Notes on Functional Analysis

Ziquan Yang

October 12, 2015

Contents

1 Preliminaries

In algebra, we know that if there is a group homomorphism $h : A \to B$, where B is abelian, then h will factor through the abelianization of A. B.L.T theorem says that if $T : \langle V_1, \| \cdot \|_1 \rangle \to \langle V_2, \| \cdot \|_2 \rangle$ is a bounded linear transformation where V_2 is complete, then T will factor through the completion of V_1.

A main point in the philosophy of functional analysis is that in order to define something on a normed linear space, it is often convenient to define it on a dense set and extend it by B.L.T. For example, this idea can be used to define Riemann integral on $PC[a, b]$ very quickly. ($PC[a, b]$ is the set of piecewise continuous functions.) To extend the obvious integral functional on step functions.

Recall that Borel sets of $\mathbb{R}$ is the smallest family of subsets of $\mathbb{R}$ that is closed under complements and countable union, and contains the open intervals.

Theorem 1.1. *Let $f_n \geq 0$ be an increasing sequence of measurable functions and suppose $f_n \to f$ pointwise. Then*

A concise proof of Riesz-Fisher on L^1.

2 Geometry of Hilbert Space

2.1 Riesz lemma

A complex inner product space is a complex vector space V, together with a function $(\cdot, \cdot) : V \times V \to \mathbb{C}$ such that

1. $(x, x) \geq 0$ and $(x, x) = 0$ if and only if $x = 0$.

2. $(x, y + z) = (x, y) + (x, z)$.

3. $(x, \alpha y) = \alpha(x, y)$.

4. $(y, x) = \overline{(x, y)}$.

A consequence of these definitions is that

$$(\alpha x, y) = \overline{(y, \alpha x)} = \overline{\alpha(y, x)} = \overline{\alpha}(x, y)$$

We say $(\cdot, \cdot)$ is conjugate-linear in the first entry.

Theorem 2.1. Pythagorean theorem *Let $\{x_i\}_{i=1}^{N}$ be an orthonormal set in an inner product space V. Then for any $x \in V$,*

$$\|x\|^2 = \sum_{i=1}^{N} |(x, x_i)|^2 + \|x - \sum_{i=1}^{N}(x, x_i)x_i\|^2$$

Proof. We may write

$$x = \sum_{i=1}^{N}(x, x_i)x_i + (x - \sum_{i=1}^{N}(x, x_i)x_i)$$

We readily verify that

$$(\sum_{i=1}^{N}(x, x_i)x_i), \ (x - \sum_{i=1}^{N}(x, x_i)x_i)$$

are orthogonal. Therefore

$$(x, x) = \|\sum_{i=1}^{N}(x, x_i)x_i\|^2 + \|x - \sum_{i=1}^{N}(x, x_i)x_i\|^2$$

2

We can easily verify using induction that

$$\| \sum_{i=1}^{N} (x, x_i) x_i \|^2 = \sum_{i=1}^{N} |(x, x_i)|^2$$

$\square$

Of course, Pythagorean theorem is well known. Some of its corollaries are less obvious at first sight, especially in an infinite dimensional space.

Corollary 2.2. (Bessel's inequality) *Same setting as before,*

$$\|x\|^2 \geq \sum_{i=1}^{N} |(x, x_i)|^2$$

Corollary 2.3. (Schwarz inequality) *For any x, y in an inner product space V,*
$$|(x, y)| \leq \|x\| \|y\|$$

Another useful geometric inequality is parallelogram law:

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

If we draw a parallelogram, then the law is saying that the sum of squares of the diagonals equal to that of the four edges.

Definition 2.4. A *Hilbert space* is a complete inner product space.

Examples $\mathcal{L}^2[a, b]$ the square-integrable functions with inner product $(f, g) = \int \overline{f} g$. ℓ^2 the square summable sequences with inner product $((a_i), (b_i)) = \sum \overline{a}_i b_i$. A *closed* subspace of a Hilbert space is a Hilbert space with the natural inner product.

Lemma 2.5. *Let $\mathcal{H}$ be a Hilbert space, M be a closed subspace. Then for each $x \in \mathcal{H}$, there exists a unique $z \in M$, such that $\|x - z\|$ is smallest.*

Proof. Set
$$d = \inf_{y \in M} \|x - y\|$$

We want to find a z such that $\|x - z\| = d$. Take a sequence $\{y_n\}$ in M such that $\|x - y_n\| \to d$ as $n \to \infty$. Now we show that the sequence is Cauchy:

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

The above equation is given by parallelogram law. Note that

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

Therefore

$$\|y_n - y_m\|^2 \leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4d^2$$

RHS clearly converges to zero as $n, m \to \infty$. Therefore $\{y_n\}$ converges and we can take z to be this limit. $\qquad\square$

Remark 2.6. The above lemma is rather intuitive if the vector space is finite dimensional. It reassures us that the minimum distance of a subspace to a point is actually achieved, except we need to make sure the subspace is closed. Of course, in the finite dimensional case, any subspace is closed. But for infinite dimensions it is possible, say, to have a dense proper subspace. For example, consider $C[a, b] \subseteq \mathcal{L}^1[a, b]$.

Now denote the set of bounded linear transformations from $\mathscr{H}$ to $\mathscr{H}'$ by $\mathscr{L}(\mathscr{H}, \mathscr{H}')$. It is a normed linear space with the usual supreme norm:

$$\|T\| = \sup_{\|x\|_{\mathscr{H}}=1} \|Tx\|_{\mathscr{H}'}$$

Consider the special case when $\mathscr{H}' = \mathbb{C}$. $\mathscr{L}(\mathscr{H}, \mathbb{C})$ is called the *dual space* of $\mathscr{H}$ and is denoted $\mathscr{H}^*$. The elements of $\mathscr{H}^*$ are called continuous linear functionals.

For each $x \in \mathscr{H}$ we associate $T_x \in \mathscr{H}'$ given by $T_x(y) = (x, y)$. T_x is bounded and in fact $\|T_x\| = \|x\|$ since $(x, y) \leq \|x\|\|y\|$ is always true and $(x, x) = \|x\|\|x\|$. We will show that these are the only elements that we have. In other words, *Hilbert spaces are self-dual*.

Lemma 2.7. (Riesz lemma) *For each $T \in \mathscr{H}^*$, there is a unique element $y_T \in \mathscr{H}$ such that $T(x) = (y_T, x)$ for all $x \in \mathscr{H}$. In addition, $\|T\| = \|y_T\|$.*

Proof. Let $N = \ker T$. It is a closed subspace since T is continuous and $N = T^{-1}(0)$. Therefore $\mathscr{H} = N \oplus N^{\perp}$. Find a nonzero $x_0 \in N^{\perp}$, if $N \neq \mathscr{H}$. y_T should satisfy

$$(y_T, x_0) = T(x_0)$$
$$(y_T, x) = 0, \forall x \in N$$

Set

$$y_T = \frac{\overline{T(x_0)}}{\|x_0\|^2} x_0$$

4

This element clearly satisfy the previous two conditions. This suffices, since for each y, we can write

$$y = (y - \frac{T(y)}{T(x_0)}x_0) + \frac{T(y)}{T(x_0)}x_0$$

Uniqueness: The association $\mathscr{H} \to \mathscr{H}^*$ is injective. $\qquad\square$

A sesquilinear form $B(\cdot, \cdot) : \mathscr{H} \times \mathscr{H} \to \mathbb{C}$ is linear in the second component and conjugate linear in the first, and is bounded in the sense that there is some C such that $|B(x, y)| \leq C\|x\|\|y\|$.

Corollary 2.8. *For each sesquilinear form $B(\cdot, \cdot) : \mathscr{H} \times \mathscr{H} \to \mathbb{C}$, there exists a unique bounded linear transformation A such that $B(x, y) = (Ax, y)$.*

Proof. For each x, then $T_x(y) = B(x, y)$ is a bounded linear functional. Find $\widetilde{x}$ such that $T_x = (\widetilde{x}, \cdot)$. Define A by $Ax = \widetilde{x}$. $\qquad\square$

Now we want to extend the idea of basis set for Hilbert spaces. If S is orthonormal such that no other orthonormal set contains S as a proper subset, then we call it an orthonormal basis for $\mathscr{H}$.

By Zorn's lemma, every Hilbert space $\mathscr{H}$ has an orthonormal basis.

Theorem 2.9. *If $S = \{x_n\}$ is a basis, then for all $y \in \mathscr{H}$,*

$$y = \sum_{x_n \in S}(x_n, y)x_n$$

and $\|y\|^2 \geq \sum |(x_n, y)|^2$. Conversely, if $\sum |c_n|^2 < \infty$, $c_n \in \mathbb{C}$, then $\sum c_n x_n$ converges in $\mathscr{H}$.

We will see soon that there are only countably many x_n such that $(x_n, y) \neq 0$. Therefore the sum in the theorem makes sense.

If $(x_n, y) \neq 0$ for a countable subset of S, then let $y_n = \sum_{j=1}^n (x_j, y)x_j$. We see that

$$\|y_m - y_n\| = \sum_{j=n}^m |(x_j, y)|^2$$

converges to 0. Hence y_n is a Cauchy sequence. (finish the proof)

Gram-Schmidt orthogonalization: if the set of vectors is countable we have a constructive proof.

2.2   Orthonormal basis

Last time
$$u \sum_{v_i \in S} (v_i, u) v_i$$

A natural question is to ask how many of such v_i we need.

Definition 2.10. A metric space which has a countable dense subset is said to be *separable*.

Theorem 2.11. $\mathscr{H}$ *is separable if and only if it has a countable orthonormal basis.*

If S contains a finite number of elements, then the Hilbert space is isomorphic as a vector space to $\mathbb{C}^N$, $N = |S|$. If it has countably infinite elements, then $\mathscr{H} \simeq \ell^2$.

Proof. Let $S = \{v_i\}$ be a basis. For each $u \in \mathscr{H}$, we expand it to
$$u = \sum_i (v_i, u) v_i$$

Let $X = \{\sum_j c_j v_j : v_j \in \mathbb{Q}[i]\}$. Then X is dense and countable in $\mathscr{H}$. Conversely, suppose $X \subseteq \{u_i\}$ is dense countable set. X has to span $\mathscr{H}$. Otherwise suppose there is some nonzero $w \perp \operatorname{Span} X$. We can find a sequence $u_i \to w$. Then $(w, u_i) \to (w, w)$. But each $(w, u_i) = 0$, so $(w, w) = 0$, contradiction. We apply Gram-Schmidt to X to obtain an orthonormal subset, which must be a basis. $\qquad\square$

A basis $\{v_i\}$ gives an isomorphism $\mathscr{H} \simeq \ell^2$. Define $T : \mathscr{H} \to \ell^2$ by $T(u) = \{(v_i, u)\}$. Bessel's inequality says that $\{(v_i, u)\} \in \ell^2$.

Let $\mathscr{H} = L^2[0, 1]$ with norm defined by
$$(f, g) = \int_0^1 \overline{f(x)} g(x) dx$$

It is easily checked that $\{v_n = e^{2\pi nix}\}$ is an orthonormal basis.

It is not the case to study separable Hilbert spaces it suffices to study ℓ^2 space since the basis itself has additional structures.

3   Banach Spaces

3.1   Hahn-Banach theorem

Recall that a complete complex normed linear space is called a *Banach* space.

Example $\mathcal{L}^\infty(\mathbb{R}), C_b(\mathbb{R})$

Non-example $\kappa(\mathbb{R}) = \{continuous\ functions\ with\ compact\ support\}$, but its completion (is it true?) $C_\infty(\mathbb{R}) = \{continuous\ functions\ that\ vanish\ at\ infinity\}$ is a Banach space.

Theorem 3.1. $\mathscr{L}(X,Y)$ *is a Banach space with operator norm if* Y *is complete.*

Proof. If A, B are two operators, then

$$\|(A + B)x\| = \|Ax + Bx\| \le \|Ax\|_Y + \|Bx\|_Y \le \|A\|\|x\| + \|B\|\|x\|$$

Now let $\{A_n\}$ be a Cauchy sequence, i.e. $\|A_n - A_m\| \to 0$. Let $x \in X$ be a point. Consider the sequence $\{A_n(x)\}$. $\|A_n x - A_m x\| \le \|A_n - A_m\|\|x\| \to 0$, so it is a Cauchy sequence. Since Y is complete, for every x we can define $A = \lim A_n x$. We want to show that $A \in \mathscr{L}(X,Y)$. Linearity of A is clear. We only need to verify that A is bounded.

$$\|Ax\|_Y = \lim_{n\to\infty} \|Ax\|_Y \le \limsup_{n\to\infty} \|A_n\|\|x\|$$

$$\frac{\|(A - A_n)(x)\|}{\|x\|} \le \frac{\|\lim_{m\to\infty} A_m x - A_n x\|}{\|x\|_X} \le \lim_{m\to\infty} \|A_m - A_n\| = 0$$

$\square$

Definition 3.2. A bounded linear operator from a normed linear space X to another normed linear space Y is called an *isomorphism* if it is a bijection which is continuous and has a continuous inverse. If it also norm preserving, then it is called an isometric isomorphism.

We can define different norms to the same space. Consider $\mathbb{C}^N$, we can have

$$\|(x_1, \cdots, x_n)\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$$

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

$$\|(x_1, \cdots, x_n)\|_\infty = \max_{1 \le i \le n} |x_i|$$

In principle different norms should give different Banach spaces, but we want to identify norms that induce the same metric topology on the underlying space. In orhter words, two norms are equivalent if and only if the identity map on the underlying space is an isomorphism. All norms on finite dimensional space $\mathbb{R}^n, \mathbb{C}^n$ are equivalent.

Recall that $\mathscr{L}(X, \mathbb{C})$ is Banach since $\mathbb{C}$ is complete. We denoted it by X^*. In fact, we call elements of X^* bounded linear functionals.

Example $(\mathcal{L}^2(\mathbb{R}))^* = \mathcal{L}^2(\mathbb{R})$. In general $\mathscr{H}^* = \mathscr{H}$.

Example Let $p, q > 1$ be conjugates, i.e. $1/p + 1/q = 1$. By Riesz representation theorem, we see that $(\mathcal{L}^p(\mathbb{R}))^* = \mathcal{L}^q(\mathbb{R})$. Apply this twice we see that $(\mathcal{L}^p(\mathbb{R}))^{**} = \mathcal{L}^p(\mathbb{R})$. In general, if $X^{**} = X$, then X is called *reflexive*.

Non-example $(\mathcal{L}^1(\mathbb{R}))^* = \mathcal{L}^\infty(\mathbb{R})$, but $\mathcal{L}^1(\mathbb{R}) \subset \mathcal{L}^\infty(\mathbb{R})^{**}$.

Examples of sequence spaces[1] $c_0 = \{\{a_n\} : a_n \to 0\}$, $\ell^1 = \{\{a_n\} : \sum |a_n| < \infty\}$, $\ell^\infty = \{\{a_n\} : \sup a_n < \infty\}$. We also have $\ell^1 \subset C_0 \subset \ell^\infty$. $(c_0)^* = l^1$ and $(\ell^1)^* = l^\infty$. Hence $(c_0)^{**} = \ell^\infty \neq c_0$, so c_0 is not reflexive.

We show that $(\ell^1)^* = \ell^\infty$. For all $\lambda \in \ell^\infty$, we define

$$T_\lambda a = \sum_{n=1}^\infty \lambda_n a_n$$

We can easily check that the sum converges. Conversely, given $T \in \mathscr{L}(\ell^1, \mathbb{C})$, we need to find $\lambda \in l^\infty$, such that $T_\lambda = T$. Set $\lambda_k = T(e_k)$, we can recover the sequence λ. By construction $T_\lambda = T$, and $\|\lambda\|_\infty \leq \|T\|$.

Theorem 3.3. *Let X be a Banach space. For each $x \in X$, we can define $\widetilde{x}(\cdot)$ a linear functional on X^* by $\widetilde{x}(\lambda) = \lambda(x), \forall \lambda \in X^*$. Then the map $J : x \mapsto \widetilde{x}$ is an isometric isomorphism of X onto a (possibly proper) subspace of X^{**}.*

Proof. $\|\widetilde{x}(\lambda)\| = \|\lambda(x)\| \leq \|\lambda\|_{X^*} \|x\|$ Therefore $\|\widetilde{x}\|_{\mathscr{L}(X^*, \mathbb{C})} \leq \|x\|$ and $J : X \to X^{**}$ is continuous. To show that J is isometric, we want to find $\lambda \in X^*$ such that $\|\widetilde{x}(\lambda)\| = \|\lambda\| \|x\|$. $\qquad \square$

Theorem 3.4. (Hahn-Banach Theorem) *Let X be a real vector space. $P : X \to \mathbb{R}$ is a real valued function satisfying $P(\alpha x + (1 - \alpha)y) \leq \alpha P(x) + (1 - \alpha)P(y)$, $|\alpha| \leq 1$. Suppose λ is a linear funcitonal defined on a subspace $Y \subseteq X$, which satisfies $\lambda(x) \leq P(x)$. Then there exists a linear functional Λ defined on X such that $\Lambda(x) \leq P(x)$ for all $x \in X$ and $\Lambda(x) = \lambda(x)$ for all $x \in Y$.*

Proof. First consider extension by one vector $z \in X - Y$. Let $\widetilde{Y} = \text{span}\{Y, z\}$. We want to extend λ to $\widetilde{\lambda} : Y \to \mathbb{R}$. By linearity

$$\widetilde{\lambda}(az + y) = a\widetilde{\lambda}(z) + \lambda(y)$$

so it suffices to specify $\lambda(y)$. Additionally, we need it $\leq P(az + y)$, i.e.

$$\lambda(z) \leq \frac{1}{a}(P(az + y) - \lambda(y))$$

---

[1]Lecture 5

when $a > 0$, and the sign is reversed when $a < 0$. Therefore it suffices to choose $\lambda(z)$ such that

$$\sup_{y \in Y, \alpha > 0} \frac{1}{\alpha}(-P(y - az) + \lambda(y)) \leq \lambda(z) \leq \inf_{y \in Y, \beta > 0} \frac{1}{\beta}(P(az + y) - \lambda(y))$$

To show this is possible, it suffices to show that for aribitrary $y_1, y_2 \in Y, \alpha, \beta > 0$,

$$\frac{1}{\alpha}(-P(y_1 - az) + \lambda(y)) \leq \frac{1}{\beta}(P(az + y_2) - \lambda(y))$$

i.e.

$$\beta\lambda(y_1) + \alpha\lambda(y_2) \leq \beta P(y_1 - \alpha y_2) + \alpha P(y_2 + \beta z)$$

This is true since

$$\beta P(y_1 - \alpha z) + \alpha P(y_2 + \beta z) \geq (\alpha + \beta)(\frac{\beta}{\alpha + \beta}P(y_1 - \alpha z) + \frac{\alpha}{\alpha + \beta}P(y_2 + \beta z))$$

$$\geq (\alpha + \beta)(P(\frac{\beta}{\alpha + \beta}(y_1 - \alpha z) + \frac{\alpha}{\alpha + \beta}(y_2 + \beta z))$$

$$= (\alpha + \beta)(P(\frac{\beta}{\alpha + \beta}y_1 + \frac{\alpha}{\alpha + \beta}y_2)$$

$$\geq (\alpha + \beta)(\lambda(\frac{\beta}{\alpha + \beta}y_1 + \frac{\alpha}{\alpha + \beta}y_2)$$

$$\geq \lambda(\beta y_1 + \alpha y_2)$$

Apply Zorn's lemma to complete the proof. $\square$

In the complex version, we replace the assumption on P by

$$P(\alpha x + \beta y) \leq |\alpha|P(x) + |\beta|P(y), \forall x, y \in X, |\alpha| + |\beta| = 1$$

$\lambda(x) \leq P(x)$ is replaced by $|\lambda(x)| \leq P(x)$.

Proof. Note that λ may be complex valued, but we wish to apply the real version. Let $l(x) = \operatorname{Re}\lambda(x)$. $l(x)$ extends to $L(x)$. $\operatorname{Im}\lambda(x) = -\operatorname{Re}\lambda(ix)$. We have $\lambda(x) = l(x) - il(ix)$, so we define $\Lambda(x) = L(x) - iL(ix)$. We can check that Λ is the desired extension. $\square$

Corollary 3.5. *Let X be a normed linear space and Y a subspace. $\lambda \in Y^*$. Then there exists $\Lambda \in X^*$ such that $\Lambda(x) = \lambda(x), \forall x \in Y$ and $\|\Lambda\|_{X^*} = \|\lambda\|_{Y^*}$.*

As an example, since c_0 is a subspace of ℓ^∞, we have that $\ell^1 = c_0^* \subseteq \ell^\infty$.

Let $p(x) = \|\lambda\|_{Y^*}\|x\|, \forall x \in X$. Hahn-Banach says that we have an extension Λ, such that $|\Lambda(x)| \leq p(x)$. Therefore we have inequality $\|\Lambda\|_{X^*} = \|\lambda\|_{Y^*}$.

If y is an element of a normed linear space X, then there is a bounded linear functional $\Lambda \in X^*$ such that $\Lambda(y) = \|\Lambda\|_{X^*}\|y\|$. That is, any vector can be the element where Λ achieves its supreme norm.

Hahn-Banach theorem roughly says that linear functionals serve as "coordinates" of infinite dimensional linear space.

Let Z be a subspace of X. Suppose there is a vector $y \in X - Z$ whose distance from subspace Z is d. Then there exists $\Lambda \in X^*$ such that $\|\Lambda\| \leq 1$ and $\Lambda(y) = d$, $Z \subseteq \ker \Lambda$. That is, we can find a linear functional that separates a subspace and a vector. For finite dimensional vector spaces, this of course is trivial, but Hahn-Banach handles the infinite dimensional case.

Theorem 3.6. *Let X be a Banach space. If X^* is separable, then X is also separable.*

This gives another proof $\ell^1 \subset (\ell^\infty)^*$ since ℓ^1 is separable.

Proof. Let $\{\lambda_n\}$ be a dense subseteq of X^*. Choose $\{x_n\} \subset X$, such that $\|x_n\| = 1$ and

$$|\lambda_n(x_n)| \geq \frac{\|\lambda_n\|}{2}$$

Let D be the set of all finte linear combinations of $\{x_n\}$ with rational coefficients. D is countable by construction. We want to check that it is dense. Suppose not. Then we can find a vector $y \in X - D$ and $\lambda \in X^*$ such that $\lambda(y) \neq 0$, $\lambda(x) = 0, \forall x \in D$. However, this is impossible. Since $\{\lambda_n\}$ is dense in X^*, we can find a subsequence convergent to λ. $\qquad\square$

3.2 Consequences of Baire category theorem

Theorem 3.7. *A complete vector space is never a countable union of nowhere dense subsets.*

Usually the Baire category theorem is used as follows: if we can write a complete vector space X as a union of closed subsets $\bigcup S_n$, then at least one of them has nonempty interior.

(Principle of uniform boundedness)

Theorem 3.8. *Let X be a Banach space and suppose we have a family $\mathscr{F}$ of bounded linear transformations $X \to Y$ for some normed linear space Y. Suppose for each $x \in X$, $\{\|Tx\|_Y : T \in \mathscr{F}\}$ is bounded, then $\{\|T\|_{\mathscr{L}(X,Y)} : T \in \mathscr{F}\}$ is also bounded.*

So how is Baire category theorem related to boundedness?

Lemma 3.9. *Let X, Y be two normed linear spaces and T be a linear map $X \to Y$. T is bounded if and only if $T^{-1}(\{y \in Y : \|y\| \leq 1\})$ has a nonempty interior.*

Proof. One way is easy: If T is bounded, say by M, then $\{x \in X : \|x\| \leq 1/M\} \subseteq T^{-1}(\{y \in Y : \|y\| \leq 1\})$.

Now assume that $T^{-1}(\{y \in Y : \|y\| \leq 1\})$ has a nonempty interior, so we assume it contains a small ball $\{x \in X : \|x - x_0\| < \varepsilon\}$. Then for all x with $\|x'\| < \varepsilon$

$$\|Tx'\| = \|T(x' + x_0) - T(x_0)\| < \|T(x' + x_0)\| + \|Tx_0\| < 1 + \|Tx_0\|_Y$$

This implies that the operator norm of T is smaller than

$$\frac{1 + \|T(x)\|}{\varepsilon}$$

$\square$

Lecture 8

1. Principle of uniform boundedness

2. Open mapping theorem

3. Inverse mapping theorem

4. Closed graph theorem

Theorem 3.10. *Let X, Y be Banach spaces and $T : X \to Y$ be a linear map. T is bounded if and only if $\Gamma(T) \subseteq X \oplus Y$ is closed.*

That T is continuous means if $x_n \to x \in X$ as $n \to \infty$, $\{T(x_n)\} \to y = T(x) \in Y$. That $\Gamma(T)$ is closed means that if $\langle x_n, T(x_n)\rangle \to \langle x, y\rangle$, then $y = T(x)$. One way is trivial: if T is continuous, then its graph is clearly closed. To see the converse, we first note that $\Gamma(T)$ is a subspace of $X \oplus Y$. If it is closed, it is a Banach space. (We use Taxi-cab norm on $X \oplus Y$. Consider two maps π_1, π_2 that project $\Gamma(T)$ to X, Y respectively. π_1 is bijective since $\Gamma(T)$ is a graph. By inverse mapping theorem, π_1 has a continuous inverse. (Open mapping theorem only holds for Banach spaces, so we need the assumption that $\Gamma(T)$ is closed.) Note that $T = \pi_2 \circ \pi_1^{-1}$, so T is continuous.

We have seen the convergence of the operator norm. $T_n \to T$ if and only if $\|T_n - T\| \to 0$. A topology gives us a notion of convergence, and we call this topology "norm topology".

Given a sequence of operators, we say that the sequence $\{T_n\}$ converges *strongly* to an operator T denoted by $T_n \xrightarrow{s} T$ if for each $x \in X$, $\|T_n x - Tx\| \to 0$. This is analogous to the concept of pointwise convergence. However, since T_n's are linear, this condition is in fact quite strong. We say $\{T_n\}$ converges weakly to T, denoted by $T_n \xrightarrow{s} T$, if $|L(T_n x) - L(Tx)| \to 0$ as $n \to \infty$, for each $x \in X, L \in Y^*$. $T_n \to Y$ is stronger than $T_n \xrightarrow{s} T$, which is stronger than $T_n \xrightarrow{w} T$.

Example   Consider $\mathscr{L}(\ell^2, \ell^2) = \mathscr{L}(\ell^2)$. Let S_n by

$$(\xi_1, \xi_2, \cdots) \mapsto (0, 0, 0, \cdots, 0, \xi_{n+1}, \xi_{n+2}, \cdots)$$

For each n, $\|S_n\| = 1$, so S_n does not converge to 0. However, $S_n \xrightarrow{s} 0$. $S_n \xi \to 0$ since $\xi \in \ell^2$.

Example   Let W_n is the "shifting operator" given by

$$(\xi_1, \xi_2, \cdots) \mapsto (0, \cdots, 0, \xi_1, \xi_2, \cdots)$$

(n zeroes are appended in the front.) $W_n \xrightarrow{w} 0$. For fixed $l \in (\ell)^*$, $\xi \in \ell^2$, $|l(W_n \xi) - l(0 \cdot \xi)| \to 0$.

Theorem 3.11. *Let $\{T_n\} \subseteq \mathscr{L}(\mathscr{H})$. Suppose that $(T_n x, y)$ converges as $n \to \infty$ for all $x, y \in \mathscr{H}$. Then there exists $T \in \mathscr{L}(\mathscr{H})$ such that $T_n \xrightarrow{w} T$.*

Step 1   Prove $\forall x \in \mathscr{H}$, $\|Tx\|$ is a uniformly bounded. For each $y \in \mathscr{H}$, $(T_n x, y)$ converges, so that

$$\sup_n |(T_n x, y)| < \infty$$

We can show $(T_n x, y) \in \mathscr{H}^*$. This implies the operator norm is uniformly bounded.

Step 2 $\|T_n\|_{\mathscr{L}(H)}$ is uniformly bounded.

Step 3   Define $B(x, y) = \lim_{n \to infty}(T_n x, y)$. B is sesquilinear and bounded. $|B(x, y)| = |\lim_{n \to \infty}(T_n x, y)| \leq \sup \|T_n\| \|x\| \|y\|$. By Riesz lemma, there exists $T \in \mathscr{L}(\mathscr{H})$, $B(x, y) = (Tx, y)$.

If T is linear, then its kernel and range are both vector spaces. The kernel is always closed, but the range may not be.

Let X, Y be Banach spaces. $T \in \mathscr{L}(X, Y)$. The Banach space adjoint of T, denoted by T', is a bounded linear operator from Y^* to X^* defined by $(T'l)(x) = l(Tx)$.

4 Spectral Theory

Definition 4.1. A complex number λ is said to be in the resolvent set $\rho(T)$ if $\lambda I - T$ is a bijection with a bounded inverse. $\sigma(T) = \mathbb{C} - \rho(T)$.

$\rho(T)$ is open. The resolvent R_λ is analytic on $\rho(T)$ if $\lambda \geq \|T\|$, then $\lambda \in \rho(T)$.

Definition 4.2. Let $r(T) = \sup_{\lambda \in \sigma(T)} |\lambda|$. $r(T)$ is called the spectral radius of T. $r(T) \leq \|T\|$.

Theorem 4.3. *Let $T \in \mathscr{L}(X)$. Then*

$$\lim_{n \to \infty} \sqrt[n]{\|T\|^n}$$

exists and equals $r(T)$. If X is a Hilbert space, and A is self-adjoint, then $r(A) = \|A\|$.

Proof. To compute $(\lambda I - T)^{-1}$, we can use Laurent series

$$\frac{1}{\lambda} \sum_{n=0}^{\infty} \frac{1}{\lambda^n} T^n$$

We want to know when the series converges. We know that $\|T^n\| \leq \|T\|^n$. Hadamard theorem says that $r(T) = \sqrt[n]{\|T\|^n}$.

If A is self-adjoint on a Hilbert space, $\|A^*A\| = \|A\|\|A^*\| = \|A\|^2$. If we take $n = 2^k$, then we see that $\|A^n\| = \|A\|^n$ by replacing A by A^*'s. Since the whole sequence converges, it has to converge to A. $\qquad\square$

Example Let T be the left-shift operator on ℓ_1. We consider T' on ℓ_∞. We have see that $\sigma(T) = \sigma(T') = \{\lambda : |\lambda| \leq 1\}$. $\{\lambda : |\lambda| < 1\}$ is point spectrum of T and residual spectrum of T'.

Theorem 4.4. *Let X be a Banach space, $T \in \mathscr{L}(X)$. Then $\sigma(T) = \sigma(T')$ and $R_\lambda(T') = R_\lambda(T)'$. If $\mathscr{H}$ is a Hilbert space, then $\sigma(T^*) = \{\lambda : \overline{\lambda} \in \sigma(T)\}$.*

Proof. Recall that T^* is characterized by $(T^*x, y) = (x, Ty)$. Inner product is conjugate linear in the first and linear in the second, so it makes sense that $\sigma(T^*) = \{\lambda : \overline{\lambda} \in \sigma(T)\}$.

Suppose $\lambda \in \rho(T)$, $R_\lambda(T) \in \mathscr{L}(X)$. $R_\lambda(T)' \in \mathscr{L}(X^*)$ acts on X^* by

$$R_\lambda(T)'(l)(x) = l(R_\lambda(T)(x)) = l((\lambda I - T)^{-1}x)$$

This is simply $(\lambda I - T')^{-1}$ acting on l. $\qquad\square$

Let us think about $|\lambda| = 1$ in the previous example. If there would by a nonzero element, $l \in \ell_\infty = (\ell_1)^*$, such that

$$l(\lambda I - T)(x) = 0, \forall x$$

If the image of $\lambda I - T$ is dense, then the above is not possible. Otherwise, Hahn-Banach theorem says we can find such l. Then

$$(\lambda I - T')(l)(x) = 0$$

which implies that $(\lambda I - T')(l) = 0$. l is an eigenvector and λ is an eigenvalue, since $T'l = \lambda l$. We easily see that T' does not have a point spectrum. Therefore $\mathrm{Ran}\,(\lambda I - T)$ is dense and λ is not the residual spectrum. This gives us an example of $\lambda \in \sigma(T)$ but lies in neither point spectrum nor residual spectrum.

If λ lies in the residual spectrum of T, then λ lies in the point spectrum of T'.

We now show that $\{\lambda : |\lambda| = 1\}$ is the residual spectrum of T'. If $a, b \in \ell_\infty$, $a = (\lambda I - T')b = \lambda b - T'b$ and $a_0 = \lambda b_0, a_1 = \lambda b_1 - b_0, \cdots, a_n = \lambda b_n - b_{n-1}$. We can easily solve for b from a:

$$b_n = (\bar{\lambda})^{n+1} \sum_{m=0}^{n} \lambda^m a_m$$

Let $c = (c_1, c_2, \cdots)$ where $c_n = \bar{\lambda}^n$. Suppose $d \in \ell_\infty$ and

$$\|d - c\|_{\ell_\infty} \leq \frac{1}{2}$$

We get $\Re\{\lambda^n d_n\} \geq \Re\{\lambda^n c_n\} = \|d - c\|_{\ell_\infty} \geq 1/2$. Thus if $(\lambda I - T')e = d$, $e_n = \bar{\lambda}^{n+1} \sum_{m=0}^{n} \lambda^m d_m$, $e_n \geq n/2$ then $e_n \notin \ell_\infty$. $\mathrm{Ran}\,(\lambda I - T')$ does not contain the ball of radius of radius $1/2$ at c and hence λ is in the residual spectrum of T'. Altough T is injective, its inverse is not bounded.

Theorem 4.5. *Let T be a self-adjoint operator on a Hilbert space $\mathscr{H}$. Then T has no residual spectrum, and $\sigma(T) \subseteq R$. Eigenvectors of distinct eigenvalues are orthogonal.*

Proof. Consider $T - (\lambda + i\mu)I$. Expand the inner product

$$\|(T - (\lambda + i\mu))x\|^2 \leq \langle (T - (\lambda + i\mu))x, (T - (\lambda + i\mu))x \rangle = \|(T - \lambda)x\|^2 + \|\mu x\|^2$$

Note that

$$\langle -i\mu x, (T - \lambda)x \rangle = i\mu \langle x, (T - \lambda)x \rangle$$

and

$$\langle (T - \lambda)x, -i\mu x \rangle = -i\langle (T - \lambda)x, x \rangle$$

so they cancell each other. We have shown that $(\lambda + i\mu) \in \rho(T)$ if $\mu \neq 0$. If $\lambda \in \mathbb{R}$ in the residual spectrum of T, then λ lies in the point spectrum of $T = T^*$, a contradiction. Therefore λ does not have residual spectrum. $\qquad\square$

Lecture hat

We can diagonalize a Hermitian $n \times n$ matrix A using eigenvalues and eigenvectors $A\phi_i = \lambda\phi_i$. Let $\Phi = (\phi_1, \cdots, \phi_n)$, then $A\Phi = \Phi\Lambda$, or equivalently, $\Phi^* A\Phi = \Lambda$ where $\Lambda = \text{diag}[\lambda_1, \cdots, \lambda_n]$.

We are going to study bounded self-adjoint operators, which are analogues of Hermitian matrices. We are going to estabilish that bounded self-adjoint operators look at "diagonal operators".

Let A be a bounded self-adjoint operator on a Hilbert space $\mathscr{H}$. We can always find a measure space (M, μ) and a unitary operator $U : \mathscr{H} \to L^2(M, d\mu)$ such that $(UAU^{-1}f)(x) = F(x)f(x)$ where $F(x)$ is some bounded measurable function on (M, μ). (Since F is bounded, it is clearly in L^2.)

Let f be a continuous function. We want to define $f(A)$ (f acting on operators). If $f = \sum_{n=0}^{N} a_n x^n$, then we know how to do it. If f is a power series with radius of convergence R, then we can similarly define

$$f(A) = \sum_{n=0}^{\infty} a_n A^n$$

We can using dominated convergence to see that the sum converges.

(Continuous functional calculus)

Theorem 4.6. *Let A be a bounded self-adjoint operator on a Hilbert space $\mathscr{H}$. There is a unique map $\phi : C(\sigma(A)) \to \mathscr{L}(\mathscr{H})$, where $C(\sigma(A))$ is the set of continuous functions on A, satisfying the following properties:*

1. *ϕ is an algebraic $*$-homomorphism. It means it preserves the polynomial relations between functions. E.g. $\phi(fg) = \phi(f) \circ \phi(g)$, $\phi(\lambda f) = \lambda\phi(f)$, and $\phi(\overline{f}) = (\phi(f))^*$.*

2. *ϕ is continuous. $\|\phi(f)\|_{\mathscr{L}(\mathscr{H})} \leq C\|f\|_{C(\sigma(A))}$ for some constant C.*

3. *$\phi(x) = A$. (These three properties already uniquely determine the function ϕ)*

4. *If $A\psi = \lambda\psi, \phi(f)\psi = f(\lambda)\psi$.*

5. *$\sigma[\phi(f)] = \{f(\lambda) : \lambda \in \sigma(A)\}$.*

6. *If $f \geq 0$, then $\phi(f) \geq 0$. (Recall what it means for an operator to be ≥ 0.)*

7. *$\|\phi(f)\|_{\mathscr{L}(\mathscr{H})} = \|f\|_\infty$.*

Suppose A is a Hermitian matrix and $A = \Phi \Lambda \Phi^*$. Then $f(\Lambda) = \mathrm{diag}[f(\lambda_1), \cdots, f(\lambda_n)]$. Therefore the natural definition of $f(A)$ is $f(A) = \Phi f(\Lambda) \Phi^*$. If we assume $A \geq 0$ so that all its eigenvalues ≥ 0, then we can canonically define $A = \Phi \sqrt{\Lambda} \Phi^*$.

We begin with two lemmas.

Lemma 4.7. *Let*

$$P(x) = \sum_{n=0}^{N} a_n x^n, \quad P(A) = \sum_{n=0}^{N} a_n A^n$$

. Then

$$\sigma(P(A)) - \{P(\lambda) : \lambda \in \sigma(A)\}$$

Proof. Let $\lambda \in \sigma(A)$. Let us look at the polynomial $P(x) - P(\lambda)$. It certainly contains λ as a root, so $P(x) - P(\lambda) = (x - \lambda)Q(x)$ where Q is some other polynomial. Now

$$P(A) - P(\lambda)I = (A - \lambda I)Q(A) = Q(A)(A - \lambda I)$$

Assume $\mu \in \sigma(P(A))$, we let $\lambda_1, \cdots, \lambda_n$ be all the roots to $P(x) - \mu$.

$$P(x) - \mu = a(x - \lambda_1) \cdots (x - \lambda_n)$$

and hence

$$P(A) - \mu = a(A - \lambda_1) \cdots (A - \lambda_n)$$

Once of $\{\lambda_i\}$ lies in $\sigma(A)$ and $\mu = P(\lambda_i)$. $\qquad \square$

Lemma 4.8. *Let A be a bounded self-adjoint operator. Then*

$$\|P(A)\| = \sup_{\lambda \in \sigma(A)} |P(\lambda)|$$

Proof. The proof uses the fact that A is self-adjoint. It is not true if A is not.

$$\|P(A)\|^2 = \|P(A)^* P(A)\| = \|\overline{P}(A)P(A)\| = \|(\overline{P}P)(A)\| = \sup_{\lambda \in \overline{P}P(\sigma(A))} |\lambda| = \sup_{\mu \in \sigma(A)} |\overline{P}P(\mu)| = (\ \text{su}$$

$$\square$$

If $T(f) = \int f d\mu$, then $T \in C(\sigma(A))^*$. Find ψ $(\psi, f(A)\psi) : C(\sigma(A)) \to \mathbb{C}$ by $\int_{\sigma(A)} f d\mu_A$.

Definition 4.9. The measure μ_ψ is called the spetral measure assoicated with the vector ψ.

Using μ_ψ $g \in \mathscr{B}(\mathbb{R})$ is some Borel function on $\mathbb{R}$. Define $g(A)$ such that $\psi g(A)\psi = \int_{\sigma(A)} g(\lambda)\mu_\psi(d\lambda)$

This defines $g(A)$ by polarization identity.

If these are true, then we have showed existence of ϕ on a dense subset. Our plan is to use BLT to extend the map.

Let A be a bounded self-adjoint linear operator on $\mathscr{H}$. Then there is a unique map $\widehat{\phi} : \mathscr{B}(\mathbb{R}) \to \mathscr{L}(\mathscr{H})$ so that

1. $\widehat{\phi}$ is algebraically $*$-homomorphism.

2. $\widehat{\phi}$ is norm continuous.

3. $\widehat{\phi}(x) = A$.

4. If we have sequence of functions $f_n \in \mathscr{B}(\mathbb{R}) \to f$ pointwise, then $\widehat{\phi}(f_n) \to \widehat{\phi}(f)$ strongly.

Moreover, if ψ is an eigenvector, then $\widehat{\phi}(f)\psi = f(\lambda)\psi$. If $f \geq 0$, then $\widehat{\phi}(f) \geq 0$. If $BA = AB$, then $\widehat{\phi}(f)B = B\widehat{\phi}(f)$.

4.1 Compact operators

Definition 4.10. Let X, Y be Banach spaces. $T \in \mathscr{L}(X, Y)$ is compact if T maps bounded sets in X to precompact sets in Y. Equivalently, T is compact if and only if for every bounded sequence $\{x_n\}$, the image $\{Tx_n\}$ has a convergent subsequence.

Example Let $K(x, y) \in C[0, 1]^2$. The operator $K \in \mathscr{L}(C[0, 1])$ defined by

$$K\phi(x) = \int_0^1 K(x, y)\phi(y)dy$$

is a compact operator. The key is to show that the image of each $\{\phi \in C[0, 1] : \|\phi\| \leq M\}$ is equicontinuous. Ascoli theorem says that the image has a convergent subsequence.

Example A finite rank operator (i.e. an operator whose range is of finite dimension) is automatically compact. A bounded operator will send a bounded set to a bounded set, but a bounded set in a finite dimensional space is automatically precompact by Heine-Borel.

Theorem 4.11. *A compact operator maps weakly convergent sequences onto norm convergent sequences.*

Proof. The intuition, I think, is that weak convergence and convergence in norm coincides in finite dimensional spaces. Compact operators are not far from finite rank operators. We need to show that a compact operator maps a weakly convergent sequence to a weakly convergent sequence. Let $x_n \to_w x$ and $Tx_n = y_n$, $Tx = y$. Let $l \in Y^*$,

$$l(y_n) - l(y) = (T'l)(x_n - x)$$

Note that $T'l \in X^*$, by definition of weak convergence, $y_n \to_w y$. Now suppose that y_n does not converge to y in norm. Then we can find a sequence y_{n_k} such that $\|y_{n_k} - y\| \geq \varepsilon$. This subsequence will contain a subsequence that converges in norm to some point other than y, so that the sequence y_n contains subsequences that converge weakly to different points, contradiction. □

Theorem 4.12. *Let $\mathcal{H}$ be a separable Hilbert space, then a compact operator T is the norm limit of a sequence of finite rank operators.*