*$}$$

and so, in particular, for any i,

$$|x_i^{(n)} - x_i^{(m)}| < \epsilon.$$

Thus, for each fixed i, $\{x_i^{(n)}\}_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{C}$ and so has a limit $x_i \in \mathbb{C}$.

Next, note that, for any $M \geq 1$ and $n, m \geq N$, we have

$$\sum_{i=1}^M |x_i^{(n)} - x_i^{(m)}| \leq \sum_{i=1}^\infty |x_i^{(n)} - x_i^{(m)}| < \epsilon.$$

If we let $m \to +\infty$, this gives

$$\sum_{i=1}^M |x_i^{(n)} - x_i| \leq \epsilon, \tag{$**$}$$

for any $M \geq 1$ and $n \geq N$. We have

$$\sum_{i=1}^{M} |x_i| \leq \sum_{i=1}^{M} |x_i^{(N)} - x_i| + \sum_{i=1}^{M} |x_i^{(N)}|$$
$$\leq \epsilon + \|(x_i^{(N)})_{i=1}^{\infty}\|_1.$$

Letting $M \to +\infty$, we see that $\sum_{i=1}^{\infty} |x_i|$ is finite, so $(x_i)_{i=1}^{\infty} \in l^1$.

Finally, letting $M \to +\infty$ in $(**)$ gives

$$\sum_{i=1}^{\infty} |x_i^{(n)} - x_i| \leq \epsilon,$$

for all $n \geq N$, so that $\lim_{n\to+\infty} \|(x_i^{(n)})_{i=1}^{\infty} - (x_i)_{i=1}^{\infty}\|_1 = 0$, as required. □

Theorem 2.8. *$(l^p, \|\cdot\|_p)$ $(p \geq 1)$ and $(l^\infty, \|\cdot\|_\infty)$ are Banach spaces.*

Proof. Omitted. **(Extra reading for MATH41001/MATH61001.)** □

Exercise. $C([0,1], \mathbb{R}), \|\cdot\|_1)$ (where $\|f\|_1 = \int_0^1 |f(x)|dx$) is *not* a Banach space.

Exercise. $(l^1, \|\cdot\|_2)$ is *not* a Banach space.

Inner Products and Hilbert Spaces

Hilbert spaces are special types of Banach spaces in which the norm is defined by an inner product. A particular feature of these spaces is that one has the notion of orthogonality. They are named after David Hilbert (1862-1943): he and Poincaré were the most important mathematicians of the early 20th century.

Definition. Let H be a vector space over $\mathbb{R}$ or $\mathbb{C}$. An *inner product* is a map $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ (or $\mathbb{C}$) such that, for all $x, y, z \in H$ and scalars λ, μ,

(1) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (complex conjugation);
(2) $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$; and
(3) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.

Of course, if the vector space is over $\mathbb{R}$ then (1) is just $\langle x, y \rangle = \langle y, x \rangle$.

This definition generalizes the familiar inner product (= dot product) on $\mathbb{R}^n$.

Example 1.

(a) $H = \mathbb{R}^n$, $\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$ is an inner product.
(b) $H = \mathbb{C}^n$, $\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y}_i$ is an inner product.

With the above example in mind, the following result is a more abstract version of the Cauchy-Schwart inequality.

Lemma 2.9 (Cauchy-Schwarz Inequality for Inner Products). *Let $\langle\cdot,\cdot\rangle$ be an inner product on H. Then*

$$|\langle x,y\rangle| \le \langle x,x\rangle^{1/2}\langle y,y\rangle^{1/2},$$

for all $x, y \in H$.

Proof. If $x=0$ or $y=0$, the result is immediate.

Suppose $x \neq 0$ and $y \neq 0$ and put

$$\lambda = -\frac{\overline{\langle x,y\rangle}}{\langle x,x\rangle}.$$

Then

$$\begin{aligned} 0 &\le \langle \lambda x + y, \lambda x + y\rangle \\ &= \left|\frac{\overline{\langle x,y\rangle}}{\langle x,x\rangle}\right|^2 \langle x,x\rangle - \frac{\overline{\langle x,y\rangle}\langle x,y\rangle}{\langle x,x\rangle} - \frac{\langle x,y\rangle\langle y,x\rangle}{\langle x,x\rangle} + \langle y,y\rangle \\ &= \frac{|\langle x,y\rangle|^2}{\langle x,x\rangle} - \frac{|\langle x,y\rangle|^2}{\langle x,x\rangle} - \frac{|\langle x,y\rangle|^2}{\langle x,x\rangle} + \langle y,y\rangle \\ &= -\frac{|\langle x,y\rangle|^2}{\langle x,x\rangle} + \langle y,y\rangle \end{aligned}$$

(using $\langle y,x\rangle = \overline{\langle x,y\rangle}$). Hence

$$|\langle x,y\rangle|^2 \le \langle x,x\rangle\langle y,y\rangle,$$

as required. □

Lemma 2.10. *Suppose that $\langle\cdot,\cdot\rangle$ is an inner product on H. Then $\|\cdot\|$ defined by $\|x\| = \langle x,x\rangle^{1/2}$ is a norm on H.*

Proof. It is immediate from the definition of inner product that $\|x\| \ge 0$ and $\|x\| = 0$ iff and only if $x = 0$. We also have

$$\|\lambda x\| = \langle \lambda x, \lambda x\rangle^{1/2} = \big(\lambda\overline{\lambda}\langle x,x\rangle\big)^{1/2} = |\lambda|\langle x,x\rangle^{1/2} = |\lambda|\|x\|.$$

Finally, we need to show that $\|x+y\| \le \|x\| + \|y\|$ (triangle inequality). We have

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + 2\Re\langle x,y\rangle + \langle y,y\rangle \\ &\le \langle x,x\rangle + 2\langle x,x\rangle^{1/2}\langle y,y\rangle^{1/2} + \langle y,y\rangle \\ &= (\langle x,x\rangle^{1/2} + \langle y,y\rangle^{1/2})^2 = (\|x\| + \|y\|)^2 \end{aligned}$$

(using the Cauchy-Schwarz inequality and $\Re z \le |z|$), as required. □

In view of this, we may restate the Cauchy-Schwarz inequality in a more convenient form.

Lemma 2.9' (Restatement of the Cauchy-Schwarz Inequality for Inner Products). *Let $\langle \cdot, \cdot \rangle$ be an inner product on H and let $\| \cdot \|$ be the associated norm. Then*

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

for all $x, y \in H$.

Corollary. *The map $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ is continuous (with respect to the associated norm).*

Proof. Suppose that $(x_n, y_n) \to (x, y)$ in $H \times H$, as $n \to +\infty$. Then $\lim_{n\to+\infty} \|x_n - x\| = 0$ and $\lim_{n\to+\infty} \|y_n - y\| = 0$. Using the Cauchy-Schwarz inequality,

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle| \\ &= |\langle x_n, y_n - y\rangle + \langle x_n - x, y\rangle| \\ &\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\| \to 0, \text{ as } n \to +\infty \end{aligned}$$

(since x_n converges, $\|x_n\|$ is bounded).

Now, at last, we can give the definition of a Hilbert space.

Definition. A *Hilbert space* is a vector space H with an inner product $\langle \cdot, \cdot \rangle$ such that H is complete with respect to the associated norm $\|x\| = \langle x, x\rangle^{1/2}$.

Another way of putting this is that a Hilbert space is a Banach space where the norm is given by an inner product.

Example 2. l^2 is a Hilbert space, with the inner product given by

$$\langle (x_i)_{i=0}^{\infty}, (y_i)_{i=0}^{\infty}\rangle = \sum_{i=0}^{\infty} x_i \overline{y}_i.$$

One easily sees that the associated norm is the 2-norm:

$$\langle (x_i)_{i=0}^{\infty}, (x_i)_{i=0}^{\infty}\rangle^{1/2} = \|(x_i)_{i=0}^{\infty}\|_2.$$

(We use the Cauchy-Schwarz inequality to ensure that the inner product is finite.)

Example 3. Let $H = M_n(\mathbb{C}) = n \times n$ complex matrices. This is a Hilbert space with respect to the inner product

$$\langle A, B\rangle = \text{trace}(AB^*),$$

where B^* is the matrix with (i, j) entry $B^*(i, j) = \overline{B(j, i)}$.

The following is a useful property of Hilbert spaces.

Lemma 2.11 (Parallelogram law). *Let H be a Hilbert space. For $x, y \in H$, we have the identity*

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

Proof. We have the identities

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

and

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

Adding them together gives the result. □

Orthogonal Complements in Hilbert Spaces

Let H be a Hilbert space over $\mathbb{R}$ (or $\mathbb{C}$) and let L be a linear subspace of H.

Definition. The *orthogonal complement* $L^\perp$ of L is defined by

$$L^\perp = \{x \in H : \langle x, y\rangle = 0 \text{ for all } y \in L\}.$$

Recall that if U, W are two subspaces of the vector space H then we write $H = U \oplus W$ (H is the direct sum of U and W) if $U \cap W = \{0\}$ and every $x \in H$ can be written $x = u + w$, with $u \in U$ and $w \in W$. Put another way, $H = U \oplus W$ if every $x \in H$ has a *unique* decomposition $x = u+w$, with $u \in U$ and $w \in W$. (Why are these two formulations equivalent?)

Lemma 2.12. *Let H be a Hilbert space and let L be a linear subspace. Then $L^\perp$ is a closed linear subspace of H and $H = \overline{L} \oplus L^\perp$. (Here $\overline{L}$ denotes the closure of L.)*

Proof. First we show that $L^\perp$ is linear. Suppose that $x, y \in L^\perp$, λ, μ are scalars and that $z \in L$. Then

$$\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle = 0 + 0 = 0,$$

so $\lambda x + \mu y \in L^\perp$.

Next we show that $L^\perp$ is closed. Suppose that $x_n \in L^\perp$ and that $x_n \to x$ in H, as $n \to +\infty$. For any $z \in L$, we have

$$\begin{aligned} |\langle x, z\rangle| &= |\langle x, z\rangle - \langle x_n, z\rangle| = |\langle x - x_n, z\rangle| \\ &\leq \|x - x_n\|\|z\| \to 0, \text{ as } n \to +\infty. \end{aligned}$$

thus $\langle x, z\rangle = 0$, for all $z \in L$, i.e., $x \in L^\perp$.

We omit the proof that $H = \overline{L} \oplus L^\perp$. □

Separable Hilbert Spaces

Definition. A metric space (X, d) is called *separable* if it contains a coutable dense subset (i.e. there exists $\{x_n\}_{n=1}^\infty \subset X$ such that, for all $x \in X$ and for all $\epsilon > 0$, there exists n such that $d(x, x_n) < \epsilon$.)

Examples. The following are all separable:

(1) $\mathbb{R}$, $\mathbb{C}$;
(2) $\mathbb{R}^n$, $\mathbb{C}^n$;
(3) $C([0,1], \mathbb{R})$ (use polynomials with rational coefficients – why are these dense?);
(4) l^p, $1 \leq p < \infty$ (exercise).

In particular, l^2 is a separable Hilbert space.

Exercise. l^∞ is *not* separable.

A set $\{e_n\}_{n=1}^\infty$ in a Hilbert space H is called *orthonormal* if

$$\langle e_i, e_j\rangle = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}.$$

(So, in particular, $\|e_i\| = 1$.)

Let H be an infinite dimensional Hilbert space and let $\{f_n\}_{n=1}^{\infty} \subset H$ be a linearly independent subset. (We are going to apply what follows in the case where H is separable but at the moment it doesn't need to be.) Using a procedure called the *Gram-Schmidt algorithm*, it is possible to construct an orthonormal set $\{e_n\}_{n=1}^{\infty} \subset H$ with the same span as $\{f_n\}_{n=1}^{\infty}$. (Here we are going to allow infinite combinations in our spans, so, for example, the span of $\{e_n\}_{n=1}^{\infty}$ will contain any sum $\sum_{n=1}^{\infty} a_n e_n$ provided it converges in H.)

Gram-Schmidt algorithm. We first let $e_1 = f_1/\|f_1\|$. Next we set

$$g_2 = f_2 - \langle f_2, e_1\rangle e_1$$

and let $e_2 = g_2/\|g_2\|$. Then we set

$$g_3 = f_3 - \langle f_3, e_1\rangle e_1 - \langle f_3, e_2\rangle e_2$$

and let $e_3 = g_3/\|g_3\|$. In general, at the nth stage, we set

$$g_n = f_n - \sum_{k=1}^{n-1} \langle f_n, e_k\rangle e_k$$

and let $e_n = g_n/\|g_n\|$. It is an exercise to check that $\{e_n\}_{n=1}^{\infty}$ has the required properties.

Definition. A set $\{e_n\}_{n=1}^{\infty} \subset H$ is a *complete orthonormal* set if

(1) it is orthonormal (see above); and
(2) it is complete, in the sense that if $x \in H$ has $\langle x, e_n\rangle = 0$, for all $n \geq 1$, then $x = 0$. (This is *different* from the notion of completeness in metric spaces.)

We shall show that every separable Hilbert space contains a complete orthonormal set. To show this we need the following inequality.

Proposition 2.13 (Bessel's inequality). *Let $\{e_n\}_{n=1}^{\infty} \subset H$ be an orthonormal set. Let $x \in H$ and write $a_n = \langle x, e_n\rangle$, $n \geq 1$. Then, for all $N \geq 1$,*

$$\sum_{n=1}^{N} |a_n|^2 \leq \|x\|^2.$$

In particular, $\sum_{n=1}^{\infty} |a_n|^2 \leq \|x\|^2$.

Proof. We may write

$$\begin{aligned} 0 &\leq \left\| x - \sum_{n=1}^{N} a_n e_n \right\|^2 \\ &= \langle x, x\rangle - \left\langle x, \sum_{n=1}^{N} a_n e_n \right\rangle - \left\langle \sum_{n=1}^{N} a_n e_n, x \right\rangle + \left\langle \sum_{n=1}^{N} a_n e_n, \sum_{n=1}^{N} a_n e_n \right\rangle \\ &= \|x\|^2 - \sum_{n=1}^{N} a_n \bar{a}_n - \sum_{n=1}^{N} a_n \bar{a}_n + \sum_{n=1}^{N} a_n \bar{a}_n \\ &= \|x\|^2 - \sum_{n=1}^{N} a_n \bar{a}_n. \end{aligned}$$

This gives the required inequality. $\square$

The next result shows that if $\{e_n\}_{n=1}^{\infty}$ is a *complete* orthonormal set then any $x \in H$ may be written in terms of the a_n.

Proposition 2.14. *Let $\{e_n\}_{n=1}^{\infty} \subset H$ be a complete orthonormal set. If $x \in H$ and $a_n = \langle x, e_n \rangle$ then*

$$x = \lim_{N \to +\infty} \sum_{n=1}^{N} a_n e_n = \sum_{n=1}^{\infty} a_n e_n.$$

Proof. Let $w_N = \sum_{n=1}^{N} a_n e_n$. Then, for $N \geq M$,

$$\|w_N - w_M\|^2 = \left\| \sum_{n=M+1}^{N} a_n e_n \right\|^2 = \sum_{n=M+1}^{N} |a_n|^2.$$

By Bessel's inequality, $\sum_{n=1}^{\infty} |a_n|^2 < +\infty$, so

$$\sum_{n=M+1}^{N} |a_n|^2 \to 0, \text{ as } N, M \to +\infty,$$

giving that $\{w_N\}$ is a Cauchy sequence. Hence, w_N converges to some $w \in H$, as $N \to +\infty$. In other words, $w = \sum_{n=1}^{\infty} a_n e_n$.

Now, for all $n \geq 1$, we have

$$\langle w, e_n \rangle = a_n = \langle x, e_n \rangle.$$

Thus $\langle w - x, e_n \rangle = 0$. By the completeness of $\{e_n\}_{n=1}^{\infty}$, this means that $w - x = 0$, i.e., $w = x$. $\square$

Theorem 2.15. *Every infinite dimensional separable Hilbert space contains a complete orthonormal set $\{e_n\}_{n=1}^{\infty}$. Furthermore, each $x \in H$ can be written uniquely in the form*

$$x = \sum_{n=1}^{\infty} a_n e_n, \quad \textit{where } a_n = \langle x, e_n \rangle.$$

Proof. Since H is separable it contains a countable dense subset $\{g_n\}_{n=1}^{\infty}$ and we have $\mathrm{span}(\{g_n\}_{n=1}^{\infty})$ is dense in H. One can show (using Zorn's Lemma) that $\{g_n\}_{n=1}^{\infty}$ has a linearly independent subset $\{f_n\}_{n=1}^{\infty}$ such that

$$\mathrm{span}(\{f_n\}_{n=1}^{\infty}) = \mathrm{span}(\{g_n\}_{n=1}^{\infty}).$$

We can use the Gram-Schmidt algorithm to construct a orthonormal set $\{e_n\}_{n=1}^{\infty}$ such that

$$\mathrm{span}(\{e_n\}_{n=1}^{\infty}) = \mathrm{span}(\{f_n\}_{n=1}^{\infty}),$$

so $\text{span}(\{e_n\}_{n=1}^{\infty})$ is dense in H. Now suppose that $x \in H$ has $\langle x, e_n \rangle = 0$, for all $n \geq 1$. Then $\langle x, v \rangle = 0$ for any $v \in \text{span}(\{e_n\}_{n=1}^{\infty})$. We know that

$$H \to \mathbb{C} : v \mapsto \langle x, v \rangle$$

is continuous: since it is zero on a dense set, it is zero everywhere, i.e,. $\langle x, v \rangle = 0$ for all $v \in H$. Thus $x = 0$ and so $\{e_n\}_{n=1}^{\infty}$ is complete.

By Proposition 2.14, any $x \in H$ can be represented as $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$. If $x = \sum_{n=1}^{\infty} b_n e_n$ then, for all $m \geq 1$,

$$0 = \left\langle \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n - \sum_{n=1}^{\infty} b_n e_n, e_m \right\rangle = \langle x, e_m \rangle - b_m,$$

so the representation is unique. $\square$