MATH602
Real Analysis II

Pingbang Hu

September 28, 2022

Abstract

This is a graduate level functional analysis taught by Joseph Conlon. The prerequisites include linear algebra, complex analysis and also real analysis. We'll use Peter Lax[Lax02] and Reed-Simon[RS80] as textbooks.

This course is taken in Fall 2022, and the date on the covering page is the last updated time.

Contents

Chapter 1

Banach and Hilbert Spaces

Lecture 1: Introduction

We first briefly review different kinds of vector spaces. 30 Aug. 14:30

1.1 Linear Space

Definition 1.1.1 (Linear vector space). A *linear vector space* E over a field $\mathbb{F}$ is a set with operations of addition and multiplication (by a scalar) such that it's closed under operations, and also the addition and scalar multiplication obey

(a) $u + v = v + u$ for $u, v \in E$

(b) $u + (v + w) = (u + v) + w$ for $u, v, w \in E$

(c) $\exists 0 \in E$ such that $0 + u = u + 0 = u$ for $u \in E$

(d) $\forall u \in E$, $\exists - u \in E$ such that $u + (-u) = 0$

(e) $\lambda(u + v) = \lambda u + \lambda v$ for $u, v \in E$, $\lambda \in \mathbb{F}$

(f) $(\lambda + \mu)u = \lambda u + \mu u$ for $u \in E$, $\lambda, \mu \in \mathbb{F}$

(g) $\lambda(\mu u) = (\lambda\mu)u$ for $u \in E$, $\lambda, \mu \in \mathbb{F}$

Remark. If $v, w \in E$, $\lambda, \mu \in \mathbb{R}$ (or $\mathbb{C}$), then $\lambda v + \mu w \in E$.

Notation (Real and complex vector space). If E is over $\mathbb{F} = \mathbb{C}$, we usually call E a *complex vector space*; if $\mathbb{F} = \mathbb{R}$, we say E is a *real vector space*.

Example. $\mathbb{R}^n$ an n dimensional real linear vector space, $\mathbb{C}^n$ an n dimensional complex linear vector space.

We concentrate on ∞ dimensional linear vector space.

Example. Let K is a compact Hausdorff space, then

$$E = \{f : K \to \mathbb{R} \mid f(\cdot) \text{ is continuous}\}.$$

We then see that E is a ∞ dimensional **real** linear vector space.

1.2 Quotient Space

Observe that a linear vector space can have many subspaces. Say E is a linear vector space, and $E_1 \subset E$ where E_1 is a proper subspace, i.e., $E_1 \neq E$.

Definition 1.2.1 (Quotient Space)**.** The *quotient space* E / E_1 of two linear vector spaces E, E_1 such that $E_1 \subseteq E$ is the set of equivalence classes of vectors in E where equivalence is given by $x \sim y$ if $x - y \in E_1$. Additionally, denote $[x]$ as the equivalence class of $x \in E$, i.e., $[x] = x + E_1$.

Remark. A quotient space E / E_1 is a linear vector space

Proof. Since if $x_1 + x_2 \in E$, $[x_1] + [x_2] = [x_1 + x_2]$, and also, $\lambda[x] = [\lambda x]$ for $\lambda \in \mathbb{R}$ or $\mathbb{C}$, i.e., $v, w \in E / E_1$, $\lambda, \mu \in \mathbb{R}$ or $\mathbb{C}$ implies $\lambda v + \mu w \in E$. ⊛

Turns out that the way of defining dimensions for finite dimensional vector spaces doesn't work here: since we may encounter something like $\frac{\infty}{\infty}$. Hence, we introduce Definition 1.2.2.

Definition 1.2.2 (Codimension)**.** If E / E_1 has finite dimension, then the dimension of E / E_1 is called the *codimension* of E_1 in E, denoted as $\operatorname{codim}(E_1)$.

Example. There exists the case that $\dim(E) = \infty$, $\dim(E_1) < \infty$ where $\dim(E / E_1) < \infty$.

Proof. Let $E = \{f \colon K \to \mathbb{R} \mid f(\cdot) \text{ continuous}\}$, and $E_1 = \{f \in E \colon f(k_1) = 0\}$ where $k_1 \in K$ is fixed. We see that the dimension of E / E_1 is exactly 1 since E / E_1 is the set of constant functions. ⊛

Theorem 1.2.1. If E is finite dimensional, then $\operatorname{codim}(E_1) + \dim(E_1) = \dim(E)$

Definition 1.2.3 (Linear operator)**.** A map $T \colon E \to F$ between 2 linear spaces is a *linear operator* if it preserves the properties of addition and multiplication by a scalar, i.e., for $v, w \in E$ and $\lambda, \mu \in \mathbb{R}$ or $\mathbb{C}$,

$$T(\lambda v + \mu w) = \lambda T(v) + \mu T(w).$$

Definition. Given a linear operator $T \colon E \to F$ we have the following.

Definition 1.2.4 (Kernel)**.** The *kernel* of T is the subspace $\ker(T) = \{x \in E \mid Tx = 0\}$.

Definition 1.2.5 (Image)**.** The *image* of T is the subspace $\operatorname{Im}(T) = \{Tx \in F \mid x \in E\}$.

1.3 Normed Spaces

We review some basic notions.

Definition 1.3.1 (Norm)**.** Let E be a linear vector space. A *norm* $\|\cdot\| : E \to \mathbb{R}$ on E is a function from E to $\mathbb{R}$ with the properties:

(a) $\|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$.

(b) $\|\lambda x\| = |\lambda| \, \|x\|$, $\lambda \in \mathbb{R}$ or $\mathbb{C}$.

(c) $\|x + y\| \leq \|x\| + \|y\|$.

Notation (Dilation). We say that the second condition is the *dilation* property.

Definition 1.3.2 (Normed vector space). A linear vector space E equipped with a norm $\|\cdot\|$ is called a *normed vector space*.

Remark (Induced metric space). A normed vector space E induces a *metric space* with metric $d(x,y) = \|x-y\|$, where the metric has properties

(a) $d(x,y) \geq 0$. Also, $d(x,x) = 0$ and $d(x,y)$ implies $x = y$.

(b) $d(x,y) = d(y,x)$.

(c) $d(x,z) \leq d(x,y) + d(y,z)$.

Example (Bounded sequences ℓ_∞). Let ℓ_∞ be the space of bounded sequences $x = (x_1, x_2, \ldots)$ with $x_i \in \mathbb{R}$ for $i = 1, 2, \ldots$. Then we define $\|x\| = \|x\|_\infty = \sup_{i\geq 1} |x_i|$.

Example (Absolutely summable sequences ℓ_1). Let ℓ_1 be the space of absolutely summable sequences $x = (x_1, x_2, \ldots)$ and $\sum_{i=1}^\infty |x_i| < \infty$. Then we define $\|x\| = \|x\|_1 = \sum_{i=1}^\infty |x_i| < \infty$.

Example (Continuous functions $C(k)$). The space $C(k)$ of continuous functions $f\colon K \to \mathbb{R}$ where K is compact Hausdorff. Then we define $\|f\| = \|f\|_\infty = \sup_{x\in K} |f(x)|$.

1.3.1 Geometry of Normed Spaces

Definition 1.3.3 (Ball). A (closed) *ball* centered at a point $x_0 \in E$ with radius $r > 0$ is the set $B(x_0, r) = \{x \in E \mid \|x - x_0\| \leq r\}$.

Definition 1.3.4 (Sphere). The *sphere* centered at x_0 with radius $r > 0$ is the set $S(x_0, r) = \{x \in E \mid \|x - x_0\| = r\}$.

Remark. We see that $S(x_0, r)$ is the **boundary** of $B(x_0, r)$, i.e., $S(x_0, r) = \partial B(x_0, r)$.

Note (Infinite dimensional geometry). We know that in finite dimensional, all norms are equivalent, which is not true for infinite dimensional vector spaces. This has something to do with the geometry of balls.

Explicitly, balls can have different geometries depending on the properties of the norms. We see that a $\|\cdot\|_\infty$ can have multiple supporting hyperplane at the corner, while for a $\|\cdot\|_2$ can have only one at each point.

Also, unit balls for $\|\cdot\|_1$ is also a **square**, where we have

$$B(0,1) = \{x = (x_1, x_2, \ldots) \mid -1 < y_\epsilon < 1 \forall \epsilon\}$$

such that $y_\epsilon = \sum_{i=1}^\infty \epsilon_i x_i$, $\epsilon_i = \pm 1$ and $\epsilon = (\epsilon_1, \epsilon_2, \ldots)$.

We see that different norms give different geometry, but they have important common features, most notably, convexity properties.

Definition 1.3.5 (Convex set). Given E a linear vector space, a set $K \subset E$ is *convex* if for $x, y \in K$ and $0 \leq \lambda \leq 1$,

$$\lambda x + (1-\lambda) y \in K.$$

Definition 1.3.6 (Convex function). Given E a linear vector space, a function $f\colon E \to \mathbb{R}$ is called *convex* if for $x, y \in E$ and $0 \le \lambda \le 1$,

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y).$$

Remark. If $f\colon E \to \mathbb{R}$ is a convex function, then for any $M \in \mathbb{R}$ the set $\{x \in E \mid f(x) \le M\}$ is convex.

The upshot is that norms are convex, and the unit balls are convex as well.

Lecture 2: Banach Spaces and Completion

Let's first see a proposition. 01 Sep. 14:30

Proposition 1.3.1. Let $(E, \|\cdot\|)$ be a normed linear space, then the norm is convex and continuous.

Proof. Let $f\colon E \to \mathbb{R}$ be $f(x) = \|x\|$. Then $f(x) - f(y) = \|x\| - \|y\| \le \|x - y\|$, which implies $|f(x) - f(y)| \le \|x - y\|$ for $x, y \in E$, i.e., f is Lipschitz continuous hence continuous. For convexity, let $0 < \lambda < 1$, we have

$$f(\lambda x + (1-\lambda)y) = \|\lambda x + (1-\lambda)y\| \le \|\lambda x\| + \|(1-\lambda)y\| = \lambda \|x\| + (1-\lambda)\|y\| = \lambda f(x) + (1-\lambda) f(y).$$

■

Note. Note that $f(\cdot)$ is continuous implies the closed ball

$$B(x_0, r) = \{x \in E \mid \|x - x_0\| \le r\} = \{x \in E \mid f(x - x_0) \le r\}$$

is closed in topology of E. Also, $f(\cdot)$ is convex implies $B(x_0, r)$ is convex.

Remark. If $f\colon E \to \mathbb{R}$ is convex, then the sets $\{x \in E \mid f(x) \le M\}$ is also convex. However, it's possible to have non-convex functions f such that all sets $\{x \in E \mid f(x) \le M\}$ are convex.

Proof. Take $f(x) = |x|^p$ for $x \in \mathbb{R}$ and $p > 0$. We see that f is convex if $p > 1$, and non-convex if $p < 1$. However, the sets $\{x \in \mathbb{R} \mid f(x) \le M\}$ all convex since it's independent of p. ⊛

Lemma 1.3.1. Suppose $x \mapsto \|x\|$ satisfies

(a) $\|x\| \ge 0$ and $\|x\| = 0 \Leftrightarrow x = 0$.

(b) $\|\lambda x\| = |\lambda| \|x\|$, $\lambda \in \mathbb{R}$ or $\mathbb{C}$.

(c) The unit ball $B(0,1)$ is convex.

Then $f(x) = \|x\|$ satisfies the triangle inequality $\|x + y\| \le \|x\| + \|y\|$.

Proof. We see that if the third condition is true, the for $u, v \in B(0,1)$ and $0 < \lambda < 1$, we have $\lambda u + (1-\lambda)v \in B(0,1)$. Let $x, y \in E$, and

$$\lambda = \frac{\|x\|}{\|x\| + \|y\|} \Rightarrow 1 - \lambda = \frac{\|y\|}{\|x\| + \|y\|}.$$

By letting $u = x/\|x\|$, $v = y/\|y\|$ we see that

$$\lambda u + (1-\lambda)v = \frac{\|x\|}{\|x\| + \|y\|}\frac{x}{\|x\|} + \frac{\|y\|}{\|x\| + \|y\|}\frac{y}{\|y\|} \in B(0,1) \Rightarrow \left\| \frac{x}{\|x\| + \|y\|} + \frac{y}{\|x\| + \|y\|} \right\| \le 1.$$

From the second condition, it follows that $\|x + y\| \le \|x\| + \|y\|$, which is the triangle inequality. ■

Remark. If $x \mapsto \|x\|$ satisfies the first two condition and is convex, then it satisfies the triangle inequality.

Proof. Since $\frac{1}{2}\|x+y\| = \left\|\frac{x}{2}+\frac{y}{2}\right\| \leq \frac{1}{2}\|x\| + \frac{1}{2}\|y\|$. ⊛

Now, given a quotient space $E \,/\, E_1$, the question is can we try to define a norm?

Problem 1.3.1. On $E \,/\, E_1$, is $\|[x]\| \coloneqq \inf_{y\in E_1}\|x+y\|$ a norm?

Answer. We see that if $x \in \overline{E}_1 \setminus E_1$, then $\|[x]\| = 0$ but $0 \neq [x] \in E \,/\, E_1$. ⊛

Note. Notice the difference from finite dimensional situation. All finite dimensional spaces E_1 are closed but not in general if E_1 has ∞ dimensions.

Example. Let $\ell_1(\mathbb{R})$ be the sequence of x_n for $n \geq 1$ in $\mathbb{R}$ such that $\sum_{i=1}^{\infty}|x_i| \leq \infty$. Define

$$\|x\|_1 \coloneqq \sum_{i=1}^{\infty}|x_i|\,,$$

and let E_1 be all sequences with finite number of the x_n are nonzero. We see that $\overline{E}_1 = \ell_1(\mathbb{R})$ is infinite dimensional.

Proposition 1.3.2. Let $(E, \|\cdot\|)$ be a normed space and $E_1 \subseteq E$, E_1 is closed. Then

$$\|\cdot\| : {}^{E}\!\big/_{E_1} \to \mathbb{R}, \quad \|[x]\| = \inf_{y\in E_1}\|x+y\|$$

is a norm on $E \,/\, E_1$.

Proof. If $\|[x]\| = 0$, then $\inf_{y\in E_1}\|x-y\| = 0$, which implies $x \in E_1$ since E_1 is closed, so $[x] = 0$. Also, since

$$\|\lambda[x]\| = \inf_{y\in E_1}\|\lambda x + y\| = \inf_{z\in E_1}\|\lambda x + \lambda z\| = |\lambda|\inf_{z\in E_1}\|x+z\| = |\lambda|\,\|[x]\|\,,$$

the dilation property is satisfied. Finally, for triangle inequality, we have

$$\|[x]+[y]\| = \inf_{x_1,y_1\in E}\|x+y+x_1+y_1\| \leq \inf_{x_1\in E_1}\|x+x_1\| + \inf_{y_1\in E_1}\|y+y_1\| = \|[x]\| + \|[y]\|\,.$$

■

Remark. This shows that the only obstacle for this kind of norm being an actual norm is the closeness of E_1.

1.4 Banach Spaces

Definition 1.4.1 (Banach space)**.** A linear normed space is a *Banach space* if it's complete, i.e., every Cauchy sequence converges.

This implies that given a Banach space $(E, \|\cdot\|)$, if $\{x_n \in E : n \geq 1\}$ is a sequence with property such that $\lim\limits_{m\to\infty}\sup\limits_{n\geq m}\|x_n - x_m\| = 0$, then $\exists x_\infty \in E$ such that $\lim\limits_{n\to\infty}\|x_n - x_m\| = 0$ as well.

Example. The spaces ℓ_1, ℓ_∞ and $C(K)$ are Banach spaces.

We want to give a different criterion for showing $(E, \|\cdot\|)$ is Banach. Let E be a linear normed space and $\{x_\ell \mid \ell \geq 1\}$ a sequence in E.

1.4.1 Completion of Normed Space

We now show an important theorem which characterizes completeness in terms of convergence of series rather than sequences. We first see the definition.

Definition 1.4.2 (Absolutely summable). A sequence is *absolutely summable* if $\sum_{i=1}^{\infty} \|x_i\| < \infty$.

Then, we have the following.

Theorem 1.4.1 (Criterion for completeness). A normed space $(E, \|\cdot\|)$ is a Banach space if and only if every absolutely summable series in E converges.

Proof. We need to prove two directions.

($\Rightarrow$) Suppose E is a Banach space and $\{x_k \mid x \geq 1\}$ an absolutely summable series. Set $s_n = \sum_{k=1}^{n} x_k$, $n \geq 1$, we want to show s_n is Cauchy, and if this is the case, completeness of E implies $\exists s_\infty$ and $\lim\limits_{n\to\infty} \|s_n - s_\infty\| = 0$. Let $n > m$, we see that

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^{n} x_k \right\| \leq \sum_{k=m+1}^{n} \|x_k\| \leq \sum_{k=m+1}^{\infty} \|x_k\| .$$

Observe that $\lim\limits_{m\to\infty} \sum\limits_{k=m+1}^{\infty} \|x_k\| = 0$, we see that the sequence $\{s_n\}$ is Cauchy.

($\Leftarrow$) Conversely, suppose E is **not** complete. Then there exists a Cauchy sequence $\{x_n \mid n \geq 1\}$ which does not converge. Furthermore, no subsequence of $\{x_n \mid n \geq 1\}$ converges.[a] We now construct an absolutely summable series which does not converge.

Define $n(1) \geq 1$ such that $\left\|x_n - x_{n(1)}\right\| \leq \frac{1}{2}$ if $n \geq n(1)$, similarly, let $n(2) > n(1)$ be such that $\left\|x_n - x_{n(2)}\right\| \leq \frac{1}{2^2}$ if $n > n(2)$. In all, we have $n(1) < n(2) < n(3) < \ldots$ such that $\left\|x_n - x_{n(k)}\right\| \leq \frac{1}{2^k}$ if $n > n(k)$. Define $w_j \coloneqq x_{n(j+1)} - x_{n(j)}$ for $j = 1, 2, \ldots$. We see that

$$x_{n(m)} = x_{n(1)} + \sum_{j=1}^{m} w_j$$

for $m = 1, 2, \ldots$, and $\{x_{n(m)}\}$ does not converge, hence so does the series $\sum_{j=1}^{\infty} w_j$. However, $\sum_{j=1}^{\infty} \|w_j\| \leq \sum_{j=1}^{\infty} \frac{1}{2^j} = 1$, which implies $\{w_j\}$ is absolutely summable. ∎

[a] Otherwise, the whole sequence converges by the fact that it's Cauchy.

Theorem 1.4.2 (Completion). Suppose E is a normed space. Then there exists a Banach space $\hat{E}$ called a*the completion* of E with the following properties:

(a) There exists a linear map $\iota \colon E \to \hat{E}$ such that $\|\iota x\| = \|x\|$.[a]

(b) $\operatorname{Im}(\iota)$ is dense in $\hat{E}$, and $\hat{E}$ is the smallest Banach space containing image of E.

[a] This is called an *isometric embedding* of E into $\hat{E}$.

Lecture 3: Banach, Inner Product Spaces

Notice that ℓ_1 and ℓ_∞ are Banach, and we want to generalize to ℓ_p with $1 < p < \infty$. For $x = \{x_n, n \geq 1\} \in \ell_p$ and if $\sum_{n=1}^{\infty} |x_n|^p < \infty$, for $\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p}$, we want to show that $x \to \|x\|_p$ 06 Sep. 14:30

satisfies properties of a norm. The first two properties of a norm is easy check. As for triangle inequality, we have the following.

Lemma 1.4.1 (Minkowski inequality)**.** Let $1 \leq p < \infty$, for $x, y \in \ell_p$,

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p .$$

Proof. Recall that from Lemma 1.3.1, we only need to show that $B(0,1)$ is convex, where

$$B(0,1) = \left\{ x = \{x_n : n \geq 1\} \mid f(x) = \sum_{n=1}^{\infty} |x_n|^p \leq 1 \right\}.$$

But $f(x)$ is convex since $x \mapsto |x|^p$, $x \in \mathbb{R}$ is convex if $p \geq 1$, we're done. Hence, $\|x+y\|_p \leq \|x\|_p + \|y\|_p$, i.e.,

$$\left(\sum_{j=1}^{\infty} |x_j + y_j|^p \right)^{1/p} \leq \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{1/p} \left(\sum_{j=1}^{\infty} |y_j|^p \right)^{1/p} .$$

■

Lemma 1.4.2 (Hölder's inequality)**.** Let $1 < p < \infty$, for $x \in \ell_p$, $y \in \ell_q$, we have

$$\|x \cdot y\|_1 \leq \|x\|_p \|y\|_q$$

where $1/p + 1/q = 1$.

Proof. Note first that we can assume without loss of generality, $\sum_{j=1}^{\infty} |x_j|^p = \sum_{j=1}^{\infty} |y_j|^q = 1$. Then, result follows from the Young's inequality,

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

for $x, y > 0$, $x, y \in \mathbb{R}$.

> **Remark** (Legendre transform and the inequality)**.** Young's inequality is a special case of the inequality
>
> $$xy \leq f(x) + \mathcal{L}f(y)$$
>
> where $\mathcal{L}f(\cdot)$ is the *Legendre transform* of $f(\cdot)$, i.e., $\mathcal{L}f(y) = \sup_x [xy - f(x)]$.
>
> If f is convex, then the function $xy \mapsto xy - f(x)$ is concave so has unique maximum. And $\mathcal{L}f(\cdot)$ always convex even if $f(\cdot)$ is not. In particular, if $f(x) = x^p/p$, then $\mathcal{L}f(y) = y^q/q$.

■

Note. Minkowski inequality is usually proved via the Hölder's inequality. To show this, since

$$\sum_{j=1}^{\infty} |x_j + y_j|^p \leq \sum_{j=1}^{\infty} |x_j| |x_j + y_j|^{p-1} + \sum_{j=1}^{\infty} |y_j| |x_j + y_j|^{p-1} .$$

Then Holder inequality implies

$$\sum_{j=1}^{\infty} |x_j| |x_j y_j|^{p-1} \leq \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{1/p} \left(\sum_{j=1}^{\infty} |x_j + y_j|^{(p-1)q} \right)^{1/q} ,$$

and we're done.[a]

[a]Note that $(p-1)q = p$.

The above argument applies to more general spaces of p integrable functions. Let (Ω, Σ, μ) be a measure space and $L_p(\Omega, \Sigma, \mu)$ where all Σ measure functions $f\colon \Omega \to \mathbb{R}$ (or $\mathbb{C}$) such that $\int_\Omega |f|^p \,\mathrm{d}\mu < \infty$. Then, $L_p(\Omega, \Sigma, \mu)$ is a normed space with norm

$$\|f\|_p = \left(\int_\Omega |f|^p \,\mathrm{d}\mu\right)^{1/p}.$$

It's more tricky to show that L^p is a Banach space, but it's indeed still the case.

Theorem 1.4.3 (Riesz-Fisher)**.** The space $L_p(\Omega, \Sigma, \mu)$ is a Banach space for $1 \leq p < \infty$.

Proof. Toward using Theorem 1.4.1, let $\{f_n\colon n \geq 1\}$ be an absolutely summable sequence in L_p. Then the norm satisfies

$$\left\|\sum_{k=1}^N f_k\right\|_p \overset{!}{\leq} \sum_{k=1}^N \|f_k\|_p \leq C < \infty \Rightarrow \int_\Omega \left|\sum_{k=1}^N f_k\right|^p \,\mathrm{d}\mu \leq C^p.$$

- Assume all f_k are non-negative. From monotone convergence theorem, we have

$$\lim_{N\to\infty} \int_\Omega \left(\sum_{k=1}^N f_k\right)^p \,\mathrm{d}\mu = \int_\Omega \left(\sum_{k=1}^\infty f_k\right)^p \,\mathrm{d}\mu \leq C^p.$$

  Hence, $g = \sum_{k=1}^\infty f_k \in L_p$. We now want to show that $\sum_{k=1}^N f_k \to g$ in L_p. Set $r_n = \sum_{k=n+1}^\infty f_k$ where r_n is a decreasing sequence where $r_n \to 0$ a.e. and also

$$\int_\Omega r_1^p \,\mathrm{d}\mu < \infty.$$

  This means that $\lim\limits_{n\to\infty} \|r_n\|_p = 0$ by dominate convergence theorem.

- For arbitrary $f_k\colon \Omega \to \mathbb{R}$, write $f_k = f_k^+ + f_k^-$ where $f_k^+ = \sup(f_k, 0)$ and $f_k^- = \inf(f_k, 0)$. The sequence $\left\{f_k^+\colon k \geq 1\right\}$ are absolutely summable, and we just proceed as before. Similarly, if $f_k\colon \Omega \to \mathbb{C}$, we have the similar result.

■

1.5 Inner Product Spaces

Definition 1.5.1 (Inner product)**.** Let E be a linear space over $\mathbb{C}$. An *inner product* $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{C}$ is a function which has the following properties:

(a) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.

(b) $\langle ax + by, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle$ for $a, b \in \mathbb{C}$.

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

Remark (Real inner product)**.** We can also define inner products of spaces over $\mathbb{R}$ with no extra conjugation in the last property.

Definition 1.5.2 (Inner product space)**.** An *inner product space* is a linear space E with an inner product $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{C}$.

Definition 1.5.3 (Orthogonal). Given a linear space E, $x, y \in E$ are *orthogonal* if $\langle x, y\rangle = 0$, denote as $x \perp y$.

Theorem 1.5.1 (Cauchy-Schwarz inequality). Let $x, y \in E$ and an inner product $\langle \cdot, \cdot\rangle$, then

$$|\langle x, y\rangle| \leq \langle x, x\rangle^{\frac{1}{2}} \langle y, y\rangle^{\frac{1}{2}}.$$

Proof. Define $Q(t)$ by $Q(t) = \langle x+ty, x+ty\rangle = \langle y, y\rangle t^2 + 2t \operatorname{Re}\langle x, y\rangle + \langle x, x\rangle$ if $t \in \mathbb{R}$. Then we see that $Q(t) \geq 0$ with $t \in \mathbb{R}$, by looking at the discriminant, we have $(\operatorname{Re}\langle x, y\rangle)^2 \leq \langle x, x\rangle \langle y, y\rangle$. Finally, the result follows by choosing $\theta \in \mathbb{R}$ such that $\langle x, y\rangle = \operatorname{Re}\left\langle xe^{i\theta}, y\right\rangle$, we then see that

$$|\langle x, y\rangle| = \left|\operatorname{Re}\left\langle xe^{i\theta}, y\right\rangle\right| = |\operatorname{Re}\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle \langle y, y\rangle},$$

proving the result. ■

Corollary 1.5.1. The function $x \mapsto \|x\| := \langle x, x\rangle^{\frac{1}{2}}$ is a norm on E.

Proof. The triangle inequality is a consequence of Theorem 1.5.1 such that

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + 2\operatorname{Re}\langle x, y\rangle + \|y\|^2 \overset{!}{\leq} \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

■

Remark (Pythagorean theorem). The calculation in Corollary 1.5.1 clearly implies *Pythagorean theorem*: if $x \perp y$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Example. The space ℓ_2 of square summable sequences $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$,

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

defines an inner product.

Example (Canonical inner product on L_2). The space $L_2(\Omega, \Sigma, \mu)$ of square integrable functions f, g,

$$\langle f, g\rangle = \int_\Omega f(x)\overline{g}(x) \,\mathrm{d}\mu(x)$$

defines an inner product. Furthermore, $\|f\|_2 = \langle f, f\rangle^{1/2}$.

Proof. The only non-trivial fact to prove is that $\langle f, g\rangle$ is finite, i.e., $f\overline{g}$ is integrable. Firstly, f^2, $\overline{f}^2$ and $(f+g)^2$ are all integrable since f, $\overline{g}$ and $f+\overline{g}$ are all in L_2, hence $f\overline{g}$ is also integrable. ⊛

Example. The space of $m \times n$ matrices $A = (a_{ij})$, $1 \leq i \leq m, 1 \leq j \leq n$. Then

$$\langle A, B\rangle = \operatorname{tr}(AB^*)$$

defines an inner product, where B^* is the Hermitian adjoint of B, i.e., for $B = (b_{ij})$, then $B^* = (b^*_{ij})$ for $b^*_{ij} = \overline{b}_{ji}$.

Remark (Hilbert-Schmidt (Frobenius) norm). Specifically, the norm corresponding to this inner product is

$$\|A\|_{\mathrm{HS}} := \left(\sum_{i,j}^{\infty} |a_{ij}|^2 \right)^{1/2},$$

which is known as the *Hilbert-Schmidt* or *Frobenius* norm.

Note (Angle). Recall that in Euclidean space $\mathbb{R}^n$, the inner product can be computed by the formula

$$\langle x, y \rangle = \|x\| \, \|y\| \cos \theta(x, y)$$

where $\theta(x, y)$ denotes the angle between x and y. We see that we can similarly define the angle between x, y in an inner product space by

$$\cos \theta(x, y) := \frac{\langle x, y \rangle}{\|x\| \, \|y\|} \in [-1, 1]$$

where the range is ensured by Theorem 1.5.1, so it's well-defined. Though this concept this rarely used anyway.

Other geometry result also follows in an inner product space: since inner product can be expressed in terms of the norm. This is because both parallelogram law and polarization identity hold.

Lemma 1.5.1 (Parallelogram law). Given E an inner product space, we have

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

Proof. Recall that $\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2$ and similarly, $\|x - y\|^2 = \|x\|^2 - 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2$, hence the result follows. ∎

Lemma 1.5.2 (Polarization identity). Given E an inner product space, we have

$$\langle x, y \rangle = \frac{1}{4} \left\{ \|x + y\|^2 - \|x - y\|^2 + i \|x + iy\|^2 - i \|x - iy\|^2 \right\}$$

Proof. The proof is just to expand the right-hand side in terms of inner product. ∎

Remark. Polarization identity shows that the function $x \mapsto \|x\|^2$ determines the inner product.

Lecture 4: Orthogonality and Projection

1.6 Hilbert Spaces

08 Sep. 14:30

Definition 1.6.1 (Hilbert space). A complete inner product space is called a *Hilbert space*.

Example. We have seen that ℓ_2 and $L^2(\Omega, \Sigma, \mu)$ are complete, hence are Hilbert space.

1.6.1 Orthogonality

We'll soon see that the key notion in Hilbert space theory is orthogonality.

Definition 1.6.2 (Orthogonal complement). Let $A \subseteq \mathcal{H}$ where $\mathcal{H}$ is a Hilbert space. Then the *orthogonal complement* $A^{\perp}$ of A is

$$A^{\perp} := \{x \in \mathcal{H} \colon \langle x, y \rangle = 0 \text{ for } y \in A\}.$$

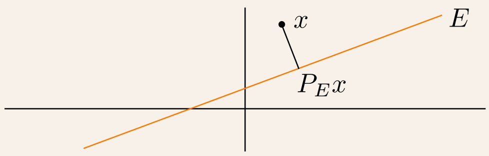
Remark. $A^{\perp}$ is also a Hilbert space, in particular, closed and $A^{\perp} \cap A \subseteq \{0\}$.

Proof. Since $A^{\perp}$ is closed linear subspace of $\mathcal{H}$, where the closure follows from the continuity of the function $x \mapsto \langle x, y \rangle$ for $x \in \mathcal{H}$ by looking at the inverse image of $\{0\}$. Finally, for $x \in A^{\perp} \cap A$, $\langle x, x \rangle = 0$ implies $x = 0$. ⊛

Theorem 1.6.1 (Orthogonality principle). Assume $E \subseteq \mathcal{H}$ is a closed linear subspace of the Hilbert space $\mathcal{H}$ and $x \in \mathcal{H}$. Then we have the following.

(a) Then there exists a unique closest point $y = P_E x \in E$ to x, i.e., $\|x - P_E x\| = \inf_{y' \in E} \|x - y'\|$.

(b) The point $y = P_E x \in E$ is the unique vector such that $x - y \in E^{\perp}$.



Proof. Note that the function $y' \mapsto \|x - y'\|$ for $y' \in E$ is convex. We expect a minimizer y'.

(a) Let $y_n \in E$ for $n = 1, 2, \ldots$ be a minimizing sequence, i.e.,

$$\lim_{n \to \infty} \|x - y_n\| = \inf_{y' \in E} \|x - y'\| =: d.$$

From parallelogram law, we have

$$\|y_n - y_m\|^2 + 4\|x - (y_n + y_m)/2\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2.$$

As $n, m \to \infty$, the right-hand side goes to $4d^2$. But since $\frac{1}{2}(y_n + y_m) \in E$, we have $\left\|x - \frac{1}{2}(y_n - y_m)\right\| \geq d$, so

$$\lim_{m \to \infty} \sup_{m \geq n} \|y_n - y_m\|^2 = 0,$$

which further implies $\{y_n\}$ is a Cauchy sequence. As $\mathcal{H}$ is complete, we see that $y_n \to y_\infty \in E$, with $\|x - y_\infty\| = d$.

Now, with the fact that E is closed, we set $y_\infty = P_E x$ where y_∞ is unique since if $\|x - y_\infty\| = \|x - y'_\infty\| = d$, again by the parallelogram law where we now plug in y_∞ and y'_∞ instead of y_n and y_m as above, we see that $\|y_\infty - y'_\infty\| = 0$. In all, $P_E x \in E$ is uniquely defined.

(b) We now show $P_E x$ is the unique vector $y \in E$ such that $x - y \perp E$, i.e., $x - y \in E^{\perp}$. Let $y' \in E$ and let $Q(t)$ be the quadratic

$$Q(t) := \langle x - P_E x + t y', x - P_E x + t y' \rangle = \|x - P_E x + t y'\|^2.$$

Since $t \mapsto Q(t)$ has a **strict** minimum at $t = 0$, which implies $Q'(0) = 0$, i.e., $\operatorname{Re}(x - P_E x, y') = 0$ for all $y' \in E$, which further implies $\langle x - P_E x, y' \rangle = 0$ for all $y' \in E$. This shows that $x - P_E x \in E^{\perp}$.

Finally, we need to show $P_E x \in E$ is the unique vector such $x - P_E x \in E^{\perp}$. This can be seen from $Q(t) = \|x - P_E x\|^2 + t^2 \|y'\|^2$ for any $y' \in E$.

■

Note. To show this exists, we typically need

1. Compactness properties
2. Non-degeneracy properties for uniqueness

Here by using parallelogram law, we don't need compactness.

Remark. Theorem 1.6.1 shows that the minimizer for the function $y' \mapsto \|x - y'\|$ for $y' \in E$ is characterized by the orthogonality condition, i.e., $x - y \perp E$ for some $y \in E$.

Definition 1.6.3 (Orthogonal projection). Let $\mathcal{H}$ be a Hilbert space ad let $E \subseteq \mathcal{H}$ be a closed subspace. The *orthogonal projection operator* $P_E \colon \mathcal{H} \to E$ is given by $x \mapsto P_E x$ where $P_E x$ is defined uniquely via $x - P_E x \in E^\perp$.

Definition 1.6.4 (Bounded linear map). Given a mapping $A \colon \mathcal{B} \to \mathcal{B}$ on a Banach space $\mathcal{B}$, we say it's a *bounded linear map* if it's bounded and linear.

Definition 1.6.5 (Linear map). The operator A is *linear* if for $x, y \in \mathcal{B}$, $a, b \in \mathbb{C}$,

$$A(ax + by) = aA(x) + bB(y).$$

Definition 1.6.6 (Bounded map). The operator A is *bounded* if

$$\|A\| \coloneqq \sup_{\|x\|=1} \|Ax\| < \infty.$$

Remark. Note that $\|Ax\| \leq \|A\| \, \|x\|$ for $x \in \mathcal{B}$.

We see that $P_E x$ is a bounded linear operator $P_E \colon \mathcal{H} \to E$ with the properties $P_E^2 g P_E$ and $\|x\|^2 = \|P_E x\|^2 + \|(I - P_E)x\|^2$ since $(I - P_E)x \perp P_E x$. The latter property shows that

$$\|P_E\| \leq 1, \quad \|(I - P_E)\| \leq 1,$$

and fact, $\|P_E\| = \|I - P_E\| = 1$. Also, $I - P_E$ is also an orthogonal projection onto $E^\perp$.

1.7 Fourier Series

Hilbert space gives a geometric framework for studying Fourier series. The classical Fourier analysis studies situations where a function $f \colon [-\pi, \pi] \to \mathbb{C}$ can be expanded as Fourier series

$$f(t) = \sum_{k=-\infty}^{\infty} \hat{f}(k) \frac{1}{\sqrt{2\pi}} e^{ikt}$$

with the Fourier coefficients

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(t) e^{-ikt} \,\mathrm{d}t.$$

In order to make Fourier analysis rigorous, we have to understand what functions f can be written as Fourier series, and in what sense the Fourier series converges. To do so, it's of great advantage to depart from this specific situation and carry out Fourier analysis in an abstract Hilbert space. Let $f(t)$ be a vector in the function space $L_2[-\pi, \pi]$, and the exponential functions e^{-ikt} will form a set of orthogonal vectors in this space. Then, Fourier series will become an orthogonal decomposition of a vector f w.r.t. an orthogonal system of coordinates.

1.7.1 Orthogonal Systems

We first give the definition.

Definition 1.7.1 (Orthogonal system)**.** A sequence $\{x_k : k \geq 1\}$ of non-zero vectors in a Hilbert space $\mathcal{H}$ is *orthogonal* if $\langle x_k, x_\ell \rangle = 0$ for all $\ell \neq k$.

Definition 1.7.2 (Orthonormal system)**.** An orthogonal system is called an *orthonormal system* if in addition, we have $\|x_k\| = 1$ for all k.

Remark (Equivalence definition of orthonormal system)**.** $\{x_k : k \geq 1\}$ is orthonormal if $\langle x_k, x_\ell \rangle = \delta_{k,\ell}$ where δ is the Kronecker delta.

We first see an immediate generation given the remark.

Theorem 1.7.1 (Pythagorean theorem)**.** Let $\{x_k\}_k$ be an orthogonal system in a Hilbert space $\mathcal{H}$. Then for every $n \in \mathbb{N}$,

$$\left\| \sum_{k=1}^{n} x_k \right\|^2 = \sum_{k=1}^{n} \|x_k\|^2$$

Proof. From orthogonality,

$$\left\langle \sum_{k=1}^{n} x_k, \sum_{k=1}^{n} x_k \right\rangle = \sum_{k,j=1}^{n} \langle x_k, x_j \rangle = \sum_{k=1}^{n} \langle x_k, x_k \rangle ,$$

proving the result ■

We now see some examples.

Example (Canonical basis of ℓ_2)**.** In the space ℓ_2, $x_k = (0, 0, \ldots, 1, 0, \ldots, 0) \in \ell_2$ for $k = 1, 2, \ldots$ is orthonormal system in ℓ_2.

Example (Fourier basis in L_2)**.** In the space $L_2[-\pi, \pi]$, consider the exponentials

$$e_k(t) = \frac{1}{\sqrt{2\pi}} e^{ikt}$$

for $t \in [-\pi, \pi]$. The set $\{e_k\}_{k=-\infty}^{\infty}$ is an orthonormal-system in $L_2[-\pi, \pi]$.

1.7.2 Fourier Series

We can further generalize Fourier series to any Hilbert space by letting $\{x_k : k \geq 1\}$ be an orthonormal set in $\mathcal{H}$ as follows.

Definition. Consider an orthonormal-system $\{x_k\}_{k=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ and a vector $x \in \mathcal{H}$.

Definition 1.7.3 (Fourier series)**.** The *Fourier series* of x w.r.t. $\{x_k\}$ is the formal series

$$\sum_{k=1}^{\infty} \langle x, x_k \rangle x_k .$$

Definition 1.7.4 (Fourier coefficient)**.** The coefficient $\langle x, x_k \rangle$ in the Fourier series are called *Fourier coefficients* of x.

To understand the convergence of Fourier series, we first focus on the finite case and study the partial sums of Fourier series. For $n = 1, 2, \ldots$, we define $S_n : \mathcal{H} \to E_n$ such that

$$S_n(x) = \sum_{k=1}^{n} \langle x, x_k \rangle x_k$$

for $x \in \mathcal{H}$ where $E_n = \operatorname{span}\{x_1, \ldots, x_n\}$. We see that S_n is a linear operator and $S_n = P_{E_n}$ is the orthogonal projection onto E_n since $\langle x - S_n(x), x_k \rangle = 0$ for $k = 1, \ldots, n$ and $S_n(x) \in E_n$, $x - S_n(x) \perp E_n$.

Theorem 1.7.2 (Bessel's inequality). Let $\{x_k\}_k$ be an orthogonal system in a Hilbert space $\mathcal{H}$. then for every $x \in \mathcal{H}$,

$$\sum_k |\langle x, x_k \rangle|^2 \leq \|x\|^2 .$$

Proof. To estimate the size of $S_n(x)$, since $x - S_n(x)$, by Theorem 1.7.1,

$$\|S_n(x)\|^2 + \|x - S_n(x)\|^2 = \|x\|^2 \Rightarrow \|S_n(x)\|^2 \leq \|x\|^2 .$$

On the other hand, again by Theorem 1.7.1 and orthogonality,

$$\|S_n(x)\|^2 = \sum_{k=1}^{n} \|\langle x, x_k \rangle x_k\|^2 = \sum_{k=1}^{n} |\langle x, x_k \rangle|^2 .$$

We see that by combining these two inequalities and let $n \to \infty$, we have the result. ■

Remark. In particular, we see that $\|S_n(x)\|^2 = \sum_{k=1}^{n} |\langle x, x_k \rangle|^2$, with $S_n = P_{E_n}$ and $\|P_{E_n} x\|^2 \leq \|x\|^2$, we have

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

for $x \in \mathcal{H}$.

Theorem 1.7.3. Let $\{x_k\}_k$ be an orthonormal system in a Hilbert space $\mathcal{H}$. Then the corresponding Fourier series $S_n(x) = \sum_{k=1}^{n} \langle x, x_k \rangle x_k$ converges, i.e., $\lim\limits_{n \to \infty} S_n(x) = S_\infty(x)$ exists for $x \in \mathcal{H}$. Furthermore, $S_n = P_{E_n}$ for every n where E_n is the space spanned by $\{x_i\}_{i=1}^{n}$.[a]

[a] This includes $n = \infty$, where E_∞ is the **closure** of the space spanned by $\{x_i\}_i$.

Proof. We show that the sequence $S_n(x)$ for $n = 1, 2, \ldots$ is Cauchy. This is because

$$\|S_n(x) - S_m(x)\|^2 = \sum_{k=m+1}^{n} |\langle x, x_k \rangle|^2 ,$$

and Bessel's inequality implies $\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \leq \|x\|^2$. Hence, for any $\epsilon > 0$, there exists $m(\epsilon)$ such that

$$\sum_{k=m(\epsilon)+1}^{\infty} |\langle x, x_k \rangle|^2 < \epsilon,$$

which implies $\|S_n(x) - S_m(x)\|^2 < \epsilon$ if $n > m(\epsilon)$, hence $\{S_n(x) : n \geq 1\}$ is Cauchy, and $\lim\limits_{n \to \infty} S_n(x) = S_\infty(x) \in \mathcal{H}$. Also, $S_\infty = P_{E_\infty}$ where E_∞ is the closure of the linear space generated by the sequence $\{x_k : k \geq 1\}$. ■

Remark. Note that the closeness of E_∞ makes sense since the self-dual of a set's orthogonal complement is itself if it's closed in the first place.

Lecture 5: Abstract Fourier Series

1.7.3 Orthonormal Bases

13 Sep. 14:30

Follows from the discussion, we note that the Fourier series of x needs not converge to x, but we can still compute the point where it converges. However, we can now identify an extra condition so that the Fourier series of every vector x converges to x.

Definition 1.7.5 (Complete system). A system of vector $\{x_k\}_k$ in Hilbert space $\mathcal{H}$ is *complete* if the space spanned by $\{x_k\}_k$ is **dense** in $\mathcal{H}$, i.e., $\overline{\operatorname{span}(\{x_k\}_k)} = \mathcal{H}$.

Definition 1.7.6 (Orthonormal basis). A complete orthonormal system in a Hilbert space $\mathcal{H}$ is called an *orthonormal basis* of $\mathcal{H}$.

Theorem 1.7.4 (Fourier expansions). Let $\{x_k\}_k$ be an orthonormal basis of a Hilbert space $\mathcal{H}$. Then every vector $x \in \mathcal{H}$ can be expanded in its Fourier series

$$x = \sum_k \langle x, x_k\rangle x_k.$$

This is sometimes called Fourier inversion formula.

Proof. If an orthogonal set $\{x_k\}_k$ is complete, then $E_\infty = \mathcal{H}$, $P_{E_\infty} = I$. This implies $x = \sum_{k=1}^{\infty} \langle x, x_k\rangle x_k$ for $x \in \mathcal{H}$. ∎

Remark (Parseval's identity). From Theorem 1.7.4, we have $\|x\|^2 = \|P_{E_n} x\|^2 + \|(I - P_{E_n})\|^2$. By letting $n \to \infty$, we have

$$\|x\|^2 = \lim_{n\to\infty} \|P_{E_n} x\|^2 = \lim_{n\to\infty} \sum_{k=1}^{n} |\langle x, x_k\rangle|^2 = \sum_{k=1}^{\infty} |\langle x, x_k\rangle|^2 .$$

1.7.4 Gram-Schmidt Orthogonalization

Suppose $x_1, x_2, \ldots \in \mathcal{H}$ is a set of vectors and $E_n = \operatorname{span}(\{x_1, \ldots, x_n\})$. Then we can find an orthonormal set $\{y_k \in \mathcal{H} \colon k \geq 1\}$ such that $E_n = \operatorname{span}(\{y_1, y_2, \ldots, y_{m(n)}\})$ where $m(n) \leq n$.

Firstly, set $y_1 = x_1 / \|x_1\|$, and

$$y_n = \frac{(I - P_{E_{n-1}})x_n}{\|(I - P_{E_{n-1}})x_n\|}$$

if $x_n \notin E_{n-1}$, i.e., E_{n-1} is properly contained in E_n.

Remark. Proving completeness of a set of vectors $\{x_k \colon k \geq 1\}$ in $\mathcal{H}$ can be **non**-trivial.

Example (Haar basis). We consider the *Haar basis* for $L^2([0,1])$. Let $h\colon (0,1) \to \mathbb{R}$ where

$$h(t) = \begin{cases} 1, & \text{if } 0 < t < 1/2; \\ -1, & \text{if } 1/2 < t < 1. \end{cases}$$

Extend $h(\cdot)$ by zero outside $(0,1)$, we get $h\colon \mathbb{R} \to \mathbb{R}$, $h(t) = 0$ if $t \notin (0,1)$, otherwise it's the same as above. The function $t \mapsto h(2^k t)$ has support in interval $0 < t < 2^{-k}$. Move the support to interval $\ell 2^{-k} < t < (\ell+1)2^{-k}$ by translation. Set

$$h_{k,\ell}(t) = h(2^k t - \ell), \quad \ell = 0, 1, \ldots, 2^k - 1.$$

The constant function plus functions $h_{k,\ell}$, $k = 0, 1, 2, \ldots$, $0 \leq \ell \leq 2^k - 1$ are a complete orthogonal set for $\mathcal{H} = L^2([0,1])$.

Proof. The span of the Haar functions includes characteristics functions χ_F for all dyadic intervals $[2^{-k}\ell, 2^{-k}(\ell+1)]$ for $\ell = 0, 1, \ldots, 2^{k-1}$, $k = 0, 1, \ldots$. If the set is **not** complete, then there exists $f \in L^2([0,1])$ such that

$$\int_F f \,\mathrm{d}t = 0$$

for all dyadic intervals F. Since we can approximate any measurable set $E \subseteq (0,1)$ by a union of dyadic intervals.

Intuition. An easy way to see this is to consider

$$\left\{ F \in \mathcal{B} \colon \int_F f \,\mathrm{d}t = 0 \right\},$$

which is the Borel subalgebra of $\mathcal{B}$, which indeed is a Borel algebra on $(0,1)$. Then observe that dyadic intervals generate all open intervals.

Hence, we see that $\int_F f \,\mathrm{d}t = 0$ for all measurable $F \subseteq (0,1)$. Let $F = \{t \in (0,1) \colon f(t) > 0\}$, if $m(F) > 0$, then

$$\int_F f \,\mathrm{d}t > 0.$$

Hence, a contradiction, so $m(F) = 0$. ⊛

Example (Fourier basis). Consider the Fourier basis $e_k(t) = \frac{1}{\sqrt{2\pi}} e^{ikt}$ for $k \in \mathbb{Z}$, $-\pi < t < \pi$. This is complete in $L^2([-\pi,\pi])$.

Proof. We use Stone-Weierstrass theorem and apply it to Fourier basis. All $e_k(\cdot)$ are in $C[-\pi,\pi]$, i.e., continuous functions $f \colon [-\pi,\pi] \to \mathbb{C}$. We know that $C([-\pi,\pi])$ is a Banach space with supremum norm $\|f\| \coloneqq \sup_{t\in[-\pi,\pi]} |f(t)|$. Stone-Weierstrass theorem implies density of the space spanned by $e_k(\cdot)$, $k \in \mathbb{Z}$ in $C([-\pi,\pi])$, hence the completeness in $L^2([-\pi,\pi])$ follows from the density of continuous functions in $L^2([-\pi,\pi])$. ⊛

Proposition 1.7.1. Let $\{x_k\}_k$ be a linear independent system in a Hilbert space $\mathcal{H}$. Then the system $\{y_k\}_k$ obtained by Gram-Schmidt orthogonalization of $\{x_k\}_k$ is an orthonormal system in $\mathcal{H}$, and

$$\operatorname{span}(\{h_k\}_{k=1}^n) = \operatorname{span}(\{x_k\}_{k=1}^n)$$

for all $n \in \mathbb{N}$.

Proof. The system $\{y_k\}_k$ is orthonormal by construction, and we obviously have the inclusion $\operatorname{span}(\{y_k\}_k) \subseteq \operatorname{span}(\{x_k\}_k)$. Furthermore, since the dimensions of these subspaces both equal n by construction, so they're indeed equal. ∎

1.7.5 Existence of Orthogonal Bases

We see that from Proposition 1.7.1, we'll obtain that every Hilbert space that is not *too large* has an orthonormal basis. We call this Hilbert space separable.

Definition 1.7.7 (Separable). A metric space is *separable* if it contains a countable dense subset.

Remark (Banach space). For Banach space, separability follows from finding a countable set of vectors $\{x_k\}_k$ such that the span of $\{x_k\}_k$ is dense in E.

Chapter 2

Bounded Linear Operators

In this chapter we study certain transformations of Banach spaces. Because these spaces are linear, the appropriate transformations to study will be linear operators. Furthermore, since Banach spaces carry topology, it is most appropriate to study continuous transformations, i.e. continuous linear operators. They are also called bounded linear operators for the reasons that will become clear shortly.

2.1 Bounded Linear Functionals

2.1.1 Continuity and Boundedness

Definition. Let E be a linear space over $\mathbb{R}$ or $\mathbb{C}$.

Definition 2.1.1 (Linear functional)**.** A *linear functional* on E is a linear operator $f\colon E \to \mathbb{R}$ of $\mathbb{C}$ such that

$$f(ax + by) = af(x) + bf(y)$$

for $x, y \in E$, $a, b \in \mathbb{R}$ or $\mathbb{C}$.

Definition 2.1.2 (Bounded linear functional)**.** We say a linear functional $f(\cdot)$ is a *bounded linear functional* if

$$\|f\| \coloneqq \sup_{\|x\|=1} |f(x)| < \infty$$

by dilation and additive.

Clearly, the boundedness of $f(\cdot)$ implies $|f(x-y)| \leq \|f\| \|x-y\|$ for $x, y \in E$. Hence, $f(\cdot)$ is continuous and in fact Lipschitz continuous.

Remark. Conversely, if a linear functional is continuous then it is bounded.

Proof. Suppose $f(\cdot)$ is not bounded, then there exists a sequence $x_n \in E$ such that $|f(x_n)| \geq n\|x_n\|$ for $n = 1, 2, \ldots$. By linearity,

$$\left| f\left(\frac{x_n}{n\|x_n\|}\right) \right| \geq 1, \quad n = 1, 2, \ldots.$$

But we know $\lim\limits_{n\to\infty} \frac{x_n}{n\|x_n\|} = 0$ and $f(0) = 0$, hence $f(\cdot)$ is not continuous at 0. ⊛

2.1.2 Dual Spaces and Hyperplanes

Definition 2.1.3 (Dual space)**.** Let E be a normed space. The space of all bounded linear functionals $f(\cdot)$ on E is known as the *dual space* E^* of E.

The dual space is also a normed space with norm $\|f\| := \sup_{\|x\|=1} |f(x)|$, which is in fact a Banach space. And it is a Banach space even if the original E is not. This definition implies $|f(x)| \leq \|f\| \|x\|$ for $x \in E$, $f \in E^*$. Also, $\|f\|$ is the smallest number in this inequality that makes it valid for all $x \in X$.

Definition 2.1.4 (Hyperplane)**.** Let E be a linear space and $H \subseteq E$ is a subspace. Say H is a *hyperplane* if $\operatorname{codim}(H) = 1$, i.e., $\dim(E / H) = 1$.

The goal is to make an equivalence between bounded linear functionals on E and *closed* hyperplanes in E.

Problem 2.1.1. Does there exist a **non**-closed hyperplane?

Answer. We know that this is not the case in finite dimension. And this question is analogous to asking *does there exist a subset $F \subseteq \mathbb{R}$ which is **not** Lebesgue measurable?* The answer to this is yes in both cases. However, construction uses axiom of choice. ⊛

Turns out that there is a canonical correspondence between the lienar functionals and the hyperplanes in E. This is clarified in Proposition 2.1.1.

Proposition 2.1.1 (Linear functionals and hyperplanes)**.** Let E be a linear space.

(a) For every linear functional on E, $\ker(f)$ is a hyperplane in E. If E is a Banach space, and $f(\cdot)$ is bounded, then $\ker(f) = H$ is closed.

(b) If $f, g \neq 0$ are linear functionals on E such that $\ker(f) = \ker(g)$, then $f = ag$ for some $a \neq 0$.

(c) For every hyperplane $H \subseteq E$, there exists a linear functional $f \neq 0$ on E such that $\ker(f) = H$. If E is a Banach space, and $\ker(f) = H$ is closed, then $f(\cdot)$ is bounded.

Lecture 6: Riesz Representation Theorem

Let's first see the proof of Proposition 1.7.1. 15 Sep. 14:30

Proof of Proposition 1.7.1. We prove them in order.

(a) Let $x, y \notin \ker(f)$, then $f(x), f(y) \neq 0$, meaning that there exists a scalar $\lambda \neq 0$ such that $f(x) = \lambda f(y)$, i.e., $x - \lambda y \in \ker(f)$. Hence, if $[x], [y] \in E / \ker(f)$, $[x] = \lambda[y]$, implying $\dim(E / \ker(f)) = 1$. Now, if f is bounded, then f is continuous, so $\ker(f) = f^{-1}(\{0\})$ is closed.

(b) Consider the induced functionals $\widetilde{f}, \widetilde{g} \colon E / H \to \mathbb{R}$ or $\mathbb{C}$ where $H = \ker(f) = \ker(g)$. This implies

$$\dim \left({}^{E}\!\Big/_{H} \right) = 1 \Rightarrow \widetilde{f} = a\widetilde{g} \text{ for some } a \neq 0 \Rightarrow f = ag.$$

(c) Assume $\dim(E / H) = 1$, so $E / H = \{a[x_0] \colon a \in \mathbb{C} \text{ (or } \mathbb{R})\}$ for some $x_0 \in E$. Then, for any $x \in E$, $[x] = a(x)[x_0]$ for some $a(x) \in \mathbb{C}$ or $\mathbb{R}$. Define $f(x) := a(x)$, we see that f is linear and $\ker(f) = H$. Now, if E is a Banach space and H is closed with $\dim(E / H) = 1$. Recall that E / H is also a Banach space with norm $\|[x]\| = \inf_{y \in H} \|x + y\|$ for $x \in E$.[a] Let $\widetilde{f}$ be a linear functional on E / H. Since $\dim(E / H)$ is finite, $\widetilde{f}$ is continuous, implying $|\widetilde{f}([x])| \leq A \|[x]\|$ for all $x \in E$ for some scalar A. Finally, we define $f(x) = \widetilde{f}([x])$ for $x \in E$, then $\ker(f) = H$ and $|f(x)| \leq A \|[x]\| \leq A \|x\|$.

■

[a] We see now why we need the closure: otherwise we'll get a non-zero function with norm 0.

2.2 Representation Theorems

In concrete Banach spaces, the bounded linear functionals usually have a specific and useful form. Generally speaking, all linear functionals on function spaces (such as L_p and $C(K)$) act by integration of the function (with respect to some weight or measure). Similarly, all linear functionals on sequence spaces (such as ℓ_p) act by summation with weights.

We now start by characterizing bounded linear functionals on a Hilbert space $\mathcal{H}$.

Theorem 2.2.1 (Riesz representation theorem)**.** Let $\mathcal{H}$ be a Hilbert space. Then we have the following.

(a) For every $y \in \mathcal{H}$, then function $f(x) = \langle x, y\rangle$ for $x \in \mathcal{H}$ is a bounded linear functional on $\mathcal{H}$.

(b) If $f\colon \mathcal{H} \to \mathbb{C}$ or $\mathbb{R}$ is a bounded linear functional on $\mathcal{H}$, then there exists $y \in \mathcal{H}$ such that $f(x) = \langle x, y\rangle$ for all $x \in \mathcal{H}$. Hence, the dual $\mathcal{H}^*$ of $\mathcal{H}$ is isometric to $\mathcal{H}$.

Proof. We prove this in order.

(a) $f(x) = \langle x, y\rangle$ is clearly a linear functional. Boundedness follows form Cauchy-Schwarz inequality such that

$$\lvert f(x)\rvert = \lvert\langle x, y\rangle\rvert \leq \lVert x\rVert \, \lVert y\rVert ,$$

and we can achieve $\lVert f\rVert = \lVert y\rVert$ by setting $x = y / \lVert y\rVert$.

> **Note.** Note that there exists x_f such that $\lVert x_f\rVert = 1$ since $\lVert f\rVert = \sup_{\lVert x\rVert=1} \lvert f(x)\rvert = f(x_f)$, i.e., the supremum is achieved, although we're working on an infinite dimensional space. This property does not always hold for bounded linear functionals on Banach space since the unit ball can be not compact. But this holds for Hilbert space.

(b) Let $f\colon \mathcal{H} \to \mathbb{C}$ or $\mathbb{R}$ be a bounded linear functional on $\mathcal{H}$. Let $H = \ker(f)$, which is closed from Proposition 1.7.1. Let $H^{\perp}$ be the orthogonal complement of H, i.e., $\mathcal{H} = H \oplus H^{\perp}$. Then $\dim(\mathcal{H} \,/\, H) = 1 \Rightarrow \dim(H^{\perp}) = 1$. Choose $y' \in H^{\perp}$ such that $g(x) = \langle x, y'\rangle$, which is in $\mathcal{H}^*$ from (i). Furthermore, we see that $\ker(g) = \ker(f)$, so from Proposition 1.7.1, f and g are equal up to a constant $\lambda \in \mathbb{C}$ or $\mathbb{R}$, i.e., $f = \lambda g$. It follows that

$$f(x) = \lambda g(x) = \lambda \langle x, y'\rangle = \langle x, \lambda y'\rangle =: \langle x, y\rangle$$

for $y := \lambda y'$, hence we're done.[a]

■

[a] We can even show that y here is unique.

In a concise form, Riesz representation theorem can be realized as $\mathcal{H}^* = \mathcal{H}$. Given a Hilbert space $\mathcal{H}$, Riesz representation theorem identifies the dual space $\mathcal{H}^*$, which can be used to show Radon-Nikodym theorem.

Theorem 2.2.2 (Radon-Nikodym theorem)**.** Let μ, ν be two finite measures such that $v \ll \mu$, i.e., ν is absolutely continuous w.r.t, μ.[a] Then there exists $g \geq 0$ such that g is μ-integrable and

$$\nu(A) = \int_A g \,\mathrm{d}\mu$$

for A measurable.

[a] This means $\mu(A) = 0 \Rightarrow \nu(A) = 0$.

Proof. Consider the linear functional $F\colon L^2(\mu) \to \mathbb{R}$ or $\mathbb{C}$ such that

$$F(f) = \int_\Omega f \,\mathrm{d}\mu.$$

Then we have $\|F(f)\| \leq \|f\|_2 \sqrt{\mu(\Omega)}$, i.e., F is also a bounded linear functional on $L^2(\mu+\nu)$, hence by Theorem 2.2.1, there exists $h \in L^2(\mu+\nu)$ such that

$$F(f) = \int_\Omega fh \,\mathrm{d}(\mu+\nu)$$

for $f \in L^2(\mu+\nu)$, i.e.,

$$\int_\Omega f \,\mathrm{d}\mu = \int_\Omega fh \,\mathrm{d}\mu + \int_\Omega fh \,\mathrm{d}\nu \tag{2.1}$$

if $f \in L^2(\mu+\nu)$. This further implies

$$\int_\Omega fh \,\mathrm{d}\nu = \int_\Omega f[1-h] \,\mathrm{d}\mu \tag{2.2}$$

for $f \in L^2(\mu+\nu)$.

Claim. Such h satisfies $0 < h \leq 1$ μ-a.e., moreover, $(\mu+\nu)$-a.e.

Proof. We first note that $\mu(A) = 0 \Leftrightarrow \mu(A) + \nu(A) = 0$. Let $A = \{h \leq 0\}$, $f = \mathbb{1}_A$ be the characteristic function on A. Then Equation 2.1 implies

$$\int_A h(\,\mathrm{d}\mu + \,\mathrm{d}\nu) \leq 0 \Rightarrow \mu(A) = 0 \Rightarrow h > 0 \ \mu \text{ a.e.}$$

But since g is a positive function, so we also need $h \leq 1$. Again, set $B = \{h > 1\}$, $f = \mathbb{1}_B$. Then Equation 2.1 implies

$$\mu(B) = \int_B h\,(\,\mathrm{d}\mu + \,\mathrm{d}\nu) > \mu(B)$$

unless $\mu(B) = 0$. ⊛

Now, by using monotone convergence theorem, we conclude[a] that Equation 2.2 holds for all $f \geq 0$, $f \in L^2(\mu+\nu)$.[b] Finally, let $A \subseteq \Omega$ measurable and $hf = \chi_A$, from Equation 2.2,

$$\nu(A) = \int_A \frac{1-h}{h} \,\mathrm{d}\mu.$$

By letting $g := 1 - h/h \Rightarrow g = \mathrm{d}\nu/\mathrm{d}\mu$, we're done. ■

[a] Consider $f_n(t) := \min(f(t), n)$ and let $n \to \infty$.
[b] Both sides could be ∞.

Notation (Radon-Nikodym derivative). g in Theorem 2.2.2 is referred to as the *Radon-Nikodym derivative* where $g := \mathrm{d}\nu/\mathrm{d}\mu$.

Note (Uniqueness). The uniqueness of Radon-Nikodym derivatives can be shown via

$$\int_A g \,\mathrm{d}\mu = 0$$

for all μ-measurable A, i.e., $g = 0$ μ-a.e.

Another useful application of Theorem 2.2.1 is to characterize L_p and ℓ_p spaces and their dual L_p^* and ℓ_p^*. We first see the follwoing.

Remark. Consider spaces $L^p(\Omega, \mu)$ for $1 \leq p \leq \infty$, $L^q(\Omega, \Sigma, \mu) \subseteq (L^p(\Omega, \Sigma, \mu))^*$ where $1/p + 1/q = 1$.

Proof. The easy part is that $g \in L^q$ induces a bounded linear functional on L^p by setting

$$F(f) = \int_\Omega fg \,\mathrm{d}\mu.$$

By Hölder's inequality, $|F(f)| \leq \|f\|_p \|g\|_q$, hence $\|F\| \leq \|g\|_q$. To show the equality and $\sup_{\|f\|_p} |F(f)|$ is attained for $1 < p < \infty$, we choose $f = g^{q-1}\operatorname{sgn}(g)$ since

$$F(f) = \int_\Omega |g|^q \,\mathrm{d}\mu = \|g\|_q^q,$$

and from $1/p + 1/q = 1 \Rightarrow q - 1 = q/p$, we have

$$\|f\|_p^p = \int |f|^p \,\mathrm{d}\mu = \int_\Omega |g|^q \,\mathrm{d}\mu = \|g\|_q^q \Rightarrow \|f\|_p = \|g\|_q^{q/p} = \|g\|_q^{q-1}.$$

This implies

$$F(f) = \int_\Omega |g|^q \,\mathrm{d}\mu \Rightarrow \|g\|_q^q = \|g\|_q \|f\|_p.$$

> **Note.** We see that $\sup_{\|f\|_p=1} |F(f)|$ is attained by taking $f = \operatorname{sgn}(g)$.

⊛

In particular, we have the following.

Theorem 2.2.3 ($L_p^* = L_q$). Consider the space $L_p = L_p(\Omega, \Sigma, \mu)$ with finite measure of σ-finite measure μ. Then for $1 \leq p < \infty$ and the conjugate exponent q of q.

(a) For every weight function $g \in L_q$, integration with weight

$$\int_\Omega fg \,\mathrm{d}\mu$$

for $f \in L_p$ is a bounded linear functional on L_p, and its norm is $\|G\| = \|g\|_q$.

(b) Conversely, every bounded linear functional $G \in L_p^*$ can be represented as integration with weight for some unique weight function $g \in L_q$. Moreover, $\|G\| = \|g\|_q$.

Lecture 7: Hahn-Banach Theorem

20 Sep. 14:30

> **Remark.** When $p = 1$, the supremum is not attained necessarily. Consider $g \in L_\infty$, $F(f) := \int fg \,\mathrm{d}\mu$ is dual of L_1. If $g(\cdot)$ is continuous on $\mathbb{R}$ with unique maximum, then the supremum $\sup_{\|f\|_1} |F(f)|$ is not attained. In all, for $1 \leq p \leq \infty$, L_q contained in the dual of L_p. If $1 < p \leq \infty$, then $\sup_{\|f\|_p=1} |F(f)|$ is attained. For $p = 1$, the supremum is not necessarily attained.

Now, we're ready to prove Theorem 2.2.3.

Proof of Theorem 2.2.3. To show that the dual of L_p^* is L_q if $1 \leq p < \infty$ where $1/p + 1/q = 1$, we use Theorem 2.2.2. Suppose $E = L_p(\Omega, \Sigma, \mu)$ with $1 \leq p < \infty$ and $f \in E^*$. Just consider finite measure space, i.e., $\mu(\Omega) < \infty$. We define a measure ν on Σ by $\nu(A) := F(\chi_A)$ for $A \in \Sigma$, where χ_A is the characteristic function of A. We see that

$$\mu(A) = 0 \Rightarrow \nu(A) = 0 \Rightarrow \nu \ll \mu,$$

and Theorem 2.2.2 implies

$$\nu(A) = \int_A g \,\mathrm{d}\mu$$

for some $g =: \frac{d\nu}{d\mu} \in L_1(\Omega, \Sigma, \mu)$. Note that g may not be in L_q since $q > 1$. Hence, $F(f) = \int_\Omega fg \, d\mu$ for all simple function f assuming $g \geq 0$. Set $f = g^{q-1}$ with the fact that $\|F(f)\| \leq \|F\|_p \|f\|_p$. Recall that $q - 1 = q/p$, hence

$$\int g^q \, d\mu \leq \|F\|_p \left(\int g^q \, d\mu \right)^{1/p} \Rightarrow \|g\|_q^q \leq \|g\|_q^q \leq \|F\|_p \|g\|_q^{q/p} = \|F\|_p \|g\|_q^{q-1},$$

hence $\|g\|_q \leq \|F\|_p$.

Note. We assume $g \geq 0$ is because ν is a sign measure, then if we have a bounded variation function, we can just break it into $\nu^+ + \nu^-$.

■

Remark. L_1 is a subset of L_∞^* but not equal to it. If $F \colon L_\infty(\mu) \to \mathbb{C}$ is bounded linear functional, then if $\Omega = K$ is a compact Hausdorff space, F induces a bounded linear functional on $C(K)$, i.e., the space of continuous functions on K. We see that $C(K) \subseteq L_\infty(K, \Sigma, \mu)$ where Σ is the Borel algebra on K.

Theorem 2.2.4. Let $E = C(K)$ be the space of continuous functions on compact Hausdorff space K. Then we have the following.

(a) For every Borel regular signed measure on K, the functional $F(f) = \int_K f \, d\mu$ is a bounded linear functional on K.

(b) Every bounded linear functional on $C(K)$ can be expressed as $F(f) = \int_K f \, d\mu$ for some measure μ, and $\|F\| = |\mu|(K)$, i.e., $TV(K)$.

2.3 Hahn-Banach Theorem

Hahn-Banach theorem allows one to extend continuous linear functionals f from a subspace to the whole normed space, while preserving the continuity of f. Hahn-Banach theorem is a major tool in functional analysis. Together with its variants and consequences, this result has applications in various areas of mathematics, computer science, economics and engineering.

Theorem 2.3.1 (Hahn-Banach theorem)**.** Let E_0 be a subspace of a Banach space E. Then every $f_0 \colon E_0 \to \mathbb{R}$ or $\mathbb{C}$ has a continuous extension $f \colon E \to \mathbb{R}$ or $\mathbb{C}$ such that $\|f\| = \|f_0\|$.

Before proving this, let's first see some implications.

Theorem 2.3.2 (Supporting functional)**.** Let E be a Banach space. For every $x \in E$, there exists $f \in E^*$ such that $\|f\| = 1$, $f(x) = \|x\|$. i.e., $\sup_{\|y\|=1} |f(y)|$ attained at $y = x$.

Proof. Consider dimension 1 space $E_0 = \operatorname{span}(x) = \{tx, t \in \mathbb{R} \text{ or } \mathbb{C}\}$. Define $f_0 \colon E_0 \to \mathbb{R}$ or $\mathbb{C}$ such that $f_0(tx) = t\|x\|$. We see that $\|f_0\| = 1$, and Theorem 2.3.1 implies there exists $f \in E^*$, $f(x) = \|x\|$ such that $\|f\| = 1$. ■

Remark (Geometric interpretation)**.** Let B be a unit ball $\{x \in E \colon \|x\| \leq 1\}$ in a real Banach space E. Choose $x_0 \in \partial B$ such that $\|x_0\| = 1$. Then there exists $f \in E^*$, $\|f\| = 1$, $f(x) = \|x\|$. Let $H = \ker(f) + x_0$ where H intersects B at x_0, we see that H divides E into 2 disjoint subsets, while B lies in one of which.

Proof. Since $x \in B$ and $\|x\| < 1$ implies $|f(x)| \leq \|x\| < 1$, we have $f(x) < 1$, i.e., $B \subseteq \{x \colon f(x) < 1\}$ and $E = \{x \colon f(x) < 1\} \cup H \cup \{x \colon f(x) > 1\}$. ⊛

Note. Notice that we don't have uniqueness (as we don't have it in Theorem 2.3.1) since a unit ball in L^∞ has corner, which will give multiple hyperplanes...

Lecture 8: Proof of Hahn-Banach Theorem and Duality

We now see the proof of Hahn-banach theorem. 20 Sep. 14:30

Proof of Theorem 2.3.1. We assume E is separable, otherwise we need transfinite induction.

Note. Separability allows us to extend f_0 one dimension at a time.

Let $\{x_n : n \geq 1\}$ have the property that its span is dense in E. Now, if we can extend f_0 such that $E_0 \to E_0 + \{x_1\} \to E_0 + \{x_1, x_2\} \to \ldots \to E_0 + \operatorname{span}(\{x_n : n \geq 1\})$, then we can have $\|f\| = \|f_0\|$, with the final space is dense in E, we can extend f to E by continuity.

To extend f by 1 dimension, i.e., $E \to E + \{x_1\}$. Note that extension is determined by a single number $\gamma = f(x_1)$ since f is a linear functional. Firstly, we want that $\|f\| = \|f_0\|$ such that the linear functional $f_0 \colon E_0 \to \mathbb{R}$ extends to $f \colon D_0 + \{x_1\} \to \mathbb{R}$, i.e., we want

$$|f_0(x_0) + \lambda\gamma| \leq \|x_0 + \lambda x_1\|$$

for $x_0 \in E$, $\lambda \in \mathbb{R}$. By dividing the inequality by $\lambda \neq 0$, it's sufficient to find γ such that $|f_0(x_0) + \gamma| \leq \|x_0 + x_1\|$, $x_0 \in E_0$.

Suppose f_0 is a real-valued function, we need

$$-\|x_0 + x_1\| \leq f_0(x_0) + \gamma \leq \|x_0 + x_1\|$$

for all $x_0 \in E_0$. Such a γ exists, provides $\|x_0 + x_1\| - f_0(x_0) \geq -\|x_0' + x_1\| - f_0(x_0')$ for all $x_0, x_0' \in E_0$. Furthermore, this is equivalent to write

$$f_0(x_0 - x_0') \leq \|x_0 + x_1\| + \|x_0' + x_1\|$$

for all $x_0, x_0' \in E_0$, i.e., $f_0(x_0 - x_0') \leq \|x_0 + x_1\| + \|-x_1 - x_0'\|$ for $x_0, x_0' \in E_0$. Recall that $\|f_0\| = 1$, we have

$$f_0(x_0 - x_0') \leq \|x_0 - x_0'\| \leq \|x_0 + x_1\| + \|-x_1 - x_0'\|.$$

For complex valued f, consider $f \colon E \to \mathbb{C}$ be a linear functional over $\mathbb{C}$ and let $g(x) = \operatorname{Re} f(x)$. Then $g \colon E \to \mathbb{R}$ is a real-valued linear functional. We see that $f(x) = g(x) - ig(ix)$ for all $x \in E$.[a] Conversely, if $g \colon E \to \mathbb{R}$ is a real linear functional on Banach space E over $\mathbb{C}$, then $f \colon E \to \mathbb{C}$ defined by $f(x) = g(x) - ig(ix)$, $x \in E$ is a complex linear functional on E.

But we need to be a bit careful since when we extend $f_0 \colon E_0 \to \mathbb{C}$, we're extending 2 real dimensions since for $g_0 = \operatorname{Re} f_0$, we need to do $E_0 \to E_0 + \{x_1\} \to E_0 \to \{x_1, x_2\}$. Again, define $f(\cdot) = g(\cdot) - ig(i\cdot)$, we want to show $|f| = \|f_0\|$. We use the fact that for $x \in E_0 + \{\lambda x_0 \colon \lambda \in \mathbb{C}\}$,

$$e^{i\theta} f(x) = f(xe^{i\theta})$$

for $\theta \in \mathbb{R}$. Choose θ such that $f(xe^{i\theta}) = g(xe^{i\theta})$, and since we already have $\left|g(xe^{i\theta})\right| \leq \|f_0\| \left\|xe^{i\theta}\right\|$, we see that $|f(x)| \leq \|f_0\| \|x\|$ for $x \in E_0 + \{\lambda x_1 \colon \lambda \in \mathbb{C}\}$. ■

[a] Since $f(ix) = if(x)$, hence $g(ix) = -\operatorname{Im} f(x)$.

Before we end this section, we see some corollaries of Hahn-Banach theorem. From Theorem 2.3.2, we see that for every vector x, we indeed attain its norm on some functional $f \in E^*$, i.e., their supporting functional. But recall that the norm of a functional $f \in E^*$ is defined as

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

and in general, f will not attain its norm on some vector x. Above observation leads to the following.

Corollary 2.3.1. For every vector x in a normed space E,

$$\|x\| = \max_{f \neq} \frac{|f(x)|}{\|f\|}$$

where the maximum is taken over all non-zero functionals $f \in E^*$.

Hahn-Banach theorem implies that there are enough bounded linear functionals $f \in E^*$ on every space E. One manifestation of this is the following.

Corollary 2.3.2 (Separation of points). For every two vectors $x_1 \neq x_2$ in a normed space E, there exists a functional $f \in E^*$ such that $f(x_1) \neq f(x_2)$.

Proof. The supporting functional $f \in E^*$ of the vector $x = x_1 - x_2$ must satisfy

$$f(x_1 - x_2) = \|x_1 - x_2\| \neq 0,$$

as required. ∎

2.4 Second Dual Space

Let E be a normed space, then the functionals f^* are designed to act on vectors $x \in E$ via

$$f : x \mapsto f(x).$$

But indeed, we can instead say that *vectors $x \in E$ act on functionals $f \in E^*$* via

$$x : f \mapsto f(x).$$

Thus, a vector $x \in E$ can itself be considered as a function from E^* to $\mathbb{R}$. Furthermore, this function x is clearly linear, so we may consider x as a linear functional on E^*. Also, the inequality

$$|f(x)| \leq \|x\| \, \|f\|$$

shows that this functional is bounded, so $x \in E^{**}$. We may instead write x as x^{**} for clarity. Note that the norm of x^{**} as a functional is $\|x^{**}\|_{E^{**}} \leq \|x\|$ since

$$\|x^{**}\| = \sup_{\substack{\|f\|=1 \\ f \in E^*}} |x^{**}(f)| = \sup_{\substack{\|f\|=1 \\ f \in E^*}} |f(x)| \leq \|x\|,$$

implying that $\|x^{**}\| \leq \|x\|$ for all $x \in E$. But from supporting functional $f \in E^*$ of x, we actually have

$$\|x^{**}\| = \|x\|,$$

i.e., we have a *canonical embedding* of E into E^{**}. The above discussion leads to Theorem 2.4.1.

Theorem 2.4.1 (Second dual space). Let E be a normed space. Then E can be considered as a linear subspace of E^{**}. For this, a vector $x \in E$ is considered as a bounded linear functional on E^* via the action

$$x : f \mapsto f(x), \quad f \in E^*.$$

2.5 Reflexive Spaces

Definition 2.5.1 (Reflexive space). A normed space E is called *reflexive space* if $E \to E^{**}$.[a]

[a] Under the canonical embedding.

Example. L_p spaces for $1 < p < \infty$ are reflexive spaces.

Proof. We know that $L_p^* = L_q$ where $1 \leq p < \infty$ for q being the conjugate index of p. ⊛

Example. L_p spaces for $p = 1$ or ∞ are not reflexive spaces

Proposition 2.5.1. Let E be a reflexive space, then every linear functional $f \in E^*$ attains its norm on E.

Proof. By reflexivity, the supporting functional of f is a vector $x \in E^{**} = E$, thus $\|x\| = 1$ and $f(x) = \|f\|$, as required. ∎

Remark (James' theorem)**.** The converse of Proposition 2.5.1 is also true, i.e., if every functional $f \in E^*$ on a Banach space E attains its norm, then E is are reflexive. This is the so-called *James' theorem*.

Lecture 9: Hahn-Banach Theorem for Sublinear Functions

From Proposition 2.5.1, we see that to show a Banach space E is not reflexive, it's sufficient to find $f \in E^*$ such that $\sup_{\|x\|=1} |f(x)|$ is not attained. 27 Sep. 14:30

Example. Let $C([0, 1])$ be the space of continuous functions $g\colon [0, 1] \to \mathbb{C}$ with $\|g\| := \sup_{0 \leq t \leq 1} |g(t)|$. Then for $f \in E^*$,

$$f(g) = \int_0^1 h(x)g(x)\,\mathrm{d}x$$

for

$$h(x) = \begin{cases} -1, & \text{if } 0 < x < \frac{1}{2}; \\ 1, & \text{if } \frac{1}{2} < x < 1. \end{cases}$$

Then we have $\|f\| = 1 = \sup_{\|g\|=1} |f(g)|$, but the supremum is not attained since g needs to be continuous.

2.6 Separation of Convex Sets

In this section, we can extend supporting functional theorem such that we now have it for arbitrary convex sets other than the unit ball. We see that supporting functional theorem depends on Hahn-Banach theorem, so we should first generalize Hahn-Banach theorem.

2.6.1 Sublinear Functions

By looking into the proof of Hahn-Banach theorem, we see that we only used positive homogeneity and triangle inequality of the axiom of norm, which suggests we define the following.

Definition 2.6.1 (Sublinear)**.** Let E be a linear vector space. a function $\|\cdot\| : E \to [0, \infty)$ is *sublinear* if it satisfies

(a) $\|\lambda x\| = \lambda \|x\|$ for $\lambda \in \mathbb{R}^+$, $x \in E$.

(b) $\|x + y\| \leq \|x\| + \|y\|$, $x, y \in E$.

Remark (Differences from norm)**.** Note that for a sublinear function to be a norm, we need

(a) $\|-x\| = \|x\|$, $x \in E$

(b) $\|x\| = 0 \Rightarrow x = 0$.

Theorem 2.6.1 (Hahn-Banach theorem for sublinear functions)**.** Let E_0 be a subspace of a linear vector space over $\mathbb{R}$. Let $\|\cdot\|$ be a sublinear functional on E, and $f_0 \colon E_0 \to \mathbb{R}$ be a linear functional on E_0 satisfying $f_0(x) \leq \|x\|$ for $x \in E_0$. Then f_0 admits an extension f to E such that $f(x) \leq \|x\|$ for $x \in E$.

Proof. The idea is the same from Theorem 2.3.1. ■

2.6.2 Geometric Properties of Sublinear Functions

We see that by considering sublinear functionals instead of norms offers us more flexibility in geometric applications. In particular, sublinear functionals arise as Minkowski functionals of convex sets.

Definition 2.6.2 (Absorbing)**.** A subset K of a linear vector space is *absorbing* if

$$E = \bigcup_{t \geq 0} tK$$

where $tK \coloneqq \{tk \colon k \in K\}$.

Definition 2.6.3 (Minkowski functional)**.** Let K be an absorbing convex subset of a linear vector space E such that $0 \in K$. Then the *Minkowski functional* $\|\cdot\|_K$ is defined as

$$\|x\|_K \coloneqq \inf \{t > 0 \colon x/t \in K\}.$$

Proposition 2.6.1. Let K be an absorbing convex subset of a linear vector space E such that $0 \in K$. Then Minkowski functional $\|x\|_K$ is a sublinear functional on E. Conversely, let $\|\cdot\|$ be a sublinear functional on a linear vector space E, then the sub-level set

$$K = \{x \in E \colon \ \|x\| \leq 1\}$$

is an absorbing convex set, and $0 \in K$.

Proof. To prove the forward direction, the main observation is that since $0 \in K$ and K is convex, then $x \in K \Rightarrow tx \in K$ if $0 \leq t < 1$. To show dilation, for $\lambda > 0$,

$$\|\lambda x\| = \inf \left\{ t > 0 \colon x \in \frac{t}{\lambda} K \right\} = \lambda \inf \{s > 0 \colon x \in sK\} = \lambda \|x\|.$$

To show triangle inequality, suppose $x \in tK$, $y \in sK$, then $x = tk_1$, $y = sk_2$ for some $k_1, k_2 \in K$. We then have

$$x + y = (t+s)\left(\frac{t}{t+s} k_1 + \frac{s}{t+s} k_2 \right) = (t+s)k$$

for some $k \in K$ since K is convex, hence $x + y \in (t+s)K$, we then have $\|x+y\| \leq \|x\| + \|y\|$.

Now, if $\|\cdot\|$ is sublinear, then $K = \{x \in E \colon \ \|x\| \leq 1\}$ is absorbing, convex and $0 \in K$.[a] ■

[a] $0 \in K$ since $\|0\| = 0$, while the convexity comes from the triangle inequality.

Remark. If $K \neq -K$, then $\exists x \in E$ with $\|x\| \neq \|-x\|$. If $K = E$, then $\|\cdot\| \equiv 0$.

2.6.3 Separation of Convex Sets

Theorem 2.6.2 (Separation of a point from a convex set)**.** Let K be an open convex subset of a normed space E and $x_0 \notin K$. Then there exists a continuous linear functional $f\colon E \to \mathbb{R}$ with $f \equiv 0$ and $f(x) < f(x_0)$ for $x \in K$.

Proof. By translation, we can assume without loss of generality that $0 \in K$. Since K is open, it is absorbing. Now, let $\|\cdot\|_K$ be the Minkowski functional, then

$$\|x\|_K \leq \frac{1}{r}\|x\|$$

for $x \in E$ if $B(0, r) \subseteq K$. Proceed as in Theorem 2.3.2 for unit ball, we define f_0 on $\operatorname{span}(x_0)$ by

$$f_0(tx_0) = t\|x_0\|$$

for $t \in \mathbb{R}$. Then if $E_0 = \{\lambda x_0 \colon \lambda \in \mathbb{R}\}$, $f_0(x) \leq \|x\|_K$ for $x \in E_0$ since $\|tx\|_K = t\|x\|_K$ for $t \geq 0$. By Theorem 2.3.1, we can extend f_0 to $f\colon E \to \mathbb{R}$ such that

$$f(x) \leq \|x\|_K \leq \frac{1}{r}\|x\|$$

for $x \in E$, hence f is a bounded linear functional. For separation, we see that

$$f(x) \leq \|x\|_K \leq 1 \leq \|x_0\|_K = f_0(x_0) = f(x_0)$$

for $x \in K$, hence $f(x) \leq f(x_0)$. Now, since K is open, so $x + tv \in K$ for some $t > 0$ and all v with $\|v\| = 1$. Hence, we have

$$f(x + tv) \leq f(x_0)$$

for $t = t_x > 0$, $\|x\| = 1$, which further implies

$$f(x) + t \sup_{\|x\|=1} f(x) \leq f(x_0) \Rightarrow \sup_{\|x\|=1} > 0 \Rightarrow f(x) < f(x_0).$$

■

Theorem 2.6.3 (Separation of convex sets)**.** Let A, B be disjoint convex subsets of a Banach space E.

(a) If A is open, then there exists a bounded linear functional $f\colon E \to \mathbb{R}$ such that $f(a) < f(b)$ for $a \in A$, $b \in B$.

(b) If A, B are closed and B is compact, then $\sup_{a \in A} f(a) < \inf_{b \in B} f(b)$.

Proof. We have the following.

(a) Let $K = A - B = \{a - b \colon a \in A, b \in B\}$. Then K is open, convex and $0 \notin K$. By Theorem 2.6.2, there exists $f \in E^*$ such that

$$f(a - b) < f(0) = 0$$

for $a \in A$, $b \in B$, hence $f(a) < f(b)$ for $a \in A$, $b \in B$.

(b) Let A be closed, B be compact. Then we have

$$d(A, B) = \inf\{\|x - y\| : x \in A, y \in B\} = r > 0.$$

Define $A_\delta \coloneqq \{x \in E \colon d(x, A) < \delta\}$ where A_δ is open. By setting $\delta \coloneqq r/2$, we have $A_\delta \cap B = 0$. From (a), we see that there exists $f \in E^*$ such that $f(x) < f(y)$ for $x \in A_\delta$, $y \in B$. Then $a \in A$ implies $a + \delta/2v \in A_\delta$ if $\|v\| = 1$, hence

$$f(a + \delta/2v) < f(b)$$

for $b \in B$. So

$$f(a) + \frac{\delta}{2} f(v) < f(b)$$

for $b \in B$, $\|v\| = 1$. Take the supremum over $\|v\| = 1$, we have $\sup_{\|v\|=1} |f(v)| = \delta > 0$, implying $f(a) < f(b) - \delta$, $a \in A$, $b \in B$. Finally, we have

$$\sup_{a \in A} f(a) < \inf_{b \in B} f(b).$$

■

Appendix

Appendix A

Additional Proofs

Bibliography

[Lax02] P.D. Lax. *Functional Analysis*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2002. ISBN: 9780471556046. URL: https://books.google.com/books?id=18VqDwAAQBAJ.

[RS80] M. Reed and B. Simon. *I: Functional Analysis*. Methods of Modern Mathematical Physics. Elsevier Science, 1980. ISBN: 9780125850506. URL: https://books.google.com/books?id=hInvAAAAMAAJ.