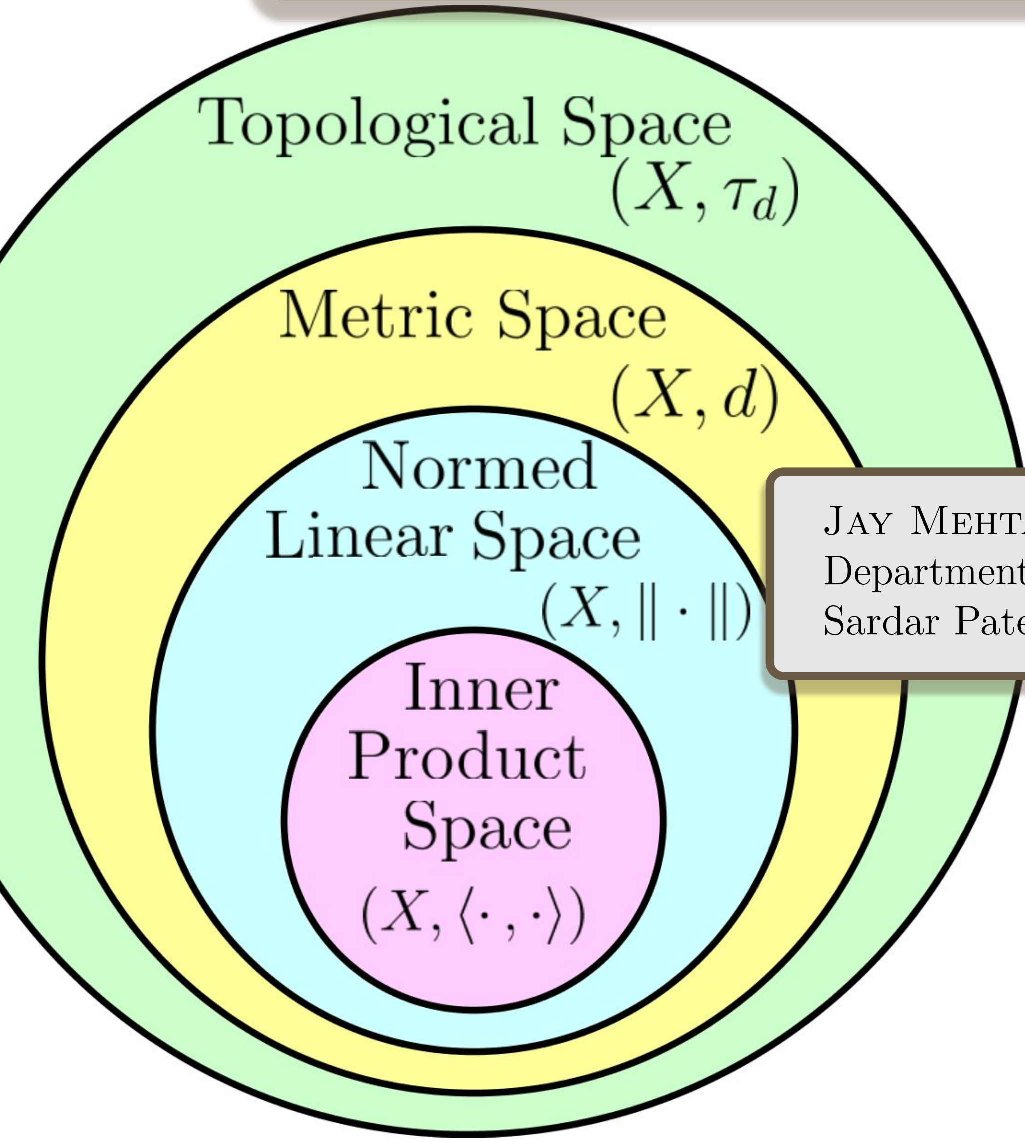
Lecture notes on

Functional Analysis - I

PS02CMTH24



Jay Mehta
Department of Mathematics,
Sardar Patel University.

$\langle x, y \rangle$

ℓ^2

$Y \bigoplus Y^{\perp}$

$\sigma(T)$

Semester - II
2020-21

Preface and Acknowledgments

These lecture notes are the outcome of our course "Functional Analysis-I" offered to M.Sc. students at Department of Mathematics, Sardar Patel University in the year 2016-17. These are aimed to provide reading material to the students besides other reference books and literature mentioned in the syllabus so as to save their time in taking down the notes during the classroom discussion. These notes are tailored for the Functional Analysis - I (PS02CMTH24) syllabus of M.Sc. Semester-II of the University and do not cover all the topics from the subject. Solutions to some of the exercises are not provided in the notes as they were given as student seminar exercises or as assignments and discussed in the seminar sessions during the semester.

.

These handouts are prepared from the recommended reference books as well as lecture notes of past years and not my original work. We mostly followed the book "Functional Analysis" by B. V. Limaye. The lectures are supplemented by many problems and exercises, which contain a lot of additional material. Students are strongly encouraged to refer to the seminar problems and other references along with these notes.

In 2017-18, I did not get the chance to teach this subject again and it was taught by Dr. D. J. Karia who enhanced the notes. These notes may still contain a few errors and it will improve in subsequent repetition of the course. Students and interested readers are welcome to give their valuable suggestions, comments, criticism or point out errors, if they find any.

Acknowledgment.

Thanks are due to Dr. D. J. Karia for revising the notes by making corrections, adding exercises and some more matter to it.

Date: November 25, 2018 JAY MEHTA

Contents

Syllabus

PS02CMTH24: Functional Analysis - I

Unit I: Inner product spaces, normed linear spaces, Banach spaces, examples of inner product spaces, Polarization identity, Schwarz inequality, parallelogram law, uniform convexity of the norm induced by inner product, orthonormal sets, Pythagoras theorem, Gram-Schmidt othonormalization, Bessel's inequality, Riesz-Fischer theorem. Hilbert spaces, orthonormal basis, characterization of orthonormal basis, separable Hilbert spaces.

Unit II: Uniqueness of best approximation from a convex subset of inner product space to a point, orthogonality and best approximation, Gram matrix and its applications, existence and uniqueness of best approximation from a convex subset of a Hilbert space to a point, continuity of a linear mapping, projection theorem and Riesz representation theorem, reflexivity of a Hilbert space. Unique Hahn-Banach extension theorem, weak convergence and weak boundedness.

Unit III: Bounded operators, equivalence of boundedness and continuity of an operator, boundedness of the operator associated to an infinite matrix, adjoint of a bounded operator, properties of adjoint, relations between zero space and the range of operators, normal, unitary and self-adjoint operators, examples, characterizations and results pertaining to these operators, positive operators and generalized Schwarz inequality.

Unit IV: Spectrum, eigenspectrum, approximate eigenspectrum, definition and characterization, spectrum of a normal operator, numerical range, relations of numerical range and different spectra, spectral theorem for a normal/self-adjoint operator on a finite dimensional Hilbert space, compact operators, properties of compact operators, Hilbert-Schmidt operator and its properties, spectrum of a compact operator, spectral theorem for a compact self-adjoint operator.

Text Book

1. Limaye B.V., Functional Analysis, New Age International Publ. Ltd., New Delhi, 1996.
   Chapter 6: Sections 21, 22, 23, 24, Chapter 7: Sections 25, 26, 27, 28.

Reference Book

1. Simmons, G.F., Introduction to Topology and Modern Analysis, McGraw-Hill Co., Tokyo, 1963.
2. Thumban Nair, Functional Analysis: A First Course, Prentice-Hall of India, New Delhi, 2002.

CHAPTER 1

Hilbert Spaces

In this unit, we shall learn Inner product spaces, normed linear spaces, Banach spaces, examples of inner product spaces, Polarization identity, Schwarz inequality, parallelogram law, uniform convexity of the norm induced by inner product, orthonormal sets, Pythagoras theorem, Gram-Schmidt othonormalization, Bessel's inequality, Riesz-Fischer theorem. Hilbert spaces, orthonormal basis, characterization of orthonormal basis, separable Hilbert spaces.

1.1 Inner Product Spaces

1.1.1 Normed Linear Space

Definition 1.1.1. Let X be a linear (vector) space over $\mathbb{K}$ (where $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$). A function $\|\cdot\|: X\to\mathbb{R}$ is called a *norm* on X if it satisfies the following properties.

1. $\|x\|\geq 0$ for all $x\in X$ and $\|x\|=0$ if and only if $x=0$.
2. $\|x+y\|\leq\|x\|+\|y\|$ for all $x,y\in X$.
3. $\|\lambda x\|=|\lambda|\|x\|$ for all $x\in X,\ \lambda\in\mathbb{K}$.

$(X,\|\cdot\|)$ is called a *normed linear space* or a *normed space*. If $\mathbb{K}=\mathbb{R}$, then X is also called a *real normed linear space*. If $\mathbb{K}=\mathbb{C}$, then X is also called a *complex normed linear space*.

Examples 1.1.2. 1. Let $X=\mathbb{K}^n$, where $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$. Define

$$\|x\|_p=\begin{cases}\left(\sum_{i=1}^{n}|x(i)|^p\right)^{\frac{1}{p}} & \text{if } 1\leq p<\infty\\ \sup\limits_{1\leq i\leq n}|x(i)| & \text{if } p=\infty,\end{cases}$$

where $x=(x(1),x(2),\dots,x(n))\in X=\mathbb{K}^n$. Then $(X,\|\cdot\|_p)$ is a normed linear space for $1\leq p\leq\infty$.

2. Let $X = \ell^p = \{x = (x(1), x(2), \ldots) : x(i) \in \mathbb{K} \text{ and } \sum_{i=1}^{\infty} |x(i)|^p < \infty\}$ if $1 \le p < \infty$ and $\ell^\infty = \{x = (x(1), x(2), \ldots) : x(i) \in K \text{ and } \sup_{i \ge 1} |x(i)| < \infty\}$. Define
$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^{\infty} |x(i)|^p \right)^{\frac{1}{p}} & \text{if } 1 \le p < \infty \\ \sup_{i \ge 1} |x(i)| & \text{if } p = \infty, \end{cases}$$
where $x = (x(1), x(2), \ldots) \in \ell^p$. Then $(X, \| \cdot \|_p)$ is a normed linear space for $1 \le p \le \infty$.
3. Let $C[a, b]$ denote the collection of all continuous $f : [a, b] \to \mathbb{K}$ for $1 \le p \le \infty$. Define
$$\|f\|_p = \left(\int_a^b |f(t)|^p dt \right)^{\frac{1}{p}}$$
and
$$\|f\|_\infty = \sup\{|f(t)| : t \in [a, b]\}.$$
Then $\| \cdot \|_p$ is a norm on $C[a, b]$ for $1 \le p \le \infty$.

Note: A complete normed linear space is called a *Banach space*. In the above example $C[a, b]$ is a Banach space with the $\| \cdot \|_\infty$ norm but not a Banach space with $\| \cdot \|_p$ norm.

Definition 1.1.3. Let $(X, \| \cdot \|)$ be a normed linear space. Define $d(x, y) = \|x - y\|$ for all $x, y \in X$, then $d(\cdot, \cdot)$ is a metric on $(X, \| \cdot \|)$. This metric is called the *metric induced by the norm* $\| \cdot \|$.

Remark 1.1.4. Whenever we have a normed linear space $(X, \|.\|)$, we get a metric (induced by the norm) which makes (X, d) a metric space. Thus, every normed linear space is a metric space. The *converse* is not true as X may not have a vector space structure at all, i.e. $x + y$ may not be defined for $x, y \in X$. For example, any non-empty set X with a discrete metric.

1.1.2 Inner Product Space

Definition 1.1.5. Let X be a linear space over the field $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. A function $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{K}$ is called an *inner product* on X if it satisfies following properties.

1. **(Positive-definiteness)** $\langle x, x \rangle \ge 0$ for all $x \in X$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.
2. **(Linearity in the first variable)**: $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ and $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ for every $x, y, z \in X$ and $\alpha \in \mathbb{K}$.
3. **(Conjugate symmetry)** $\langle y, x \rangle = \overline{\langle x, y \rangle}$ for every $x, y \in X$.

A linear (vector) space together with an inner product is called an *inner product space*.

Remark 1.1.6. From the conjugate symmetry, it follows that an inner product is *conjugate linear* in the second variable. That is,
$$\langle x, y + z \rangle = \overline{\langle y + z, x \rangle} = \overline{\langle y, x \rangle} + \overline{\langle z, x \rangle} = \langle x, y \rangle + \langle x, z \rangle \qquad \text{for all } x, y, z \in X$$

and

$$\langle x, \alpha y\rangle = \overline{\langle \alpha y, x\rangle} = \overline{\alpha}\overline{\langle y, x\rangle} = \overline{\alpha}\langle x, y\rangle \qquad \text{for all } x, y \in X,\ \alpha \in \mathbb{K}.$$

Examples 1.1.7. 1. Let $X = K^n$. For $x = (x(1), x(2), \ldots, x(n)),\ y = (y(1), y(2), \ldots, y(n)) \in \mathbb{K}^n$, define

$$\langle x, y\rangle = \sum_{i=1}^{n} x(i)\overline{y(i)}.$$

Then $\langle \cdot\,, \cdot\rangle$ is an inner product on $X = \mathbb{K}^n$.

2. Let $X = c_{00}$ be the linear space of all real (or complex) sequences each with only finitely many non-zero terms. For $x = (x(1), x(2), \ldots, x(n), 0, 0, \ldots)$ and $y = (y(1), y(2), \ldots, y(n), 0, 0, \ldots)$ in $X = c_{00}$, define

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x(i)\overline{y(i)}.$$

Then it is easy to see that $\langle \cdot\,, \cdot\rangle$ is an inner product on c_{00}.

3. Let $X = \ell^2$. For $x = (x(1), x(2), \ldots),\ y = (y(1), y(2), \ldots) \in X$, define

$$\langle x, y\rangle = \sum_{n=1}^{\infty} x(n)\overline{y(n)}.$$

Then $\langle \cdot\,, \cdot\rangle$ is an inner product on ℓ^2.

4. Let $X = C[a, b]$. For $f, g \in X$, define

$$\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}\, dt.$$

Then $\langle \cdot\,, \cdot\rangle$ is an inner product on $X = C[a, b]$.

Proposition 1.1.8. *Let $\langle \cdot\,, \cdot\rangle$ be an inner product on a vector space X.*

1. **(Polarization identity)** *For all $x, y \in X$,*

$$4\langle x, y\rangle = \langle x+y, x+y\rangle - \langle x-y, x-y\rangle + i\langle x+iy, x+iy\rangle - i\langle x-iy, x-iy\rangle.$$

2. Let $x \in X$. Then $\langle x, y\rangle = 0$ for all $y \in X$ if and only if $x = 0$.

3. **(Schwarz inequality)** *For all $x, y \in X$,*

$$|\langle x, y\rangle|^2 \le \langle x, x\rangle\langle y, y\rangle, \tag{1.1}$$

and the equality holds if and only if x and y are linearly dependent.

Proof. (1) Due to linearity of $\langle \cdot\,, \cdot\rangle$ in the first variable and conjugate-linearity in the second variable, the right hand side can be reduced to left hand side as follows.

$$\begin{aligned}
&\langle x+y, x+y\rangle - \langle x-y, x-y\rangle + i\langle x+iy, x+iy\rangle - i\langle x-iy, x-iy\rangle \\
&= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle - \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle - \langle y, y\rangle + i\langle x, x\rangle \\
&\quad - i^2\langle x, y\rangle + i^2\langle y, x\rangle - i^3\langle y, y\rangle - i\langle x, x\rangle - i^2\langle x, y\rangle + i^2\langle y, x\rangle + i^3\langle y, y\rangle \\
&= 4\langle x, y\rangle.
\end{aligned}$$

(2) If $x = 0$, then
$$\langle 0, y\rangle = \langle 0+0, y\rangle = \langle 0, y\rangle + \langle 0, y\rangle.$$
Therefore, $\langle 0, y\rangle = 0$.

Conversely, assume that $\langle x, y\rangle = 0$ for all $y \in X$. In particular, taking $y = x$, we get $\langle x, x\rangle = 0$. Hence, by the positive-definiteness of inner product (Definition 1.1.5), $x = 0$.
(3) For $x, y \in X$, consider $z = \langle y, y\rangle x - \langle x, y\rangle y$. Then

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

Now, if $\langle y, y\rangle > 0$, then $\langle x, x\rangle\langle y, y\rangle - |\langle x, y\rangle|^2 \geq 0$ and the Schwarz inequality follows.

If $\langle y, y\rangle = 0$, then by the definition of inner product, $y = 0$ and hence by (2) above, we have $\langle x, y\rangle = 0$. Hence, $|\langle x, y\rangle|^2 = 0 = \langle x, x\rangle\langle y, y\rangle$.

Now assume that equality holds in the Schwarz inequality (1.1). Then $\langle z, z\rangle = 0$ implies $z = 0$. Hence,
$$\langle y, y\rangle x - \langle x, y\rangle y = z = 0.$$
Thus, x and y are linearly dependent.

Conversely if x and y are linearly dependent, then $y = \alpha x$ for some $\alpha \in \mathbb{K}$. Then

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

and
$$|\langle x, y\rangle|^2 = |\alpha|^2 \langle x, x\rangle^2 = \langle x, x\rangle\langle y, y\rangle.$$
Hence the equality holds in the Schwarz Inequality 1.1. □

Theorem 1.1.9. *Let $\langle \cdot, \cdot \rangle$ be an inner product on a linear space X. For $x \in X$, define $\|x\| = \sqrt{\langle x, x\rangle}$, the non-negative square root of $\langle x, x\rangle$. Then*
$$|\langle x, y\rangle| \leq \|x\|\|y\| \text{ for all } x, y \in X$$
and $\|\cdot\|$ is a norm on X, i.e. the function $\|\cdot\| : X \to K$ is a norm function.

Proof. By the Schwarz inequality, we have, for $x, y \in X$,
$$|\langle x, y\rangle|^2 \leq \|x\|^2\|y\|^2; \tag{1.2}$$
and therefore $|\langle x, y\rangle| \leq \|x\|\|y\|$. Now, we verify that $\|\cdot\|$ is a norm on X.

- $\|x\| = \sqrt{\langle x, x\rangle} \geq 0$ for all $x \in X$ since $\langle x, x\rangle \geq 0$ for all x. Also,
$$\|x\| = 0 \Leftrightarrow \langle x, x\rangle = 0 \Leftrightarrow x = 0.$$

- For all $x, y \in X$, we have,

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x,x\rangle + \langle y,x\rangle + \langle x,y\rangle + \langle y,y\rangle \\ &= \|x\|^2 + \overline{\langle x,y\rangle} + \langle x,y\rangle + \|y\|^2 \\ &= \|x\|^2 + 2\,\mathrm{Re}\langle x,y\rangle + \|y\|^2 \\ &\le \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\|+\|y\|)^2 \quad \text{(by 1.2)}.\end{aligned}$$

  Therefore, $\|x+y\| \le \|x\| + \|y\|$ for all $x, y \in X$.
- For all $x \in X$ and $\alpha \in K$, we have

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

  Therefore, $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in X$ and $\alpha \in K$.

□

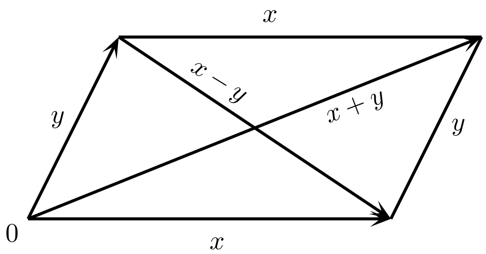
The norm $\|\cdot\|$ defined above is called the *the norm induced by the inner product* or the norm defined by the inner product or norm generated by the inner product.

Remark 1.1.10. From the above theorem, we can say that, "every inner product space is a normed linear space." However, the converse is not true. We will address to the converse very soon but first we recall the law of parallelogram.

Law of Parallelogram

Recall that the parallelogram law states that the sum of the squares of the lengths of four sides of a parallelogram is equal to the sum of the squares of its diagonals.

We have the following theorem:



Theorem 1.1.11 (Parallelogram law). *Let X be an inner product space. Then $\|\cdot\|$ induced by the inner product satisfies*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \ \text{ for all } \ x, y \in X.$$

Proof. Let $x, y \in X$. Then

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

□

Theorem 1.1.12 (Polarization identity). *Suppose X is an inner product space. Then for $x, y \in X$,*

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

Proof. Exercise. □

Question 1.1.13. Is it true that every normed linear space is an inner product space? The answer to this question is given by the following theorem.

Theorem 1.1.14 (Jordan and von Neumann). *Let $\|\cdot\|$ be a norm on a linear space X which satisfies the parallelogram law. Define $\langle\cdot\,,\cdot\rangle : X \times X \to K$ by*

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

for all $x, y \in X$. Then $\langle\cdot\,,\cdot\rangle$ is the unique inner product on X satisfying $\sqrt{\langle x, x\rangle} = \|x\|$ for all $x \in X$.

Proof. Seminar. □

Remark 1.1.15. By the above result, we can say that a normed linear space is an inner product space if the norm satisfies the parallelogram law. The following proposition makes it more clear.

Proposition 1.1.16. *The normed linear space $(\ell^p, \|\cdot\|_p)$ is an inner product space if and only if $p = 2$.*

Proof. Define the inner product on ℓ^2 by

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x(n)\overline{y(n)}, (x = (x(1), x(2), \dots),\ y = (y(1), y(2), \dots) \in \ell^2).$$

Then clearly, (verify!) $\langle\cdot\,,\cdot\rangle$ is an inner product on ℓ^2 making ℓ^2 an inner product space. Also, the norm is defined as

$$\|x\| = \sqrt{\langle x, x\rangle} = \left(\sum_{i=1}^{\infty} |x(n)|^2\right)^{1/2} = \|x\|_2\ .$$

Conversely, assume that $(\ell^p, \|\cdot\|_p)$ is an inner product space. Then the norm $\|\cdot\|_p$ satisfies the parallelogram law, i.e. for $x, y \in \ell^p$,

$$\|x+y\|_p^2 + \|x-y\|_p^2 = 2(\|x\|_p^2 + \|y\|_p^2) \tag{1.3}$$

must hold. Now, take $x = (1, 0, 0, \dots)$ and $y = (0, 1, 0, \dots)$ in ℓ^p. Then

$$x + y = (1, 1, 0, \dots)$$

$$x - y = (1, -1, 0, \ldots).$$

Therefore, $\|x+y\|_p = 2^{\frac{1}{p}}$, $\|x-y\|_p = 2^{\frac{1}{p}}$, $\|x\|_p = 1$ and $\|y\|_p = 1$. Thus by (1.3), we get

$$\begin{aligned} 2^{\frac{2}{p}} + 2^{\frac{2}{p}} &= 2(1+1) \\ \Rightarrow 2^{\frac{2}{p}} &= 2 \\ \Rightarrow &\boxed{p = 2}. \end{aligned}$$

□

Proposition 1.1.17. *Let X be an inner product space. If $\{x_n\}$ and $\{y_n\}$ are sequences in X such that $\|x_n - x\| \to 0$ and $\|y_n - y\| \to 0$ in X. Then*

1. *$\langle x_n, y_n\rangle \to \langle x, y\rangle$ i.e. inner product is jointly continuous.*
2. *$\langle x_n, z\rangle \to \langle x, z\rangle$ for all $z \in X$.*

Proof.

1. $$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle| \\ &\le |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\le \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\| \qquad \text{(by the Schwarz inequality)} \\ &\to 0. \end{aligned}$$
2. $$\begin{aligned} |\langle x_n, z\rangle - \langle x, z\rangle| &= |\langle x_n - x, z\rangle| \\ &\le \|x_n - x\|\|z\| \qquad \text{(by the Schwarz inequality)} \\ &\to 0. \end{aligned}$$

□

Definition 1.1.18. Let V be a vector space over K. A subset C of V is said to be *convex* if for each $x, y \in C$ and $0 \le t \le 1$,

$$tx + (1-t)y \in C.$$

That is, the line segment joining x and y is also in C.

Example 1.1.19. Every subspace of a vector space is convex.

Definition 1.1.20. Let X be a normed linear space. Then

$$S_1(0) = \{x \in X : \|x\| < 1\}$$

is called the *open unit ball of X* and

$$\overline{S_1(0)} = \{x \in X : \|x\| \le 1\}$$

is called the *closed unit ball of X*.

Example 1.1.21. The open unit ball in X, $S_1(0)$ is convex.

Solution. Let $x, y \in S_1(0)$. Then $\|x\| < 1$ and $\|y\| < 1$. Now,

$$\begin{aligned}\|tx + (1-t)y\| &\leq \|tx\| + \|(1-t)y\| \\ &= t\|x\| + (1-t)\|y\| \qquad \text{(since } t \in [0,1]) \\ &< t(1) + (1-t)(1) \\ &= 1.\end{aligned}$$

So, the (open) unit ball in X is convex. Similarly, the closed unit ball in X is also convex. □

Definition 1.1.22. A normed linear space X is called *uniformly convex* if for every $\epsilon > 0$ there exists $\delta > 0$ such that for each $x, y \in X$ with $\|x\| \leq 1$, $\|y\| \leq 1$ and $\|x - y\| \geq \epsilon$,

$$\left\|\frac{x+y}{2}\right\| \leq 1 - \delta.$$

Theorem 1.1.23. *Let X be an inner product space. Then the normed linear space X with the induced norm is uniformly convex.*

Proof. Let $\epsilon > 0$ be given and let $x, y \in X$ such that $\|x\| \leq 1$, $\|y\| \leq 1$ and $\|x - y\| \geq \epsilon$. By parallelogram law,

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

Therefore,

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

Therefore,

$$\left\|\frac{x+y}{2}\right\|^2 \leq \frac{4-\epsilon^2}{4} \quad \text{or} \quad \left\|\frac{x+y}{2}\right\| \leq \sqrt{1 - \frac{\epsilon^2}{4}}.$$

Take $\delta = 1 - \sqrt{1 - \frac{\epsilon^2}{4}}$. Then $\left\|\frac{x+y}{2}\right\| \leq \sqrt{1 - \frac{\epsilon^2}{4}} = 1 - \delta$. Thus, X with the norm induced from the inner product is uniformly convex. □

Seminar Topics 1.

1. Let X be a nonzero vector space and $B = \{v_i : i \in I\}$ be a basis. For $v = \sum_{i \in I} \alpha_i v_i \in X$, define $\|x\| = \sum_{i \in I} |\alpha_i|$. Show that $\|\cdot\|$ is a norm on X.
2. For $x = (x(1), x(2), \ldots, x(n)) \in K^n$ and $p \in [1, \infty)$, let $\|x\|_p = \left(\sum_{i=1}^{n} |x(i)|^p\right)^{\frac{1}{p}}$. Show that $\|\cdot\|_p$ is a norm on K^n.
3. For $x = (x(1), x(2), \ldots, x(n)) \in K^n$ and $p \in [1, \infty)$, let $\|x\|_\infty = \sup_{1 \leq i \leq n} |x(i)|$. Show that $\|\cdot\|_\infty$ is a norm on K^n.
4. Show that for $1 \leq p \leq \infty$, $(\ell^p, \|\cdot\|_p)$ is a normed linear space.
5. Show that $(\ell^p, \|\cdot\|_p)$ is complete for $1 \leq p \leq \infty$.

6. For $1 \le p_1 \le p_2 \le \infty$, show that $\ell^{p_1} \subset \ell^{p_2}$.

7. Let $c_0 = \{x = (x(1), x(2), \dots) : x(i) \in K \text{ and } \lim_{n\to\infty} x(n) = 0\}$. For $x \in c_0$ define $\|x\|_\infty = \sup_{n\in\mathbb{N}} |x(n)|$. Show that $\|\cdot\|_\infty$ is a norm on c_0.

8. Show that $(c_0, \|\cdot\|_\infty)$ is complete.

9. Let $\mathcal{P}[0,1]$ denote the set of all polynomials with complex coefficients. For $p(x) = a_0 + a_1 x + \cdots a_n x^n \in \mathcal{P}[0,1]$, define

$$\|p(x)\|_\infty = \sup_{0\le x\le 1} |p(x)| \tag{1.4}$$

$$\|p(x)\|_{sup} = \sup_{0\le i\le n} |a_i| \tag{1.5}$$

$$\|p(x)\|_{sum} = \sum_{i=0}^{n} |a_i|. \tag{1.6}$$

Show that all these define norms on $\mathcal{P}[0,1]$.

10. Let $B[0,1] = \{f : [0,1] \to K : f \text{ is bounded }\}$. For $f \in B[0,1]$, define $\|f\|_\infty = \sup_{0\le t\le 1} |f(t)|$. Show that $\|\cdot\|_\infty$ is a complete norm on $B[0,1]$.

11. Let $C[0,1] = \{f : [0,1] \to K : f \text{ is continuous}\}$. For $1 \le p \le \infty$, show that $\|\cdot\|_p$ is a norm on $C[0,1]$. Show that $\|\cdot\|_\infty$ is a complete norm on $C[0,1]$.

12. Show that the sequence $\{f_n\}$ in $C[0,1]$ defined by $f_n(t) = t^n$, $(t \in [0,1], n \in \mathbb{N})$, converges pointwise but does not converge in the supnorm $\|\cdot\|_\infty$ on $C[0,1]$.

13. Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space. For a scalar λ, define $\langle x, y\rangle_\lambda = \lambda\langle x, y\rangle$, $(x, y \in X)$. Show that $\langle\cdot,\cdot\rangle_\lambda$ is an inner product if and only if $\lambda > 0$.

14. Prove the Polarization identity (Theorem 1.1.12).

15. Prove Jordan and von Neumann identity.

16. On ℓ^2 define

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x(n)\overline{y(n)}, (x = (x(1), x(2), \dots),\ y = (y(1), y(2), \dots) \in \ell^2).$$

Show that $\langle\cdot,\cdot\rangle$ is an inner product on ℓ^2.

17. Show that the open unit ball in X, $S_1(0)$ is convex.

18. Let $(X, \|\cdot\|)$ be a normed linear space, $x_0 \in X$ and $r > 0$. Show that

$$S_r(x_0) = \{x \in X : \|x - x_0\| < r\}$$

is convex.

19. Show that $(C[0,1], \|\cdot\|_\infty)$ is not uniformly convex.

1.2 Orthonormal sets

Definition 1.2.1. Let X be a normed linear space. Two elements $x, y \in X$ are said to be *orthogonal* if $\langle x, y\rangle = 0$. In this case, we write $x \perp y$ (read x perp y) i.e. x is orthogonal to y or x is perpendicular to y.

Examples 1.2.2. 1. In $X = \mathbb{R}^2$, the elements $x = (1,0)$ and $y = (0,0)$ are orthogonal as $\langle x, y\rangle = 0$. In fact, 0 is orthogonal to every element.

2. Take $X = \mathbb{R}^2$ and $x = (2,0)$, $y = (0,-7)$. Then $x \perp y$.

Two nonzero vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, are orthogonal if and only if they are perpendicular in the usual sense.

Definition 1.2.3. A non-empty subset E of X is said to be *orthogonal subset* if for every $x, y \in E$ such that $x \neq y$, then $\langle x, y\rangle = 0$.

Examples 1.2.4.
1. In $X = \mathbb{R}^2$, the set $E = \{(1,0),(0,8),(0,0)\}$ is an orthogonal subset of X.
2. Take $X = \mathbb{R}^2$ and $E = \{(4,18),(9,-2)\}$. Then E is orthogonal.

Definition 1.2.5. An orthogonal subset E of X is said to be *orthonormal* if $\|x\| = 1$ for all $x \in E$.

Examples 1.2.6.
1. In $X = \mathbb{R}^3$, the set $E = \{(1,0,0),(0,1,0),(0,0,1)\}$ is an orthonormal subset of X.
2. Take $X = \mathbb{C}^2$ and $E = \{(\frac{i}{2}, \frac{\sqrt{3}}{2}), (\frac{i\sqrt{3}}{2}, \frac{1}{2})\}$. Then E is an orthonormal subset of $\mathbb{C}^2$.

Remark 1.2.7. An orthonormal set E will never contain the zero element since $\|x\| = 1$ for all $x \in E$. So $0 \notin E$. Also, every orthonormal set is an orthogonal set.

Theorem 1.2.8 (Pythagoras theorem)**.** *Let X be an inner product space and $x_1, x_2, \ldots, x_n \in X$ be orthogonal. Then*

$$\|x_1 + x_2 + \cdots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2.$$

i.e.

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

Proof. Since $x_1, x_2, \ldots, x_n \in X$ are orthogonal, if $j \neq i$, then $\langle x_j, x_i\rangle = 0$ and if $j = i$, then $\langle x_j, x_i\rangle = \|x_j\|^2$. So, we have,

$$\begin{aligned}\left\|\sum_{j=1}^{n} x_j\right\|^2 &= \left\langle \sum_{j=1}^{n} x_j, \sum_{i=1}^{n} x_i \right\rangle \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} \langle x_j, x_i\rangle \\ &= \sum_{j=1}^{n} \|x_j\|^2.\end{aligned}$$

□

Theorem 1.2.9. *Let X be an inner product space and $E \subset X$ be orthogonal such that $0 \notin E$. Then E is a linearly independent set. In particular, if E is orthonormal, then E is linearly independent. In fact, if E has more than one element, then the diameter of E is $\sqrt{2}$.*

Proof. Let $x_1, x_2, \ldots, x_n \in E$ and $\alpha_1, \alpha_2, \ldots, \alpha_n \in K$ such that

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_j x_j + \cdots + \alpha_n x_n = 0.$$

Then for $j = 1, 2, \ldots, n$, we have

$$\begin{aligned} 0 &= \langle 0, x_j \rangle \\ &= \langle \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_j x_j + \cdots + \alpha_n x_n, x_j \rangle \\ &= \alpha_1 \langle x_1, x_j \rangle + \alpha_2 \langle x_2, x_j \rangle + \cdots + \alpha_j \langle x_j, x_j \rangle + \cdots + \alpha_n \langle x_n, x_j \rangle \\ &= \alpha_j \|x_j\|^2 \qquad (\because \langle x_i, x_j \rangle = 0 \text{ if } i \neq j). \end{aligned}$$

Since $0 \notin E$, $\|x_j\| \neq 0$ and hence, $\alpha_j = 0$ for all $j = 1, 2, \ldots, n$. Therefore, E is linearly independent.

Clearly, if E is orthonormal, then E is orthogonal and $0 \notin E$. Hence, E is linearly independent by the same argument as above.

Now, suppose E is an orthonormal set and it has more than one element. Then for any $x, y \in E$, $x \neq y$

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

Therefore, the diameter of E is $\operatorname{diam}(E) = \sup\{\|x - y\| : x, y \in E\} = \sqrt{2}$. □

Remark 1.2.10. Thus, we have seen that (by the above result), every orthogonal set which does not contain 0 and so, every orthonormal set is a linearly independent set. Then we have the following question asking about its converse.

Question 1.2.11. Is the converse of above true? That is, if $E \subset X$ is linearly independent, then is it true that E is orthogonal or orthonormal. The answer is **NO** in general. Consider the following example.

Example 1.2.12. Let $X = K^2$ and $E = \{(1, 0), (1, 1)\}$. Then clearly, (Check!) E is a linearly independent set but E is not orthogonal (or orthonormal).

Remark 1.2.13. We have seen so far that an orthogonal set not containing 0 or an orthonormal set is always linearly independent but the converse is not true. However, given any linearly independent set, we can always find an orthonormal set such that they span the same set. This result (given below) is well-known as the *Gram-Schmidt orthonormalization theorem* and the process by which we obtain the required orthonormal set is called the *Gram-Schmidt orthonormalization process*. More precisely, we have the following theorem.

Theorem 1.2.14 (Gram-Schmidt orthonormalization)**.** *Let X be an inner product space and $\{x_1, x_2, \ldots\}$ be a linearly independent subset of X. Then there exists an orthonormal subset $\{u_1, u_2, \ldots\}$ of X such that for each $k = 1, 2, \ldots,$*

$$L(\{u_1, u_2, \ldots, u_k\}) = L(\{x_1, x_2, \ldots, x_k\}).$$

In fact, the above set can be obtained as follows:

Define $y_1 = x_1$, $u_1 = \frac{y_1}{\|y_1\|}$ *and for* $j = 2, 3, \ldots,$ *define*

$$y_j = x_j - \langle x_j, u_1\rangle u_1 - \langle x_j, u_2\rangle u_2 - \cdots - \langle x_j, u_{j-1}\rangle u_{j-1}, \qquad u_j = \frac{y_j}{\|y_j\|}.$$

Proof. We prove this result by the principle of mathematical induction on j.
Case: $j = 1$. Since $\{x_1\} = \{y_1\}$ is linearly independent as $x_1 \neq 0$ and since $u_1 = \frac{y_1}{\|y_1\|}$, $\|u_1\| = 1$, and hence clearly $\{u_1\}$ is an orthonormal set and $L(\{u_1\}) = L(\{x_1\})$.

For understanding only, not required to prove

Case: $j = 2$. Note that $y_2 = x_2 - \langle x_2, u_1\rangle u_1$. If $y_2 = 0$, then $x_2 = \frac{\langle x_2, x_1\rangle}{\|x_1\|^2} x_1 \in L(\{x_1\})$, which is not possible as the $\{x_1, x_2\}$ is a linearly independent set. Now,

$$\begin{aligned}\langle y_2, u_1\rangle &= \langle x_2 - \langle x_2, u_1\rangle u_1, u_1\rangle \\ &= \langle x_2, u_1\rangle - \langle x_2, u_1\rangle\langle u_1, u_1\rangle \\ &= 0 \qquad (\because \langle u_1, u_1\rangle = \|u_1\|^2 = 1).\end{aligned}$$

Also, $u_2 = \frac{y_2}{\|y_2\|}$. Then $\|u_2\| = 1$ and from the above, we have $\langle u_2, u_1\rangle = \frac{1}{\|y_2\|}\langle y_2, u_1\rangle = 0$, i.e. $\{u_1, u_2\}$ is orthonormal. Also since $u_2 \in L(\{x_1, x_2\})$, we have

$$L(\{u_1, u_2\}) = L(\{x_1, u_2\}) \subset L(\{x_1, x_2\}).$$

Since dimension of both the spaces $L(\{u_1, u_2\})$ and $L(\{x_1, x_2\})$ is 2 (same), we have

$$L(\{u_1, u_2\}) = L(\{x_1, x_2\}).$$

Induction Hypothesis: $j = k$. Assume that the result holds for $j = k$, i.e. y_k and u_k defined above are such that $\{u_1, u_2, \ldots, u_k\}$ is an orthonormal set and

$$L(\{u_1, u_2, \ldots, u_k\}) = L(\{x_1, x_2, \ldots, x_k\}).$$

Case: $j = k + 1$. Now,

$$y_{k+1} = x_{k+1} - \langle x_{k+1}, u_1\rangle u_1 - \langle x_{k+1}, u_2\rangle u_2 - \cdots - \langle x_{k+1}, u_k\rangle u_k.$$

If $y_{k+1} = 0$, then $x_{k+1} \in L(\{u_1, u_2, \ldots, u_k\}) = L(\{x_1, x_2, \ldots, x_k\})$, which is not possible since $\{x_1, x_2, \ldots, x_k, x_{k+1}\}$ is a linearly independent set. Hence, $y_{k+1} \neq 0$. Now, for $i \leq k$

$$\begin{aligned}\langle y_{k+1}, u_i\rangle &= \langle x_{k+1} - \langle x_{k+1}, u_1\rangle u_1 - \cdots - \langle x_{k+1}, u_k\rangle u_k, u_i\rangle \\ &= \langle x_{k+1}, u_i\rangle - \langle x_{k+1}, u_1\rangle\langle u_1, u_i\rangle - \cdots - \langle x_{k+1}, u_k\rangle\langle u_k, u_i\rangle \\ &= \langle x_{k+1}, u_i\rangle - \langle x_{k+1}, u_i\rangle \qquad (\because \langle u_j, u_i\rangle = 0,\ j \neq i \text{ and } \langle u_i, u_i\rangle = 1) \\ &= 0.\end{aligned}$$

Take $u_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|}$, Then $\|u_{k+1}\| = 1$ and for $\leq i \leq k$, $\langle u_{k+1}, u_i\rangle = \frac{1}{\|y_{k+1}\|}\langle y_{k+1}, u_i\rangle = 0$. Hence, $\{u_1, u_2, \ldots, u_{k+1}\}$ is an orthonormal set. Also,

$$L(\{u_1, u_2, \ldots, u_{k+1}\}) = L(\{x_1, x_2, \ldots, x_k, u_{k+1}\}) = L(\{x_1, x_2, \ldots, x_{k+1}\}).$$

(since dimension of the above spaces is same).

This completes the proof. □

Example 1.2.15. Let $X = \ell^2$. For $n = 1, 2, \ldots$, let $x_n = (\underbrace{1, \ldots, 1}_{n \text{ times}}, 0, 0, \ldots)$ i.e. 1 occurs only in the first n entries. It can be easily seen that by Gram-Schmidt orthonormalization process, we get an orthonormal set $\{u_1, u_2, \ldots\}$, where

$$y_n = (0, \ldots, 0, \underbrace{1}_{n^{\text{th}}}, 0, 0, \ldots) = u_n,$$

where 1 occurs only in the n^{th} entry.

Lemma 1.2.16 (Bessel's inequality)**.** *Let X be an inner product space and $\{u_1, u_2, \ldots\}$ be a countable orthonormal subset of X. Then for each $x \in X$*

$$\sum_{n=1}^{\infty} |\langle x, u_n\rangle|^2 \le \|x\|^2,$$

where the equality holds if and only if $x = \sum_{n=1}^{\infty} \langle x, u_n\rangle u_n$.

Proof. Let $x \in X$ and for $m = 1, 2, \ldots$, let

$$x_m = \sum_{n=1}^{m} \langle x, u_n\rangle u_n.$$

Then,

$$\begin{aligned}
\langle x_m, x\rangle &= \left\langle \sum_{n=1}^{m} \langle x, u_n\rangle u_n, x \right\rangle \\
&= \sum_{n=1}^{m} \langle x, u_n\rangle \langle u_n, x\rangle \\
&= \sum_{n=1}^{m} |\langle x, u_n\rangle|^2.
\end{aligned}$$

Since the above entity is a real number, we have

$$\langle x, x_m\rangle = \overline{\langle x_m, x\rangle} = \sum_{n=1}^{m} |\langle x, u_n\rangle|^2.$$

Also,

$$\begin{aligned}
\langle x_m, x_m\rangle &= \left\langle \sum_{n=1}^{m} \langle x, u_n\rangle u_n, \sum_{k=1}^{m} \langle x, u_k\rangle u_k \right\rangle \\
&= \sum_{n=1}^{m} \sum_{k=1}^{m} \langle x, u_n\rangle \overline{\langle x, u_k\rangle} \langle u_n, u_k\rangle \\
&= \sum_{n=1}^{m} \langle x, u_n\rangle \overline{\langle x, u_n\rangle} \qquad (\because \{u_1, u_2, \ldots\} \text{ is orthonormal}) \\
&= \sum_{n=1}^{m} |\langle x, u_n\rangle|^2.
\end{aligned}$$

Thus,

$$\langle x_m, x_m\rangle = \langle x, x_m\rangle = \langle x_m, x\rangle = \sum_{n=1}^{m} |\langle x, u_n\rangle|^2. \tag{1.7}$$

Now,

$$\begin{aligned} 0 \le \|x - x_m\|^2 &= \langle x - x_m, x - x_m\rangle \\ &= \langle x, x\rangle - \langle x, x_m\rangle - \langle x_m, x\rangle + \langle x_m, x_m\rangle \\ &= \langle x, x\rangle - \sum_{n=1}^{m} |\langle x, u_n\rangle|^2 \qquad \text{(by (1.7))}. \end{aligned} \tag{1.8}$$

Thus, for each $m = 1, 2, \ldots$

$$\sum_{n=1}^{m} |\langle x, u_n\rangle|^2 \le \langle x, x\rangle = \|x\|^2.$$

Taking limit as $m \to \infty$, we get

$$\sum_{n=1}^{\infty} |\langle x, u_n\rangle|^2 \le \|x\|^2.$$

($S_m = \sum_{n=1}^{m} |\langle x, u_n\rangle|^2$ increasing and bounded above by $\|x^2\|$. So, it is convergent). By equation (1.8), the equality holds if and only if

$$\begin{aligned} &\lim_{m\to\infty} \|x - x_m\|^2 = 0 \quad \text{if and only if} \\ &\lim_{m\to\infty} x_m = x \quad \text{if and only if} \\ &\sum_{n=1}^{\infty} \langle x, u_n\rangle u_n = x. \end{aligned}$$

□

Example 1.2.17. Derive Schwarz inequality using Bessel's inequality.

Solution. We know that the Schwarz inequality is trivially true if $y = 0$. So, we assume that $y \neq 0$. Take $u = \frac{y}{\|y\|}$, then $\{u\}$ is orthonormal subset and by the Bessel's inequality, we have

$$\begin{aligned} |\langle x, u\rangle|^2 &\le \|x\|^2 \\ \Rightarrow \left|\left\langle x, \frac{y}{\|y\|}\right\rangle\right|^2 &\le \|x\|^2 \\ \Rightarrow \frac{1}{\|y\|^2}|\langle x, y\rangle|^2 &\le \|x\|^2 \\ \Rightarrow |\langle x, y\rangle|^2 &\le \|x\|^2\|y\|^2. \end{aligned}$$

Therefore, $\boxed{|\langle x, y\rangle|^2 \le \langle x, x\rangle\langle y, y\rangle}$. Hence, we have deduced the Schwarz inequality from the Bessel's inequality. □

Seminar Topics 2.

1. Let $X = K^3$, where $K = \mathbb{R}$ or $\mathbb{C}$. Let $x_1 = (1, 0, 0)$, $x_2 = (1, 1, 0)$ and $x_3 = (1, 1, 1)$. Orthonormalize the set $\{x_1, x_2, x_3\}$.
2. Let $a = (1 + 2i, 4, 7) \in \mathbb{C}$. Show that

$$a \perp \{x = (x(1), x(2), x(3)) \in \mathbb{C} : a \perp x\}$$

is a subspace of $\mathbb{C}^3$. Find the dimension of $a^{\perp}$.

3. Let $\{u_1, u_2, u_3\}$ be an orthonormal set in an inner product space X over $\mathbb{C}$. Show that $\{\alpha_1 u_1, \alpha_2 u_2, \alpha_3 u_3\}$ is orthonormal iff $|\alpha_i| = 1$.
4. Orthonormalize the following set in respective inner product spaces.
   (i) $\{(1,0,1),(1,0,2),(1,1,1)\}$ in $\mathbb{R}^3$.
   (ii) $\{(1,0,2),(1,0,1),(1,1,1)\}$ in $\mathbb{R}^3$.
   (iii) $\{(1,0,0),(1,1,0),(1,1,1)\}$ in $\mathbb{R}^3$.

1.3 Hilbert spaces

Definition 1.3.1. A complete inner product space is called a *Hilbert space.*

Theorem 1.3.2 (Riesz-Fischer theorem)**.** *Let H be a Hilbert space and $\{u_1, u_2, \ldots\}$ be a countable orthonormal subset of a Hilbert space H. Suppose $\{\alpha_n\}$ is a sequence in K. Then $\sum_{n=1}^{\infty} \alpha_n u_n$ converges to some $x \in H$ if and only if $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$. In this case, $\alpha_n = \langle x, u_n \rangle$ for all n.*

Proof. Suppose $\sum_{n=1}^{\infty} \alpha_n u_n = x \in H$. Then for $m = 1, 2, \ldots$,

$$\begin{aligned}
\langle x, u_m \rangle &= \left\langle \sum_{n=1}^{\infty} \alpha_n u_n, u_m \right\rangle \\
&= \sum_{n=1}^{\infty} \alpha_n \langle u_n, u_m \rangle \qquad \text{(since inner product is a continuous function)} \\
&= \alpha_m .
\end{aligned}$$

Therefore, by the Bessel's inequality,

$$\sum_{n=1}^{\infty} |\alpha_n|^2 = \sum_{n=1}^{\infty} |\langle x, u_n \rangle|^2 \le \|x\|^2 .$$

Next suppose that $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$. Let $S_m = \sum_{n=1}^{m} \alpha_n u_n$ for $m = 1, 2, \ldots$. Then for $k \le m$,

$$S_m - S_k = \sum_{k+1}^{m} \alpha_n u_n \qquad \text{and so}$$

$$\begin{aligned}
\|S_m - S_k\|^2 &= \left\| \sum_{n=k+1}^{m} \alpha_n u_n \right\|^2 \\
&= \left\langle \sum_{n=k+1}^{m} \alpha_n u_n, \sum_{l=k+1}^{m} \alpha_l u_l \right\rangle \\
&= \sum_{n=k+1}^{m} \sum_{l=k+1}^{m} \alpha_n \overline{\alpha_l} \langle u_n, u_l \rangle
\end{aligned}$$

$$= \sum_{n=k+1}^{m} \alpha_n \overline{\alpha_n}$$
$$= \sum_{n=k+1}^{m} |\alpha_n|^2 \to 0 \text{ as } n, m \to 0. \quad (1.9)$$

Therefore, $\{S_m\}$ is a Cauchy sequence in H. Since H is complete, $\{S_m\}$ converges to some $x \in H$, i.e.

$$\sum_{n=1}^{\infty} \alpha_n u_n = x \in H.$$

□

Definition 1.3.3. An orthonormal subset E of a Hilbert space H is called a *maximal orthonormal set* if for every orthonormal set $F \subset H$, $E \subset F \Rightarrow E = F$. A maximal orthonormal subset of a Hilbert space H is called an *orthonormal basis* for H.

Examples 1.3.4.

1. Let H be finite dimensional, i.e. $\dim H < \infty$. If $x_1, x_2, \ldots, x_n \in H$ are linearly independent such that $L(\{x_1, x_2, \ldots, x_n\}) = H$, then Gram-Schmidt orthonormalization yields an orthonormal set $\{u_1, u_2, \ldots, u_n\}$ such that

$$L(\{u_1, u_2, \ldots, u_n\}) = L(\{x_1, x_2, \ldots, x_n\}) = H.$$

Note that there no linearly independent superset of $\{u_1, u_2, \ldots, u_n\}$. Thus, $\{u_1, u_2, \ldots, u_n\}$ is an orthonormal basis for H.

2. $H = \ell^2 = \{\{x_n\} | \sum_{n=1}^{\infty} |x_n|^2 < \infty\}$. In this case, $\{e_1, e_2, \ldots\}$ is orthonormal, where $e_1 = (1, 0, 0, \ldots)$, $e_2 = (0, 1, 0, \ldots), \ldots$.

Theorem 1.3.5. *Let X be an inner product space and E be an orthonormal subset of X. Then for each $x \in X$ the set*

$$E_x = \{u \in E : \langle x, u \rangle \neq 0\}$$

is countable. Suppose $E_x = \{u_1, u_2, \ldots\}$ (countable) and X is complete. Then $\sum_{n=1}^{\infty} \langle x, u_n \rangle u_n$ converges to $y \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n$ in X such that $x - y \perp E$.

Proof. Let $x \in X$. If $x = 0$, then $E_x = \emptyset$. So assume that $x \neq 0$. For $j = 1, 2, \ldots$, consider the set

$$F_j = \{u \in E : \|x\| \leq j |\langle x, u \rangle|\}.$$

Fix j. Let $u_1, u_2, \ldots, u_m \in F_j$, then

$$\|x\| \leq j |\langle x, u_i \rangle| \qquad i = 1, 2, \ldots, m.$$

Therefore, by Bessel's inequality, we have

$$\sum_{i=1}^{m} \|x\|^2 \leq j^2 \sum_{i=1}^{m} |\langle x, u_i \rangle| \leq j^2 \|x\|^2.$$

Therefore, $m \leq j^2$. Thus, F_j has at most j^2 elements (i.e., it is a finite set).

Claim: $E_x = \bigcup_{j=1}^{\infty} F_j$.

If $u \in F_j$ for some j, then

$$0 < \|x\| \leq j|\langle x, u\rangle|.$$

Therefore, $\langle x, u\rangle \neq 0$ and so $u \in E_x$. Thus,

$$\bigcup_{j=1}^{\infty} F_j \subset E_x.$$

Now suppose $u \in E_x$. Then $\langle x, u\rangle \neq 0$. So, there exists $j_0 \in \mathbb{N}$ such that

$$\|x\| \leq j_0|\langle x, u\rangle|, \text{ i.e.}$$

$$u \in F_{j_0}.$$

Therefore, $E_x \subset \bigcup_{j=1}^{\infty} F_j$ and so

$$E_x = \bigcup_{j=1}^{\infty} F_j.$$

Hence, E_x is countable.

Take $E_x = \{u_1, u_2, \ldots\}$. Now, by Bessel's inequality, since X is complete,

$$\sum_{n=1}^{\infty} |\langle x, u_n\rangle|^2 \leq \|x\|^2 < \infty.$$

Therefore, by Riesz-Fischer theorem, $\sum_{n=1}^{\infty}\langle x, u_n\rangle u_n$ converges in X. Suppose $y = \sum_{n=1}^{\infty}\langle x, u_n\rangle u_n$ Now for $u \in E$,

$$\begin{aligned}
\langle y, u\rangle &= \left\langle \sum_{n=1}^{\infty}\langle x, u_n\rangle u_n, u \right\rangle \\
&= \sum_{n=1}^{\infty}\langle x, u_n\rangle\langle u_n, u\rangle \\
&= \langle x, u\rangle \qquad (\text{if } u \neq u_n \text{ then } \langle u, u_n\rangle = 0 \text{ otherwise it is } 1).
\end{aligned}$$

Therefore, $\langle x - y, u\rangle = 0$. That is, $(x - y) \perp E$. □

Theorem 1.3.6. *Le H be a Hilbert space and $E \subset H$ be an orthonormal set. Then the following are equivalent:*

1. *E is an orthonormal basis for H.*
2. **(Fourier expansion):** *For each $x \in H$,*

$$x = \sum_{n=1}^{\infty}\langle x, u_n\rangle u_n,$$

where $E_x = \{u_1, u_2, \ldots\}$.

3. **(Parseval's identity):** *For each* $x \in H$,

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, u_n\rangle|^2,$$

where $E_x = \{u_1, u_2, \ldots\}$.
4. $\overline{L(E)} = H$.
5. *If for* $x \in H$ *such that* $\langle x, u\rangle = 0$ *for all* $u \in E$, *then* $x = 0$.

Proof. **(1)** $\Rightarrow$ **(2)** Suppose E is an orthonormal basis for H, i.e. E is a maximal orthonormal set in H. Let $x \in H$ and $E = \{u_1, u_2, \ldots\}$ then by the previous theorem $\sum_{n=1}^{\infty}\langle x, u_n\rangle u_n$ converges to some y in H. Let

$$y = \sum_{n=1}^{\infty}\langle x, u_n\rangle u_n \in H.$$

If $y = x$, then the required equality holds. If $y \neq x$, then by the last theorem $(x - y) \perp E$. Take

$$v = \frac{x-y}{\|x-y\|}.$$

Then $v \perp E$ and so $v \notin E$. Take $E_0 = E \cup \{v\}$. Then E_0 is an orthonormal subset of H and $E \subsetneq E_0$, which contradicts our assumption that E is a maximal orthonormal set. Therefore,

$$x = y = \sum_{n=1}^{\infty}\langle x, u_n\rangle u_n.$$

(2) $\Leftrightarrow$ **(3)** This is the proved in the equality case of Bessel's inequality.

(2) $\Rightarrow$ **(4)** Assume that for every $x \in H$, we have $x = \sum\limits_{n=1}^{\infty}\langle x, u_n\rangle u_n$, where $E_x = \{u_1, u_2, \ldots\} = \{u : \langle x, u\rangle \neq 0\}$. Take

$$x_m = \sum_{n=1}^{m}\langle x, u_n\rangle u_n.$$

Then $x_m \in E$ and clearly by our assumption, $x_m \to x$. Hence

$$\overline{L(E)} = H.$$

(4) $\Rightarrow$ **(5)** $\overline{L(E)} = H$. Let $x \in H$ such that $\langle x, u\rangle = 0$ for all $u \in E$. Consider a sequence $\{x_m\}$ in $L(E)$ such that $x_m \to x$ ($\because \overline{L(E)} = H$). Since $x_m \in L(E)$, it is of the form

$$x_m = \alpha_{1m}u_1 + \alpha_{2m}u_2 + \cdots + \alpha_{nm}u_n,$$

where $u_1, u_2, \ldots, u_n \in E$ and $\alpha_{1m}, \alpha_{2m}, \ldots, \alpha_{nm} \in K$. Then by our assumption

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

We know that $x_m \to x \Rightarrow \langle x_m, x\rangle \to \langle x, x\rangle$. Therefore, $\langle x, x\rangle = 0$ and hence $x = 0$

(5) $\Rightarrow$ **(1)** Assume that (5) holds, then we have to prove that E is a maximal orthonormal subset of H. Suppose E is not maximal, then there exists an orthonormal subset E_0 of H

such that $E \subsetneq E_0$. Let $x \in E_0$ such that $x \notin E$. Since E_0 is orthonormal and $x \notin E$, we have

$$\langle x, u \rangle = 0 \quad \forall\, u \in E.$$

Then by (5), $x = 0$ which is contradiction to our assumption that $x \in E_0$ since E_0 is orthonormal and an orthonormal set does not contain 0. Therefore, E must be a maximal orthonormal subset of H, i.e. E is an orthonormal basis for H. □

Theorem 1.3.7. *Let H be an n-dimensional Hilbert space. Then H is isometrically isomorphic to $(K^n, \|\cdot\|_2)$.*

Proof. Since $\dim H = n$, consider a basis $\{x_1, x_2, \ldots, x_n\}$ of H, i.e. the set $\{x_1, x_2, \ldots, x_n\}$ is linearly independent and it spans H. By Gram-Schmidt orthonormalization, there exists an orthonormal subset $\{u_1, u_2, \ldots, u_n\}$ of H such that

$$L(\{u_1, u_2, \ldots, u_n\}) = L(\{x_1, x_2, \ldots, x_n\}) = H.$$

Then by the previous theorem, $\{u_1, u_2, \ldots, u_n\}$ is an orthonormal basis for H. Define $T : H \to K^n$ by

$$T(x) = (\langle x, u_1 \rangle, \langle x, u_2 \rangle, \ldots, \langle x, u_n \rangle) \qquad x \in H.$$

Then T is homomorphism (i.e. linear). Now, since $u_1, u_2, \ldots, u_n$ is an orthonormal basis, by Parseval's identity, we have

$$\|T(x)\|_2^2 = \sum_{j=1}^{n} |\langle x, u_j \rangle|^2 = \|x\|^2.$$

Therefore, T is isometry. Now, let $y = (y_1, y_2, \ldots, y_n) \in K^n$. Take

$$x = y_1 u_1 + y_2 u_2 + \cdots + y_n u_n = \sum_{j=1}^{n} y_j u_j.$$

Then

$$\begin{aligned}
\langle x, u_i \rangle &= \left\langle \sum_{j=1}^{n} y_j u_j, u_i \right\rangle \\
&= \sum_{j=1}^{n} y_j \langle u_j, u_i \rangle \\
&= y_i \qquad\qquad (\because \langle u_j, u_i \rangle = 0,\ j \neq i \text{ and } \langle u_i, u_i \rangle = 1).
\end{aligned}$$

Therefore,

$$T(x) = (\langle x, u_1 \rangle, \langle x, u_2 \rangle, \ldots, \langle x, u_n \rangle) = y.$$

Thus T is an onto linear isometry. □

Definition 1.3.8. A metric space, in particular a normed linear space H, is said to be *separable* if it has a countable dense subset.

Exercise 1.3.9. Show that ℓ^p is separable for $1 \leq p < \infty$ but ℓ^∞ is not separable.

Solution. Seminar exercise. □

Theorem 1.3.10. *Let H be an infinite dimensional Hilbert space Then the following are equivalent:*

1. *H has a countable orthonormal basis.*
2. *H is isometrically isomorphic to ℓ^2.*
3. *H is separable (i.e. H has a countable dense subset).*

Proof. **(1) ⇒ (2)** Suppose H has a countable orthonormal basis, say $\{u_1, u_2, \ldots\}$. Define $T : H \to \ell^2$ by

$$T(x) = (\langle x, u_1\rangle, \langle x, u_2\rangle, \ldots) \qquad \text{for } x \in H.$$

Note that, by Bessel's inequality $\sum_{n=1}^{\infty} |\langle x, u_n\rangle|^2 < \infty$ and hence $T(x) \in \ell^2$. Also, T is (clearly) a homomorphism (i.e. T is linear). Then by the Parseval's identity, we have

$$\|T(x)\|_2^2 = \sum_{n=1}^{\infty} |\langle x, u_n\rangle|^2 = \|x\|^2.$$

Therefore, T is isometry. Now, let $y = (y_1, y_2, \ldots) \in \ell^2$, i.e. $\sum_{n=1}^{\infty} |y_n|^2 < \infty$. Then by Riesz-Fischer theorem,

$$\sum_{n=1}^{\infty} y_n u_n$$

converges in H. Suppose $x = \sum\limits_{n=1}^{\infty} y_n u_n$. Then for each $i = 1, 2\ldots$,

$$\begin{aligned} \langle x, u_i\rangle &= \left\langle \sum_{n=1}^{\infty} y_n u_n, u_i \right\rangle \\ &= \sum_{n=1}^{\infty} y_n \langle u_n, u_i\rangle \\ &= y_n \end{aligned}$$

$$T(x) = (\langle x, u_1\rangle, \langle x, u_2\rangle, \ldots) = (y_1, y_2, \ldots) = y.$$

Thus, $T : H \to \ell^2$ is an onto linear isometry. In other words, H is isometrically isomorphic to ℓ^2.

(2) ⇒ (3): Let $T : H \to \ell^2$ be a linear onto isometry. Since ℓ^2 is separable, ℓ^2 has a dense subset D. Then $T^{-1}(D)$ is a countable dense subset of H and therefore H is separable.

(3) ⇒ (1): Assume that H is separable. So it has a countable dense subset. Suppose $D = \{z_1, z_2, \ldots\}$ is a countable dense subset of H. Let i_1 be the first integer such that $z_{i_1} \neq 0$. Let $x_1 = \{z_{i_1}\}$. Clearly,

$$L(\{z_1, z_2, \ldots, z_{i_1}\}) = L(\{x_1\}).$$

Let i_2 be the first integer such that x_1 and z_{i_2} are linearly independent. Take $x_2 = z_{i_2}$. Then

$$L(\{x_1, x_2\}) = L(\{z_1, z_2, \ldots, z_{i_1}, \ldots, z_{i_2}\}).$$

Continuing this way, inductively we can choose a linearly independent subset $\{x_1, x_2, \ldots, x_n\}$ of H such that for each $n = 1, 2, \ldots$, we have

$$L(\{x_1, x_2, \ldots, x_n\}) = L(\{z_1, \ldots, z_{i_1}, \ldots, z_{i_2}, \ldots, z_{i_n}\}). \tag{1.10}$$

Then by Gram-Schmidt orthonormalization, there exists an orthonormal subset $\{u_1, u_2, \ldots\}$ of H such that

$$L(\{x_1, x_2, \ldots\}) = L(\{u_1, u_2, \ldots\}).$$

Also, since $\overline{D} = H$, $\overline{L(D)} = H$. But then

$$\begin{aligned} \overline{L(\{u_1, u_2, \ldots\})} &= \overline{L(\{x_1, x_2, \ldots\})} \\ &= \overline{L(D)} \qquad \text{(by (1.10))} \\ &= H. \end{aligned}$$

Thus, $L(\{u_1, u_2, \ldots\})$ is dense in H and hence $\{u_1, u_2, \ldots\}$ is a countable orthonormal basis for H. □

Seminar Topics 3.

1. Prove that every non-zero Hilbert space has an orthonormal basis.
2. Show that ℓ^p is separable for $1 \leq p < \infty$ but ℓ^∞ is not separable.
3. Show that $\{e_n : n \in \mathbb{N}\}$ is an orthonormal basis of ℓ^2 but it is not a basis of ℓ^2.

Chapter 2

Approximations and Riesz representation theorem

2.1 Approximation and Optimization

Definition 2.1.1. Let X be an inner product space, $E \neq \emptyset$, $E \subset X$ and $x \in X$. An element $y \in E$ is said to be *a best approximation* from E to x if

$$\|x - y\| \leq \|x - z\| \qquad \forall\ z \in E.$$

i.e.

$$\|x - y\| = \text{dist}(x, E).$$

Naturally, there are three questions, one may ask here.

1. Does best approximation always exist?
2. If it does, is it unique?
3. How does one find a best approximation?

The following remark answers the first two questions. For the answer to the third question, in what follows, we prove certain results.

Remarks 2.1.2.

1. In general, best approximation may not exist. For example, take $X = \mathbb{R}$ and $E = (0, 1) \cap \mathbb{Q}$. Then best approximation does not exist for say $x = 2$.
2. In general, best approximation may not be unique. For example, take $X = \mathbb{R}^2$, $E = \{z \in X : \|z\| = 1\}$ and $x = (0, 0)$. Then all the points of the set E are best approximations from E to $x = 0$.

Proposition 2.1.3. *Let X be an inner product space. If $E \subset X$ and $x \in \overline{E}$, then there is a best approximation from E to x if and only if $x \in E$.*

Proof. If $x \in E$ then $y = x$ (i.e. x itself) is a best approximation from E to x. In fact,

$$\operatorname{dist}(x, E) = \|x - y\| = 0.$$

Conversely, let $x \in \overline{E}$ and suppose that $y \in E$ is a best approximation from E to x. Then

$$\|x - y\| = \operatorname{dist}(x, E) = 0 \qquad (\because x \in \overline{E})$$

i.e., $x = y$ and hence $x \in E$. □

Proposition 2.1.4. *Let X be an inner product space. If $E \subset X$ is convex and $x \in X$, then there exists at most one best approximation from E to x.*

Proof. Suppose $y_1 \in E$ and $y_2 \in E$ are two best approximations from E to x, i.e.

$$\|x - y_1\| = \|x - y_2\| = \operatorname{dist}(x, E) = d.$$

Now, by Parallelogram law,

$$\begin{aligned} &\|(x - y_1) + (x - y_2)\|^2 + \|(x - y_1) - (x - y_2)\|^2 = 2\|(x - y_1)\|^2 + 2\|(x - y_2)\|^2 \\ \Rightarrow\ &\|2x - (y_1 + y_2)\|^2 + \|y_1 - y_2\|^2 = 2\|(x - y_1)\|^2 + 2\|(x - y_2)\|^2 \end{aligned}$$

Therefore

$$\begin{aligned} \|y_1 - y_2\|^2 &= 2\|(x - y_1)\|^2 + 2\|(x - y_2)\|^2 - 4\left\|x - \left(\frac{y_1 + y_2}{2}\right)\right\| \\ &\le 2d^2 + 2d^2 - 4d^2 \qquad \left(\because E \text{ is convex and } y_1, y_2 \in E \Rightarrow -\left\|x - \left(\frac{y_1 + y_2}{2}\right)\right\| \le -d\right) \\ &= 0. \end{aligned}$$

Therefore, $\boxed{y_1 = y_2}$. □

Proposition 2.1.5. *Let X be an inner product space, Y be a subspace of X and $x \in X$. Then $y \in Y$ is a best approximation from Y to x if and only if $(x - y) \perp Y$.*

Proof. Suppose $y \in Y$ such that $(x - y) \perp Y$. Then for any $z \in Y$,

$$(x - y) \perp z \qquad \text{i.e. } \langle x - y, z\rangle = 0.$$

Also, since Y is a subspace and $y, z \in Y$, $y - z \in Y$ and so

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

Therefore by Pythagoras theorem,

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

Therefore, $\|x - y\| \le \|x - z\|$ for all $z \in Y$. Hence, y is a best approximation from Y to x.

Conversely, assume that $y \in Y$ is a best approximation from Y to x. Let $z \in Y$ be such that $\|z\| = 1$. Consider $w = y + \langle x - y, z\rangle z$. Then $w \in Y$. Therefore,

$$x - w = (x - y) - \langle x - y, z\rangle z.$$

Now,

$$\begin{aligned}
\|x - y\|^2 &\le \|x - w\|^2 \qquad (\because y \text{ is best approx. and } w \in Y)\\
&= \langle x - w, x - w\rangle\\
&= \langle (x - y) - \langle x - y, z\rangle z, (x - y) - \langle x - y, z\rangle z\rangle\\
&= \|x - y\|^2 - \langle (x - y), \langle x - y, z\rangle z\rangle - \langle \langle x - y, z\rangle z, (x - y)\rangle + |\langle x - y, z\rangle|^2 \langle z, z\rangle\\
&= \|x - y\|^2 - \overline{\langle x - y, z\rangle}\langle x - y, z\rangle - \langle x - y, z\rangle\langle z, x - y\rangle + |\langle x - y, z\rangle|^2\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\because \langle z, z\rangle = 1)\\
&\le \|x - y\|^2 - |\langle x - y, z\rangle|^2.
\end{aligned}$$

Therefore, $\langle x - y, z\rangle = 0$ and hence $(x - y) \perp z$ for all $z \in Y$, i.e.

$$\boxed{(x - y) \perp Y}.$$

□

Definition 2.1.6 (Gram matrix)**.** Let X be an inner product space and $x_1, x_2, \ldots, x_n \in X$. The matrix

$$G(x_1, x_2, \ldots, x_n) = \begin{bmatrix} \langle x_1, x_1\rangle & \langle x_2, x_1\rangle & \cdots & \langle x_n, x_1\rangle\\ \langle x_1, x_2\rangle & \langle x_2, x_2\rangle & \cdots & \langle x_n, x_2\rangle\\ \vdots & \vdots & \ddots & \vdots\\ \langle x_1, x_n\rangle & \langle x_2, x_n\rangle & \cdots & \langle x_n, x_n\rangle \end{bmatrix}$$

is known as the *Gram matrix* of $x_1, x_2, \ldots, x_n$.

Remarks 2.1.7.

1. $x_1, x_2, \ldots, x_n$ are orthogonal if and only if the Gram matrix is a diagonal matrix.
2. $x_1, x_2, \ldots, x_n$ are orthonormal if and only if the Gram matrix is the identity matrix.

Lemma 2.1.8. *Let X be an inner product space and $x_1, x_2, \ldots, x_n \in X$ be linearly independent. Then the Gram matrix of $x_1, x_2, \ldots, x_n$ is regular.*

Proof. The Gram matrix of $x_1, x_2, \ldots, x_n$ is

$$M = \begin{bmatrix} \langle x_1, x_1\rangle & \langle x_2, x_1\rangle & \cdots & \langle x_n, x_1\rangle\\ \langle x_1, x_2\rangle & \langle x_2, x_2\rangle & \cdots & \langle x_n, x_2\rangle\\ \vdots & \vdots & \ddots & \vdots\\ \langle x_1, x_n\rangle & \langle x_2, x_n\rangle & \cdots & \langle x_n, x_n\rangle \end{bmatrix}.$$

Claim: Column rank of M is n.

Let $a_1, a_2, \ldots, a_n \in K$ be such that

$$a_1 \begin{bmatrix} \langle x_1, x_1\rangle \\ \langle x_1, x_2\rangle \\ \cdots \\ \langle x_1, x_n\rangle \end{bmatrix} + a_2 \begin{bmatrix} \langle x_2, x_1\rangle \\ \langle x_2, x_2\rangle \\ \cdots \\ \langle x_2, x_n\rangle \end{bmatrix} + \cdots + a_n \begin{bmatrix} \langle x_n, x_1\rangle \\ \langle x_n, x_2\rangle \\ \cdots \\ \langle x_n, x_n\rangle \end{bmatrix} = 0.$$

Then for each $i = 1, 2, \ldots, n$,

$$\sum_{j=1}^{n} a_j \langle x_j, x_i\rangle = 0. \tag{2.1}$$

Now,

$$\begin{aligned} \left\| \sum_{j=1}^{n} a_j x_j \right\|^2 &= \left\langle \sum_{j=1}^{n} a_j x_j, \sum_{i=1}^{n} a_i x_i \right\rangle \\ &= \sum_{i=1}^{n} \overline{a_i} \left\langle \sum_{j=1}^{n} a_j x_j, x_i \right\rangle \\ &= \sum_{i=1}^{n} \overline{a_i} \left(\sum_{j=1}^{n} a_j \langle x_j, x_i\rangle \right) \\ &= 0 \qquad\qquad \text{(by (2.1))}. \end{aligned}$$

Therefore,

$$\sum_{j=1}^{n} a_j x_j = 0.$$

Since $x_1, x_2, \ldots, x_n$ are linearly independent, $a_1 = a_2 = \cdots = a_n = 0$. Therefore, the columns of M are linearly independent which means that the column rank of M is n. Hence, M is regular. □

Theorem 2.1.9. *Let X be an inner product space and $x_1, x_2, \ldots, x_n \in X$ be linearly independent and $x \in X$. Let $Y = L(\{x_1, x_2, \ldots, x_n\})$, then $y = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$ is a best approximation from Y to x, where $\alpha_1, \alpha_2, \ldots, \alpha_n$ form the unique solution of the normal equations.*

$$\begin{aligned} \alpha_1\langle x_1, x_1\rangle + \alpha_2\langle x_2, x_1\rangle + \cdots + \alpha_n\langle x_n, x_1\rangle &= \langle x, x_1\rangle \\ \alpha_1\langle x_1, x_2\rangle + \alpha_2\langle x_2, x_2\rangle + \cdots + \alpha_n\langle x_n, x_2\rangle &= \langle x, x_2\rangle \\ \vdots \qquad\qquad &\qquad \vdots \\ \alpha_1\langle x_1, x_n\rangle + \alpha_2\langle x_2, x_n\rangle + \cdots + \alpha_n\langle x_n, x_n\rangle &= \langle x, x_n\rangle \end{aligned} \tag{2.2}$$

Proof. Consider the normal equations (2.2). Now if $y = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \in Y$ is a best approximation from Y to x, then by Proposition 2.1.5 $(x - y) \perp Y$. That is,

$$\langle (x - y), x_j\rangle = 0 \qquad \text{for } j = 1, 2, \ldots, n.$$

That is,

$$\langle y, x_j\rangle - \langle x, x_j\rangle = 0 \qquad \text{for } j = 1, 2, \ldots, n.$$

That is,

$$\alpha_1\langle x_1, x_j\rangle + \alpha_2\langle x_2, x_j\rangle + \cdots + \alpha_n\langle x_n, x_j\rangle = \langle x, x_j\rangle.$$

That is, $\alpha_1, \alpha_2, \ldots, \alpha_n$ are solutions of the normal equations (by (2.2)). The solution is unique (since the Gram matrix M is regular). □

Theorem 2.1.10. *Let X be an inner product space, $x_1, x_2, \ldots, x_n \in X$ be linearly independent, $c_1, c_2, \ldots, c_n \in K$ and $x \in X$. Consider the set*

$$E = \{y \in X : \langle y, x_i\rangle = c_i,\ i = 1, 2, \ldots, n\}.$$

Then the unique best approximation from E to x is given by

$$y = x + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n, \tag{2.3}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n \in K$ form the unique solution of the equations

$$\begin{array}{rcl} \alpha_1\langle x_1, x_1\rangle + \alpha_2\langle x_2, x_1\rangle + \cdots + \alpha_n\langle x_n, x_1\rangle & = & c_1 - \langle x, x_1\rangle \\ \alpha_1\langle x_1, x_2\rangle + \alpha_2\langle x_2, x_2\rangle + \cdots + \alpha_n\langle x_n, x_2\rangle & = & c_2 - \langle x, x_2\rangle \\ \vdots \qquad\qquad \vdots & & \vdots \\ \alpha_1\langle x_1, x_n\rangle + \alpha_2\langle x_2, x_n\rangle + \cdots + \alpha_n\langle x_n, x_n\rangle & = & c_n - \langle x, x_n\rangle \end{array} \tag{2.4}$$

Proof. Since $x_1, x_2, \ldots, x_n$ are linearly independent, the Gram matrix for $x_1, x_2, \ldots, x_n$ is regular. So the system (2.4) has a unique solution, say $\alpha_1, \alpha_2, \ldots, \alpha_n$, i.e. for $i = 1, 2, \ldots, n$

$$\alpha_1\langle x_1, x_i\rangle + \alpha_2\langle x_2, x_i\rangle + \cdots + \alpha_n\langle x_n, x_i\rangle = c_i - \langle x, x_i\rangle.$$

If $y = x + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$, then for $i = 1, 2, \ldots, n$

$$\begin{aligned} \langle y, x_i\rangle &= \langle x + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n, x_i\rangle \\ &= \langle x, x_i\rangle + \alpha_1\langle x_1, x_i\rangle + \cdots + \alpha_n\langle x_n, x_i\rangle \\ &= \langle x, x_i\rangle + c_i - \langle x, x_i\rangle \qquad (\because \alpha_1, \ldots, \alpha_n \text{ is the solution of (2.4)}) \\ &= c_i. \end{aligned}$$

Therefore $y \in E$.
Claim: $E - y$ is a subspace of X.

Let $z_1, z_2 \in E - y$, then there exists $u_1, u_2 \in E$ such that $z_1 = u_1 - y$, $z_2 = u_2 - y$ and $\langle u_1, x_i\rangle = c_i$ and $\langle u_2, x_i\rangle = c_i$, $i = 1, 2, \ldots, n$. So, $z_1 + z_2 = u_1 + u_2 - 2y$. Therefore, for $i = 1, 2, \ldots, n$,

$$\begin{aligned} \langle z_1 + z_2 + y, x_i\rangle &= \langle u_1 + u_2 - y, x_i\rangle \\ &= \langle u_1, x_i\rangle + \langle u_2, x_i\rangle - \langle y, x_i\rangle \\ &= c_i + c_i - c_i \\ &= c_i. \end{aligned}$$

Thus, $z_1 + z_2 + y \in E$ and so $z_1 + z_2 \in E - y$. Now, let $\alpha \in K$ and $z = u - y \in E - y$ for $u \in E$. For $i = 1, 2, \ldots, n$,

$$\langle \alpha z + y, x_i\rangle = \langle \alpha u - \alpha y + y, x_i\rangle$$

$$
\begin{aligned}
&= \langle \alpha u + (1-\alpha)y, x_i\rangle \\
&= \alpha\langle u, x_i\rangle + (1-\alpha)\langle y, x_i\rangle \\
&= \alpha c_i + (1-\alpha)c_i \\
&= c_i.
\end{aligned}
$$

Therefore, $\alpha z + y \in E$, i.e. $\alpha z \in E - y$ and hence $E - y$ is a subspace of X.

Now, y is a best approximation from E to x if and only if 0 is a best approximation form $E - y$ to $x - y$.

$$(\because \|x - y\| = \operatorname{dist}(E, x) = \operatorname{dist}(E - y, x - y) = \|(x - y) - 0\|).$$

Now for $z \in E$,

$$
\begin{aligned}
\langle z - y, x - y\rangle &= \langle z - y, -\alpha_1 x_1 - \alpha_2 x_2 - \cdots - \alpha_n x_n\rangle && \text{(by (2.3))} \\
&= \langle z, -\alpha_1 x_1 - \alpha_2 x_2 - \cdots - \alpha_n x_n\rangle - \langle y, -\alpha_1 x_1 - \alpha_2 x_2 - \cdots - \alpha_n x_n\rangle \\
&= -\alpha_1 c_1 - \alpha_2 c_2 - \cdots - \alpha_n c_n + \alpha_1 c_1 + \alpha_2 c_2 + \cdots + \alpha_n c_n && (\because y \in E) \\
&= 0,
\end{aligned}
$$

i.e. $(x - y) \perp E - y$. Hence by Proposition 2.1.5, 0 is a best approximation from $E - y$ to $x - y$. Since, $E - y$ is a subspace, it is convex and hence 0 is the unique best approximation from $E - y$ to $x - y$ or y is the unique best approximation from E to x. □

Theorem 2.1.11. *Let H be a Hilbert space, $x_1, x_2, \ldots \in H$ be linearly independent and $x \in H$. Let*

$$Y = \overline{L(\{x_1, x_2, \ldots\})}$$

and let $\{u_1, u_2, \ldots\}$ be orthonormal subset of H obtained by applying Gram-Schmidt process to $x_1, x_2, \ldots$. For $m = 1, 2, \ldots$, consider the subspace

$$Y_m = L(\{x_1, x_2, \ldots, x_m\})$$

and

$$y_m = \sum_{n=1}^{m} \langle x, u_n\rangle u_n.$$

Then y_m is a unique best approximation from Y_m to x. Suppose

$$y = \sum_{n=1}^{\infty} \langle x, u_n\rangle u_n.$$

Then y is a unique best approximation from Y to x. Also,

$$\operatorname{dist}(x, Y) = \left(\|x\|^2 - \sum_{n=1}^{\infty} |\langle x, u_n\rangle|^2\right)^{\frac{1}{2}}.$$

Proof. For $m = 1, 2, \ldots$, we have

$$Y_m = L(\{x_1, x_2, \ldots, x_m\}) = L(\{u_1, u_2, \ldots, x_m\}).$$

Now, for $k = 1, 2, \ldots, m$,

$$\begin{aligned}
\langle x - y_m, u_k \rangle &= \langle x, u_k \rangle - \langle y_m, u_k \rangle \\
&= \langle x, u_k \rangle - \left\langle \sum_{n=1}^{m} \langle x, u_n \rangle u_n, u_k \right\rangle \\
&= \langle x, u_k \rangle - \sum_{n=1}^{m} \langle x, u_n \rangle \langle u_n, u_k \rangle \\
&= \langle x, u_k \rangle - \langle x, u_k \rangle \\
&= 0.
\end{aligned}$$

Therefore, $(x - y_m) \perp Y_m$ for $m = 1, 2, \ldots$. Thus, y_m is a unique best approximation from Y_m to x (since Y_m a subspace and hence it is convex). Also,

$$Y = L(\{x_1, x_2, \ldots\}) = L(\{u_1, u_2, \ldots\}).$$

By Theorem 1.3.5 (using Bessel's inequality and Riesz-Fischer theorem),

$$y = \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n$$

converges in H (since H is Hilbert space).

Now, for $k = 1, 2, \ldots$,

$$\begin{aligned}
\langle x - y, u_k \rangle &= \langle x, u_k \rangle - \langle y, u_k \rangle \\
&= \langle x, u_k \rangle - \left\langle \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n, u_k \right\rangle \\
&= \langle x, u_k \rangle - \sum_{n=1}^{\infty} \langle x, u_n \rangle \langle u_n, u_k \rangle \\
&= \langle x, u_k \rangle - \langle x, u_k \rangle \\
&= 0.
\end{aligned}$$

Therefore, $(x - y) \perp Y$. This, y is a unique best approximation from Y to x (since Y is a subspace). In fact,

$$\begin{aligned}
\text{dist}(x, Y)^2 &= \|x - y\|^2 \\
&= \langle x - y, x - y \rangle \\
&= \langle x, x - y \rangle - \langle y, x - y \rangle \\
&= \langle x, x - y \rangle && (\because (x - y) \perp Y) \\
&= \left\langle x, x - \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n \right\rangle \\
&= \|x\|^2 - \sum_{n=1}^{\infty} |\langle x, u_n \rangle|^2.
\end{aligned}$$

□

Theorem 2.1.12. *Let H be a Hilbert space, $E \subset H$ be a closed convex subset of H and $x \in H$. Then there is a unique best approximation from E to x.*

Proof. Since $d = \text{dist}(x, E) = \inf\{\|y - x\| : y \in E\}$, there is a sequence $\{y_n\}$ in E such that $\|y_n - x\| \to d$. Now for $n, m = 1, 2, \ldots$, by the Parallelogram law

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

Therefore,

$$\begin{aligned}\|y_n - y_m\|^2 &= 2(\|x - y_n\|^2 + \|x - y_m\|^2) - \|2x - (y_n + y_m)\|^2 \\ &= 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4\left\|x - \left(\frac{y_n + y_m}{2}\right)\right\|^2.\end{aligned}$$

Since E is convex and $y_n, y_m \in E$, $\frac{y_n+y_m}{2} \in E$ (taking $t = \frac{1}{2}$) and so by the above equation,

$$\left\|x - \left(\frac{y_n + y_m}{2}\right)\right\|^2 \geq d^2.$$

Therefore,

$$\begin{aligned}\|y_n - y_m\|^2 &\leq 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4d^2 \\ &\to 2d^2 + 2d^2 - 4d^2 = 0 \qquad \text{as } n, m \to \infty.\end{aligned}$$

Hence, $\{y_n\}$ is a Cauchy sequence in E. Since E is a closed subset of a complete (Hilbert) space H, E is complete. Then there is $y \in E$ such that $y_n \to y$ in E. Therefore,

$$\|x - y\| = \lim_{n\to\infty} \|x - y_n\| = d = \text{dist}(x, E).$$

Therefore, y is a best approximation from E to x. Since E is convex, the best approximation from E to x is unique. □

Corollary 2.1.13. *Let H be a Hilbert space and E be a closed convex subset of H. Then E contains a unique vector y of minimum norm.*

Proof. Take $x = 0$. Then by Theorem 2.1.12, there exists $y \in E$ which is the unique best approximation from E to x, i.e.,

$$\begin{aligned}\|y\| = \|y - x\| &= \text{dist}(x, E) \\ &= \inf\{\|x - z\| : z \in E\} \\ &= \inf\{\|z\| : z \in E\} \qquad (\because x = 0).\end{aligned}$$

□

2.2 Projection

Definition 2.2.1. Let $(X, \|\cdot\|), (y, \|\cdot\|)$ be two normed linear spaces and $T : X \to Y$ be a linear transformation. The *kernel* or *zero space* of T is

$$\ker(T) = \{x \in H : Tx = 0\}$$

and the *range of* T is

$$R(T) = \{Tx : x \in X\}.$$

Definition 2.2.2. Let H be a Hilbert space. A linear transformation $T : H \to H$ is said to be a *projection* (or *idempotent*) if $T^2 = T$.

If T is a projection, then

$$R(T) = \text{Range of } T = \{x \in H : Tx = x\}$$

Indeed, for $x \in R(T)$, there exists $y \in H$ such that $Ty = x$. Therefore,

$$x = Ty = T^2y = T(Ty) = Tx.$$

On the other hand if $x = Tx$, then clearly, $x \in R(T)$.

Definition 2.2.3. A projection on a Hilbert space H is called *orthogonal projection* if $R(T) \perp \ker(T)$ i.e., if $y \in R(T)$ and $x \in \ker(T)$ then $\langle x, y\rangle = 0$.

For a subset E of a Hilbert space H, the set

$$E^\perp = \{y \in H : \langle x, y\rangle = 0 \ \forall\ x \in E\}$$

is called the *orthogonal complement* of E.

Note that if $E = \emptyset$, then $E^\perp = H$.

Proposition 2.2.4. *Let H be a Hilbert space and $E \subset H$, then $E^\perp$ is a closed subspace of H.*

Proof. Since $\emptyset^\perp = H$ is closed, we assume that $E \neq \emptyset$. Let $x_1, x_2 \in E^\perp$. Then $\langle x_1, y\rangle = 0$ and $\langle x_2, y\rangle = 0$ for all $y \in E$. Therefore,

$$\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle = 0 \qquad \forall\ y \in E.$$

Also, for $\alpha \in K$ and $x \in E^\perp$,

$$\langle \alpha x, y\rangle = \alpha\langle x, y\rangle = 0 \qquad \forall\ y \in E.$$

Thus $E^\perp$ is a subspace of H.

Now, suppose $\{x_n\}$ is a sequence in $E^\perp$ such that $x_n \to x$ in H. Then for all $y \in E$,

$$0 = \langle x_n, y\rangle \to \langle x, y\rangle \Rightarrow \langle x, y\rangle = 0 \Rightarrow x \in E^\perp.$$

Therefore, $E^\perp$ is closed in H. □

Example 2.2.5. Consider the line $\ell : y = 2x$ and $m : y = \frac{3}{4}x$. Now the map $P : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $P(x, y) = (\frac{3}{5}x + \frac{4}{5}y, \frac{6}{5}x + \frac{8}{5}y)$ is a projection.

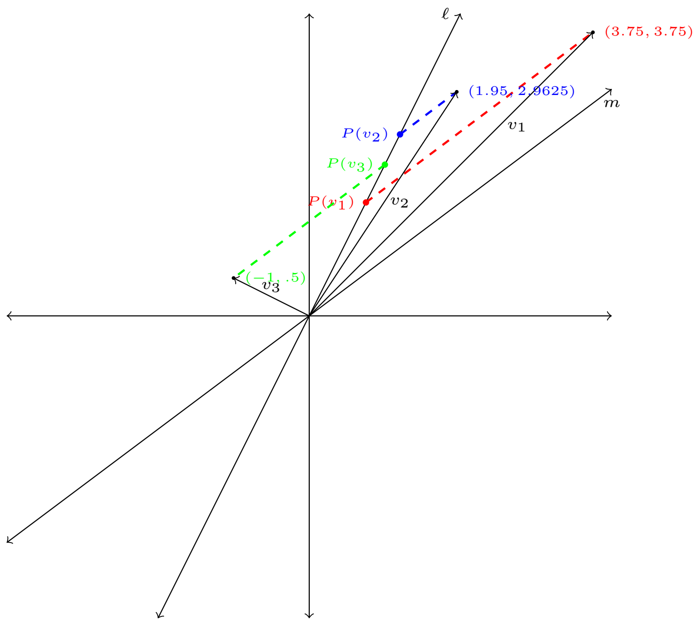


Figure 2.1: Projection on ℓ along m

Theorem 2.2.6 (Projection theorem)**.** *Let H be a Hilbert space and Y be a closed subspace of H. Then*

$$Y \oplus Y^{\perp} = H \quad \text{and} \quad Y^{\perp\perp} = Y,$$

where $Y^{\perp\perp} = (Y^{\perp})^{\perp}$.

Proof. Clearly, $H^{\perp} = \{0\}$ and $\{0\}^{\perp} = H$. So we take $Y \neq \{0\}$. Since Y is a closed subspace of a Hilbert space H, Y is a non-zero Hilbert space. So, Y has an orthonormal basis E.

Now let $x \in H$. Then the set

$$E_x = \{u \in E : \langle x, u\rangle \neq 0\}$$

is countable, say $E_x = \{u_1, u_2, \ldots\}$. Also, since H is a Hilbert space, $\sum_{n=1}^{\infty}\langle x, u_n\rangle u_n$ converges to some y in Y and $(x - y) \perp E$. Since E is an orthonormal basis of Y, we have

$$(x - y) \perp Y,$$

i.e., $z = (x - y) \in Y^{\perp}$, so that $x = y + z$ with $y \in Y$ and $z \in Y^{\perp}$. Therefore,

$$H = Y + Y^{\perp}.$$

Now, suppose $x \in Y \cap Y^{\perp}$. Then $x \in Y^{\perp}$ and hence $\langle x, y\rangle = 0$ for all $y \in Y$. Since $x \in Y$, in particular taking $y = x$, we get $\langle x, x\rangle = 0 \Rightarrow x = 0$. So, $Y \cap Y^{\perp} = \{0\}$. Thus,

$$\boxed{H = Y \oplus Y^{\perp}}.$$

Now, we show that $Y = Y^{\perp\perp}$. If $x \in Y$, then $\langle x, z\rangle = 0$ for all $z \in Y^{\perp}$. Therefore, $x \in (Y^{\perp})^{\perp}$, i.e.

$$Y \subset Y^{\perp\perp}.$$

Now, suppose $x \in Y^{\perp\perp}$. Since $H = Y \oplus Y^{\perp}$, there exists $y \in Y$ and $z \in Y^{\perp}$ such that

$$x = y + z.$$

Since $y \in Y \subset Y^{\perp\perp}$ and $x \in Y^{\perp\perp}$, we have

$$(x - y) = z \in Y^{\perp\perp}.$$

Therefore,

$$z \in Y^{\perp} \cap Y^{\perp\perp} = \{0\} \Rightarrow z = 0 \Rightarrow x = y.$$

Thus, $x = y \in Y$ and so $Y^{\perp\perp} \subset Y$. Hence,

$$\boxed{Y^{\perp\perp} = Y}.$$

□

Proposition 2.2.7. *Let H be a Hilbert space and Y be a closed subspace of H. Then there is an orthogonal projection P on H such that $R(P) = Y$ and $\ker(P) = Y^{\perp}$.*

Proof. Since, $H = Y \oplus Y^{\perp}$, for $x \in H$ there are unique $x_1 \in Y$ and $x_2 \in Y^{\perp}$ such that $x = x_1 + x_2$. Define $P : H \to H$ by $Px = x_1$. Since x_1 is uniquely associated with x, P is well-defined.

Now, let $x, x' \in H$. Then there exist $x_1, x_1{}' \in Y$ and $x_2.x_2{}' \in Y^{\perp}$ such that

$$x = x_1 + x_2 \quad \text{and} \quad x' = x_1' + x_2'.$$

Then, $x + x' = (x_1 + x_1') + (x_2 + x_2')$. Therefore,

$$P(x + x') = x + x_1' = Px + Px'.$$

Similarly, $P(\alpha x) = \alpha x_1 = \alpha P x$ for $\alpha \in K$ and $x \in H$. Therefore, $P : H \to H$ is linear.

Now, let $x = x_1 + x_2 \in H$ with $x_1 \in Y$ and $x_2 \in Y^{\perp}$. Observe that $x_1 \in (Y \subset)H$. So, $x_1 = x_1 + 0$, where we consider $0 \in Y^{\perp}$ and so by the definition of P, $Px_1 = x_1$. Therefore, we have

$$P^2 x = P(Px) = P(x_1) = x_1 = Px,$$

i.e. $P^2 = P$ and hence P is a projection on H.

Now, clearly $R(P) \subset Y$. If $x_1 \in Y$, then as before $x_1 = x_1 + 0$ and so $Px_1 = x_1 \in R(P)$. Therefore $R(P) = Y$. Also,

$$\begin{aligned}\ker P &= \{x \in H : P(x) = 0\} \\ &= \{x \in H : x = x_1 + x_2, x_1 \in Y,\ x_2 \in Y^{\perp},\ x_1 = 0\} \\ &= \{x \in H : x = x_2,\ x_2 \in Y^{\perp}\} = Y^{\perp}.\end{aligned}$$

□

Note: This P in the above proposition is called the *orthogonal projection* associated to a closed subspace Y of a Hilbert space H.

2.2.1 Continuous linear functionals

Definition 2.2.8. Let X and Y be normed linear spaces. A linear transformation $T : X \to Y$ is called *bounded* if there exists $\beta > 0$ such that

$$\|Tx\| \leq \beta\|x\| \qquad \forall\, x \in X.$$

The set of all bounded linear transformation from X to Y is denoted by $BL(X, Y)$. In this case, we define,

$$\|T\| = \sup\{\|Tx\| : x \in X,\ \|x\| \leq 1\}.$$

$BL(X, X)$ is denoted by $BL(X)$. Also, $BL(X, \mathbb{K})$ is denoted by X', called the dual of X. Elements of X' are called *bounded linear functionals on* X.

Remark 2.2.9. Let X and Y be normed linear spaces. and $T \in BL(X, Y)$. Let $0 \neq x \in X$ and $y = \frac{x}{\|x\|}$. Then $\|y\| = 1$ and so, $\|Ty\| \leq \|T\|$. This gives

$$\|Tx\| \leq \|T\|\|x\| \qquad \forall\, x \in H.$$

Proposition 2.2.10. *Let X and Y be normed linear spaces and $T : X \to Y$ be a linear map. Then T is bounded if and only if T is continuous at 0. (In fact, T is uniformly continuous).*

Proof. Suppose T is a bounded linear map. Then there exists $\beta > 0$ such that

$$\|Tx\| \leq \beta\|x\| \qquad \forall\, x \in X.$$

Therefore,

$$\|Tx - Ty\| = \|T(x - y)\| \leq \beta\|x - y\| \ \forall\, x, y \in X.$$

(Take $\|x - y\| < \frac{\epsilon}{\beta}$. Then $\|Tx - Ty\| < \epsilon$). Therefore, T is uniformly continuous. In particular, T is continuous at 0.

Next, suppose that T is continuous at 0. Then for $\epsilon > 0$, there exists $\delta > 0$ such that $\|Tx\| < \epsilon$ whenever $x \in X$ and $\|x\| < \delta$ (since $T(0) = 0$). Now, let $x \in X$, $x \neq 0$. Take $y = \frac{x}{2\|x\|}\delta$ then $\|y\| = \frac{\delta}{2} < \delta$. Therefore,

$$\|Ty\| < \epsilon.$$

$$\therefore \|Tx\| < \frac{2\epsilon}{\delta}\|x\| \qquad \forall\, x \in X.$$

Thus $\|Tx\| \leq \beta\|x\|$ for all $x \in X$. Therefore, T is a bounded linear map. □

Definition 2.2.11. Let X and Y be normed linear spaces. The collection of all bounded linear transformations $T : X \to Y$ is denoted by $BL(X, Y)$.

Exercise 2.2.12. $(BL(X, Y), \|\cdot\|)$ is a normed linear space.

Solution. Seminar exercise. □

Remarks 2.2.13. 1. If $Y = X$, then we denote $BL(X, X)$ by $BL(X)$.

2. For $S, T \in BL(X)$,

$$\|ST(x)\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|.$$

Therefore, $ST \in BL(X)$ and $\|ST\| \leq \|S\|\|T\|$.

Exercise 2.2.14. If $\{S_n\}$ and $\{T_n\}$ are sequences in $BL(X)$ such that $S_n \to S$ and $T_n \to T$ then

1. $S_n + T_n \to S + T$.
2. $S_n T_n \to ST$ and $\alpha S_n \to \alpha S$, $\alpha \in K$.

Proposition 2.2.15. *Let X be an inner product space and $T \in BL(X)$. Then*

$$\|T\| = \sup\{|\langle Tx, y\rangle| : x, y \in X,\ \|x\| \leq 1,\ \|y\| \leq 1\}.$$

Proof. **Claim:** $\|Tx\| = \sup\{|\langle Tx, y\rangle| : y \in X, \|y\| \leq 1\}$, for $x \in X$.

For if $y \in X$ with $\|y\| \leq 1$, then for all $x \in X$,

$$|\langle Tx, y\rangle| \leq \|Tx\|\|y\| \leq \|Tx\|.$$

Now if $Tx \neq 0$, take $y = \frac{Tx}{\|Tx\|}$, then $\|y\| = 1$ and

$$\begin{aligned}\langle Tx, y\rangle &= \left\langle Tx, \frac{Tx}{\|Tx\|}\right\rangle \\ &= \frac{1}{\|Tx\|}\langle Tx, Tx\rangle \\ &= \frac{\|Tx\|^2}{\|Tx\|} = \|Tx\|.\end{aligned}$$

Therefore,

$$\|Tx\| = \sup\{|\langle Tx, y\rangle| : y \in X,\ \|y\| \leq 1\}.$$

Therefore,

$$\begin{aligned}\|T\| =&\ \sup\{\|Tx\| : x \in X, \|x\| \leq 1\} \\ =&\ \sup\{\sup\{\|\langle Tx, y\rangle| : y \in X,\ \|y\| \leq 1\} : x \in X,\ \|x\| \leq 1\} \\ =&\ \sup\{|\langle Tx, y\rangle| : x, y \in X,\ \|x\| \leq 1,\ \|y\| \leq 1.\end{aligned}$$

□

Notation: Let X be a normed linear space. We denote $BL(X, K)$ by X' and is known as the continuous dual of X, i.e. the elements of X' are bounded (and hence continuous) linear functional on X.

Proposition 2.2.16. *Let X be an inner product space and $y \in X$. Define $f : X \to K$ by*

$$f(x) = \langle x, y\rangle \qquad x \in X.$$

Then f is a bounded linear functional on X and $\|f\| = \|y\|$.

Proof. Clearly, $f : X \to K$ is a linear functional on X. Also for all $x \in X$,

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

Therefore, f is bounded linear functional on X and $\|f\| \leq \|y\|$. If $y \neq 0$, then take $x = \frac{y}{\|y\|}$. Then $\|x\| = 1$ and

$$f(x) = \langle x, y\rangle = \left\langle \frac{y}{\|y\|}, y \right\rangle = \|y\|.$$

Therefore, $\|f\| = \|y\|$. □

Seminar Topics 4.

1. Let H be a Hilbert space and $P : H \to H$ be a projection. Show that $I - P$ is also a projection.
2. Consider the standard basis $B_1 = \{e_1, e_2, \ldots\}$ of c_{00}. Let B be a basis of ℓ^2 such that $B_1 \subset B$. Show that there is a unique projection $P : \ell^2 \to \ell^2$ such that $P(x) = \begin{cases} 0, & \text{if } x \in B_1 \\ 1, & \text{if} x \in B \setminus B_1. \end{cases}$
3. Find the zero space of above projection, and hence, show that it is discontinuous.
4. Define orthogonal projections $\mathbb{R}^3 \to \mathbb{R}^3$ with the following zero spaces.
   (i) $\{(x(1), x(2), x(3)) \in \mathbb{R}^3 : x(1) + x(2) = 0\}$
   (ii) $\{(x(1), x(2), x(3)) \in \mathbb{R}^3 : x(1) + x(2) + x(3) = 0\}$
   (iii) $\{(x(1), x(2), x(3)) \in \mathbb{R}^3 : x(1) = x(2) + x(3) = 0\}$
5. Show that every orthogonal projection is continuous.
6. Show that there is a unique nonzero orthogonal projection $P : \mathbb{R}^3 \to \mathbb{R}^3$ such that $P(x) = 0$ for every $x \in \{(x(1), x(2), x(3)) \in \mathbb{R}^3 : x(1) + x(2) = 1, x(3) = 0\}$.
7. For normed linear spaces X, Y, show that $(BL(X, Y), \| \cdot \|)$ is a normed linear space.
8. For normed linear spaces X, Y and $T \in BL(X, Y)$, show that

$$\|T\| = \inf \beta > 0 : \|Tx\| \leq \beta\|x\| textforallx \in X.$$

9. Let X be a normed linear space and $\lambda \in \mathbb{K}$. Define $T : X \to X$ by $T(x) = \lambda x$, $(x \in X)$. Show that $T \in BL(X)$.
10. Show that composition of two bounded linear transformations, if exists, is a bounded linear transformation.
11. Let X be a normed linear space and $\{S_n\}$ and $\{T_n\}$ be sequences in $BL(X)$ such that $S_n \to S$ and $T_n \to T$. Show that
    (i) $S_n + T_n \to S + T$.
    (ii) $S_n T_n \to ST$ and $\alpha S_n \to \alpha S$, $\alpha \in K$.

2.3 Riesz-Representation Theorem

Theorem 2.3.1 (Riesz-representation theorem). *Let H be a Hilbert space and $f \in H'$ (i.e. f is (continuous) bounded linear functional). Then there is a unique $y \in H$ such* *that $f(x) = \langle x, y \rangle$, $x \in H$. In fact,*

$$y = \frac{\overline{f(z)}\, z}{\|z\|^2}$$

for some $z \in (\ker f)^\perp$.

Proof. Let $Y = \ker f$. Since f is continuous and linear, clearly Y is a closed subspace of H (since $\{0\}$ is closed, $Y = f^{-1}(\{0\})$ is closed). If $f = 0$ then take $y = 0$. So, we assume that $f \neq 0$ and so $Y \neq H$. Then by the projection theorem,

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

As $Y \neq H$, we have $Y^\perp \neq \{0\}$. Consider an element $z \in Y^\perp$ such that $z \neq 0$. Let $x \in H$. Take $w = f(x)z - f(z)x$. Then since f is linear,

$$f(w) = f(x)f(z) - f(z)f(x) = 0,$$

i.e. $w \in \ker f$ and therefore $\langle w, z \rangle = 0$ ($\because w \in Y,\ z \in Y^\perp$). Therefore,

$$\begin{aligned} \langle f(x)z - f(z)x, z \rangle &= 0 \\ \Rightarrow f(x)\langle z, z \rangle - f(z)\langle x, z \rangle &= 0. \end{aligned}$$

Therefore,

$$f(x) = \frac{f(z)}{\|z\|}\langle x, z \rangle = \left\langle x, \frac{\overline{f(z)}\, z}{\|z\|^2} \right\rangle.$$

Take $y = \frac{\overline{f(z)}\, z}{\|z\|^2}$, then $f(x) = \langle x, y \rangle$ for $x \in H$.

Now, we show uniqueness of y. Suppose there exists $y_1 \in H$ such that $f(x) = \langle x, y_1 \rangle$ for $x \in H$. Then

$$\langle y - y_1, y \rangle = f(y - y_1) = \langle y - y_1, y_1 \rangle.$$

Therefore, $\langle y - y_1, y - y_1 \rangle = 0$ and hence $y = y_1$. Thus, y is unique. □

Notation: The unique y (in the above theorem) corresponding to $f \in H'$ is called the *representor of f* and it is denoted by y_f.

Proposition 2.3.2. *Let H be a Hilbert space and $f \in H'$. Let y_f be the representor of f. Then $\|f\| = \|y_f\|$.*

Proof. For $x \in H$,

$$\begin{aligned} |f(x)| &= |\langle x, y_f \rangle| \\ &\leq \|x\|\|y_f\|. \end{aligned}$$

Hence,

$$\|f\| \leq \|y_f\|.$$

On the other hand

$$f\left(\frac{y_f}{\|y_f\|}\right) = \frac{1}{\|y_f\|}\langle y_f, y_f\rangle$$
$$= \|y_f\|.$$

Thus,

$$\|y_f\| \leq \|f\|.$$

□

Example 2.3.3. Let $H = (K^n, \|\cdot\|_2)$. If $f \in H'$ (i.e. $f : H \to K$ is continuous linear functional) then by the Riesz representaion theorem there exists $y = (y_1, y_2, \ldots, y_n) \in K^n$ such that for $x = (x_1, x_2, \ldots, x_n) \in H$,

$$f(x) = \langle x, y\rangle = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$$
$$= \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n,$$

where $\alpha_i = \overline{y_i}$.

Example 2.3.4. Completeness of the space is essential in the Projection theorem.

Solution. Consider the space $X = c_{00}$, the space of all sequence in K having finitely many non-zero terms. For $x, y \in c_{00}$, define

$$\langle x, y\rangle = \sum_{n=1}^{\infty} x_n\overline{y(n)}.$$

Then $\langle\cdot,\cdot\rangle$ is an inner product on X.

Note that the sequence $(1, \frac{1}{2}, \ldots, \frac{1}{n}, 0, 0, \ldots)$ is a Cauchy sequence but it is not convergent. Therefore, the space $X = c_{00}$ is not complete. Define $f : X \to K$ by

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n}x_n \qquad x = (x_n) \in c_{00}.$$

Then f is a linear functional on X. Also,

$$|f(x)|^2 = \left|\sum_{n=1}^{\infty}\frac{1}{n}x_n\right|^2 = \left(\sum_{n=1}^{\infty}\frac{1}{n}|x_n|\right)^2$$
$$= \left(\sum_{n=1}^{\infty}\frac{1}{n^2}\right)\left(\sum_{n=1}^{\infty}|x_n|^2\right) \qquad \text{(by Hölder's inequality)}$$
$$= \frac{\pi^2}{6}\|x\|_2^2.$$

Therefore, $|f(x)| \leq \frac{\pi}{\sqrt{6}}\|x\|_2$ for all $x \in X = c_{00}$ and hence f is a bounded linear functional on X. Therefore, $Y = \ker f$ is a closed subspace of $X = c_{00}$. Since $f(e_1) = 1$, $f \neq 0$, where $e_1 = (1, 0, 0, \ldots)$. Therefore $Y(= \ker f) \neq X$.
Claim: $Y^\perp = \{0\}$.

Let $z = (z(1), z(2), \ldots) \in Y^\perp \subset c_{00}$. Then there exists $m \in N$ such that $z(j) = 0$ for all $j > m$, i.e. $z = (z(1), z(2), \ldots, z(m), 0, 0, \ldots)$. For $1 \leq n \leq m$, take

$$x_n(j) = \begin{cases} 1 & \text{if } j = n \\ -\frac{(m+1)}{n} & \text{if } j = m+1 \\ 0 & \text{otherwise,} \end{cases}$$

i.e. $x_n = \Big(0, \ldots, 0, \underbrace{1}_{n^{\text{th}} \text{ place}}, 0, \ldots, \underbrace{0}_{m^{\text{th}}\text{place}}, -\frac{(m+1)}{n}, 0, 0, \ldots\Big)$. Therefore,

$$f(x_n) = \frac{1}{n} + \frac{-\frac{(m+1)}{n}}{(m+1)} = \frac{1}{n} - \frac{1}{n} = 0.$$

Therefore, $x_n \in \ker f = Y$ and since $z \in Y^\perp$, we have $\langle z, x_n \rangle = 0$. But

$$\langle z, x_n \rangle = z(n).$$

Therefore $z(n) = 0 \; \forall \; n \Rightarrow z = 0$. Thus, $Y^\perp = \{0\}$. Since $Y \neq X$, $Y \oplus Y^\perp \neq X$. Thus, completeness of the space X is necessary for the projection theorem to hold. □

Example 2.3.5. Completeness of the space is necessary in the Reisz-representation theorem.

Solution. Consider $X = c_{00}$ and f as in the last example, then f is a bounded linear functional on X. Suppose, if possible, there exists $y \in X = c_{00}$ such that for every $x \in X$,

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

For $m = 1, 2, \ldots$, take $e_m = (0, 0, \ldots, 0, \underbrace{1}_{m^{\text{th}}\text{place}}, 0, 0, \ldots)$, then $f(e_m) = \frac{1}{m}$ and clearly $\langle e_m, y \rangle = \overline{y(m)}$. Now,

$$\frac{1}{m} = f(e_m) = \langle e_m, y \rangle = \overline{y(m)},$$

i.e. $y = (1, \frac{1}{2}, \ldots, \frac{1}{m}, \frac{1}{m+1}, \ldots) \notin c_{00} = X$. Thus, Riesz-representation theorem does not hold without completeness of X. □

Theorem 2.3.6 (Unique Hahn-Banach extension theorem)**.** *Let H be a Hilbert space and X be a subspace of H. Let $g \in X'$, i.e. $g : X \to K$ is bounded (continuous) linear functional on X. Then there exists a unique $f \in H'$ such that $f\big|_X = g$ and $\|f\| = \|g\|$.*

Proof. Let $g \in X'$ and $Y = \overline{X}$. Then Y is a closed subspace of H. Let $x \in Y = \overline{X}$. Then there is a sequence $\{x_n\}$ in X such that $x_n \to x$. Since g is bounded, $|g(x)| \leq \|g\|\|x\|$ for every $x \in X$. Therefore,

$$\begin{aligned} |g(x_n) - g(x_m)| &= |g(x_n - x_m)| \qquad \text{(since } g \text{ is linear)} \\ &\leq \|g\|\|x_n - x_m\|. \end{aligned}$$

Since $\{x_n\}$ is Cauchy, $\{g(x_n)\}$ is Cauchy and since K is complete, $\{g(x_n)\}$ is convergent. Let

$$\alpha = \lim_{n \to \infty} g(x_n).$$

Now, suppose $\{z_n\}$ is a sequence in X such that $z_n \to x$. Then

$$\begin{aligned}|g(x_n) - g(z_n)| &= |g(x_n - z_n)| \qquad \text{(since } g \text{ is bounded linear)}\\ &\le \|g\|\|x_n - z_n\|\\ &= \|g\|\|x_n - x - z_n + x\|\\ &\le \|g\|\big(\|x_n - x\| + \|z_n - x\|\big)\\ &\to 0 \quad \text{as } n \to \infty.\end{aligned}$$

Therefore,

$$\lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} g(z_n) = \alpha.$$

Define $\overline{g}(x) = \alpha = \lim_{n\to\infty} g(x_n)$. Then $\bar{g} : Y \to K$ and $\bar{g}$ is clearly bounded linear (functional) and satisfies $\|g\| = \|\bar{g}\|$. Thus, $\bar{g} \in Y'$. Since Y is a closed subspace of a Hilbert space, Y is Hilbert space.

Then by Riesz-representation theorem there is $y \in Y$ such that

$$\bar{g}(x) = \langle x, y\rangle \quad \forall\, x \in Y$$

and

$$\|g\| = \|y\|.$$

Define $f : H \to K$ by $f(x) = \langle x, y\rangle$ for $x \in H$. Then $f \in H'$, $\|f\| = \|y\|$ and $f\big|_X = \bar{g}$ and $f\big|_X = \bar{g}\big|_X = g$. Therefore,

$$\|f\| = \|y\| = \|\bar{g}\| = \|g\|.$$

Hence, $\|f\| = \|g\|$.

To prove the uniqueness of extension f of g, consider $h \in H'$ such that $h\big|_X = g$ and $\|h\| = \|g\|$. Since h is continuous and X is closed in Y.

$$\|h\| = \|\bar{g}\| \quad \text{and} \quad h\big|_Y = \bar{g}.$$

As $h \in H'$, there exists $z \in H$ such that $h(x) = \langle x, z\rangle$, $x \in H$ and $\|h\| = \|z\|$. Therefore

$$\langle y, z\rangle = h(y) = \bar{g}(y) = \langle y, y\rangle = \|y\|^2.$$

Now,

$$\begin{aligned}\|y - z\|^2 &= \|y\|^2 - 2\,\mathrm{Re}\langle y, z\rangle + \|z\|^2\\ &= \|y\|^2 - 2\|y\|^2 + \|y\|^2\\ &= 0 \qquad (\because \|z\| = \|h\| = \|\bar{g}\| = \|y\|).\end{aligned}$$

Therefore, $z = y$ and hence $h = f$. □

Definition 2.3.7. Let H be a Hilbert space. We say that a sequence $\{x_n\}$ in H is *weakly convergent* or *converges weakly* to x in H if

$$\langle x_n, y\rangle \to \langle x, y\rangle \quad \forall\, y \in H.$$

In this case, we write, $x_n \to x$ weakly or $x_n \xrightarrow{w} x$.

Remark 2.3.8. Is every weakly convergent sequence, convergent? The answer is not true in general. Consider the following example, where we show that a weakly convergent sequence may not be convergent.

Example 2.3.9. Let H be an infinite dimensional Hilbert space and $\{u_1, u_2, \dots\}$ be orthonormal basis of H. Then by Bessel's inequality, for each $y \in H$,

$$\sum_{n=1}^{\infty} |\langle y, u_n \rangle|^2 \leq \|y\|^2.$$

Therefore,

$$\langle y, u_n \rangle \to 0 \qquad \forall\, y \in H.$$

$$\therefore \langle u_n, y \rangle \to \langle 0, y \rangle \qquad \forall\, y \in H.$$

Thus, $u_n \to 0$ weakly. But for $m \neq n$,

$$\begin{aligned} \|u_m - u_n\|^2 &= \langle u_m - u_n, u_m - u_n \rangle \\ &= \|u_m\|^2 + \|u_n\|^2 \\ &= 1 + 1 = 2, \end{aligned}$$

i.e. $\|u_m - u_n\| = \sqrt{2}$. Therefore, $\{u_n\}$ is not Cauchy and hence it is not convergent.

Theorem 2.3.10. *Let H be a Hilbert space and $\{x_n\}$ be a sequence in H. Then $x_n \to x$ if and only if $x_n \to x$ weakly and $\|x_n\| \to \|x\|$.*

Proof. Suppose $x_n \to x$, i.e. $\|x_n - x\| \to 0$. Therefore, for each $y \in H$,

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

$$\therefore \langle x_n, y \rangle \to \langle x, y \rangle \text{ for each } y \in H.$$

Therefore $x_n \to x$ weakly. Also since $x_n \to x$, clearly $\|x_n\| \to \|x\|$.

Conversely, suppose that $x_n \to x$ weakly and $\|x_n\| \to \|x\|$. Then

$$\begin{aligned} \|x_n - x\|^2 &= \|x_n\|^2 + \|x\|^2 - 2\,\mathrm{Re}\langle x_n, x \rangle \\ &\to \|x\|^2 + \|x\|^2 - 2\|x\|^2 \qquad (\because \langle x_n, y \rangle \to \langle x, y \rangle) \\ &= 0. \end{aligned}$$

Therefore $x_n \to x$. □

Theorem 2.3.11. *Let H be a Hilbert space and $\{x_n\}$ be a bounded sequence in H. Then $\{x_n\}$ has a weakly convergent subsequence.*

Proof. Since $\{x_n\}$ is a bounded sequence in H, there exists $\alpha > 0$ such that $\|x_n\| \leq \alpha$ for all n. Then by Schwarz's inequality,

$$|\langle x_n, x_1\rangle| \leq \|x_n\|\|x_1\| \leq \alpha^2 \qquad \forall\ n.$$

Therefore $\{\langle x_n, x_1\rangle\}$ is a bounded sequence in K and hence by Bolzano-Weierstrass theorem for K, the sequence $\{\langle x_n, x_1\rangle\}$ has a convergent subsequence, say $\{\langle x_{n,1}, x_1\rangle\}$. Observe that the sequence $\{\langle x_{n,1}, x_2\rangle\}$ is bounded because

$$|\langle x_{n,1}, x_2\rangle| \leq \|x_{n,1}\|\|x_2\| \leq \alpha^2 \qquad \forall\ n.$$

Again, therefore, the bounded sequence $\{\langle x_{n,1}, x_2\rangle\}$ has a convergent subsequence $\{\langle x_{n,2}, x_2\rangle\}$ and so on. Thus, for each m we get a convergent subsequence $\{\langle x_{n,m}, x_m\rangle\}$ such that $\{\langle x_{n,m}, x_j\rangle\}$ converges for each $j = 1, 2, \ldots, m$.

Consider the convergent subsequence $\{\langle x_{n,n}, x_i\rangle\}$ for $i = 1, 2, \ldots$ ($\because$ for $n > m$ $\{\langle x_{n,n}, x_m\rangle\}$ is a subsequence of the convergent subsequence $\{\langle x_{n,m}, x_m\rangle\}$).

If $y \in \{x_1, x_2, \ldots\}$ then $\{\langle x_{n,n}, y\rangle\}$ converges. As a result, if $y \in L(\{x_1, x_2, \ldots\})$, then $\{\langle x_{n,n}, y\rangle\}$ converges in K. Let $Y = \overline{L(\{x_1, x_2, \ldots\})}$. If $y \in Y$, then there is a sequence $\{y_k\}$ in $L(\{x_1, x_2, \ldots\})$ such that $y_k \to y$. Fix $k_0 \in \mathbb{N}$ such that $\|y_k - y\| < \frac{\epsilon}{4\alpha}$. Fix $k > k_0$. Since $\{\langle x_{n,n}, y_k\rangle\}$ converges, it is Cauchy. Consequently, there is $n_0 \in \mathbb{N}$ such that $|\langle x_{n,n} - x_{m,m}, y_k\rangle| = |\langle x_{n,n}, y_k\rangle - \langle x_{n,n} - x_{m,m}, y_k\rangle| \leq \frac{\epsilon}{2}$. As a conclusion to all this, for all $n, m \geq n_0$,

$$\begin{aligned}
|\langle x_{n,n}, y\rangle - \langle x_{m,m}, y\rangle| &= |\langle x_{n,n} - x_{m,m}, y\rangle| \\
&\leq |\langle x_{n,n} - x_{m,m}, y - y_k\rangle| + |\langle x_{n,n} - x_{m,m}, y_k\rangle| \\
&\leq \|x_{n,n} - x_{m,m}\|\|y - y_k\| + |\langle x_{n,n} - x_{m,m}, y_k\rangle| \\
&\leq (\|x_{n,n}\| + \|x_{m,m}\|)\|y - y_k\| + |\langle x_{n,n} - x_{m,m}, y_k\rangle| \\
&\leq 2\alpha\|y_k - y\| + |\langle x_{n,n} - x_{m,m}, y_k\rangle| \\
&< 2\alpha\frac{\epsilon}{4\alpha} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}$$

Therefore, $\{\langle x_{n,n}, y\rangle\}$ is a Cauchy sequence in K and hence, it converges in K for all $y \in Y$. Since Y is a closed subspace of H, by the projection theorem, we have

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

Therefore every $x \in H$ can be written as $x = y + z$ with $y \in Y$ and $z \in Y^{\perp}$. Then

$$\begin{aligned}
\langle x_{n,n}, x\rangle &= \langle x_{n,n}, y\rangle + \langle x_{n,n}, z\rangle \\
&= \langle x_{n,n}, y\rangle \qquad (\because x_{n,n} \in Y,\ z \in Y^{\perp} \Rightarrow \langle x_{n,n}, z\rangle = 0).
\end{aligned}$$

Therefore $\{\langle x_{n,n}, x\rangle\}$ converges in K for each $x \in H$. Now, for $x \in H$, take

$$f(x) = \lim_{n\to\infty} \langle x, x_{n,n}\rangle.$$

Then $f : H \to K$ is a linear functional on H. Also, since

$$|f(x)| = \lim_{n\to\infty} |\langle x, x_{n,n}\rangle| \leq \alpha\|x\|, \qquad x \in H$$

we see that f is a bounded (continuous) linear functional on H. By Riesz representation theorem there exists $y \in H$ such that $f(x) = \langle x, y\rangle$, for all $x \in H$. Therefore

$$\lim_{n\to\infty} \langle x, x_{n,n}\rangle = f(x) = \langle x, y\rangle, \qquad \forall\ x \in H,$$

i.e. $\langle x_{n,n}, x\rangle \to \langle y, x\rangle\ \forall\ x \in H$.
Thus, $x_{n,n} \to y$ weakly, where $\{x_{n,n}\}$ is a subsequence of $\{x_n\}$. □

Definition 2.3.12. Let H be a Hilbert space and $E \subset H$. We say that E is *weakly bounded* if for each $y \in H$ such that $\alpha_y \geq 0$ such that

$$|\langle x, y\rangle| \leq \alpha_y \quad \forall\ x \in E.$$

Remark 2.3.13. Let $E \subset H$ and $f \in H'$. Then there is $y \in H$, let $f = f_y$, where f_y denotes the bounded linear functional on H defined by $f_y(x) = \langle x, y\rangle$, $(x \in H)$. Clearly, $f(E) = \{\langle x, y\rangle : x \in E\}$. Consequently, E is weakly bounded if and only if $f(E)$ is bounded for all $f \in H'$.

Lemma 2.3.14. *Let H be a Hilbert space and Y be a finite dimensional subspace of H. Let P_Y denote the orthogonal projection of H on Y. If E is weakly bounded subset of H then the set $\{P_Y(x) : x \in E\}$ is bounded.*

Proof. Let $B = \{y_1, y_2, \dots, y_n\}$ be an orthonormal basis of Y. Define $P_Y : H \to Y$ by

$$P_Y(x) = \langle x, y_1\rangle y_1 + \langle x, y_2\rangle y_2 + \cdots + \langle x, y_n\rangle y_n, \quad (x \in H). \tag{2.5}$$

Clearly,

$$\begin{aligned} {P_Y}^2 = P_Y(P_Y(x)) &= P_Y(\langle x, y_1\rangle y_1 + \langle x, y_2\rangle y_2 + \cdots + \langle x, y_n\rangle y_n) \\ &= \langle x, y_1\rangle P_Y(y_1) + \langle x, y_2\rangle P_Y(y_2) + \cdots + \langle x, y_n\rangle P_Y(y_n) \\ &= \langle x, y_1\rangle y_1 + \langle x, y_2\rangle y_2 + \cdots + \langle x, y_n\rangle y_n. \qquad (\because \text{by } (2.5), P_Y(y_i) = y_i) \\ &= P_Y(x). \end{aligned}$$

Therefore $P_Y^2 = P_Y$, i.e. P_Y is idempotent (projection). Also, the range of P_Y is $R(P_Y) = Y$. Also,

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

Then, since P is a projection, $R(I - P) = \ker P = Y^{\perp}$.

Since E is weakly bounded, there exist $\alpha_{y_1}, \alpha_{y_2}, \dots, \alpha_{y_n} \geq 0$ such that

$$|\langle x, y_j\rangle| \leq \alpha_{y_j} \qquad \forall\ x \in E,\ \forall\ j = 1, 2, \dots, n.$$

Therefore for $x \in E$,

$$\begin{aligned} \|P_Y x\|^2 &= |\langle x, y_1\rangle|^2 + |\langle x, y_2\rangle|^2 + \cdots + |\langle x, y_n\rangle|^2 \qquad (\text{by Pythagoras theorem applied to } (2.5)) \\ &= \alpha_{y_1}^2 + \alpha_{y_2}^2 + \cdots + \alpha_{y_n}^2. \end{aligned}$$

Hence for all $x \in E$, $\|P_Y x\| \leq N$, where $N = \sqrt{\alpha_{y_1}^2 + \alpha_{y_2}^2 + \cdots + \alpha_{y_n}^2}$. Thus, the set $\{P_Y x : x \in E\}$ is bounded. □

Theorem 2.3.15. *Let H be a Hilbert space and E be a subset of H then E is bounded if and only if E is weakly bounded.*

Proof. Suppose E is bounded. Then there exists an $M \geq 0$ such that $\|x\| \leq M$ for all $x \in E$. Then for each $y \in H$,

$$|\langle x, y\rangle| \leq \|x\|\|y\| \leq M\|y\|, \quad x \in E.$$

Therefore E is weakly bounded.

Conversely, assume that E is weakly bounded, i.e. for each $y \in H$ there is $\alpha_y \geq 0$ such that $|\langle x, y\rangle| \leq \alpha_y$ for all $x \in E$.

Suppose, if possible, E is unbounded. Then there exists $x_1 \in E$ such that $\|x_1\| \geq 1$. Let $z_1 = x_1$ and $Y_1 = L(\{z_1\})$. Take $P_1 = P_{Y_1}$. Since, $\dim Y_1 < \infty$, i.e. Y_1 is finite dimensional, the set $\{P_1(x) : x \in E\}$ is bounded. Hence the set $\{x - P_1(x) : x \in E\}$ is unbounded (otherwise E is bounded).

Since $2\left(2 + \frac{\alpha_{z_1}}{\|\alpha_{z_1}\|}\right) > 0$, there exists $x_2 \in E$ such that $\|x_2 - P_1(x_2)\| > 2\left(2 + \frac{\alpha_{z_1}}{\|\alpha_{z_1}\|}\right)$. Let $z_2 = x_2 - P_1(x_2)$. Then $\|z_2\| > 2\left(2 + \frac{\alpha_{z_1}}{\|\alpha_{z_1}\|}\right)$ and $z_2 \perp Y_1$. Therefore,

$$z_1 \perp z_2.$$

Let $Y_2 = L(\{z_1, x_2, z_2\})$. So $\dim Y_2 < \infty$. Let $P_2 = P_{Y_2}$. Then the set $\{P_2(x) : x \in E\}$ is bounded. Therefore the set $\{x - P_2(x) : x \in E\}$ is unbounded. Since $3\left(3 + \frac{\alpha_{z_1}}{\|\alpha_{z_1}\|} + \frac{\alpha_{z_2}}{\|\alpha_{z_2}\|}\right) > 0$, there exists $x_3 \in E$ such that

$$\|x_3 - P_2(x_3)\| > 3\left(3 + \frac{\alpha_{z_1}}{\|\alpha_{z_1}\|} + \frac{\alpha_{z_2}}{\|\alpha_{z_2}\|}\right).$$

Take $z_3 = x_3 - P_2(x_3)$. Thus, $\|z_3\| > 3\left(3 + \frac{\alpha_{z_1}}{\|\alpha_{z_1}\|} + \frac{\alpha_{z_2}}{\|\alpha_{z_2}\|}\right)$ and $z_3 \perp Y_2$. So, z_1, z_2, z_3 are orthogonal. Continuing this way, suppose that $z_1, x_2, z_2, \ldots, x_m, z_m$ are chosen such that $z_1, z_2, \ldots, z_m$ are orthogonal. Take

$$Y_m = L(\{z_1, x_2, z_2, \ldots, x_m, z_m\}.$$

Then $\dim Y_m < \infty$. Take $P_m = P_{Y_m}$. Then the set $\{P_m(x) : x \in E\}$ is bounded and hence the set $\{x - P_m(x) : x \in E\}$ is not bounded. Since $(m+1)\left(m + 1 + \sum_{j=1}^{m} \frac{\alpha_{z_j}}{\|\alpha_{z_j}\}}\right) > 0$, there exists $x_{m+1} \in E$ such that

$$\|x_{m+1} - P_m(x_{m+1})\| > (m+1)\left(m + 1 + \sum_{j=1}^{m} \frac{\alpha_{z_j}}{\|\alpha_{z_j}\|}\right).$$

Take $z_{m+1} = x_{m+1} - P_m(x_{m+1})$. Then

$$\|z_{m+1}\| > (m+1)\left(m + 1 + \sum_{j=1}^{m} \frac{\alpha_{z_j}}{\|\alpha_{z_j}\|}\right) \tag{2.6}$$

and $z_{m+1} \perp Y_m$. Therefore $z_1, z_2, \ldots, z_m, z_{m+1}$ are orthogonal such that none of them is 0. Then

$$\|z_{m+1}\|^2 = \langle z_{m+1}, z_{m+1}\rangle$$

$$
\begin{aligned}
&= \langle x_{m+1}, z_{m+1}\rangle - \langle P_m(x_{m+1}), z_{m+1}\rangle \\
&= \langle x_{m+1}, z_{m+1}\rangle \qquad (\because P_m(x_{m+1}) \in Y_m,\ z_{m+1} \perp Y_m). \qquad (2.7)
\end{aligned}
$$

Now,

$$\langle x_{m+1}, z_n\rangle = 0 \qquad \forall\ n \geq m+2 \qquad (\because z_n \perp Y_{m+1},\ x_{m+1} \in Y_{m+1}). \qquad (2.8)$$

Take $u_j = \frac{z_j}{\|z_j\|}$, $\quad j = 1, 2, \ldots$. Therefore $u_1, u_2, \ldots$ are orthonormal. Since, $\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$, by Riesz-Fischer theorem,

$$\sum_{n=1}^{\infty} \frac{1}{n} u_n \text{ converges in } H.$$

Suppose $y = \sum_{n=1}^{\infty} \frac{1}{n} u_n$. Then

$$
\begin{aligned}
|\langle x_{m+1}, y\rangle| &= \left|\left\langle x_{m+1}, \sum_{n=1}^{\infty} \frac{1}{n} u_n \right\rangle\right| \\
&= \left|\left\langle x_{m+1}, \sum_{n=1}^{m+1} \frac{1}{n} u_n \right\rangle\right| && \text{(by (2.8))} \\
&\geq \left|\left\langle x_{m+1}, \frac{u_{m+1}}{m+1} \right\rangle\right| - \left|\left\langle x_{m+1}, \sum_{n=1}^{m} \frac{1}{n} u_n \right\rangle\right| && (\because |\alpha+\beta| \geq |\alpha| - |\beta|) \\
&\geq \left|\left\langle x_{m+1}, \frac{z_{m+1}}{(m+1)\|z_{m+1}\|} \right\rangle\right| - \sum_{n=1}^{m} \frac{\alpha_{z_n}}{\|z_n\|} \\
&= \frac{\|z_{m+1}\|^2}{(m+1)\|z_{m+1}\|} - \sum_{n=1}^{\infty} \frac{\alpha_{z_n}}{\|z_n\|} && \text{(by (2.7))}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
|\langle x_{m+1}, y\rangle| &\geq \frac{\|z_{m+1}\|}{m+1} - \sum_{n=1}^{m} \frac{\alpha_{z_n}}{\|z_n\|} \\
&> (m+1) + \sum_{n=1}^{m} \frac{\alpha_{z_n}}{\|z_n\|} - \sum_{m=1}^{m} \frac{\alpha_{z_n}}{\|z_n\|} && \text{(by (2.6))} \\
&= m+1
\end{aligned}
$$

which is not possible since $|\langle x, y\rangle| \leq \alpha_y$, $\forall\ x \in E$. Therefore E must be bounded. □

Seminar Topics 5. In these exercises, H will denote a Hilbert space.

1. Let E, F be subsets of H. If E, F are bounded, then show that $E + F, E \cup F$ are also bounded.
2. Let Y be a subspace of H. Show that Y is bounded if and only if $Y = \{0\}$.
3. $E \subset H$ and P be an orthogonal projection on H. Show that E is bounded if and only if $P(E)$ as well as $(I - P)(E)$ is bounded.
4. Let $E \subset H$ be convex. Show that $\overline{E}$ is also convex.
5. Show that a subspace of H is convex.

CHAPTER 3

Bounded Operators on Hilbert spaces

3.1 Adjoints of Bounded Operators

In this section we discuss the adjoint of a bounded linear operator (i.e. bounded linear map). Before we define it formally, consider the following example:

Example 3.1.1. Let $H = \ell^2$ be the Hilbert space of square summable sequences. Let $S : H \to H$ and $T : H \to H$ be the left-shift and the right-shift operators respectively, i.e. for $x = (x(1), x(2), \ldots)$, $y = (y(1), y(2), \ldots) \in \ell^2$,

$$S(y) = (y(2), y(3), \ldots) \qquad \text{and} \qquad T(x) = (0, x(1), x(2), \ldots).$$

Clearly, S and T are linear maps. It is also easy to see that S and T are bounded, as

$$\|Sy\| = \left(\sum_{i=1}^{\infty} |(Sy)(i)|^2\right)^{\frac{1}{2}} = \left(\sum_{i=2}^{\infty} |y(i)|^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{\infty} |y(i)|^2\right)^{\frac{1}{2}} = \|y\|$$

and

$$\|Tx\| = \left(\sum_{i=1}^{\infty} |(Tx)(i)|^2\right)^{\frac{1}{2}} = \left(\sum_{i=1}^{\infty} |x(i)|^2\right)^{\frac{1}{2}} = \|x\| \qquad (\because (Tx)(1) = 0).$$

Thus, $S, T \in BL(H)$. Also, observe that

$$\begin{aligned}
\langle Tx, y\rangle &= \langle (0, x(1), x(2), \ldots), (y(1), y(2), y(3), \ldots)\rangle \\
&= x(1)\overline{y(2)} + x(2)\overline{y(3)} + \cdots \\
&= \langle (x(1), x(2), \ldots), (y(2), y(3), \ldots)\rangle \\
&= \langle x, Sy\rangle.
\end{aligned}$$

Thus, $S, T \in BL(H)$ are opposite of each other in the sense that for every $x, y \in H = \ell^2$,

$$\langle Tx, y\rangle = \langle x, Sy\rangle.$$

Remark 3.1.2. Now, we may have two questions here. First question is: given a bounded linear map T on a Hilbert space H, i.e. $T \in BL(H)$, does there always exists another bounded linear operator $S \in BL(H)$ such that $\langle Tx, y\rangle = \langle x, Sy\rangle$ for all $x, y \in H$? Secondly, if such operator (map) S exists, then is it unique?

The answer to both the questions posed above is affirmative in case of Hilbert space but not true for every inner product space (which are not complete). First we prove the following theorem which affirms the existence of unique operator S such that $\langle Tx, y\rangle = \langle x, Sy\rangle$ for all $x, y \in H$. After proving the theorem, we give a counter example which shows that it need not be true in an inner product space which is not complete.

Theorem 3.1.3. *Let H be a Hilbert space and $T \in BL(H)$. There there is a unique $S \in BL(H)$ such that $\langle Tx, y\rangle = \langle x, Sy\rangle$ for every $x, y \in H$ and $\|S\| \leq \|T\|$.*

Proof. For $y \in H$, define $f_y : H \to K$ by

$$f_y(x) = \langle Tx, y\rangle \qquad \text{for all } x \in H.$$

Then f_y is a linear functional on H (Verify!). Also,

$$\begin{aligned} |f_y(x)| &= |\langle Tx, y\rangle| \\ &\leq \|Tx\|\|y\| && \text{(Schwarz inequality)} \\ &\leq \|T\|\|x\|\|y\| && (\because T \text{ is bounded}) \\ &= (\|T\|\|y\|)\|x\| \end{aligned}$$

Therefore, f_y is bounded and $\|f_y\| \leq \|T\|\|y\|$. Then by the Riesz-representation theorem there is a unique $z \in H$ such that

$$f_y(x) = \langle x, z\rangle \qquad x \in H$$

and

$$\|f_y\| = \|z\|.$$

Define $S : H \to H$ by $Sy = z$, $(y \in H)$. Then

$$\langle Tx, y\rangle = f_y(x) = \langle x, z\rangle = \langle x, Sy\rangle, \qquad x \in H.$$

Then $S : H \to H$ is linear for if $y_1, y_2 \in H$ then for all $x \in H$,

$$\begin{aligned} \langle x, S(y_1 + y_2)\rangle &= \langle Tx, y_1 + y_2\rangle \\ &= \langle Tx, y_1\rangle + \langle Tx, y_2\rangle \\ &= f_{y_1}(x) + f_{y_2}(x) \\ &= \langle x, Sy_1\rangle + \langle x, Sy_2\rangle \\ &= \langle x, S(y_1 + y_2)\rangle. \end{aligned}$$

Therefore

$$S(y_1 + y_2) = Sy_1 + Sy_2 \quad \forall\, y_1, y_2 \in H.$$

Similarly, (Check!)

$$S(\alpha y) = \alpha Sy \quad \forall\, y \in H,\ \alpha \in K.$$

Thus, S is a linear map. Now, for $y \in H$

$$\|Sy\| = \|z\| = \|f_y\| \leq \|T\|\|y\|.$$

Therefore S is bounded and taking supremum over all y with $\|y\| \leq 1$, we have

$$\|S\| \leq \|T\|.$$

Now to show the uniqueness of S, suppose $S' \in BL(H)$ such that for all $x, y \in H$,

$$\langle x, Sy\rangle = \langle Tx, y\rangle = \langle x, S'y\rangle.$$

Then, $\langle x, (S - S')(y)\rangle = 0$ for all $x, y \in H$ and hence $S = S'$. □

Definition 3.1.4. Let H be a Hilbert space and $T \in BL(H)$. The (unique) operator $S \in BL(H)$ such that $\langle Tx, y\rangle = \langle x, Sy\rangle$ for all $x, y \in H$ is known as the *adjoint* of T and it is denoted by T^*. Thus,

$$\langle Tx, y\rangle = \langle x, T^*y\rangle \text{ for all } x, y \in H.$$

In Example 3.1.1, we saw that the right-shift operator on ℓ^2 is the adjoint of the left-shift operator on ℓ^2. Let us give one more example of adjoint of a bounded linear operator.

Example 3.1.5. Consider the Hilbert space $H = \mathbb{C}^2$. Let $T \in BL(H)$ by defined as $T(x, y) = (x + iy, iy)$ for $(x, y) \in \mathbb{C}^2$. Then one can see that its adjoint $T^* \in BL(H)$ is given by $T^*(x, y) = (x, -ix - iy)$ for all $(x, y) \in H = \mathbb{C}^2$ as for $(z_1, w_1), (z_2, w_2) \in \mathbb{C}^2$,

$$\begin{aligned}
\langle T(z_1, w_1), (z_2, w_2)\rangle &= \langle (z_1 + iw_1, iw_1), (z_2, w_2)\rangle \\
&= (z_1 + iw_1)\bar{z}_2 + iw_1\bar{w}_2 \\
&= z_1\bar{z}_2 + iw_1(\bar{z}_2 + \bar{w}_2) \\
&= \langle (z_1, w_1), (z_2, -iz_2 - iw_2)\rangle \\
&= \langle (z_1, w_1), T^*(z_2, w_2)\rangle.
\end{aligned}$$

Thus, for all $x, y \in \mathbb{C}^2$ we have

$$\langle Tx, y\rangle = \langle x, Ty\rangle.$$

Exercise 3.1.6. Show by an example that the completion of the space is necessary for the existence of the adjoint of a bounded operator.

Solution. Let $X = c_{00}$. We have already seen how inner product is defined on c_{00} and that it is not a complete space. Define $T : X \to X$ by

$$Tx = \left(\sum_{n=1}^{\infty} \frac{x(n)}{n}, 0, 0, \dots\right) \qquad \text{for } x = (x(1), x(2), \dots) \in X. \tag{3.1}$$

Then $T : X \to X$ is a linear map (Verify!). Now, for all $x \in H$

$$\begin{aligned}
\|Tx\| &= \left|\sum_{n=1}^{\infty} \frac{x(n)}{n}\right| \\
&\leq \sum_{n=1}^{\infty} \frac{1}{n}|x(n)|
\end{aligned}$$

$$\leq \left(\sum_{n=1}^{\infty} \frac{1}{n^2}\right)^{\frac{1}{2}} \left(\sum_{n=1}^{\infty} |x(n)|^2\right)^{\frac{1}{2}} \qquad \text{(Holder's inequality)}$$
$$= \frac{\pi^2}{6}\|x\|.$$

Therefore $T : X \to X$ is a bounded linear operator, i.e. $T \in BL(X)$.

Suppose there exists $S \in BL(X)$ such that $\langle Tx, y\rangle = \langle x, Sy\rangle$ for all $x, y \in X = c_{00}$. Take $u_n = (0, 0, \ldots, 0, \underbrace{1}_{n^{\text{th}}\text{place}}, 0, \ldots)$. Then

$$\begin{aligned} \overline{(Su_1)(n)} &= \langle u_n, Su_1\rangle \\ &= \langle Tu_n, u_1\rangle \qquad \text{(by (3.1))} \\ &= \frac{1}{n} \end{aligned}$$

Therefore $(Su_1)(n) = \frac{1}{n}$ for $n = 1, 2, \ldots$. Hence, $Su_1 \notin X = c_{00}$ which is a contradiction as $S \in BL(X)$. ☐

Proposition 3.1.7. *Let H be a Hilbert space and $S, T \in BL(H)$, $\alpha \in K$. Then*

1. $(S + T)^* = S^* + T^*$
2. $(\alpha S)^* = \bar{\alpha} S^*$
3. $(ST)^* = T^* S^*$
4. $(S^*)^* = S$

Proof.

1. For $y \in H$,

$$\begin{aligned} \langle x, (S+T)^* y\rangle &= \langle (S+T)x, y\rangle \\ &= \langle Sx, y\rangle + \langle Tx, y\rangle \\ &= \langle x, S^* y\rangle + \langle x, T^* y\rangle \\ &= \langle x, (S^* + T^*)y\rangle \qquad x \in H. \end{aligned}$$

Therefore,

$$(S+T)^* = S^* + T^*.$$

2. For $y \in H$,

$$\begin{aligned} \langle x, (\alpha S)^* y\rangle &= \langle (\alpha S)x, y\rangle \\ &= \alpha\langle Sx, y\rangle \\ &= \alpha\langle x, S^* y\rangle \\ &= \langle x, \bar{\alpha} S^* y\rangle \qquad x \in H. \end{aligned}$$

Therefore,

$$(\alpha S)^* = \bar{\alpha} S^*.$$

3. For $y \in H$,

$$\begin{aligned} \langle x, (ST)^* y\rangle &= \langle (ST)x, y\rangle \\ &= \langle Tx, S^* y\rangle \end{aligned}$$

$$= \langle x, T^*S^*y \rangle \qquad x \in H.$$

Therefore,

$$(ST)^* = T^*S^*.$$

4. For $y \in H$,

$$\begin{aligned}\langle x, (S^*)^*y \rangle &= \langle S^*x, y \rangle \\ &= \overline{\langle y, S^*x \rangle} \\ &= \overline{\langle Sy, x \rangle} \\ &= \langle x, Sy \rangle \qquad x \in H.\end{aligned}$$

Therefore,

$$(S^*)^* = S.$$

□

Corollary 3.1.8. *Let H be a Hilbert space and $S \in BL(H)$ be invertible in $BL(H)$. Then S^* is invertible in $BL(H)$ and $(S^*)^{-1} = (S^{-1})^*$.*

Proof. Since S is invertible in $BL(H)$, there exists $S^{-1} \in BL(H)$ such that

$$SS^{-1} = S^{-1}S = I.$$

Taking adjoint, we get

$$(SS^{-1})^* = (S^{-1}S)^* = I^* = I.$$

Therefore,

$$(S^{-1})^*S^* = S^*(S^{-1})^* = I.$$

Hence, S^* is invertible in $BL(H)$ and $(S^*)^{-1} = (S^{-1})^*$. □

Proposition 3.1.9. *Let H be a Hilbert space and $T \in BL(H)$. Then $\|T^*\| = \|T\|$ and $\|T^*T\| = \|T\|^2$.*

Proof. We have seen (in Proposition 2.2.15) that in a Hilbert space H, norm of $T \in BL(H)$ is defined by

$$\|T\| = \sup\{|\langle Tx, y \rangle| : x, y \in X,\ \|x\| \le 1,\ \|y\| \le 1\}.$$

Now, for $x, y \in H$,

$$|\langle Tx, y \rangle| = |\langle y, Tx \rangle| = |\langle T^*y, x \rangle|.$$

By taking supremum over $x, y \in H$ with $\|x\| \le 1$, $\|y\| \le 1$, we get

$$\|T\| = \|T^*\|.$$

We also know (by Remark 2.2.13) that $\|T^*T\| \le \|T^*\|\|T\| = \|T\|^2$.

Now for $x \in H$,

$$\|Tx\|^2 = \langle Tx, Tx \rangle$$

$$= \langle x, T^*Tx \rangle$$
$$\leq \|x\| \|T^*Tx\|.$$

Taking supremum over $x \in H$ with $\|x\| \leq 1$, we get $\|T\|^2 \leq \|T^*T\|$. Therefore

$$\|T\|^2 = \|T^*T\|.$$

□

Remark 3.1.10. Suppose H is a separable Hilbert space with orthonormal basis $u_1, u_2, \ldots$ and $T \in BL(H)$. If the matrix of T with respect to this orthonormal basis is

$$m(T) = (\alpha_{ij}).$$

Then

$$\alpha_{ij} = \langle Tu_j, u_i \rangle.$$

Now if $m(T^*) = (\beta_{ij})$ with respect to orthonormal basis, then $\beta_{ij} = \langle T^* u_j, u_i \rangle$. Then

$$\begin{aligned} \beta_{ij} &= \langle T^* u_j, u_i \rangle \\ &= \langle u_j, Tu_i \rangle \\ &= \overline{\langle Tu_i, u_j \rangle} = \bar{\alpha}_{ij}. \end{aligned}$$

Thus, $m(T^*)$ is the complex conjugate of the transpose of the matrix $m(T)$.

Remark 3.1.11. Note that this is the case in Example 3.1.5. The bounded linear operator T is defined by the matrix

$$m(T) = \begin{bmatrix} 1 & i \\ 0 & i \end{bmatrix}$$

and its adjoint $T^* \in BL(\mathbb{C}^2)$ as defined in Example 3.1.5 is given by the adjoint (conjugate transpose) of the matrix of T as follows:

$$m(T^*) = \begin{bmatrix} 1 & 0 \\ -i & -i \end{bmatrix}.$$

Thus, for all $(x, y) \in \mathbb{C}^2$,

$$T(x, y) = (x + iy, iy) \qquad \text{and} \qquad T^*(x, y) = (x, -ix - iy)$$

are adjoints of each other as bounded linear operators.

Example 3.1.12. Let $H = \ell^2$. As seen in Example 3.1.1, let T be the right-shift operator on ℓ^2, i.e. define for $x = (x(1), x(2), \ldots)$, $T : \ell^2 \to \ell^2$ by

$$T(x(1), x(2), \ldots) = (0, x(1), x(2), \ldots).$$

Then $\|Tx\|^2 = \|x\|^2$ for all $x \in \ell^2$, i.e. T is isometry. Therefore T is bounded and $\|T\| = 1$. Consider the orthonormal basis $\{e_n : n \in \mathbb{N}\}$ of ℓ^2, where $e_n = (0, 0, \ldots, 0, \underbrace{1}_{n^{\text{th}} \text{ place}}, 0, \ldots)$.

So, by definition, $T(e_n) = e_{n+1}$ for all $n = 1, 2, \ldots$. If

$$m(T) = (\alpha_{ij})$$

is the matrix of T with respect to this orthonormal basis then

$$\alpha_{ij} = \langle Te_j, e_i \rangle = \langle e_{j+1}, e_i \rangle = \delta_{(j+1)(i)}.$$

Therefore,

$$m(T) = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \qquad \text{and} \qquad m(T^*) = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Then $T^*(x(1), x(2), x(3), \ldots) = (x(2), x(3), \ldots)$.

Note: The operator T on ℓ^2 defined in the above example is known as unilateral right-shift and T^* is known as unilateral left-shift.

Theorem 3.1.13. *Let H be a Hilbert space and $T \in BL(H)$. Then*
(a) $\ker(T) = R(T^*)^\perp$ *and* $\ker(T^*) = R(T)^\perp$.
(b) $\ker(T)^\perp = \overline{R(T^*)}$ *and* $\ker(T^*)^\perp = \overline{R(T)}$.

Proof. (a) First we show that $\ker(T) = R(T^*)^\perp$.

$$\begin{aligned} x \in \ker(T) &\Leftrightarrow Tx = 0 \\ &\Leftrightarrow \langle Tx, y \rangle = 0 \qquad \forall\, y \in H \\ &\Leftrightarrow \langle x, T^*y \rangle = 0 \qquad \forall\, y \in H \\ &\Leftrightarrow x \in R(T^*)^\perp. \end{aligned}$$

By replacing T by T^* and using $(T^*)^* = T$, we get $\ker(T^*) = R(T)^\perp$.

(b) Taking $\perp$ (complement) on both sides of (a) and also using the result that if Y is a subspace (not necessarily closed) of H then $Y^{\perp\perp} = \overline{Y}$, we get

$$(\ker T)^\perp = R(T^*)^{\perp\perp} = \overline{R(T)}.$$

By replacing T by T^* and using $(T^*)^* = T$, we get

$$\ker(T^*)^\perp = \overline{R(T)} \qquad (\because T^{**} = T).$$

□

Corollary 3.1.14. *Let H be Hilbert space and $T \in BL(H)$. Then*
(a) T is injective i.e. T is one-one if and only if $R(T^*)$ is dense in H.
(b) T^* is one-one if and only if $R(T)$ is dense in H.

Proof. We know that T is one-one if and only if $\ker T = \{0\}$ if and only if

$$H = \{0\}^\perp = (\ker T)^\perp = \overline{R(T^*)},$$

i.e. if and only if $R(T^*)$ is dense in H.

For (b) part, replace T by T^*. □

Definition 3.1.15. Let H be a Hilbert space and $T \in BL(H)$. T is called *bounded below* if there exists $\beta > 0$ such that $\|Tx\| \geq \beta\|x\|$ for all $x \in H$.

Remarks 3.1.16.

1. If $T \in BL(H)$ is isometry then T is bounded below.
   If T is isometry then $\|Tx\| = \|x\|$, so in this case taking $\beta = 1$, we conclude that T is bounded below
2. If $T \in BL(H)$ is bounded below then T is one-one.
   Let $x \in H$ such that $Tx = 0$ then because T is bounded below, there exits $\beta > 0$ such that
   $$0 = \|Tx\| \geq \beta\|x\| \Rightarrow \|x\| \leq 0 \Rightarrow x = 0.$$
   Therefore, T is one-one.

Proposition 3.1.17. *Let H be a Hilbert space and $T \in BL(H)$ be bounded below. Then $R(T)$ is closed in H.*

Proof. Let $y \in \overline{R(T)}$. Then there is a sequence $\{x_n\}$ in H such that $Tx_n \to y$. Therefore, $\{Tx_n\}$ is Cauchy. Since T is bounded below, there exists $\beta > 0$ such that

$$\|Tx\| \geq \beta\|x\| \qquad \forall\ x \in H.$$

Now, for $m, n \in \mathbb{N}$,

$$\beta\|x_n - x_m\| \leq \|Tx_n - Tx_m\|.$$

Since $\{Tx_n\}$ is Cauchy, we get that $\{x_n\}$ is a Cauchy sequence in H. Since H is complete, $x_n \to x$ in H. As T is continuous, $Tx_n \to Tx$. But we have $Tx_n \to y$ and hence by uniqueness of limit, we have $y = Tx \in R(T)$. Therefore, $R(T)$ is closed. □

Theorem 3.1.18. *Let H be a Hilbert space and $T \in BL(H)$. Then $R(T) = H$ (i.e. T is onto) if and only if T^* is bounded below. Hence, $R(T^*) = H$ if and only if T is bounded below.*

Proof. Suppose $R(T) = H$ (i.e. T is onto) then we have to show that T^* is bounded below. Suppose T^* is not bounded below. Then for each n, there exists $x_n \in H$ such that

$$\frac{1}{n}\|x_n\| > \|T^*x_n\|. \tag{3.2}$$

Take $y_n = n\frac{x_n}{\|x_n\|}$. Then $\|y_n\| = n$. Now,

$$\begin{aligned} \|T^*y_n\| &= \frac{n}{\|x_n\|}\|T^*x_n\| \\ &< \frac{n}{\|x_n\|}\frac{\|x_n\|}{n} \qquad \text{(by (3.2))} \\ &< 1 \qquad \forall\ n. \end{aligned}$$

Let $y \in H = R(T)$. Then there exists $x \in H$ such that $y = Tx$. Now,

$$|\langle y_n, y\rangle| = |\langle y_n, Tx\rangle|$$

$$
\begin{aligned}
&= |\langle T^* y_n, x\rangle| \\
&\leq \|T^* y_n\| \|x\| \\
&< \|x\| \qquad (\because \|T^* y_n\| < 1).
\end{aligned}
$$

Therefore the set $\{y_n : n \in \mathbb{N}\}$ is weakly bounded (taking $\alpha_y = \|x\|$) and hence it is bounded as we know (by Theorem 2.3.15) that a set E is weakly bounded if and only if it is bounded. But $\|y_n\| = n$ which is a contradiction. Hence, T^* must be bounded below.

Conversely, assume that T^* is bounded below. Then there exists $\beta > 0$ such that

$$\|T^* x\| \geq \beta \|x\| \qquad \forall\, x \in H.$$

Then by the last proposition, $R(T^*)$ is a closed subspace of a Hilbert space. Hence, $R(T^*)$ is a Hilbert space.

Now, since T^* is bounded below (by Remark 3.1.16), T^* is one-one. Hence, for each $z \in R(T^*)$ there is a unique $w \in H$ such that $T^* w = z$. Let $y \in H$. Define $g : R(T^*) \to K$ by

$$g(z) = g(T^* w) = \langle w, y\rangle, \qquad w \in H. \tag{3.3}$$

Then, clearly g is well-defined linear functional on $R(T^*)$. Now, for all $z \in R(T^*)$

$$
\begin{aligned}
|g(z)| &= \| \langle w, y\rangle| \\
&= \|w\| \|y\| \\
&\leq \frac{1}{\beta} \|T^* w\| \|y\| \qquad \text{(since } T^* \text{ is bounded below)} \\
&= \frac{1}{\beta} \|z\| \|y\| = \left(\frac{1}{\beta} \|y\| \right) \|z\|.
\end{aligned}
$$

Thus, $g : R(T^*) \to K$ is a bounded-linear functional on the Hilbert space $R(T^*)$. So, by Riesz-representation theorem, there exists $x \in R(T^*)$ such that $g(z) = \langle z, x\rangle,\ z \in R(T^*)$. Now, for all $w \in H$, we have

$$g(T^* w) = \langle T^* w, x\rangle = \langle w, Tx\rangle.$$

But by (3.3), we have $g(T^* w) = \langle w, y\rangle,\ \forall\, w \in H$. Thus,

$$\langle w, Tx\rangle = \langle w, y\rangle, \qquad \forall\, w \in H.$$

Therefore $y = Tx$ and hence $R(T) = H$. □

Remark 3.1.19. Summing up the above remarks and results, we observed here that

1. T is bounded below $\Rightarrow$ T is one-one (by Remark 3.1.16).
2. T is bounded below $\Rightarrow$ $R(T)$ is closed subspace of H (by above Proposition).
3. T^* is bounded below $\Leftrightarrow$ $R(T) = H$, i.e. T is onto (by above Theorem).
   Also, T is bounded below $\Leftrightarrow$ T^* is onto, i.e. $R(T^*) = H$.

3.2 Normal, Unitary and Self-adjoint operators

Definition 3.2.1. Let H be a Hilbert space and $T \in BL(H)$. Then

1. T is said to be a normal operator if $T^*T = TT^*$.
2. T is said to be unitary if $T^*T = I = TT^*$.
3. T is said to be self-adjoint if $T^* = T$.

Remarks 3.2.2. 1. T is normal if and only if

$$\langle T^*x, T^*y\rangle = \langle Tx, Ty\rangle, \qquad \forall\ x, y \in H.$$

2. T is unitary if and only if

$$\langle T^*x, T^*y\rangle = \langle x, y\rangle = \langle Tx, Ty\rangle, \qquad \forall\ x, y \in H.$$

3. T is self-adjoint if and only if

$$\langle Tx, y\rangle = \langle x, Ty\rangle, \qquad \forall\ x, y \in H.$$

Note that remark (2) above implies that unitary operator preserves the inner product, i.e. it preserves the geometric structure. It is clear that every unitary operator is normal and very self-adjoint operator is also normal.

Now, we give an example of a normal operator. We show that the diagonal operator is normal.

Example 3.2.3. Let H be a separable Hilbert space. Then (by Theorem 1.3.10) H has a countable orthonormal basis. Let $u_1, u_2, \ldots$ be orthonormal basis for H. Let $\{\alpha_n\}$ be a bounded sequence in K. Define $T : H \to H$ by

$$Tx = \sum_{n=1}^{\infty} \alpha_n \langle x, u_n\rangle u_n, \qquad x \in H.$$

Then for $j = 1, 2, \ldots,$

$$\begin{aligned} Tu_j &= \alpha_j u_j \\ &= 0u_1 + 0u_2 + \cdots + 0u_{j-1} + \alpha_j u_j + 0u_{j+1} + \cdots . \end{aligned}$$

This operator T is thus called a diagonal operator. Since, $\{\alpha_n\}$ is a bounded sequence, the operator T is bounded. Also,

$$T^*x = \sum_{n=1}^{\infty} \bar{\alpha}_n \langle x, u_n\rangle u_n, \qquad x \in H.$$

Thus, T^* is also diagonal and

$$\begin{aligned} T^*(Tx) &= \sum_{n=1}^{\infty} \bar{\alpha}_n \langle Tx, u_n\rangle u_n \\ &= \sum_{n=1}^{\infty} \bar{\alpha}_n \left\langle \sum_{m=1}^{\infty} \alpha_m \langle x, u_m\rangle u_m \right\rangle u_n \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \bar{\alpha}_n \alpha_m \langle x, u_m\rangle \langle u_m, u_n\rangle u_n \end{aligned}$$

$$= \sum_{n=1}^{\infty} |\alpha_n|^2 \langle x, u_n\rangle u_n \qquad x \in H.$$

Similarly, (Show!) we have

$$TT^*x = \sum_{n=1}^{\infty} |\alpha_n|^2 \langle x, u_n\rangle u_n, \quad x \in H.$$

Therefore $TT^* = T^*T$, i.e. the diagonal operator is a normal operator.

Note: From above example, it follows that, unitary diagonal operator is self-adjoint if and only if all the diagonal entries are ± 1.

Now, we shall derive the condition on a matrix of an operator T for T to be normal, unitary and self-adjoint. Consider the following example.

Example 3.2.4. Let H be a separable of Hilbert space and $u_1, u_2, \ldots$ be orthonormal basis for H. Let $T \in BL(H)$ and $M = (\alpha_{ij})$ be the matrix of T with respect to this orthonormal basis, i.e.

$$\alpha_{ij} = \langle Tu_j, u_i\rangle \qquad \forall\ i, j$$

and

$$Tu_j = \sum_{i=1}^{\infty} \alpha_{ij} u_i.$$

So, $T^*u_k = \sum_{m=1}^{\infty} \beta_{mk} u_m$, where $\beta_{mk} = \bar{\alpha}_{km}$. Now,

$$\begin{aligned}
\langle T^*Tu_j, u_i\rangle &= \langle Tu_j, Tu_i\rangle \\
&= \left\langle \sum_{n=1}^{\infty} \alpha_{nj} u_n, \sum_{m=1}^{\infty} \alpha_{mi} u_m \right\rangle \\
&= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_{nj}\bar{\alpha}_{mi} \langle u_n, u_m\rangle \\
&= \sum_{n=1}^{\infty} \alpha_{nj}\bar{\alpha}_{ni}.
\end{aligned}$$

Similarly, $\langle TT^*u_j u_i\rangle = \sum_{n=1}^{\infty} \bar{\alpha}_{jn}\alpha_{in}$. Thus,

1. T is normal if and only if for each i, j

$$\sum_{n=1}^{\infty} \bar{\alpha}_{jn}\alpha_{in} = \sum_{n=1}^{\infty} \alpha_{nj}\bar{\alpha}_{ni}.$$

2. T is unitary if and only if for each i, j

$$\sum_{n=1}^{\infty} \bar{\alpha}_{jn}\alpha_{in} = \delta_{ij} = \sum_{n=1}^{\infty} \alpha_{nj}\bar{\alpha}_{ni},$$

where $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$

3. An operator T is self-adjoint if and only if the matrix of T is conjugate symmetry (i.e. matrix of T is same as the conjugate of its transpose).

Theorem 3.2.5. *Let H be a Hilbert space and $T \in BL(H)$ be self-adjoint. Then*

$$\|T\| = \sup\{|\langle Tx, x\rangle| : x \in H, \|x\| \leq 1\}.$$

Proof. Let $\alpha = \sup\{|\langle Tx, x\rangle| : x \in H,\ \|x\| \leq 1\}$. Then clearly $\alpha \leq \|T\|$ as

$$\begin{aligned}\alpha = &\ \sup\{|\langle Tx, x\rangle| : x \in H,\ \|x\| \leq 1\} \\ \leq &\ \sup\{|\langle Tx, y\rangle| : x, y \in H,\ \|x\| \leq 1,\ \|y\| \leq 1\} = \|T\|.\end{aligned}$$

Now, we show that $\|T\| \leq \alpha$. For $x \in H$, $x \neq 0$, take $y = \frac{x}{\|x\|}$. Then $\|y\| = 1$ and by definition of α, we have

$$|\langle Ty, y\rangle| = \left|\left\langle \frac{Tx}{\|x\|}, \frac{x}{\|x\|}\right\rangle\right| \leq \alpha.$$

Therefore, for all $x \in H$, we have

$$|\langle Tx, x\rangle| \leq \alpha \|x\|^2. \tag{3.4}$$

Now, for $x, y \in H$,

$$\begin{aligned}\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle &= \langle Tx, x\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle + \langle Ty, y\rangle \\ &\quad - [\langle Tx, x\rangle - \langle Tx, y\rangle - \langle Ty, x\rangle + \langle Ty, y\rangle] \\ &= 2[\langle Tx, y\rangle + \langle Ty, x\rangle] \\ &= 2[\langle Tx, y\rangle + \langle y, Tx\rangle] \qquad (\because T \text{ is self-adjoint}) \\ &= 2[\langle Tx, y\rangle + \overline{\langle Tx, y\rangle}] \\ &= 4\,\mathrm{Re}\langle Tx, y\rangle.\end{aligned}$$

Now,

$$\begin{aligned}4\,\mathrm{Re}\langle Tx, y\rangle &\leq |\langle T(x+y), x+y\rangle| + |\langle T(x-y), x-y\rangle| \\ &\leq \alpha(\|x+y\|^2 + \|x-y\|^2) \qquad \text{(by (3.4))} \\ &= 2\alpha(\|x\|^2 + \|y\|^2) \quad \forall\ x, y \in H \qquad \text{(by Parellogram law).}\end{aligned}$$

Thus, if $x, y \in H$ with $\|x\| \leq 1$, $\|y\| \leq 1$, then $4\,\mathrm{Re}\langle Tx, y\rangle \leq 4\alpha$ or

$$\mathrm{Re}\langle Tx, y\rangle \leq \alpha \tag{3.5}$$

for all $x, y \in H$ with $\|x\| \leq 1$ and $\|y\| \leq 1$. Take $x, y \in H$ with $\|x\| \leq 1$, $\|y\| \leq 1$ and $\langle Tx, y\rangle = re^{i\theta}$, where $r = |\langle Tx, y\rangle|$. Take $x_0 = e^{-i\theta}x$. Then $\|x_0\| = \|x\| \leq 1$ and

$$\langle Tx_0, y\rangle = e^{-i\theta}\langle Tx, y\rangle = r = |\langle Tx, y\rangle|.$$

Therefore, by equation (3.5), since $\|x_0\| \leq 1$ and $\|y\| \leq 1$, we have

$$|\langle Tx, y\rangle| = \langle Tx_0, y\rangle = \mathrm{Re}\langle x_0, y\rangle \leq \alpha.$$

Taking supremum over all $x, y \in H$ with $\|x\| \leq 1$, $\|y\| \leq 1$, we obtain

$$\|T\| \leq \alpha.$$

$$\therefore \|T\| = \sup\{|\langle Tx, x\rangle| : x \in H, \|x\| \leq 1\}.$$

□

Corollary 3.2.6. *Let H be a Hilbert space and $T \in BL(H)$ be self-adjoint. Then $\langle Tx, x\rangle = 0$ for all $x = 0$ if and only if $T = 0$.*

Proof. By above theorem, we have

$$\begin{aligned} & \|T\| = 0 \\ \Leftrightarrow\ & \sup\{|\langle Tx, x\rangle| : x \in H,\ \|x\| \leq 1\} = 0 \\ \Leftrightarrow\ & \langle Tx, x\rangle = 0,\ \forall\ x \in H. \end{aligned}$$

□

Consider the following example in which $\langle Tx, x\rangle = 0$ for all $x \in H$ but $T \neq 0$. Then we show that T is not self-adjoint.

Example 3.2.7. Take $H = \mathbb{R}^2$ and define $T : H \to H$ by $T(x(1), x(2)) = (-x(2), x(1))$ for all $x = (x(1), x(2)) \in \mathbb{R}^2 = H$. Then

$$\begin{aligned} \langle Tx, x\rangle &= \langle(-x(2), x(1)), (x(1), x(2))\rangle \\ &= -\,x(2)x(1) + x(1)x(2) = 0. \end{aligned}$$

Thus, $\langle Tx, x\rangle = 0$ for all $x \in H$. But notice that $T \neq 0$ as

$$T(1, 0) = (0, 1) \neq (0, 0).$$

Then by above corollary, T cannot be self-adjoint. Consider the matrix of T given by

$$m(T) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Then the matrix of T^* is

$$m(T^*) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Thus, $T \neq T^*$ and so T is not self-adjoint.

Proposition 3.2.8. *Let H be a Hilbert space and $T \in BL(H)$. Then*

1. *T is isometry if and only if $T^*T = I$.*
2. *T is unitary if and only if T is an onto isometry. In that case, $\|T^{-1}(x)\| = \|x\|$ for all $x \in H$.*
3. *T is normal if and only if $\|Tx\| = \|T^*x\|$ for all $x \in H$.*

Proof. 1. For every $x \in H$,

$$\begin{aligned} & \|Tx\|^2 = \|x\|^2 \\ \Leftrightarrow\ & \langle Tx, x\rangle = \langle x, x\rangle \\ \Leftrightarrow\ & \langle T^*Tx, x\rangle = \langle x, x\rangle \\ \Leftrightarrow\ & \langle (T^*T - I)x, x\rangle = 0 \\ \Leftrightarrow\ & T^*T - I = 0 \qquad (\because T^*T - I \text{ is self-adjoint, by Corollary 3.2.6}) \\ \Leftrightarrow\ & T^*T = I. \end{aligned}$$

2. If T is unitary, then $T^*T = I = TT^*$. As $T^*T = I$ by (1) above, T is an isometry. Let $y \in H$ and let $x = T^*y$. Then
$$Tx = T(T^*y) = Iy = y.$$
Thus, T is onto. Since, T is unitary, $T^{-1} = T^*$. Then,
$$\|x\| = \|TT^*(x)\| = \|T(T^*x)\| = \|T^*x\| = \|T^{-1}x\| \qquad (\because T^* \text{ is isometry}).$$
Conversely, assume that T is an onto isometry. Since, T is an isometry, clearly T is one-one. Thus, $T : H \to H$ is one-one and onto and hence T is invertible. Also, since T is isometry by (1) above, $T^*T = I$. Now,
$$\begin{aligned} TT^* &= (TT^*)(TT^{-1}) \\ &= T(T^*T)T^{-1} \\ &= TIT^{-1} = I. \end{aligned}$$
Therefore, $TT^* = I = T^*T$, i.e. T is unitary.
3. For every $x \in H$,
$$\begin{aligned} & \|Tx\|^2 = \|T^*x\|^2 \\ \Leftrightarrow\ & \langle Tx, Tx\rangle = \langle T^*x, T^*x\rangle \\ \Leftrightarrow\ & \langle T^*Tx, x\rangle = \langle TT^*x, x\rangle \\ \Leftrightarrow\ & \langle (T^*T - TT^*)x, x\rangle = 0 \\ \Leftrightarrow\ & T^*T - TT^* = 0 \qquad (\because T^*T - TT^* \text{ is self-adjoint, by Corollary 3.2.6}) \\ \Leftrightarrow\ & T^*T = TT^*. \end{aligned}$$
Therefore, T is normal if and only if $\|Tx\| = \|T^*x\|$ for all $x \in H$.

□

Corollary 3.2.9. *Let H be a Hilbert space and $T \in BL(H)$ be normal. Then*
$$\|T^2\| = \|T^*T\| = \|T\|^2 = \|T^*\|^2 = \|(T^*)^2\|.$$

Proof. Since T is normal, by above theorem, for $x \in H$
$$\|T^2x\| = \|T(Tx)\| = \|T^*(Tx)\|.$$
Taking supremum over all $x \in H$ with $\|x\| \le 1$, we get
$$\|T^2\| = \|T^*T\| = \|T\|^2 \qquad \text{(the last equality by Proposition 3.1.9)}.$$
Replacing T by T^*, since T is normal, we get
$$\|(T^*)^2\| = \|T^*T\| = \|TT^*\| = \|T^*\|^2.$$
Hence, the result. □

Now, we investigate whether and under what conditions sums, products (compositions) and limits of self-adjoint, normal and unitary operators are self-adjoint, normal and unitary respectively.

Theorem 3.2.10. *Let H be a Hilbert space.*

(a) Let S and T be self-adjoint. Then $S+T$ is self-adjoint. Also, ST is self-adjoint if and only if S and T commutes.

(b) Let S and T be unitary. Then ST is unitary. Also, $S+T$ is unitary if and only if it is surjective and $\operatorname{Re}\langle Sx, Tx\rangle = -\frac{1}{2}$ for every $x \in H$ with $\|x\| = 1$.

(c) Let S and T be normal. If S commutes with T^* and (hence) T commutes with S^* then $S+T$ and ST are normal.

Proof. (a) Suppose S and T are self-adjoint, i.e. $S = S^*$ and $T = T^*$. Then

$$(S+T)^* = S^* + T^* = S + T.$$

Thus, $S+T$ is self-adjoint. Also, ST is self-adjoint if and only if

$$(ST) = (ST)^* = T^*S^* = TS.$$

Thus, ST is self-adjoint if and only if S and T commutes.

(b) S and T are unitary. Therefore,

$$SS^* = I = S^*S \qquad \text{and} \qquad TT^* = I = T^*T.$$

Then

$$(ST)^*(ST) = (T^*S^*)ST = T^*(S^*S)T = T^*T = I$$

and

$$ST(ST)^* = ST(T^*S^*) = S(TT^*)S^* = SS^* = I.$$

Thus, ST is unitary. Since S and T are unitary, by (2) of Proposition 3.2.8, S and T are surjective isometry. Then,

$$\begin{aligned} \|(S+T)x\|^2 &= \langle (S+T)x, (S+T)x\rangle \\ &= \langle Sx, Sx\rangle + \langle Tx, Tx\rangle + \langle Sx, Tx\rangle + \langle Tx, Sx\rangle \\ &= \|x\|^2 + \|x\|^2 + 2\operatorname{Re}\langle Sx, Tx\rangle \qquad (\because S, T \text{ are isometry}). \end{aligned}$$

Thus, by (2) of Proposition 3.2.8, $S+T$ is unitary if and only if $S+T$ is surjective and it is isometry, i.e. $\|(S+T)x\| = \|x\|$. That is, $S+T$ is unitary if and only if $S+T$ is surjective and $\|x\|^2 = \|(S+T)x\|^2 = \|x\|^2 + \|x\|^2 + 2\operatorname{Re}\langle Sx, Tx\rangle$ or $\|x\|^2 + 2\operatorname{Re}\langle Sx, Tx\rangle = 0$. Thus, if $x \in H$ with $\|x\| = 1$ then $\operatorname{Re}\langle Sx, Tx\rangle = -\frac{1}{2}$.

(c) Suppose S, T are normal and S commutes with T^*, i.e. $ST^* = T^*S$ and T commutes with S^*, i.e. $TS^* = S^*T$. Then,

$$\begin{aligned} (S+T)^*(S+T) &= (S^* + T^*)(S+T) \\ &= S^*S + T^*S + S^*T + T^*T \\ &= SS^* + ST^* + TS^* + TT^* \\ &= S(S^* + T^*) + T(S^* + T^*) \\ &= (S+T)(S^* + T^*). \end{aligned}$$

Thus, $S+T$ is normal. Also, ST is normal, as

$$(ST)^*(ST) = (T^*S^*)(ST)$$

$$
\begin{aligned}
&= T^*(S^*S)T \\
&= T^*(SS^*)T && (\because S \text{ is normal}) \\
&= (T^*S)(S^*T) \\
&= (ST^*)(TS^*) && (\text{by assumption}) \\
&= S(TT^*)S^* && (\because T \text{ is normal}) \\
&= (ST)(ST)^*.
\end{aligned}
$$

□

Theorem 3.2.11. *Let H be a Hilbert space. Then the set of all normal operators, the set of all unitary operators and the set of all self-adjoint operators in $BL(H)$ are closed in $BL(H)$.*

Proof. Let H be a Hilbert space and consider a sequence of operators $\{S_n\}$ in $BL(H)$ such that $S_n \to S$, i.e. $\|S_n - S\| \to 0$. Then $S_n^* \to S^*$.

- If $\{S_n\}$ is a sequence of normal operators, then

$$SS^* = \lim S_n S_n^* = \lim S_n^* S_n = S^*S.$$

  Thus, S is normal.
- If $\{S_n\}$ is a sequence of unitary operators, then

$$SS^* = \lim S_n S_n^* = I = \lim S_n^* S_n = S^*S.$$

  Thus, S is unitary.
- If $\{S_n\}$ is a sequence of self-adjoint operators, then

$$S^* = \lim S_n^* = I = \lim S_n = S.$$

  Thus, S is self-adjoint.

□

Theorem 3.2.12. *Let H be a Hilbert space over $K = \mathbb{C}$ and $S \in BL(H)$. Then there are unique self-adjoint operators A and B in $BL(H)$ such that $S = A + iB$.*

Proof. Let

$$A = \frac{S + S^*}{2} \qquad \text{and} \qquad B = \frac{S - S^*}{2i}.$$

Then,

$$A + iB = \left(\frac{S + S^*}{2}\right) + i\left(\frac{S - S^*}{2i}\right) = S.$$

It is easy to see that (Check!) A and B are self-adjoint. Now to prove uniqueness, let A_1 and B_1 be self-adjoint operators in $BL(H)$ such that $S = A_1 + iB_1$. Then $S^* = A_1 - iB_1$ and

$$A = \frac{S + S^*}{2} = \frac{(A_1 + iB_1) + (A_1 - iB_1)}{2} = A_1.$$

Similarly,

$$B = \frac{S - S^*}{2i} = \frac{(A_1 + iB_1) - (A_1 - iB_1)}{2i} = B_1.$$

Thus, there are unique self-adjoint operators A and B in $BL(H)$ such that $S = A+iB$. □

Exercise 3.2.13. In the above theorem, show that

1. S is normal if and only if $AB = BA$.
2. S is unitary if and only if $AB = BA$ and $A^2 + B^2 = I$.
3. S is self-adjoint if and only if $B = 0$.

Solution. Seminar exercise. □

3.3 Positive Operators

Definition 3.3.1. Let H be a Hilbert space. An operator $S \in BL(H)$ is called *positive* if S is self-adjoint and $\langle Sx, x\rangle \geq 0$ for all $x \in H$. In this case, we write $S \geq 0$.

Note: If $S, T \in BL(H)$ are self-adjoint then $S - T$ and $T - S$ are self-adjoint. Further, if $S - T \geq 0$ then we may write $S \geq T$ or $T \leq S$.

Exercise 3.3.2. What is a partial order? Show that the above relation "$\geq$" on the set of self-adjoint operators on H is a partial order.

Next, we give couple of examples of positive operators.

Example 3.3.3. Let H be a separable Hilbert space and $u_1, u_2, \ldots$ be orthonormal basis for H. For $n = 1, 2, \ldots$, define

$$P_n(x) = \sum_{j=1}^{n} \langle x, u_j\rangle u_j, \qquad x \in H.$$

Then for $x, y \in H$,

$$\begin{aligned}
\langle P_n(x), y\rangle &= \left\langle \sum_{j=1}^{n} \langle x, u_j\rangle u_j, \sum_{i=1}^{\infty} \langle y, u_i\rangle u_i \right\rangle \\
&= \sum_{j=1}^{n} \sum_{i=1}^{\infty} \langle x, u_j\rangle \overline{\langle y, u_i\rangle} \langle u_j, u_i\rangle \\
&= \sum_{j=1}^{n} \langle x, u_j\rangle \overline{\langle y, u_j\rangle},
\end{aligned}$$

where $y = \sum_{i=1}^{\infty} \langle y, u_i\rangle u_i$ is the Fourier expansion of y. Similarly,

$$\langle x, P_n(y)\rangle = \sum_{j=1}^{n} \langle x, u_j\rangle \overline{\langle y, u_j\rangle}.$$

Hence, for all $x, y \in H$

$$\langle P_n(x), y\rangle = \langle x, P_n(y)\rangle.$$

Thus, P_n is self-adjoint. Also,

$$\langle P_n(x), x\rangle = \sum_{j=1}^{n} |\langle x, u_j\rangle|^2 \geq 0, \qquad \forall\ x \in H.$$

Therefore, P_n is a positive operator, i.e. $P_n \geq 0$.

Example 3.3.4. Let $\{\alpha_n\}$ be a sequence of real numbers. Define $S : H \to H$ by

$$Sx = \sum_{n=1}^{\infty} \alpha_n \langle x, u_n\rangle u_n, \qquad x \in H.$$

Then S is bounded and self-adjoint (as seen before). Also, for all $x \in H$,

$$\begin{aligned}\langle Sx, x\rangle &= \left\langle \sum_{n=1}^{\infty} \alpha_n \langle x, u_n\rangle u_n, x \right\rangle \\ &= \sum_{n=1}^{\infty} \alpha_n \langle x, u_n\rangle \langle u_n, x\rangle \\ &= \sum_{n=1}^{\infty} \alpha_n |\langle x, u_n\rangle|^2.\end{aligned}$$

If $\alpha_n \geq 0$ for all n then $\langle Sx, x\rangle \geq 0$ for all $x \in H$, i.e. $S \geq 0$. Conversely, if $S \geq 0$ then $\langle Sx, x\rangle \geq 0$ for all $x \in H$ and hence $\alpha_n = \langle Su_n, u_n\rangle \geq 0$ for all n.

Theorem 3.3.5 (Generalized Schwarz inequality)**.** *Let H be a Hilbert space and $S \in BL(H)$. Then S or $-S$ is positive if and only if*

$$|\langle Sx, y\rangle|^2 \leq \langle Sx, x\rangle\langle Sy, y\rangle, \qquad \forall\ x, y \in H.$$

Proof. Suppose S is a positive operator, i.e. $\langle Sx, x\rangle \geq 0$ for all $x \in H$. For $x, y \in H$, define

$$\langle x, y\rangle_S = \langle Sx, y\rangle.$$

Then (Show that)

- $\langle x, x\rangle_S \geq 0$ for all $x \in H$.
- The function $\langle \cdot, \cdot\rangle_S$ from $H \times H$ to K is linear in first variable.
- The function $\langle \cdot, \cdot\rangle_S$ is conjugate symmetry ($\because$ S is self-adjoint).

We have to prove that for all $x, y \in H$,

$$|\langle x, y\rangle_S|^2 \leq \langle x, x\rangle_S \langle y, y\rangle_S.$$

The proof of the above follows exactly as in Schwarz inequality provided that $\langle y, y\rangle_S \neq 0$. If $\langle y, y\rangle_S = 0$ but $\langle x, x\rangle_S \neq 0$ then we can interchange the role of x and y to have the above inequality. Now, it remains to show that above inequality is true for $\langle x, x\rangle_S = 0 = \langle y, y\rangle_S$. Then, in this case

$$\langle x + y, x + y\rangle_S + \langle x - y, x - y\rangle_S = 2\langle x, x\rangle_S + \langle y, y\rangle_S = 0.$$

Therefore, $\langle x+y, x+y\rangle_S = 0 = \langle x-y, x-y\rangle_S$. Replacing y by iy, we get $\langle x+iy, x+iy\rangle_S = 0 = \langle x-iy, x-iy\rangle_S$. Hence,

$$\begin{aligned} 4\langle x, y\rangle_S &= \langle x+y, x+y\rangle_S - \langle x-y, x-y\rangle_S \\ &\quad + i\langle x+iy, x+iy\rangle_S - i\langle x-iy, x-iy\rangle_S \\ &= 0. \end{aligned}$$

Thus, in any case, for all $x, y \in H$, we have

$$|\langle x, y\rangle_S|^2 \leq \langle x, x\rangle_S \langle y, y\rangle_S,$$

provided that S is a positive operator. If, in case, $-S$ is positive, then by the above case (as proved earlier)

$$\begin{aligned} |\langle Sx, y\rangle|^2 &= |\langle(-S)x, y\rangle|^2 \\ &\leq \langle(-S)x, x\rangle\langle(-S)y, y\rangle \\ &= \langle Sx, x\rangle\langle Sy, y\rangle \end{aligned}$$

for all $x, y \in H$.

Conversely, assume that

$$|\langle Sx, y\rangle|^2 \leq \langle Sx, x\rangle\langle Sy, y\rangle, \qquad \forall\, x, y \in H.$$

Then either $\langle Sx, x\rangle \geq 0$ for all $x \in H$ or $\langle Sx, x\rangle \leq 0$ for all $x \in H$, i.e. either S is positive or $-S$ is a positive operator. □

Definition 3.3.6. A self-adjoint operator S on a Hilbert space H is said to be *positive-definite* if $\langle Sx, x\rangle > 0$ for every non-zero $x \in H$.

Note: If S is a positive-definite operator on H, then equality holds in the generalized Schwarz inequality, in above theorem, if and only if x and y are linearly dependent. This follows by observing that

$$\langle x, y\rangle_S = \langle Sx, y\rangle, \qquad \forall\, x, y \in H,$$

defines an inner product on H in this case.

Proposition 3.3.7. *Let H be a Hilbert space and $T \in BL(H)$. Then T is not bounded below if and only if there is a sequence $\{x_n\}$ in H such that $\|x_n\| = 1$ and $Tx_n \to 0$.*

Proof. Suppose T is not bounded below. Then for each $n \in \mathbb{N}$ there exists $x_n \in H$ such that

$$\|Tx_n\| < \frac{1}{n}.$$

Thus, $Tx_n \to 0$.

Conversely, assume that there is a sequence $\{x_n\}$ in H such that $\|x_n\| = 1$ and $Tx_n \to 0$. Suppose, if possible, T is bounded below. Then there exists $\beta > 0$ such that $\beta\|x\| \leq \|Tx\|$ for all $x \in H$. Then

$$0 < \beta = \beta\|x_n\| \leq \lim \|Tx_n\| = 0$$

which is a contradiction and hence no such $\beta > 0$ exists. Therefore, T is not bounded below. □

CHAPTER 4

Spectrum and Numerical Range

4.1 Spectrum of a bounded operator

Definition 4.1.1. Let H be a Hilbert space over K and $T \in BL(H)$. The set

$$\sigma(T) = \{\lambda \in K : T - \lambda I \text{ is not invertible in } BL(H)\}$$

is called the *spectrum* of T.

Elements of $\sigma(T)$ are known as *spectral values* of T.

Definition 4.1.2. Let H be a Hilbert space over K and $T \in BL(H)$. The set

$$\begin{aligned}\sigma_e(T) &= \{\lambda \in K : T - \lambda I \text{ is not one-one}\}\\ &= \{\lambda \in K : \exists\, x \in H,\ \|x\| = 1 \text{ and } (T - \lambda I)x = 0\}.\end{aligned}$$

is known as the *eigen spectrum* of T.

Elements of $\sigma_e(T)$ are known as *eigenvalues* (or characteristic roots) of T.

Definition 4.1.3. Let H be a Hilbert space over K and $T \in BL(H)$. The set

$$\sigma_a(T) = \{\lambda \in K : T - \lambda I \text{ is not bounded below}\}$$

is known as the *approximate eigen spectrum* of T and the elements of $\sigma_a(T)$ are known as *approximate eigenvalues* of T.

By the last Proposition 3.3.7, we have

$$\sigma_a(T) = \{\lambda \in K : \exists \text{ a sequence } \{x_n\} \in H \text{ such that } \|x_n\| = 1 \text{ and } (T - \lambda I)x_n \to 0\}.$$

Proposition 4.1.4. *Let H be a Hilbert space and $T \in BL(H)$. Then $\lambda \in \sigma(T)$ if and only if $\bar{\lambda} \in \sigma(T^*)$, i.e.*

$$\sigma(T) = \{\bar{\mu} : \mu \in \sigma(T^*)\}.$$

Proof. For $T \in BL(H)$,

$$\begin{aligned} \lambda \in \sigma(T) &\Leftrightarrow (T - \lambda I) \text{ is not invertible in } BL(H) \\ &\Leftrightarrow (T - \lambda I)^* \text{ is not invertible in } BL(H) \\ &\Leftrightarrow (T^* - \bar{\lambda} I) \text{ is not invertible in } BL(H) \\ &\Leftrightarrow \bar{\lambda} \in \sigma(T^*). \end{aligned}$$

□

Theorem 4.1.5. *Let H be a Hilbert space and $T \in BL(H)$. Then*

1. $\sigma_e(T) \subset \sigma_a(T)$.
2. $\sigma(T) = \sigma_a(T) \cup \{\bar{\mu} : \mu \in \sigma_e(T^*)\}$.

Proof.

1. Let $\lambda \in \sigma_e(T)$. Then there exists $x \in H$ with $\|x\| = 1$ such that $(T - \lambda I)x = 0$. Take $x_n = x$ for all n, then

$$0 = (T - \lambda I)x_n \to 0 \Rightarrow \lambda \in \sigma_a(T).$$

Thus, $\sigma_e(T) \subset \sigma_a(T)$.

2. Let $\lambda \notin \sigma_a(T) \cup \{\bar{\mu} : \mu \in \sigma_e(T^*)\}$. Then $\lambda \notin \sigma_a(T)$ and $\lambda \not\in \{\bar{\mu} : \mu \in \sigma_e(T^*)\}$. Therefore, $(T - \lambda I)$ is bounded below and $T^* - \bar{\lambda} I$ is one-one. Therefore (by Proposition 3.1.17) $R(T - \lambda I)$ is closed in H and (by Theorem 3.1.13)

$$\overline{R(T - \lambda I)} = \ker(T^* - \bar{\lambda} I)^{\perp} = \{0\}^{\perp} = H.$$

Therefore $R(T - \lambda I) = H$, i.e. $T - \lambda I$ is onto.
Since, $(T - \lambda I)$ is bounded below it is one-one. Thus, $(T - \lambda I)$ is one-one and onto and hence it is invertible, i.e. $(T - \lambda I)^{-1}$ exists and it is linear.
Now, since $T - \lambda I$ is bounded below, there exists $\beta > 0$ such that $\|(T - \lambda I)x\| \geq \beta \|x\|$ for all $x\ inH$. Let $y \in H$. Take $x = (T - \lambda I)^{-1} y$. Then

$$\begin{aligned} \|(T - \lambda I)^{-1} y\| &= \|x\| \\ &= \frac{1}{\beta} \|(T - \lambda I)x\| \\ &= \frac{1}{\beta} \|y\|. \end{aligned}$$

Therefore, $(T - \lambda I)^{-1} \in BL(H)$ which implies $\lambda \notin \sigma(T)$. Hence

$$\sigma(T) \subset \sigma_a(T) \cup \{\bar{\mu} : \mu \in \sigma_e(T^*)\}.$$

Now, consider $\lambda \notin \sigma(T)$. Then $(T - \lambda I)^{-1} \in BL(H)$. Therefore for $x \in H$

$$\|x\| = \|(T - \lambda I)^{-1}(T - \lambda I)x\|$$

$$= \|(T - \lambda I)^{-1}\| \|(T - \lambda I)x\|$$
$$\beta \|x\| = \|(T - \lambda I)x\|,$$

where $\beta = \frac{1}{\|(T-\lambda I)^{-1}\|}$. Therefore, $T - \lambda I$ is bounded below, i.e. $\lambda \notin \sigma_a(T)$ and hence

$$\sigma_a(T) \subset \sigma(T). \tag{4.1}$$

Now, let $\lambda \in \{\bar{\mu} : \mu \in \sigma_e(T^*)\}$. Then

$$\bar{\lambda} \in \sigma_e(T^*) \subset \sigma_a(T^*) \subset \sigma(T^*).$$

Therefore $\bar{\lambda} \in \sigma(T^*) \Rightarrow \lambda \in \sigma(T)$. Therefore, $\{\bar{\mu} : \mu \in \sigma_e(T^*)\} \subset \sigma(T)$. Hence,

$$\sigma(T) = \sigma_a(T) \cup \{\bar{\mu} : \mu \in \sigma_e(T^*)\}.$$

□

Remark 4.1.6. From the above theorem and from equation (4.1), we have

$$\sigma_e(T) \subset \sigma_a(T) \subset \sigma(T).$$

If H is finite dimensional, i.e. $\dim H < \infty$ then

$$\sigma_e(T) = \sigma_a(T) = \sigma(T).$$

However, in general, it is possible that

$$\sigma_e(T) \subsetneq \sigma_a(T) \subsetneq \sigma(T).$$

Consider the following examples.

Example 4.1.7. Define $T : \ell^2 \to \ell^2$ by

$$T(x(1), x(2), \ldots) = \left(x(1), \frac{x(2)}{2}, \frac{x(3)}{3}, \ldots\right).$$

Then observe that $Tx = 0 \Rightarrow x = 0$, i.e. there does not exists a $x \neq 0,\ x \in \ell^2$ such that $Tx = 0$. In other words, $T = T - 0I$ is one-one. Therefore

$$0 \notin \sigma_e(T).$$

Now, $Te_n = \frac{1}{n} e_n$ for all n, where $e_n = (0, 0, \ldots, 0, \underbrace{1}_{n^{\text{th}} \text{ place}}, 0, \ldots)$.

Since $\|e_n\| = 1$ and $Te_n \to 0$, by definition we have $0 \in \sigma_a(T)$. Thus,

$$\sigma_e(T) \subsetneq \sigma_a(T).$$

Example 4.1.8. Consider the right shift operator on ℓ^2, $T : \ell^2 \to \ell^2$ defined by

$$T(x(1), x(2), \ldots) = (0, x(1), x(2), \ldots).$$

Then, $\|Tx\| = \|x\|$. Therefore, there does not exist a sequence $\{x_n\}$ such that $\|x_n\| = 1$ and $Tx_n \to 0$ and hence $0 \notin \sigma_a(T)$.

On the other hand, observe that $e_1 \notin R(T)$, i.e. T is not onto. Hence, T is not invertible. Therefore $0 \in \sigma(T)$. Thus,

$$\sigma_a(T) \subsetneq \sigma(T).$$

Proposition 4.1.9. *Let H be a Hilbert space and $T \in BL(H)$ be normal. Then*

$$\lambda \in \sigma_e(T) \text{ if and only if } \bar{\lambda} \in \sigma_e(T^*).$$

In fact, if $x \in H$, $\|x\| = 1$ such that $(T - \lambda I)x = 0$ if and only if $(T^* - \bar{\lambda} I)x = 0$.

Proof. Since T is normal, $T - \lambda I$ is normal. So, by a previous result (Proposition 3.2.8)

$$\|(T - \lambda I)x\| = \|(T^* - \bar{\lambda} I)x\| \qquad \forall\, x \in H.$$

Thus, $\lambda \in \sigma_e(T)$ if and only if there exists $x \in H$ with $\|x\| = 1$ such that $(T - \lambda I)x = 0$
if and only if (by above) $(T^* - \bar{\lambda} I)x = (T - \lambda I)^* x = 0$
if and only if $\bar{\lambda} \in \sigma_e(T^*)$. □

Corollary 4.1.10. *Let H be a Hilbert space and $T \in BL(H)$ be normal. Then*

$$\sigma(T) = \sigma_a(T).$$

Proof. By previous theorem, we know that

$$\sigma(T) = \sigma_a(T) \cup \{\bar{\mu} : \mu \in \sigma_e(T^*)\}.$$

Now, we show that if T is normal then $\{\bar{\mu} : \mu \in \sigma_e(T^*)\} \subset \sigma_a(T)$. Then we are done (as this will give $\sigma(T) = \sigma_a(T)$ from above theorem).

Let $\lambda \in \{\bar{\mu} : \mu \in \sigma_e(T^*)\}$, i.e. $\bar{\lambda} \in \sigma_e(T^*)$. Then, since T is normal, by previous proposition

$$\lambda \in \sigma_e(T) \subset \sigma_a(T).$$

Thus, $\{\bar{\mu} : \mu \in \sigma_e(T^*)\} \subset \sigma_a(T)$ and hence $\sigma(T) = \sigma_a(T)$ by above theorem. □

Example 4.1.11. In this example we show that above proposition is not true if T is not normal, i.e. $\lambda \in \sigma_e(T) \not\Rightarrow \bar{\lambda} \in \sigma_e(T^*)$ in general (if T is not normal).

Define $T : K^2 \to K^2$ by $T(x(1), x(2)) = (ix(1) + x(2), ix(2))$. Then

$$m(T) = \begin{bmatrix} i & 1 \\ 0 & i \end{bmatrix}$$

with respect to the orthonormal basis $\{e_1, e_2\}$ and

$$m(T^*) = \begin{bmatrix} -i & 0 \\ 0 & -i \end{bmatrix}$$

Therefore, $T^*(x(1), x(2)) = (-ix(1), x(1) - ix(2))$.

Now, $T(x(1), x(2)) = i(x(1), x(2))$ if and only if $(ix(1) + x(2), ix(2)) = i(x(1), x(2))$ if and only if $x(2) = 0$. Thus,

$$T(1, 0) = i(1, 0) \qquad \text{i.e. } Te_1 = ie_1.$$

On the other hand, $T * (x(1), x(2)) = -i(x(1), x(2))$ if and only if $(-ix(1), x(1) - ix(2)) = -i(x(1), x(2))$ if and only if $x(1) = 0$. Thus,

$$T(1, 0) = i(1, 0) \;\text{ but }\; T^*(1, 0) \neq \bar{i}(1, 0).$$

Thus, $i \in \sigma_e(T)$ but $\bar{i} = -i \notin \sigma_e(T^*)$. Note that here $T^*(0, 1) = \bar{i}(0, 1)$.

Consider one more example of the same, i.e. $\lambda \in \sigma_e(T) \not\Rightarrow \bar{\lambda} \in \sigma_e(T^*)$ in general (when T is not normal).

Example 4.1.12. Consider the left-shift operator on ℓ^2, $S : \ell^2 \to \ell^2$ defined by

$$S(x(1), x(2), \ldots) = (x(2), x(3), \ldots).$$

Then, $Se_1 = 0$, where $e_1 = (1, 0, 0, \ldots) \neq 0$. Thus, $S = S - 0I$ is not one-one, i.e.

$$0 \in \sigma_e(T).$$

As seen in Example 3.1.1, S^* is the right-shift operator defined by

$$S^*(x(1), x(2), \ldots = (0, x(1), x(2), \ldots).$$

Then $S^*e_1 = e_2 \neq 0$. Observe that $S^* = S^* - 0I$ is one-one, i.e. $\bar{\lambda} = 0 \notin \sigma_e(S^*)$ but $\lambda = 0 \in \sigma_e(S)$.

Proposition 4.1.13. *Let H be a Hilbert space and $T \in BL(H)$ be normal. If $x \in H$ such that $(T - \lambda I)^2 x = 0$ then $(T - \lambda I)x = 0$.*

Proof. Suppose $x \in H$ such that $(T - \lambda I)^2 x = 0$. Then

$$\begin{aligned}
\|(T - \lambda I)x\|^2 &= \langle (T - \lambda I)x, (T - \lambda I)x \rangle \\
&= \langle (T - \lambda I)^*(T - \lambda I)x, x \rangle \\
&\leq \|(T - \lambda I)^*(T - \lambda I)x\| \|x\| && \text{(by Schwarz inequality)} \\
&= \|(T - \lambda I)(T - \lambda I)x\| \|x\| && (\because T - \lambda I \text{ is normal}) \\
&= \|(T - \lambda I)^2 x\| = 0.
\end{aligned}$$

□

In the following example, we show that the above result is not true if T is not normal.

Example 4.1.14. We show that $T^2 x = 0$ then $Tx = 0$ is not true in general, i.e. if T is not normal. Define $T : K^2 \to K^2$ by

$$T(x(1), x(2)) = (0, x(1)).$$

Then $T^2(x(1), x(2)) = T(0, x(1)) = 0$ for all $(x(1), x(2)) \in K^2$. But

$$T(1, 0) = (0, 1) \neq 0.$$

Proposition 4.1.15. *Let H be a Hilbert space and $T \in BL(H)$ be normal. Then the eigenvectors corresponding to distinct eigenvalues of T are orthogonal i.e. if $\lambda, \mu \in \sigma_e(T)$, $\lambda \neq \mu$ and $x \neq 0, y \neq 0$ are such that $Tx = \lambda x$ and $Ty = \mu y$ then $\langle x, y \rangle = 0$.*

Proof. Suppose $\lambda, \mu \in \sigma_e(T)$, $\lambda \neq \mu$ and $x \neq 0, y \neq 0$ are such that $Tx = \lambda x$ and $Ty = \mu y$ then $\langle x, y\rangle = 0$. then

$$\begin{aligned}\lambda\langle x, y\rangle &= \langle \lambda x, y\rangle \\ &= \langle Tx, y\rangle \\ &= \langle x, T^* y\rangle \\ &= \langle x, \bar{\mu} y\rangle \qquad \text{(by Proposition 4.1.9)} \\ &= \mu\langle x, y\rangle.\end{aligned}$$

Thus,

$$(\lambda - \mu)\langle x, y\rangle = \lambda\langle x, y\rangle = \mu\langle x, y\rangle = 0.$$

Therefore, $\langle x, y\rangle = 0$ as λ and μ are distinct eigenvalues, i.e. $\lambda \neq \mu$. □

In the following example, we show that the above result is not true if T is not normal.

Example 4.1.16. We show that eigenvectors corresponding to distinct eigenvalues need not be orthogonal in general, i.e. if T is not normal. Define $T : K^2 \to K^2$ by

$$T(x(1), x(2)) = (x(1) + x(2), 2x(2)).$$

Then $Te_1 = T(1, 0) = (1, 0)$, i.e. 1 is eigenvalue of T and $(1, 0)$ is the corresponding eigenvector.

Also, $T(1, 1) = (2, 2)$, i.e. $(1, 1)$ is the eigenvector corresponding to eigenvalue 2 of T. But the eigenvectors are not orthogonal as

$$\langle (1, 0), (1, 1)\rangle = 1 \neq 0.$$

4.2 Numerical range of a bounded operator

Definition 4.2.1 (Numerical range)**.** Let H be a Hilbert space and $T \in BL(H)$. The set

$$W(T) = \{\langle Tx, x\rangle : x \in H,\ \|x\| = 1\}$$

is called *numerical range* of T.

Remark 4.2.2. For $x \in H$ with $\|x\| = 1$,

$$|\langle Tx, x\rangle| \leq \|Tx\|\|x\| \leq \|T\|\|x\|^2 = \|T\|.$$

Thus, the numerical range $W(T)$ is bounded by $\|T\|$.

However, it is not closed but it is convex.

Proposition 4.2.3. *Let H be a Hilbert space and $T \in BL(H)$. Then*

1. *$\lambda \in W(T)$ if and only if $\bar{\lambda} \in W(T^*)$.*
2. *$\sigma_e(T) \subset W(T)$.*

Proof. 1. Let $\lambda \in W(T)$. Then there exists $x \in H$ such that $\|x\| = 1$ and $\lambda = \langle Tx, x\rangle$. Therefore

$$\bar{\lambda} = \overline{\langle Tx, x\rangle} = \overline{\langle x, T^*x\rangle} = \langle T^*x, x\rangle \in W(T^*).$$

Thus, $\lambda \in W(T) \Rightarrow \bar{\lambda} \in W(T^*)$. For the converse part, replace λ by $\bar{\lambda}$.

2. Let $\lambda \in \sigma_e(T)$. Then there exists $x \in H$, $\|x\| = 1$ such that $(T - \lambda I)x = 0$, i.e. $Tx = \lambda x$. Then

$$\lambda = \lambda\langle x, x\rangle = \langle \lambda x, x\rangle = \langle Tx, x\rangle \in W(T).$$

Therefore, $\sigma_e(T) \subset W(T)$.

□

Proposition 4.2.4. *Let H be a Hilbert space and $T \in BL(H)$. Then $\sigma_a(T) \subset \overline{W(T)}$.*

Proof. Let $\lambda \in \sigma_a(T)$. Then there exists a sequence $\{x_n\}$ in H with $\|x_n\| = 1$ for all n such that $(T - \lambda I)x_n \to 0$. Therefore

$$\begin{aligned} |\langle Tx_n, x_n\rangle - \lambda| &= |\langle Tx_n, x_n\rangle - \lambda\langle x_n, x_n\rangle| \\ &= |\langle (T - \lambda I)x_n, x_n\rangle| \\ &\leq \|(T - \lambda I)x_n\| \to 0. \end{aligned}$$

Thus, $\langle Tx_n, x_n\rangle \to \lambda$ in K and hence $\lambda \in \overline{W(T)}$ ($\because \langle Tx_n, x_n\rangle \in W(T)$ for all n). □

Corollary 4.2.5. *Let H be a Hilbert space. If $T \in BL(H)$ then $\sigma(T) \subset W(T)$.*

Proof. We know that

$$\sigma(T) \subset \sigma_a(T) \cup \{\bar{\mu} : \mu \in \sigma_e(T^*)\}.$$

Let $\lambda \in \{\bar{\mu} : \mu \in \sigma_e(T^*)\}$. Then $\bar{\lambda} \in \sigma_e(T^*)$. By previous result (2. of Proposition 4.2.3), we have $\bar{\lambda} \in W(T^*)$ and hence $\lambda \in W(T)$. By above proposition, we already have $\sigma_a(T) \subset \overline{W(T)}$ and hence we conclude that $\sigma(T) \subset \overline{W(T)}$. □

Example 4.2.6. By above corollary we have $\sigma(T) \subset \overline{W(T)}$. It is not true in general that $\sigma(T) \subset W(T)$. Consider the diagonal operator on ℓ^2 having diagonal entries $1, \frac{1}{2}, \frac{1}{3}, \ldots$, i.e. $T : \ell^2 \to \ell^2$ defined by

$$T(x(1), x(2), x(3), \ldots) = \left(x(1), \frac{x(2)}{2}, \frac{x(3)}{3}, \ldots\right).$$

Then $Te_n = \frac{e_n}{n}$ for all $n = 1, 2, \ldots$. Then $\|e_n\| = 1$ and $Te_n \to 0$ ($\because \|Te_n\| = \frac{1}{n} \to 0$). Therefore,

$$0 \in \sigma_a(T) \subset \sigma(T).$$

But for $x = (x(1), x(2), \ldots) \in \ell^2$ with $\|x\| = 1$, we have

$$\langle Tx, x\rangle = \sum_{n=1}^{\infty} \frac{x(n)}{n}\bar{x}(n) = \sum_{n=1}^{\infty} \frac{|x(n)|^2}{n} \neq 0.$$

Therefore $0 \notin W(T)$.

Remark 4.2.7. Let H be a Hilbert spae and $T \in BL(H)$ be self-adjoint. Then

1. $W(T) \subset \mathbb{R}$.
   This is true, indeed if $x \in H$, $\|x\| = 1$ then
   $$\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}.$$
   Therefore, $\langle Tx, x\rangle \in \mathbb{R}$ and hence $W(T) \subset \mathbb{R}$.
2. $\sigma(T) \subset \mathbb{R}$.
   Indeed $\sigma(T) \subset \overline{W(T)} \subset \mathbb{R}$.

Notations: Let $T \in BL(H)$ be self-adjoint. Consider

$$m_T = \inf\{\lambda : \lambda \in W(T)\}$$
$$M_T = \sup\{\lambda : \lambda \in W(T)\}$$

The inf and sup exists because $W(T)$ is a bounded subset of $\mathbb{R}$.

Theorem 4.2.8. *Let H be a Hilbert space ($H \neq \{0\}$) and $T \in BL(H)$ be self-adjoint. Then*
$$\{m_T, M_T\} \in \sigma_a(T) = \sigma(T) \subset [m_T, M_T].$$

Proof. By definition of m_T there exists a sequence $\{x_n\}$ in H with $\|x_n\| = 1$ for all n such that
$$\langle Tx_n, x_n\rangle \to m_T.$$
Now, since T is self-adjoint, $T - m_T I$ is self-adjoint. Also, for all $x \in H$ with $\|x\| = 1$,
$$\langle (T - m_T I)x, x\rangle = \langle Tx, x\rangle - m_T\langle x, x\rangle = \langle Tx, x\rangle - m_T \geq 0.$$
This is because, by definition of m_T, $\langle Tx, x\rangle \geq m_T$ for all $x \in H$ with $\|x\| = 1$. Thus, $T - m_T I$ is a positive operator, i.e.
$$(T - m_T I) \geq 0.$$
Take $S = T - m_T I$ then by the generalized Schwarz inequality, we have
$$|\langle Sx, y\rangle|^2 \leq \langle Sx, x\rangle\langle Sy, y\rangle, \qquad \forall\ x, y \in H.$$
Therefore, taking $x = x_n$ and $y = Sx_n$ in the above inequality, we get
$$\begin{aligned}\|Sx_n\|^4 &= |\langle Sx_n, Sx_n\rangle|^2 \\ &\leq \langle Sx_n, x_n\rangle\langle S^2x_n, Sx_n\rangle \\ &\leq \langle Sx_n, x_n\rangle\|S\|^3 \qquad (\because \|x_n\| = 1).\end{aligned}$$
Now,
$$\langle Sx_n, x_n\rangle = \langle (T - m_T I)x_n, x_n\rangle = \langle Tx_n, x_n\rangle - m_T \to 0.$$
Therefore from above, we get
$$\|Sx_n\| \to 0, \quad \text{i.e. } Sx_n \to 0.$$

and hence $(T - m_T I)x_n \to 0$, where $\|x_n\| = 1$. Therefore,

$$m_T \in \sigma_a(T).$$

Similarly, one can prove that $M_T \in \sigma_a(T)$ by taking $S = T - M_T I$ and observing that $-S$ is a positive operator.

Since, T is self-adjoint, it is normal and hence $\sigma_a(T) = \sigma(T)$. Thus,

$$m_T, M_T \in \sigma_a(T) = \sigma(T) \subset \overline{W(T)} \subset [m_T, M_T].$$

□

Corollary 4.2.9. *Let H be a Hilbert space and $T \in BL(H)$ be self-adjoint. Then*

1. $\|T\| = \max\{|m_T|, |M_T|\} = \sup\{|\lambda| : \lambda \in \sigma(T)\}$.
2. $\|T\| = \sup\{|\lambda|^{\frac{1}{2}} : \lambda \in \sigma(T^*T)\}$.

Proof. 1. Since T is self-adjoint, by previous theorem, we have

$$m_T, M_T \in \sigma_a(T) = \sigma(T) \subset \overline{W(T)} \subset [m_T, M_T].$$

We know that

$$\|T\| = \sup\{|\langle Tx, x\rangle| : x \in H, \ \|x\| \leq 1\}.$$

Therefore, $\|T\| = \max\{|m_T|, |M_T|\}$. Also, from above theorem it follows that

$$\|T\| = \max\{|m_T|, |M_T|\} = \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

2. Since T is self-adjoint, T^*T is self-adjoint and $\|T\|^2 = \|T^*T\|$. Also, from above theorem

$$\|T^*T\| = \sup\{\lambda : \lambda \in \sigma(T^*T)\}$$

and hence $\|T\| = \sup\{\sqrt{\lambda} : \lambda \in \sigma(T^*T)\}$.

□

Theorem 4.2.10 (Ritz Method). *Let H be a Hilbert space ($H \neq \{0\}$) and $T \in BL(H)$ be self-adjoint. Consider $x_1, x_2, \ldots$ in H. For $n = 1, 2, \ldots$, let*

$$Y_n = L(\{x_1, x_2, \ldots, x_n\}).$$

Take

$$\alpha_n = \inf\{\langle Tx, x\rangle : x \in Y_n, \|x\| = 1\}$$

and

$$\beta_n = \sup\{\langle Tx, x\rangle : x \in Y_n, \|x\| = 1\}.$$

Then

$$m_T \leq \alpha_{n+1} \leq \alpha_n \leq \cdots \leq \alpha_1 \leq \beta_1 \leq \beta_2 \leq \cdots \leq \beta_n \leq \beta_{n+1} \leq M_T.$$

If $L(\{x_1, x_2, \ldots\})$ is dense in H, then

$$m_T = \lim_{n\to\infty} \alpha_n \qquad \text{and} \qquad M_T = \lim_{n\to\infty} \beta_n.$$

Proof. Since $Y_n \subset Y_{n+1}$, it is clear that

$$m_T \leq \alpha_{n+1} \leq \alpha_n \leq \cdots \leq \beta_n \leq \beta_{n+1} \leq M_T.$$

Since $\{\alpha_n\}$ is a non-increasing sequence which is bounded below, it converges. Suppose

$$m_0 = \lim_{n\to\infty} \alpha_n.$$

Then, clearly $m_T \leq m_0$. Suppose if possible, $m_T < m_0$. Then there exists $x \in H$, $\|x\| = 1$ such that

$$m_T < \langle Tx, x\rangle < m_0.$$

Now, since $L(\{x_1, x_2, \ldots\})$ is dense in H, there exists a sequence $\{y_n\}$ in $L(\{x_1, x_2, \ldots\})$ such that $y_n \to x$ in H. Then

$$\|y_n\| \to \|x\| = 1.$$

For sufficiently large n, take $z_n = \frac{y_n}{\|y_n\|}$ then $\|z_n\| = 1$ and $z_n \to x$ ($\because \|x\| = 1$).

Now, since $y_n \in L(\{x_1, x_2, \ldots\})$, there exists an integer j_n such that $y_n \in Y_{j_n} = L(\{x_1, x_2, \ldots, x_{j_n}\})$. Therefore, $z_n \in Y_{j_n}$. Then letting $n \to \infty$, we have

$$m_0 \leq \alpha_{j_n} \leq \langle Tz_n, z_n\rangle \to \langle Tx, x\rangle < m_0$$

which is a contradiction and hence $m_T = m_0 = \lim\limits_{n\to\infty} \alpha_n$.

In the same way, it follows that $\{\beta_n\}$ is a non-decreasing sequence which is bounded above and hence it converges. As above, one can show that $M_T = \lim_{n\to\infty} \beta_n$. Since, $\alpha_n \leq \beta_n$ for each n, the proof is complete. □

4.3 Compact Self-Adjoint Operators

Definition 4.3.1. Let H be a Hilbert space and $T : H \to H$ be a linear map. T is said to be *compact* if for every bounded sequence $\{x_n\}$ in H, $\{Tx_n\}$ has a convergent subsequence.

Example 4.3.2. Every bounded linear operator $T : K \to K$ is compact. This is because if $\{x_n\}$ is a bounded sequence in K then $\{Tx_n\}$ is bounded sequence in K ($\because T \in BL(K)$). Therefore by Bolzano-Weiertrass theorem, $\{Tx_n\}$ has a convergent subsequence. Due to Bolzano-Weiertrass property, this is true for any bounded linear map on K^n. More generally, we shall show that any bounded linear operator with finite-dimensional range is compact.

Proposition 4.3.3. *Let H be a Hilbert space and $T : H \to H$ be compact linear transformation. Then T is bounded.*

Proof. Suppose T is compact on H. To show that T is bounded it suffices to show that there exists $\alpha > 0$ such that

$$\|Tx\| \leq \alpha \quad \forall\, x \in H,\ \|x\| \leq 1.$$

Suppose T is not bounded. Then for each $\beta > 0$ there exists $x_\beta \in H$, $\|x_\beta\| \leq 1$ such that

$$\|Tx_\beta\| > \beta.$$

Therefore for $\beta =$ there exists $x_1 \in H$, $\|x_1\| \leq 1$ such that

$$\|Tx_1\| > 1.$$

Now, take $\beta = 1 + \|Tx_1\|$ then there exists $x_2 \in H$, $\|x_2\| \leq 1$ such that

$$\|Tx_2\| > 1 + \|Tx_1\|.$$

Taking $\beta = 1 + \max\{\|Tx_1\|, \|Tx_2\|\}$ then there exists $x_3 \in H$, $\|x_3\| \leq 1$ such that

$$\|Tx_3\| > 1 + \max\{\|Tx_1\|, \|Tx_2\|\}.$$

Continuing this way, we get $x_1, x_2, \ldots, x_n$ such that $\|x_j\| \leq 1$ for all j and

$$\|Tx_n\| > 1 + \max\{\|Tx_1\|, \|Tx_2\|, \ldots, \|Tx_{n-1}\|\}.$$

Therefore for $n < m$, $\|Tx_m\| > 1 + \|Tx_n\|$. Therefore,

$$\|Tx_m\| - \|Tx_n\| > 1.$$

Hence,

$$\|Tx_m - Tx_n\| \geq \|Tx_m\| - \|Tx_n\| > 1.$$

Therefore $\{Tx_n\}$ has no convergent subsequence and hence T is not compact, which is a contradiction. Hence, T must be bounded. □

The converse of above result is not true, i.e. a bounded linear operator in general need not be compact. Consider the following example of identity operator on an infinite dimensional Hilbert space. It is bounded but not compact.

Example 4.3.4. Suppose H is an infinite dimensional Hilbert space. Let $\{u_1, u_2, \ldots\}$ be infinite orthonormal subset of H. Then clearly $\{u_n\}$ is bounded ($\because \|u_n\| = 1,\ \forall\ n$). Let I be the identity operator on H. Then

$$\|Iu_n = Iu_m\| = \|u_m - u_n\| = \sqrt{2}, \qquad m \neq n.$$

Therefore, $\{Iu_n\} = \{u_n\}$ has no convergent subsequence and hence I is not compact.

Proposition 4.3.5. *Let H be a Hilbert space and $T \in BL(H)$ be such that $R(T)$ is finite dimensional (i.e. rank of T is finite). Then T is compact.*

Proof. Suppose $\dim(R(T)) = m$. Let $\{u_1, u_2, \ldots, u_m\}$ be orthonormal basis of $R(T)$. Define $\phi : R(T) \to (K^m, \|\cdot\|_2)$ by

$$\phi(x) = (\langle x, u_1\rangle, \langle x, u_2\rangle, \cdots, \langle x, u_m\rangle), \qquad x \in R(T).$$

Then ϕ is a linear onto isometry (Verify!).

Now, if $\{x_n\}$ is a bounded sequence in H, then since T is bounded, $\{Tx_n\}$ is bounded. Since ϕ is an isometry, $\{\phi(Tx_n)\}$ is bounded in K^m. Then by Bolzano-Weierstrass property of K^m, $\{\phi(Tx_n)\}$ has a convergent subsequence, say $\{\phi(Tx_{n,1})\}$. Hence, $\{Tx_{n,1}\}$ is convergent subsequence of $\{Tx_n\}$. Therefore, T is compact. □

Corollary 4.3.6. *Let H be a finite dimensional Hilbert space and $T \in BL(H)$. Then T is compact.*

Proof. As $\dim H < \infty$, $\dim R(T) < \infty$. Therefore T is compact by above theorem. □

Remark 4.3.7. The identity operator on a Hilbert space H is compact if and only if H is finite dimensional.

Theorem 4.3.8. *Let H be a Hilbert space and $\{Tx_n\}$ be a sequence of compact operators in $BL(H)$ such that $T_n \to T$ in $BL(H)$ (i.e., $T \in BL(H)$ and $\|T_n - T\| \to 0$). Then T is compact.*

Proof. Suppose $\{x_n\}$ is a bounded sequence in H. Then there exists $\alpha > 0$ such that $\|x_n\| < \alpha$, $\forall\ n$.

Since T_1 is compact and $\{x_n\}$ is bounded, $\{T_1 x_n\}$ has a convergent subsequence, say $\{T_1 x_{n,1}\}$. Now, as $\{x_{n,1}\}$ is bounded (being subsequence of $\{x_n\}$) and T_2 is compact, $\{T_2 x_{n,1}\}$ has a convergent subsequence, say $\{T_2 x_{n_2}\}$.

Note that here the sequence $\{x_{n,2}\}$ is a subsequence of $\{x_{n_1}\}$ and $\{T_1 x_{n,1}\}$ is convergent. Therefore, $\{T_1 x_{n,2}\}$ is also convergent.

Continuing this way, we get convergent sequence $\{T_k x_{n,k}\}$ such that $\{T_j x_{n,k}\}$ is convergent for all $j = 1, 2, \ldots, k$. Therefore $\{T_n x_{k,k}\}$ converges for each n. Now, for $m, k \in \mathbb{N}$

$$\begin{aligned}
\|Tx_{k,k} - Tx_{m,m}\| &\le \|Tx_{k,k} - T_n x_{k,k}\| + \|T_n x_{k,k} - T_n x_{m,m}\| + \|T_n x_{m,m} - Tx_{m,m}\| \\
&\le \|T - T_n\|\|x_{k,k}\| + \|T_n x_{k,k} - T_n x_{m,m}\| + \|T - T_n\|\|x_{m,m}\| \\
&\le 2\alpha\|T - T_n\| + \|T_n x_{k,k} - T_n x_{m,m}\| \\
&\to 0 \text{ as } k, m \to \infty \text{ and } n \to \infty.
\end{aligned}$$

Therefore $\{Tx_{k,k}\}$ is Cauchy in H and since H is a Hilbert space $\{Tx_{k,k}\}$ converges in H which is a subsequence of $\{Tx_n\}$ where $\{x_n\}$ is bounded. Hence, T is compact. □

Theorem 4.3.9. *Let H be a Hilbert space and $T \in BL(H)$ be compact. then T^* is compact.*

Proof. Suppose $\{x_n\}$ is a bounded sequence in H. Then there exists $\alpha > 0$ such that $\|x_n\| \le \alpha$ for all n. Since, $T^* \in BL(H)$, i.e. since T^* is bounded, $\{T^* x_n\}$ is a bounded sequence in H.

Let $y_n = T^* x_n$ for $n = 1, 2, \ldots$. Since T is compact, $\{Ty_n\}$ has a convergent subsequence, say $\{Ty_{n_j}\}$. Then for $j, k \in \mathbb{N}$,

$$\begin{aligned}
\|T^* x_{n_j} - T^* x_{n_k}\|^2 &= \langle T^* x_{n_j} - T^* x_{n_k}, T^* x_{n_j} - T^* x_{n_k} \rangle \\
&= \langle y_{n_j} - y_{n_k}, T^*(x_{n_j} - x_{n_k}) \rangle \\
&= \langle T(y_{n_j} - y_{n_k}), x_{n_j} - x_{n_k} \rangle \\
&= \langle Ty_{n_j} - Ty_{n_k}, x_{n_j} - x_{n_k} \rangle \\
&\le \|Ty_{n_j} - Ty_{n_k}\|\|x_{n_j} - x_{n_k}\|
\end{aligned}$$

$$\leq 2\alpha \|Ty_{n_j} - Ty_{n_k}\|$$
$$\to 0 \text{ as } j,k \to \infty.$$

Therefore, $\{T^*x_{n_j}\}$ is a Cauchy sequence in H and since H is complete, $\{T^*x_{n_j}\}$ converges in H. Hence, T^* is compact. □

Proposition 4.3.10. *Let H be a Hilbert space. Then*

1. *If S, T are compact on H, then $S+T$ is compact and αS is compact $\forall\ \alpha \in K$.*
2. *If S is compact on H and $T \in BL(H)$, then ST and TS are compact.*

Proof.

1. Let $\{x_n\}$ be a bounded sequence in H. Since S is compact $\{Sx_n\}$ has a convergent subsequence $\{Sx_{n_j}\}$. Since $\{x_{n_j}\}$ is bounded and T is compact, $\{Tx_{n_j}\}$ has a convergent subsequence $\{Tx_{n_{j_k}}\}$.
   Now, since $\{Sx_{n_j}\}$ converges and $\{Sx_{n_{j_k}}\}$ is a subsequence of $\{Sx_{n_j}\}$, then $\{Sx_{n_{j_k}}\}$ is convergent. Therefore $\{(S+T)x_{n_{j_k}}\}$ is convergent. Hence, $S+T$ is compact.
   Let $\{x_n\}$ be a bounded sequence in H. Since S is compact $\{Sx_n\}$ has a convergent subsequence $\{Sx_{n_j}\}$. Therefore $\{\alpha Sx_{n_j}\}$ is convergent for all $\alpha \in K$. Hence, αS is compact
2. Let $\{x_n\}$ be a bounded sequence in H. Since T is bounded $\{Tx_n\}$ is bounded. Since S is compact, $\{S(Tx_n)\} = \{(ST)x_n\}$ has a convergent subsequence. Therefore, ST is compact.
   Next we show that TS is compact. Suppose $\{x_n\}$ is a bounded sequence in H. Since S is compact, $\{Sx_n\}$ has a convergent subsequence $\{Sx_{n_j}\}$. As $T \in BL(H)$, i.e. T is continuous linear functional, $\{TSx_{n_j}\}$ is convergent. Therefore TS is compact.

□

Remark 4.3.11. Let H be a Hilbert space and $\mathcal{K}(H)$ be the set of all compact operators on H. Then $\mathcal{K}(H)$ is a closed two-sided ideal in $BL(H)$.

Remark 4.3.12. Let H be a Hilbert space and $T \in BL(H)$ be compact. Then T^{-1} is bounded, i.e. $T^{-1} \in BL(H)$ (T is invertible) if and only if H is finite dimensional.

Indeed this is true because T is compact and $T^{-1} \in BL(H)$ implies that $I = TT^{-1}$ is compact. Therefore, Hilbert space H must be finite dimensional.

4.3.1 Hilbert-Schmidt Operators

Definition 4.3.13. Let H be a separable Hilbert space. $T \in BL(H)$ is said to be *Hilbert-Schmidt operator* if

$$\sum_{n=1}^{\infty} \|Tu_n\|^2 < \infty,$$

where $\{u_1, u_2, \ldots\}$ is an orthonormal basis of H.

Theorem 4.3.14. *Let H be a separable Hilbert space and T be a Hilbert-Schmidt operator on H. Then*

1. *T is compact.*

2. T^* is Hilbert-Schmidt.

Proof. 1. Let $\{u_1, u_2, \ldots\}$ be orthonormal basis of H ($\because H$ is separable) such that

$$\sum_{n=1}^{\infty} \|Tu_n\|^2 < \infty.$$

Since $\{u_1, u_2, \ldots\}$ is orthonormal basis of H, each $x \in H$ has a Fourier expansion written as follows:

$$x = \sum_{n=1}^{\infty} \langle x, u_n \rangle u_n.$$

Therefore,

$$Tx = \sum_{n=1}^{\infty} \langle x, u_n \rangle T u_n.$$

Now, for $m = 1, 2, \ldots$ define

$$T_m(x) = \sum_{n=1}^{m} \langle x, u_n \rangle T u_n.$$

Then $\dim R(T_m) \leq m$, i.e. T_m is a finite rank operator for all m. Therefore T_m is compact. Now, for each $x \in H$,

$$\begin{aligned}
\|(T - T_m)x\|^2 &= \|Tx - T_m x\|^2 \\
&= \left\| \sum_{n=m+1}^{\infty} \langle x, u_n \rangle u_n \right\|^2 \\
&\leq \left(\sum_{n=m+1}^{\infty} |\langle x, u_n \rangle| \|Tu_n\| \right)^2 \\
&\leq \sum_{n=m+1}^{\infty} |\langle x, u_n \rangle|^2 \sum_{n=m+1}^{\infty} \|Tu_n\|^2 && \text{(by Holder's inequality)} \\
&\leq \|x\|^2 \sum_{n=m+1}^{\infty} \|Tu_n\|^2 && \text{(by Bessel's inequality).}
\end{aligned}$$

Therefore

$$\|T - T_m\| \leq \sum_{n=m+1}^{\infty} \|Tu_n\|^2 \to 0 \text{ as } n \to \infty.$$

Hence, T is compact.

2. Suppose $\{u_1, u_2, \ldots\}$ is orthonormal basis of H such that

$$\sum_{n=1}^{\infty} \|Tu_n\|^2 < \infty.$$

Now,

$$\begin{aligned}
\sum_{n=1}^{\infty} \|T^* u_n\|^2 &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |\langle T^* u_n, u_m \rangle|^2 && \text{(by Parseval's identity)} \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\langle u_n, T u_m \rangle|^2
\end{aligned}$$

$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\langle Tu_m, u_n \rangle|^2$$
$$= \sum_{m=1}^{\infty} \|Tu_m\|^2 < \infty \qquad \text{(by Parseval's identity)}.$$

Therefore T^* is Hilbert-Schmidt.

□

Remark 4.3.15. Unlike the set of all compact operators, in general, the set of Hilbert-Schmidt operators is not closed in $BL(H)$, i.e. if $\{T_n\}$ is a sequence of Hilbert-Schmidt operators such that $T_n \to T$, then T need not be Hilbert-Schmidt.

Next we show that there is no significance of the chosen orthonormal basis $\{u_1, u_2, \ldots\}$ in the definition of Hilbert-Schmidt operator. In other words, the condition $\sum_{n=1}^{\infty} \|Tu_n\|^2 < \infty$ is independent of the choice of the orthonormal basis $\{u_1, u_2, \ldots\}$.

Proposition 4.3.16. *Let H be a separable Hilbert space and $T \in BL(H)$ be Hilbert-Schmidt. Suppose $\{u_1, u_2, \ldots\}$ and $\{v_1, v_2, \ldots\}$ be two orthonormal bases of H. Then*

$$\sum_{n=1}^{\infty} \|Tu_n\|^2 = \sum_{n=1}^{\infty} \|Tv_n\|^2.$$

Proof.

$$\sum_{n=1}^{\infty} \|Tu_n\|^2 = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |\langle Tu_n, v_m \rangle|^2 \qquad \text{(Parseval's identity)}$$
$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\langle u_n, T^* v_m \rangle|^2$$
$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\langle T^* v_m, u_n \rangle|^2$$
$$= \sum_{m=1}^{\infty} \|T^* v_m\|^2 \qquad \text{(Parseval's identity)}$$
$$= \sum_{m=1}^{\infty} \|Tv_m\|^2 \qquad \text{(by above)}.$$

□

Exercise 4.3.17. The set of all Hilbert-Schmidt operators on H is a linear space (vector space).

Exercise 4.3.18. Let $C_2(H)$ denote the set of all Hilbert-Schmidt operators on H. For $T \in C_2(H)$ define

$$\|T\|_2 = \left(\sum_{n=1}^{\infty} \|Tu_n\|^2 \right)^{\frac{1}{2}},$$

where $\{u_1, u_2, \ldots\}$ is orthonormal basis of H. Then $\| \cdot \|_2$ is a norm on $C_2(H)$.

By above proposition, it is clear that $\|T\|_2$ is invariant of the choice of orthonormal basis of H. Now, we give an example of a Hilbert-Schmidt operator.

Example 4.3.19. Suppose (α_{ij}) is an infinite matrix such that

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|\alpha_{ij}|^2 < \infty.$$

Let T be a operator on ℓ^2 defined by (α_{ij}), i.e. $Te_j = \sum\limits_{i=1}^{\infty}\alpha_{ij}e_j$. Then

$$\|Te_j\|^2 = \left\|\sum_{i=1}^{\infty}\alpha_{ij}e_j\right\|^2 = \sum_{i=1}^{\infty}|\alpha_{ij}|^2.$$

Therefore,

$$\sum_{j=1}^{\infty}\|Te_j\|^2 = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty}|\alpha_{ij}|^2 < \infty.$$

Hence, T is a Hilbert-Schmidt operator on ℓ^2.

Theorem 4.3.20. *Let H be Hilbert space and $T \in BL(H)$ be compact. Then*

$$\sigma_a(T) \setminus \{0\} = \sigma_e(T) \setminus \{0\},$$

i.e. non-zero approximate eigenvalue of T is eigenvalue of T.

If $0 \neq \lambda \in \sigma_e(T)$, then the corresponding eigenspace $\ker(T-\lambda I)$ is finite dimensional.

Proof. Suppose $0 \neq \lambda \in \sigma_a(T)$. Then there exists a sequence $\{x_n\}$ such that $\|x_n\| = 1$ for all n and $(T-\lambda I)x_n \to 0$. Since $\{x_n\}$ is bounded and T is compact, $\{Tx_n\}$ has a convergent subsequence $\{Tx_{n_j}\}$.

Suppose $\lim_{j\to\infty} Tx_{n_j} = x$. Then $\lim_{j\to\infty}\lambda x_{n_j} = x$ ($\because (T-\lambda I)x_n \to 0$). Therefore,

$$\lim_{j\to\infty}\|\lambda x_{n_j}\| = \|x\|.$$

Since $\|x_{n_j}\| = 1$, $|\lambda| = \|x\|$ and since $\lambda \neq 0$, $x \neq 0$. Now,

$$\begin{aligned} Tx &= T\left(\lim_{j\to\infty}\lambda x_{n_j}\right) \\ &= \lambda \lim_{j\to\infty} Tx_{n_j} && (\because T \in BL(H)) \\ &= \lambda x. \end{aligned}$$

Thus, we have $x \neq 0$ such that $(T-\lambda I)x = 0$. Hence, $\lambda \in \sigma_e(T)$.

Next, suppose for $0 \neq \lambda \in \sigma_e(T)$, the corresponding eigen space $\ker(T-\lambda I)$ is not finite dimensional. Then by Gram-Schmidt orthonormalization, we have an infinite orthonormal subset $\{u_1, u_2, \ldots\}$ of $\ker(T-\lambda I)$. Therefore

$$Tu_n = \lambda u_n \qquad \forall\ n.$$

Therefore for $n \neq m$,

$$\|Tu_n - Tu_m\|^2 = |\lambda|^2\|u_n - u_m\|^2 = 2|\lambda|^2.$$

Thus, $\{u_n\}$ is a bounded sequence for which $\{Tu_n\}$ does not have a convergent subsequence. This is contradiction since T is compact. Therefore

$$\dim\ker(T-\lambda I) < \infty.$$

□

Proposition 4.3.21. *Let H be a Hilbert space and $T \in BL(H)$ be compact self-adjoint. Then $\|T\|$ or $-\|T\|$ is eigenvalue of T.*

Proof. Since T is self-adjoint, $m_T, M_T \in \sigma_a(T)$ and (by Corollary 4.2.9)

$$\|T\| = \max\{|m_T|, |M_T|\}.$$

Now, if $M_T + m_T \geq 0$, then $M_T \geq 0$ and $M_T \geq |m_T|$. Therefore

$$\|T\| = M_T \in \sigma_a(T).$$

If $M_T + m_T < 0$, then $|m_T| > |M_T|$ and so $|m_T| = \|T\|$. Therefore

$$-\|T\| = m_T \in \sigma_a(T).$$

As T is compact, $\|T\|$ or $-\|T\| \in \sigma_a(T) \setminus \{0\} = \sigma_e(T) \setminus \{0\}$. Therefore, $\|T\|$ or $-\|T\|$ is eigenvalue of T. □

Now, we state the following result (without proof) about the spectrum of a compact self-adjoint operator which is known as *Spectral theorem for compact self-adjoint operators.*

Theorem 4.3.22 (Spectral theorem for compact self-adjoint operators)**.** *Let H be a Hilbert space and $T \in BL(H)$ be a non-zero compact self-adjoint operator. Then there exists a finite or infinite sequence $\{s_n\}$ of real numbers with $|s_1| \geq |s_2| \geq \cdots$ and orthonormal subset $\{u_1, u_2, \ldots\}$ of H such that and*

$$Tx = \sum_{n=1}^{\infty} s_n \langle x, u_n \rangle u_n. \tag{4.2}$$

Further, if the set $\{u_n\}$ is infinite, then $s_n \to 0$ as $n \to \infty$.

Corollary 4.3.23. *Let T be a non-zero self-adjoint Hilbert-Schmidt operator on H. If $\{s_n\}$ is a sequence of non-zero eigenvalues of T as given in the above theorem, i.e. $|s_1| \geq |s_2| \geq \cdots$, then*

$$\sum_{n=1}^{\infty} |s_n|^2 < \infty.$$

Proof. Since T is Hilbert-Schmidt operator, it is compact. Let

$$T(x) = \sum_{n=1}^{\infty} s_n \langle x, u_n \rangle u_n, \qquad x \in H,$$

as in above theorem. Then $T(u_n) = s_n u_n$ for $n = 1, 2, \ldots$. Since T is Hilbert-Schmidt operator, we have

$$\sum_{n=1}^{\infty} |s_n|^2 = \sum_{n=1}^{\infty} \|T(u_n)\|^2 < \infty.$$

□

Index

R

S

U

W