Study Notes of Functional Analysis

Pei Zhong

January 3, 2024

Contents

I Hilbert Spaces 3

1 Elementary Properties and Examples 4
- 1.1 Elementary Properties and Examples . . . 4
- 1.2 Homework . . . 7
- 1.3 Reference . . . 10

2 Orthogonality 11
- 2.1 Homework . . . 15
- 2.2 Reference . . . 15

3 The Riesz Representation Theorem 17
- 3.1 Homework . . . 19
- 3.2 Reference . . . 20

4 Orthonormal Sets of Vectors and Bases 21
- 4.1 Reference . . . 26

5 Isomorphic Hilbert Spaces and the Fourier Transform for the Circle 27
- 5.1 Reference . . . 28

II Operators on Hilbert Spaces 29

6 Elementary Properties and Examples 30
- 6.1 Reference . . . 32

7 The Adjoint of an Operator 33

8 Projections and Idempotents; Invariant and Reducing Subspaces 37
- 8.1 Reference . . . 39

9 Compact Operators **40**
9.1 Homework 41
9.2 Reference 42

III Banach Spaces **43**

10 Elementary Properties and Examples **44**
10.1 Reference 45

11 Linear Operators on Normed Space **46**
11.1 Reference 47

12 Finite Dimensional Normed Spaces **48**

13 Quotients and Products of Normed Spaces **49**
13.1 Reference 50

14 Linear Functional **51**
14.1 Reference 51

15 The Hahn-Banach Theorem **52**
15.1 Homework 55
15.2 Reference 56

16 The Dual of a Quotient Space and a Subspace **57**

17 Reflexive Space **58**

18 The Open Mapping and Closed Graph Theorems **59**

19 The Principle of Uniform Roundedness **60**

Preface

The notes mainly refer to:

- A Course in Functional Analysis 2nd
- lecture notes from Michigan university
- lecture notes from cuhk
- Functional Analysis Problems with Solutions
- lecture notes by Mihai Nica
- Functional Analysis
- Companion to Functional Analysis
- (Functional Analysis

Homework:
I_1: T2, T6
I_2: T2, T3
I_3: T3, T5
I_4: T13, T19
I_5: T6, T9
II_1: T9, T11
II_2: T12, T15
II_3: T4, T6
II_4: T4, T8
III_1: T4, T5
III_2: T1, T4
III_3: T2
III_4: T1, T3
III_5: T2
III_6: T1, T2

Part I

Hilbert Spaces

Chapter 1

Elementary Properties and Examples

1.1 Elementary Properties and Examples

Remark. $\mathbb{F}$ will mean either $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.1

If $\mathscr{X}$ is a vector space over $\mathbb{F}$, a semi-inner product on $\mathscr{X}$ is a function $u : \mathscr{X} \times \mathscr{X} \to \mathbb{F}$ such that for all α, β in $\mathbb{F}$, and x, y, z in $\mathscr{X}$, then following are satisfied:
(a) $u(\alpha x + \beta y, z) = \alpha u(x, z) + \beta u(y, z)$,
(b) $u(x, \alpha y + \beta z) = \overline{\alpha} u(x, y) + \overline{\beta} u(x, z)$,
(c) $u(x, x) \geqslant 0$,
(d) $u(x, y) = \overline{u(y, x)}$.

Definition 1.2

If $\mathscr{X}$ is a vector space over $\mathbb{F}$. An inner product on $\mathscr{X}$ is a mapping $u : \mathscr{X} \times \mathscr{X} \to \mathbb{F}$ such that $\forall x, y, z \in X, \alpha, \beta \in F$,
(1) $u(\alpha x + \beta y, z) = \alpha u(x, z) + \beta u(y, z)$
(2) $u(x, y) = \overline{u(y, x)}$
(3) $u(x, x) \geqslant 0, < x, x >= 0$ iff $x = 0$
$\mathscr{X}$ together with such a funtion u is called an inner product space.

Remark. We always denote our innner product by $\langle x, y \rangle := u(x, y)$.
Remark. By def1.2(1)(2), we can get $u(x, \alpha y + \beta z) = \overline{\alpha} u(x, y) + \overline{\beta} u(x, z)$.

Definition 1.3

$\mathscr{X}$ is an inner product space. The mapping

$$\begin{aligned} ||\cdot|| : \mathscr{X} &\to [0, +\infty] \\ x &\mapsto \langle x, x \rangle \end{aligned}$$

defines a norm on $\mathscr{X}$. This norm is called norm induced by the inner space or simply induced norm.

Before checking the induced norm well definited, we introduce the Cauchy-Schwarz inequality.

Theorem 1.1

If $\mathscr{X}$ is an inner product space. Then

$$|\langle x, y\rangle| \leqslant ||x|| \cdot ||y||.$$

equality occurs iff $x = ky(k \neq 0)$.

Proof. write when available. □

Now let's check the three properties of induced norm.
(a) $||x|| = \sqrt{\langle x, x\rangle} \geqslant 0$; $||x|| = 0 \Leftrightarrow \langle x, x\rangle = 0 \Leftrightarrow x = 0$.
(b) $||\alpha x|| = \sqrt{< \alpha x, \alpha x >} = \sqrt{\alpha^2 < x, x >} = |\alpha|||x||$.
(c)

$$\begin{aligned} ||x+y||^2 &= < x+y, x+y > \\ &= < x, x > + < x, y > + < y, x > + < y, y > \\ &= ||x||^2 + \mathrm{Re} < x, y > + ||y||^2. \end{aligned}$$

Since $\mathrm{Re} < x, y > \leqslant | < x, y > | \leqslant ||x||||y||$, $||x+y||^2 \leqslant ||x||^2 + 2||x||||y|| + ||y||^2$. Then $||x+y||^2 \leqslant (||x|| + ||y||)^2$. Hence, $||x+y|| \leqslant ||x|| + ||y||$.

Proposition 1.1: Continuity of the Inner Product

Let $\mathscr{X}$ be an inner product space with induced norm $||\cdot||$. Then $< \cdot, \cdot >: \mathscr{X} \times \mathscr{X} \to \mathbb{F}$ is continuous.

Proposition 1.2: Parallelogram Law

Let $\mathscr{X}$ be an inner product space. Then $\forall x, y \in \mathscr{X}$

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

Proposition 1.3: Polarization Identity

Let $\mathscr{X}$ be an inner product space. Then $\forall x, y \in \mathscr{X}$

$$< x, y > = \frac{1}{4}(||x+y||^2 - ||x-y||^2 - \mathrm{i}||y+\mathrm{i}x||^2 + \mathrm{i}||y-\mathrm{i}x||^2).$$

Remark. In an inner product space, the inner product determines the norm. The polarization identity shows that the norm determines the inner product. But not every norm on a vector space X is induced by an inner product, for example $(\ell^\infty, ||\cdot||_\infty)$ is not an inner space. The following theorem shows that when a norm is induced by inner product.

Theorem 1.2

Suppose $(\mathscr{X}, ||\cdot||)$ is a normed linear space. The norm $||\cdot||$ is induced by an inner product iff the parallelogram law holds in $(\mathscr{X}, ||\cdot||)$.

Proof. ($\Rightarrow$):See the proof of prop1.2.
($\Leftarrow$): We need to show that the inner product determined by the norm is well definited. Use the polarization identity to define $<\cdot,\cdot>$. Then immediately

$$< x, x > = \frac{1}{4}(||x+x||^2) = ||x||^2 \Rightarrow < x, x > \geqslant 0, < x, x >= 0 \text{ iff } x = 0,$$
$$< x, y > = \overline{< y, x >}$$

Use parallelogram law, we can get

$$< x, y+z > =< x, y > + < x, z >,$$
$$< \alpha x, y > = \alpha < x, y >$$

The inner product satisfys three properties in def1.2 and so well definited. □

Definition 1.4: Cauchy Sequence, Convergent Sequence

Let $\mathscr{X}$ be a normed space, and let $\{x_n\}$ be a sequence of points in X.
(1) We say that $\{x_n\}$ is a Cauchy sequence if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ so that

$$i, j \geqslant N \Rightarrow ||x_i - x_j|| < \epsilon.$$

(2) We say that $\{x_n\}$ converges to a point $x \in \mathscr{X}$ if

$$\lim_{n\to\infty} ||x_n - x|| = 0.$$

Remark. In a non empty space $\mathscr{X}$, any constant sequence is a cauchy sequence, so you can always find a cauchy sequence.

Proposition 1.4

Cauchy sequence is always bounded.

Proof. Suppose $\{x_n\}$ is a Cauchy sequence. Then for $\epsilon = 1$, $\exists N$, s.t., when $n > N$: $||x_n - x_N|| < 1$. Then by triangle inequality, we have

$$||x_n|| = ||x_n - x_N + x_N|| \leqslant ||x_n - x_N|| + ||x_N|| < 1 + ||x_N||.$$

Let $M = \max\{||x_1||, ||x_2||, ..., ||x_N||, 1 + ||x_N||\}$, then $\forall n$, $||x_n|| \leqslant M$. Hence, $\{x_n\}$ is bounded. □

Proposition 1.5: Convergent Sequence are Cauchy

If $\mathscr{X}$ is a normed space, then every convergent sequene in $\mathscr{X}$ is a Cauchy sequence.

Proof. Let $\{x_n\}$ be a sequence that converges to a point $x \in X$, and let $\epsilon > 0$. Since $\lim_{n\to\infty} ||x_n - x|| = 0$, there exists an $N \in \mathbb{N}$ so that

$$n \geqslant N \Rightarrow ||x_n - x|| < \frac{\epsilon}{2}.$$

If $i, j \geqslant N$, it follows that

$$||x_i - x_j|| \leqslant ||x_i - x|| + ||x_j - x|| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

which proves that $\{x_n\}$ is a Cauchy sequence. □

Though every convergent sequence is Cauchy, it is not necessarily the case that every Cauchy sequence in a normed space converges. For example, let $\mathbb{Q}$ be the normed space of all rational numbers under the usual norm: $||q_1 - q_2|| = |q_1 - q_2|$. Then there are many Cauchy sequences in $\mathbb{Q}$ that do not converge to any point in $\mathbb{Q}$. For example, the sequence

$$3, 3.1, 3.14, 3.141, 3.1415, 3.14159, \ldots$$

is a Cauchy sequence in $\mathbb{Q}$, but it does not converge to any point in $\mathbb{Q}$.

Definition 1.5: Complete Normed Space

A normed space $\mathscr{X}$ is said to be complete if every Cauchy sequence in $\mathscr{X}$ converges to a point in $\mathscr{X}$.

Definition 1.6: Hilbert Space

An inner product space which is complete with respect to the norm induced by the inner product is called a Hilbert space.

Proposition 1.6

If $\mathscr{X}$ is a vector space and $< \cdot, \cdot >_{\mathscr{X}}$ is an inner product on $\mathscr{X}$. The completion of $\mathscr{X}$ is a Hilbert space.

Proof. You can construct the completion of $\mathscr{X}$, denoted by $\overline{X}$, by referring to lecture notes from nw. Then You can construct the inner product in $\overline{X}$ by referring to answer in mathexchage. □

1.2 Homework

Exercise 1.1: l1 T2

Let I be any set and let $l^2(A)$ denote the set of all functions $x : I \to \mathbb{F}$ such that $x(i) = 0$ for all but a countable number of i and $\sum\limits_{i\in I}|x(i)|^2<\infty$. For $x,y\in l^2(I)$ define

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

Then $l^2(I)$ is a Hilbert space.

Proof. Firstly, we show that $l^2(I)$ is a vector space. For $x,y\in l^2(I)$ and $c\in\mathbb{F}$, $(x+y)(i)=x(i)+y(i)=0$ for all but a countable number of i and by Minkowski's Inequality($||a+b||_p\leqslant||a||_p+||b||_p$), $\sum\limits_{i\in I}|(x+y)(i)|^2=\sum\limits_{i\in I}|x(i)+y(i)|^2\leqslant\sum\limits_{i\in I}|x(i)|^2+\sum\limits_{i\in I}|y(i)|^2<\infty$. Hence, $x+y\in l^2(I)$. Simlarly, cx is in $l^2(I)$ and so $l^2(I)$ is a vector space over $\mathbb{F}$.

Secondly, we show that $l^2(I)$ is a inner product space.

(1) $\langle\alpha x+\beta y,z\rangle=\sum\limits_i(\alpha x+\beta y)(i)\overline{z(i)}=\alpha\sum\limits_i x(i)\overline{z(i)}+\beta\sum\limits_i y(i)\overline{z(i)}=\alpha\langle x,z\rangle+\beta\langle y,z\rangle$.

(2) $\langle x,y\rangle=\sum\limits_i x(i)\overline{y(i)}=\overline{\sum\limits_i y(i)\overline{x(i)}}=\overline{\langle y,x\rangle}$.

(3) $\langle x,x\rangle=\sum\limits_i x(i)\overline{x(i)}=\sum\limits_i|x(i)|^2\geqslant 0$. And $\langle x,x\rangle=0\Leftrightarrow x=0$.

Hence, $l^2(I)$ is an inner product space.

Thirdly, we show that $l^2(I)$ is complete. Note that the induced norm in $l^2(I)$ is

$$||x||=(\sum_i|x(i)|^2)^{\frac{1}{2}}.$$

Let $\{x_n\}$ be Cauchy sequence in $l^2(I)$ and Then,

$$|x_m(i)-x_n(i)|\leqslant(\sum_i|x_m(i)-x_n(i)|^2)^{\frac{1}{2}}=||x_n-x_m||\to 0.$$

Thus, $x_n(i)$ is a Cauchy sequence over the real(or complex) numbers. By the completness of $\mathbb{R}$(or $\mathbb{C}$), it must converge to a limit, denoted by $xl(i)$. Now we show that $x_n\to x(x:i\mapsto xl(i))$ when $n\to\infty$.

Since $\{x_n\}$ is Cauchy, $\forall\epsilon$, $\exists N$, s.t., when $m,n>N,\forall k:$

$$\sum_i^k|x_m(i)-x_n(i)|^2\leqslant||x_m-x_n||^2<\sqrt{\frac{\epsilon}{2}}.$$

Let $m\to\infty$ and $n>N$, we have

$$\sum_i^k|x(i)-x_n(i)|^2\leqslant||x-x_n||^2\leqslant\sqrt{\frac{\epsilon}{2}}.$$

Let $k \to \infty$ and $n > N$, we have

$$(\sum_i^\infty |x(i) - x_n(i)|^2)^{\frac{1}{2}} \leqslant ||x - x_n|| \leqslant \frac{\epsilon}{2} < \epsilon.$$

Hence, $x_n \to x$ when $n \to \infty$.Then, we show that $x \in l^2(A)$.

We first show that $x(i) = 0$ for all but a countable number of i. Suppose the contrary. Then for any x_n, $||x - x_n||^2$ being a sum of uncountably many nonzero elements, which cannot be bounded. This is a contradiction as $||x - x_n|| \to 0$ $(n \to \infty)$.

Finally, we show that $\sum\limits_{i \in I} |x(i)|^2 < \infty$. Since $\{x_n\}$ is Cauchy, $\{x_n\}$ is bounded. Then $\exists M > 0$, $\forall n$, we have $||x_n|| < M$. Then $\forall n, k$, we have

$$\sum_i^k |x_n(i)|^2 \leqslant ||x_n||^2 \leqslant M^2.$$

Let $n \to \infty$, we have

$$\sum_i^k |x_n(i)|^2 \leqslant ||x||^2 \leqslant M^2.$$

Let $k \to \infty$, we have

$$\sum_i |x_n(i)|^2 \leqslant M^2.$$

Hence, $x \in l^2(I)$. Thus, every Cauchy sequence in $l^2(I)$ is convergent and so $l^2(I)$ is a Hilbert space. □

If $I = \mathbb{N}$, $l^2(I)$ is usually denoted by l^2. Let (X, Ω, μ) be a measure space consisting of a set X, a σ-algebra Ω of subsets of X, and a countably additive measure μ defined on Ω with values in the non-negative extended real numbers. Note that if Ω= the set of all subsets of A and for E in Ω, $\mu(E) := \infty$ if E is infinite and $\mu(E)$=the cardinality of E if E is finite, then $l^2(A)$ and $L^2(A, \Omega, \mu)$ are equal.

Exercise 1.2: I1 T6

Let u be a semi-inner product on $\mathscr{X}$ and put $\mathscr{N} = \{x \in \mathscr{X} : u(x, x) = 0\}$.
(a) Show that $\mathscr{N}$ is a linear subspace of $\mathscr{X}$.
(b) Show that if

$$\langle x + \mathscr{N}, y + \mathscr{N} \rangle \equiv u(x, y)$$

for all $x + \mathscr{N}$ and $y + \mathscr{N}$ in the quotient space $\mathscr{X}/\mathcal{N}$, then $\langle \cdot, \cdot \rangle$ is a well-definited inner product on $\mathscr{X}/\mathcal{N}$.

Proof. (a) For $x, y \in \mathscr{N}$, $u(x+y, x+y) = u(x, x) + u(x, y) + u(y, x) + u(y, y) = u(x, y) + u(y, x)$. To see that $x + y \in \mathscr{N}$, it suffices to show that $u(x, y) = 0$. But, this follows from Cauchy-Schwarz (which is true for semi-inner products). By Cauchy-Schwarz we have that

$$|u(x,y)| \leqslant \sqrt{u(x,x)}\sqrt{u(y,y)} = 0$$

Since $u(x,y) \geqslant 0$, it follows that $u(x,y) = 0$, and similarly $u(y,x) = 0$. Then $u(x+y, x+y) = 0$ and so $x + y \in \mathscr{N}$. For $c \in \mathbb{F}$, $u(cx, cx) = c\bar{c}u(x,x) = 0$. Then $cx \in \mathscr{N}$. Hence, $\mathscr{N}$ is a linear subspace of $\mathscr{X}$.

(b) Suppose $x_1 + \mathscr{N} = x_2 + \mathscr{N}$ and $y_1 + \mathscr{N} = y_2 + \mathscr{N}$ in the space $\mathscr{X}/\mathscr{N}$, i.e. we have that $x_1 - x_2 \in \mathscr{N}$ and $y_1 - y_2 \in \mathscr{N}$. To show that $\langle \cdot, \cdot \rangle$ is well defined, we need to show that $\langle x_1 + \mathscr{N}, y_1 + \mathscr{N} \rangle = \langle x_2 + \mathcal{N}, y_2 + \mathscr{N} \rangle$, i.e. $u(x_1, y_1) = u(x_2, y_2)$. By (a), for $a \in \mathscr{N}, b \in \mathscr{X}$, $|u(a,b)| \leqslant \sqrt{u(a,a)}\sqrt{u(b,b)} = 0$. Then $u(a,b) = 0$ as $|u(a,b)| \geqslant 0$. Now we have

$$\begin{aligned}
\langle x_1 + \mathscr{N}, y_1 + \mathscr{N} \rangle = u(x_1, y_1) &= u(x_2 + n_1, y_2 + n_2) \text{ where } n_1, n_2 \in \mathscr{N} \\
&= u(x_2, y_2) + u(x_2, n_2) + u(n_1, y_2) + u(n_1, n_2) \\
&= u(x_2, y_2) = \langle x_2 + \mathscr{N}, y_2 + \mathscr{N} \rangle
\end{aligned}$$

Hence, $\langle \cdot, \cdot \rangle$ is well defined. To show it is an inner product, it suffices to show that $\langle x+\mathscr{N}, x+\mathscr{N} \rangle = 0$ implies $x + \mathscr{N} = \mathscr{N}$. To see this note that if $\langle x + \mathscr{N}, x + \mathscr{N} \rangle = u(x,x) = 0$, then $x \in \mathscr{N}$ and so $x + \mathscr{N} = \mathscr{N}$. Hence, $\langle \cdot, \cdot \rangle$ is a inner product.

□

1.3 Reference

- lecture notes from brmh
- lecture notes from washington university
- lecture notes from cornell
- the completness of L^p space
    * lectrue notes from princeton
    * lecture notes from cornell
    * lecture notes from ndsu
- I1 T2
- I1 T6
- I1 T6

Chapter 2

Orthogonality

Definition 2.1

If $\mathscr{H}$ is a Hilbert space, we say $x, y \in \mathscr{H}$ are orthogonality if $< x, y >= 0$, in which case we write $x \perp y$. We say subsets A, B are orthogonal if $x \perp y$ for all $x \in A$ and $y \in B$, and we write $A \perp B$.

Theorem 2.1: The Pythagorean Theorem

If $x_1, x_2, ..., x_n \in \mathscr{H}$ and $x_j \perp x_k$ for $j \neq k$, then

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

Proof. When $n = 2$, expanding the left side. Then use induction. □

Remark. The pythagorean theorem is strongly related to orthogonality.

Definition 2.2

For any subset $A \subset \mathscr{H}$, define $A^{\perp} = \{x \in \mathscr{H} | \langle y, x \rangle = 0, \forall y \in A\}$.

Definition 2.3: Linear Span, Closed Linear Span

If $A \subset \mathscr{H}$, let $span(A)$:= the intersetion of all linear subspaces of $\mathscr{H}$ that contain A. $span(A)$ is called the linear span of A. Let $\overline{span(A)}$:= the intersetion of all closed linear subspaces of $\mathscr{H}$ that contain A. $\overline{span(A)}$ is called the closed linear span of A.

Proposition 2.1

(1) $span(A)$ is the smallest linear subspace of $\mathscr{H}$ that contains A.

(2) $span(A) = \{\sum_{k=1}^{n} \alpha_k x_k | n \geqslant 1, \alpha_k \in \mathbb{F}, x_k \in A\}$.

For convenience, we use the notation $A \leqslant \mathscr{H}$ to signify that A is a closed linear subspace of $\mathscr{H}$. and $\overline{span(A)}$ is denoted by $\vee A$.

Proposition 2.2

(1) $\vee A \leqslant \mathscr{H}$
(2) $\vee A$ is the smallest closed linear subspace of $\mathscr{H}$ that contains A.
(3) $\vee A =$ the closure of $\{\sum\limits_{k=1}^{n} \alpha_k x_k | n \geqslant 1, \alpha_k \in \mathbb{F}, x_k \in A\}$.

Definition 2.4: Direct Sum

$\mathscr{M}_1, \mathscr{M}_2 \leqslant \mathscr{H}$. We write

$$\mathscr{H} = \mathscr{M}_1 \bigoplus \mathscr{M}_2,$$

if $\mathscr{M}_1 \cap \mathscr{M}_2 = \{0\}$ and for every $x \in \mathscr{H}$ there exist $v_i \in \mathscr{M}_i$ with $v_1 + v_2 = x$.

Proposition 2.3

If $\mathscr{M} \leqslant \mathscr{H}$, then

$$\mathscr{H} = \mathscr{M} \bigoplus \mathscr{M}^{\perp}$$

Definition 2.5: Dense Subset

Let $(\mathscr{X}, || \cdot ||)$ be a normed space and $A \subset \mathscr{X}$. Then A is dense in $\mathscr{X}$ if $\overline{A} = \mathscr{X}$.

In Conway textbook, non-closed subspaces are called ”linear manifolds”.

Proposition 2.4

Let $\mathscr{Y}$ be a linear manifold in $\mathscr{H}$. Then $\mathscr{Y}$ is dense iff $\mathscr{H}^{\perp} = \{0\}$.

Definition 2.6: Convex Sets

A subset A of a vector space $\mathscr{X}$ is called convex if

$$\forall x, y \in A, t \in [0,1], tx + (1-t)y \in A.$$

Proposition 2.5

(1) Every subspace is convex.
(2) In a normed linear space, the open ball $B(x, \epsilon)$ is convex.
(3) If $A \subset \mathscr{X}$ is convex and $x \in \mathscr{X}$, then $A + x := \{y + x | y \in A\}$ is convex.

Proof. (3) A is convex, then $\forall a, b \in A, t \in [0,1], ta + (1-t)b \in A$. Then $t(a+x) + (1-t)(b+x) = ta + (1-t)b + x \in A + x$. Hence $A + x$ is convex. $\square$

Theorem 2.2

Every none-empty closed convex subset A of $\mathscr{H}$ has a unique element of smallest norm, i.e.

$$\exists! a \in A, ||a|| = \inf_{b \in A} ||b||.$$

Proof. Let $\delta = \inf\{||x|| | x \in A\}$. Since A is convex, $x + y \in A (\forall x, y \in A)$. Then, by the parallelogram law,

$$||x - y||^2 = 2(||x||^2 + ||y||^2) - ||x + y||^2 \leqslant 2(||x||^2 + ||y||^2) - 4\delta^2.$$

Existence follows: by the definition of infimum, we can choose $\{y_n\}_{n=1}^{\infty} \subset A$ for which $||y_n|| \to \delta$. Then when $n, m \to 0$,

$$||y_n - y_m||^2 \leqslant 2(||y_n||^2 + ||y_m||^2) - 4\delta^2 \to 0.$$

Hence, $\{y_n\}$ is Cauchy. By completeness, $\exists y \in \mathscr{H}$ for which $y_n \to y$, and since A is closed, $y \in A$. Also $||y|| = \lim_{n\to\infty} ||y_n|| = \delta$.
Uniqueness follows: if $||x|| = ||y|| = \delta$, then $||x - y||^2 \leqslant 4\delta^2 - 4\delta^2 = 0$, so $x = y$. □

Corollary 2.1

If A is a nonempty closed convex set in $\mathscr{H}$ and $x \in \mathscr{H}$, then there exists a unique closest element of A to x.

Proof. Since A is closed and convex, $A - x$ is closed and convex. By thm2.2, let z_0 be the unique element of smallest norm in $A - x$ and let $y_0 = z_0 + x$. Then $y_0 \in A$ and $||y_0 - x|| = ||z_0|| = \inf_{z \in A - x} ||z|| = \inf_{y \in A} ||y - x||$. Hence, y is the unique closest element of A to x. □

Since the closet point from closed convex set to point can be obtained, we can define the distance about these two object.

Definition 2.7

If A is a nonempty closed convex set in $\mathscr{H}$ and $x \in \mathscr{H}$, we define the distance from A to x:

$$\operatorname{dist}(x, A) := \inf_{y \in A} ||y - x||.$$

If the convex set in cor2.1 is in fact a closed linear subspace of $\mathscr{H}$, more can be said.

Theorem 2.3

If $\mathscr{M}$ is a closed linear subspace of $\mathscr{H}$, $x \in \mathscr{H}$, and y_0 is the unique element of $\mathscr{M}$ such that $||y_0 - x|| = \operatorname{dist}(x, \mathscr{M})$, then $x - y_0 \perp \mathscr{M}$. Conversely, if $y_0 \in \mathscr{M}$ such that $x - y_0 \perp \mathscr{M}$, then $||x - y_0|| = \operatorname{dist}(x, \mathscr{M})$.

Proof. You can refer to the lecture notes from brmh. □

Note that Theorem2.3, together with the uniqueness statement in corollary2.1, shows that if $\mathscr{M}$ is a closed linear subspace of $\mathscr{H}$ and $x \in \mathscr{H}$, then there is a unique element $y_0 \in A$ such that $x - y_0 \perp \mathscr{M}$. Thus a function $P : \mathscr{H} \to \mathscr{M}$ can be defined by $Px = y_0$.

Definition 2.8

If $\mathscr{M} \leqslant \mathscr{H}$ and $P : \mathscr{H} \to \mathscr{M}$ given by $Px = y_0$, where y_0 is the unique element in $\mathscr{M}$ such that $x - y_0 \perp \mathscr{M}$. Then P is called the orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}$. If we wish to show this dependence of P on $\mathscr{M}$, we will denote the orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}$ by $P_{\mathscr{M}}$.

Remark. P is onto.

Theorem 2.4

If P is the orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}$, then
(1) P is a linear transformation on $\mathscr{H}$,
(2) $||Px|| \leqslant ||x||$ for every x in $\mathscr{H}$,
(3) $P^2 = P$,
(4) $N(P) = \mathscr{M}^\perp$ and $R(P) = \mathscr{M}$.

Corollary 2.2

If $\mathscr{M} \leqslant \mathscr{H}$, then $(\mathscr{M}^\perp)^\perp = \mathscr{M}$.

Proof. Let $P = P_{\mathscr{M}}$, then $I - P$ is the orthogonal projection onto $\mathscr{M}^\perp$, then $\ker(I-P) = (\mathscr{M}^\perp)^\perp$, then $(\mathscr{M}^\perp)^\perp = \ker(I-P) = \operatorname{ran} P = \mathscr{M}$. □

Proposition 2.6

For any $A \subset \mathscr{H}$,
(1) $A^\perp \leqslant \mathscr{H}$
(2) $(span(A))^\perp = A^\perp$
(3) $\overline{A}^\perp = A^\perp$
(4) $A^{\perp\perp} = \vee A$.

Proof. (4) Since $A \subseteq ((A)^\perp)^\perp$ and $\vee A$ is the smallest closed linear subspace containing A, it follows that $\vee A \subseteq (A^\perp)^\perp$. On the other hand, since $C \subseteq D \Rightarrow D^\perp \subseteq C^\perp$ and $A \subseteq \vee A$, it follows that $A \subseteq \vee A \Rightarrow (\vee A)^\perp \subseteq A^\perp \Rightarrow A^{\perp\perp} \subseteq (\vee A)^{\perp\perp}$. Since $\vee A$ is closed, it follows that $(\vee A)^{\perp\perp} = \vee A$. Then $A^{\perp\perp} \subseteq \vee A$. Hence, $\vee A = A^{\perp\perp}$. □

2.1 Homework

Exercise 2.1: I2 T2

If $\mathscr{M} \leqslant \mathscr{H}$ and $P = P_{\mathscr{M}}$. show that $I - P$ is the orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}^{\perp}$.

Proof. Since $\mathscr{M} \leqslant \mathscr{H}$, it follows that $\mathscr{M} = (\mathscr{M}^{\perp})^{\perp}$. As $P = P_{\mathscr{M}}$, for $h \in \mathscr{H}$, $Ph \in \mathscr{M}$ and $h - ph = (I - P)h \in \mathscr{M}^{\perp}$. Also $h - (h - ph) = ph \perp \mathscr{M}^{\perp}$. Since Ph is unique, it follows that $(I - P)h$ is the unique element in $\mathscr{M}^{\perp}$ such that $h - (I - p)h = ph \perp \mathscr{M}^{\perp}$. Hence, $I - P$ is the orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}^{\perp}$. $\square$

Exercise 2.2: I2 T3

(a) If $\mathscr{M} \leqslant \mathscr{H}$, show that $\mathscr{M} \cap \mathscr{M}^{\perp} = (0)$ and every h in $\mathscr{H}$ can be written as $h = f + g$ where $f \in \mathscr{M}$ and $g \in \mathscr{M}^{\perp}$.
(b) If $\mathscr{M} + \mathscr{M}^{\perp} \equiv \{(f, g) : f \in \mathscr{M}, g \in \mathscr{M}^{\perp}\}$ and $T : \mathscr{M} + \mathscr{M}^{\perp} \to \mathscr{H}$ is defined by $T(f, g) = f + g$, show that T is a linear bijection and a homeomorphism if $\mathscr{M} + \mathscr{M}^{\perp}$ is given the product topology. (This is usually phrased by stating the $\mathscr{M}$ and $\mathscr{M}^{\perp}$ are toplogically complementary in $\mathscr{H}$.)

Proof. (a) $\forall h \in \mathscr{M} \cap \mathscr{M}^{\perp}$, we have $h \perp h$, then $\langle h, h\rangle = 0$ and so $h = 0$. Hence, $\mathscr{M} \cap \mathscr{M}^{\perp} = (0)$. Let $P = P_{\mathscr{M}}$, then ph is the unique element in $\mathscr{M}$ such that $h - ph \perp \mathscr{M}$, then $h - ph \in \mathscr{M}^{\perp}$. Let $f = ph, g = h - ph$, then $f \in \mathscr{M}, g \in \mathscr{M}^{\perp}$ and $h = f + g$.
(b) Firstly, we show that T is linear. For $c \in \mathbb{F}$ and $(f_1, g_1), (f_2, g_2) \in \mathscr{M} + \mathscr{M}^{\perp}$, $T(f_1 + f_2, g_1 + g_2) = f_1 + f_2 + g_1 + g_2 = f_1 + g_1 + f_2 + g_2 = T(f_1, g_1) + T(f_2, g_2)$ and $T(cf_1, cg_1) = cf_1 + cg_1 = cT(f_1, g_1)$. Hence, T is linear.

Next, we show that T is bijective. We know that T is well-defined as $(f_1, g_1) = (f_2, g_2)$, $T(f_1, g_1) = f_1 + g_1 = f_2 + g_2 = T(f_2, g_2)$. By (a), we know that T is surjective. Now, we show that T is injective. If $T(f_1, g_1) = T(f_2, g_2)$, then $f_1 + g_1 = f_2 + g_2$ and so $f_1 - f_2 = g_1 - g_2$. Hence, $f_1 - f_2 = g_1 - g_2 \in \mathscr{M} \cap \mathscr{M}^{\perp} = (0)$. Then $f_1 = f_2, g_1 = g_2$ and so $(f_1, g_1) = (f_2, g_2)$. Hence, T is a bijection.

Now we show that T is homeomorphism. It suffices to show that T and T^{-1} is continuous. For a sequence $\{(f_n, g_n)\} \in \mathscr{M} + \mathscr{M}^{\perp}$ such that $(f_n, g_n) \underset{n\to\infty}{\longrightarrow} (f, g)$, we have $f_n \underset{n\to\infty}{\longrightarrow} f$, $g_n \underset{n\to\infty}{\longrightarrow} g$ and so $T(f_n, g_n) = f_n + g_n \underset{n\to\infty}{\longrightarrow} f + g$. Hence, T is continuous. For a sequence $\{h_n\} \in \mathscr{H}$ such that $h_n \underset{n\to\infty}{\longrightarrow} h$, there are sequences $\{f_n\} \subseteq \mathscr{M}$, $\{g_n\} \subseteq \mathscr{M}^{\perp}$ and $f \in \mathscr{M}, g \in \mathscr{M}^{\perp}$ such that $h_n = f_n + g_n \underset{n\to\infty}{\longrightarrow} h = f + g$. Since $f_n \perp g_n$, $f_n \underset{n\to\infty}{\longrightarrow} f$ and $g_n \underset{n\to\infty}{\longrightarrow} g$. Then $T^{-1}(h_n) = T^{-1}(f_n + g_n) = (f_n, g_n) \underset{n\to\infty}{\longrightarrow} (f, g) = T^{-1}(f + g) = T^{-1}(h)$. Hence, T^{-1} is continuous. Hence, T is a homeomorphism. $\square$

2.2 Reference

- lecture notes from washington university
- lecture notes from brmh

- lecture notes from umich
- lecture notes from msu

Chapter 3

The Riesz Representation Theorem

Definition 3.1: Linear Functional

(1) A linear functional on a vector space $\mathscr{X}$ is a linear mapping $L : \mathscr{X} \to$ Ł:

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y), \forall x, y \in \mathscr{X}, \alpha, \beta \in Ł.$$

(2) A linear functional L on a normed space $(\mathscr{X}, ||\cdot||)$ is called a bouned linear functional if there exists $C \geqslant 0$ such that $|L(x)| \leqslant C||x||$ for each $x \in \mathscr{X}$.

Proposition 3.1

Let $\mathscr{H}$ be a Hilbert space and $L : \mathscr{H} \to$ Ł a linear functional. The following statements are equivalent.
(1) L is continuous at $h = 0$.
(2) L is continuous.
(3) L is continuous at some points.
(3) L is bounded.
(4) Ker L is closed in $\mathscr{H}$.

Note that the bounded linear functionals forms a vector space $\mathscr{H}^*$: $0 \in \mathscr{H}^*$, if $f_i \in \mathscr{H}^*$ and $\alpha_i \in$ Ł then $\alpha_1 f_1 + \alpha_2 f_2 \in \mathscr{H}^*$. We will now explain how to define a norm on $\mathscr{H}^*$.

Definition 3.2

For a bounded linear functional $L \in \mathscr{H}^*$, its norm is defined as

$$||L||_{\mathscr{H}^*} = \sup_{||h||_{\mathscr{H}} \leqslant 1} |L(x)|.$$

For convenience, the following content will follow the convention: $||L|| = ||L||_{\mathscr{H}^*}, ||x|| = ||x||_{\mathscr{H}}$. Let's check three properties fo the norm:
(1) $||L|| \geqslant 0$; $||L|| = 0 \Leftrightarrow L = 0$
(2) $||\alpha L|| = |\alpha|||L||$

(3)

$$\begin{aligned}||f_1+f_2|| &= \sup_{||h||\leqslant 1}|(f_1+f_2)(h)| \\ &= \sup_{||h||\leqslant 1}|f_1(h)+f_2(h)| \\ &\leqslant \sup_{||h||\leqslant 1}|f_1(h)| + \sup_{||h||\leqslant 1}|f_2(h)| \\ &= ||f_1|| + ||f_2||\end{aligned}$$

Proposition 3.2

If L is a bounded linear functional, then
(1) $|L(h)| \leqslant ||L|| \cdot ||h||$ for every $h \in \mathscr{H}$.
(2)

$$\begin{aligned}||L|| &= \sup_{||h||=1}|L(x)| \\ &= \sup_{h\in\mathscr{H}\setminus\{0\}}\frac{|L(x)|}{||x||} \\ &= \inf\{c>0 : |L(h)| \leqslant c||h||, h \in \mathscr{H}\}.\end{aligned}$$

Proposition 3.3

If $h_0 \in \mathscr{H}$, then $L_{h_0} : \mathscr{H} \to \mathbb{F}$ given by

$$L_{h_0}(h) = \langle h, h_0\rangle \tag{3.1}$$

is a bounded linear funcional on $\mathscr{H}$, with $||L_{h_0}|| = ||h_0||$.

Proof. For $h_1, h_2 \in \mathscr{H}$, $\alpha, \beta \in \mathbb{F}$,

$$L_{h_0}(\alpha h_1 + \beta h_2) = \langle \alpha h_1 + \beta h_2, h_0\rangle = \alpha\langle h_1, h_0\rangle + \beta\langle h_2, h_0\rangle = \alpha L_{h_0}(h_1) + \beta L_{h_0}(h_2).$$

Hence, L_{h_0} is linear. By the Cauchy-Schwarz inequality, for $h \in \mathscr{H}$,

$$|L_{h_0}(h)| = |\langle h, h_0\rangle| \leqslant ||h|| \cdot ||h_0||.$$

Hence, L_{h_0} is bounded and $||L_{h_0}|| = \sup_{h\in\mathscr{H}\setminus\{0\}}\frac{|L_{h_0}(h)|}{||h||} \leqslant ||h_0||$. Then $||L_{h_0}|| = \sup_{||h||=1}|L(h)| \leqslant ||h_0||$. Since $||\frac{h_0}{||h_0||}|| = 1$ and $L_{h_0}(\frac{h_0}{||h_0||}) = \langle\frac{h_0}{||h_0||}, h_0\rangle = ||h_0||$, it follows that $||L_{h_0}|| \geqslant ||h_0||$. Hence, $||L_{h_0}|| = ||h_0||$. $\square$

The following theorem shows that all bounded linear functionals in Hilbert space are of the form (3.1). In other words, every bounded linear functional on $\mathscr{H}$ can be identified with a unique point in the space itself.

Theorem 3.1

If L is a bounded linear functional on $\mathscr{H}$, then there is a unique vector $h_0 \in \mathscr{H}$ such that

$$L(h) = \langle h, h_0 \rangle, \forall h \in \mathscr{H}. \tag{3.2}$$

Moreover, $||L|| = ||h_0||$

Proof. By proposition3.1, $\mathscr{M} = \mathrm{Ker}L$ is closed in $\mathscr{H}$. If $\mathscr{M} = \mathscr{H}$, $L(h) = 0$, $h_0 = 0$ is desired requested. If $\mathscr{M} \neq \mathscr{H}$, then $\mathscr{M}^\perp \neq (0)$ and so there exists some $u \in \mathscr{M}^\perp$ such that $L(u) \neq 0$, and we take $f_0 = \frac{f_0}{L(f_0)} \in \mathscr{M}^\perp$ so that $L(f_0) = 1$. Then $L(h - L(h)f_0) = L(h) - L(h)L(f_0) = L(h) - L(h) = 0$ for all $h \in \mathscr{H}$. Hence, $h - L(h)f_0 \in \mathscr{M}$ and so we have $0 = \langle h - L(h)f_0, f_0 \rangle$ and so $\langle h, f_0 \rangle = L(h)||f_0||^2$. Let $h_0 = f_0/||f_0||^2$ now seals the deal.
Uniquness follows because if $L(h) = \langle h, h_0 \rangle = \langle h, h_0' \rangle$ then $h_0 - h_0' \perp \mathscr{H}$ and so $h_0 - h_0' \in \mathscr{H}^\perp = \{0\}$. $||L|| = ||h_0||$ is shown in proposition3.3 □

3.1 Homework

Exercise 3.1: I3 T3

Let $\mathscr{H} = l^2(\mathbb{N} \cup \{0\})$.

(a) Show that if $\{\alpha_n\} \in \mathscr{H}$, then the power series $\sum\limits_{n=0}^{\infty} \alpha_n z^n$ has radius of convergence $\geqslant 1$.

(b) If $|\lambda| < 1$ and $L : \mathscr{H} \to \mathbb{F}$ is defined by $L(\{\alpha_n\}) = \sum\limits_{n=0}^{\infty} \alpha_n \lambda^n$, find the vector h_0 in $\mathscr{H}$ such that $L(h) = \langle h, h_0 \rangle$ for every h in $\mathscr{H}$.

(c) What is the norm of the linear functional L defined in (b)?

Proof. (a) If $\{\alpha_n\} \in \mathscr{H}$, then $\sum\limits_n |\alpha_n|^2 < \infty$, i.e. $\sum\limits_n |\alpha_n|^2$ is absolutely convergent. Then by root test, $\lim_{n\to\infty} \sqrt[n]{|\alpha_n|^2} \leqslant 1$. Then $\lim_{n\to\infty} \sqrt[n]{|\alpha_n|} \leqslant 1$. Hence, the radius of convergence is $\frac{1}{\lim_{n\to\infty} \sqrt[n]{|\alpha_n|}} \geqslant 1$.
(b)(c) Since $|\lambda| < 1$, then for $h = \{\alpha_n\} \in \mathscr{H}$, $|L(h)| = \sum\limits_n \lambda^n \alpha_n$ converges and so is bounded. Then L is a bounded linear functional on $\mathscr{H}$, so by the Riesz representation theorem, there exists a vector h_0 such that $L(h) = \langle h, h_0 \rangle$ for every h in $\mathscr{H}$. Assume $h_0 = \{\beta_n\}$, then $\langle h, h_0 \rangle = \sum\limits_n \alpha_n \overline{\beta_n} = L(h) = \sum\limits_n \lambda^n \alpha_n = \sum\limits_n \alpha_n \overline{(\overline{\lambda})^n}$. Hence, we can let $h_0 = (1, \overline{\lambda}, (\overline{\lambda})^2, \ldots)$. Since $||h_0||_{l^2}^2 = \sum\limits_n |(\overline{\lambda})^n|^2 = \sum\limits_n |\lambda|^{2n} = \lim_{n\to\infty} \frac{1(1-(|\lambda|^2)^n)}{1-|\lambda|^2} = \frac{1}{1-|\lambda|^2}$, it follows that $||h_0|| = \frac{1}{\sqrt{1-|\lambda|^2}}$ and by Riesz representation theorem $||L|| = ||h_0|| = \frac{1}{\sqrt{1-|\lambda|^2}}$. □

Exercise 3.2: I3 T5

Let $\mathscr{H}$ = the collection of all absolutely continuous functions $f : [0, 1] \to \mathbb{F}$ such that $f(0) = 0$, $f' \in L^2(0, 1)$ and for $f, g \in \mathscr{H}$, $\langle f, g \rangle = \int_0^1 f'(t)\overline{g'(t)}dt$. By Example 1.8, we know that $\mathscr{H}$ is a Hilbert space. If $0 < t \leqslant 1$, define $L : \mathscr{H} \to \mathbb{F}$ by $L(h) = h(t)$. Show that L is a bounded linear funcional, find $||L||$, and find the vector h_0 in $\mathscr{H}$ such that $L(h) = \langle h, h_0 \rangle$ for all h in $\mathscr{H}$.

Proof. Firstly, we show that L is linear. For $\alpha, \beta \in \mathbb{F}$, $h_1, h_2 \in \mathscr{H}$, then $L(\alpha h_1 + \beta h_2) = (\alpha h_1 + \beta h_2)(t) = \alpha h_1(t) + \beta h_2(t) = \alpha L(h_1) + \beta L(h_2)$. Hence, L is linear. Then, we show that L is continuous. For a sequence $\{h_n\} \in \mathscr{H}$ such that $h_n \underset{n\to\infty}{\longrightarrow} h$, if $0 < t \leqslant 1$ we have

$$\lim_{n\to\infty} L(h_n) = \lim_{n\to\infty} h_n(t) \underset{h_n \text{ is absolutely continuous}}{=} h(t) = L(h).$$

Hence, L is a continuous linear functional and so a bounded linear funcional. Then by Riesz representation theorem, there exists a vector $h_0 \in \mathscr{H}$ such that $L(h) = \langle h, h_0 \rangle$ for all h in $\mathscr{H}$. and $||L|| = ||h_0||$. Then $\langle h, h_0 \rangle = \int_0^1 h'(t)\overline{h_0'(t)}dt = L(h) = h(t) = \int_0^t h'(x)dx$. Hence, we can let $\overline{h_0'(t)} = \begin{cases} 1 & 0 < x \leqslant t \\ 0 & t < x \leqslant 1. \end{cases}$ and so $h_0(t) = \begin{cases} x & 0 < x \leqslant t \\ 0 & t < x \leqslant 1. \end{cases}$ Then $||L|| = ||h_0|| = \sqrt{\langle h_0, h_0 \rangle} = \sqrt{\int_0^1 h_0'(x)\overline{h_0'(x)}dx} = \sqrt{\int_0^t 1 \cdot 1 dx} = \sqrt{t}$. $\square$

3.2 Reference

- lecture notes from brmh
- Riesz Representation Theorem geometric intuition
- lecture notes from msu
- lecture notes from mit

Chapter 4

Orthonormal Sets of Vectors and Bases

It will be shown in this chapter that, as in Euclidean space, each Hilbert space can be coordinatized. The vehicle for introducing the coordinates is an orthonormal basis. The corresponding vectors in $\mathbb{F}^d$ are the vectors $\{e_1, e_2, ..., e_d\}$, where e_k is the d-tuple having a 1 in the kth place and zeros elsewhere.

Subsets $\mathscr{E}$ mentioned below can be finite, countably infinite and uncoutably infinite.

Definition 4.1

A subset $\mathscr{E} \subset \mathscr{H}$ is said to orthogonal if $\langle e_1, e_2 \rangle = 0$, for all $e_1 \neq e_2 \in \mathscr{E}$. We say it is orthonormal if we further regive $||e|| = 1$ for all $e \in \mathscr{E}$.

Remark. Orthogonal set can contain zero vector, but orthonormal set can not.

Definition 4.2

$\{h_1, ..., h_k\}$ is said to be linearly dependent, if there exist scalars $a_1, a_2, ..., a_k$, not all zero, such that

$$a_1 h_1 + ... + a_k h_k = 0$$

where 0 denotes the zero vector. An infinite set of vectors is linearly independent if every nonempty finite subset is linearly independent.

Corollary 4.1

Every orthonormal set is linearly independent.

Proof. If the set is finite. Suppose $\{h_1, ..., h_k\}$ is an orthonormal set. Assume $b_1 h_1 + ... b_k h_k = 0$, then $b_1 u_1 = -\sum_{i=2}^{k} b_i u_i$. When $2 \leqslant i \leqslant n$, $0 = \langle h_i, -b_1 h_1 \rangle = \sum_{j=2}^{k} b_i \langle h_i, h_j \rangle = b_i \langle h_i, h_i \rangle = b_i$. Then $b_1 h_1 = 0$. Since $h_1 \neq 0$, it follows that $b_1 = 0$. Hence, $b_i = 0$ for all $1 \leqslant i \leqslant k$ and so $\{h_1, ..., h_k\}$ is linearly independent. If the set is infinite, the above proof can show that every nonempty finite subset is linearly independent and so the set is linearly independent. $\square$

Now we consider the question: Does the orthonormal set always exist? The following theorem tell us we can always construct an orthonormal set.

Theorem 4.1: The Gram-Schmidt Orthogonalization Process

If $\{h_1, ..., h_k\}$ is a linearly independent subset of $\mathscr{H}$, then there is an orthonormal set $\{e_1, ..., e_k\}$ such that $span(\{e_1, ..., e_k\}) = span(\{h_1, ..., h_k\})$.

Proof. Define $\{e_k\}$ inductivey. Start with $e_1 = \frac{h_1}{||h_1||}$. Suppose for $k-1$, $e_1, ..., e_{k-1}$ exist. Let $v_k = h_k - \sum_{j=1}^{k-1} < e_j, h_k > e_j$ and $e_k = \frac{v_k}{||v_k||}$. Then $\langle e_k, e_j \rangle = \frac{1}{||v_k||}(\langle h_k, e_j \rangle - \langle e_j, h_k \rangle \langle e_j, e_j \rangle) = 0$ and $||e_k|| = 1$. Hence, $\{e_1, ..., e_k\}$ is an orthonormal set. from the construct process of $e_j (1 \leqslant j \leqslant k)$, we can know $e_j \in span(\{h_1, ..., h_k\})$ and so $span(\{e_1, ..., e_k\}) \subset span(\{h_1, ..., h_k\})$. Since $\{e_1, ..., e_k\}$ is linearly independent, $dim(span(\{e_1, ..., e_k\})) = n$. Hence, $span(\{e_1, ..., e_k\}) = span(\{h_1, ..., h_k\})$. □

Proposition 4.1

If $\mathscr{H}$ is a Hilbert space and $\{h_n : n \in \mathbb{N}\}$ is a linearly independent subset of $\mathscr{H}$, then there is an orthonormal set $\{e_n : n \in \mathbb{N}\}$ such that for every n, the linear span of $\{e_1, ..., e_n\}$ equals the linear span of $\{h_1, ..., h_n\}$.

Remember that $\vee A$ is the closed linear span of A. Now we consider the following question: How can we determine Ph when h and the subspace $\mathscr{M}$ are given? When $\mathscr{M}$ is finite, we have

Proposition 4.2

Let $\{e_1, ..., e_n\}$ be an orthonormal set in $\mathscr{H}$ and let $\mathscr{M} = \vee\{e_1, ..., e_n\}$. If P is the orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}$, then

$$Ph = \sum_{k=1}^{n} \langle h, e_k \rangle e_k$$

for all h in $\mathscr{H}$.

Proof. Let $Qh = \sum_{k=1}^{n} \langle h, e_k \rangle e_k$. it suffices to show that Qh is the unique element in $\mathscr{M}$ such that $h - Qh \perp \mathscr{M}$. For $1 \leqslant j \leqslant n$, $\langle Qh, e_j \rangle = \sum_{k=1}^{n} \langle h, e_k \rangle \langle e_k, e_j \rangle = \langle h, e_j \rangle$ as $e_k \perp e_j$ when $k \neq j$. Then $\langle h - Qh, e_j \rangle = 0$ and so $h - Qh \perp \mathscr{M}$. Since $Qh \in \mathscr{M}$, it follows that Qh is the unique element in $\mathscr{M}$ such that $h - Qh \perp \mathscr{M}$. Hence, by the property of Ph, $Ph = Qh$ for every h in $\mathscr{H}$. □

Proposition 4.3: Bessel's inequality

If $\{e_n : n \in \mathbb{N}\}$ is an orthonormal set and $h \in \mathscr{H}$, then

$$\sum_{n=1}^{\infty} |\langle h, e_n \rangle|^2 \leqslant ||h||^2.$$

Proof. Let $h_n = h - \sum_{k=1}^{n} \langle h, e_k \rangle e_k$. Then $\langle h_n, e_k \rangle = \langle h, e_k \rangle - \langle h, e_k \rangle = 0$. By the Pythagorean Theorem,

$$\begin{aligned} ||h||^2 &= ||h_n||^2 + ||\sum_{k=1}^{n} \langle h, e_k \rangle e_k||^2 \\ &= ||h_n||^2 + \sum_{k=1}^{n} |\langle h, e_k \rangle|^2 \\ &\geqslant \sum_{k=1}^{n} |\langle h, e_k \rangle|^2. \end{aligned}$$

Since n was arbitrary, the result is proved. □

Corollary 4.2

If $\mathscr{E}$ is an orthonormal set in $\mathscr{H}$ and $h \in \mathscr{H}$, then $\langle h, e \rangle \neq 0$ for at most countable number of vectors e in $\mathscr{E}$.

Proof. Let $n \in \mathbb{N}_+$. We claim that the subset $\mathscr{E}_n$

$$\mathscr{E}_n = \{e \in \mathscr{E} : \frac{1}{n} \leqslant |\langle h, e \rangle|\}$$

of $\mathscr{E}$ is a finite set. Pick $e_1, ..., e_N$ from $\mathscr{E}_n$, then by Bessel's inequality,

$$||h||^2 \geqslant \sum_{k=1}^{N} |\langle h, e_k \rangle|^2 \geqslant \frac{N}{n^2},$$

It follows that the cardinality of $\mathscr{E}_n$ cannot exceed $n^2||h||^2$. Let $\mathscr{E}_x$ be the subset of $\mathscr{E}$ consisting of all e such that $\langle h, e \rangle$ is non-zero. Since $\mathscr{E}_x = \cup_{n \in \mathbb{N}_+} \mathscr{E}_n$, it follows that $\mathscr{E}_x$ is a countable union of finite sets and so $\mathscr{E}_x$ is countable. □

Corollary 4.3

If $\mathscr{E}$ is an orthonormal set and $h \in \mathscr{H}$, then

$$\sum_{e \in \mathscr{E}} |\langle h, e \rangle|^2 \leqslant ||h||^2.$$

Proof. Restrict our attention to the $\{e \in \mathscr{E} : \langle h, e\rangle \neq 0\}$ which is countable by the last corollary, and now it is just a straight up use of Bessel's ineq in proposition4.3 □

Let $\mathscr{F}$ be the collection of all finite subsets of I and order by inclusion, so $\mathscr{F}$ becomes a directed set. For each F in $\mathscr{F}$, define $h_F = \sum\{h_i : i \in F\}$. Since this is a finite sum, h_F is a well-defined element of $\mathscr{H}$. Now $\{h_F : F \in \mathscr{F}\}$ is a net in $\mathscr{H}$.

Definition 4.3

With the notation above, the sum $\sum\{h_i : i \in I\}$ converges if the net $\{h_F : F \in \mathscr{F}\}$ converges; the value of the sum is the limit of the net.

Now Corollary4.3 can be given its precise meaning; namely, $\sum\limits_{e\in\mathscr{E}} |\langle h, e\rangle|^2$ converges and the value $\leqslant ||h||^2$.

Lemma 4.1

If $\mathscr{E}$ is an orthonormal set and $h \in \mathscr{H}$, then $\sum\{\langle h, e\rangle e : e \in \mathscr{E}\}$ converges in $\mathscr{H}$.

Now we can determine Ph even if $\mathscr{M}$ is infinite or possibly uncountable.

Corollary 4.4

Let $\mathscr{E}$ be an orthonormal subset of $\mathscr{H}$ and let $\mathscr{M} = \vee\mathscr{E}$. If P is the orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}$, show that $Ph = \sum\{\langle h, e\rangle e : e \in \mathscr{E}\}$ for every h in $\mathscr{H}$.

In particular, if $\vee\mathscr{E} = \mathscr{H}$, then every $h \in \mathscr{H}$ may be expanded in terms of elements of $\mathscr{E}$. The following theorem gives equivalent conditions for this property of $\vee\mathscr{E}$.

Theorem 4.2

If $\mathscr{E}$ is an orthonormal subset of $\mathscr{H}$, then the following conditions are equivalent:
(1) If $h \in \mathscr{H}$ and $h \perp \mathscr{E}$, then $h = 0$.($\mathscr{E}$ is a maximal orthonormal set).
(2) $\vee\mathscr{E} = \mathscr{H}$.
(3) If $h \in \mathscr{H}$, then $h = \sum\{\langle h, e\rangle e : e \in \mathscr{E}\}$.
(4) If $g, h \in \mathscr{H}$, then $\langle g, h\rangle = \sum\{\langle g, e\rangle\langle e, h\rangle : e \in \mathscr{E}\}$.
(5) If $h \in \mathscr{H}$, then $||h||^2 = \sum\{|\langle h, e\rangle|^2 : e \in \mathscr{E}\}$ (Parseval's Identity)

Proof. (1)$\Rightarrow$ (2): Since $\{u_n\}$ is maximal, $\underline{span(\{u_n\})^\perp} = \{0\}$. Let $\mathscr{M} = \overline{span(\{u_n\})}$, then $\mathscr{M}^\perp = \{0\}$. Since $\mathscr{M}^\perp \bigoplus \mathscr{M} = \mathscr{H}$, $\mathscr{M} = \overline{span(\{u_n\})} = \mathscr{H}$.
(2)$\Rightarrow$ (3): □

Definition 4.4

An orthonormal set $\mathscr{E}$ in $\mathscr{H}$ satisfying any of the equivalent conditions (1)-(5) in theorem4.2 is called a complete orthonormal set(or a complete orthonormal system) or an orthonormal basis in $\mathscr{H}$.

Remark. If $\mathscr{H}$ is infinite dimensional, an orthonormal basis is not a basis in the usual definition of a basis for a vector space (i.e., each $h \in \mathscr{H}$ has a unique representation as a finite linear combination of basis elements). Such a basis in this context is called a Hamel basis.

The following theorem shows that orthonormal basis always exists in Hilbert space.

Theorem 4.3

Every Hilbert space $\mathscr{H}$ has an orthonormal basis. If $\mathscr{E}$ is an orthonormal set, then $\mathscr{E}$ has an orthonormal basis containing $\mathscr{E}$.

Proof. referring to lecture notes from ucdavis □

Just as in finite dimensional spaces, a basis in Hilbert space can be used to define a concept of dimension. For this purpose the next result is pivotal.

Proposition 4.4

If $\mathscr{H}$ is a Hilbert space, any two bases have the same cardinality.

Definition 4.5

The dimension of a Hilbert space is the cardinality of a basis and is denoted by $\dim \mathscr{H}$.

Proposition 4.5

If $\mathscr{H}$ is an infinite dimensional Hilbert space, then $\mathscr{H}$ is separable if and only if $\mathscr{H}$ has a countable basis.

Exercise 4.1: I4 T13

Let $\mathscr{E}$ be an orthonormal subset of $\mathscr{H}$ and let $\mathscr{M} = \vee\mathscr{E}$. If P is the orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}$, show that $Ph = \sum\{\langle h, e\rangle e : e \in \mathscr{E}\}$ for every h in $\mathscr{H}$.

Proof. Let $Qh = \sum\{\langle h, e\rangle e : e \in \mathscr{E}\}$. it suffices to show that Qh is the unique element in $\mathscr{M}$ such that $h - Qh \perp \mathscr{M}$. By Lemma 4.12, there are vectors $e_1, e_2, \ldots$ in $\mathscr{E}$ such that $\{e \in \mathscr{E} : \langle h, e\rangle \neq 0\} = e_1, e_2, \ldots$ and $Qh = \sum_{n=1}^{\infty} \langle h, e_n\rangle e_n$. By Bessel's Inequality, $\sum_{n=1}^{\infty} \langle h, e_n\rangle^2 \leqslant ||h||^2 < \infty$. and thus $\{\langle h, e_n\rangle\}$ is l^2 and the sum for Qh converges. Since $\langle Qh, e_n\rangle = \langle h, e_n\rangle$ for all n, it follows that $\langle h - Qh, e_n\rangle = 0$ for all n. Then by the continuity of inner product, $\langle s, h - Qh\rangle = 0$ for any s in l^2-span of $\{e_n\}$. Hence, $h - Qh \perp \mathscr{M}$. Since $Qh \in \mathscr{M}$, it follows that Qh is the unique element in $\mathscr{M}$ such that $h - Qh \perp \mathscr{M}$. Hence, by the property of Ph, $Ph = Qh$ for every h in $\mathscr{H}$. □

Exercise 4.2: I4 T19

If $\{h \in \mathscr{H} : ||h|| \leqslant 1\}$ is compact, show that $\dim \mathscr{H} < \infty$.

Proof. Suppose that $\mathscr{H}$ is not finite dimensional. We want to show that $\overline{B(0;1)}$ is not compact. It suffices to show that $\overline{B(0,1)}$ is not sequentially compact. To do this, we construct a sequence in $\overline{B(0,1)}$ which have no convergent subsequence. We will uset the following fact usually known as Riesz's Lemma: Let $\mathscr{M}$ be a closed subspace of a Banach space $\mathscr{X}$. Given any $r \in (0,1)$, there exists

an $x \in \mathscr{X}$ such that $||x|| = 1$ and $d(x, \mathscr{M}) \geqslant r$.
Pick $x_1 \in \mathscr{X}$ such that $||x_1|| = 1$. Let $\mathscr{M}_1 = \vee\{x_1\}$. Then $\mathscr{M}_1$ is closed. Then accroding to Riesz's Lemma, there exists $x_2 \in \mathscr{H}$ such that $||x_2|| = 1$ and $d(x_2, \mathscr{M}_1) \geqslant \frac{1}{2}$. Now consider the subspace $\mathscr{M}_2 = \vee\{x_1, x_2\}$. Since $\mathscr{H}$ is infinite dimensional, $\mathscr{M}_2$ is a proper closed subspace of $\mathscr{H}$, and we can apply the Riesz's Lemma to find an $x_3 \in \mathscr{H}$ such that $||x_3|| = 1$ and $d(x_3, \mathscr{M}_2) \geqslant \frac{1}{2}$. If we continue to proceed this way, we will have a sequence (x_n) and a sequence of closed subspaces $(\mathscr{M}_n)$ such that for all $n \in \mathbb{N}$: $||x_n|| = 1$ and $d(x_{n+1}, \mathscr{M}_n) \geqslant \frac{1}{2}$. It is clear that the sequence (x_n) is in $\overline{B(0,1)}$, and for $m > n$ we have $x_m \in \mathscr{M}_m \subset \mathscr{M}_{n-1}$ and $||x_n - x_m|| \geqslant d(x_n, \mathscr{M}_{n-1}) \geqslant \frac{1}{2}$. Therefore, no subsequence of (x_n) can form a Cauchy sequence. Thus, $\overline{B(0,1)}$ is not compact. □

4.1 Reference

- lecture notes from brmh
- lecture notes from mit
- lecture notes from ucdavis
- lecture notes from washington
- lecture notes from cornell
- lecture notes from cuhk
- I4 T13 P5
- I4 T19: P17

Chapter 5

Isomorphic Hilbert Spaces and the Fourier Transform for the Circle

Definition 5.1

If $\mathscr{H}$ and $\mathscr{K}$ are Hilbert spaces, an isomorphism between $\mathscr{H}$ and $\mathscr{K}$ is linear surjection $U : \mathscr{H} \to \mathscr{K}$ such that

$$\langle Uh, Ug \rangle = \langle h, g \rangle$$

for all $h, g \in \mathscr{H}$. In this case $\mathscr{H}$ and $\mathscr{K}$ are said to be isomorphic.

Remark. Many call what we call an isomorphism a unitary operator

from the previous definition, we know that isomorphism preserves inner product. Now we claim that isomorphism preserves distance and completeness.

Proposition 5.1

If $V : \mathscr{H} \to \mathscr{K}$ is a linear map between Hilbert space, then V is an isometry iff $< Vh, Vg >=< h, g >$ for all $h, g \in \mathscr{H}$.

Exercise 5.1: I5 T6

Let $\mathscr{C} = \{f \in C[0, 2\pi] : f(0) = f(2\pi)\}$ and show that $\mathscr{C}$ is dense in $L^2[0, 2\pi]$.

Proof. In a Hilbert space $\mathscr{H}$, $A \subset \mathscr{H}$ is dense if for $h \in \mathscr{H}$, there exists a sequence $(a_n) \subseteq A$ such that $a_n \underset{n\to\infty}{\longrightarrow} h$. Since $C[0, 2\pi]$ is dense in $L^2[0, 2\pi]$ and $\mathscr{C} \subset C[0, 2\pi]$, it suffices to show that $\mathscr{C}$ is dense in $C[0, 2\pi]$. For $f \in C[0, 2\pi]$, then there exists $M > 0$ such that $|f(x)| \leqslant M$, $\forall x \in [0, 2\pi]$. Let $f_n(x) = \begin{cases} f(x), & 0 \leqslant x \leqslant 2\pi - \frac{1}{n} \\ f(2\pi - \frac{1}{n}) + \frac{f(0)-f(2\pi-\frac{1}{n})}{\frac{1}{n}}(x - 2\pi + \frac{1}{n}), & 2\pi - \frac{1}{n} < x \leqslant 2\pi \end{cases}$. Then $f_n(0) = f(0)$, $f_n(2\pi) = f(2\pi - \frac{1}{n}) + \frac{f(0)-f(2\pi-\frac{1}{n})}{\frac{1}{n}}(2\pi - 2\pi + \frac{1}{n}) = f(0)$. Hence, $f_n \in \mathscr{C}$. Since $|f|, |f_n| \leqslant M$, it follows that $||f - f_n|| = (\int_{[0,2\pi]} |f - f_n|^2 dx)^{1/2} = (\int_{[2\pi - \frac{1}{n}, 2\pi]} |f - f_n|^2)^{1/2} \leqslant (\frac{(2M)^2}{n})^{1/2} = \frac{2M}{\sqrt{n}} \underset{n\to\infty}{\longrightarrow} 0$. Then $\lim_{n\to\infty} f_n = f$. Hence, $\mathscr{C}$ is dense in $C[0, 2\pi]$ and so is dense in $L^2[0, 1]$. $\square$

Exercise 5.2: I5 T9

If $\mathscr{H}$ and $\mathscr{K}$ are Hilbert spaces and $U : \mathscr{H} \to \mathscr{K}$ is surjective function such that $\langle Uf, Ug\rangle = \langle f, g\rangle$ for all vectors f and g in $\mathscr{H}$, then U is linear.

Proof. It suffices to show that $U(f+g) = U(f) + U(g)$ and $U(\alpha f) = \alpha U(f)$ for $\alpha \in \mathbb{F}$. For $f, g \in \mathscr{H}$, let $a := U(f+g), b := U(f)+U(g)$. Then $\langle a-b, a-b\rangle = \langle a, a\rangle - \langle a, b\rangle - \langle b, a\rangle + \langle b, b\rangle$.
(1) $\langle a, a\rangle = \langle f+g, f+g\rangle = \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle = \langle U(f), U(f)\rangle + \langle U(f), U(g)\rangle + \langle U(g), U(f)\rangle + \langle U(g), U(g)\rangle = \langle U(f)+U(g), U(f)+U(g)\rangle = \langle b, b\rangle$.
(2) $\langle a, b\rangle = \langle U(f+g), U(f)+U(g)\rangle = \langle f+g, f\rangle + \langle f+g, g\rangle = \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle = \langle U(f)+U(g), U(f)+U(g)\rangle = \langle b, b\rangle$.
(3) $\langle b, a\rangle = \langle U(f)+U(g), U(f+g)\rangle = \langle f, f+g\rangle + \langle g, f+g\rangle = \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle = \langle U(f)+U(g), U(f)+U(g)\rangle = \langle b, b\rangle$. Hence, $\langle a-b, a-b\rangle = 0$ and so $a-b=0$. Then $U(f+g) = U(f) = U(g)$.
And $\langle U(\alpha f) - \alpha U(f), U(\alpha f) - \alpha U(f)\rangle = \langle U(\alpha f), U(\alpha f)\rangle - \langle U(\alpha f), \alpha U(f)\rangle - \langle \alpha U(f), U(\alpha f)\rangle + \langle \alpha U(f), \alpha U(f)\rangle = \langle \alpha f, \alpha f\rangle - \overline{\alpha}\langle \alpha f, f\rangle - \alpha\langle f, \alpha f\rangle + \alpha\overline{\alpha}\langle f, f\rangle = 0$. Hence, $U(\alpha f) = \alpha U(f)$. Hence, U is linear. □

5.1 Reference

- C([0, 1]) is dense in L2([0, 1]); P5

Part II

Operators on Hilbert Spaces

Chapter 6

Elementary Properties and Examples

Let $\mathscr{H}$ and $\mathscr{K}$ be two Hilbert space over $\mathbb{F}$. Recall that a map $A : \mathscr{H} \to \mathscr{K}$ is a linear transformation if for all $x_1, x_2 \in \mathscr{H}$ and $\alpha, \beta \in \mathbb{F}$,

$$A(\alpha x_1 + \beta x_2) = \alpha A(x_1) + \beta A(x_2).$$

Then $N(A) = \{x \in \mathscr{H} : Ax = 0\}$ and $R(A)$ are subspaces of $\mathscr{H}$ and $\mathscr{K}$ respectively.

The collection of all linear operators from $\mathscr{H}$ to $\mathscr{K}$ forms a vector space $\mathcal{L}(\mathscr{H}, \mathscr{K})$ under pointwise addition and scalar multiplication of functions.

Let's recall the definition of bounded linear transformation and the norm of bounded linear transformation. The proof of the next proposition is similar to the proofs of the corresponding results for linear funtionals in proposition3.1.

Proposition 6.1

Let $\mathscr{H}$ and $\mathcal{K}$ be Hilbert spaces and $A : \mathscr{H} \to \mathcal{K}$ be a linear transformation. Then the following are equivalent
(1) A is continuous.
(2) A is continuous at 0.
(3) A is continuous at some point.
(4) There is a constant $c > 0$ such that $||Ah|| \leqslant c||h||$ for all h in $\mathscr{H}$.

As in definition3.2, if A is a bounded linear transformation, the norm of A defined as

$$||A|| = \sup_{x \in \mathscr{H}, ||x|| \leqslant 1} ||Ax||.$$

And then

$$\begin{aligned} ||A|| &= \sup_{||x||=1} ||Ax|| \\ &= \sup_{x \neq 0} \frac{||Ax||}{||x||} \\ &= \inf\{c > 0 : ||Ax|| \leqslant C||x||, x \in \mathscr{H}\}. \end{aligned}$$

Also,

$$||Ax|| \leqslant ||x||.$$

We denote the collection of all bounded linear operators from $\mathscr{H}$ to $\mathscr{K}$ by $\mathcal{B}(\mathscr{H}, \mathscr{K})$. It is a subspace of $\mathcal{L}(\mathscr{H}, \mathscr{K})$. They coincide when $\mathscr{H}$ and $\mathscr{K}$ are of finite dimension, of course. For $\mathscr{K} = \mathscr{H}$, $\mathcal{B}(\mathscr{H}) := \mathcal{B}(\mathscr{H}, \mathscr{H})$. Note that $\mathcal{B}(\mathscr{H}, \mathbb{F})$= all the bounded linear functionals on $\mathscr{H}$.

Proposition 6.2: (

) For $A, B \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, $A + B \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, and $||A + B|| \leqslant ||A|| + ||B||$.
(2) For $\alpha \in \mathbb{F}$, $A \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, then $\alpha A \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ and $||\alpha A|| = |\alpha| \cdot ||A||$.
(3) For $A \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, $B \in \mathscr{B}(\mathscr{K}, \mathscr{L})$, $BA \in \mathscr{B}(\mathscr{H}, \mathscr{L})$ and $||BA|| \leqslant ||B|| \cdot ||A||$.

Now we introduce some example of bounded linear transformation.

Exercise 6.1: II1 T9

(Schur test) Let $\{\alpha_{ij}\}_{i,j=1}^{\infty}$ be an infinite matrix such that $\alpha_{ij} \geqslant 0$ for all i, j and such that there are scalars $p_i > 0$ and $\beta, \gamma > 0$ with

$$\sum_{i=1}^{\infty} \alpha_{ij} p_i \leqslant \beta p_j,$$
$$\sum_{j=1}^{\infty} \alpha_{ij} p_j \leqslant \gamma p_i$$

for all $i, j \geqslant 1$. Show that there is an operator A on $l^2(\mathbb{N})$ with $\langle Ae_j, e_i \rangle = \alpha_{ij}$ and $||A||^2 \leqslant \beta\gamma$.

Proof. Let $\mathscr{H} = l^2(\mathbb{N})$ and $\{e_j\}$ be orthnormal basis of $\mathscr{H}$. Then $\forall x \in \mathscr{H}$, $x = \sum\limits_{j=1}^{\infty} \lambda_j e_j$. Define

$A : \mathscr{H} \to \mathscr{H}$ given by $Ae_j = \sum_{i=1}^{\infty} \alpha_{ij} e_i$, then $\langle Ae_j, e_i \rangle = \alpha_{ij}$ and $Ax = \sum_{i=1}^{\infty} (\sum_{j=1}^{\infty} \alpha_{ij}\lambda_j) e_i$. Hence,

$$
\begin{aligned}
||Ax||^2 = \langle Ax, Ax \rangle &= \langle \sum_{i=1}^{\infty} (\sum_{j=1}^{\infty} \alpha_{ij}\lambda_j) e_i, \sum_{i=1}^{\infty} (\sum_{j=1}^{\infty} \alpha_{ij}\lambda_j) e_i \rangle \\
&= \sum_{i=1}^{\infty} \langle (\sum_{j=1}^{\infty} a_{ij}\lambda_j) e_i, (\sum_{j=1}^{\infty} a_{ij}\lambda_j) e_i \rangle = \sum_{i=1}^{\infty} |\sum_{j=1}^{\infty} \alpha_{ij}|^2 \\
&= \sum_{i=1}^{\infty} |\sum_{j=1}^{\infty} \alpha_{ij}^{1/2} p_j^{1/2} \alpha_{ij}^{1/2} \frac{\lambda_j}{p_j^{1/2}}|^2 \\
&\leqslant \sum_{i=1}^{\infty} [(\sum_{j=1}^{\infty} \alpha_{ij} p_j)(\sum_{j=1}^{\infty} \alpha_{ij} \frac{\lambda_j^2}{p_j})] && \text{(Cauchy-Schwarz inequality)} \\
&\leqslant \sum_{i=1}^{\infty} (\gamma p_i \sum_{j=1}^{\infty} \alpha_{ij} \frac{\lambda_j^2}{p_j}) && (\sum_{j=1}^{\infty} \alpha_{ij} p_j \leqslant \gamma p_i) \\
&= \sum_{j=1}^{\infty} (\gamma \frac{\lambda_j^2}{p_j} \sum_{i=1}^{\infty} \alpha_{ij} p_i) \\
&\leqslant \sum_{j=1}^{\infty} (\gamma \frac{\lambda_j^2}{p_j} \beta p_j) && (\sum_{i=1}^{\infty} \alpha_{ij} p_i \leqslant \beta p_j) \\
&= \sum_{j=1}^{\infty} \gamma\beta\lambda_j^2 = \gamma\beta ||x||^2
\end{aligned}
$$

Hence, $||A||^2 \leqslant \beta\gamma$. □

Exercise 6.2: II1 T11

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, put $\alpha = [|a|^2 + |b|^2 + |c|^2 + |d|^2]^{1/2}$ and show that $||A|| = \frac{1}{2}(\alpha^2 + \sqrt{\alpha^4 - 4\delta^2})$, where $\delta^2 = \det A^* A$.

Proof. A^* is Hermitian conjugate of A. □

6.1 Reference

- matrix norm; P11

Chapter 7

The Adjoint of an Operator

Definition 7.1

If $\mathscr{H}$ and $\mathscr{K}$ are Hilbert spaces, a function $u : \mathscr{H} \times \mathscr{K} \to \mathbb{F}$ is a sesquilinear form if for $h, g \in \mathscr{H}$, $k, l \in \mathscr{K}$,and $\alpha, \beta \in \mathbb{F}$,
(1) $u(\alpha h + \beta g, k) = \alpha u(h, k) + \beta u(g, k)$;
(2) $u(h, \alpha k + \beta l) = \overline{\alpha} u(h, k) + \overline{\beta} u(h, l)$.

The prefix "sesqui" is used because the function is linear in one variable but (for $\mathbb{F} = \mathbb{C}$) only conjugate linear in the other. ("Sesqui" means "one-and -a-half.")

A sesquilinear form is bounded if there is a constant M such that $|u(h, k)| \leqslant M||h|| \cdot ||k|||$ for all h in $\mathscr{H}$ and k in $\mathscr{K}$. The constant M is called a bound for u.

Proposition 7.1: I

$A \in \mathcal{B}(\mathscr{H}, \mathscr{K})$, then $u(h, k) := \langle Ah, k \rangle$ is a bounded sesquilinear form.

Proof.

$$\begin{aligned} u(\alpha h + \beta g, k) &= \langle A(\alpha h + \beta g), k \rangle \\ &= \langle \alpha Ah + \beta Ag, k \rangle \\ &= \alpha \langle Ah, k \rangle + \beta \langle Ag, k \rangle \\ &= \alpha u(h, k) + \beta u(g, k), \end{aligned}$$

$$\begin{aligned} u(h, \alpha k + \beta l) &= \langle Ah, \alpha k + \beta l \rangle \\ &= \overline{\alpha} \langle Ah, k \rangle + \overline{\beta} \langle Ag, l \rangle \\ &= \overline{\alpha} u(h, k) + \overline{\beta} u(g, l), \end{aligned}$$

$$|u(h, k)| = |\langle Ah, k \rangle| \leqslant ||Ah|| \cdot ||k|| \leqslant ||h|| \cdot ||k||.$$

□

Also, if $B \in \mathcal{B}(\mathscr{H}, \mathscr{H})$, $u(h,k) = \langle h, Bk \rangle$ is a bounded sesquilinear form. Are there any more? Are these two forms related?

Theorem 7.1

If $u : \mathscr{H} \times \mathscr{H} \to \mathbb{F}$ is a bounded sesquilinear form with bound M, then there are unique operators A in $\mathscr{B}(\mathscr{H}, \mathscr{K})$ and B in $\mathscr{B}(\mathscr{K}, \mathscr{H})$ such that

$$u(h,k) = \langle Ah, k \rangle = \langle h, Bk \rangle \tag{7.1}$$

for all h in $\mathscr{H}$ and k in $\mathscr{K}$ and $||A||, ||B|| \leqslant M$.

Proof. □

Definition 7.2

If $A \in \mathcal{B}(\mathscr{H}, \mathscr{K})$, then the unique operator B in $\mathscr{B}(\mathscr{K}, \mathscr{H})$ satisfying (7.1) is call the adjoint of A and is denote by $B = A^*$.

Proposition 7.2

If $U \in \mathcal{B}(\mathscr{H}, \mathscr{K})$, then U is an isomorphism iff U is invertible and $U^{-1} = U^*$.

Proof. U is isomorphism, then for $h, g \in \mathscr{H}$,

$$< h, g >=< Uh, Ug >=< h, U^*Ug > .$$

So, $U^*U = I$. Since, U is surjection, U is invertible and $U^{-1} = U^*$.

Conversely, let U be invertible with $U^{-1} = U^*$. Then, u is a surjection, and

$$< Ux, Uy >=< x, U^*Uy >=< x, Iy >=< x, y > .$$

□

From now on we will examine and prove results for the adjoint of operators in $\mathcal{B}(\mathscr{H})$. Often, as in the next proposition, there are analogous results for the adjoint of operators in $\mathcal{B}(\mathscr{H}, \mathscr{K})$.

Proposition 7.3

If $A, B \in \mathscr{B}(\mathscr{H})$ and $\alpha \in \mathbb{F}$, then:
(1) $(\alpha A + B)^* = \overline{\alpha} A^* + B^*$.
(2) $(AB)^* = B^* A^*$.
(3) $A^{**} = (A^*)^* = A$.
(4) If A is invertible in $\mathcal{B}(\mathscr{H})$ and A^{-1} is its inverse, then A^* is invertible and $(A^*)^{-1} = (A^{-1})^*$.

Proof. (1) $< (\alpha A + B)h, g >= \alpha < Ah, g > + < Bh, g >= \alpha < h, A^*g > + < h, B^*g >=< h, (\overline{\alpha} A^* + B^*)g >$.
(2)

(3)
(4) □

Proposition 7.4

If $A \in \mathcal{B}(\mathscr{H})$, $||A|| = ||A^*|| = \sqrt{||A^*A||}$.

Definition 7.3: I

$A \in \mathcal{B}(\mathscr{H})$, then
(1) A is hermitian or self-adjoint if $A^* = A$.
(2) A is normal if $AA^* = A^*A$.

In the analogy between the adjoint and the complex conjugate, hermitian operators become the analogues of real numbers and, unitaries are the analogues of complex numbers of modulus 1. Normal operators, as we shall see, are the true analogues of complex numbers. Notice that hermitian and unitary operators are normal.

Proposition 7.5

If $\mathscr{H}$ is a $\mathbb{C}$-Hilbert space and $A \in \mathcal{B}(\mathscr{H})$, then A is self-adjoint iff $\langle Ah, h\rangle \in \mathbb{R}$ for all h in $\mathscr{H}$.

Proposition 7.6

If $A = A^*$, then

$$||A|| = \sup_{||h||=1} |\langle Ah, h\rangle|.$$

Corollary 7.1

If $A = A^*$ and $\langle Ah, h\rangle = 0$ for all h, then $A = 0$.

Proposition 7.7

If $\mathscr{H}$ is $\mathbb{C}$-Hilbert space and $A \in \mathcal{B}(\mathscr{H})$ such that $\langle Ah, h\rangle = 0$ for all h in $\mathscr{H}$, then $A = 0$.

Exercise 7.1: II2 T12

Let $\sum_{n=0}^{\infty} \alpha_n z^n$ be a power series with radius of convergence R, $0 < R \leqslant \infty$. If $A \in \mathscr{B}(\mathscr{H})$ and $||A|| < R$, show that there is an operator T in $\mathscr{B}(\mathscr{H})$ such that for any $h, g \in \mathscr{H}$, $\langle Th, g\rangle = \sum_{n=0}^{\infty} \alpha_n \langle A^n h, g\rangle$. [If $f(z) = \sum \alpha_n z^n$, the operator T is usually denoted by $f(A)$.]

Exercise 7.2: II2 T15

If A is a normal operator on $\mathscr{H}$, show that A is injective if and only if A has dense range. Give an example of an operator B such that $\ker B = (0)$ but $\operatorname{ran} B$ is not dense. Give an example of an operator C such that C is surjective but $\ker C \neq (0)$.

Chapter 8

Projections and Idempotents; Invariant and Reducing Subspaces

Definition 8.1

An idempotent is a bounded linear operator E so that $E^2 = E$. A orthogonal projection is an idempotent P such that $\mathrm{Ker}P = (\mathrm{ran}P)^{\perp}$. We actually use the word projection to refer only to orthogonal projections.

Remark. An non-orthogonal projection is an idempotent, but it is not a proejction. For example, take an basis (not)

Proposition 8.1

E is a idempotent.
(1) E is an idempotent $\Leftrightarrow$ $I - E$ is an idempotent.
(2) If E is an idempotent, $\mathrm{ran}(E) = \ker(I - E)$ and $\ker(E) = \mathrm{ran}(I - E)$ and $\mathrm{ran}(E)$ is a closed linear subspace of $\mathscr{H}$.
(3) If $\mathscr{M} = \mathrm{ran}E$ and $\mathscr{N} = \ker E$, then $\mathscr{M} \cap \mathscr{N} = (0)$ and $\mathscr{M} + \mathscr{N} = \mathscr{H}$.

Proof. (1) ($\Rightarrow$): $(I - E)^2 = I^2 - 2E + E^2 = I^2 - E = I - E$.
($\Leftarrow$): $E^2 = (I - E - I)^2 = (I - E)^2 - 2I(I - E) + I^2 = I^2 - 2E + E^2 - 2I^2 + 2E + I^2 = E$.
(2) For $h \in \ker(I - E)$, $(I - E)h = 0 \Rightarrow h = Eh \Rightarrow h \in \mathrm{ran}(E)$. For $h \in \mathrm{ran}(E)$, $h = Ea = E(Ea) = Eh$, then $(I - E)h = 0$.
For $h \in \mathrm{ran}(I - E)$, $h = (I - E)b = (I - E)^2 b = (I - E)h$, then $Eh = 0$. For $h \in \ker(E)$, $Eh = 0 \Rightarrow (I - E)h = h$.
For any linear bounded operator A, $\ker(A)$ is closed. Then $\mathrm{ran}(E) = \ker(I - E)$ is closed.
(3) For $h \in \mathrm{ran}(E) \cap \ker(E)$, $h = Ea = E(Ea) = Eh = 0$, $x = Ex + (I - E)x$. □

Lemma 8.1

If E is idempotent then: $h \in \mathrm{ran}(E) \Leftrightarrow h = Eh$.

Exercise 8.1: II3 T4

Let P and Q be projections. Show:
(1) $P+Q$ is a projection if and only if $\text{ran}P \perp \text{ran}Q$. If $P+Q$ is a projection, then $\text{ran}(P+Q) = \text{ran}P + \text{ran}Q$ and $\ker(P+Q) = \ker P \cap \ker Q$.
(2) PQ is a projection if and only if $PQ = QP$. If PQ is a projection, then $\text{ran}PQ = \text{ran}P \cap \text{ran}Q$ and $\ker PQ = \ker P + \ker Q$.

Proof. We claim if $\text{ran}P \perp \text{ran}Q \Leftrightarrow PQ = QP = 0$. In fact, if $\text{ran}P \perp \text{ran}Q$, then $\text{ran}P \subset (\text{ran}Q)^{\perp}$, which implies $(\text{ran}Q)^{\perp\perp} \subset (\text{ran}P)^{\perp}$. Since Q is projection, $\text{ran}Q$ is closed. Then $\text{ran}Q = (\text{ran}Q)^{\perp\perp}$. Since P is projection, $(\text{ran}P)^{\perp} = \ker P$. Then $\text{ran}Q \subset \ker P$. So $PQ = 0$. SImilarly, $QP = 0$.
If $PQ = 0$, then $\text{ran}Q \subset \ker P$, then $(\text{ran}Q)^{\perp} \supset (\ker P)^{\perp}$. Since P is projection, $(\ker P)^{\perp} = \text{ran}P$, then $\text{ran}P \subset (\text{ran}Q)^{\perp}$, then $\text{ran}P \perp \text{ran}Q$
(1) ($\Rightarrow$): Show that $PQ = QP = 0$.
Since $P+Q$ is a projection, it follows that $(P+Q)^2 = P+Q$, then for $h \in \mathscr{H}$,

$$\begin{aligned}
(P+Q)^2(h) &= (P+Q)(P(h)+Q(h)) \\
&= P(Ph+Qh) + Q(Ph+Qh) \\
&= P^2h + PQh + QPh + Q^2h \\
&= Ph + PQh + QPh + Qh \\
&= (P+Q)h + PQh + QPh.
\end{aligned} \tag{8.1}$$

This means that $0 = PQh + QPh$. Hence, $PQ = -QP$. For $x \in \text{ran}PQ$, then $-x \in \text{ran}(QP)$, then $x \in \text{ran}P$ and $-x \in \text{ran}(Q)$ then $x = Px$ and $-x = Q(-x) = -Qx$. So $x = Px = PQx = -QPx = -Qx = -x$, then $x = 0$. So $PQ = QP = 0$.
($\Leftarrow$): By (8.1), if $PQ = QP = 0$, then $(P+Q)^2 = P+Q$. Since P, Q are projection, it follows that $P^* = P$ and $Q^* = Q$. Then $(P+Q)^* = P^* + Q^* = P+Q$, which means that $P+Q$ is hermitian. SO $P+Q$ is projection.
If $P+Q$ is projection, then for $h \in \mathscr{H}$, $(P+Q)^2(h) = (P+Q)h$. Then $Ph + Qh \in \text{ran}(P+Q)$ and $(P+Q)h = Ph + Qh \in \text{ran}P + \text{ran}Q$, then $\text{ran}P + \text{ran}Q \subset \text{ran}(P+Q)$ and $\text{ran}P + \text{ran}Q \supset \text{ran}(P+Q)$. Hence, $\text{ran}(P+Q) = \text{ran}P + \text{ran}Q$. For $h \in \ker P \cap \ker Q$, $Ph = 0 = Qh$, then $(P+Q)h = 0$. Then $\ker P \cap \ker Q \subset \ker(P+Q)$. For $h \in \ker(P+Q)$, $(P+Q)h = 0$, then $Ph = -Qh$. Since $\text{ran}P \perp \text{ran}Q$, $0 = QPh = -Q^2h = -Q^h$ and $0 = -PQh = -P^2h = -Ph$. Then $Qh = Ph = 0$ and so $h \in \ker P \cap \ker Q$. Then $\ker(P+Q) = \ker P \cap \ker Q$.
(2) ($\Rightarrow$): PQ is a projection, then PQ is hermitian, then $PQ = (PQ)^* = Q^*P^* = QP$.
($\Leftarrow$): If $PQ = QP$, then $(PQ)^* = Q^*P^* = QP = PQ$ and so PQ is hermitian. And for $h \in \mathscr{H}$, $(PQ)^2h = (PQ)(PQh) = (PQ)(QPh) = P(Q^2)Ph = PQPh = PPQh = PQh$. Then PQ is idempotent. Hence, PQ is a projection.
Since $\text{ran}PQ \subset \text{ran}P$ and $\text{ran}PQ = \text{ran}QP \subset \text{ran}Q$, it follows that $\text{ran}PQ \subset \text{ran}P \cap \text{ran}Q$. On the other hand, for $h \in \text{ran}P \cap \text{ran}Q$, $Px = Qx = x$, then $PQx = x$. So $\text{ran}P \cap \text{ran}Q \subset \text{ran}PQ$. Hence, $\text{ran}PQ = \text{ran}P \cap \text{ran}Q$. For $h \in \ker PQ$, then either $x \in \ker Q$ or $x \in (\ker Q)^{\perp}$. If $x \in \ker Q$, then $x = 0 + x \in \ker P + \ker Q$. If $x \in (\ker Q)^c$, then $Qx \neq 0$. Since $0 = PQx$, $Qx \in \ker P$. Then $x = Qx + x - Qx$. Since $Q(x - Qx) = 0$, $x = Qx + x - Qx \in \text{ran}P + \text{ran}Q$. So $\ker PQ \subset \ker P + \ker Q$. For $h \in \ker P + \ker Q$, we can write x as $x = u + v$ where $u \in \ker P$ and $v \in \ker Q$, then $(PQ)h = PQ(u+v) = PQu + PQv = PQu = QPu = 0$. So $\ker P + \ker Q \subset \ker PQ$. Hence, $\ker PQ = \ker P + \ker Q$. $\square$

Exercise 8.2: II3 T6

If P and Q are projection, then the following statements are equivalent.
(1) $P-Q$ is a projection.
(2) $\text{ran}Q \subseteq \text{ran}P$.
(3) $PQ = Q$.
(4) $QP = Q$.
If $P-Q$ is a projection, then $\text{ran}(P-Q) = (\text{ran}P) \ominus (\text{ran}Q)$ and $\ker(P-Q) = \text{ran}Q + \ker P$.

Proof. $(2) \Rightarrow (3)$: For $h \in \mathscr{H}$, since $\text{ran}Q \subseteq \text{ran}P$, it follows that $Qh \in \text{ran}P$. Then $\exists x \in \mathscr{H}$, such that $Qh = Px$. Then $PQh = PPx = Px = Qh$. Hence, $PQ = Q$.
$(3) \Leftrightarrow (4)$: Suppose $PQ = Q$, then $QP = Q^*P^* = (PQ)^* = Q^* = Q$. Suppose $QP = Q$, then $PQ = P^*Q^* = (QP)^* = Q^* = Q$.
$(4) \Rightarrow (1)$: Since $(P-Q)^* = P^* - Q^* = P - Q$, it follows that $P-Q$ is hermitian. Since $(P-Q)^2 = P^2 - PQ - QP + Q^2 = P - Q - Q + Q = P - Q$, it follows that $P-Q$ is idempotent. Hence, $P-Q$ is projection.
Now, If $P-Q$ is projection, we will show that $\text{ran}(P-Q)$ =($\text{ran}P$) $\ominus$ ($\text{ran}Q$)=$\text{ran}P \cap (\text{ran}Q)^\perp = \text{ran}P \cap \ker Q$ and $\ker(P-Q) = \text{ran}Q + \ker P$.
For $h \in \text{ran}(P-Q)$, since $P-Q$ is projection, $h = (P-Q)h = Ph - Qh$. By (3), $h = Ph - PQh = P(h - Qh) \in \text{ran}P$. And by (4), $Qh = QPh - Q^2h = Qh - Qh = 0$, then $h \in \ker Q$. Hence, $\text{ran}(P-Q) \subseteq \text{ran}P \cap \ker Q$. For $h \in \text{ran}P \cap \ker Q$, $h = Ph$ as P is projection and $Qh = 0$, then $h = Ph - Qh = (P-Q)h \in \text{ran}(P-Q)$. Then $\text{ran}P \cap \ker Q \subseteq \text{ran}(P-Q)$. Hence, $\text{ran}(P-Q) = \text{ran}P \cap \ker Q$.
For $h \in \ker(P-Q)$, then $(P-Q)h = Ph - Qh = 0$ and either $h \in \ker P$ or $h \in (\ker P)^\perp$. If $h \in \ker P$, then $h = 0 + h$. If $h \in (\ker P)^\perp$, $h \in \text{ran}P$. Then $h = Ph = Qh + Ph - Qh$. Since $P(Ph - Qh) = Ph - Qh = 0$, $Ph - Qh \in \ker P$. Hence $h \in \text{ran}Q + \ker P$ and $\ker(P-Q) \subseteq \text{ran}Q + \ker P$. For $h \in \text{ran}Q + \ker P$, we can write $h = u + v$ where $u \in \text{ran}Q$ and $v \in \ker P$. Then $u = Qu$ as Q is projection and $Pv = 0$. Then $(P-Q)h = (P-Q)(u+v) \overset{P,Q\text{is linear}}{=} (P-Q)u + (P-Q)v = Pu - u + 0 - Qv = Pu - u - Qv$. Since $Qv = QPv = 0$ and $Pu = PQu = Qu = u$, it follows that $(P-Q)h = 0$. Then $h \in \ker(P-Q)$ and so $\text{ran}Q + \ker P \subseteq \ker(P-Q)$. Hence, $\ker(P-Q) = \text{ran}Q + \ker P$.

□

8.1 Reference

- Operators in Hilbert spaces

Chapter 9

Compact Operators

Let ball $\mathscr{H}$ denote the closed unit ball $\{h \in \mathscr{H} : ||h|| \leqslant 1\}$ in $\mathscr{H}$. Now let me explain the definition of a few symbols that appear in this chapter.

Symbol	**Name**	**Definition**
$\mathscr{B}(\mathscr{H}, \mathscr{K})$	Bounded Operators	$\|\|T\|\| < \infty$ i.e. T(ball $\mathscr{H}$) is bounded
$\mathscr{B}_0(\mathscr{H}, \mathscr{K})$	Compact Operators	cl[T(ball $\mathscr{H}$)] is compact
$\mathscr{B}_{00}(\mathscr{H}, \mathscr{K})$	Finite rank Operators	ran(T) is finite dimensional

Proposition 9.1

In a normed space $\mathscr{X}$ and $S \subseteq \mathscr{X}$, the following "compactness" are all equivalent:
(1) S is compact.
(2) S is sequentially compact.
(3) S is complete and totally bounded.

Proposition 9.2

$T \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ is compact iff for every bounded sequence $\{h_n\}$ in $\mathscr{H}$, $\{Th_n\}$ has a convergent subsequence in $\mathscr{K}$.

Proof. 1 $\square$

Proposition 9.3

(1) $\mathscr{B}_0(\mathscr{H}, \mathscr{K}) \subseteq \mathscr{B}(\mathscr{H}, \mathscr{K})$.
(2) $\mathscr{B}_0(\mathscr{H}, \mathscr{K})$ is a linear space and if $\{T_n\} \subseteq \mathscr{B}_0(\mathscr{H}, \mathscr{K})$ and $T \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ such that $||T_n - T|| \to 0$, then $T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$ ($\mathscr{B}_0(\mathscr{H}, \mathscr{K})$ is closed).
(3) If $A \in \mathscr{H}$, $B \in \mathscr{K}$ and $T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$, then TA and $BT \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$.

Proof. (1) For $T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$, cl[$T$(ball$\mathscr{H}$)] is compact, then cl[$T$(ball$\mathscr{H}$)] is totally bounded. Hence, $||T|| = \sup_{||h||=1} ||Th|| = \sup_{h \in \text{ball } \mathscr{H}} ||Th|| < \infty$ and so T is bounded.
(2) To show $T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$, we should show that cl[T(ball $\mathscr{H}$)] is compact. Since $\mathscr{K}$ is complete, it follows that we should show that T(ball $\mathscr{H}$) is totally bounded and so cl[T(ball $\mathscr{H}$)] is totally bounded. Since $||T_n - T|| \to 0$, for $\epsilon > 0$, $\exists n \in \mathbb{N}$ such that $||T_n - T|| < \frac{\epsilon}{3}$. Since T_n is compact,

there are vectors $h_1, \ldots, h_m$ in ball $\mathscr{H}$ such that $T_n(\text{ball } \mathscr{H}) \subseteq \cup_{j=1}^m B(T_n h_j, \frac{\epsilon}{3})$. So if $h \in$ ball $\mathscr{H}$, there exists an h_j such that $||T_n h_j - T_n h|| < \frac{\epsilon}{3}$. Thus:

$$\begin{aligned} ||Th_j - Th|| &\leqslant ||Th_j - T_n h_j|| + ||T_n h_j - T_n h|| + ||T_n h - Th|| \\ &< ||T - T_n|| + \frac{\epsilon}{3} + ||T - T_n|| \\ &< \epsilon. \end{aligned}$$

Hence, $T(\text{ball } \mathscr{H}) \subseteq \cup_{i=1}^m B(Th_j, \epsilon)$. Hence, T (ball $\mathscr{H}$) is totally bounded.
(3) □

Proposition 9.4

If $T \in \mathscr{B}(\mathscr{H}, \mathscr{K})$. Then T is compact iff there is a sequence $\{T_n\}$ of operators of finite rank such that $||T - T_n|| \to 0$.

Lemma 9.1

$T \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, then dim(ran$T$)=dim(ran$T^*$). In particular, T is finite rank iff T^* is finite rank.

Proposition 9.5

If $T \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, then T is compact iff T^* is compact.

Proof. ($\Rightarrow$): If T is compact, then there is sequence of $\{T_n\}$ of operators of finite rank such that $||T - T_n|| \to 0$. Since $||(T_n - T)^*|| = ||T_n - T||$, it follows that $||T_n^* - T^*|| \to 0$. Since $T_n^* \in \mathscr{B}_{00}(\mathscr{H}, \mathscr{K})$ and $T^* \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, by proposition9.4, $T^* \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$.
($\Leftarrow$): Since $T = (T^*)^*$, result is clear by the same proof above. □

Definition 9.1

If $A \in \mathscr{B}(\mathscr{H})$, a scalar α is an eigenvalue of A if ker$(A - \alpha) \neq (0)$. If h is a nonzero vector in ker$(A - \alpha)$, h is called an eigenvector for α; thus $Ah = \alpha h$. Let $\sigma_p(A)$ denote the set of eigenvalues of A.

Proposition 9.6

If $T \in \mathscr{B}_0(\mathscr{H})$, $\lambda \in \sigma_p(A)$ and $\lambda \neq 0$, then the eigenspave ker$(T - \lambda)$ is finite dimensional.

9.1 Homework

Exercise 9.1: II4 T4

Show that an idempotent is compact if and only if it has finite rank.

Proof. (⇒): We need to show that ran(E) is finite dimensional.
If ran$(E) = \{0\}$, obviously ran$E$ is finite dimensional. If ran$(E) \neq \{0\}$, for $0 \neq h \in$ ran$E$, $h = Eh$ since E is idempotent. Then $1 \in \sigma_p(A)$, since E is compact, by proposition9.6, ker$(E - I) =$ ran$E$ is finite dimensional.
(⇐): Since E is finite rank, it follows that ran$E$ is finite dimensional. Since cl[E(ball $\mathscr{H}$)] is a closed subset of ran$E$ and ran$E$ is Hausdorff, it follows that cl[E(ball $\mathscr{H}$)] is compact. Hence E is compact. □

Exercise 9.2: II4 T8

If $h, g \in \mathscr{H}$, define $T : \mathscr{H} \to \mathscr{H}$ by $Tf = \langle f, h\rangle g$.
(1) Show that T has rank 1 [that is, dim(ran$T$)= 1].
(2) Moreover, every rank 1 operator can be so represented.
(3) Show that if T is a finite rank operator, then there are orthonormal vectors $e_1, ..., e_n$ and vectors $g_1, ..., g_n$ such that $Th = \sum_{j=1}^{n} \langle h, e_j\rangle g_j$ for all h in $\mathscr{H}$.
(4) In this case show that T is normal if $g_j = \lambda_j e_j$ for some scalars $\lambda_1, ..., \lambda_n$.
(5) Find $\sigma_p(T)$.

Proof. (1) If $h, g \neq 0 \in \mathscr{H}$, then $Th = \langle h, h\rangle g \neq 0$. Hence, dim(ran$T$)$\geqslant 1$. For $Tf_1, Tf_2 \neq 0 \in$ ran$T$, $\langle f_2, h\rangle Tf_1 - \langle f_1, h\rangle Tf_2 = \langle f_2, h\rangle\langle f_1, h\rangle g - \langle f_1, h\rangle\langle f_2, h\rangle g = 0$. Then Tf_1 and Tf_2 is linear dependent and so dim(ran$T$)$\leqslant 1$. Hence, dim(ran$T$)= 1.
(2) If dim(ran$T$)= 1, let g be a basis of ran(T) and $||g|| = 1$, then for $f \in \mathscr{H}$, assume $Tf = \alpha g$, then $\langle Tf, g\rangle = \langle \alpha g, g\rangle = \alpha$. Then $Tf = \langle Tf, g\rangle g$. Since $Tf \in \mathscr{H}$, let $Tf = h$ and $Tf = \langle h, g\rangle g$.
(3) □

9.2 Reference

- lectures notes from tqft
- lecture notes from washington
- lecture notes from mit
- Compact Operators

Part III

Banach Spaces

Chapter 10

Elementary Properties and Examples

Definition 10.1: I

$\mathscr{X}$ is a vector space over $\mathbb{F}$, a seminorm is a function $p : \mathscr{X} \to [0, \infty)$ having the properties:
(1) $p(x + y) \leqslant p(x) + p(y)$ for all x, y in $\mathscr{X}$.
(2) $p(\alpha x) = |\alpha| p(x)$ for all α in $\mathbb{F}$ and x in $\mathscr{X}$.

Exercise 10.1: III1 T4

If $1 \leqslant p \leqslant \infty$ and $(x_1, x_2) \in \mathbb{R}^2$, define $||x||_p \equiv (|x_1|^p + |x_2|^p)^{1/p}$ and $||x||_\infty \equiv \sup\{|x_1|, |x_2|\}$, graph $\{x \in \mathbb{R}^2 : ||x||_p = 1\}$. Note that if $1 < p < \infty$, $||x||_p = ||y||_p = 1$ and $x \neq y$, then for $0 < t < 1$, $||tx + (1-t)y||_p < 1$. The same cannot be said for $p = 1, \infty$.

Proof. By Minkowski inequality, $||tx + (1-t)y||_p \leqslant ||tx||_p + ||(1-t)y||_p = |t|||x||_p + |1-t|||y||_p = t + 1 - t = 1$. If $p = 1$, for $x = (-1, 0), y = (0, 1), t = \frac{1}{2}$, then $||tx + (1-t)y||_1 = ||(-t, 0) + (0, 1 - t)||_1 = ||(-\frac{1}{2}, 0) + (0, \frac{1}{2})||_1 = ||(-\frac{1}{2}, \frac{1}{2})||_1 = 1$. If $p = \infty$, for $x = (-1, 1), y = (1, 1), t = \frac{1}{2}$, then $||tx + (1-t)y||_\infty = ||(-t, t) + (1-t, 1-t)||_\infty = ||(1-2t, 1)||_\infty = ||(0, 1)||_\infty = 1$. □

Exercise 10.2: III1 T5

Let $c =$ the set of all sequences $\{\alpha_n\}_1^\infty$, α_n in $\mathbb{F}$, such that $\lim \alpha_n$ exists. Show that c is closed subspace of l^∞ and hence is a Banach space.

Proof. We know that $c = \{(\alpha_n)_{n \in \mathbb{N}_+} | \alpha_n \in \mathbb{C}$ and the sequences converges$\}$ and $l^\infty = \{(\alpha_n)_{n \in \mathbb{N}_+} | \alpha_n \in \mathbb{C}, \sup_n |\alpha_n| < \infty\}$. Then, for (α_n) in c, $\exists M > 0$ s.t. $a_n \leqslant M$ for all n. Then, c is a subspace of l^∞. Now, we show that c is closed. Let $x = (\alpha_1, \alpha_2, \ldots) \in l^\infty$ and $\{x^n\}$ is a sequences in c convering to x in the $|\cdot||_\infty$ norm. For $n \in \mathbb{N}$, write $x^n = (x_1^n, x_2^n, \ldots)$ so that x_i^n is the i i-th term of the sequence x^n in c. We claim that $x \in c$. Let $\epsilon > 0$. As $x^n \underset{n\to\infty}{\longrightarrow} x$, there exists $N \in \mathbb{N}$ such that $||x^N - x||_\infty < \frac{\epsilon}{3}$. As x^N is in c, it follows that x^N is Cauchy, so there exists $K \in \mathbb{N}$ such that $|x_i^N - x_j^N| < \frac{\epsilon}{3}$ for all

$i, j \geqslant K$. Then for such i, j, we have

$$\begin{aligned}|x_i - x_j| &\leqslant |x_i - x_i^N| + |x_i^N - x_j^N| + |x_j^N - x_j| \\ &\leqslant ||x - x^N||_\infty + |x_i^N - x_j^N| + ||x^N - x||_\infty \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}.\end{aligned}$$

Then x is Cauchy and so converges. Hence, x is in c. So c is a closed subspace of l^∞ and so a Banach space. □

10.1 Reference

- lp norm strictly convex when $1 < p < \infty$
- c is a closed subspace of l^∞

Chapter 11

Linear Operators on Normed Space

Exercise 11.1: III2 T1

Show that for $\mathscr{B}(\mathscr{X},\mathbb{F}) \neq (0)$, $\mathscr{B}(\mathscr{X},\mathscr{Y})$ is a Banach space if and only if $\mathscr{Y}$ is a Banach space.

Proof. ($\Rightarrow$): Let $x_0 \in \mathscr{X}$ with $||x_0|| = 1$, then by Hahn-Banach Theorem, there exists $f \in \mathscr{X}^*$ such that $f(x_0) = ||x_0|| = 1$. ($\Leftarrow$): We have to prove that $\mathscr{B}(\mathscr{X},\mathscr{Y})$ is complete. Let (T_n) be a Cauchy sequence in $\mathscr{B}(\mathscr{X},\mathscr{Y})$. For each $x \in \mathscr{X}$, we have

$$||T_n x - T_m x|| = ||(T_n - T_m)x|| \leqslant ||T_n - T_m|| \cdot ||x||,$$

which shows that $(T_n x)$ is a Cauchy sequence in $\mathscr{Y}$. Since $\mathscr{Y}$ is complete, there is an unique $\mathscr{Y} \in \mathscr{Y}$ such that $T_n x \underset{n\to\infty}{\longrightarrow} y$. Define $T : \mathscr{X} \to \mathscr{Y}$ given by $Tx = y$. Then T is well-defined and linear. We show that $||T_n - T|| \underset{n\to\infty}{\longrightarrow} 0$ and T is bounded, i.e. $\sup_{||\mathscr{X}||=1} ||T_n x - Tx|| \underset{n\to\infty}{\longrightarrow} 0$ and $\sup_{||\mathscr{X}||=1} ||Tx|| < \infty$.

For any $\epsilon > 0$, since (T_n) is Cauchy, it follows that there exists $N_1 > 0$ such that $||T_n x - T_m x|| \leqslant ||T_n - T_m|| < \epsilon/2$, for all $n, m > N_1$. Since $T_n x \underset{n\to\infty}{\longrightarrow} Tx$, there exists $N_2 > 0$ such that $||T_m x - Tx|| < \epsilon/2$, for all $m > N_2$. Let $N = \max\{N_1, N_2\}$ and $m_0 > N$, then for $||\mathscr{X}|| = 1$,

$$||T_n x - Tx|| \leqslant ||T_n x - T_{m_0} x|| + ||T_{m_0} x - Tx|| < \epsilon.$$

Hence, $||T_n - T|| \to 0, n \to \infty$. And

$$||Tx|| \leqslant ||T_{m_0} x|| + ||Tx - T_{m_0} x|| \leqslant ||T_{m_0} x|| + \epsilon.$$

Since T_{m_0} is bounded, T is bounded. Hence, $T \in \mathscr{B}(\mathscr{X},\mathscr{Y})$ and so $\mathscr{B}(\mathscr{X},\mathscr{Y})$ is complete. $\square$

Exercise 11.2: III2 T4

If $(\mathscr{X}, \Omega, \mu)$ is a σ-finite measure space and $\phi \in L^\infty(\mathscr{X}, \Omega, \mu)$, define $M_\phi : L^p(\mathscr{X}, \Omega, \mu) \to L^p(\mathscr{X}, \Omega, \mu)$, $1 \leqslant p \leqslant \infty$, by $M_\phi f = \phi f$ for all f in $L^p(\mathscr{X}, \Omega, \mu)$. Then $M_\phi \in \mathscr{B}(L^p(\mathscr{X}, \Omega, \mu))$ and $||M_\phi|| = ||\phi||_\infty$.

11.1 Reference

- $\mathscr{Y}$ is a Banach space, then $B(\mathscr{X}, \mathscr{Y})$ is a Banach space. P110
- $B(\mathscr{X}, \mathscr{Y})$ is a Banach space, then $\mathscr{Y}$ is Banach space
-

Chapter 12

Finite Dimensional Normed Spaces

Exercise 12.1: III3 T2

If $\mathscr{X}$ is a finite dimensional vector space over $\mathbb{F}$, define $||x||_\infty \equiv \max\{|x_j| : 1 \leqslant j \leqslant d\}$. Show that $||\cdot||_\infty$ is a norm.

Proof. It is clearly that $||\cdot||_\infty$ is well defined. Now, we show that $||\cdot||$ satisfys norm axioms.
(1) $||x|| \geqslant 0$ and $||x||_\infty = 0 \Leftrightarrow x = 0$. In fact, $||x||_\infty \geqslant 0$ is clear and

$$\begin{aligned} ||x||_\infty = \max_j |x_j| = 0 &\Leftrightarrow 0 \leqslant |x_j| \leqslant 0, \forall 1 \leqslant j \leqslant d \\ &\Leftrightarrow x = 0 \end{aligned}$$

(2) For $\alpha \in \mathbb{F}$, $||\alpha x||_\infty = |\alpha| ||x||_\infty$. In fact,

$$\begin{aligned} ||\alpha x||_\infty = \max_j |\alpha x_j| &= \max_j |\alpha||x_j| \\ &= |\alpha| \max_j |x_j| \\ &= |\alpha| ||x|| \end{aligned}$$

(3) $||x+y||_\infty \leqslant ||x||_\infty + ||y||_\infty$. In fact,

$$\begin{aligned} ||x+y||_\infty &= \max\{|x_j + y_j| : 1 \leqslant j \leqslant d\} \\ &\leqslant \max\{|x_j| + |y_j| : 1 \leqslant j \leqslant d\} \\ &\leqslant \max_j |x_j| + \max_j |y_j| \\ &= ||x||_\infty + ||y||_\infty \end{aligned}$$

$\square$

Chapter 13

Quotients and Products of Normed Spaces

Exercise 13.1: III4 T1

Let $\mathscr{X}$ be a normed space, let $\mathscr{M}$ be a linear manifold in $\mathscr{X}$, and let $Q : \mathscr{X} \to \mathscr{X}/\mathscr{M}$ be the natural map $Qx = x + \mathscr{M}$. show $||x + \mathscr{M}|| = \inf\{||x + y|| : y \in \mathscr{M}\}$ is a norm in $\mathscr{X}/\mathscr{M}$.

Proof. Firstly, we show that the $||x + \mathscr{M}||$ is well defined i.e. if $x, x' \in \mathscr{X}$ such that $Q(x) = Q(x')$ then $||x + \mathscr{M}|| = ||x' + \mathscr{M}||$. If $Q(x) = Q(x')$, by Q is linear transformation, $Q(x - x') = 0$ and so $x - x' \in \ker Q = \mathscr{M}$. So, $x - x' + \mathscr{M} = \mathscr{M}$. Then

$$\begin{aligned}
||x + \mathscr{M}|| &= \inf\{||x + y|| : y \in \mathscr{M}\} \\
&= \inf\{||x + (z - (x - x'))|| : z \in x - x' + \mathscr{M}\} \\
&= \inf\{||x' + z|| : z \in \mathscr{M}\} \\
&= ||x' + \mathscr{M}||.
\end{aligned}$$

Secondly, we show that the $||x + \mathscr{M}||$ satisfys norm axioms.
(1) $||x + \mathscr{M}|| \geqslant 0$ and $||x + \mathscr{M}|| = 0 \Leftrightarrow x + \mathscr{M} = 0_{x/\mathscr{M}} = \mathscr{M}$. In fact, $||x + \mathscr{M}|| \geqslant 0$ is clear and

$$\begin{aligned}
x + \mathscr{M} = 0_{\mathscr{X}/\mathscr{M}} = \mathscr{M} \Rightarrow ||x + \mathscr{M}|| &= ||\mathscr{M}|| \\
&= \inf_{y \in \mathscr{M}} ||y|| \\
&= 0 \qquad\qquad (0 \in \mathscr{M}, ||0|| = 0).
\end{aligned}$$

If $||x + \mathscr{M}|| = \inf_{y \in \mathscr{M}} ||x + y|| \overset{z=-y}{=} \inf_{z \in \mathscr{M}} ||x - z|| = 0$, then For each $n \in \mathbb{N}$, $\exists z_n \in \mathscr{M}$ s.t. $0 \leqslant ||x - z_n|| \leqslant \frac{1}{n}$. Then $\lim_{n \to \infty} ||x - z_n|| = 0$ and so $z_n \underset{n \to \infty}{\longrightarrow} x$. Since $\mathscr{M}$ is closed, it follows that $x \in \mathscr{M}$. Then $x + \mathscr{M} = Q(x) = 0_{\mathscr{X}/\mathscr{M}} = \mathscr{M}$.
(2) For $\alpha \in \mathbb{F}$, $||\alpha(x + \mathscr{M})|| = |\alpha| \cdot ||x + \mathscr{M}||$. In fact, equality clearly holds when $\alpha = 0$, if $\alpha \neq 0$

$$\begin{aligned}
||\alpha(x + \mathscr{M})|| = ||\alpha x + \mathscr{M}|| = \inf_{y \in \mathscr{M}} ||\alpha x + y|| &\overset{z=\frac{y}{\alpha}}{=} \inf_{z \in \mathscr{M}} ||\alpha x + \alpha z|| \\
&= |\alpha| \inf_{z \in \mathscr{M}} ||x + z|| = |\alpha| \cdot ||x + \mathscr{M}||
\end{aligned}$$

(3) For $x+\mathscr{M}, y+\mathscr{M}$, $||x+y+\mathscr{M}|| \leqslant ||x+\mathscr{M}|| + ||y+\mathscr{M}||$. In fact,

$$\begin{aligned} ||x+y+\mathscr{M}|| &= \inf_{z\in\mathscr{M}} ||x+y+z|| \overset{z=z_1+z_2}{=} \inf_{z_1,z_2\in\mathscr{M}} ||x+z_1+y+z_2|| \\ &\leqslant \inf_{z_1,z_2\in\mathscr{M}} (||x+z_1|| + ||y+z_1||) \\ &= \inf_{z_1\in\mathscr{M}} ||x+z_1|| + \inf_{z_2\in\mathscr{M}} ||y+z_1|| \\ &= ||x+\mathscr{M}|| + ||y+\mathscr{M}||. \end{aligned}$$

□

Exercise 13.2: III4 T3

Show that if (X,d) is a metric space and $\{x_n\}$ is a Cauchy sequence such that there is a subsequence $\{x_{n_k}\}$ that converges to x_0, then $x_n \to x_0$.

Proof. Since the subsequence $\{x_{n_k}\}$ converges to $x \in X$, we know that for all $\epsilon > 0$, there exists $N_1 \in \mathbb{N}$ such that if $n_k \geqslant N_1$ then $d(x_{n_k}, x) < \frac{\epsilon}{2}$. Moreover, since $\{x_n\}$ is Cauchy, for all $\epsilon > 0$, there exists $N_2 \in \mathbb{N}$ such that if $n, m \geqslant N_2$, then $d(x_n, x_m) < \frac{\epsilon}{2}$. Choose $s \in \mathbb{N}$ such that $n_s > \max\{N_1, N_2\}$. Then, if $n \geqslant \max\{N_1, N_2\}$, we have that $d(x_n, x) \leqslant d(x_n, x_{n_s}) + d(x_{n_s}, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$, that is, x_n converges to x. □

13.1 Reference

- Quotient Norm is Norm
- III4 T3

Chapter 14

Linear Functional

Exercise 14.1: III5 T2

Show that $\mathscr{X}^*$ is a normed space.

Proof. For $f \in \mathscr{X}^*$, $||f|| = \sup_{||x||=1} |fx|$. It suffices to show that operator norm is a norm.
(1) $||f|| \geqslant 0$ and $||f|| = 0 \Leftrightarrow f = 0$. In fact, since $|fx| \geqslant 0$, it follows that $||f|| \geqslant 0$. And, the zero operator indeed has norm 0 (because $||fx|| = 0$ for all x). On the other hand, suppose that $Tx = 0$ for all $||x|| = 1$. Then rescaling tells us that $0 = fx' = ||x'||f(\frac{x'}{||x'||}) = 0$ for all $x' \neq 0$, so f is indeed the zero operator.
(2) For $\alpha \in \mathbb{F}$, $||\alpha x|| = |\alpha| \cdot ||x||$. In fact,

$$||\alpha f|| = \sup_{||x||=1} ||\alpha f x|| = \sup_{||x||=1} |\alpha| \cdot ||fx|| = |\alpha| \sup_{||x||=1} ||fx|| = |\alpha| \cdot ||f||.$$

(3) For $f, g \in \mathscr{X}^*$, $||f+g|| \leqslant ||f|| + ||g||$. In fact, for x with $||x|| = 1$,

$$\begin{aligned} ||f+g|| = \sup_{||x||=1} ||(f+g)x|| = \sup_{||x||=1} ||fx+gx|| &\leqslant \sup_{||x||=1} ||fx|| + ||gx|| \\ &\leqslant \sup_{||x||=1} ||fx|| + \sup_{||x||=1} ||gx|| = ||f|| + ||g||. \end{aligned}$$

Hence, the operator norm is a norm and so $\mathscr{X}^*$ is a normed space. $\square$

14.1 Reference

- Show that X* is a normed space
- operator norm is a norm
- The operator norm is a norm

Chapter 15

The Hahn-Banach Theorem

If there is a "fundamental theorem of functional analysis," it is the Hahn-Banach theorem. The theorem is somewhat abstract-looking at first, but its importance will be clear after studying some of its corollaries.

To state and prove the Hahn-Banach Extension Theorem, we first work in the setting $\mathbb{F} = \mathbb{R}$, then extend the results to the complex case.

Definition 15.1

If $\mathscr{X}$ is a vector space, a sublinear functional is a function $q : \mathscr{X} \to \mathbb{R}$ such that
(1) $q(x+y) \leqslant q(x) + q(y)$ for all x, y in $\mathscr{X}$;
(2) $q(\alpha x) = \alpha q(x)$ for x in $\mathscr{X}$ and $\alpha \geqslant 0$.

Remark. Note that every seminorm is a sublinear functional, but not conversely.

Theorem 15.1: The Hahn-Banach Theorem on $\mathbb{R}$

Let $\mathscr{X}$ be a vector space over $\mathbb{R}$ and let q be a sublinear functional on $\mathscr{X}$. If $\mathscr{M}$ is a linear mainfold in $\mathscr{X}$ and $f : \mathscr{M} \to \mathbb{R}$ is a linear funcional such that $f(x) \leqslant q(x)$ for all x in $\mathscr{M}$, then there is a linear functional $F : \mathscr{X} \to \mathbb{R}$ such that
(1) $F|\mathscr{M} = f$ (F extends f)
(2) $F(x) \leqslant q(x)$ for all x in $\mathscr{X}$ (F is dominated by q).

The proof will invoke Zorns Lemma, a result that is equivalent to the axiom of choice (as well as ordering principal and the Hausdorff maximality principal). A partial order $\preceq$ on a set is a relation that is reflexive, symmetric and transitive; that is
(1) $x \preceq x$ for all $x \in S$;
(2) for $x, y \in S$, if $x \preceq y$ and $y \preceq x$, then $x = y$;
(3) for $x, y, z \in S$, if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

We call S, or more precisely, $(S, \preceq)$ a patially ordered set or poset. A subset T of S is totally ordered, if for each $x, y \in T$ either $x \preceq y$ or $y \preceq x$. A totally ordered subset T is often called a chain. An upper bound z for a chain T is an element $z \in S$ such that $t \preceq z$ for all $t \in T$. A maximal element for S is $w \in S$ that has no successor; that is there does not exist an $s \in S$ such that $s \neq w$ and $w \preceq s$.

Theorem 15.2: Zorn's Lemma

Suppose S is a partially ordered set. If every chain in S has an upper bound, then S has a maximal element.

Now, let us prove the Hahn-Banach Theorem on $\mathbb{R}$.

Proof. The idea is to show that the extension can be done one dimension at a time and then infer the existence of an extension to the whole space by appeal to Zorn's lemma. We may of course assume $\mathscr{M} \neq \mathscr{X}$. So, fix a vector $x \in \mathscr{X} \setminus \mathscr{M}$ and consider the subspace $\mathscr{M} + \mathbb{R}x \subset \mathscr{X}$. For any $m_1, m_2 \in \mathscr{M}$, by hypothesis,

$$f(m_1) + f(m_2) = f(m_1 + m_2) \leqslant q(m_1 + m_2) \leqslant q(m_1 - x) + q(m_2 + x).$$

Rearranging gives, for $m_1, m_2 \in \mathscr{M}$,

$$f(m_1) - q(m_1 - x) \leqslant q(m_2 + x) - f(m_2)$$

and thus

$$\sup_{m \in \mathscr{M}} \{f(m) - q(m - x)\} \leqslant \inf_{m \in \mathscr{M}} \{q(m + x) - f(m)\}.$$

Now choose any real number λ satisfying

$$\sup_{m \in \mathscr{M}} \{f(m) - q(m - x)\} \leqslant \lambda \leqslant \inf_{m \in \mathscr{M}} \{q(m + x) - f(m)\}.$$

In particular, for $m \in \mathscr{M}$,

$$\begin{aligned} f(m) - \lambda &\leqslant q(m - x) \\ f(m) + \lambda &\leqslant q(m + x) \end{aligned} \tag{15.1}$$

Let $\mathscr{N} = \mathscr{M} + \mathbb{R}x$ and define $F : \mathscr{N} \to \mathbb{R}$ by $F(m + tx) = f(m) + t\lambda$ for $m \in \mathscr{M}$ and $t \in \mathbb{R}$. Thus F is linear and agrees with F on $\mathscr{M}$ by definition. We now check that $F(y) \leqslant q(y)$ for all $y \in \mathscr{M} + \mathbb{R}x$. Accordingly, suppose $m \in \mathscr{M}, t \in \mathbb{R}$ and let $y = m + tx$. If $t = 0$ there is nothing to prove. If $t > 0$, then, in view of inequation (15.1),

$$F(y) = F(m + tx) = t(f(\frac{m}{t}) + \lambda) \leqslant t(q(\frac{m}{t} + x)) = q(m + tx) = q(y). \tag{15.2}$$

and a similar estimate shows that $F(m + tx) \leqslant q(m + tx)$for $t < 0$.

We have thus successfully extended f to $\mathscr{M} + \mathbb{R}x$. To finish the proof, let $\mathscr{S}$ be the collcetion of all pairs $(\mathscr{M}_1, f_1)$, where $\mathscr{M}_1$ is a linear manifold in $\mathscr{X}$ such that $\mathscr{M}_1 \supset \mathscr{M}$ and $f_1 : \mathscr{M} \to \mathbb{R}$ is a linear functional with $f_1|\mathscr{M} = f$ and $f_1 \leqslant q$ on $\mathscr{M}_1$. If $(\mathscr{M}_1, f_1)$ and $(\mathscr{M}_2, f_2) \in \mathscr{S}$, define $(\mathscr{M}_1, f_1) \preceq \mathscr{M}_2, f_2$ to mean that $\mathscr{M}_1 \subseteq \mathscr{M}_2$ and $f_2|\mathscr{M}_1 = f_1$. So $(\mathscr{S}, \preceq)$ is a partially ordered set. Suppose $\mathscr{C} = \{(\mathscr{M}_i, f_i) : i \in I\}$ is a chain in $\mathscr{S}$. If $\mathscr{N} \equiv \cup\{\mathscr{M}_i : i \in I\}$, then the fact that $\mathscr{C}$ is a chain implies that $\mathscr{N}$ is a linear mainfold. Define $F : \mathscr{N} \to \mathbb{R}$ by setting $F(x) = f_i(x)$ if $x \in \mathscr{M}_i$. Then F is well defined, linear, and satisfies $F \leqslant q$ on $\mathscr{N}$. So $(\mathscr{N}, F) \in \mathscr{S}$ and $(\mathscr{N}, F)$ is an upper bound for $\mathscr{C}$. By Zorn's Lemma, $\mathscr{S}$ has a maximal element $(\mathscr{Y}, F)$. Since it always possible to extend to a strictly larger subspace, the maximal element must have $\mathscr{N} = \mathscr{X}$, and the proof is finished. $\square$

The proof is a typical application of Zorn's lemma - one knows how to carry out a construction one step a time, but there is no clear way to do it all at once.

Before obtaining further corollaries, we extend these results to the complex case. First, if X is a vector space over $\mathbb{C}$, then trivially it is also a vector space over $\mathbb{R}$, and there is a simple relationship between the $\mathbb{R}$- and $\mathbb{C}$-linear functionals.

Lemma 15.1

Let $\mathscr{X}$ be a vector space over $\mathbb{C}$. If $g : \mathscr{X} \to \mathbb{C}$ is a $\mathbb{C}$-linear functional, then $f(x) = \text{Re } g(x)$ defines an $\mathbb{R}$-linear functional on $\mathscr{X}$ and $g(x) = f(x) - if(ix)$. Conversely, if $f : \mathscr{X} \to \mathbb{R}$ is $\mathbb{R}$-linear then $g(x) = f(x) - if(ix)$ is $\mathbb{C}$-linear. If in addition $p : \mathscr{X} \to \mathbb{R}$ is a seminorm, then $|f(x)| \leqslant p(x)$ for all $x \in \mathscr{X}$ if and only if $|g(x)| \leqslant p(x)$ for all $x \in \mathscr{X}$.

Theorem 15.3: The Hahn-Banach Theorem on $\mathbb{F}$

Let $\mathscr{X}$ be a vector space over $\mathbb{C}$ and let $p : \mathscr{X} \to [0, \infty)$ be a seminorm. If $\mathscr{M}$ is a linear mainfold in $\mathscr{X}$ and $f : \mathscr{M} \to \mathbb{C}$ is a linear funcional such that $|f(x)| \leqslant p(x)$ for all x in $\mathscr{M}$, then there is a linear functional $F : \mathscr{X} \to \mathbb{C}$ such that
(1) $F|\mathscr{M} = f$ (F extends f)
(2) $|F(x)| \leqslant p(x)$ for all x in $\mathscr{X}$ (F is dominated by p).

Proof. Case 1: Note that $f(x) \leqslant |f(x)| \leqslant p(x)$, by theorem15.1, there exists $F : \mathscr{X} \to \mathbb{R}$ such that $F|\mathscr{M} = f$ and $F(x) \leqslant p(x)$ for all $x \in \mathscr{X}$. Hence, $-F(x) = F(-x) \leqslant p(-x) = p(x)$ and so $|F(x)| \leqslant p(x)$.
Case 2: Let $f_1 = \text{Re } f$. By lemma15.1, $|f_1| \leqslant p$, by Case 1, there exist F_1 such $F_1|\mathscr{M} = f_1$ and $|F_1(x)| \leqslant p(x)$ for x in $\mathscr{X}$. Let $F(x) = F_1(x) - iF_1(ix)$, by lemma15.1, $F|\mathscr{M} = f$ and $|F(x)| \leqslant p(x)$ for all x in $\mathscr{X}$. $\square$

The following corollaries are quite important, and when the Hahn-Banach theorem is applied it is usually in one of the following forms:

Corollary 15.1

Let $\mathscr{X}$ be a normed vector space.
(1) If $\mathscr{M}$ is a linear manifold in $\mathscr{X}$, and $f : \mathscr{M} \to \mathbb{F}$ is a bounded linear functional, then there is an F in $\mathscr{X}^*$ such that $F|\mathscr{M} = f$ and $||F|| = ||f||$.
(2) (Linear functionals detect norms) If $x \in \mathscr{X}$ is nonzero, there exists $F \in \mathscr{X}^*$ with $||F|| = 1$ such that $F(x) = ||x||$.
(3) (Linear functionals separate points) If $x \neq y$ in $\mathscr{X}$, there exists $F \in \mathscr{X}^*$ such that $F(x) \neq F(y)$.

Proof. (1) Consider the seminorm $p(x) = ||f|| \cdot ||x||$. Then $|f(x)| \leqslant ||f|| \cdot ||x|| = p(x)$ for $x \in \mathscr{M}$. Hence, by Hahn-Banach Theorem, there exists a linear funcional F on $\mathscr{X}$ such that $F|\mathscr{M} = f$ and $|F(x)| \leqslant p(x)$ for all $x \in \mathscr{X}$. Then $\frac{|F(x)|}{||x||} \leqslant ||f||$ for $x \neq 0$ and so $||F|| \leqslant ||f||$. On the other hand, $||F|| \geqslant ||f||$ since F agree with f on $\mathscr{M}$. Hence, $||F|| = ||f||$.
(2) Let $\mathscr{M}$ be a one-dimensional linear mainfold of $\mathscr{X}$ spanned by x. Define a functional $f : \mathscr{M} \to \mathbb{F}$

by $f(t\frac{x}{||x||}) = t$. Then $|f(y)| = ||y||$ for $y \in \mathscr{M}$ and thus $||f|| = \sup_{y\neq 0} \frac{|f(y)|}{||y||} = 1$. By (1), the funcional f can extend to a funcional F on $\mathscr{X}$ such that $F|\mathscr{M} = f$, $||F|| = ||f||$ and so $F(x) = ||x||$.
(3) Let F be as in (2), then $F(x) - F(y) = F(x-y) = ||x-y|| \neq 0$ and so $F(x) \neq F(y)$. □

Corollary 15.2

If $\mathscr{X}$ is a normed space, $\{x_1, x_2, ..., x_d\}$ is a linearly independent subset of $\mathscr{X}$, and $\alpha_1, \alpha_2, ..., \alpha_d$ are arbitrary scalars, then there is an f in $\mathscr{X}$ such that $f(x_j) = \alpha_j$ for $1 \leqslant j \leqslant d$.

Proof. Let $\mathscr{M}$ be the linear span of $x_1, ..., x_d$ and define $g : \mathscr{M} \to \mathbb{F}$ by $g(\sum_j \beta_j x_j) = \sum_j \beta_j \alpha_j$. Then g is linear. Since $\mathscr{M}$ is finite dimensional, g is bounded. Then by Hahn-Banach Theorem, there is $f : \mathscr{X} \to \mathbb{F}$ such that $f|\mathscr{M} = g$ and so $f(x_j) = \alpha_j, 1 \leqslant j \leqslant d$. □

Corollary 15.3

If x is a normed space and $x \in \mathscr{X}$, then

$$||x|| = \sup\{|f(x)| : f \in \mathscr{X}^* \text{ and } ||f|| \leqslant 1\}.$$

Proof. Let $\alpha = \sup\{|f(x)| : f \in \mathscr{X}^* \text{ and } ||f|| \leqslant 1\}$. Since $|f(x)| \leqslant ||f|| \cdot ||x||$ and $||f|| \leqslant 1$, it follows that $|f(x)| \leqslant ||x||$. Then $\alpha \leqslant ||x||$. On the other hand, there is $f \in \mathscr{X}^*$ such that $||f|| = 1$ and $f(x) = ||x||$. Then $\alpha \geqslant ||x||$. Hence, $\alpha = ||x||$. □

Corollary 15.4: I

$\mathscr{X}$ is normed space, $\mathscr{M} \leqslant \mathscr{X}$, $x_0 \in \mathscr{X} \setminus \mathscr{M}$, and $d = \text{dist}(x_0, \mathscr{M})$, then there is an f in $\mathscr{X}$ such that $f(x_0) = 1$, $f|\mathscr{M} = 0$, and $||f|| = d^{-1}$.

Proof. Since $||x_0 + \mathscr{M}|| = d$, then there is $g \in (\mathscr{X} \setminus \mathscr{M})^*$ such that $g(x_0 + \mathscr{M}) = d$ and $||g|| = 1$. Let $Q : \mathscr{X} \to \mathscr{X}/\mathscr{M}$ be the natural map and $f = d^{-1} g \circ Q : \mathscr{X} \to \mathbb{F}$. Then f is continuous, $f|\mathscr{M} = 0$ since $\ker f = \mathscr{M}$ and $f(x_0) = 1$. Also, $|f(x)| = d^{-1}|g(Q(x))| \leqslant d^{-1}||g|| \cdot ||Q(x)|| \leqslant d^{-1}||x||$; hence $||f|| \leqslant d^{-1}$. On the other hand, $||g|| = 1$ so there is a sequence $\{x_n\}$ such that $|g(x_n + \mathscr{M})| \to 1$ and $||x_n + \mathscr{M}|| < 1$ for all n. Let $y_n \in \mathscr{M}$ such that $||x_n + y_n|| < 1$. Then $|f(x_n + y_n)| = d^{-1}|g(x_n + \mathscr{M})| \to d^{-1}$, so $||f|| \geqslant d^{-1}$. So $||f|| = d^{-1}$. □

15.1 Homework

Exercise 15.1: III6 T1

Let $\mathscr{X}$ be a vector space over $\mathbb{C}$.
(a) If $f : \mathscr{X} \to \mathbb{R}$ is an $\mathbb{R}$-linear Functional, then $\tilde{f}(x) = f(x) - if(ix)$ is a $\mathbb{C}$-linear functional and $f = \text{Re } \tilde{f}$.
(b) If $g : \mathscr{X} \to \mathbb{C}$ is $\mathbb{C}$-linear, $f = \text{Re } g$, and $\tilde{f}$ is defined as in (1), then $\tilde{f} = g$.
(c) If p is a seminorm on $\mathscr{X}$ and f and $\tilde{f}$ are as in (a), then $|f(x)| \leqslant p(x)$ for all x if and only if $|\tilde{f}(x)| \leqslant p(x)$ for all x.
(d) If $\mathscr{X}$ is a normed space and f and $\tilde{f}$ are as in (a), then $||f|| = ||\tilde{f}||$.

Proof. (a) By the construction of $\tilde{f}$, it is clearly that Re $\tilde{f} = f$. Define $h : \mathscr{X} \to \mathbb{C}$ given by $h(x) = ix$. Then h is $\mathbb{R}$-linear. Since composition of two linear functions is a linear function, it follows that $\tilde{f} = f(x) - g \circ f \circ g(x)$ is $\mathbb{R}$-linear. So we only need to check $\tilde{f}(ix) = i\tilde{f}(x)$.

$$\begin{aligned}\tilde{f}(ix) &= f(ix) - if(iix) \\ &= f(ix) - if(-x) = f(ix) + if(x) \\ &= i(f(x) - if(ix)) = i\tilde{f}(x).\end{aligned}$$

Hence, $\tilde{f}$ is $\mathbb{C}$-linear.
(b) If g is $\mathbb{C}$-linear, then $f =$ Re g is $\mathbb{R}$-linear. Since Im $z = -$Re $(iz), \forall z \in \mathbb{C}$, we have

$$\begin{aligned}g(x) &= \text{Re } g(x) + i\text{Im } g(x) \\ &= \text{Re } g(x) - i\text{Re}(ig(x)) \\ &= \text{Re } g(x) - i\text{Re}(g(ix)) \\ &= f(x) - if(ix) = \tilde{f}(x).\end{aligned}$$

(c) ($\Rightarrow$): Now assume that $|f(x)| \leqslant p(x)$ and fix $x \in \mathscr{X}$ and choose θ such that $\tilde{f}(x) = e^{i\theta}|\tilde{f}(x)|$. Hence,

$$\begin{aligned}|\tilde{f}(x)| &= \text{Re } |\tilde{f}(x)| \\ &= \text{Re}(e^{-i\theta}\tilde{f}(x)) = \text{Re } \tilde{f}(e^{-i\theta}x) = f(e^{-i\theta}x) \\ &\leqslant |f(e^{-i\theta}x)| = |f(x)| \leqslant p(x).\end{aligned}$$

($\Leftarrow$): Suppose $|\tilde{f}(x)| \leqslant p(x)$. Then $f(x) =$ Re $\tilde{f}(x) \leqslant |\tilde{f}(x)| \leqslant p(x)$. Also, $-f(x) =$ Re $\tilde{f}(-x) \leqslant |\tilde{f}(-x)| \leqslant p(x)$. Hence $|f(x)| \leqslant p(x)$.
(d) For all $x \in \mathscr{X}$, by (c), $|\tilde{f}(x)| \leqslant |f(x)|$ and $|f(x)| = |\text{Re } \tilde{f}(x)| \leqslant |\tilde{f}(x)|$. Hence, $||f|| \leqslant ||\tilde{f}||$ and $||\tilde{f}|| \leqslant ||f||$. So $||f|| = ||\tilde{f}||$. □

Exercise 15.2: III6 T2

If $\mathscr{X}$ is a normed space, $\mathscr{M}$ is a linear manifold in $\mathscr{X}$, and $f : \mathscr{M} \to \mathbb{F}$ is a bounded linear functional, then there is an F in $\mathscr{X}^*$ such that $F|\mathscr{M} = f$ and $||F|| = ||f||$.

Proof. Consider the seminorm $p(x) = ||f|| \cdot ||x||$. Then $|f(x)| \leqslant ||f|| \cdot ||x|| = p(x)$ for $x \in \mathscr{M}$. Hence, there exists a linear funcional F on $\mathscr{X}$ such that $F|\mathscr{M} = f$ and $|F(x)| \leqslant p(x)$ for all $x \in \mathscr{X}$. Then $\frac{|F(x)|}{||x||} \leqslant ||f||$ for $x \neq 0$ and so $||F|| \leqslant ||f||$. On the other hand, $||F|| \geqslant ||f||$ since F agree with f on $\mathscr{M}$. Hence, $||F|| = ||f||$. □

15.2 Reference

- The Hahn-Banach Theorem
- III6 T1; P32
- III6 T2; P31

Chapter 16

The Dual of a Quotient Space and a Subspace

Chapter 17

Reflexive Space

Chapter 18

The Open Mapping and Closed Graph Theorems

Chapter 19

The Principle of Uniform Roundedness