LOVELY PROFESSIONAL UNIVERSITY

FUNCTIONAL ANALYSIS

Edited by
Dr. Sachin Kaushal

CONTENT

Unit 1: Banach Space: Definition and Some Examples

CONTENTS

Objectives

After studying this unit, you will be able to:

- Know about Banach spaces.
- Define Banach spaces.
- Solve problems on Banach spaces.

Introduction

Banach space is a linear space, which is also, in a special way, a complete metric space. This combination of algebraic and metric structures opens up the possibility of studying linear transformations of one Banach space into another which have the additional property of being continuous. The concept of a Banach space is a generalization of Hilbert space. A Banach space assumes that there is a norm on the space relative to which the space is complete, but it is not assumed that the norm is defined in terms of an inner product. There are many examples of Banach spaces that are not Hilbert spaces, so that the generalization is quite useful.

1.1 Banach Spaces

1.1.1 Normed Linear Space

Definition: Let N be a complex (or real) linear space. A real valued function $n : N \to R$ is said to define, a norm on N if for any $x, y \in N$ and any scalar (complex number) α, the following conditions are satisfied by n:

(i) $n(x) \geq 0$, $n(x) = 0, \Leftrightarrow x = 0$;

(ii) $n(x + y) \leq n(x) + n(y)$; and

(iii) $n(\alpha x) = |\alpha|\, n(x)$

It is customary to denote n (x) by

$n(x) = \|x\|$ (read as norm x)

With this notation the above conditions (i) – (iii) assume the following forms:

(i) $\|x\| \geq 0$, $\|x\| = 0 \Leftrightarrow x = 0$;

(ii) $\|x + y\| \leq \|x\| + \|y\|$; and

(iii) $\|\alpha x\| = |\alpha|\, \|x\|$.

A linear space N together with a norm defined on it, i.e., the pair $(N, \| \ \|)$ is called a *normed linear space* and will simply be denoted by N for convenience.

Notes

1. The condition (ii) is called subadditivity and the condition (iii) is called absolute homogeneity.
2. If we drop the condition viz. $\|x\| = 0 \Leftrightarrow x = 0$, then $\| \ \|$ is called a semi norm (or pseudo norm) or N and the space N is called a semi-normed linear space.

Theorem 1: If N is a normed linear space and if we define a real valued function $d : N \times N \to R$ by $d(x, y) = \|x - y\|$ $(x, y \in N)$, then d is a metric on N.

Proof: We shall verify the conditions of a metric

(i) $d(x, y) \geq 0$, $d(x, y) = 0 \Leftrightarrow \|x - y\| = 0 \Leftrightarrow x = y$;

(ii) $d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1|\, \|y - x\| = \|y - x\| = d(y, x)$;

(iii) $d(x, y) = \|x - y\| = \|x - z + z - y\|$ $\quad (z = N)$

$$\leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$$

Hence, d defines a metric on N. Consequently, every normed linear space is automatically a metric space.

This completes the proof of the theorem.

Notes

1. The above metric has the following additional properties:

   (i) If $x, y, z \in N$ and α is a scalar, then

   $d(x + z, y + z) = \|(x + z) - (y + z)\| = \|x - y\| = d(x, y)$.

   (ii) $d(\alpha x, \alpha y) = \|\alpha x - \alpha y\| = \|\alpha(x - y)\|$

   $= |\alpha|\, \|x - y\| = |\alpha|\, d(x, y)$.

2. Since every normed linear space is a metric space, we can rephrase the definition of convergence of sequences by using this metric induced by the norm.

Notes

1.1.2 Convergent Sequence in Normed Linear Space

Definition: Let (N, || ||) be a normed linear space. A sequence (x_n) is N is said to converge to an element x in N if given $\varepsilon > 0$, there exists a positive integer n_o such that

$$\| x_n - x \| < \varepsilon \text{ for all } n \geq n_o.$$

If x_n converges to x, we write $\lim_{n \to \infty} x_n = x$.

or $x_n \to x$ as $n \to \infty$

It follows from the definition that

$$x_n \to x \Leftrightarrow \| x_n - x \| \to 0 \text{ as } n \to \infty$$

Theorem 2: If N is a normed linear space, then

$$\left| \| x \| - \| y \| \right| \leq \| x - y \| \text{ for any } x, y \in N$$

Proof: We have

$$\| x \| = \| (x - y) + y \|$$

$$\leq \| x - y \| + \| y \|$$

$$\Rightarrow \quad \| x \| - \| y \| \leq \| x - y \| \qquad \ldots (1)$$

Using (1), we have

$$-(\| x \| - \| y \|) = \| y \| - \| x \| \leq \| y - x \|$$

But $\| y - x \| = \| (-1)(x - y) \| = |-1| \, \| x - y \|$

Therefore

$$-(\| x \| - \| y \|) \leq \| x - y \| \text{ so that}$$

$$\| x \| - \| y \| \geq - \| x - y \| \qquad \ldots (2)$$

From (1) and (2) we get

$$\left| \| x \| - \| y \| \right| \leq \| x - y \|$$

This completes the proof of the theorem.

1.1.3 Subspace of a Normed Linear Space

Definition: A subspace M of a normed linear space is a subspace of N consider as a vector space with the norm obtain by restricting the norm of N to the subset M. This norm on M is said to be induced by the norm on N. If M is closed in N, then M is called a closed subspace of N.

Theorem 3: Let N be a normed linear space and M is a subspace of N. Then the closure $\overline{M}$ of M is also a subspace of N.

(Note that since $\overline{M}$ is closed, $\overline{M}$ is a closed subspace).

Proof: To prove that $\overline{M}$ is a subspace of N, we must show that any linear combination of element in $\overline{M}$ is again in M. That is if x and y $\in \overline{M}$, then $\alpha x + \beta y \in \overline{M}$ for any scalars α and β.

Since x, y ∈ $\overline{M}$, there exist sequences (x_n) and (y_n) in M such that

$$x_n \to x \text{ and } y_n \to y,$$

By joint continuity of addition and scalar multiplication in M.

$\alpha x_n + \beta y_n \to \alpha x + \beta y$ for every scalars α and β.

Since $\alpha x_n + \beta y_n \in M$, we conclude that

$\alpha x + \beta y \in \overline{M}$ and consequently $\overline{M}$ is a subspace of N.

This completes the proof of the theorem.

Notes

1. The scalars α, β can be assumed to be non-zero.

   For if α = 0 = β, then

   $\alpha x + \beta y = 0 \in M \subset \overline{M}$

2. In a normed linear space, the smallest closed subspace containing a given set of vectors S is just the closure of the subspace spanned by the set S. To see this, let S be the subset of a normed linear space N and let M be the smallest closed subspace of N, containing S. We show that $M = [\overline{S}]$, where [S] is the subspace spanned by S.

   By theorem, $[\overline{S}]$ is a closed subspace of N and it contains S.

   Since M is the smallest closed subspace containing S, we have

   $M \subset [\overline{S}]$.

   But $[S] \subset M$ and $M = \overline{M}$, we must have

   $[\overline{S}] \subset \overline{M} = M$ so that $[\overline{S}] \subset M$.

   Hence $[\overline{S}] = M$.

1.1.4 Complete Normed Linear Space

Definition: A normed linear space N is said to be complete if every Cauchy sequence in N converges to an element of N. This means that if $\| x_m - x_n \| \to 0$ as $m, n \to \infty$, then there exists x ∈ N such that

$$\| x_n - x \| \to 0 \text{ as } n \to \infty.$$

1.1.5 Banach Space

Definition: A complete normed linear space is called a Banach space.

OR

A normed linear space which is complete as a metric space is called a Banach space.

In the definition of a Banach space completeness means that if

$\| x_m - x_n \| \to 0$ as m, n $\to \infty$, where $(x_n) \subset N$, then

$\exists$ a $x \in N$ such that

$\| x_n - x \| \to 0$ as $n \to \infty$.

Note A subspace M of a Banach space B is a subspace of B considered as a normed linear space. We do not require M to be complete.

Theorem 4: Every complete subspace M of a normed linear space N is closed.

Proof: Let $x \in N$ be any limit point of M.

We have to show that $x \in M$.

Since x is a limit point of M, there exists a sequence (x_n) in M and $x_n \to x$ as $n \to \infty$.

But, since (x_n) is a convergent sequence in M, it is Cauchy sequence in M.

Further M is complete $\Rightarrow$ (x_n) converges to a point of M so that $x \in M$.

Hence M is closed.

This completes the proof of the theorem.

Theorem 5: A subspace M of a Banach space B is complete iff the set M is closed in B.

Proof: Let M be a complete subspace of a Banach space M. They be above theorem, M is closed (prove it).

Conversely, let M be a closed subspace of Banach space B. We shall show that M is complete.

Let $x = (x_n)$ be a Cauchy sequence in M. Then

$x_n \to x$ in B as B is complete.

We show that $x \in M$.

Now $x \in \overline{M} \Rightarrow x \in M$ $\quad (\because$ M being closed $\Rightarrow M = \overline{M})$

Thus every Cauchy sequence in M converges to an element of M. Hence the closed sequence M of B is complete. This completes the proof of the theorem.

Example 1: The linear space R of real numbers or C of complex numbers are Banach spaces under the norm defined by

$$\| x \| = |x|, x \in R \text{ (or C)}$$

Solution: We have

$$\| x \| = |x| > 0 \text{ and } \| x \| = 0 \Leftrightarrow |x| = \Leftrightarrow x = 0$$

Further, let $z_1, z_2 \in C$ and let $\bar{z}_1$ and $\bar{z}_2$ be their complex conjugates, then

$$|z_1 + z_2|^2 = (z_1 + z_2)(\bar{z}_1 + \bar{z}_2)$$

$$= z_1\bar{z}_1 + z_1\bar{z}_2 + z_2\bar{z}_1 + z_2\bar{z}_2$$

$$\leq |z_1|^2 + 2|z_1\bar{z}_2| + |z_2|^2$$

Notes

$$= |z_1|^2 + 2|z_1||z_2| + |z_2|^2 \qquad \left[\because\ |z_1, \bar{z}_2| + |z_1, z_2| = |z_1||z_2|\right]$$

$$= (|z_1| + |z_2|)^2$$

$\therefore \qquad |z_1 + z_2| \le |z_1| + |z_2|$

or $\qquad \|z_1 + z_2\| \le \|z_1\| + \|z_2\| \qquad (\because \|x\| = |x|)$

Also $\qquad \|\alpha x\| = |\alpha x| = |\alpha|\ |x| = |\alpha|\ \|x\|$

Hence all the conditions of normed linear space are satisfied. Thus both C or R are normed linear space. And by Cauchy general principle of convergence, R and C are complete under the matrices induced by the norm. So R and C are Banach spaces.

Example 2: ***Euclidean and Unitary spaces:*** The linear space R^n and C^n of all n-tuples $(x_1, x_2 \ldots, x_n)$ of real and complex numbers are Banach spaces under the norm

$$\|x\| = \left\{\sum_{i=1}^{n} |x_i|^2\right\}^{1/2}$$

[Usually called Euclidean and unitary spaces respectively].

Solution: (i) Since each $|x_i| \ge 0$, we have

$$\|x\| \ge 0$$

and $\|x\| = 0 \Leftrightarrow \sum_{i=1}^{n} |x_i|^2 = 0 \Leftrightarrow x_i = 0,\ i = 1, 2, \ldots, n$

$$\Leftrightarrow (x_1, x_2, \ldots x_n) = 0$$

$$\Leftrightarrow x = 0$$

(ii) Let $x = (x_1, x_2, \ldots, x_n)$

and $y = (y_1, y_2, \ldots y_n)$ be any two numbers of C^n (or R^n). Then

$$\|x + y\|^2 = \|(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots y_n)\|^2$$

$$= \|(x_1 + x_1), (x_2 + y_2), \ldots, (x_n + y_n)\|^2$$

$$= \sum_{i=1}^{n} |x_i + y_i|^2$$

$$\le \sum_{i=1}^{n} |x_i + y_i|(|x_i| + |y_i|)$$

$$\le \sum_{i=1}^{n} |x_i + y_i||x_i| + \sum_{i=1}^{n} |x_i + y_i||y_i|$$

Usually Cauchy inequality for each sum, we get

$$\|x + y\|^2 = \left\{\sum_{i=1}^{n} |x_i + y_i|\right\}^2 \left\{\sum_{i=1}^{n} |x_i|^2\right\}^{1/2} + \left\{\sum_{i=1}^{n} |x_i + y_i|^2\right\}^{1/2} \left\{\sum_{i=1}^{n} |y_i|^2\right\}^{1/2}$$

$$= \| x + y \| \ \| x \| + \| x + y \| \ \| y \|$$

$$= (\| x \| + \| y \|)\ (\| x + y \|).$$

If $\| x + y \| = 0$, then the above inequality is evidently true.

If $\| x + y \| \neq 0$, we can divide both sides by it to obtain

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

(iii)
$$\| \alpha x \| = \left\{ \sum_{i=1}^{n} | \alpha x_i |^2 \right\}^{1/2} = | \alpha | \left\{ \sum_{i=1}^{n} | x_i |^2 \right\}^{1/2}$$

$$= | \alpha | \ \| x \|.$$

This proves that R^n or C^n are normed linear spaces.

Now we show the completeness of C^n (or R^n).

Let $< x_1, x_2, \ldots x_n >$ be a Cauchy sequence in C^n (or R^n). Since each x_m is an n-tuple of complex (or real) numbers, we shall write

$$x_m = \left(x_1^{(m)}, x_2^{(m)}, \ldots\ldots, x_n^{(m)} \right)$$

So that $x_k^{(m)}$ is the k^{th} coordinate of x_m.

Let $\varepsilon > 0$ be given, since $<x_m>$ is a Cauchy sequence, there exists a positive integer m_o, such that

$$\ell, m \geq m_o \Rightarrow \left\| x_m - x_\ell \right\| < \varepsilon$$

$$\Rightarrow \left\| x_m - x_\ell \right\|^2 < \varepsilon^2$$

$$\Rightarrow \sum_{i=1}^{n} \left| x_i^{(m)} - x_i^{(\ell)} \right| < \varepsilon^2 \qquad \ldots (1)$$

$$\Rightarrow \left| x_i^{(m)} - x_i^{(\ell)} \right| < \varepsilon^2 \ \ (i = 1, 2, \ldots\ldots, n)$$

$$\Rightarrow \left| x_i^{(m)} - x_i^{(\ell)} \right| < \varepsilon$$

Hence $\left\langle x_i^{(m)} \right\rangle_{m=1}^{\infty}$ is a Cauchy sequence of complex (or real) numbers for each fixed but arbitrary i.

Since C (or R) is complete, each of these sequences converges to a point, say 2_i in C (or R) so that

$$\lim_{m \to \infty} x_i^{(m)} = z_i \ (i = 1, 2, \ldots, n) \qquad \ldots (2)$$

Now we show that the Cauchy sequence $<x_m>$ converges to the point $z = (z_1, z_2, \ldots\ldots, z_n) \in C^n$ (or R^n).

To prove this let $\ell \to \infty$ in (1). Then by (2) we have

$$\sum_{i=1}^{n} \left| x_i^{(m)} - z_i \right|^2 < \varepsilon^2$$

$\Rightarrow \quad \| x_m - z \|^2 < \varepsilon^2$

$\Rightarrow \quad \| x_m - z \| < \varepsilon$

It follows that the Cauchy sequence $<x_m>$ converges to $z \in C^n$ (or R^n).

Hence C^n or R^n are complete spaces and consequently they are Banach spaces.

1.2 Summary

- A linear space N together with a norm defined on it, i.e. the pair (N, $\| \ \|$) is called a normed linear space.
- Let (N, $\| \ \|$) be a normed linear space. A sequence (x_n) in N is said to converge to an element x in N if given $\varepsilon > 0$, there exists a positive integer n_o such that

  $$\| x_n - x \| < \varepsilon \quad \text{for all } n \geq n_o.$$
- If N is a normed linear space, then

  $\big| \| x \| - \| y \| \big| \leq \| x - y \|$ for any x, y $\in$ N.
- A normed linear space N is said to be complete if every Cauchy sequence in N converges to an element of N.
- A complete normed linear space is called a Banach space.

1.3 Keywords

A Subspace M of a Normed Linear Space: A subspace M of a normed linear space is a subspace of N consider as a vector space with the norm obtain by restricting the norm of N to the subset M. If norm on M is said to be induced by the norm on N. If M is closed in N, then M is called a closed subspace of N.

Banach Space: A complete normed linear space is called a Banach space.

Complete Normed Linear Space: A normed linear space N is said to be complete if every Cauchy sequence in N converges to an element of N. This means that if $\| x_m - x_n \| \to 0$ as m, n $\to \infty$, then there exists x $\in$ N such that

$$\| x_n - x \| \to 0 \text{ as } n \to \infty.$$

Normed Linear: A linear space N together with a norm defined on it, i.e., the pair (N, $\| \ \|$) is called a *normed linear space* and will simply be denoted by N for convenience.

1.4 Review Questions

1. Let N be a non-zero normed linear space, prove that N is a Banach space $\Leftrightarrow$ $\{x : \| x \| = 1\}$ is complete.
2. Let a Banach space B be the direct sum of the linear subspaces M and N, so that B = M $\oplus$ N. If z = x + y is the unique expression of a vector z in B as the sum of vectors x and y in M and N, then a new norm can be defined on the linear space B by $\| z \|' = \| x \| + \| y \|$.

   Prove that this actually is a norm. If B′ symbolizes the linear space B equipped with this new norm, prove that B′ is a Banach space of M and N are closed in B.

Notes

1.5 Further Readings

Books

Bourbaki, Nicolas (1987), *Topological Vector Spaces, Elements of Mathematics*, Berlin: Springer-Verlag.

Beauzamy, Bernard (1985), *Introduction to Banach Spaces and their Geometry* (Second revised ed.), North-Holland.

Online links

mathword.wolfram.com?Calculus and Analysis>Functional Analysis

homepage.ntlworld-com/ivan.wilde/notes/fal/fal.pdf

www.math.ucdavis.edu/~ hunter/book.chs.pdf

Unit 2: Continuous Linear Transformations

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand continuous linear transformation
- Define bounded linear functional and norm of a bounded linear functional
- Understand theorems on continuous linear transformations.

Introduction

In this unit, we obtain the representation of continuous linear functionals on some of Banach spaces.

2.1 Continuous Linear Transformation

2.1.1 Continuous Linear Functionals Definition

- Let N be a normed linear space. Then we know the set R of real numbers and the set C of complex numbers are Banach spaces with the norm of any $x \in R$ or $x \in C$ given by the absolute value of x. Thus with our previous notations, β (N, R) or β (N, C) denote respectively the set of all *continuous linear transformations* from N into R or C.
- We denote the Banach space β (N, R) or β (N, C) by N* and call it by the conjugate space (or dual space or adjoint space) of N.
- The elements of N* will be referred to as continuous linear functionals or simply functionals on N.

> *Note* The conjugate space $(N^*)^*$ of N^* is called the second conjugate space of N and shall be denoted by N^{**}. Also note that N^{**} is complete too.

Theorem 1: The conjugate space N^* is always a Banach space under the norm

$$\|f\| = \sup\left\{\frac{|f(x)|}{\|x\|} : x \in N, x \neq 0\right\} \quad \ldots \text{(i)}$$

$$= \sup\left\{|f(x)| : \|x\| < 1\right\}$$

$$= \inf\left\{k, k \geq 0 \text{ and } |f(x)| \leq k\|x\| \forall x\right\}$$

Proof: As we know that if N, N′ are normed linear spaces, β (N, N′) is a normed linear space. If N′ is a Banach space, β (N, N′) is Banach space. Hence β (N, R) or β (N, C) is a Banach space because R and C are Banach spaces even if N is not complete.

This completes the proof of the theorem.

Theorem 2: Let f be a linear functional on a normed linear space. If f is continuous at $x_o \in N$, it must be continuous at every point of N.

Proof: If f is continuous at $x = x_o$, then

$$x_n \to x_o \Rightarrow f(x_n) \to f(x)$$

To show that f is continuous everywhere on N, we must show that for any $y \in N$,

$$y_n \to y \Rightarrow f(y_n) \to f(y)$$

Let $y_n \to y$ as $n \to \infty$

Now $\quad f(y_n) = f(y_n - y + x_o + y - x_o)$

since f is linear.

$\therefore \quad f(y_n) = f(y_n - y + x_o) + f(y) - f(x_o) \quad \ldots (1)$

As $\quad y_n \to y \Rightarrow y_n - y + x_o \to x_o$ by hypothesis

Also f is continuous, $\quad f(y_n - y + x_o) \to f(x_o) \quad \ldots (2)$

From (1) and (2), it follows that

$$f(y_n) \to f(y) \text{ as } n \to \infty.$$

⇒ f is continuous at $y \in N$ and consequently as it is continuous everywhere on N.

Hence proved.

2.1.2 Bounded Linear Functional

A linear functional on a normed linear space N is said to be bounded, if there exists a constant k such that

$$|f(x)| \leq K\|x\| \quad \forall x \in N \cdot$$

Note

We may find many K's satisfying the above condition for a given bounded function. If it is satisfied for one K, it is satisfied for a $K_1 > K$.

Theorem 3: Let f be a linear functional defined on a normed linear space N, then f is bounded $\Leftrightarrow$ f is continuous.

Proof: Let us first show that continuity of f $\Rightarrow$ boundedness of f.

If possible let f is continuous but not bounded. Therefore, for any natural number n, however large, there is some point x_n such that

$$|f(x_n)| \geq n \, \|x_n\| \quad \ldots (1)$$

Consider the vector, $y_n = \dfrac{x_n}{n\|x_n\|}$ so that

$$\|y_n\| = \frac{1}{n}.$$

$\Rightarrow \quad \|y_n\| \to 0$ as $n \to \infty$

$\Rightarrow \quad y_n \to 0$ in the norm.

Since any continuous functional maps zero vector into zero and f is continuous $f(y_n) \to f(0) = 0$.

But $$|f(y_n)| = \frac{1}{n\|x_n\|} f(x_n) \quad \ldots (2)$$

It now follows from (1) & (2) that $|f(y_n)| > 1$, a contradiction to the fact that $f(y_n) \to 0$ as $n \to \infty$.

Thus if f is bounded, then f is continuous.

Conversely, let f is bounded. Then for any sequence (x_n), we have

$$|f(x_n)| \leq K \, \|x_n\| \qquad \forall\, n = 1, 2, \ldots, \text{ and } K \geq 0.$$

Let $x_n \to 0$ as $n \to \infty$ then

$f(x_n) \to 0 \Rightarrow$ f is continuous at the origin and consequently it is continuous everywhere.

This completes the proof of the theorem.

Note The set of all bounded linear function on N is a vector space denoted by N*. As in the case of linear operators, we make it a normed linear space by suitably defining a norm of a functional f.

2.1.3 Norm of a Bounded Linear Functional

If f is a bounded linear functional on a normed space N, then the norm of f is defined as:

$$\|f\| = \sup_{\|x\| \neq 0} \frac{|f(x)|}{\|x\|} \quad \ldots (1)$$

Notes

We first note that the above norm is well defined. Since f is bounded, we have

$$|f(x)| \leq M \, \|x\|, \quad M \geq 0.$$

Let M′ be the set of real numbers M satisfying this relation. Then the set $\left\{\frac{|f(x)|}{\|x\|}; x \neq 0\right\}$ is bounded above so that it must possess a supremum. Let it be $\|f\|$. So $\|f\|$ is well defined and we must have

$$\frac{|f(x)|}{\|x\|} \leq \|f\| \quad \forall \; x \neq 0.$$

or $$|f(x)| \leq \|f\| \; \|x\|.$$

Let us check that $\| \;\; \|$ defined by (1) is truly a norm on N*:

If f, g ∈ N*, then

$$\|f+g\| = \sup_{\|x\| \neq 0} \frac{|f(x)+g(x)|}{\|x\|}$$

$$\leq \sup_{\|x\| \neq 0} \frac{|f(x)|}{\|x\|} + \sup_{\|x\| \neq 0} \frac{|g(x)|}{\|x\|}$$

⇒ $$\|f+g\| \leq \|f\| + \|g\|.$$

Similarly, we can see that $\|\alpha f\| = |\alpha||f|$.

16.1.4 Equivalent Methods of Finding || F ||

If f is a bounded linear functional on N, then

$$|f(x)| \leq M\|x\|, M \geq 0.$$

(I) $\|f\|$ = inf {M : M ∈ M′} where M′ is the set of all real numbers satisfying

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

Since $\|f\|$ ∈ M′ and M′ is the set of all non-negative real numbers, it is bounded below by zero so that it has an infimum. Hence

$$\|f\| \geq \inf \{M : M \in M'\} \qquad \ldots (2)$$

For x ≠ 0 and M ∈ M′ we have $\frac{f(x)}{\|x\|} \leq M$. Since M is the only upper bound then from definition (2), we have

$$M \geq \sup_{\|x\| \neq 0} \frac{|f(x)|}{\|x\|} = \|f\| \text{ for any } M \in M'.$$

Notes

Since M′ is bounded below by $\|f\|$, it has an infimum so that we have

$$\inf_{M \in M''} M = \inf \{M : M \in M'\} \geq \|f\| \qquad \ldots (3)$$

From (2) and (3), it follows that

$$\|f\| = \inf \{M : M \in M'\}$$

(II) $$\|f\| = \sup_{\|x\| \neq 0} |f(x)|$$

Let us consider $\|x\| \leq 1$. Then

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

Therefore, we have

$$\sup_{\|x\| \leq |} |f(x)| \leq \|f\|. \qquad \ldots (4)$$

Now by definition,

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

It follows from the property of the supremum that, given $\in > 0$, $\exists$ an $x' \in N$ such that

$$\frac{|f(x')|}{\|x'\|} > (\|f\| - \in) \qquad \ldots (5)$$

Define

$\bar{x} = \dfrac{x'}{\|x\|}$. Then $\bar{x}$ is a unit vector.

Since $\{\|\bar{x}\| = 1\} \subset \{\|\bar{x}\| \leq 1\}$, we have

$$\sup_{\|x\| \leq 1} |f(x)| \geq |f(\bar{x})| = \frac{1}{\|x'\|} |f(x')| > (\|f\| - \in) \qquad \text{[by (2)]}$$

Hence $\in > 0$ is arbitrary, we have

$$\sup_{\|x\| \leq 1} |f(x)| > \|f\| \qquad \ldots (6)$$

From (4) and (6), we obtain

$$\sup_{\|x\| \leq 1} |f(x)| = \|f\|.$$

(III) $$\|f\| = \sup_{\|x\| = 1} |f(x)|.$$

Consider $\|x\| = 1$, we have

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

So that

$$\sup_{\|x\|=1} |f(x)| \le \|f\| \qquad \dots (7)$$

Now consider

$$\|f\| = \sup_{\|x\| \ne 0} \frac{|f(x)|}{\|x\|}$$

By supremum property, given $\in > 0$, $\exists\, x' \ne 0$

Such that $|f(x')| > (\|f\| - \in)\, \|x'\|$

Define $\bar{x} = \dfrac{x'}{\|x\|}$.

Since f is continuous in $\|x\| \le 1$ and reaches its maximum on the boundary $\|x\| = 1$,

We get

$$\sup_{\|x\|=1} |f(x)| \ge f(\bar{x}) = \frac{1}{\|x'\|} f(x') > \|f\| - \in .$$

$$\Rightarrow \sup_{\|x\|=1} |f(x)| > \|f\| - \in .$$

The arbitrary character of $\in$ yields that

$$\sup_{\|x\|=1} |f(x)| \ge \|f\| \qquad \dots (8)$$

Hence from (7) and (8), we get

$$\|f\| = \sup_{\|x\|=1} |f(x)|.$$

> *Note* If N is a finite dimensional normed linear space, all linear functions are bounded and hence continuous. For, let N be of dimension n so that any $x \in N$ is of the form $\sum_{i=1}^{n} \alpha_i x_i$, where $x_1, x_2, \dots, x_n$ is a basis of N and $\alpha_1, \alpha_2, \dots, \alpha_n$ are scalars uniquely determined by the basis.

Since f is linear, we have

$$f(x) = \sum_{i=1}^{n} \alpha_i\, f(x_i) \text{ so that}$$

$$|f(x)| \le \sum_{i=1}^{n} |\alpha_i|\, |f(x_i)| \qquad \dots (1)$$

We have from (1) by using the notation of the Zeroth norm in a finite dimensional space,

$$|\, f(x) \,| \;\leq\; \| x \|_0 \sum_{i=1}^{n} |f(x_i)| \qquad \dots (2)$$

If $\sum_{i=1}^{n} |f(x_i)|$ = M, then from (2), we have

$$|\, f(x) \,| \;\leq M \| x \|_0 .$$

Hence f is bounded with respect to $\| \;\|_0$.

Since any norm $\| \;\|$ on N is equivalent to $\| \;\|_0$, f is bounded with respect to any norm on N. Consequently, f is continuous on N.

16.1.5 Representation Theorems for Functionals

We shall prove, in this section, the representation theorems for functionals on some concrete Banach spaces.

Theorem 4: If L is a linear space of all n-tuples, then (i) $\left(\ell_p^n\right)^* = \ell_q^n$.

Proof: Let $(e_1, e_2, \dots, e_n)$ be a standard basis for L so that any $x = (x_1, x_2, \dots, x_n) \in$ L can be written as

$$x \;= x_1e_1 + x_2e_2 + \dots + x_ne_n.$$

If f is a scalar valued linear function defined on L, then we get

$$f(x) \;= x_1 f(e_1) + x_2 f(e_2) + \dots + x_n f(e_n) \qquad \dots (1)$$

⇒ f determines and is determined by n scalars

$$y_i \;= f(e_i).$$

Then the mapping

$$y \;= (y_1, y_2, \dots, y_n) \to f$$

where $f(x) = \sum_{i=1}^{n} x_i y_i$ is an isomorphism of L onto the linear space L′ of all function f. We shall establish (i) – (iii) by using above given facts.

(i) If we consider the space

L = ℓ_p^n $(1 \leq p < \infty)$ with the p^{th} norm, then f is continuous and L′ represents the set of all continuous linear functionals on ℓ_p^n so that

$$L' \;= \left(\ell_p^n\right)^* .$$

Now for y → f as an isometric isomorphism we try to find the norm for y's.

For $1 < p < \infty$, we show that

$$\left(\ell_p^n\right)^* = \ell_q^n .$$

Notes

For $x \in \ell_p^n$, we have defined

$$\|x\| = \left\{\sum_{i=1}^{n} |x_i|^p\right\}^{\frac{1}{p}}$$

Now $|f(x)| = \left|\sum_{i=1}^{n} x_i y_i\right| \leq \sum_{i=1}^{n} |x_i||y_i|$

By using Hölder's inequality, we get

$$\sum_{i=1}^{n} x_i y_i \leq \left\{\sum_{i=1}^{n} |x_i|^p\right\}^{\frac{1}{p}} \left\{\sum_{i=1}^{n} |y_i|^q\right\}^{\frac{1}{q}}$$

so that

$$|f(x)| \leq \left\{\sum_{i=1}^{n} |y_i|^q\right\}^{\frac{1}{q}} \left\{\sum_{i=1}^{n} |x_i|^p\right\}^{\frac{1}{p}}$$

Using the definition of norm for f, we get

$$\|f\| \leq \left\{\sum_{i=1}^{n} |y_i|^q\right\}^{\frac{1}{q}} \qquad \ldots (2)$$

Consider the vector, defined by

$$x_i = \frac{|y_i|^q}{y_i},\ y_i \neq 0 \text{ and } x_i = 0 \text{ if } y_i = 0 \qquad \ldots (3)$$

Then

$$\|x\| = \left\{\sum_{i=1}^{n} |x_i|^p\right\}^{\frac{1}{p}} = \left[\left\{\sum_{i=1}^{n} \frac{|y_i|^q}{|y_i|}\right\}^p\right]^{\frac{1}{p}} \qquad \ldots (4)$$

Since q = p (q – 1) we have from (4),

$$\|x\| = \left\{\sum_{i=1}^{n} |y_i|^q\right\}^{\frac{1}{p}} \qquad \ldots (5)$$

Now

$$|f(x)| = \left|\sum_{i=1}^{n} x_i y_i\right| = \left|\sum_{i=1}^{n} \frac{|y_i|^q}{y_i} y_i\right|$$

$$= \sum_{i=1}^{n} |y_i|^q, \qquad \text{(By (3))}$$

So that

Notes

$$\sum_{i=1}^{n}|y_i|^q = |f(x)| \leq \|f\|\|x\| \qquad \ldots (6)$$

From (5) and (6), we get

$$\left\{\sum_{i=1}^{n}|y_i|^q\right\}^{1-\frac{1}{p}} \leq \|f\|$$

$$\Rightarrow \qquad \left\{\sum_{i=1}^{n}|y_i|^q\right\}^{\frac{1}{q}} \leq \|f\| \qquad \ldots (7)$$

Also from (2) and (7), we have

$$\|f\| = \left\{\sum_{i=1}^{n}|y_i|^q\right\}^{\frac{1}{q}}, \text{ so that}$$

$y \to f$ is an isometric isomorphism.

Hence $\boxed{\left(\ell_p^n\right)^* = \ell_q^n}$.

(ii) Let L = ℓ_1^n with the norm defined by $\|x\| = \sum_{i=1}^{n}|x_i|$.

Now f defined in (1), above is continuous as in (i) and L′ here represents the set of continuous linear functional on ℓ_1^n so that

$$L' = \left(\ell_1^n\right)^*.$$

We now determine the norm of y's which makes $y \to f$ an isometric isomorphism.

Now,

$$|f(x)| = \left|\sum_{i=1}^{n} x_i y_i\right|$$

$$\leq \sum_{i=1}^{n}|x_i||y_i|$$

But $\sum_{i=1}^{n}|x_i||y_i| \leq \max.\{|y_i|\}\sum_{i=1}^{n}|x_i|$ so that $|f(x)| \leq \max.\{|y_i|\}\sum_{i=1}^{n}|x_i|$.

From the definition of norm for f, we have

$$\|f\| = \max.\{|y_i| : i = 1, 2, \ldots, n\} \qquad \ldots (8)$$

Now consider the vector defined as follows:

If $|y_i| = \max_{1 \leq i \leq n}\{|y_i|\}$, let us consider vector x as

Notes

$$x_i = \frac{|y_i|}{y_i} \text{ when } |y_i| = \max_{1 \le i \le n}\{|y_i|\}$$

and $x_i = 0$ otherwise.

From the definition, $x_k = 0 \ \forall \ k \neq i$. So that we have

$$\|x\| = \left|\frac{y_i}{y}\right| = 1$$

$$\text{Further } |f(x)| = \left|\sum_{i=1}^{n}(x_i y_i)\right| = |y_i|$$

$$\text{Hence } |y_i| = |f(x)| \le \|f\| \, \|x\|$$

$\Rightarrow$ $$|y_i| \le \|f\| \text{ or } \max.\{|y_i|\} \qquad [\because \ ||x|| = 1]$$

$$\le ||f|| \qquad \ldots (10)$$

From (8) and (10), we obtain

$$||f|| = \max.\{|y_i|\} \text{ so that}$$

$y \to f$ is an isometric isomorphism of L′ to $\left(\ell_1^n\right)^*$.

Hence $\boxed{\left(\ell_1^n\right)^* = \ell_\infty^n}$.

(iii) Let L = ℓ_∞^n with the norm

$\|x\| = \max\{|x_i| : i = 1, 2, 3, \ldots, n\}$.

Now f defined in (1) above is continuous as in (1).

Let L′ represents the set of all continuous linear functionals on ℓ_∞^n so that

$$L' = \left(\ell_\infty^n\right)^*.$$

Now we determine the norm of y's which makes $y \to f$ as isometric isomorphism

$$|f(x)| = \left|\sum_{i=1}^{n} x_i y_i\right| \le \sum_{i=1}^{n}|x_i||y_i|.$$

But $\sum_{i=1}^{n}|x_i||y_i| \le \max(|x_i|)\sum_{i=1}^{n}|y_i|$

Hence we have

$$|f(x)| \le \left\{\sum_{i=1}^{n}|y_i|\right\}(\|x\|) \text{ so that}$$

$$(\|f\|) \le \sum_{i=1}^{n}|y_i| \qquad \ldots (11)$$

Consider the vector x defined by

$$x_i = \frac{|y_i|}{y_i} \text{ when } y_i \neq 0 \text{ and } x_i = 0 \text{ otherwise.} \qquad \dots (12)$$

Hence $\| x \| = \max\left\{\frac{|y_i|}{|y_i|}\right\} = 1$.

and $| f(x) | = \left|\sum_{i=1}^{n} |x_i||y_i|\right| = \sum_{i=1}^{n} |y_i|$.

Therefore $\sum_{i=1}^{n} |y_i| = |f(x)| \leq \| f \| \, \| x \| = \| f \|$.

$$\Rightarrow \quad \sum_{i=1}^{n} |y_i| \leq \| f \| \qquad \dots (13)$$

It follows now from (11) and (13) that $\| f \| = \sum_{i=1}^{n} |y_i|$ so that $y \to f$ is an isometric isomorphism.

Hence, $(\ell_\infty^n)^* = \ell_1^n$.

This completes the proof of the theorem.

Theorem 5: The conjugate space of ℓ_p is ℓ_q, where

$$\frac{1}{p} + \frac{1}{q} = 1 \text{ and } 1 < p < \infty.$$

or $\qquad \ell_p^* = \ell_q$.

Proof: Let $x = (x_n) \in \ell_p$ so that $\sum_{n=1}^{\infty} |x_n|^p < \infty$. $\qquad \dots (1)$

Let ℓ_n = (0, 0, 0, …, 1, 0, 0, …) where 1 is in the m[th] place.

$e_n \in \ell_p$ for n = 1, 2, 3, …

We shall first determine the form of f and then establish the isometric isomorphism of ℓ_p^* onto ℓ_q.

By using (e_n), we can write any sequence $(x_1, x_2, \dots, x_n, 0, 0, 0, \dots)$ in the form $\sum_{k=1}^{n} x_k e_k$ and

$$x - \sum_{k=1}^{n} x_k e_k = (0, 0, 0, \dots, x_{n+1}, x_{n+2}, \dots).$$

Now $\left\| x - \sum_{k=1}^{n} x_k e_k \right\| = \left\{ \sum_{k=n+1}^{\infty} |x_k|^p \right\}^{\frac{1}{p}}$ … (2)

The R.H.S. of (2) gives the remainder after n terms of a convergent series (1).

Hence $\left\{ \sum_{k=n+1}^{\infty} |x_k|^p \right\}^{\frac{1}{p}} \to 0$ as $n \to \infty$ … (3)

From (2) and (3), it follows that

$$x = \sum_{k=1}^{\infty} x_k e_k . \quad \text{… (4)}$$

Let $f \in \ell_p^*$ and $s_n = \sum_{k=1}^{n} x_k e_k$ then

$s_n \to x$ as $n \to \infty$. (Using (4))

since f is linear, we have

$$f(s_n) = \sum_{k=1}^{n} x_k f(e_k) .$$

Also f is continuous and $s_n \to x$, we have

$$f(s_n) \to f(x) \quad \text{as } n \to \infty$$

$$\Rightarrow \quad f(x) = \sum_{k=1}^{n} x_k f(e_k) \quad \text{… (5)}$$

which gives the form of the functional on ℓ_p.

Now we establish the isometric isomorphism of ℓ_p^* onto ℓ_q, for which we proceed as follows:

Let $f(e_k) = \alpha_k$ and show that the mapping

$$T : \ell_p^* \to \ell_q \text{ given by} \quad \text{… (6)}$$

$T(f) = (\alpha_1, \alpha_2, \ldots, \alpha_k, \ldots)$ is an isometric isomorphism of ℓ_p^* onto ℓ_q.

First, we show that T is well defined.

For let $x \in \ell_p$, where $x = (\beta_1, \beta_2, \ldots, \beta_n, 0, 0, \ldots)$

where $\beta_k = \begin{cases} |\alpha_k|^{g-1} \operatorname{sgn} \bar{\alpha}_k, & 1 \le k \le n \\ 0 & \forall\, n > k \end{cases}$

$$\Rightarrow \quad |\beta_k| = |\alpha_k|^{q-1} \text{ for } 1 \le k \le n.$$

$$\Rightarrow \quad |\beta_k|^p = |\alpha_k|^{(q-1)p} = |\alpha_k|^q. \qquad \left(\because \frac{1}{p} + \frac{1}{q} = q \Rightarrow p(q-1) = q \right)$$

Notes

$$\text{Now } \alpha_k\ \beta_k = \alpha_k |\alpha_k|^{q-1} \operatorname{sgn} \bar{\alpha}_k = |\alpha_k|^{q-1} \alpha_k \operatorname{sgn} \bar{\alpha}_k$$

$$\Rightarrow \quad \alpha_k\ \beta_k = |\alpha_k|^q = |\beta_k|^p \quad \text{(Using property of sgn function)} \quad \dots (7)$$

$$\Rightarrow \quad \|x\| = \left\{\sum_{k=1}^{n} |\beta_k|^p\right\}^{\frac{1}{p}}$$

$$= \left\{\sum_{k=1}^{n} |\alpha_k|^q\right\}^{\frac{1}{q}} \quad \dots (8)$$

Since we can write

$$x = \sum_{k=1}^{n} \beta_k e_k \text{, we get}$$

$$f(x) = \sum_{k=1}^{n} \beta_k\ f(e_k) = \sum_{k=1}^{n} \alpha_k\ \beta_k$$

$$\Rightarrow \quad f(x) = \sum_{k=1}^{n} |\alpha_k|^q \quad \text{(Using (7))} \quad \dots (9)$$

We know that for every $x \in \ell_p$

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

which upon using (8) and (9), gives

$$|f(x)| \le \sum_{k=1}^{n} |\alpha_k|^q \le \|f\| \left\{\sum_{k=1}^{n} |\alpha_k|^q\right\}^{\frac{1}{p}}$$

which yields after simplification.

$$\left\{\sum_{k=1}^{n} |\alpha_k|^q\right\}^{\frac{1}{p}} \le \|f\| \quad \dots (10)$$

since the sequence of partial sums on the L.H.S. of (10) is bounded, monotonic increasing, it converges. Hence

$$\left\{\sum_{k=1}^{n} |\alpha_k|^q\right\}^{\frac{1}{q}} \le \|f\| \quad \dots (11)$$

so the sequence (α_k) which is the image of f under T belongs to ℓ_q and hence T is well defined.

We next show that T is onto ℓ_q.

Let $(\beta_k) \in \ell_q$, we shall show that there is a $g \in \ell_p^*$ such that T maps g into (β_k).

Notes

Let $x \in \ell_p$ so that

$$x = \sum_{k=1}^{\infty} x_k e_k .$$

We shall show that

$$g(x) = \sum_{k=1}^{\infty} x_k \beta_k \text{ is the required g.}$$

Since the representation for x is unique, g is well defined and moreover it is linear on ℓ_p. To prove it is bounded, consider

$$|g(x)| = \left|\sum_{k=1}^{\infty} \beta_k x_k\right| \leq \sum_{k=1}^{n} |\beta_k x_k|$$

$$\leq \left\{\sum_{k=1}^{\infty} |x_k|^p\right\}^{\frac{1}{p}} \left\{\sum_{k=1}^{\infty} |\beta_k|^q\right\}^{\frac{1}{q}} \qquad \text{(Using Hölder's inequality)}$$

$\Rightarrow$ $$|g(x)| \leq \|x\| \left\{\sum_{k=1}^{\infty} |\beta_k|^q\right\}^{\frac{1}{q}} .$$

$\Rightarrow$ g is bounded linear functional on ℓ_p.

since $e_k \in \ell_p$ for k = 1, 2, …, we get

$g(e_k) = \beta_k$ for any k so that

$T_g = (\beta_k)$ and T is on ℓ_p^* onto ℓ_p.

We next show that

$\|Tf\| = \|f\|$ so that T is an isometry.

Since $Tf \in \ell_q$, we have from (6) and (10) that

$$\left\{\sum_{k=1}^{\infty} |\alpha_k|^q\right\}^{\frac{1}{q}} = \|Tf\| \leq \|f\| \qquad \ldots (12)$$

Also, $x \in \ell_p \Rightarrow x = \sum_{k=1}^{\infty} x_k e_k$. Hence

$$f(x) = \sum_{k=1}^{\infty} x_k (e_k) = \sum_{k=1}^{\infty} x_k \alpha_k .$$

$\Rightarrow$ $$|f(x)| \leq \sum_{k=1}^{\infty} |x_k||\alpha_k|$$

$$\le \left\{\sum_{k=1}^{\infty}|\alpha_k|^q\right\}^{\frac{1}{q}} \left\{\sum_{k=1}^{\infty}|x_k|^p\right\}^{\frac{1}{p}} \qquad \text{(Using Hölder's inequality)}$$

or $$|f(x)| \le \left\{\sum_{k=1}^{\infty}|\alpha_k|^q\right\}^{\frac{1}{q}} \|x\| \quad \forall x \in \ell_p.$$

Hence, we have

$$\sup_{x\neq 0}\left\{\frac{|f(x)|}{\|x\|}\right\} \le \left\{\sum_{k=1}^{\infty}|\alpha_k|^q\right\}^{\frac{1}{q}} = \|Tf\| \qquad \text{(Using (6))}$$

which upon using definition of norm yields.

$$\|f\| \le \|Tf\| \qquad \dots (13)$$

Thus $$\|f\| = \|Tf\| \qquad \text{(Using (12) and (13))}$$

From the definition of T, it is linear. Also since it is an isometry, it is one-to-one and onto.

Hence T is an isometric isomorphism of ℓ_p^* onto ℓ_p, i.e.

$$\boxed{\ell_p^* = \ell_q}$$

Theorem 6: Let N and N′ be normed linear and let T be a linear transformation of N into N′. Then the inverse T^{-1} exists and is continuous on its domain of definition if and only if there exists a constant $m > 0$ such that

$$m\|x\| \le \|T(x)\| \;\forall\; x \in N. \qquad \dots (1)$$

Proof: Let (1) holds. To show that T^{-1} exists and is continuous.

Now T^{-1} exists iff T is one-one.

Let $x_1, x_2, \in N$. Then

$$T(x_1) = T(x_2) \Rightarrow T(x_1) - T(x_2) = 0$$

$$\Rightarrow T(x_1 - x_2) = 0$$

$$\Rightarrow \quad x_1 - x_2 = 0 \text{ by (1)}$$

$$\Rightarrow \quad x_1 = x_2$$

Hence T is one-one and so T^{-1} exists. Therefore to each y in the domain of T^{-1}, there exists x in N such that

$$T(x) = y \Leftrightarrow T^{-1}(y) = x \qquad \dots (2)$$

Hence (1) is equivalent to

$$m\|T^{-1}(y)\| \le \|y\| \Rightarrow \|T^{-1}(y)\| \le \frac{1}{m}\|y\|$$

$\Rightarrow$ T^{-1} is bounded $\Rightarrow$ T^{-1} is continuous converse.

Let T^{-1} exists and be continuous on its domain T(N). Let x be an arbitrary element in N. Since T^{-1} exists, there is $y \in T(N)$ such that $T^{-1}(y) = x \Leftrightarrow T(x) = y$.

Notes

Again since T^{-1} is continuous, it is bounded so that there exists a positive constant k such that

$$\| T^{-1} (y) \| \leq k \| y \| \Rightarrow \| x \| \leq k \| T(x) \|$$

$\Rightarrow m \| x \| \leq \| T(x) \|$ where $m = \frac{1}{k} > 0$.

This completes the proof of the theorem.

Theorem 7: Let $T : N \to N'$ be a linear transformation. Then T is bounded if and only if T maps bounded sets in N onto bounded set in N′.

Proof: Since T is a bounded linear transformation,

$$\| T (x) \| \leq k \| x \| \text{ for all } x \in N.$$

Let B be a bounded subset of N. Then

$$\| x \| \leq k_1 k \ \forall \ x \in B.$$

We now show that T(B) is bounded subset of N′.

From above we see that

$$\| T (x) \| \leq k_1 \ \forall \ x \in B.$$

$\Rightarrow$ T (B) is bounded in N′.

Conversely, let T map bounded sets in N into bounded sets in N′. To prove that T is a bounded linear transformation, let us take the closed unit sphere S [0, 1] in N as a bounded set. By hypothesis, its image T (S[1, 0]) must be bounded set in N′.

Therefore there is a constant k_1 such that

$$\| T (x) \| \leq k_1 \text{ for all } x \in S [0, 1]$$

Let x be any non-zero vector in N. Then $\left(\frac{x}{\|x\|}\right) \in S[0,1]$ and so we get

$$\left\| T\left(\frac{x}{\|x\|}\right) \right\| \leq k_1$$

$$\Rightarrow \qquad \| T (x) \| \leq k_1 \| x \|.$$

Since this is true for x = 0 also, T is a bounded linear transformation.

This completes the proof of the theorem.

2.2 Summary

- Let N be a normed linear space. Then we know the set R of real numbers and the set C of complex numbers are Banach spaces with the norm of any $x \in R$ or $x \in C$ be the absolute value of X. β (N, R) or β (N, C) denote respectively the set of all ***continuous linear transformations*** from N into R or C.
- A linear functional on a normed linear space N is said to be bounded, if there exists a constant k such that

$$| f (x) | \leq k \| x \| \ \forall \ x \in N.$$

- If f is a bounded linear functional on a normed space N, then the norm of f is defined as:

$$\| f \| = \sup_{\|x\| \neq 0} \left\{ \frac{|f(x)|}{\|x\|} \right\}$$

2.3 Keywords

Bounded Linear Functional: A linear functional on a normed linear space N is said to be bounded, if there exists a constant k such that

$$|f(x)| \leq K \|x\| \quad \forall x \in N \cdot$$

Continuous Linear Transformations: Let N be a normed linear space. Then we know the set R of real numbers and the set C of complex numbers are Banach spaces with the norm of any $x \in R$ or $x \in C$ given by the absolute value of x. Thus with our previous notations, β (N, R) or β (N, C) denote respectively the set of all continuous linear transformations from N into R or C.

Norm of a Bounded Linear Functional: If f is a bounded linear functional on a normed space N, then the norm of f is defined as:

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

Second Conjugate: The conjugate space (N*)* of N* is called the second conjugate space of N .

2.4 Review Questions

1. Prove that the conjugate space of ℓ_1 is ℓ_∞,

   i.e. $\ell_1^* = \ell_\infty$.

2. Prove that the conjugate space of c_o is ℓ_1.

   or $c_o^* = \ell_1$

3. Let p > 1 with $\frac{1}{p} + \frac{1}{q} = 1$ and let $g \in L_q$ (X).

   Then prove that the function defined by

   $$F(f) = \int_X fg \, d\mu \text{ for } f \in Lp(X)$$

   is a bounded linear functional on Lp (X) and

   $$\| F \| = \| g \|_q$$

4. Let N be any n dimensional normed linear space with a basis $B = \{x_1, x_2, ..., x_n\}$. If $(r_1, r_2, ..., r_n)$ is any ordered set of scalars, then prove that, there exists a unique continuous linear functional f on N such that

   $$f(x_i) = r_i \text{ for } i = 1, 2, \ldots, n$$

5. If T is a continuous linear transformation of a normed linear space N into a normed linear space N′, and if M is its null space, then show that T induces a natural linear transformation T′ of N/M into N′ and that || T′ || = || T ||.

2.5 Further Readings

Books

JB Conway (1990), *A Course in Functional Analysis*.

E Hille (1957), *Functional Analysis and Semigroups*.

Online links

pt.scribd.com/doc/86559155/14/Continuous-Linear-Transformations

www.math.psu.edu/bressan/PSPDF/fabook.pdf

Notes

Unit 3: The Hahn-Banach Theorem

CONTENTS

Objectives

After studying this unit, you will be able to:

- State the Hahn-Banach theorem
- Understand the proof of the Hahn-Banach theorem
- Solve problems related to it.

Introduction

The Hahn-Banach theorem is one of the most fundamental and important theorems in functional analysis. It is most fundamental in the sense that it asserts the existence of the linear, continuous and norm preserving extension of a functional defined on a linear subspace of a normed linear space and guarantees the existence of non-trivial continuous linear functionals on normed linear spaces. Although there are many forms of Hahn-Banach theorem, however we are interested in Banach space theory, in which we shall first prove Hahn-Banach theorem for normed linear spaces and then prove the generalised form of this theorem. In the next unit, we shall discuss some important applications of this theorem.

3.1 The Hahn-Banach Theorem

3.1.1 Theorem: The Hahn-Banach Theorem – Proof

Let N be a normed linear space and M be a linear subspace of N. If f is a linear functional defined on M, then f can be extended to a functional f_o defined on the whole space N such that

$$\| f_o \| = \| f \|.$$

Proof: We first prove the following lemma which constitutes the most difficult part of this theorem.

Notes

Lemma: Let M be a linear subspace of a normed linear space N let f be a functional defined on M. If $x_o \in N$ such that $x_o \notin M$ and if $M_o = M + [x_o]$ is the linear subspace of N spanned by M and x_o, then f can be extended to a functional f_o defined on M_o s.t.

$$\| f_o \| = \| f \|.$$

Proof: We first prove the following lemma which constitutes the most difficult part of this theorem.

Lemma: Let M be a linear subspace of a normed linear space N let f be a functional defined on M. If $x_o \in N$ such that $x_o \notin M$ and if $M_o = M + [x_o]$ is the linear subspace of N spanned by M and x_o, then f can be extended to a functional f_o defined on M_o s.t.

$$\| f_o \| = \| f \|.$$

Proof: The lemma is obvious if f = o. Let then $f \neq 0$.

Case I: Let N be a real normed linear space.

Since $x_o \notin M$, each vector y in M_o is uniquely represented as

$$y = x + \alpha x_o, x \in M \text{ and } \alpha \in R.$$

This enables us to define

$$f_o : M_o \to R \text{ by}$$

$$f_o (y) = f_o (x + \alpha x_o) = f(x) + \alpha r_o,$$

where r_o is any given real number ... (1)

We show that for every choice of the real number r_o, f_o is not only linear on M but it also extends f from M to M_o and

$$\| f_o \| = \| f \|.$$

Let $x_1, y_1 \in M_o$. Then these exists x and $y \in M$ and real scalars α and β such that

$$x_1 = x + \alpha x_o \text{ and } y_1 = y + \beta x_o,$$

Hence,
$$f_o (x_1 + y_1) = f_o (x + \alpha x_o + y + \beta x_o)$$
$$= f_o (x + y + (\alpha + \beta) x_o)$$
$$= f (x + y) + (\alpha + \beta) r_o, r_o \text{ is a real scalar} \quad \dots (2)$$

$$\text{Since f is linear M, } f (x + y) = f (x + y) \quad \dots (3)$$

From (2) and (3) it follows after simplification that

$$f_o (x_1 + y_1) = f (x) + \alpha r_o + f (y) + \beta r_o$$
$$= f_o (x + \alpha x_o) + f_o (y + \beta x_o)$$
$$= f_o (x_1) + f_o (y_1)$$

$$\Rightarrow \quad f_o (x_1 + y_1) = f_o (x_1) + f_o (y_1) \quad \dots (4)$$

Let k be any scalar. Then if $y \in M_o$, we have

$$f_o (ky) = f_o [k (x + \alpha x_o)]$$
$$= f_o (k_x + k \alpha x_o)$$

But
$$f_o (k_x + k\alpha x_o) = f (kx) + k\alpha r_o$$
$$= k f (x) + k \alpha r_o$$

Notes

Hence $$f_o(ky) = k[f(x) + \alpha r_o] = k f_o(y) \quad \dots (5)$$

From (4) and (5) it follows that f_o is linear on M_o.

If $y \in M$, then $\alpha = 0$ in the representation for y so that

$$y = x.$$

Hence $$f_o(x) = f(x) \ \forall \ x \in M.$$

$\Rightarrow$ f_o extends f from M to M_o.

Next we show that

$$\| f_o \| = \| f \|.$$

If $\alpha = 0$ this is obvious. So we consider when $\alpha \neq 0$. Since M is a subspace of M_o we then have

$$\begin{aligned} \| f_o \| &= \sup \{ | f_o(x) | : x \in M_o, \| x \| \leq 1 \} \\ &\geq \sup \{ | f_o(x) | : x \in M, \| x \| \leq 1 \} \\ &= \sup \{ | f(x) | : x \in M, \| x \| \leq 1 \} \qquad (\because f_o = f \text{ on } M) \\ &= \| f \|. \end{aligned}$$

Thus, $$\| f_o \| \geq \| f \| \quad \dots (A)$$

So our problem now is to choose r_o such that $\| f_o \| \leq \| f \|$.

Let $x_1, x_2 \in M$. Then we have

$$\begin{aligned} f(x_2) - f(x_1) &= f(x_2 - x_1) \\ &\leq | f(x_2 - x_1) | \\ &\leq \| f \| \, \| (x_2 + x_o) - (x_1 + x_o) \| \\ &\leq \| f \| (\| x_2 + x_o \| + \| -(x_1 + x_o) \|) \\ &= \| f \| \, \| x_2 + x_o \| + \| f \| \, \| x_1 + x_o \| \end{aligned}$$

Thus $$-f(x_1) - \| f \| \, \| x_1 + x_o \| \leq -f(x_2) + \| f \| \, \| x_2 + x_o \| \quad \dots (6)$$

Since this inequality holds for arbitrary $x_1, x_2 \in M$, we see that

$$\sup_{y \in M} \{ -f(y) - \| f \| (\| y + x_o \|) \} \leq \operatorname{int}_{y \in M} \{ -f(y) + \| f \| (\| y + x_o \|) \}$$

Choose r_o to be any real number such that

$$\sup_{y \in M} \{ -f(y) - \| f \| (\| y + x_o \|) \} \leq r_o$$

$$\leq \inf_{y \in M} \{ -f(y) + \| f \| (\| y + x_o \|) \}$$

From this, we get for all $y \in M$

$$\begin{aligned} \sup \{ -f(y) - \| f \| (\| y + x_o \|) \} &\leq r_o \\ &\leq \inf \{ -f(y) + \| f \| (\| y + x_o \|) \} \end{aligned}$$

Let us take $y = \dfrac{x}{\alpha}$ in the above inequality, we have

$$\sup_{y \in M} \left\{ -f\left(\frac{x}{\alpha}\right) - \| f \| \left(\left\| \frac{x}{\alpha} + x_o \right\| \right) \right\} \leq r_o$$

$$\leq \inf_{y \in M} \left\{ -f\left(\frac{x}{\alpha}\right) + \| f \| \left(\left\| \frac{x}{\alpha} + x_o \right\| \right) \right\} \qquad \ldots (7)$$

If $\alpha > 0$ the right hand side of (7) becomes

$$r_o \leq \frac{-1}{\alpha} f(x) + \frac{1}{\alpha} \| f \| \| x + \alpha x_o \| \text{ which implies that}$$

$$f(x) + \alpha r_o = f_o(x + \alpha x_o) \leq \| f \| \| x + \alpha x_o \|$$

If $z = x + \alpha x_o \in M_o$ then we get from above

$$| f_o(z) | \leq \| f \| \| z \| \qquad \ldots (8)$$

If $\alpha < 0$, then from L.H.S. of (7) we have

$$-f\left(\frac{x}{\alpha}\right) - \| f \| \left\| \frac{x}{\alpha} + x_o \right\| \leq r_o$$

$$-\frac{1}{\alpha} f(x) + \frac{1}{\alpha} \| f \| \| x + \alpha x_o \| \leq r_o, \text{ since } \alpha < 0, \left| \frac{1}{\alpha} \right| = \frac{-1}{\alpha}.$$

$\Rightarrow$ $$f(x) + \alpha r_o \geq \| f \| \| x + \alpha x_o \|$$

$\Rightarrow$ $$f_o(z) \geq \| f \| \| z \| \text{ for every } z \in M_o \qquad \ldots (9)$$

Replacing z by –z in (9) we get

$$-f_o(z) \leq \| f \| \| z \|, \text{ since } f_o \text{ is linear on } M_o \qquad \ldots (10)$$

Hence we get from (9) and (10).

$$| f_o(z) | \leq \| f \| \| z \| \qquad \ldots (11)$$

Since f is functional on M, || f || is bounded.

Thus (||) shows that f_o is a functional on M_o.

Since $\| f_o \| = \sup \{ | f_o(z) | : z \in M_o, \| z \| \leq | \}$, it follows from (||) that

$$\| f_o \| \leq \| f \|$$

We finally obtain from (A) and (B) that

$$\| f_o \| = \| f \|$$

This power the lemma for real normed linear space.

Case II: Let N be a complex normed linear space.

Let N be a normed linear space over C and f be a complex valued functional on a subspace M of N.

Let g = Re (f) and h = Im (f) so that we can write

f (x) = g (x) + i h (x). We show that g (x) and h (x) are real valued functionals.

Since f is linear, we have

$$f(x + y) = f(x) + f(y)$$

$$\Rightarrow \quad g(x+y) + i\,h(x+y) = g(x) + i\,h(x) + g(y) + i\,h(y)$$
$$= g(x) + g(y) + i\,(h(x) + h(y))$$

Equating the real and imaginary parts, we get

$$g(x+y) = g(x) + g(y)$$

and

$$h(x+y) = h(x) + h(y)$$

If $\alpha \in R$, then we have

$$f(\alpha x) = g(\alpha x) + i\,h(\alpha x)$$

Since f is linear

$$f(\alpha x) = \alpha f(x) = \alpha g(x) + \alpha i\,h(x)$$

$$\Rightarrow \quad f(\alpha x) = \alpha g(x) \text{ and } h(\alpha x) = \alpha h(x)$$

(equating real and imaginary parts)

$\Rightarrow$ g, h are real linear functions on M.

Further $|g(x)| \le |f(x)| \le \|f\|\ \|x\|$

$\Rightarrow$ If f is bounded on M, then g is also bounded on M.

Similarly h is also bounded on M.

Since a complex linear space can be regarded as a real linear space by restricting the scalars to be real numbers, we consider M as a real linear space. Hence g and h are real functional on real space M.

For all x in M we have

$$f(i\,x) = i\,f(x) = i\,\{g(x) + i\,h(x)\}$$

or

$$g(i\,x) + i\,h(i\,x) = -h(x) + i\,g(x)$$

Equating real and imaginary parts, we get

$$g(i\,x) = -h(x) \text{ and } h(i\,x) = g(x)$$

Therefore we can express f (x) either only by g or only h as follows:

$$f(x) = g(x) - i\,g(i\,x)$$
$$= h(i\,x) + i\,h(x).$$

Since g is a real functional on M, by case I, we extend g to a real functional g_o on the real space M_o such that $\|g_o\| = \|g\|$. For $x \in M_o$, we define

$$f_o(x) = g_o(x) - i\,g_o(i\,x)$$

First note that f_o is linear on the complex linear space M_o. Such that $f_o = f$ on M.

Now

$$f_o(x+y) = g_o(x+y) - i\,g_o(i\,x + i\,y)$$
$$= g_o(x) + g_o(y) - i\,g_o(i\,x) - i\,g_o(i\,y)$$
$$= f_o(x) + f_o(y).$$

Now for $a, b \in R$, we have

$$f_o((a + i\,b)\,x) = g_o(ax + i\,bx) - i\,g_o(-bx + i\,ax)$$
$$= a\,g_o(x) + b\,g_o(i\,x) - i\,(-b)\,g_o(x) - i\,a\,g_o(i\,x)$$
$$= (a + ib)\,\{g_o(x) - i\,g_o(i\,x)\}$$

So that

$$f_o((a + i b) x) = (a + i b) f_o(x)$$

$\Rightarrow f_o$ is linear on M_o and also $g_o = g$ on M.

$\Rightarrow f_o = f$ on M.

Now have to show that $\| f_o \| = \| f \|$ on M_o.

Let $x \in M_o$ and $f_o(x) = re^{i\theta}$

$$| f_o(x) | = r = e^{-i\theta} re^{i\theta} = e^{-i\theta} f_o(x) \qquad \dots (12)$$

Since $f_o(x)$ is linear,

$$e^{-i\theta} f_o(x) = f_o(e^{-i\theta} x) \qquad \dots (13)$$

So we get from (12) and (13) that

$$| f_o(x) | = r = f_o(r\, e^{-i\theta}x).$$

Thus the complex valued functional f_o is real and so it has only real part so that

$$| f_o(x) | = g_o(e^{-i\theta}x) \leq | g_o(e^{-i\theta}x) |$$

But $\quad | g_o(e^{-i\theta}x) | \leq \| g_o \| \| e^{-i\theta}x \|$

$$= \| g_o \| \| x \|,$$

We get $\quad | f_o(x) | \leq \| g_o \| \| x \|$

Since g_o is the extension of g, we get

$$\| g_o \| \| x \| = \| g \| \| x \| \leq \| f \| \| x \|.$$

Therefore

$| f_o(x) \leq \| f \| \| x \|$ so that from the definition of the norm of f_o, we have

$$\| f_o \| \leq \| f \|$$

As in case I, it is obvious that $\| f \| \leq \| f_o \|$

Hence $\quad \| f_o \| = \| f \|.$

This completes the proof of the theorem.

3.1.2 Theorems and Solved Examples

Theorem: *The generalized Hahn-Banach Theorem for Complex Linear Space.*

Let L be a complex linear space. Let p be a real valued function defined on L such that

$$p(x + y) \leq p(x) + p(y)$$

and $\quad p(\alpha x) = |\alpha| \, p(x) \;\; \forall \;\; x \in L$ and scalar α.

If f is a complex linear functional defined on the subspace M such that $| f(x) | \leq p(x)$ for $x \in M$, then f can be extended to a complex linear functional to be defined on L such that $| f_o(x) | \leq p(x)$ for every $x \in L$.

Proof: We have from the given hypothesis that f is a complex linear functional on M such that

$$| f(x) \leq p(x) \;\; \forall \;\; x \in M.$$

Let g = Re (f) then $g(x) \leq |f(x)| \leq p(x)$.

So by the generalised Hahn-Banach Theorem for Real Linear space, can be extended to a linear functional g_o on L into R such that $g_o = g$ on M and $g_o(x) \leq p(x)\ \forall\ x \in L$.

Define $f_o(x) = g_o(x) - i\, g_o(i\,x)$ for $x \in L$ as in the Hahn-Banach Theorem, f_o is linear functional on L such that $f_o = f$ on M.

To complete the proof we have to prove that

$$|f_o(x)| \leq p(x)\ \forall\ x \in L.$$

Let $x \in L$ and $f_o(x) = r\,e^{i\theta}$, r > o and θ real. Then

$$|f_o(x)| = r = e^{-\theta}\, re^{i\theta} = e^{-i\theta} f_o(x)$$
$$= f_o(e^{-i\theta} x).$$

Since $r = f_o(e^{-i\theta}x)$, f_o is real so that we can take

$$|f_o(x)| = r = f_o(e^{-i\theta} x) = g_o(e^{-i\theta}x) \quad \dots (1)$$

Since $g_o(x) \leq p(x)$, $g_o(e^{-i\theta}x) \leq p(e^{-i\theta}x)$ for $x \in L$.

But $p(e^{-i\theta}x) = |e^{-i\theta}|\ p(x)$ so that $g_o(e^{-i\theta}x) \leq p(x)$ … (2)

It follows from (1) and (2) that

$$|f_o(x)| \leq p(x)$$

This completes the proof of the theorem.

Corollary 1: Deduce the Hahn-Banach theorem for normed linear spaces from the generalised Hahn-Banach theorem.

Proof: Let $p(x) = \|f\|\,\|x\|$ for $x \in N$.

We first note that $p(x) \geq 0$ for all $x \in N$.

Then for any $x, y \in N$, we have

$$p(x + y) = \|f\|\ \|x + y\|$$
$$\leq \|f\|\,(\|x\| + \|y\|)$$
$$= \|f\|\ \|x\| + \|f\|\ \|y\|$$
$$= p(x) + p(y)$$

$\Rightarrow$ $$p(x + y) \leq p(x) + p(y)$$

Also $$p(\alpha x) = \|f\|\ \|\alpha x\| = |\alpha|\ \|f\|\ \|x\| = |\alpha|\ p(x).$$

Hence p satisfies all the conditions of the generalized Hahn-Banach Theorem for Complex Linear space. Therefore ∃ a functional f_o defined on all of N such that $f_o = f$ on M and $|f_o(x)| \leq p(x) = \|f\|\ \|x\|\ \forall x \in N$.

$\Rightarrow$ $$\|f_o\| \leq \|f\| \quad \dots (3)$$

Since f_o is the extension of f from a subspace M, we get

$$\|f\| \leq \|f_o\| \quad \dots (4)$$

From (3) and (4) it follows that

$$\|f_o\| = \|f\|$$

Notes Let L be a linear space. A mapping $p : L \to R$ is called a sub-linear functional on L if it satisfies the following two properties namely,

(i) $p(x + y) \le p(x) + p(y) \;\; \forall \; x, y \in L$ (sub additivity)

(ii) $p(\alpha x) = \alpha p(x), \alpha \ge 0$ (positive homogeneity)

Thus p defined on L in the above theorems is a sub-linear functional on L.

Some Applications of the Hahn-Banach Theorem

Theorem: If N is a normed linear space and $x_o \in N$, $x_o \ne 0$ then there exists a functional $f_o \in N^*$ such that

$$f_o(x_o) = \| x_o \| \text{ and } \| f_o \| = 1.$$

Proof: Let M denote the subspace of N spanned by x_o, i.e.,

$$M = \{\alpha x_o : \alpha \text{ any scalar}\}.$$

Define $f : M \to F$ (R or C) by

$$f(\alpha x_o) = \alpha \| x_o \|.$$

We show that f is a functional on M with $\| f \| = 1$.

Let $x_1, x_2 \in M$ so that

$x_1 = \alpha_1 x_o$ and $x_2 = \alpha_2 x_o$. Then

$$f(x_1 + x_2) = f(\alpha_1 x_o + \alpha_2 x_o)$$

$$= (\alpha_1 + \alpha_2) \| x_o \|$$

But $$(\alpha_1 + \alpha_2) \| x_o \| = \alpha_1 \| x_o \| + \alpha_2 \| x_o \|$$

$$= f(x_1) + f(x_2)$$

Hence $$f(x_1 + x_2) = f(x_1) + f(x_2) \quad \dots (1)$$

Let k be a scalar (real or complex). Then if $x \in M$, then $x = \alpha x_o$.

Now $f(kx) = f(k \alpha x_o) = k \alpha \| x_o \| = k f(x)$ … (2)

If follows from (1) and (2) that f is linear.

Further, we note that since $x_o \in M$ with $\alpha = 1$, we get

$$f(x_o) = \| x_o \|.$$

For any $x \in M$, we get, $| f(x) | = |\alpha| \, | | x_o | | = \| \alpha x_o \| = \| x \|$

$\Rightarrow$ $$| f(x) | = \| x \|$$

$\Rightarrow$ f is bounded and we have

$$\sup \frac{| f(x) |}{\| x \|} = 1 \text{ for } x \in M \text{ and } x \ne 0.$$

So by definition of norm of a functional, we get

$$\| f \| = 1.$$

Hence by Hahn-Banach theorem, f can be extended to a functional $f_o \in N^*$ such that $f_o(M) = f(M)$ and $\| f_o \| = \| f \| = 1$, which in particular yields that

$$f_o(x_o) = f(x_o) = \| x_o \| \text{ and } \| f_o \| = 1.$$

This completes the proof of the theorem.

Corollary 2: N* separates the vector (points) in N.

Proof: To prove the cor. it suffices to show that if $x, y \in N$ with $x \neq y$, then there exists a $f \in N^*$ such that $f(x) \neq f(y)$.

Since $x \neq y \Rightarrow x - y \neq 0$.

So by above theorem there exists a functional $f \in N^*$ such that

$$f(x - y) = f(x) - f(y) \neq 0$$

and hence $f(x) \neq f(y)$.

This shows that N* separates the point of N.

Corollary 3: If all functional vanish on a given vector, then the vector must be zero, i.e.

if $f(x) = 0 \ \forall \ f \in N^*$ then $x = 0$.

Proof: Let x be the given vector such that $f(x) = 0 \ \forall \ f \in N^*$.

Suppose $x \neq 0$. Then by above theorem, there exists a function $f \in N^*$ such that

$f(x) = \| x \| > 0$, which contradicts our supposition that

$f(x) = 0 \ \forall \ f \in N^*$. Hence we must have $x = 0$.

3.2 Summary

- ***The Hahn-Banach Theorem:*** Let N be a normed linear space and M be a linear subspace of N. If f is a linear functional defined on M, then f can be extended to a functional f_o defined on the whole space N such that

$$\| f_o \| = \| f \|$$

- If f is a complex linear functional defined on the subspace M such that $|f(x)| \leq p(x)$ for $x \in M$, then f can be extended to a complex linear function f_o defined on L such that $|f_o(x)| \leq p(x)$ for every $x \in L$.

3.3 Keywords

Hahn-Banach theorem: The Hahn-Banach theorem is one of the most fundamental and important theorems in functional analysis. It is most fundamental in the sense that it asserts the existence of the linear, continuous and norm preserving extension of a functional defined on a linear subspace of a normed linear space and guarantees the existence of non-trivial continuous linear functionals on normed linear spaces.

Sub-linear Functional on L: Let L be a linear space. A mapping $p : L \to R$ is called a sub-linear functional on L if it satisfies the following two properties namely,

(i) $p(x + y) \leq p(x) + p(y) \ \forall \ x, y \in L$ (sub additivity)

(ii) $p(\alpha x) = \alpha p(x), \alpha \geq 0$ (positive homogeneity)

Thus p defined on L in the above theorems is a sub-linear functional on L.

The Generalized Hahn-Banach Theorem for Complex Linear Space: Let L be a complex linear space. Let p be a real valued function defined on L such that

$$p(x + y) \leq p(x) + p(y)$$

and $p(\alpha x) = |\alpha|\ p(x)\ \forall x \in L$ and scalar α.

3.4 Review Questions

1. Let M be a closed linear subspace of a normed linear space N and x_o is a vector not in M. Then there exists a functional $f_o \in N^*$ such that

$$f_o(M) = 0 \text{ and } f_o(x_o) \neq 0$$

2. Let M be a closed linear subspace of a normed linear space N, and let x_o be a vector not in M. If d is the distance from x_o to M, then these exists a functional $f_o \in N^*$ such that $f_o(M) = 0$, $f_o(x_o) = d$, and $\| f_o \| = 1$.

3. Let M is a closed linear subspace of a normed linear space N and $x_o \in N$ such that $x_o \notin M$. If d is the distance from x_o to M, then there exists a functional $f_o \in N^*$ such that $f_o(M) = 0$, $f_o(x_o) = 1$ and $\| f_o \| = \frac{1}{d}$.

4. Let N be a normed linear space over R or C. Let $M \subset N$ be a linear subspace. Then $\overline{M} = N \Leftrightarrow f \in N^*$ is such that $f(x) = 0$ for every $x \in M$, then $f = 0$.

5. A normed linear space is separable if its conjugate (or dual) space is separable.

3.5 Further Readings

Books

Walter Rudin, *Functional Analysis* (2nd ed.). McGraw-Hill Science/Engineering/Math 1991.

Eberhard Zeidler, *Applied Functional Analysis: Main Principles and their Applications*, Springer, 1995.

Online links

mat.iitm.ac.in

www.math.ksu.edu

mizar.uwb.edu.pl/JFW/Vol5/hahnban.html

Unit 4: The Natural Imbedding of N in N**

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define the natural imbedding of N into N**.
- Define reflexive mapping.
- Describe the properties of natural imbedding of N into N*.

Introduction

As we know that conjugate space N* of a normed linear space N is itself a normed linear space. So, we can find the conjugate space (N*)* of N*. We denote it by N** and call it the second conjugate space of N. Likewise N*, N** is also a Banach space. The importance of the space N** lies in the fact that each vector x in N given rise to a functional F_x in N** and that there exists an isometric isomorphism of N into N**, called natural imbedding of N into N**.

4.1 The Natural Imbedding of N into N**

4.1.1 Definition: Natural Imbedding of N into N**

The map $J : N \rightarrow N^{**}$ defined by

$$J(x) = F_x \;\; \forall\, x \in N,$$

is called the natural imbedding of N into N**.

Since $J(N) \subset N^{**}$, N can be considered as part of N** without changing its basic norm structure. We write $N \subset N^{**}$ in the above sense.

4.1.2 Definition: Reflexive Mapping

If the map J : N → N** defined by

$$J(x) = F_x \;\; \forall x \in N,$$

is onto also, then N (or J) is said to be reflexive (or reflexive mapping). In this case we write N = N**, i.e., if N = N**, then N is reflexive.

> *Note* Equality in the above definition is in the sense of isometric isomorphism under the natural imbedding. Since N** must always be a complete normed linear space, no incomplete space can be reflexive.

4.1.3 Properties of Natural Imbedding of N into N**

I. Let N be a normed linear space. If $x \in N$, then

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

Using natural imbedding of N into N**, we have for every $x \in N$,

$$F_x(f) = f(x) \;\text{ and }\; \| F_x \| = \| x \|.$$

Now, $\| Fx \| = \sup\limits_{\|f\|=1} \{ | F_x(f) | \} = \sup\limits_{\|f\|=1} \{ | f(x) |, f \in N^* \}$

therefore, $\| x \| = \sup \{ | f(x) | : f \in N^*, \| f \| = 1 \}$.

II. Every normed linear space is a dense linear subspace of a Banach space.

Let N be a normed linear space. Let

J : N → N** be the natural imbedding of N into N**.

The image of the mapping is linear subspace $J(N) \subset N^{**}$. Let $\overline{J(N)}$ be the closure of N(N) in N**.

Since N** is a Banach space, its closed subspace $\overline{J(N)}$ is also a Banach space. Hence if we identity N with J(N), then J(N) is a dense subspace of a Banach space.

4.1.4 Theorems and Solved Examples

Theorem 1: Let N be an arbitrary normal linear space. Then each vector x in N induces a functional Fx on N* defined by

$$F_x(f) = f(x) \text{ for every } f \in N^* \text{ such that } \| F_x \| = \| x \|.$$

Further, the mapping J : N → N** : $J(x) = F_x$ for every $x \in N$ defines and isometric isomorphisms of N into N**.

Proof: To show that F_x is actually a function on N*, we must prove that F_x is linear and bounded (i.e. continuous).

We first show F_x is linear.

Notes

Let f, g ∈ N* and α, β be scalars. Then

$$F_x(\alpha f + \beta g) = (\alpha f + \beta g)\,x = \alpha f(x) + \beta g(x)$$
$$= \alpha F_x(f) + \beta F_x(g)$$

⇒ F_x is linear

F_x is bounded.

For any f ∈ N*, we have

$$|F_x(f)| = |f(x)|$$
$$\leq \|f\|\ \|x\| \qquad \dots (1)$$

Thus the constant $\|x\|$ is bounded (in the sense of a bounded linear functional) for F_x. Hence F_x is a functional on N*.

We now prove $\|F_x\| = \|x\|$

We have $$\|F_x\| = \sup\{|F_x(f)| : \|f\| \leq 1\}$$
$$\leq \sup\{\|F\|\ \|x\| : \|f\| \leq 1\} \qquad \text{(Using (1))}$$
$$\leq \|x\|$$

Hence $$\|F_x\| \leq \|x\| \qquad \dots (2)$$

To prove the reverse inequality we consider the case when x = 0. In this case (2) gives $\|F_o\| = \|0\| = 0$.

But $\|F_x\| = 0$ always. Hence $\|F_o\| = \|0\|$ i.e. $\|F_x\| = \|x\|$ for x = 0.

Not let x ≠ 0 be a vector in N. Then by theorem (If N is a normal linear space and x_o ∈ N, $x_o \neq 0$, then there exists a functional f_o ∈ N* such that

$$f_o(x_o) = \|x_o\| \text{ and } \|f_o\| = 1.)$$

∃ a functional f ∈ N* such that

$$f(x) = \|x\| \text{ and } \|f\| = 1.$$

But $$\|F_x\| = \sup\{|F_x(f)| : \|f\| \leq 1\}$$
$$= \sup\{|f(x)| : \|f\| = 1\}$$

and since $$\|x\| = \|f(x)\| \leq \sup\{|f(x)| : \|f\| = 1\}$$

we conclude that $$\|F_x\| \geq \|x\| \qquad \dots (3)$$

[Note that since $f(x) = \|x\| \geq 0$ we have $f(x) = |f(x)|$]

From (2) and (3); we have

$$\|F_x\| = \|x\| \qquad \dots (4)$$

Finally, we show that J is an isometric isomorphism of N into N**. For any x, y ∈ N and α scalar.

$$F_{x+y}(f) = f(x + y) = f(x) + f(y)$$
$$= F_x(f) + F_y(f)$$

∴ $$F_{x+y}(f) = (F_x + F_y)\,f \qquad \dots (5)$$

⇒ $$F_{x+y} = F_x + F_y \qquad \dots (6)$$

Further, $$F_{\alpha x}(f) = f(\alpha x) = \alpha f(x) = (\alpha F_x)(f)$$

Hence $F_{\alpha x} = \alpha F_x$ … (7)

Using definition of J and equations (6) and (7) we get

$$J_{(x+y)} = F_{x+y} = F_x + F_y = J(x) + J(y) \quad \dots (8)$$

and $$J_{(\alpha x)} = F_{\alpha x} = \alpha F_x = \alpha J(x) \quad \dots (9)$$

(8) and (9) ⇒ J is linear and also (4) shows that J is norm preserving.

For any x and y in N, we have

$$\| J(x) - J(y) \| = \| F_x - F_y \| = \| F_{x-y} \| = \| x - y \| \quad \dots (10)$$

Thus J preserve distances and it is an isometry. Also (10) shows that

$$J(x) - J(y) = 0 \Rightarrow J(x - y) = 0 \Rightarrow x - y = 0$$

i.e. $J(x) = J(y) \Rightarrow x = y$ so that J is one-one.

Hence J defines an isometric isomorphism of N into N**. This completes the proof of the theorem.

Example 1: The space ℓ_p^n $(1 \le p < \infty)$ are reflexive.

Solution: We know that if $1 \le p < \infty$, then

$$\left(\ell_p^n\right)^* = \ell_p^n .$$

But $$\left(\ell_q^n\right)^* = \ell_p^n$$

Hence $$\left(\ell_p^n\right)^{**} = \ell_p^n$$

Similarly we have $$\left(\ell_1^n\right)^{**} = \ell_1^n \text{ for } p = 1$$

and $$\left(\ell_\infty^n\right)^{**} = \ell_\infty^n \text{ for } p = \infty$$

So that ℓ_p^n spaces are reflexive for $1 \le p < \infty$.

Example 2: The space ℓ_p for $1 < p < \infty$ are reflexive.

Sol: We know that if $\ell^*_p = \ell_p$ and $\ell^*_q = \ell_p$

⇒ $\ell^{**}_q = \ell_p$.

⇒ ℓ_p are reflexive for $1 < p < \infty$.

A similar result can be seen to hold for $L_p(X)$.

Example 2: If N is a finite dimensional normed linear space of dimension m, then N* also has dimension m.

Solution: Since N is a finite dimensional normed linear space of dimension m then $\{x_1, x_2, \dots, x_m\}$ is a basis for N, and if $(\alpha_1\ \alpha_2 \dots \alpha_m)$ is any set of scalars, then there exists a functional f on N such that $f(x_i) = \alpha_i$, i = 1, 2, …m.

Notes

To show that N* is also of dimension m, we have to prove that there is a uniquely determined basis $(f_1, f_2, \ldots, f_m)$ in N*, with $f_j(x_i) = \delta_y$.

By the above fact, for each i = 1, 2, …, m, a unique f_j in N* exists such that $f_j(x_i) = \delta_{ij}$. We show now that $\{f_1, f_2, \ldots, f_n\}$ is a basis in N* to complete our proof.

Let us consider $\alpha_1 f_1 + \alpha_2 f_2 + \ldots\ldots \alpha_m f_m = 0$ … (1)

For all $x \in N$, we have $\alpha_1 f_1(x) + \alpha_2 f_2(x) + \ldots\ldots + \alpha_m f_m(x) = 0$.

We have $\sum_j \alpha_j f_j(x_j) = 0 = \sum_j \alpha_j \delta_{ij} = \alpha_i$ for i = 1, 2, …, m, when $x = x_i$.

⇒ $f_1, f_2, \ldots\ldots f_m$ are linearly independent in N*.

Now let $f(x_i) = \alpha_i$.

Therefore if $x = \sum_d \beta_i x_i$, we get

$$f(x) = \beta_1 f(x_1) + \beta_2 f(x_2) + \ldots\ldots + \beta_m f(x_m) \qquad \ldots (2)$$

Further $f_j(x) = \beta_1 f_j(x_1) + \ldots\ldots + \beta_i f_j(x_i) + \ldots\ldots + \beta_m f_j(x_m)$

⇒ $$f_j(x) = \beta_j$$

From (1) and (2), it follows that

$$f(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \ldots\ldots + \alpha_m f_m(x)$$

$$= (\alpha_1 f + \alpha_2 f_2 + \ldots\ldots + \alpha_m f_m)(x)$$

⇒ $(f_1, f_2, \ldots\ldots f_m)$ spans the space.

⇒ N* is m-dimensional.

4.2 Summary

- The map $J : N \to N^{**}$ defined by

  $J(x) = F_x \; \forall x \in N$,

  is called the natural imbedding of N into N**.
- If the map $J : N \to N^{**}$ defined by

  $J(x) = F_x \; \forall x \in N$,

  is onto also, then N (or J) is said to be reflexive. In this case we write N = N**, i.e., if N = N**, then N is reflexive.
- Let N be an arbitrary normal linear space. Then each vector x in N induces a functional F_x on N* defined by $F_x(f) = f(x)$ for every $f \in N^*$ such that $\| F_x \| = \| x \|$.

4.3 Keywords

Natural Imbedding of N into N**: The map $J : N \to N^{**}$ defined by

$J(x) = F_x \; \forall x \in N$,

is called the natural imbedding of N into N**.

Reflexive Mapping: If the map $J : N \to N^{**}$ defined by

$$J(x) = F_x \ \forall x \in N,$$

is onto also, then N (or J) is said to be reflexive (or reflexive mapping).

4.4 Review Questions

1. Let X be a compact Hausdorff space, and justify the assertion that C (X) is reflexive if X is finite.
2. If N is a finite-dimensional normed linear space of dimension n, show that N* also has dimension n. Use this to prove that N is reflexive.
3. If B is a Banach space, prove that B is reflexive $\Leftrightarrow$ B* is reflexive.
4. Prove that if B is a reflexive Banach space, then its closed unit sphere S is weakly compact.
5. Show that a linear subspace of a normed linear space is closed $\Leftrightarrow$ it is weakly closed.

4.5 Further Readings

Books

G.F. Simmons, *Introduction to topology and Modern Analysis*. McGraw-Hill, Kogakusha Ltd.

J.B. Conway, *A Course in Functional Analysis*. Springer-Verlag.

Online links

www.mathoverflow.net/…/natural-embedding

www.tandfonline.com

Unit 5: The Open Mapping Theorem

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- State the open mapping theorem.
- Understand the proof of the open mapping theorem.
- Solve problems on the open mapping theorem.

Introduction

In this unit, we establish the open mapping theorem. It is concerned with complete normed linear spaces. This theorem states that if T is a continuous linear transformation of a Banach space B onto a Banach space B′, then T is an open mapping. Before proving it, we shall prove a lemma which is the key to this theorem.

5.1 The Open Mapping Theorem

5.1.1 Lemma

Lemma 1: If B and B′ are Banach spaces and T is a continuous linear transformation of B onto B′, then the image of each sphere centered on the origin in B contains an open sphere centered on the origin in B′.

Proof: Let S_r and S'_r respectively denote the open sphere with radius r centered on the origin in B and B′.

We one to show that T (S_r) contains same S'_r.

However, since T (S_r) = T (r S_1) = r T (S_1), (by linearity of T).

It therefore suffices to show that T (S_1) contains some S'_r for then S'_δ, where $\delta = r^2$, will be contained in T (S_r). We first claim that $\overline{T(S_1)}$ (the closure of T (S_1)) contains some S'_r.

Notes

If x is any vector in B we can by the Archimedean property of real numbers find a positive integer n such that n > || x ||, i.e., $x \in S_n$,

therefore

$$B = \bigcup_{n=1}^{\infty} S_n$$

and since t is onto, we have

$$B' = T(B)$$

$$= T\left(\bigcup_{n=1}^{\infty} S_n\right)$$

$$= \bigcup_{n=1}^{\infty} T(S_n)$$

Now B′ being complete, Baire's theorem implies that some $\overline{T(S_{n_0})}$ possesses an interior point Z_0. This in turn yields a point $y_0 \in T(S_{n_0})$ such that y_0 is also an interior point of $\overline{T(S_{n_0})}$.

Further, maps j : B′ → B′ and g : B′ → B′

defined respectively by j (y) = y = y – y_0 and g (y) = 2 n_0 y

where n_o is a non-zero scalars, are homeomorphisms as shown below f is one-to-one and onto. To show f, f^{-1} are continuous, let $y_n \in B'$ and $y_n \to y$ in B.

Then $\quad f(y_n) = y_n - y_0 \to y - y_0 = f(y)$

and $\quad f^{-1}(y_n) = y_n + y_0 \to y + y_0 = f^{-1}(y)$

Hence f and f^{-1} are both continuous so that is a homeomorphism.

Similarly g : B′ → B′ : g (x) = $2n_0y$ is a homeomorphism for, g is one-to-one, onto and bicontinuous for $n_0 \neq 0$.

Therefore we have

(i) $\quad f(y_0) = 0$ = origin in B′ is an interior point of $f\left(\overline{T(S_n)}\right)$.

(ii) $$f\left(\overline{T(S_{n_0})}\right) = \overline{f(T(S_{n_0}))}$$

$$= \overline{T(S_{n_0}) - y_0}$$

$$\subseteq \overline{T(S_{2n_0})} \qquad (\because y_0 \in T(S_{n_0}))$$

(iii) $$\overline{T(S_{2n_0})} = \overline{T(2_{n_0}S_1)} = \overline{2n_0\, T(S_1)}$$

$$= g\,\overline{(T(S_1))} = g\left(\overline{(T(S_1))}\right)$$

$$= 2_{n_0}\overline{T(S_1)}$$

Combining (i) – (iii), it follows that origin is also an interior point of $\overline{(T(S_1))}$. Consequently, there exists ε > 0 such that

$$S'_\varepsilon \subseteq \overline{T(S_1)}$$

This justifies our claims.

We conclude the proof of the lemma by showing that

$$\bar{S}_{\varepsilon/3} \subseteq T(S_1)\text{, i.e., } S'_\varepsilon \subseteq T(S_3)$$

Let $y \in B'$ such that $\| y \| < \varepsilon$. Then $y \in \overline{T(S_1)}$ and therefore there exists a vector $x_1 \in B$ such that

$$\| x_1 \| < 1, \| y - y_1 \| < \varepsilon/2 \text{ and } y_1 = T(x_1)$$

We next observe that

$$\bar{S}_{\varepsilon/2} \subseteq \overline{T(S_{1/2})} \text{ and } y - y_1 \in S'_{\varepsilon/2}$$

Therefore there exists a vector $x_2 \in B$ such that

$$\| x_3 \| < \frac{1}{2}, \| y - y_1 - y_2 \| < \frac{\varepsilon}{2^2} \text{ and } y_2 = T(x_2)$$

Continuing this process, we obtain a sequence (x_n) in B such that

$$\| x_n \| < \frac{1}{2^{n-1}}, y_n = T(x_n) \text{ and } \| y - (y_1 + y_2 + \ldots + y_n) \| < \frac{\varepsilon}{2^2}$$

Let $s_n = x_1 + x_2 + \ldots + x_n$, then

$$\begin{aligned}
\| s_n \| &= \| x_1 + x_2 + \ldots + x_n \| \\
&\leq \| x_1 \| + \| x_2 \| + \ldots + \| x_n \| \\
&< 1 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}} \\
&\Rightarrow 2\left(1 - \frac{1}{2^n}\right) \\
&< 2
\end{aligned}$$

Also for n > m, we have

$$\begin{aligned}
\| s_n - s_m \| &= \| s_{m+1} + \ldots + x_n \| \\
&\leq \| x_{m+1} \| + \ldots + \| x_n \| \\
&< \frac{1}{2^m} + \ldots + \frac{1}{2^{n-1}} \\
&= \frac{\frac{1}{2^m}\left(1 - \frac{1}{2^{n-m}}\right)}{1 - \frac{1}{2}} \qquad \text{(summing the G.P.)} \\
&= \frac{1}{2^{m-1}} - \frac{1}{2^{n-m-1}} \\
&\to 0 \text{ as } m, n \to \infty
\end{aligned}$$

Thus (s_n) is a Cauchy sequence in B and since B is complete, $\exists$ a vector $x \in B$ such that

$$s_n \to x \text{ and therefore}$$

$$\| x \| = \left\| \lim_{n\to\infty} s_n \right\| = \lim_{n\to\infty} \| s_n \| \le 2 < 3,$$

i.e., $x \in S_3$.

It now follows by the continuity of T that

$$T(x) = T\left(\lim_{n\to\infty} s_n\right)$$

$$= \lim_{n\to\infty} T(s_n)$$

$$= \lim_{n\to\infty} (y_1 + y_2 + \ldots + y_n)$$

$$= y$$

Hence $\quad y \in T(S_3)$

Thus $\quad y \in S'_\varepsilon \Rightarrow y \in T(S_3)$. Accordingly

$$S'_\varepsilon \subseteq T(S_3)$$

This completes the proof of the lemma.

Note If B and B′ are Banach spaces, the symbol S (x; r) and S′ (x; r) will be used to denote open spheres with centre x and radius r in B and B′ respectively. Also S_r and S'_r will denote these spheres when the centre is the origin. It is easy to see that

$$S(x; r) = x + S_r \text{ and } S_r = r S_1$$

For, we have

$$y \in S(x; r) \Rightarrow \| y - x \| < r$$

$$\Rightarrow \| z \| < r \text{ and } y - x = z, z \in S_r$$

$$\Rightarrow y = x + z \text{ and } \| z \| < r$$

$$\Rightarrow y \in x + S_r$$

Thus $\quad S(x; r) = x + S_r$

and $\quad S_r = \{x : \| x \| < r\} = \left\{ x : \frac{\| x \|}{r} < 1 \right\}$

$$= \{r \,.\, y \| y \| < 1\}$$

$$= r S_1$$

Thus $\quad S_r = r S_1$

Now we prove an important lemma which is key to the proof of the open mapping theorem.

5.1.2 Proof of the Open Mapping Theorem

Statement: If T is a continuous linear transformation of a Banach space B onto a Banach space B′, then T is an open mapping.

Proof: Let G be an open set in B. We are to show that T (G) is an open set in B′

i.e. if y is any point of T (G), then there exists an open sphere centered at y and contained in T (G).

$$y \in T(G) \Rightarrow y = T(x) \text{ for some } x \in G.$$

$x \in G$, G open in B ⇒ there exists an open sphere S (x; r) with centre x and radius r such that $S(x; r) \subseteq G$.

But as remarked earlier we can write $S(x; r) = x + S_r$, where S_r is open sphere of radius r centered at the origin in B.

Thus $$x + S_r \subseteq G \qquad \ldots (1)$$

By lemma (2) (prove it),

$T(S_r)$ contains some S'_{r_1}. Therefore

$$\begin{aligned} S'(y; r_1) &= y + S'_{r_1} \\ &\subseteq y + T(S_r) \\ &= T(x) + T(S_r) \\ &= T(x + S_r) \\ &\subseteq T(G), \qquad \text{(Using (1))} \end{aligned}$$

since $$x + S_r = S(x; r) \subseteq G.$$

Thus we have shown that to each $y \in T(G)$, there exists an open sphere in B′ centered at y and contained in T (G) and consequently T (G) is an open set.

This completes the proof of the theorem.

5.1.3 Theorems and Solved Examples

Theorem 1: Let B and B′ be Banach spaces and let T be an one-one continuous linear transformation of B onto B′. Then T is a homeomorphism.

In particular, T^{-1} is automatically continuous.

Proof: We know that a one-to-one continuous open map from B onto B′ is a homeomorphism.

By hypothesis $T : B \to B'$ is a continuous one-to-one onto mapping.

By the open mapping theorem, T is open. Hence T is a homeomorphism. Since T is homeomorphism, $T^{-\prime}$ exists and continuous from B′ to B so that $T^{-\prime}$ is bounded and hence

$$T^{-1} \in \beta(B', B).$$

This completes the proof of the theorem.

Cor. 1: Let B and B′ be Banach spaces and let $T \in \beta(B, B')$. If $T : B \to B'$ is one-to-one and onto, there are positive numbers m and M such that

$$m \|x\| \le \|T(x)\| \le M \|x\|.$$

Proof: By the theorem,

$T : B \to B'$ is a homeomorphism. So that T and T^{-1} are both continuous and hence bounded. Hence by theorem,

Let N and N′ be normed linear spaces. Then N and N′ are topologically isomorphic if and only if there exists a linear transformation T of N onto N′ and positive constants m and M such that

$$m \| x \| \leq \| T(x) \| \leq M \| x \|, \text{ for every } x \in N.$$

$\exists$ constants m and M such that

$$m \| x \| \leq \| T(x) \| \leq M \| x \|.$$

Theorem 2: If a one-to-one linear transformation T of a Banach space B onto itself is continuous, then its inverse T^{-1} is continuous.

Proof: T is a homeomorphism (using theorem of B onto B. Hence T^{-1} is continuous.

This completes the proof of the theorem.

Note The following examples will show that the completeness assumption in the open mapping theorem and theorem can neither be omitted in the domain of definition of T nor in the range of T.

Example 1: Let C′ [0, 1] be the set of all continuous differentiable function on [0, 1]. We know that C′ [0, 1] is an incomplete space with the norm

$$\| f \|_{\infty} = \sup \{ | f(x) | : 0 \leq x \leq 1 \}$$

But it is complete with respect to the norm

$$\| f \| = \| f \|_{\infty} + \| f' \|_{\infty}.$$

Now let us choose $B = [C' [0, 1], \| \ \|]$ and $N = [C' [0, 1], \| \ \|_{\infty}]$.

Consider the identity mapping $I : B \to N$. The identity mapping is one-to-one onto and continuous. I^{-1} is not continuous. For, if it were continuous, then it is a homeomorphism. Mapping of a complete space into an incomplete space which cannot be. Hence I does not map open sets into open sets.

Thus the open mapping theorem fails if the range of T is not a Banach space.

Example 2: Let B′ be an infinite dimensional Banach space with a basis $\{\alpha_i : i \in I\}$ with $\| \alpha_i \| = 1$ for each $i \in I$. Let N be the set of all functions from I to C which vanish everywhere except a finite member of points in I. Then N is a linear space under addition and scalar multiplication. We can define the norm on N as

$$\| f \| = \Sigma \ | f(i) |, i \in I.$$

Then N is an incomplete normed linear space. Now consider the transformation

$T : N \to B'$ defined as follows.

For each $f \in N$, let $T(f) = \Sigma f(i) \alpha_i$.

Then T is linear and

$$\| T(f) \| \leq \sum_{i \in I} | f(i) | \, \| \alpha_i \|$$

$$= \sum_{i \in I} | f_i | = \| f \| \text{ for every } f \in N.$$

Hence T is bounded transformation from N to B′. It is also one-to-one and onto. But T does not map open subsets of N onto B′. For, if it maps, it is a linear homeomorphism from N onto B′ which cannot be since N is incomplete.

Theorem 3: Let B be a Banach space and N be a normed linear space. If T is a continuous linear open map on B onto N, then N is a Banach space.

Proof: Let (y_n) be a Cauchy sequence in N. Then we can find a sequence of positive integer (n_k) such that $n_k < n_{k+1}$ and for each k

$$\left\| y_{n_{k+1}} - y_{n_k} \right\| < \frac{1}{2^k}$$

Hence by theorem: "Let N and N′ be normed linear spaces. A linear map T : N → N′ is open and onto if and only if there is a M > 0 such that for any y ∈ N′, there is a x ∈ N such that Tx = y and $\| x \| \leq M \| y \|$."

For $\left(y_{n_{k+1}} - y_n\right) \in N$, there is a $n_k \in B$ and a constant M such that

$T(x_k) = y_{n_{k+1}} - y_n$ and $\|x_k\| \leq \left\| y_{n_{k+1}} - y_n \right\|$.

By on choice $\sum_{k=1}^{\infty} \left\| y_{n_{k+1}} - y_{n_k} \right\|$ is convergent so that $\sum_{k=1}^{\infty} x_{n_k}$ is convergent. Since B is a Banach space there is a x ∈ B such that

$$x = \lim_{n \to \infty} \sum_{k=1}^{n} x_{n_k}$$

Since T is continuous $\sum_{k=1}^{n} T(x_k) \to T(x)$ as $n \to \infty$

But $\sum_{k=1}^{\infty} T(x_k) = y_{n_{k+1}} - y_{n_1}$ so that

$$y_{n_{k+1}} - y_{n_1} \to T(x) \Rightarrow y_{n_{k+1}} \to y_{n_1} + T(x)$$

Since (y_n) is a Cauchy sequence such that every subsequences is convergent, (y_n) itself converges and $y_n \to y_{n_1} + T(x)$ in N.

Hence N is complete. Consequently, N is a Banach space.

This completes the proof of the theorem.

Example: Let N be complete in two norms $\| \ \|_1$ and $\| \ \|_2$ respectively. If there is a number $a > 0$ such that $\| x \|_1 \leq a \| x \|_2$ for all x ∈ N, then show that the two norms are equivalent.

Solution: The identity map

$i : (N, \| \ \|_2) \to (N, \| \ \|_1)$ is an one-one onto map.

Also $\| x \|_1 \leq a \| x \|_2 \Rightarrow$ i is bounded … (1)

$\Rightarrow$ i is continuous.

Hence by open-mapping theorem, i is open and so it is homeomorphism of $(N, \| x \|_2)$ onto $(N, \| x \|_1)$. Consequently i is bounded as a map from $(N, \| x \|_1) \rightarrow (N, \|x \|_2)$

Since $i^{-1}(x) = x$, $\exists a$, b s.t. $\| x \|_2 \leq b \| \ \|_1$ … (2)

$\therefore$ (1) & (2) imply that the norms are equivalent.

5.2 Summary

- If B and B′ are Banach spaces and T is a continuous linear transformation of B onto B′, then the image of each sphere centered on the origin in B contains an open sphere centered on the origin in B′.
- The open mapping theorem : If T is a continuous linear transformation of a Banach space B onto a Banach space B′, then T is an open mapping.

5.3 Keywords

Banach Space: A normed space V is said to be Banach space if for every Cauchy sequence $\{v_n\}_{n=1}^{\infty} \subset V$ then there exists an element $v \in V$ such that $\lim\limits_{n\to\infty} v_n = v$.

Homeomorphism: A map $f : (X, T) \rightarrow (Y, U)$ is said to be homeomorphism if

(i) f is one-one onto.

(ii) f and f^{-1} are continuous.

Open Sphere: Let $x_o \in X$ and $r \in R^+$. Then set $\{x \in X : p(x_o, x) < r\}$ is defined as open sphere with centre x_o and radius r.

5.4 Review Questions

1. If X and Y are Banach spaces and $A : X \rightarrow Y$ is a bounded linear transformation that is bijective, then prove that A^{-1} is bounded.

2. Let X be a vector space and suppose $\| \cdot \|_1$, and $\| \cdot \|_2$ are two norms on X and that T_1 and T_2 are the corresponding topologies. Show that if X is complete in both norms and $T_1 \supseteq T_2$, then $T_1 = T_2$.

5.5 Further Readings

Books

Walter Rudin, *Functional Analysis*, McGraw-Hill, 1973.

Jean Diendonne, *Treatise on Analysis*, Volume II, Academic Press (1970).

Online links

euclid.colorado.edu/ngwilkin/files/math 6320…/OMT_CGT.pdf

people.sissa.it/nbianchin/courses/…/lecture05.banachstein.pdf

Unit 6: The Closed Graph Theorem

CONTENTS

Objectives

After studying this unit, you will be able to:

- State the closed graph theorem.
- Understand the proof of the closed graph theorem
- Solve problems based on the closed graph theorem.

Introduction

Though many of the linear transformations in analysis are continuous and consequently bounded, there do exist linear transformation which are discontinuous. The study of such kind of transformation is much facilitated by studying the graph of transformation and using the graph of the transformation as subset in the Cartesian product space to characterise the boundedness of such transformations. The basic theorem in this regard is the closed graph theorem.

6.1 The Closed Graph Theorem

6.1.1 Graph of Linear Transformation

Definition: Let N and N′ be a normed linear space and let $T : N \rightarrow N'$ be a mapping with domain N and range N′. The graph of T is defined to be a subset of $N \times N'$ which consists of all ordered pairs (x, T (x)). It is generally denoted by G_T.

Therefore the graph of $T : N \rightarrow N'$ is

$$G_T = \{(x, T(x) : x \in N\}.$$

Notes G_T is a linear subspace of the Cartesian product $N \times N'$ with respect to coordinate-wise addition and scalar multiplications.

Notes

Theorem 1: Let N and N′ be normed linear spaces. Then N × N′ is a normed linear space with coordinate-wise linear operations and the norm.

$$\| (x, y) \| = \left(\| x \|^p + \| x \|^p\right)^{\frac{1}{p}}, \text{ where } x \in N, y \in N'$$

and | ≤ p < ∞. Moreover, this norm induces the product topology on N × N′, and N × N′ is complete iff both N and N′ are complete.

Proof:

(i) It needs to prove the triangle inequality since other conditions of a norm are immediate.

Let (x, y) and (x′, y′) be two elements of N × N′.

Then $\| (x, y) + (x', y')\| = \| (x + x', y + y')\|$

$$= \left(\| x + x' \|^p + \| y + y' \|^p\right)^{\frac{1}{p}}$$

$$= \left\{\left(\| x \| + \| x' \|\right)^p + \left(\| y \| + \| y' \|\right)^p\right\}^{\frac{1}{p}}$$

$$= \left\{\left(\| x \|^p + \| y \|^p\right)^{\frac{1}{p}} + \left(\| x' \|^p + \| y' \|^p\right)^{\frac{1}{p}}\right\}$$

(By Minkowski's inequality)

$$= \| (x, y) \| + \| (x', y') \|.$$

This establishes the triangular inequality and therefore N × N′ is a normed linear space.

Furthermore $(x_n, y_n) \to (x, y) \Leftrightarrow x_n = x$ and $y_n = y$. Hence theorem on N × N′ induces the product topology.

(ii) Next we show that N × N′ is complete ⇔ N, N′ are complete.

Let (x_n, y_n) be a Cauchy sequence in N × N′. Given ∈ > 0, we can find a n_o such that

$$\| (x_n, y_n) - (x_m, y_m) \| < \in \ \forall \ m, n \geq n_o. \quad \ldots (1)$$

$\Rightarrow \| (x_n - x_m) \| < \in$ and $\| y_n - y_m \| < \in \ \forall \ m, n \geq n_o$

$\Rightarrow (x_n)$ and (y_n) are Cauchy sequences in N and N′ respectively.

Since N, N′ are complete, let

$x_n \to x_o \in N$ and $y_n \to y_o \in N'$ in their norms,

i.e. $\| (x_n - x_o) \| < \in$ and $\| y_n - y_m \| < \in \ \forall \ m, n \geq n_o$. … (2)

since $x_o \in N$, $y_o \in N'$, $(x_o, y_o) \in N \times N'$.

Further $\| (x_n, y_n) - (x_o, y_o) \| < \in \ \forall \ n \geq n_o$ (using (2))

$\Rightarrow (x_n, y_n) \to (x_o, y_o)$ in the norm of N × N′ and $(x_o, y_o) \in N \times N'$.

⇒ N × N′ is complete.

The converse follows by reversing the above steps.

This completes the proof of the theorem.

Notes The following norms are equivalent to above norm

(i) $\| (x, y) \| = \max \{\| x \|, \| y \|\}$

(ii) $\| (x, y) \| = \| x \| + \| y \|$ (p = 1 in the above theorem)

6.1.2 Closed Linear Transformation

Definition: Let N and N′ be normed linear spaces and let M be a subspace of N. Then a linear transformation

$T : M \to N'$ is said to be closed

iff $x_n \in M$, $x_n \to x$ and $T(x_n) \to y$ imply $x \in M$ and $y = T(x)$.

Theorem 2: Let N and N′ be normed linear spaces and B be a subspace of N. Then a linear transformation $T : M \to N'$ is closed $\Leftrightarrow$ its graph G_T is closed.

Proof: Let T is closed linear transformation. We claim that its graph G_T is closed i.e. G_T contains all its limit point.

Let (x, y) be any limit point of G_T. Then $\exists$ a sequence of points in G_T, $(x_n, T(x_n))$, $x_n \in M$, converging to (x, y). But

$$(x_n, T(x_n)) \to (x, y)$$

$$\Rightarrow \quad \| x_n, T(x_n) - (x, y) \| \to 0$$

$$\Rightarrow \quad \| (x_n - x), T(x_n) - y \| \to 0$$

$$\Rightarrow \quad \| x_n - x \| + \| T(x_n) - y \| \to 0$$

$$\Rightarrow \quad \| x_n - x \| \to 0 \text{ and } \| T(x_n) - y \| \to 0$$

$$\Rightarrow \quad x_n \to x \text{ and } T(x_n) \to y \qquad (\because T \text{ is closed})$$

$$\Rightarrow \quad (x, y) \in G_T. \qquad \text{(By def. of graph)}$$

Thus we have shown that every limit point of G_T is in G_T and hence G_T is closed.

Conversely, let the graph of T, G_T is closed.

To show that T is closed linear transformation.

Let $x_n \in M$, $x_n \to x$ and $T(x_n) \to y$.

Then it can be seen that (x, y) is an adherent point of G_T so that

$(x, y) \in \overline{G}_T$. But $\overline{G}_T = G_T$ $\quad (\because G_T$ is closed)

Hence $(x, y) \in G_T$ and so by the definition of G_T we have $x \in M$ and $y = T(x)$.

Consequently, T is a closed linear transformation. This completes the proof of the theorem.

6.1.3 The Closed Graph Theorem – Proof

If B and B′ are Banach spaces and if T is linear transformation of B into B′, then T is continuous $\Leftrightarrow$ Graph of T (G_T) is closed.

Proof: *Necessary Part:*

Let T be continuous and let G_T denote the graph of T, i.e.

$$G_T = \{(x, T(x) : x \in B\} \subseteq B \times B'.$$

Notes

We shall show that $\overline{G}_T = G_T$.

Since $G_T \subset \overline{G}_T$ always, it suffices to show that $\overline{G}_T \subset G_T$.

Let $(x, y) \in \overline{G}_T$. Then there exists a sequence $(x_n, T(x_n))$ in G_T such that

$$(x_n, T(x_n)) \to (x, y)$$

$\Rightarrow$ $x_n \to x$ and $T(x_n) \to y$.

But T is continuous $\Rightarrow T(x_n) \to T(x)$ and so $y = T(x)$

$\Rightarrow$ $(x, y) = (x, T(x)) \in G_T$

$\Rightarrow$ $\overline{G}_T \subset G_T$

Hence $G_T = \overline{G}_T$ i.e. G_T is closed.

Sufficient Part:

Let G_T is closed. Then we claim that T is continuous. Let B_1 be the given linear space B renormed by $\| \ \|_1$ given by

$$\| x \|_1 = \| x \| + \| T(x) \| \text{ for } x \in B.$$

Now $\| T(x) \| \le \| x \| + \| T(x) \| = \| x \|_1$.

$\Rightarrow$ T is bounded (continuous) as a mapping from B_1 to B′.

So if B and B_1 have the same topology then T will be continuous from B to B′. To this end, we have to show that B and B_1 are homeomorphic.

Consider the identity mapping

$I : B_1 \to B$ defined by

$I(x) = x$ for every $x \in B_1$.

Then I is always one-one and onto.

Further $\| I(x) \| = \| x \| \le \| x \| + \| T(x) \| = \| x \|_1$

$\Rightarrow$ I is bounded (continuous) as a mapping from B_1 onto B.

Therefore if we show that B_1 is complete with respect to $\| \ \|_1$, then B_1 is a Banach space so by theorem.

"Let B and B′ be Banach spaces and let T be one-one continuous linear transformation of B onto B′. Then T is a homeomorphism. In particular, T^{-1} is automatically continuous."

I is homeomorphism. Therefore to complete the proof, we have to show that B_1 is complete under the norm $\| \ \|_1$.

Let (x_n) be a Cauchy sequence in B_1. Then

$$\| x_n - x_m \|_1 = \| x_n - x_m \| + \| T(x_n - x_m) \| \to 0 \text{ as } m, n \to \infty$$

$\Rightarrow$ (x_n) and $(T(x_n))$ are Cauchy sequences in B and B′ respectively.

Since B and B′ are complete, we have

$$x_n \to x \text{ in B and } T(x_n) \to T(x) \text{ in B}' \quad \ldots (1)$$

Since G_T is closed, we have

$(x_1 T(x)) \in G_T$ and if we take

$y = T(x)$; then $(x, y) \in G_T$.

Now $\quad \| x_n - x \|_1 = \| x_n - x \| + \| T(x_n - x) \|$

$$= \| x_n - x \| + \| T(x_n) - T(x) \|$$

$$= \| x_n - x \| + \| T(x_n - y) \| \to 0 \text{ as } n \to \infty. \qquad \text{(Using (1))}$$

Hence, the sequence (x_n) in $B_1 \to x \in B_1$ and consequently B_1 is complete.

This completes the proof of the theorem.

Theorem 3: Let B and B′ be Banach spaces and let $T : B \to B'$ be linear. If G_T is closed in $B \times B'$ and if T is one-one and onto, then T is a homeomorphism from B onto B′.

Proof: By closed graph theorem, T is continuous.

Let $T' = T^{-1} : B' \to B$. Then T′ is linear.

Further $(x, y) \in G_T \Leftrightarrow (y, x) \in G_{T'}$.

$\Rightarrow \quad G_{T'}$ is closed in $B' \times B$.

$\Rightarrow \quad$ T′ is continuous (By closed graph theorem)

$\Rightarrow \quad$ T is a homeomorphism on B onto B′.

This completes the proof of the theorem.

Theorem 4: Let a Banach space B be made into a Banach space B′ by a new norm. Then the topologies generated by these two norms are the same if either is stronger than the other.

Proof: Let the new norm on B′ be $\| \ \|'$. Let $\| \ \|$ is stronger than $\| \ \|'$. Then $\exists$ a constant k such that $\| x \| \le k \| x \|'$ for every $x \in B$.

Consider the identity map

$$I : B \to B'.$$

We claim that G_1 is closed.

Let $x_n \to x$ in B and $x_n \to y$ in B′.

Then $\| x \| \le k \| x \|' \Rightarrow \forall\ x \in B, I(x_n) = x_n \to y$ in $\| \ \|$ also.

Since a sequence cannot converge to two distinct points in $\| \ \|$, $y = x$. Consequently G_1 is closed.

Hence closed graph theorem, I is continuous. Therefore $\exists$ a k′s such that

$\| x \|' = \| I(x) \|' \le k' \| x \|$ for every $x \in B$. Hence $\| \ \|'$ is stronger than $\| \ \|$. Hence two topologies are same.

6.2 Summary

- Let N and N′ be a normal linear space and let $T : N \to N'$ be a mapping with domain N and range N′. The graph of T is defined to be a subset of $N \times N'$ which consist of all ordered pairs $(x, T(x))$. It is generally denoted by G_T.

- Let N and N′ be normed linear spaces and let M be a subspace of N. Then a linear transformation T : M → N′ is said to be closed iff $x_n \in M$, $x_n \to x$ and $T(x_n) \to y$ imply $x \in M$ and $y = T(x)$.
- If B and B′ are Banach spaces and if T is linear transformation of B into B′, then T is continuous ⇔ Graph of T (G_T) is closed.

6.3 Keyword

Closed Linear Transformation: Let N and N′ be normed linear spaces and let M be a subspace of N. Then a linear transformation

$$T : M \to N' \text{ is said to be closed}$$

iff $x_n \in M$, $x_n \to x$ and $T(x_n) \to y$ imply $x \in M$ and $y = T(x)$.

6.4 Review Questions

1. If X and Y are normed spaces and $A : X \to Y$ is a linear transformation, then prove that graph of A is closed if and only if whenever $x_n \to 0$ and $Ax_n \to y$, it must be that $y = 0$.
2. If P is a projection on a Banach space B, and if M and N are its range and null space, then prove that M and N are closed linear subspaces of B such that $B = M \oplus N$.

6.5 Further Readings

Books

Folland, Gerald B, *Real Analysis: Modern Techniques and their Applications (1st ed.)*, John Wiley & Sons, (1984).

Rudin, Walter, *Functional Analysis*, Tata McGraw-Hill (1973).

Online links

euclid.colorado.edu/ngwilkin/files/math6320.../OMT_CGT.pdf

mathworld.wolfram.com

Unit 7: The Conjugate of an Operator

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand the definition of conjugate of an operator.
- Understand theorems on it.
- Solve problems relate to conjugate of an operator.

Introduction

We shall see in this unit that each operator T on a normed linear space N induces a corresponding operator, denoted by T* and called the conjugate of T, on the conjugate space N*. Our first task is to define T* and our second is to investigate the properties of the mapping $T \to T^*$.

7.1 The Conjugate of an Operator

7.1.1 The Linear Function

Let N* be the linear space of all scalar-valued linear functions defined on N. Clearly the conjugate space N* is a subspace of N*. Let T be a linear transformation T′ of N* into itself as follows:

If $f \in N^+$, then T′ (f) is defined as

$$[T'(f)]x = f(T(x))$$

Since f (T (x)) is well defined, T′ is a well-defined transformation on N^+.

Theorem 1: Let $T' : N^+ \to N^+$ be defined as

$[T'(j)]\, x = f(T(x)),\ f \in N^+$, then

(a) T′ (j) is a linear junction defined on N.

(b) T′ is a linear mapping of N^+ into itself.

Notes

(c) $T'(N^*) \subseteq N^* \Rightarrow T$ is continuous, where T is a linear transformation of N into itself which is not necessarily continuous.

Proof:

(a) $x, y \in N$ and α, β be any scalars. Then

$[T'(f)](\alpha x + \beta y) = f(T(\alpha x + \beta y))$

Since T and f are linear, we get

$$f(T(\alpha x + \beta y)) = \alpha f(T(x) + \beta f(T(y))$$
$$= \alpha [T'(f)](x) + \beta [T'(f)]y$$
$$\Rightarrow \text{part (a).}$$

(b) Let $f, g \in N^+$ and α, β be any scalars. Then

$$[T'(\alpha f + \beta g)(x)] = (\alpha f + \beta g)(T(x)) = \alpha [T'(f)](x) + \beta (T'(g)](x)$$
$$\Rightarrow T' \text{ is linear on } N^+$$
$$\Rightarrow \text{part (b)}$$

(c) Let S be a closed unit sphere in N. Then we know that T is continuous $\Rightarrow$ T (S) is bounded

$\Rightarrow f(T(S))$ is bounded for each $f \in N^*$.

By definition of T', $f(T(S))$ is bounded if and only if $[T'(f)](S)$ is bounded for each f in

$$N^* = T'(f) \text{ is in } N^* \text{ for each f in } N^*.$$
$$\Rightarrow T'(N) \subseteq N^*$$
$$\Rightarrow \text{part (c)}$$

This completes the proof of the theorem.

Note: Part (c) of the above theorem enables us to restrict T' to N^* iff T is continuous. Hence by making T continuous we define an operation called the conjugate of T by restricting T' to N^*. We see it below.

7.1.2 The Conjugate of T

Definition: Let N be normed linear space and let T be a continuous linear transformation of N into itself (i.e. T is an operator). Define a linear transformation T^* of N^* into itself as follows:

If $f \in N^*$, then, $T^*(f)$ is given by

$[T^*(f)](x) = f(T(x))$

We call T^* the *conjugate of T*.

Theorem 2: If T is a continuous linear transformation on a normed linear space N, then its conjugate T^* defined by

$T^* : N^* \to N^*$ such that

$T^*(f) = f.T$ where

$[T^*(f)](x) = f(T(x)) \;\; \forall f \in N^*$ and all $x \in N$

is a continuous linear transformation on N^* and the mapping $T \to T^*$ given by

$\phi : \beta(N) \to \beta(N^*)$ such that

Notes

ϕ (T) = T* for every β (N)

is an isometric isomorphism of b (N) into b (N*) reverses products and preserves the identify transformation.

Proof: We first show that T* is linear

Let f, g $\in$ N* and α, $\in$ be any scalars

then $[T^* (\alpha j + \beta g)] (x) = (\alpha j + \beta g) (T (x))$

$= (\alpha j) T (x) + (\beta g) T (x)$

$= \alpha [j (T (x))] + \beta [g (T (x))]$

$= \alpha [T^* (j)] (x) + \beta [T^* (g)] (x)$

$= [\alpha T^* (j) + \beta T^* (g)] (x) \ \forall \ x \in N$

Hence $T^* (\alpha f + \beta g) = \beta T^* (f) + \beta T^* (g)$

$\Rightarrow$ T* is linear on N*.

To show that T* is continuous, we have to show that it is bounded on the assumption that T is bounded.

$$\| T^* \| = \sup \{ \| T^* (f) \| : \| f \| \leq 1\}$$

$$= \sup \{ \| [T^* (f)] (x) \| : \| f \| \leq 1 \text{ and } \| x \| \leq \}$$

$$= \sup \{ | f (T (x)) | : \| f \| \leq 1 \text{ and } \| x \| \leq \}$$

$$= \sup \{ \| f \| \| T \| x \| : \| x \| \leq 1 \text{ and } \| x \| \leq \} \quad \ldots (1)$$

$$\leq T$$

$\Rightarrow$ T* is a bounded linear transformation on N* into N*. Hence by application of Hahn-Banach theorem, for each non-zero x in N, $\exists$ a functional f $\in$ N* such that

$$\| f \| = 1 \text{ and } f (T (x)) = \| T (x) \| \quad \ldots (2)$$

Hence $$\| T \| = \sup \left\{ \frac{\| T(x) \|}{\| x \|} : x \neq 0 \right\}$$

$$= \sup \left\{ \frac{\| f(T(x)) \|}{\| x \|} : \| f \| = 1,\ x \neq 0 \right\} \quad \text{(by (2))}$$

$$= \sup \left\{ \frac{\| T^*(f)(x) \|}{\| x \|} : \| f \| = 1,\ x \neq 0 \right\} \quad \text{(by (1))}$$

$$\leq \sup \left\{ \frac{\| T^*(f) \| \ \| x \|}{\| x \|} : \| f \| = 1,\ x \neq 0 \right\}$$

$$= \sup \left\{ \| T^*(f) \| : \| f \| = 1 \right\} = \| T^* \| \quad \ldots (3)$$

From (1) and (3) it follows that

$$\| T \| = \| T^* \|. \quad \ldots (4)$$

Notes

Now we show that

$$\begin{Bmatrix} \phi : \beta(N) \to \beta(N^*) \text{ given by} \\ \phi(T) = T^* \end{Bmatrix} \quad \ldots (5)$$

for every $T \in \beta$ (N) is an isometric isomorphism which reverses the product and preserve the identity transformation.

The isometric character of ϕ follows by using (5) as seen below:

$$\| \phi (T) \| = \| T^* \| = \| T \|.$$

Next we show that ϕ is linear and one-to-one. Let T, $T_1 \in \beta$ (N) and α, β be any scalars. Then

$$\phi (\alpha T + \beta T_1) = (\alpha T + \beta T)^* \quad \text{by (3)}$$

But $$[(\beta T + \beta T_1)^* (f)] (x) = f (\alpha T + \beta T_1) (x)$$

$$= f (\alpha T (x) + \beta T_1 (x))$$

Since f is linear, we get

$$[(\alpha T + \beta T_1)^* (f)] = \alpha f (T (x)) + \beta f (T_1 (x))$$

$$= \alpha [T^* (f) (x) + \beta T^*_1 (f) (x)]$$

$$= \{\alpha[T^*(f)]\} + \beta[T^*_1 (f)](x)$$

$\forall$ x $\in$ N. Hence we get

$$[(\alpha T + \beta T_1)^* (f)] = \alpha [T^* (f)] + \beta [T^*_1 (f)]$$

$$= (\alpha T^* + \beta T^*_1)(f)$$

Hence $$(\alpha T + \beta T_1)^* = \alpha T^* + \beta T_1^* \quad \ldots (6)$$

Therefore $$\phi (\alpha T + \beta T_1) = (\alpha T + \beta T_1)^* = \alpha T^* + \beta T^*_1 = \alpha \phi (T) + \beta \phi (T_1)$$

$\Rightarrow \phi$ is linear.

To show ϕ is one-to-one, let $\phi (T) = \phi (T_1)$

Then $T^* = T^*_1$

$\Rightarrow \| T^* - T^*_1 \| = 0$

Using (6) by choosing $\alpha = 1$, $\beta = -1$ we get

$$\| (T - T_1)^* \| = 0 \Rightarrow \| T - T_1 \| = 0 \text{ or } T = T_1.$$

$\Rightarrow \phi$ is one-to-one.

Hence ϕ is an isometric isomorphism on β (N) onto β (N*).

Finally we show that ϕ reverses the product and preserves the identity transformation.

Now $$[(T T_1)^* (f)] (x) = f ((T T_1) (x))$$

$$= f (T (T_1 (x))$$

$$= [T^* (f)] [T_1 (x)], \text{ since } T_1 (x) \in N \text{ and } T^* (f) \in N^*.$$

Notes

$$= [T^*_1 (T^* (f))] (x)$$

$$= [(T^*_1 \; T^*) (f)] (x)$$

Hence, we get

$$(T T_1)^* = T^*_1 \; T^* \text{ so that}$$

$$\phi (T T_1) = (T T_1)^* = T^*_1 \; T.$$

$\Rightarrow \phi$ reverses the product.

Lastly if I is the identity operator on N, then

$$[I^* (f)] (x) = f (I (x)) = f (x) = (I f) (x).$$

$\Rightarrow I^* = I$ so that $\phi (I) = I^* = I$

$\Rightarrow \phi$ preserves the identity transformation.

This completes the proof of the theorem.

Theorem 3: Let T be an operator on a normal linear space N. If $N \subset N^*$ in the natural imbedding, then T** is an extension of T. If N is reflexive, then T** = T.

Proof: By definition, we have

$$(T^*)^* = T^{**}$$

Using theorem 2, we have $\| T^* \| = \| T \|$.

Hence $\| T^{**} \| = \| T^* \| = \| T \|$.

By definition of conjugate of an operator

$$T : N \to N, T^* : N^* \to N^*, T^{**} : N^{**} \to N^{**}.$$

Let $J ; x \to F_x$ be the natural imbedding of N onto N** so that

$$Fx (f) = f (x) \text{ and } J (x) = F_x.$$

Further, since T** is the conjugate operator of T*, we get

$$T^{**} (x'') x' = x'' (T^* (x')) \text{ where } x' \in N^* \text{ and } x'' \in N^{**}$$

$$T^{**} (x'') x' = T^{**} (J (x)) x'.$$

Using the definition of conjugate, we get

$$T^{**} (J (x)) x' = J (x) (T^* (x')).$$

By definition of canonical imbedding

$$J (x) (T^* (x')) = T^* (x') x.$$

$$\text{Again } T^* (x') (x) = x' (T (x)) \qquad \text{(By definition of conjugate)}$$

$$\text{Now } x' (T (x)) = J (T (x))x' \qquad \text{(By natural imbedding)}$$

Hence $$T^{**} (J (x))x' = J (T (x))x'.$$

$\Rightarrow$ $$T^{**} . J = J T$$

and so T** is the norm preserving extension of T. If N is reflexive, N = N** and so T** coincides with T.

Notes

This completes the proof of the theorem.

Theorem 4: Let T be an operator on a Banach space B. Then T has an inverse T–1 ⇔ T* has an inverse $(T^*)^{-1}$, and

$$(T^*)–1 = (T–1)^*$$

Proof: T has inverse $T^{-1} \Leftrightarrow TT^{-1} = T^{-1} T = I$

By theorem 2, the mapping $\phi : T \to T^*$ reverse the product and preserves the identity

$$\therefore \qquad (TT^{-1})^* = (T^{-1} T)^* = I^*$$

$$(T^{-1})^*T = T^* (T^{-1})^* = I$$

$\Rightarrow (T^*)^{-1}$ exists and it is given by $(T^*)^{-1} = (T^{-1})^*$. This completes the proof of the theorem.

7.2 Summary

- Let N^+ be the linear space of all scalar-valued linear functions defined on N. Clearly the conjugate space N* is a subspace of N^+. Let T be a linear transformation T′ of N^+ into itself as follows:

  If $f \in N^+$, then T′ (f) is defined as

  T′ (f) x = f (T (x))

- Let N be a normed linear space and let T be a continuous linear transformation of N into itself. Define a linear transformation T* of N* into itself as follows:

  If $f \in N^+$, then T′ (f) is given by

  T′ (f) x = f (T (x))

  We call T* the conjugate of T.

7.3 Keywords

The Conjugate of T: Let N be normed linear space and let T be a continuous linear transformation of N into itself (i.e. T is an operator). Define a linear transformation T* of N* into itself as follows:

If $f \in N^*$, then, T* (f) is given by

[T* (f)] (x) = f (T (x))

We call T* the *conjugate of T*.

The Linear Function: Let N* be the linear space of all scalar-valued linear functions defined on N. Clearly the conjugate space N* is a subspace of N*. Let T be a linear transformation T′ of N* into itself as follows:

If $f \in N^+$, then T′ (f) is defined as

$$[T' (f)]x = f (T (x))$$

Since f (T (x)) is well defined, T′ is a well-defined transformation on N^+.

7.4 Review Questions

1. Let B be a Banach space and N a normed linear space. If $\{T_n\}$ is a sequence in B (B, N) such that $T(x) = \lim T_n (x)$ exists for each x in B, prove that T is a continuous transformation.

2. Let T be an operator on a normed linear space N. If N is considered to be part of N** by means of the natural imbedding. Show that T** is an extension of T. Observe that if N is reflexive, then T** = T.

3. Let T be an operator on a Banach space B. Show that T has an inverse $T^{-1} \Leftrightarrow T^*$ has an inverse $(T^*)^{-1}$, and that in this case $(T^*)^{-1} = (T^{-1})^*$.

7.5 Further Readings

Books

James Wilson Daniel, *The Conjugate Gradient Method for Linear and Non-linear Operator Equations*.

G.O. Okikiolu, *Special Integral Operators: Poisson Operators, Conjugate Operators and related Integrals.* Vol. Okikiolu Scientific and Industrial Organization, 1981.

Online links

epubs.siam-org/sinum/resource/1/sjnamm/v9/i|/p165_s|

www.ima.umm.edu

Notes

Unit 8: The Uniform Boundedness Theorem

CONTENTS

Objectives

After studying this unit, you will be able to:

- State the uniform boundedness theorem.
- Understand the proof of this theorem.
- Solve problems related to uniform boundedness theorem.

Introduction

The uniform boundedness theorem, like the open mapping theorem and the closed graph theorem, is one of the cornerstones of functional analysis with many applications. The open mapping theorem and the closed graph theorem lead to the boundedness of T^{-1} whereas the uniform boundedness operators deduced from the point-wise boundedness of such operators. In uniform boundedness theorem we require completeness only for the domain of the definition of the bounded linear operators.

8.1 The Uniform Boundedness Theorem

8.1.1 The Uniform Boundedness Theorem – Proof

If (a) B is a Banach space and N a normed linear space,

(b) $\{T_i\}$ is non-empty set of continuous linear transformation of B into N, and

(c) {Ti (x)} is a bounded subset of N for each $x \in B$, then $\{ \| T_i \| \}$ is a bounded set of numbers, i.e. $\{T_i\}$ is bounded as a subset of β (B, N)

Proof: For each positive integer n, let

$$F_n = \{x \in B : \| T_i (x) \| \leq n \ \forall \ i\}.$$

Then F_n is a closed subset of B. For if y is any limit point of F_n, then $\exists$ a sequence (x_k) of points of F_n such that

$$x_k \to y \text{ as } k \to \infty$$

$\Rightarrow \quad T_i x_k \to T_{iy}$ as $k \to \infty$ (By continuity of T_i)

$\Rightarrow \quad \| T_i x_k \| \to \| T_{iy} \|$ as $k \to \infty$ (By continuity of norm)

$\Rightarrow \quad \| T_i y \| = \underset{k\to\infty}{\text{lt}} \| T_i x_k \|$

$\leq n \ \forall \ i$ $\quad (\because x_k \in F_n)$

$\Rightarrow \quad y \in F_n.$

Thus F_n contains all its limit point and is therefore closed. Further, if x is any element of B, then by hypothesis (c) of the theorem ∃ a real number k ≥ 0 s.t.

$$\| T_i x \| \leq k \ \forall \ i$$

Let n be a positive integer s.t. n > k. Then

$$\| T_i x \| < n \ \forall \ i$$

so that $x \in F_n$.

Consequently, we have $B = \bigcup_{n=1}^{\infty} F_n$.

Since B is complete, it therefore follows by Baire's theorem that closure of some F_n, say $\overline{F}_{n_o} = F_{n_o}$, possesses an interior point x_o. Thus we can find a closed sphere S_o with centre x_o and radius r_o such that $S_o \subseteq F_{n_o}$.

Now if y is any vector in T_i (S_o), then

$$y = T_{is_o}$$

where $\quad s_o \in S_o \subseteq F_{n_o}$.

$\therefore \quad \| y \| = \left\| T_i s_o \right\| \leq n_o.$

Thus norm of every vector in T_i (S_o) is less than or equal to n_o. We write this fact as $\| T_i (S_o) \| \leq n_o$.

Let $S = \dfrac{S_o - x_o}{r_o}$. Then S is a closed unit sphere centred at the origin in B and

$$\| T_i (S) \| = \left\| T_i \left(\frac{S_o - x_o}{r_o} \right) \right\|$$

$$= \frac{1}{r_o} \left\| T_i(S_o) - T_i(x_o) \right\|$$

$$\leq \frac{1}{r_o} \left(\left\| T_i(S_o) \right\| + \left\| T_i(x_o) \right\| \right)$$

$$\leq \frac{2n_o}{r_o} \forall i.$$

Hence $\quad \| T_i \| \leq \dfrac{2n_o}{r_o} \forall i.$

This completes the proof of the theorem.

8.1.2 Theorems and Solved Examples

Theorem 1: If B is a Banach space and (f_i (x)) is sequence of continuous linear functionals on B such that (| f_i (x) |) is bounded for every x ∈ B, then the sequence (|| f_i ||) is bounded.

Proof: Since the proof of the theorem is similar to the theorem (1), however we briefly give its proof for the sake of convenience to the readers.

For every m, let $F_m \subset B$ be the set of all x such that $|f_n(x)| \le m \;\forall\; n$

$$|f_n(x)| \le m \;\forall\; n$$

Now F_m is the intersection of closed sets and hence it is closed.

As in previous theorem, we have

$B = \bigcup_{m=1}^{\infty} F_m$. Since B is complete. It is of second category. Hence by Baire's theorem, there is a $x_o \in F_m$ and a closed sphere S [x_o, r_o] such that $|f_n(x)| \le m \;\forall\; n$.

Let x be a vector with $\|x\| \le r_o$.

Now $\quad f_n(x) = f_n(x + x_o - x_o)$

$$= f_n(x + x_o) - f_n(x_o)$$

$\therefore \quad |f_n(x)| \le |f_n(x + x_o)| + |f_n(x_o)| \quad \ldots (1)$

Since $\quad \|x + x_o - x_o\| = \|x\| < r_o$, we have $(x + r_o) \in S[x_o, r_o]$

$\therefore \quad |f_n(x + x_o)| \le m \quad \ldots (2)$

Also we have $\quad |f_n(x_o)| \le k \;\forall\; n \quad \ldots (3)$

From (1), (2) and (3), we have for $\forall\; x \in S[x_o, r]$.

$$|f_n(x)| < (m + k) \;\forall\; n.$$

Now for x ∈ B, consider the vector $\frac{r_o x}{\|x\|}$.

Then $|f_n(x)| = \left|\frac{\|x\|}{r_o} f_n\left(\frac{r_o x}{\|x\|}\right)\right| \le \frac{\|x\|}{r_o}(m+k)$ so that $\frac{|f_n(x)|}{\|x\|} \le \left(\frac{m+k}{r_o}\right)$.

In other words,

$$\|f\| \le \left(\frac{m+k}{r_o}\right).$$

This completes the proof of the theorem.

Example 1: Show that the completeness assumption in the domain of (T_i) in the uniform boundedness theorem cannot be dropped.

Solution: Consider N= space of all polynomial x

$$= x(t) = \sum_{n=0}^{\infty} a_n t_n \,,\; a_n \ne 0$$

for finitely many n's.

Notes

It we define the norm on N as

$$\|x\| = \max\{|a_n|, n = 0, 1, 2, \ldots\}$$

then N is an incomplete normed linear space.

Now define $f_n(x) = \sum_{k=0}^{n-1} a_k$, n = 1, 2, …

The functions $\{f_n\}$ are continuous linear functional on N.

If we take $x = a_0 + a_1 t + \ldots\ldots + a_m t^m$ then

$$|f_n(x)| \le (m+1)\max\{|a_k|\} = (m+1)\|x\|,$$

so that $(|f_n(x)|)$ is point-wise bounded.

Now consider $x = 1 + t + t^2 + \ldots + t^{n-1}$. Then $\|x\| = 1$ and from the definition of $|f_n(x)| = n$.

Hence $\|f_n\| \ge \frac{|f_n(x)|}{\|x\|} = n$.

$\Rightarrow (\|f_n\|)$ is unbounded.

Thus if we drop the condition of completeness in the domain of (T_i), the uniform boundedness theorem is not true anymore.

Theorem 2: Let N be a normed linear space and B be a Banach space. If a sequence $(T_n) \in \beta(B, N)$ such that $T(x) = \lim T_n(x)$ exists for each x in B, then T is a continuous linear transformation.

Proof: T is linear.

$$\begin{aligned} T(\alpha x + \beta y) &= \lim T_n(\alpha x + \beta y) \\ &= \lim \{T_n(\alpha x) + T_n(\beta y)|\} \\ &= \alpha \lim T_n(x) + \beta \lim T_n(y) \\ &= \alpha T(x) + \beta T(y) \text{ for } x, y \in B \text{ and for any scalars } \alpha \text{ and } \beta. \end{aligned}$$

since $\lim T_n(x)$ exists, $(T_n(x))$ is a convergent sequence in N. Since convergent sequences are bounded, $(T_n(x))$ is point-wise bounded.

Hence by uniform bounded theorem, $(\|T_n\|)$ is bounded so that ∃ a positive constant λ such that

$$\|T_n\| \le \lambda \ \forall\ n.$$

Now $\|T_n(x)\| \le \|T_n\|\ \|x\| \le \lambda\|x\|$.

Since $T_n(x) \to T(x)$, we have

$$\|T(x)\| \le \lambda\|x\|$$

⇒ T is bounded (continuous) linear transformation. This completes the proof of the theorem.

Corollary 1: If f is a sequence in B* such that $f(x) = \lim_{n\to\infty} f_n(x)$ exists for each $x \in B$, then f is continuous linear functional on B.

Example 2: Let (a_n) be a sequence of real or complex numbers such that for each $x = (x_n) \in c_o$, $\sum_{n=1}^{\infty} a_n x_n$ converges. Prove that $\sum_{n=1}^{\infty} |a_n| < \infty$.

Solution: For every x ∈ c_o, let $f_n = \sum_{i=1}^{n} a_i x_i$. Since each $\sum_{i=1}^{n} a_i x_i$ is a finite sum of scalars, (f_n) is sequence of continuous linear functional on c_o. Let f (x) = $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \sum_{i=1}^{n} a_i x_i$. By cor. 1, f (x) exists and bounded. $\|f\| = \sum_{n=1}^{\infty} |a_n|$. Since ‖ f ‖ is bounded, $\sum_{n=1}^{\infty} |a_n| < \infty$.

Theorem 3: A non-empty subset X of a normed linear space N is bounded ⇔ f (X) is a bounded set of numbers for each f in N*.

Proof: Let X be a bounded subset of N so that ∃ a positive constant λ_1 such that

$$\|x\| \le \lambda_1 \qquad \forall\ x \in X \qquad \dots (1)$$

To show that f (X) is bounded for each f ∈ N*. Now f ∈ N* ⇒ f is bounded.

$$\Rightarrow \qquad \exists\ \lambda_2 > 0 \text{ such that } |f(x)| \le \lambda_2 \|x\| \qquad \forall\ x \in N \qquad \dots (2)$$

It follows from (1) & (2) that

$$|f(x)| \le \lambda_1 \lambda_2 \qquad \forall\ x \in X.$$

⇒ f (X) is a bounded set of real numbers for each f ∈ N*.

Conversely, let us assume that f (X) is a bounded set of real numbers for each f ∈ N*.

To show that X is bounded. For convenience, we exhibit the vectors in X by writing X = {x_i}. We now consider the natural imbedding J from N to N** given by

$$J : x_i \to F_{x_i}$$

From the definition of this natural imbedding, we have

$$F_x(f) = f(x) \text{ for each } x \in N.$$

Hence our assumption f (X) = {f (x_i)} is bounded for each f ∈ N* is equivalent to the assumption that $\{F_{x_i}(f)\}$ is bounded set for each f ∈ N*.

Since N* is complete ⇒ $\left(F_{x_i}\right)$ is bounded subset of N** by uniform boundedness theorem.

That is, $\left(\|F_{x_i}\|\right)$ is a bounded set of numbers. Since the norms are preserved in natural imbedding, we have $\|F_{x_i}\| = \|x_i\|$ for every $x_i \in X$.

Therefore ($\|x_i\|$) is a bounded set of numbers. Hence is bounded subset of N_i.

This completes the proof of the theorem.

Theorem 4: Let N and N′ be normed linear space A linear transformation.

$$T : N \to N' \text{ is continuous} \Leftrightarrow \text{for each } f \in N^*,\ f \circ T \in N^*.$$

Proof: We first note that f o T is linear. Also f o T is well defined, since T (x) ∈ N′ for every x ∈ N and f is a functional on N′ so that f (T (x)) is well defined and f o T ∈ N*. Since T is continuous and f is continuous, f o T is continuous on N.

Conversely, let us assume that f o T is continuous for each $f \in N^*$. To show that T is continuous it suffices to show that

$$T(B) = \{Tx : x \in N, B = \| x \| \leq 1\} \text{ is bounded in } N'.$$

For each $f \in N^*$, f o T is continuous and linear on N and so (f o T) B = f (T(B)) is bounded set for every $f \in N^*$, where we have considered the unit sphere B with centre at the origin and radius 1. Since any bounded set in N can be obtained from B, T (B) is bounded by a non-empty subset X of a normed linear space N of bounded $\Leftrightarrow$ f (X) is a bounded set of number for each f in N^*.

8.2 Summary

- ***Uniform Boundedness Theorem:*** If (a) B is Banach space and N a normed linear space, (b) $\{T_i\}$ is non-empty set of continuous linear transformation of B into N and (c) $\{T_i(x)\}$ is a bounded subset of N for each $x \in B$, then $\{\| T_i \|\}$ is a bounded set of numbers, i.e. $\{T_i\}$ is bounded as a subset of β (B, N).
- If B is a Banach space and $(f_i(x))$ is sequence of continuous linear functionals on B such that $(|f_i(x)|)$ is bounded for every $x \in B$, then the sequence $(\| T_i \|)$ is bounded.

8.3 Keywords

Imbedding: Imbedding is one instance of some mathematical structure contained within another instance, such as a group that is a subgroup.

Uniform Boundedness Theorem: The uniform boundedness theorem, like the open mapping theorem and the closed graph theorem, is one of the cornerstones of functional analysis with many applications.

8.4 Review Questions

1. If X is a Banach space and $A \subseteq X^*$, then prove that A is a bounded set if and only if for every x in X, $\text{Sup}\{|f(x)| : f \in A\} < \infty$.
2. Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{E}$ be an orthonormal basis for $\mathcal{H}$. Show that a sequence $\{h_n\}$ in $\mathcal{H}$ satisfies $\langle h_n, h\rangle \to 0$ for every h in $\mathcal{H}$ if and only if $\sup\{\| h_n \| : n \geq 1\} < \infty$ and $\langle h_n, e\rangle \to 0$ for every e in $\mathcal{E}$.

8.2.5 Further Readings

Books

Bourbaki, Nicolas, *Topological vector spaces, Elements of mathematics*, Springer (1987).

Diendonne, Jean, *Treatise on Analysis, Volume 2,* Academic Press, (1970).

Rudin, Walter, *Real and Complex Analysis,* McGraw-Hill, 1966.

Online links

www.jstor.org/stable/2035429

www.sciencedirect.com/science/article/pii/S0168007211002004

Unit 9: Hilbert Spaces: The Definition and Some Simple Properties

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define inner product spaces.
- Define Hilbert space.
- Understand basic properties of Hilbert space.
- Solve problems on Hilbert space.

Introduction

Since an inner product is used to define a norm on a vector space, the inner product are special normed linear spaces. A complete inner product space is called a Hilbert space. We shall also see from the formal definition that a Hilbert space is a special type of Banach space, one which possesses additional structure enabling us to tell when two vectors are orthogonal. From the above information, one can conclude that every Hilbert space is a Banach space but not conversely in general.

We shall first define Inner Product spaces and give some examples so as to understand the concept of Hilbert spaces more conveniently.

9.1 Hilbert Spaces

9.1.1 Inner Product Spaces

Definition: Let X be a linear space over the field of complex numbers C. An inner product on X is a mapping from $X \times X \to C$ which satisfies the following conditions:

(i) $(\alpha x + \beta y, z) = \alpha\,(x, z) + \beta\,(y, z),\ \forall\, x, y, z \in X$ and $\alpha, \beta \in C$. (Linearity property)

(ii) $\overline{(x, y)} = (y, x)$ (Conjugate symmetry)

(iii) $(x, x) \geq 0,\ (x, x) = 0 \Leftrightarrow x = 0$ (Non-negativity)

A complex inner product space X is a linear space over C with an inner product defined on it.

Notes

1. We can also define inner product by replacing C by R in the above definition. In that case, we get a real inner product space.
2. It should be noted that in the above definition (x, y) does not denote the ordered pair of the vectors x and y. But it denotes the inner product of the vectors x and y.

Theorem 1: If X is a complex inner product space then

(a) $(\alpha x - \beta y, z) = \alpha\,(x, z) - \beta\,(y, z)$

(b) $(x, \beta y + \gamma z) = \bar{\beta}\,(x, y) + \bar{\gamma}\,(x, z)$

(c) $(x, \beta y - \gamma z) = \bar{\beta}\,(x, y) - \bar{\gamma}\,(x, z)$

(d) $(x, 0)$ and $(0, x) = 0$ for every $x \in X$.

Proof: (a)

$$(\alpha x - \beta y, z) = (\alpha x + (-\beta)\, y, z) = \alpha\,(x, z) + (-\beta)\,(y, z) = \alpha\,(x, z) - \beta\,(y, z).$$

(b)

$$(x, \beta y + \gamma z) = \overline{(\beta y + \gamma z, x)} = \overline{(\beta y, x) + (\gamma z, x)} = \overline{\beta\,(y, x) + \gamma(z, x)} = \bar{\beta}\,(x, y) + \bar{\gamma}(x, z)$$

(c)

$$(x, \beta y - \gamma z) = (x, \beta y + (-\gamma)\, z) = \bar{\beta}\,(x, y) + \overline{(-\gamma)}\,(x, z) = \bar{\beta}\,(x, y) - \bar{\gamma}\,(x, z)$$

(d)

$$(0, x) = (0\theta, x) = 0\,(\theta, x) = 0,$$ where θ is the zero element of x and $\overline{(x, 0)} = (\bar{0}, x) = \bar{0} = 0$.

Further note that $(x, y + z) = (x, |y + 1|z) = \bar{1}\,(x, y) + \bar{1}\,(x, z)$

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

This completes the proof of the theorem.

Notes

1. Part (b) shows an inner product is conjugate linear in the second variable.
2. If (x, y) = 0 $\forall$ x $\in$ X, then y = 0. If (x, y) = 0 $\forall$ x $\in$ X, it should be true for x = y also, so that (y, y) = 0 $\Rightarrow$ y = 0.

Example 1: The space ℓ_2^n is an inner product space.

Solution: Let x = $(x_1, x_2, \ldots\ldots, x_n)$, y = $(y_1, y_2, \ldots\ldots, y_n) \in \ell_2^n$.

Define the inner product on ℓ_2^n as follows:

$$(x, y) = \sum_{i=1}^{n} x_i \bar{y}_i$$

Now

(i)
$$(\alpha x + \beta y, z) = \sum_{i=1}^{n} (\alpha x_i + \beta \bar{y}_i)\bar{z}_i$$
$$= \sum_{i=1}^{n} \alpha x_i \bar{z}_i + \sum_{i=1}^{n} \beta y_i \bar{z}_i$$
$$= \alpha (x, z) + \beta (y, z)$$

(ii)
$$\overline{(x, y)} = \left\{ \overline{\sum_{i=1}^{n} x_i \bar{y}_i} \right\}$$
$$= \overline{(x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n)}$$
$$= (\bar{x}_1\bar{\bar{y}}_1 + \bar{x}_2\bar{\bar{y}}_2 + \ldots + \bar{x}_n\bar{\bar{y}}_n)$$
$$= \bar{x}_1 y_1 + \bar{x}_2 y_2 + \ldots + \bar{x}_n y_n$$
$$= (y, x)$$

(iii)
$$(x, x) = \sum_{i=1}^{n} x_i \bar{x}_i$$
$$= \sum_{i=1}^{n} |x_i|^2 \geq 0$$

Hence (x, x) $\geq$ 0 and (x, x) = 0 $\Leftrightarrow$ x_i = 0 for each i, i.e. (x, x) = 0 $\Leftrightarrow$ x = 0.

(i) – (iii) $\Rightarrow$ ℓ_2^n is a inner product space.

9.1.2 Hilbert Space and its Basic Properties

By using the inner product, on a linear space X we can define a norm on X, i.e. for each x $\in$ X, we define $\| x \| = \sqrt{(x, x)}$. To prove it we require the following fundamental relation known as Schwarz inequality.

Theorem 2: If x and y are any two vectors in an inner product space then

$$|(x, y)| \le \|x\| \, \|y\| \qquad \ldots (1)$$

Proof: If y = 0, we get $\|y\| = 0$ and also theorem 1 implies that $|(x, y)| = 0$ so that (1) holds.

Now, let $y \ne 0$, then for any scalar $\lambda \varepsilon C$ we have

$$0 \le \|x - \lambda y\|^2 = (x - \lambda y, x - \lambda y)$$

But

$$\begin{aligned}(x - \lambda y, x - \lambda y) &= (x, x) - (x, \lambda y) - (\lambda y, x) + (\lambda y, \lambda y)\\ &= (x, x) - \bar{\lambda}\,(x, y) - \lambda\,(y, x) + \lambda\bar{\lambda}\,(y, y)\\ &= \|x\|^2 - \lambda\,(y, x) - (x, y) + |\lambda|^2\,\|y\|^2\end{aligned}$$

$$\begin{aligned}\Rightarrow \|x\|^2 - \lambda\,(y, x) - \bar{\lambda}\,(x, y) + |\lambda|^2\,\|y\|^2 &= (x, \lambda y, x - \lambda y)\\ &= \|x - \lambda y\|^2 \ge 0 \qquad \ldots (2)\end{aligned}$$

Choose $\lambda = \dfrac{(x,y)}{\|y\|^2}, \; y \ne 0, \|y\| \ne 0.$

$\therefore$ We get from (2)

$$\|x\|^2 - \frac{(x,y)\,\overline{(x\,y)}}{\|y\|^2} - \frac{\overline{(x,y)}}{\|y\|^2}(x,y) + \frac{|(x,y)|^2}{\|y\|^4}\|y\|^2 \ge 0$$

$$\Rightarrow \quad \|x\|^2 - \frac{|(x,y)|^2}{\|y\|^2} - \frac{|(x,y)|^2}{\|y\|^2} + \frac{|(x,y)|^2}{\|y\|^2} \ge 0$$

$$\Rightarrow \quad \|x\|^2 - \frac{|(x,y)|^2}{\|y\|^2} \ge 0$$

$$\Rightarrow \quad \|x\|^2 \, \|y\|^2 \ge |(x, y)|^2$$

or $\quad |(x, y)| \le \|x\| \, \|y\|.$

This completes the proof of the theorem.

Theorem 3: If X is an inner product space, then $\sqrt{(x,x)}$ has the properties of a norm, i.e. $\|x\| = \sqrt{(x,x)}$ is a norm on X.

Proof: We shall show that $\| \;\|$ satisfies the condition of a norm.

(i) $\quad \|x\| = \sqrt{(x,x)} \Rightarrow \|x\|^2 = (x, x) \ge 0$ and $\|x\| = 0 \Leftrightarrow x = 0.$

(ii) $\quad$ Let $x, y \in X$, then

$$\begin{aligned}\|x + y\|^2 &= (x + y, x + y)\\ &= (x, x) + (x, y) + (y, x) + (y, y) \qquad \ldots (1)\\ &= \|x\|^2 + (x, y) + \overline{(x, y)} + \|y\|^2\\ &= \|x\|^2 + 2\mathrm{Re}\,(x, y) + \|y\|^2 \qquad [\because (x,y) + \overline{(x,y)} = 2\mathrm{Re}(x,y)]\\ &\le \|x\|^2 + 2\,|(x, y)| + \|y\|^2 \qquad [\because \mathrm{Re}\,(x, y) \le |(x, y)|]\end{aligned}$$

$\leq \| x \|^2 + 2 \| x \| \; \| y \| + \| y \|^2$ [using Schwarz inequality]

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

Therefore $\| x + y \| \leq \| x \| + \| y \|$

(iii) $\| \alpha x \|^2 = (\alpha x, \alpha x)$

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

$= |\alpha|^2 \; \| x \|^2$

$\Rightarrow$ $\| \alpha, x \| = \| \alpha \| \, \| x \|$

(i)-(iii) imply that $\| x \| = \sqrt{(x, x)}$ is a norm on X. This completes the proof of the theorem.

> *Note* Since we are able to define a norm on X with the help of the inner product, the inner product space X consequently becomes a normed linear space. Further if the inner product space X is complete in the above norm, then X is called a Hilbert space.

9.1.3 Hilbert Space: Definition

A complete inner product space is called a Hilbert space.

Let H be a complex Banach space whose norm arises from an inner product which is a complex function denoted by (x, y) satisfying the following conditions:

H_1 : $(\alpha x + \beta y, 2) = \alpha (x, 2) + \beta (y, 2)$,

H_2 : $\overline{(x, y)} = (y, x)$, and

H_3 : $(x, x) = \| x \|^2$,

for all x, y, z $\in$ H and for all $\alpha, \beta \in$ C.

9.1.4 Examples of Hilbert Space

1. The space ℓ_2^n is a Hilbert space.

   We have already shown in earlier example that ℓ_2^n is an inner product space. Also ℓ_2^n is a Banach space. Consequently ℓ_2^n is a Hilbert space. Moreover ℓ_2^n, being a finite dimensional, hence ℓ_2^n is a finite dimensional Hilbert space.

2. ℓ_2 is a Hilbert space.

   Consider the Banach space ℓ_2 consisting of all infinite sequence $x = (x_n)$, n = 1, 2, … of complex numbers such that $\sum_{n=1}^{\infty} |x_n|^2 < \infty$ with norm of a vector $x = (x_n)$ defined by $\| x \| =$

   $\left\{ \sum_{n=1}^{\infty} |x_n|^2 \right\}$.

Notes

We shall show that if the inner product of two vectors $x = (x_n)$ and $y = (y_n)$ is defined by $(x, y) = \sum_{n=1}^{\infty} x_n \overline{y}_n$, then ℓ_2 is a Hilbert space.

We first show that inner product is well defined. For this we are to show that for all x, y in ℓ_2 the infinite series $\sum_{n=1}^{\infty} x_n \overline{y}_n$ is convergent and this defines a complex number.

By Cauchy inequality, we have

$$\sum_{i=1}^{n} |x_i \overline{y}_i| \le \left\{\sum_{i=1}^{n} |x_i|^2\right\}^{\frac{1}{2}} \left\{\sum_{i=1}^{n} |y_i|^2\right\}^{\frac{1}{2}}$$

$$\le \left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{\frac{1}{2}} \left\{\sum_{n=1}^{\infty} |y_n|^2\right\}^{\frac{1}{2}}.$$

Since $\sum_{n=1}^{\infty} |x_n|^2$ and $\sum_{n=1}^{\infty} |y_n|^2$ are convergent, the sequence of partial sum $\sum_{i=1}^{n} |x_i \, \overline{y}_i|$ is a monotonic increasing sequence bounded above. Therefore, the series $\sum_{n=1}^{\infty} |x_i \, \overline{y}_i|$ is convergent. Hence $\sum_{n=1}^{\infty} |x_n \, \overline{y}_n|$ is absolutely convergent having its sum as a complex number.

Therefore $(x, y) = \sum_{n=1}^{\infty} x_n \, \overline{y}_n$ is convergent so that the inner product is well defined. The condition of inner product can be easily verified as in earlier example.

Theorem 4: If x and y are any two vectors in a Hilbert space, then

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

Proof: We have for any x and y

$$\| (x + y) \|^2 = (x + y, x + y) \qquad \text{(By def. of Hilbert space)}$$

$$= (x, x + y) + (y, x + y)$$

$$= (x, x) + (x, y) + (y, x) + (y, y)$$

$$= \| x \|^2 + (x, y) + (y, x) + \| y \|^2 \qquad \dots (1)$$

$$\| x - y \|^2 = (x - y, x - y)$$

$$= (x, x - y) - (y, x - y)$$

$$= (x, x) - (x, y) - (y, x) + \| y \|^2 \qquad \dots (2)$$

Adding (1) and (2), we get

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

This completes the proof of the theorem.

Notes

Theorem 5: In a Hilbert space the inner product is jointly continuous i.e.,

$$x_n \to x, y_n \to y \Rightarrow (x_n, y_n) \to (x, y)$$

Proof: We have

$$|(x_n, y_n) - (x, y)| = |(x_n, y_n) - (x_n, y) + (x_n, y) - (x,, y)|$$

$$= |(x_n, y_n - y) + (x_n - x, y)|$$

(by linearity property of inner product)

$$\leq |(x_n, y_n - y)| + |(x_n - x, y)| \qquad [\because |\alpha + \beta| \leq |\alpha| + |\beta|]$$

$$\leq \| x_n \| + \| y_n - y \| + \| x_n - x \| \| y \| \qquad \text{[By Schwarz inequality]}$$

Since $x_n \to x$ and $y_n \to y$ as $n \to \infty$.

Therefore $\| y_n - y_n \| \to 0$ and $\| x_n - x \| \to 0$ as $h \to \infty$. Also (x_n) is a continues sequence, it is bounded so that $\| x_n \| \leq M \ \forall n$.

Therefore

$$| (x_n, y_n) - (x, y) | \to 0 \text{ as } n \to \infty.$$

Hence $(x_n, y_n) \to (x, y)$ as $n \to \infty$.

This completes the proof of the theorem.

Theorem 6: A closed convex set E in a Hilbert space H continuous a unique vector of smallest norm.

Proof: Let $\delta = \inf \{\| e \|; e \in E\}$

To prove the theorem it suffices to show that there exists a unique $x \in E$ s.t. $\| x \| = \delta$.

Definition of δ yields us a sequence (x_n) in E such that

$$\lim_{n\to\infty} \| x_n \| = \delta \qquad \dots (1)$$

Convexity of E implies that $\frac{x_m + x_n}{2} \in E$. Consequently

$$\left\| \frac{x_m + x_n}{2} \right\| \geq \delta \Rightarrow \| x_m + x_n \| \geq 2\delta \qquad \dots (2)$$

Using parallelogram law, we get

$$\| x_m + x_n \|^2 + \| x_m - x_n \|^2 = 2 \| x_n \|^2 + 2 \| x_n \|^2$$

or

$$\| x_m - x_n \|^2 = 2 \| x_m \|^2 + 2 \| x_n \|^2 - \| x_m - x_n \|^2$$

$$\leq 2 \| x_m \|^2 + 2 \| x_n \|^2 - d\delta^2 \qquad \text{(Using (2))}$$

$$\to 0 \text{ as } m, n \to \infty \qquad \text{(Using (1))}$$

$\Rightarrow$ $\| x_m - x_n \|^2 \to 0$ as $m, n \to \infty$

$\Rightarrow$ (x_n) is a CAUCHY sequence in E.

$\Rightarrow$ $\exists\, x \in E$ such that $\lim_{n\to\infty} x_n = x$, since H is complete and E is a closed subset of H, therefore E is also complete and consequently (x_n) is in E is a convergent sequence in E.

Notes

Now $\|x\| = \left\|\lim_{n\to\infty} x_n\right\|$

$= \lim_{n\to\infty}\|x_n\|$ ($\because$ norm is continuous mapping)

$= \delta.$

Uniqueness of x.

Let us suppose that $y \in E$, $y \neq x$ and $\|y\| = \delta$.

Convexity of E $\Rightarrow \frac{x+y}{2} \in E$

$$\Rightarrow \quad \left\|\frac{x+y}{2}\right\| \geq \delta \quad \dots (3)$$

Also by parallelogram law, we have

$$\left\|\frac{x}{2}+\frac{y}{2}\right\|^2 = \frac{\|x\|^2}{2}+\frac{\|y\|^2}{2}-\frac{\|x-y\|^2}{2}$$

$$= \frac{\delta^2}{2}+\frac{\delta^2}{2}\left\|\frac{x-y}{2}\right\|^2 = \delta^2 - \left\|\frac{x-y}{2}\right\|^2$$

$$< \delta^2.$$

So that

$$\left\|\frac{x-y}{2}\right\|^2 < \delta, \text{ a result contrary to (3).}$$

Hence we must have $y = x$.

This completes the proof of the theorem.

Example: Give an example of a Banach space which is not an Hilbert space.

Solution: C [a, b] is a Banach space with supremum norm, i.e. if $x \in C[a, b]$ then

$$\|x\| = \text{Sup}\{|x(t)| : t \in [a, b]\}.$$

Then this norm does not satisfy parallelogram law as shown below:

Let $x(t) = 1$ and $y(t) = \frac{t-a}{b-a}$. Then $\|x\| = 1$, $\|y\| = 1$

Now $x(t) + y(t) = 1 + \frac{t-a}{b-a}$ so that $\|x + y\| = 2$

$x(t) - y(t) = 1 - \frac{t-a}{b-a}$ so that $\|x - y\| = 1$

Hence $2(\|x\|^2 - \|y\|^2) = 4$, and $\|x + y\|^2 + \|x - y\|^2 = 5$

So that $\|x + y\|^2 + \|x - y\|^2 \neq 2\|x\|^2 + 2\|y\|^2$.

$\Rightarrow$ C [a, b] is not a Hilbert space.

Notes

9.2 Summary

- Let X be a linear space over the field of complex numbers C. An inner product on X is a mapping from $X \times X \to C$ which satisfies the following conditions:

  (i) $(\alpha x + \beta y, z) = \alpha (x, z) + \beta (y, z) \ \forall \ x, y, z \in X$ and $\alpha, \beta \in C$.

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

  (iii) $(x, x) \geq 0, (x, x) = 0 \Leftrightarrow x = 0$

- A complete inner product space is called a Hilbert space.

9.3 Keywords

Hilbert Space: A complete inner product space is called a Hilbert space.

Inner Product Spaces: Let X be a linear space over the field of complex numbers C. An inner product on X is a mapping from $X \times X \to C$ which satisfies the following conditions:

(i) $(\alpha x + \beta y, z) = \alpha (x, z) + \beta (y, z), \ \forall x, y, z \in X$ and $\alpha, \beta \in C$. (Linearity property)

(ii) $\overline{(x, y)} = (y, x)$ (Conjugate symmetry)

(iii) $(x, x) \geq 0, (x, x) = 0 \Leftrightarrow x = 0$

9.4 Review Questions

1. For the special Hilbert space ℓ_2^n, use Cauchy's inequality to prove Schwarz's inequality.

2. Show that the parallelogram law is not true in ℓ_2^n (n > 1).

3. If x, y are any two vectors in a Hilbert space H, then prove that

   $4 (x, y) = \| x + y \|^2 - \| x - y \|^2 + i \| x + iy \|^2 - i \| x - iy \|^2$.

4. If B is complex Banach space whose norm obeys the parallelogram law, and if an inner product is defined on B by

   $4 (x, y) = \| x + y \|^2 - \| x - y \|^2 + i \| x + iy \|^2 - i \| x - iy \|^2$,

   then prove that B is a Hilbert space.

9.5 Further Readings

Books

Bourbaki, Nicolas (1987), *Topological vector Spaces, Elements of Mathematics,* BERLIN: Springer – Verlag.

Halmos, Paul (1982), *A Hilbert space Problem Book*, Springer – Verlag.

Online links

www.math-sinica.edu.tw/www/file_upload/maliufc/liu_ch04.pdf

mathworld.wolfram.com>Calculus and Analysis > Functional Analysis

Unit 10: Orthogonal Complements

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define Orthogonal complement
- Understand theorems on it
- Understand the Orthogonal decomposition theorem
- Solve problems related to Orthogonal complement.

Introduction

In this unit, we shall start with orthogonality. Then we shall move on to definition of orthogonal complement. Let M be a closed linear subspace of H. We know that $M^{\perp}$ is also a closed linear subspace, and that M and $M^{\perp}$ are disjoint in the sense that they have only the zero vector in common. Our aim in this unit is to prove that $H = M \oplus M^{\perp}$, and each of our theorems is a step in this direction.

10.1 Orthogonal Complement

10.1.1 Orthogonal Vectors

Let H be a Hilbert space. If $x, y \in H$ then x is said to be orthogonal to y, written as $x \perp y$, if $(x, y) = 0$.

By definition,

(a) The relation of orthogonality is symmetric, i.e.,

$$x \perp y \Rightarrow y \perp x$$

Notes

For, $x \perp y \Rightarrow (x, y) = 0$

$\Rightarrow \overline{(x, y)} = \overline{0}$

$\Rightarrow (y, x) = 0$

$\Rightarrow y \perp x$

(b) If $x \perp y$ then every scalar multiple of x is orthogonal to y i.e. $x \perp y \Rightarrow \alpha x \perp y$ for every scalar $\alpha \in C$.

For, let α be any scalar, then

$$(\alpha x, y) = \alpha(x, y) = \alpha . 0 = 0$$

$\Rightarrow$ $x \perp y \Rightarrow \alpha x \perp y$.

(c) The zero vector is orthogonal to every vector. For every vector x in H, we have

$$(0, x) = 0$$

$\therefore$ $0 \perp x$ for all $x \in H$.

(d) The zero vector is the only vector which is orthogonal to itself. For,

if $x \perp x \Rightarrow (x, x) = 0 \Rightarrow \| x \|^2 = 0 \Rightarrow x = 0$

Hence, if $x \perp x$, then x must be a zero vector.

10.1.2 Pythagorean Theorem

Statement: If x and y are any two orthogonal vectors in a Hilbert space H, then

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

Proof: Given $x \perp y \Rightarrow (x, y) = 0$, then we must have

$$y \perp x \text{ i.e. } (y, x) = 0$$

Now
$$\| x + y \|^2 = (x + y, x + y) = (x, x) + (x, y) + (y, x) + (y, y) = \| x \|^2 + 0 + 0 + \| y \|^2 = \| x \|^2 + \| y \|^2$$

Also,
$$\| x - y \|^2 = (x - y, x - y) = (x, x) - (x, y) - (y, x) + (y, y) = \| x \|^2 - 0 - 0 - \| y \|^2 = \| x \|^2 + \| y \|^2$$

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

10.1.3 Orthogonal Sets

Definition: A vector x is to be orthogonal to a non-empty subset S of a Hilbert space H, denoted by $x \perp S$ if $x \perp y$ for every y in S.

Two non-empty subsets S_1 and S_2 of a Hilbert space H are said to be orthogonal denoted by $S_1 \perp S_2$, if $x \perp y$ for every $x \in S_1$ and every $y \in S_2$.

10.1.4 Orthogonal Compliment: Definition

Let S be a non-empty subset of a Hilbert space H. The orthogonal compliment of S, written as $S^\perp$ and is read as S perpendicular, is defined as

$$S^\perp = \{x \in H : x \perp y \ \forall \ y \in S\}$$

Thus, $S^\perp$ is the set of all those vectors in H which are orthogonal to every vectors in H which are orthogonal to every vector in S.

Theorem 1: If S, S_1, S_2 are non-empty subsets of a Hilbert space H, then prove the following:

(a) $\{0\}^\perp = H$ (b) $H^\perp = \{0\}$ (c) $S \cap S^\perp \subset \{0\}$

(d) $S_1 \subset S_2 \Rightarrow S_2^\perp \subset S_1^\perp$ (e) $S \subset S^{\perp\perp}$

Proof:

(a) Since the orthogonal complement is only a subset of H, $\{0\}^\perp \subset H$.

It remains to show that $H \subset \{0\}^\perp$.

Let $x \in H$. Since $(x, 0) = 0$, therefore $x \in \{0\}^\perp$.

Thus $x \in H \Rightarrow x \in \{0\}^\perp$.

$\Rightarrow H \subset \{0\}^\perp$.

Hence $\{0\}^\perp = H$

(b) Let $x \in H$. Then by definition of H, we have

$(x, y) = 0 \ \forall \ y \in H$

Taking $y = x$, we get

$(x, x) = 0 \Rightarrow \|x\|^2 = 0 \Rightarrow x = 0$

Thus $x \in H^\perp \Rightarrow x = 0$

$\therefore H^\perp = \{0\}$

(c) $x \in S \cap S^\perp$.

Then $x \in S$ and $x \in S^\perp$

Since $x \in S^\perp$, therefore x is orthogonal to every vector in S. In particular, x is orthogonal to x because $x \in S$.

Now $(x, x) = 0 \Rightarrow \|x\|^2 = 0 \Rightarrow x = 0$.

$\Rightarrow$ 0 is the only vector which can belong to both S and $S^\perp$.

$\therefore$ $S \cap S^\perp \subset \{0\}$

If S is a subspace of H, then $0 \in S$. Also $S^\perp$ is a subspace of H. Therefore $0 \in S^\perp$. Thus, if S is a subspace of H, then $0 \in S \cap S^\perp$. Therefore, in this case $S \cap S^\perp = \{0\}$.

(d) Let $S_1 \subset S_2$, we have

$x \in S_2^{\perp} \Rightarrow x$ is orthogonal to every vector in S_2

$\Rightarrow$ x is orthogonal to every vector in S_1 because $S_1 \subset S_2$.

$\Rightarrow$ $x \in S_1^{\perp}$

$\therefore$ $S_2^{\perp} \subset S_1^{\perp}$

(e) Let $x \in S$. Then $(x, y) = 0 \ \forall \ y \in S^{\perp}$.

$\therefore$ by definition of $(S^{\perp})^{\perp}$, $x \in (S^{\perp})^{\perp}$.

Thus $x \in S \Rightarrow x \in S^{\perp\perp}$.

$\Rightarrow$ $S \subset S^{\perp\perp}$.

This completes the proof of the theorem.

Theorem 2: If S is a non-empty subset of a Hilbert space H, then $S^{\perp}$ is a closed linear subspace of H and hence a Hilbert space.

Proof: We have

$S^{\perp} = \{x \in H : (x, y) = 0 \ \forall \ y \in S\}$ by definition. Since $(0, y) = 0 \ \forall \ y \in S$, therefore at least $0 \in S^{\perp}$ and thus $S^{\perp}$ is non-empty.

Now let $x_1, x_2 \in S^{\perp}$ and α, β be scalars. Then $(x_1, y) = 0$, $(x_2, y) = 0$ for every $y \in S$.

For every $y \in S$, we have

$$\begin{aligned}(\alpha x_1 + \beta x_2, y) &= \alpha (x_1, y) + \beta (x_2, y) \\ &= \alpha (0) + \beta (0) \\ &= 0\end{aligned}$$

$\Rightarrow$ $\alpha x_1 + \beta x_2 \in S^{\perp}$

$\Rightarrow$ $S^{\perp}$ is a subspace of H.

Next we shall show that $S^{\perp}$ is a closed subset of H.

Let $(x_n) \in S^{\perp}$ and $x_n \to x$ in H.

Then we have to show that $x \in S^{\perp}$.

For this we have to prove $(x, y) = 0$ for every $y \in S$.

Since $x_n \in S^{\perp}$, $(x_n, y) = 0$ for every $y \in S$ and for $n = 1, 2, 3, \ldots$

Since the inner product is a continuous function, we get

$(x_n, y) \to (x, y)$ as $n \to \infty$

Since $(x_n, y) = 0 \ \forall \ n$, $(x, y) = 0$

$\Rightarrow$ $x \in S^{\perp}$.

Hence $S^{\perp}$ is a closed subset of H.

Now $S^{\perp}$ is a closed subspace of the Hilbert space H.

So, $S^{\perp}$ is complete and hence a Hilbert space. This completes the proof of the theorem.

Theorem 3: If M is a linear subspace of a Hilbert space H, then M is closed

$\Leftrightarrow M = M^{\perp\perp}$.

Proof: Let us assume that $M = M^{\perp\perp}$,

M being a subspace of H.

by theorem (2), $(M^{\perp})^{\perp}$ is closed subspace of H.

Therefore $M = M^{\perp\perp}$ is a closed subspace of H. Conversely, let M be a closed subspace of H. We shall show that $M = M^{\perp\perp}$.

We know that $M \subset M^{\perp\perp}$.

Now suppose that $M \neq M^{\perp\perp}$.

Now M is a proper closed subspace of Hilbert space $M^{\perp\perp}$. $\exists$ a non-zero vector z_o in $M^{\perp\perp}$ such that $z_o \perp M$ or $z_o \in M^{\perp}$.

Now $z_o \in M^{\perp}$ and $M^{\perp\perp}$ gives $z_o \in M^{\perp} \cap M^{\perp\perp}$ … (1)

Since M is a subspace of H, we have

$$M^{\perp} \cap M^{\perp\perp} = \{0\} \quad \ldots (2)$$

(by theorem 1 (iii))

From (1) and (2) we conclude that z = 0, a contradiction to the fact that z_o is a non-zero vector.

$\therefore$ $M \subset M^{\perp\perp}$ can be a proper inclusion.

Hence $M = M^{\perp\perp}$.

This completes the proof of the theorem.

Cor. If M is a non-empty subset of a Hilbert space H, then $M^{\perp} = M^{\perp\perp\perp}$.

Proof: By theorem (2), $M^{\perp}$ is a closed subspace of H. So by theorem (3),

$$M^{\perp} = (M^{\perp})^{\perp\perp} = M^{\perp\perp\perp}.$$

Theorem 4: If M and N are closed linear subspace of a Hilbert space H such that $M \perp N$, then the linear subspace $M \perp N$ is closed.

Proof: To prove: M + N is closed, we have to prove that it contains all its limit point.

Let z be a limit point of M + N,

$\exists$ a sequence (z_n) in M + N such that $z_n \to z$ in H.

Since $M \perp N$, $M \cap N = \{0\}$ and M + N is the direct sum of the subspace M and N, z_n can be written uniquely as

$$z_n = x_n + y_n \text{ where } x_n \in M \text{ and } y_n \in N.$$

Taking two points $z_m = x_m + y_m$ and $z_n = x_n + y_n$, we have

$$z_m - z_n = (x_m - x_n) + (y_m - y_n).$$

Since $x_m - x_n \in M$ and $y_m - y_n \in N$, we get

$$(x_m - x_n) \perp (y_m - y_n)$$

Notes

So, by Pythagorean theorem, we have

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

But $(x_m - x_n) + (y_m - y_n) = z_m - z_n$ so that

$$\| z_m - z_n \|^2 = \| x_m - x_n \|^2 + \| y_m - y_n \|^2 \quad \ldots (1)$$

Since (z_n) is a convergent sequence in H, it is a Cauchy sequence in H.

Hence $\| z_m - z_n \|^2 \to 0$ as $m, n \to \infty$ … (2)

Using (2) in (1), we see that

$$\| x_m - x_n \|^2 \to 0 \text{ and } \| y_m - y_n \|^2 \to 0$$

So that (x_n) and (y_n) are Cauchy sequence in M and N.

Since H is complete and M and N are closed subspace of a complete space H, M and N are complete.

Hence, the Cauchy sequence (x_n) in M converges to x in M and the Cauchy sequence (y_n) in N converges to y in N.

Now $z = \lim z_n = \lim (x_n + y_n)$

$= \lim x_n + \lim y_n$

But $\lim x_n + \lim y_n = x + y \in M + N$

Thus, $z = x + y \in M + N$

$\Rightarrow M + N$ is closed.

10.1.5 The Orthogonal Decomposition Theorem or Projection Theorem

Theorem 5: If M is a closed linear subspace of a Hilbert space H, then $H = M \oplus M^\perp$.

Proof: If M is a subspace of a Hilbert space H, then we know that $M \cap M^\perp = \{0\}$.

Therefore in order to show that

$H = M \oplus M^\perp$, we need to verify that

$H = M + M^\perp$.

Since M and $M^\perp$ are closed subspace of H, $M + M^\perp$ is also a closed subspace of H by theorem 4.

Let us take $N = M + M^\perp$ and show that N = H.

From the definition of N, we get $M \subset N$ and $M^\perp \subset N$. Hence by theorem (1), we have

$N^\perp \subset M^\perp$ and $N^\perp \subset M^{\perp\perp}$.

Hence $N^\perp \subset M^\perp \cap M^{\perp\perp} = \{0\}$.

$\Rightarrow \quad N^\perp = \{0\}$

$\Rightarrow \quad N^{\perp\perp} = \{0\}^\perp = H$ … (1)

Since $N = M + M^\perp$ is a closed subspace of H, we have by theorem (3),

$$N^{\perp\perp} = N \quad \ldots (2)$$

From (1) and (2), we have

$$N = M + M^\perp = H.$$

Since $M \cap M^{\perp} = \{0\}$ and

$$H = M + M^{\perp},$$

we have from the definition of the direct sum of subspaces,

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

This completes the proof of the theorem.

Theorem 6: Let M be a proper closed linear sub space of a Hilbert space H. Then there exists a non-zero vector z_o in H such that $z_o \perp M$.

Proof: Since M is a proper subspace of H, there exists a vector x in H which is not in M.

Let $d = d(x, M) = \inf \{\| x - y \|\} : y \in M\}$.

Since $x \notin M$, we have $d > 0$.

Also M is a proper closed subspace of H, then by theorem: "Let M be a closed linear subspace of a Hilbert space H. Let x be a vector not in M and let d = d (x, m) (or d is the distance from x to M). Then there exists a unique vector y_o in M such that $\| x - y_o \| = d$."

There exists a vector y_o in M such that

$$\| x - y_o \| = d.$$

Let $z_o = x - y_o$. We then here

$$\| z_o \| = \| x - y_o \| = d > 0.$$

$\Rightarrow z_o$ is a non-zero vector.

Now we claim that $Z_o \perp M$.

Let y be an arbitrary vector in M. We shall show that $z_o \perp y$. For any scalar α, we have

$$z_o - \alpha y = x - y_o - \alpha y = x - (y_o + \alpha y).$$

since M is a subspace of H and $y_o, y \in M$,

$$\therefore \quad y_o + \alpha M \in M.$$

Then by definition of d, we have

$$\| x - (y_o + \alpha y) \| \geq d$$

Now $\| z_o - \alpha y \| = \| x - (y_o + \alpha y) \| \geq d = \| z_o \|$

$$\therefore \quad \| z_o - \alpha y \|^2 \geq \| z_o \|^2$$

or $(z_o - \alpha y, z_o - \alpha y) - (z_o, z_o) \geq 0$

or $(z_o, z_o) - \bar{\alpha}(z_o, y) - \alpha(y, z_o) + \alpha\bar{\alpha}(y, y) - (z_o, z_o) \geq 0$

or $-\bar{\alpha}(z_o, y) - \alpha(\overline{z_o, y}) + \alpha\bar{\alpha}(y, y) \geq 0$ … (1)

The above result is true for all scalars α.

Let us take $\bar{\alpha} = \beta(\overline{z_o, y})$.

Putting the value of α, $\bar{\alpha}$ in (1), we get

$$-\beta(\overline{z_o, y})(z_o, y) - \beta(z_o, y)(\overline{z_o, y}) + \beta^2(z_o, y)(\overline{z_o, y})\| y \|^2 \geq 0$$

or $-2\beta \, |(z_o, y)|^2 + \beta^2 \, |(z_o, y)|^2 \| y \|^2 \geq 0$

or $\beta \ |(z_o, y)|^2 \{\beta \| y \|^2 - 2\} \geq 0$ … (2)

The above result is true for all real β suppose that $(z_o, y) \neq 0$. Then taking β positive and so small that $\beta \| y \|^2 < 2$, we see from (2) that $\beta \ |(z_o, y)|^2 \{\beta \| y \|^2 - 2\} < 0$.

This contradicts (2).

Hence we must have $(z_o, y_o) = 0 \Rightarrow z_o \perp y, \ \forall y \in M$.

$\therefore$ $z_o \perp M$.

This completes the proof of the theorem.

10.2 Summary

- Let H be a Hilbert space. If $x, y \in H$ then x is said to be orthogonal to y, written as $x \perp y$, if $(x, y) = 0$.
- If x and y are any two orthogonal vectors in a Hilbert space H, then

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

- Two non-empty subsets S_1 and S_2 of a Hilbert space H are said to be orthogonal denoted by $S_1 \perp S_2$, if $x \perp y$ for every $x \in S_1$ and every $y \in S_2$.
- Let S be a non-empty subsets of a Hilbert space H. The orthogonal compliment of S, written as $S^\perp$ and is read as S perpendicular, is defined as

$$S^\perp = \{x \in H : x \perp y \ \ \forall y \in S\}$$

- The orthogonal decomposition theorem: If M is a closed linear subspace of a Hilbert space H, then $H = M \oplus M^\perp$.

10.3 Keywords

Orthogonal Compliment: Let S be a non-empty subset of a Hilbert space H. The orthogonal compliment of S, written as $S^\perp$ and is read as S perpendicular, is defined as

$$S^\perp = \{x \in H : x \perp y \ \ \forall \ y \in S\}$$

Orthogonal Sets: A vector x is to be orthogonal to a non-empty subset S of a Hilbert space H, denoted by $x \perp S$ if $x \perp y$ for every y in S.

Two non-empty subsets S_1 and S_2 of a Hilbert space H are said to be orthogonal denoted by $S_1 \perp S_2$, if $x \perp y$ for every $x \in S_1$ and every $y \in S_2$.

Orthogonal Vectors: Let H be a Hilbert space. If $x, y \in H$ then x is said to be orthogonal to y, written as $x \perp y$, if $(x, y) = 0$.

Pythagorean Theorem: If x and y are any two orthogonal vectors in a Hilbert space H, then

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

10.4 Review Questions

1. If S is a non-empty subset of a Hilbert space, show that $S^\perp = S^{\perp\perp\perp}$.
2. If M is a linear subspace of a Hilbert space, show that M is closed $\Leftrightarrow M = M^{\perp\perp}$.
3. If S is a non-empty subset of a Hilbert space H, show that the set of all linear combinations of vectors in S is dense in $H \Leftrightarrow S^\perp = \{0\}$.

4. If S is a non-empty subset of a Hilbert space H, show that $S^{\perp\perp}$ is the closure of the set of all linear combinations of vectors in S.

5. If M and N are closed linear subspace of a Hilbert space h such that $M \perp N$, then the linear subspace M + N is closed.

10.5 Further Readings

Books

Halmos, Paul R. (1974), *Finite-dimensional Vector Spaces*, Berlin, New York

Paul Richard Halmos, *A Hilbert Space Problem Book*, 2nd Ed.

Online links

Itcconline.net/green/courses/203/…/orthogonal complements.html

www.math.cornell.edu/~andreim/Lec33.pdf

www.amazon.co.uk

Notes

Unit 11: Orthonormal Sets

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand orthonormal sets
- Define unit vector or normal vector
- Understand the theorems on orthonormal sets.

Introduction

In linear algebra two vectors in an inner product space are orthonormal if they are orthogonal and both of unit length. A set of vectors from an orthonormal set if all vectors in the set are mutually orthogonal and all of unit length.

In this unit, we shall study about orthonormal sets and its examples.

11.1 Orthonormal Sets

11.1.1 Unit Vector or Normal Vector

Definition: Let H be a Hilbert space. If $x \in H$ is such that $\| x \| = 1$, i.e. $(x, x) = 1$, then x is said to be a unit vector or normal vector.

11.1.2 Orthonormal Sets, Definition

A non-empty subset $\{ e_i \}$ of a Hilbert space H is said to be an orthonormal set if

(i) $i \neq j \Rightarrow e_i \perp e_j$, equivalently $i \neq j \Rightarrow (e_i, e_j) = 0$

(ii) $\| e_i \| = 1$ or $(e_i, e_i) = 1$ for every i.

Thus a non-empty subset of a Hilbert space H is said to be an orthonormal set if it consists of mutually orthogonal unit vectors.

Notes

1. An orthonormal set cannot contain zero vector because $\| 0 \| = 0$.
2. Every Hilbert space H which is not equal to zero space possesses an orthonormal set.

   Since $0 \neq x \in H$. Then $\| x \| \neq 0$. Let us normalise x by taking $e = \frac{x}{\|x\|}$, so that

   $$\| e \| = \left\| \frac{x}{\|x\|} \right\| = \frac{1}{\|x\|} \cdot \|x\| = 1.$$

   $\Rightarrow$ e is a unit vector and the set {e} containing only one vector is necessarily an orthonormal set.

3. If $\{x_i\}$ is a non-empty set of mutually orthogonal vectors in H, then $\{e_i\} = \left\{ \frac{x_i}{\|x_i\|} \right\}$ is an orthonormal set.

11.1.3 Examples of Orthonormal Sets

1. In the Hilbert space ℓ_2^n, the subset $e_1, e_2, \ldots, e_n$ where e_i is the i-tuple with 1 in the i^{th} place and O's elsewhere is an orthonormal set.

   For $(e_i, e_j) = 0 \quad i \neq j$ and $(e_i, e_j) = 1$ in the inner product $\sum_{i=1}^{n} x_i \overline{y}_i$ of ℓ_2^n.

2. In the Hilbert space ℓ_2, the set $\{e_1, e_2, \ldots, e_n, \ldots\}$ where e_n is a sequence with 1 in the n^{th} place and O's elsewhere is an orthonormal set.

11.1.4 Theorems on Orthonormal Sets

Theorem 1: Let $\{e_1, e_2, \ldots, e_n\}$ be a finite orthonormal set in a Hilbert space H. If x is any vector in H, then

$$\sum_{i=1}^{n} |(x, e_i)|^2 \leq \| x \|^2; \qquad \ldots (1)$$

further,

$$x - \sum_{i=1}^{n} (x, e_i)\, e_i \perp e_j \text{ for each } j \qquad \ldots (2)$$

Proof: Consider the vector

$$y = x - \sum_{i=1}^{n} (x, e_i)\, e_i$$

We have $\|y\|^2 = (y, y)$

$$= \left(x - \sum_{i=1}^{n} (x, e_i)\, e_i ,\; x - \sum_{i=1}^{n} (x, e_i)\, e_i \right)$$

$$= (x,x) - \sum_{i=1}^{n} (x, e_i)\,(e_i , x) - \sum_{j=1}^{n} \overline{(x, e_j)}\,(x, e_j)$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} (x, e_i)\,\overline{(x, e_i)}\,(e_i , e_j)$$

$$= \|x\|^2 - \sum_{i=1}^{n} (x, e_i)\,\overline{(x, e_i)} - \sum_{j=1}^{n} \overline{(x, e_i)}\,(x, e_j) + \sum_{i=1}^{n} (x, e_j)\,\overline{(x, e_i)}$$

On summing with respect to j and remembering that $(e_i, e_j) = 1$, $i = j$ and $(e_i, e_j) = 0$, $i \neq j$

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

$$= \|x\|^2 - \sum_{i=1}^{n} |(x, e_i)|^2$$

Now $\|y\|^2 \geq 0$, therefore $\|x\|^2 - \sum_{i=1}^{n} |(x, e_i)|^2 \geq 0$

$$\Rightarrow \quad \sum_{i=1}^{n} |(x, e_i)|^2 \leq \|x\|^2$$

$\Rightarrow$ result (1).

Further to prove result (2), we have for each j ($1 \leq j \leq n$),

$$\left(x - \sum_{i=1}^{n} (x, e_i)\, e_i , e_j \right) = (x, e_j) - \left(\sum_{i=1}^{n} (x, e_i)\, e_i , e_j \right)$$

$$= (x, e_j) - \sum_{i=1}^{n} (x, e_i)\,(e_i , e_j)$$

$= (x, e_j) - (x, e_j)$ $\qquad [\because (e_i, e_j) = 1, i \neq j\ 0, i = j]$

$= 0$

Hence $x - \sum_{i=1}^{n} (x, e_i)\, e_i \perp e_j$ for each j.

This completes the proof of the theorem.

Note The result (1) is known as Bessel's inequality for finite orthonormal sets.

Notes

Theorem 2: If $\{e_i\}$ is an orthonormal set in a Hilbert space H and if x is any vector in H, then the set $S = \{e_i : (x, e_i) \neq 0\}$ is either empty or countable.

Proof: For each positive integer n, consider the set

$$S_n = \left\{ e_i : |(x, e_i)|^2 > \frac{\|x\|^2}{n} \right\}.$$

If the set S_n contains n or more than n′ vectors, then we must have

$$\sum_{e_i \in S_n} |(x, e_i)|^2 > n \frac{\|x\|^2}{n} = \|x\|^2 \qquad \dots (1)$$

By theorem (1), we have

$$\sum_{e_i \in S_n} |(x, e_i)|^2 \leq \|x\|^2 \qquad \dots (2)$$

which contradicts (1).

Hence if (2) were to be valid, S_n should have at most (n – 1) elements. Hence for each positive n, the set S_n is finite.

Now let $e_i \in S$. Then $(x, e_i) \neq 0$. However small may be the value of $|(x, e_i)|^2$, we can take n so large that

$$|(x, e_i)|^2 > \frac{\|x\|^2}{n}.$$

Therefore if $e_i \in S$, then e_i must belong to some S_n. So, we can write $S = \bigcup_{n=1}^{\infty} S_n$.

$\Rightarrow$ S can be expressed as a countable union of finite sets.

$\Rightarrow$ S is itself a countable set.

If $(x, e_i) = 0$ for each i, then S is empty. Otherwise S is either a finite set or countable set.

This completes the proof of the theorem.

Theorem 3: *Bessel's Inequality:* If $\{e_i\}$ is an orthonormal set in a Hilbert space H, then $\Sigma |(x, e_i)|^2 \leq \|x\|^2$ for every vector x in H.

Proof: Let $S = \{e_i : (x, e_i) \neq 0\}$.

By theorem (2), S is either empty or countable.

If S is empty, then $(x, e_i) = 0 \ \forall$ i.

So if we define $\Sigma |(x, e_i)|^2 = 0$, then

$$\Sigma |(x, e_i)|^2 = 0 \leq \|x\|^2.$$

Now let S is not empty, then S is finite or it is countably infinite.

If S is finite, then we can write $S = \{e_1, e_2, \dots, e_n\}$ for some positive integer n.

In this case, we have

$$\Sigma |(x, e_i)|^2 = \sum_{i=1}^{n} |(x, e_i)|^2 \leq \|x\|^2 \qquad \dots (1)$$

which represents Bessel's inequality in the finite case.

Notes

If S is countable infinite, let S be arranged in the definite order such as $\{e_1, e_2, \dots, e_n, \dots\}$.

In this case we can write

$$\Sigma \,|\, (x, e_i) \,|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2 \qquad \dots (2)$$

The series on the R.H.S. of (2) is absolutely convergent.

Hence every series obtained from this by rearranging the terms is also convergent and all such series have the same sum.

Therefore, we define the sum $\Sigma \,|\, (x, e_i) \,|^2$ to be $\sum_{n=1}^{\infty} |(x, e_n)|^2$.

Hence the sum of $\Sigma \,|\, (x, e_i) \,|^2$ is an extended non-negative real number which depends only on S and not on the rearrangement of vectors.

Now by Bessel's inequality in the finite case, we have

$$\sum_{i=1}^{n} |(x, e_i)|^2 \le \|x\|^2 \qquad \dots (3)$$

For various values of n, the sum on the L.H.S. of (3) are non-negative. So they form a monotonic increasing sequence. Since this sequence is bounded above by $\|x\|^2$, it converges. Since the sequence is the sequence of partial sums of the series on the R.H.S. of (2), it converges and we have $e_i \in S$,

$$\Sigma \,|\, (x, e_i) \,|^2 = \sum_{n=1}^{\infty} |(x, e_i)|^2 \le \|x\|^2$$

This completes the proof of the theorem.

Note: From the Bessel's inequality, we note that the series $\sum_{n=1}^{\infty} |(x, e_n)|^2$ is convergent series.

Corollary: If $e_n \in S$, then $(x, e_n) \to 0$ as $n \to \infty$.

Proof: By Bessel's inequality, the series $\sum_{n=1}^{\infty} |(x, e_n)|^2$ is convergent.

Hence $\quad |(x, e_n)|^2 \to 0$ as $n \to \infty$.

$\Rightarrow \qquad (x, e_n) \to 0$ as $n \to \infty$.

Theorem 4: If $\{e_i\}$ is an orthonormal set in a Hilbert space H and x is an arbitrary vector in H, then

$\{x - \Sigma (x, e_i) e_i\} \perp e_j$ for each j.

Proof: Let $S = \{e_i : (x, e_i) \neq 0\}$

Then S is empty or countable. [See theorem (2)]

If S is empty, then $(x, e_i) = 0$ for every i.

In this case, we define $\Sigma (x, e_i) e_i$ to be a zero vector and so we get

$$x - \Sigma (x, e_i) e_i = x - 0 = x.$$

Notes

Hence in this case, we have to show $x \perp e_j$ for each j.

Since S is empty, $(x, e_j) = 0$ for every j.

$\Rightarrow \quad x \perp e_j$ for every j.

Now let S is not empty. Then S is either finite or countably infinite. If S is finite, let

$S = \{e_1, e_2, \ldots, e_n\}$ and we define

$$\Sigma (x, e_i)\, e_i = \sum_{i=1}^{n} (x, e_i)\, e_i,$$

and prove that $\left\{x - \sum_{n=1}^{\infty} (x, e_i)\, e_i\right\} \perp e_j$ for each j. This result follows from (2) of theorem (1).

Finally let S be countably infinite and let

$$S = \{e_1, e_2, \ldots, e_n, \ldots\}$$

Let $$s_n = \sum_{i=1}^{\infty} (x, e_i)\, e_i$$

For m > n, $$\| s_m - s_n \|^2 = \left\| \sum_{i=n+1}^{m} |(x, e_i) e_i| \right\|^2$$

$$= \sum_{i=n+1}^{m} |(x, e_i)|^2$$

By Bessel's inequality, the series $\sum_{n=1}^{\infty} |(x, e_n)|^2$ is convergent.

Hence $\sum_{i=n+1}^{\infty} |(x, e_i)|^2$ as m, n $\to \infty$.

$\Rightarrow \quad \| s_m - s_n \|^2 \to 0$ as m, n $\to \infty$.

$\Rightarrow \quad (s_n)$ is a Cauchy sequence in H.

Since H is complete $s_n \to s \in H$. Now $s \in H$ can be written as

$$s = \sum_{n+1}^{\infty} (x, e_n)\, e_n$$

Now we can define $\Sigma(x, e_i)\, e_i = \sum_{n+1}^{\infty} (x, e_n)\, e_n$.

Before, completing the proof, we shall show that the above sum is well-defined and does not depend upon the rearrangement of vectors.

For this, let the vector in S be arranged in a different manner as

$$S = \{f_1, f_2, f_3, \ldots, f_n, \ldots\}$$

Notes

Let $$s'_n = \sum_{i=1}^{n} (x, f_i) f_i$$

As shown for the case above for (s_n), let

$$s'_n \to s' \text{ in H where we can take}$$

$$s' = \sum_{n=1}^{\infty} (x, f_n) f_n .$$

We prove that $s = s'$. Given $\in > 0$ we can find n_0 such that $\forall\ n \geq n_0$.

$$\sum_{i=n_0+1}^{\infty} |x, e_i|^2 < \in^2, \| s_n - s \| < \in, \ \| s'_n - s' \| < \in \qquad \ldots (1)$$

For some positive integer $m_0 > n_0$, we can find all the terms of s_n in s'_{m_o} also.

Hence $s'_{m_o} - s_{n_o}$ contains only finite number of terms of the type $(x, e_i)\, e_i$ for $i = n_0 + 1, n_0 + 2, \ldots$

Thus, we get $\left\| s'_{m_o} - s_{n_o} \right\| \leq \sum_{i=n_0+1}^{\infty} |x, e_i|^2 < \in^2$ so that we have

$$\left\| s'_{m_o} - s'_{n_o} \right\| < \in \qquad \ldots (2)$$

Now
$$\| s' - s \| = \left\| \left(s' - s'_{m_o}\right) + \left(s'_{m_o} - s'_{n_o}\right) + \left(s_{n_o} - s\right) \right\|$$

$$\leq \left\| s' - s'_{m_0} \right\| + \left\| s'_{m_0} - s'_{n_0} \right\| + \left\| s_{n_0} - s \right\|$$

$$< \in + \in + \in = 3\in \qquad \text{(Using (1) and (2))}$$

Since $\in > 0$ is arbitrary, $s' - s = 0$ or $s = s'$.

Now consider

$$(x - \Sigma (x, e_i)\, e_i, e_j) = (x - s, e_j)$$

But
$$(x - s, e_j) = (x, e_j) - (s, e_j)$$

$$= (x, e_j) - (\lim s_n, e_j) \qquad \ldots (3)$$

By continuity of inner product, we get

$$(\lim s_n, e_j) = \lim (s_n, e_j) \qquad \ldots (4)$$

Using (3) in (4), we obtain

$$(x - \Sigma (x, e_i)\, e_i, e_j) = (x, e_j) - \lim (s_n, e_j)$$

If $e_j \notin S$, then

$$(s_n, e_j) = \left(\sum_{i=1}^{n} (x, e_i)\, e_i, e_j \right) = 0$$

$$\Rightarrow \qquad \lim_{n\to\infty} (s_n, e_j) = 0$$

Hence $(x - \Sigma (x, e_i) e_i, e_j) = (x, e_j) = 0$ since $e_j \notin S$.

If $e_j \in S$, then $(s_n, e_j) = \left(\sum_{i=1}^{n} (x, e_i) e_i, e_j\right)$ ... (5)

If $n > j$, we get $\left(\sum_{i=1}^{n} (x, e_i) e_i, e_j\right) = (x, e_j)$ ... (6)

From (5) & (6), we get

$$\lim_{n\to\infty} (s_n, e_j) = (x, e_j).$$

So, in this case

$$(x - \Sigma (x, e_i) e_i, e_j) = (x, e_j) - (x, e_j) = 0$$

Thus $(x - \Sigma (x, e_i) e_i, e_j) = 0$ for each j.

Hence $x - \Sigma (x, e_i) e_i \perp e_j$ for each j.

This completes the proof of the theorem.

Theorem 5: A Hilbert space H is separable $\Leftrightarrow$ every orthonormal set in H is countable.

Proof: Let H be separable with a countable dense subset D so that $H = \overline{D}$.

Let B be an orthonormal basis for H.

We show that B is countable.

For $\forall$ x, y $\in$ B, x $\neq$ y, we have

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

Hence the open sphere

$$S\left(x; \frac{1}{2}\right) = \left\{z : \| z - x \| < \frac{1}{2}\right\} = \text{as } x \in B \text{ are all disjoint.}$$

Since D is dense, D must contain a point in each $S\left(x, \frac{1}{2}\right)$.

Hence if B is uncountable, then B must also be uncountable and H cannot be separable contradicting the hypothesis. Therefore B must be countable.

Conversely, let B be countable and let $B = \{x_1, x_2, \ldots\}$. Then H is equal to the closure of all finite linear combinations of element of B. That is $H = \overline{L(B)}$. Let G be a non-empty open subset of H.

Then G contains an element of the form $\sum_{i=1}^{n} a_i x_i$ with $a_i \in C$. We can take $a_i \in C$. We can take a_i to be complex number with real and imaginary parts as rational numbers. Then the set

$$D = \left\{\sum_{i=1}^{n} a_i x_i, n = 1, 2, \ldots, a_i \text{ rational}\right\}$$

is a countable dense set in H and so H is separable.

This completes the proof of the theorem.

Theorem 6: A orthonormal set in a Hilbert space is linear independent.

Proof: Let S be an orthonormal set in a Hilbert space H.

To show that S is linearly independent, we have to show that every finite subset of S is linearly independent.

Let $S_1 = \{e_1, e_2, \ldots, e_n\}$ be any finite subset of S.

Now let us consider

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = 0 \qquad \ldots (1)$$

Taking the inner product with e_k $(1 \le k \le n)$,

$$\left(\sum_{i=1}^{n} \alpha_i e_i, e_k\right) = \sum_{i=1}^{n} \alpha_i (e_i, e_k) \qquad \ldots (2)$$

Using the fact that $(e_i, e_k) = 0$ for $i \neq k$ and $(e_k, e_k) = 1$, we get

$$\sum_{i=1}^{n} \alpha_i (e_j, e_k) = \alpha_k \qquad \ldots (3)$$

It follows from (2) on using (1) and (3) that

$$(0, e_k) = \alpha_k$$

$\Rightarrow \quad \alpha_k = 0 \ \forall \ k = 1, 2, \ldots, n.$

$\Rightarrow \quad S_1$ is linearly independent.

This completes the proof of the theorem.

Example: If $\{e_i\}$ is an orthonormal set in a Hilbert space H, and if x, y are arbitrary vectors in H, then $\sum \left|(x, e_i)\,\overline{(y, e_i)}\right| \le \|x\| \, \|y\|$.

Solution: Let $S = \left\{e_i : (x, e_i)\overline{(y, e_i)} \neq 0\right\}$

Then S is either empty or countable.

If S is empty, then we have

$$(x, e_i)\overline{(y, e_i)} = 0 \ \forall \ i$$

and in this case we define

$\sum \left|(x, e_i)\,\overline{(y, e_i)}\right|$ to be number 0 and we have $0 \le \|x\|^2 \ \|y\|^2$.

If S is non-empty, then S is finite or it is countably infinite. If S is finite, then we can write $S = \{e_1, e_2, \ldots, e_n\}$ for some positive integer n.

In this case we define

$$\sum \left|(x, e_i)\,\overline{(y, e_i)}\right| = \sum_{i=1}^{n} \left|(x, e_i)\,\overline{(y, e_i)}\right|$$

Notes

$$\le \left\{\sum_{i=1}^{n}|(x,e_i)|^2\right\}^{\frac{1}{2}} \left\{\sum_{i=1}^{n}|(y,e_i)|^2\right\}^{\frac{1}{2}} \quad \text{(By Cauchy inequality)}$$

$$\le \|x\|^2 \ \|y\|^2 \quad \text{(by Bessel's inequality for finite case)}$$

$$\therefore \quad \sum_{i=1}^{n}|(x,e_i)||\overline{(y,e_i)}| \le \|x\| \ \|y\| \quad \dots (1)$$

Finally let S is countably infinite. Let the vectors in S be arranged in a definite order as

$$S = \{e_1, e_2, \dots, e_n, \dots\}.$$

Let us define

$$\sum|(x,e_i)||\overline{(y,e_i)}| = \sum_{i=1}^{\infty}|(x,e_n)||\overline{(y,e_n)}|.$$

But this sum will be well defined only if we can show that the series $\sum_{n=1}^{\infty}|(x,e_n)||\overline{(y,e_n)}|$ is convergent and its sum does not change by rearranging its term i.e. by any arrangement of the vectors in the set S.

Since (1) is true for every positive integer n, therefore it must be true in the limit. So

$$\sum_{n=1}^{\infty}|(x,e_n)||\overline{(y,e_n)}| \le \|x\| \ \|y\| \quad \dots (2)$$

From (2), we see that the series $\sum_{n=1}^{\infty}|(x,e_n)||\overline{(y,e_n)}|$ is convergent. Since all the terms of the series are positive, therefore it is absolutely convergent and so its sum will not change by any rearrangement of its terms. So, we are justified in defining

$$\sum|(x,e_i)||\overline{(y,e_i)}| = \sum_{n=1}^{\infty}|(x,e_n)||\overline{(y,e_n)}|$$

and from (2), we see that this sum is $\le \|x\| \ \|y\|$.

11.2 Summary

- Two vectors in an inner product space are orthonormal if they are orthogonal and both of unit length. A set of vectors from an orthonormal set if all vectors in the set are mutually orthogonal and all of unit length.
- Examples of orthonormal sets are as follows:

  (i) In the Hilbert space ℓ_2^n, the subset $e_1, e_2, \dots, e_n$ where e_i is the i-tuple with 1 in the i^{th} place and O's elsewhere is an orthonormal set.

  For $(e_i, e_j) = 0 \quad i \ne j$ and $(e_i, e_j) = 1$ in the inner product $\sum_{i=1}^{n} x_i \overline{y}_i$ of ℓ_2^n.

(ii) In the Hilbert space ℓ_2, the set $\{e_1, e_2, \ldots, e_n, \ldots\}$ where e_n is a sequence with 1 in the n^{th} place and O's elsewhere is an orthonormal set.

11.3 Keywords

Orthonormal Sets: A non-empty subset $\{ e_i \}$ of a Hilbert space H is said to be an orthonormal set if

(i) $i \neq j \Rightarrow e_i \perp e_j$, equivalently $i \neq j \Rightarrow (e_i, e_j) = 0$

(ii) $\| e_i \| = 1$ or $(e_i, e_j) = 1$ for every i.

Unit Vector or Normal Vector: Let H be a Hilbert space. If $x \in H$ is such that $\| x \| = 1$, i.e. $(x, x) = 1$, then x is said to be a unit vector or normal vector.

11.4 Review Questions

1. Let $\{e_1, e_2, \ldots, e_n\}$ be a finite orthonormal set in a Hilbert space H, and let x be a vector in H. If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are arbitrary scalars, show that $\left\| x - \sum_{i=1}^{n} \alpha_i e_i \right\|$ attains its minimum value $\Leftrightarrow$ $\alpha_i = (x, e_i)$ for each i.

2. Prove that a Hilbert space H is separable $\Leftrightarrow$ every orthonormal set in H is countable.

11.5 Further Readings

Book

Sheldon Axler, *Linear Algebra Done Right (2nd ed.)*, Berlin, New York (1997).

Online links

www.mth.kcl.ac.uk/~jerdos/op/w3.pdf

mathworld.wolfram.com

www.utdallas.edu/dept/abp/PDF_files

Unit 12: The Conjugate Space H^*

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define the conjugate space H^*.
- Understand theorems on it.
- Solve problems related to conjugate space H^*.

Introduction

Let H be a Hilbert space. A continuous linear transformation from H into C is called a continuous linear functional or more briefly a functional on H. Thus if we say that f is a functional on H, then f will be continuous linear functional on H. The set $\beta(H,C)$ of all continuous linear functional on H is denoted by H^* and is called the conjugate space of H. The elements of H^* are called continuous linear functional or more briefly functionals. We shall see that the conjugate space of a Hilbert space H is the conjugate space H^* of H is in some sense is same as H itself. After establishing a correspondence between H and H^*, we shall establish the Riesz representation theorem for continuous linear functionals. Thereafter we shall prove that H^* is itself a Hilbert space and H is reflexive, i.e. ∃ has a natural correspondence between H and H^{**} and this natural correspondence is an isometric isomorphism of H onto H^{**}.

12.1 The Conjugate Space H^*

12.1.1 Definition

Let H be a Hilbert space. If f is a functional on H, then f will be continuous linear functional on H. The set $\beta(H,C)$ of all continuous linear functional on H is denoted by H^* and is called the conjugate space of H. The conjugate space of a Hilbert space H is the conjugate space H^* of H is in some sense is same as H itself.

12.1.2 Theorems and Solved Examples

Theorem 1: Let y be a fixed vector in a Hilbert space H and let fy be a scalar valued function on H defined by

$$fy(x) = (x, y) \forall x \in H.$$

Then fy is a functional in H^* i.e. fy is a continuous linear functional on H and $\|y\| = \|fy\|$.

Proof: From the definition

$fy : H \to C$ defined as $fy(x) = (x, y) \forall x \in H.$

We prove that fy is linear and continuous so that it is a functional.

Let $x_1, x_2 \in H$ and α, β be any two scalars. Then for any fixed $y \in H$,

$fy(\alpha x_1 + \beta x_2) = (\alpha x_1 + \beta x_2, y)$

$= \alpha(x_1, y) + \beta(x_2, y)$

$= \alpha fy(x_1) + \beta fy(x_2)$

$\Rightarrow$ fy is linear.

To show fy is continuous, for any $x \in H$

$$|fy(x)| = |(x, y)| \leq \|x\|.\|y\| \quad ...(1)$$

(Schwarz inequality)

Let $\|y\| \leq M$. Then for M > 0

$|fy(x)| \leq M\|x\|$ so that fy is bounded and hence fy is continuous.

Now let y = 0, $\|y\| = 0$ and from the definition fy = 0 so that $\|fy\| \leq \|y\|$.

Further let $y \neq 0$. Then from (1) we have $\dfrac{\text{Sup } |fy(x)|}{\|x\|} \leq \|y\|$.

Hence using the definition of the norm of a functional,

we get $\|fy\| \leq \|y\|$ ...(2)

Further $\|fy\| = \sup\{|fy(x)| : \|x\| \leq 1\}$ ...(3)

Since $y \neq 0, \left(\dfrac{y}{\|y\|}\right)$ is a unit vector.

From (3), we get

$$\|fy\| \geq \left|fy\left(\frac{y}{\|y\|}\right)\right| \quad ...(4)$$

Notes

But $fy\left(\frac{y}{\|y\|}\right) = \left(\frac{y}{\|y\|}, y\right) = \frac{1}{\|y\|}(y, y) = \|y\|$ ...(5)

Using (5) in (4) we obtain

$$\|fy\| \geq \|y\|$$

From (2) and (6) it follows that

$$\|fy\| = \|y\|$$

This completes the proof of the theorem.

Theorem 2: **(Riesz-representation Theorem for Continuous Linear Functional on a Hilbert Space):** Let H be a Hilbert space and let f be an arbitrary functional on H^*. Then there exists a unique vector y in H such that

f = fy, i.e. f(x) = (x,y) for every vector $x \in H$ and $\|f\| = \|y\|$.

Proof: We prove the following three steps to prove the theorem.

Step 1: Here we show that any $f \in H^*$ has the representation f = fy.

If f = 0 we take y = 0 so that result follows trivially.

So let us take $f \neq 0$.

We note the following properties of y in representation if it exists. First of all $y \neq 0$, since otherwise f = 0.

Further (x,y) = 0$\forall$x for which f(x) = 0. This means that if x belongs to the null space N(f) of f, then (x,y) = 0.

$\Rightarrow y \in N(f)^{\perp}$.

So let us consider the null space N(f) of f. Since f is continuous, we know that N(f) is a proper closed subspace and since $f \neq 0, N(f) \neq H$ and so $N(f)^{\perp} \neq \{0\}$.

Hence by the orthogonal decomposition theorem, $\exists\ ay_0 \neq 0 \text{ in } N(f)^{\perp}$. Let us define any arbitrary $x \in H$.

$$z = f(x)y_0 - f(y_0)x$$

Now $f(z) = f(x)f(y_0) - f(y_0)f(x) = 0$

$\Rightarrow \quad z \in N(f)$.

Since $y_0 \in N(f)^{\perp}$, we get

$$0 = (z, y_0) = \left(f(x)y_0 - f(y_0)x, y_0\right)$$

$$= f(x)(y_0, y_0) - f(y_0)(x, y_0)$$

Hence we get

$$f(x)(y_0, y_0) - f(y_0)(x, y_0) = 0 \qquad ...(3)$$

Notes

Noting that $(y_0, y_0) = \|y_0\|^2 \neq 0$, we get from (3),

$$f(x) = \left|\frac{f(y_0)}{\|y_0\|^2}\right|(x, y_0) \quad ...(4)$$

We can write (4) as

$$f(x) = \left[x, \frac{\overline{f(y_0)}}{\|y_0\|^2} y_0\right]$$

Now taking $\frac{\overline{f(y_o)}}{\|y_o\|} y_o$ as y, we have established that there exists a y such that f(x) = (x,y) for $x \in H$.

Step 2: In this step we know that

$$\|f\| = \|y\|$$

If f =0, then y = 0 and $\|f\| = \|y\|$ hold good.

Hence let $f \neq 0$. Then $y \neq 0$.

From the relation f(x) = (x,y) and Schwarz inequality we have

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

$$\Rightarrow \quad \sup_{\|x\| \neq 0} \frac{|f(x)|}{\|x\|} \leq \|y\|.$$

Using definition of norm of f, we get from above

$$\|f\| \leq \|y\| \quad ...(5)$$

Now let us take x = y in f(x) = (x,y), we get

$$\|y\|^2 = (y, y) = f(y) \leq \|f\|\|y\|$$

$$\Rightarrow \quad \|y\| \leq \|f\| \quad ...(6)$$

(5) and (6) implies that

$$\|f\| = \|y\|.$$

Step 3: We establish the uniqueness of y in f(x) = (x,y). Let us assume that y is not unique in f(x) = (x,y).

Let for all $x \in H, \exists\, y_1, y_2$ such that

$f(x) = (x,y_1) = (x,y_2)$

Then $(x,y_1) - (x,y_2) = 0$

$\Rightarrow \quad (x, y_1 - y_2) = 0 \;\forall x \in H.$

Let us choose x to be $y_1 - y_2$ so that

$$(y_1 - y_2, y_1 - y_2) = \|y_1 - y_2\|^2 = 0$$

$\Rightarrow \quad y_1 - y_2 = 0$

$\Rightarrow \quad y_1 = y_2$

$\Rightarrow$ y is unique in the representation of f(x) = (x,y)

This completes the proof of the theorem.

Note The above Riesz representation theorem does not hold in an inner product space which is not complete as shown by the example given below. In other words the completeness assumption cannot be dropped in the above theorem.

Example: Let us consider the subspace M of l_2 consisting of all finite sequences. This is the set of all scalar sequence whose terms are zero after a finite stage. It is an incomplete inner product space with inner product

$$(x, y) = \sum_{n=1}^{\infty} x_n \bar{y}_n \forall x, y \in M$$

Now let us define

$$f(x) = \sum_{n=1}^{\infty} \frac{x_n}{n} \text{ as } x = (x_n) \in M.$$

Linearity of f together with Hölder's inequality yields

$$|f(x)|^2 \leq \left(\sum_{n=1}^{\infty} \frac{|x_n|}{n}\right)^2 \leq \left(\sum_{n=1}^{\infty} \frac{1}{n^2}\right)\left(\sum_{n=1}^{\infty} |x_n|^2\right)$$

$$\leq \frac{\pi^2}{6}(x, x) = \frac{\pi^2}{6}\|x\|^2,$$

since $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

$\Rightarrow$ f is a continuous linear functional on M.

We now prove that there is no $y \in M$ such that

$$f(x) = (x, y) \forall x \in M.$$

Let us take x = e_n = (0,0,....,1,0,0,.....) where 1 is in n^{th} place.

Using the definition of f we have f(x) = $\frac{1}{n}$.

Suppose $y = (y_n) \in M$ satisfying the condition of the theorem, then

$$f(x) = (x, y) = \sum x_n \overline{y}_n = \overline{y}_n \text{ as } x = e_n.$$

Thus Riesz representation theorem is valid if and only if $y_n = \frac{1}{n} \neq 0$ for every n.

Hence $y = (y_n) \notin M$.

$\Rightarrow$ $\exists$ no $y \in M$ such that f(x,y) = (x,y) for every $x \in H$.

$\Rightarrow$ the completeness assumption cannot be left out from the Riesz-representation theorem.

Theorem 3: The mapping $\phi : H \rightarrow H^*$ defined by $\phi : H \rightarrow H^*$ defined by $\phi(y) = fy$ where fy(x) = (x,y) for every $x \in H$ is an (i) additive, (ii) one-to-one, (iii) onto, (iv) symmetry, (v) not linear.

Proof:

(i) Let us show that ϕ is additive, i.e.,

$$\phi(y_1 + y_2) = \phi(y_1) + \phi(y_2) \text{ for } y_1, y_2 \in H.$$

Now from the definition $\phi(y_1 + y_2) = f_{y1+y2}$

Hence for every $x \in H$, we get

$$f_{y_1+y_2}(x) = (x, y_1 + y_2) = (x, y_1) + (x, y_2)$$

$$= f_{y_1}(x) + f_{y_2}(x)$$

$\Rightarrow$ $f_{y_1+y_2} = \phi(y_1 + y_2) = f_{y_1} + f_{y_2} = \phi(y_1) + \phi(y_2)$

(i) ϕ is one-to-one. Let $y_1, y_2 \in H$

Then $\phi(y_1) = f_{y_1}$ and $\phi(y_2) = f_{y_2}$. Then

$\phi(y_1) = \phi(y_2) \Rightarrow f_{y_1} = f_{y_2}$

$\Rightarrow$ $f_{y_1}(x) = f_{y_2}(x) \forall x \in H.$ ...(1)

$f_{y_1}(x) = (x, y_1)$ and $f_{y_2}(x) = (x, y_2)$

$\therefore$ from (1), we get

$$(x, y_1) = (x, y_2) \Rightarrow (x, y_1) - (x, y_2) = 0$$

$\Rightarrow$ $(x, y_1 - y_2) = 0 \forall x \in H$ ...(2)

Choose x = $y_1 - y_2$ then from (2) if follows that $(y_1 - y_2, y_1 - y_2) = 0$

Notes

$\Rightarrow \quad \|y_1 - y_2\|^2 = 0$

$\Rightarrow \quad y_1 = y_2$

$\therefore \quad \phi$ is one-to-one.

(iii) ϕ is onto: Let $f \in H^*$. Then $\exists\ y \in H$ such that

$f(x) = (x,y)$

since $f(x) = (x,y)$ we get

$f = f_y$ so that $\phi(y) = f_y = f$.

Hence for $f \in H^*, \exists$ a pre-image $y \in H$. Therefore ϕ is onto.

(iv) ϕ is isometry; let $y_1, y_2 \in H$, then

$$\|\phi(y_1) - \phi(y_2)\| = \|f_{y_1} - f_{y_2}\|$$

$$= \left\|f_{y_1} - f_{(-y_2)}\right\|$$

But $\left\|f_{y_1} + f_{-y_2}\right\| = \left\|f_{y_1 - y_2}\right\| = \|y_1 - y_2\|$ (By theorem (1))

Hence $\|\phi(y_1) - \phi(y_2)\| = \|y_1 - y_2\|$.

(v) To show ϕ is not linear, let $y \in H$ and α be any scalar. Then $\phi(\alpha, y) = f\ \alpha\ y$. Hence for any $x \in H$, we get

$\Rightarrow \quad f_{\alpha y}(x) = (x, \alpha y) = \bar{\alpha}(x, y) = \bar{\alpha} f_y(x)$

$\Rightarrow \quad f_{\alpha y} = \bar{\alpha} f_y$

$\Rightarrow \quad \phi(\alpha y) = \bar{\alpha}\phi(y)$

$\Rightarrow \quad \phi$ is not linear. Such a mapping is called conjugate linear.

This completes the proof of the theorem.

Note: The above correspondence ϕ is referred to as natural correspondence between H and H^*.

Theorem 4: If H is a Hilbert space, then H^* is also an Hilbert space with the inner product defined by

$$(f_x, f_y) = (y,x) \qquad \dots (1)$$

Proof: We shall first verify that (1) satisfies the condition of an inner product.

Let $x, y \in H$ and α, β be complex scalars.

(i) We know (see Theorem 3) that

$f_{\alpha y} = \bar{\alpha} f_y$

$\Rightarrow \quad f_{\bar{\alpha} y} = \bar{\bar{\alpha}} f_y = \alpha f_y$.

Now $\left(\alpha f_x + \beta f_y, f_z\right) = \left(f_{\bar{\alpha} x} + f_{\bar{\beta} y}, f_z\right)$ ... (2)

Notes

But $\left(f_{\bar{\alpha}x}+f_{\bar{\beta}y},fz\right)=\left(z,\bar{\alpha}x,\bar{\beta}y\right)$ (by (1))

Now $\left(z,\bar{\alpha}x+\bar{\beta}y\right)=\bar{\bar{\alpha}}(z,x)+\bar{\bar{\beta}}(z,y)$

$$=\alpha\left(f_x,f_z\right)+\beta\left(f_y,f_z\right) \quad \dots (3)$$

From (2) and (3) it follows that $\left(\alpha f_x+\beta f_y,f_z\right)=\alpha\left(f_x,f_z\right)+\beta\left(f_y,f_z\right)$

(ii) $\left(\overline{f_x,f_y}\right)=\left(\overline{y,x}\right)=(x,y)=\left(f_x,f_y\right)$.

(iii) $\left(f_x,f_x\right)=(x,x)=\|x\|^2=\|f_x\|^2$ so $\left(f_x,f_x\right)\geq 0$ and $\|f_x\|=0\Leftrightarrow f_x=0$.

$(i)\rightarrow(iii)$ implies that (1) represents an inner product. Now the Hilbert space H is a complete normed linear space. Hence its conjugate space H^* is a Banach space with respect to the norm defined on H^*. Since the norm on H^* is induced by the inner product, H^* is a Hilbert space with the inner product $(f_x, f_y) = (y,x)$

This completes the proof of the theorem.

Cor. The conjugate space H^{**} of H^* is a Hilbert space with the inner product defined as follows:

If $f,g\in H^*$, let F_f and F_g be the corresponding elements of H^{**} obtained by the Riesz representation theorem.

Then $(F_f,F_g) = (g,f)$ defines the inner product of H^{**}.

Theorem 5: Every Hilbert space is reflexive.

Proof: We are to show that the natural imbedding on H and H^{**} is an isometric isomorphism.

Let x be any fixed element of H. Let F_x be a scalar valued function defined on H^* by $F_x(f) = f(x)$ for every $f\in H^*$. We have already shown in the unit of Banach spaces that $F_x\in H^{**}$. Thus each vector $x\in H$ gives rise to a functional F_x in H^{**}. F_x is called a functional on H^* induced by the vector x.

Let $J:H\rightarrow H^{**}$ be defined by $J(x)=F_x$ for every $x\in H$.

We have also shown in chapter of Banach spaces that J is an isometric isomorphism of H into H^{**}. We shall show that J maps H onto H^{**}.

Let $T_1:H\xrightarrow{\text{into}}H^*$ defined by

$T_1(x)=f_x, f_x(y)=(y,x)$ for every $y\in H$.

and $T_2:H^*\xrightarrow{\text{into}}H^{**}$ defined by

$T_2(f_x)=F_{f_x}, F_{f_x}(f)=(f,f_x)$ for $f\in H^*$.

Then $T_2.T_1$ is a composition of T_2 and T_1 from H to H^{**}. By Theorem 3, T_1,T_2 are one-to-one and onto.

Hence $T_2.T_1$ is same as the natural imbedding J.

For this we show that $J(x) = (T_2.T_1)x$ for every $x\in H$.

Now $(T_2.T_1)x = T_2(T_1(x)) = T_2(f_x) = F_{f_x}$.

By definition of J, J(x) = F_x. Hence to show $T_2.T_1$ = J, we have to prove that $F_x = F_{f_x}$.

For this let $f \in H^*$. Then f = f_y where f corresponds to y in the representation $F_{f_x}(f) = (f, f_x) = (f_y, f_x) = (x, y)$.

But (x,y) = $f_y(x)$ = f(x) = $F_x(f)$.

Thus we get $F_{f_x}(f) = F_x(f)$ for every $f \in H^*$.

Hence the mapping F_{f_x} and F_x are equal.

$\Rightarrow$ $T_2.T_1$ = J and J is a mapping of H onto H^{**}, so that H is reflexive.

This completes the proof of the theorem.

Notes

1. Since $F_x = F_{f_x} \forall x \in H$ (From above theorem)

   $\therefore \left(F_x, F_y\right) = \left(F_{f_y}, F_{f_y}\right) = \left(f_y, f_x\right) = (x, y)$ by using def. of inner product on H^{**} and by the def. of inner product on H^*.

2. Since $\exists$ an isometric isomorphism of the Hilbert space H onto Hilbert space H^{**}, therefore we can say that Hilbert space H and H^{**} are congruent i.e. they are equivalent metrically as well as algebraically. We can identify the space H^{**} with the space H.

12.2 Summary

- Let H be a Hilbert space. If f is a functional on H, then f will be continuous linear functional on H. The set $\beta(H,C)$ of all continuous linear functional on H is denoted by H^* and is called conjugate space of H. Conjugate space of a Hilbert space H is the conjugate space H^* of H.
- Riesz-representation theorem for continuous linear functional on Hilbert space:

  Let H be a Hilbert space and let f be an arbitrary functional on H^*. Then there exists a unique vector y in H such that f = fy, i.e. f(x) = (x,y) for every vector $x \in H$ and $\|f\| = \|y\|$.

12.3 Keywords

Continuous Linear Functionals: Let N be a normal linear space. Then we know that the set R of real numbers and the set C of complex numbers are Banach spaces with the norm of any $x \in R$ or $x \in C$ given by the absolute value of x. We denote the BANACH space $\beta(N,R)$ or $\beta(N,C)$ by N^*.

The elements of N^* will be referred to as continuous linear functionals on N.

Hilbert space: A complete inner product space is called a Hilbert space.

Let H be a complex Banach space whose norm arises from an inner product which is a complex function denoted by (x,y) satisfying the following conditions:

H_1: $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$

H_2: $\overline{(x, y)} = (y, x)$

H_3: $(x, x) = \|x\|^2$

for all $x, y, z \in H$ and for all $\alpha, \beta \in C$.

Inner Product: Let X be a linear space over the field of complex numbers C. An inner product on X is a mapping from X × X → C which satisfies the following conditions:

(i) $(\alpha x + \beta y, z) = \alpha (x, z) + \beta (y, z) \;\forall\; x, y, z \in X$ and $\alpha, \beta \in C$.

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

(iii) $(x, x) \geq 0$, $(x, x) = 0 \Leftrightarrow x = 0$

Riesz-representation Theorem for Continuous Linear Functional on a Hilbert Space: Let H be a Hilbert space and let f be an arbitrary functional on H^*. Then there exists a unique vector y in H such that

f = fy, i.e. f(x) = (x,y) for every vector $x \in H$ and $\|f\| = \|y\|$.

The Conjugate Space H^*: Let H be a Hilbert space. If f is a functional on H, then f will be continuous linear functional on H. The set $\beta(H, C)$ of all continuous linear functional on H is denoted by H^* and is called the conjugate space of H. The conjugate space of a Hilbert space H is the conjugate space H^* of H is in some sense is same as H itself.

12.4 Review Questions

1. Let H be a Hilbert space, and show that H* is also a Hilbert space with respect to the inner product defined by $(f_x, f_y) = (y, x)$. In just the same way, the fact that H* is a Hilbert space implies that H** is a Hilbert space whose inner product is given by $(F_f, F_g) = (g, f)$.
2. Let H be a Hilbert space. We have two natural mappings of H onto H**, the second of which is onto: the Banach space natural imbedding $x \to F_x$, where $f_x(y) = (y, x)$ and $F_{f_x}(f) = (F, f_x)$. Show that these mappings are equal, and conclude that H is reflexive. Show that $(F_x, F_y) = (x, y)$.

12.5 Further Readings

Books

Hausmann, Holm and Puppe, *Algebraic and Geometric Topology*, Vol. 5, (2005)

K. Yosida, *Functional Analysis*, Academic Press, 1965.

Online links

www.spot.colorado.edu

www.arvix.org

Unit 13: The Adjoint of an Operator

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define the adjoint of an operator.
- Understand theorems on adjoint of an operator.
- Solve problems on adjoint of an operator.

Introduction

We have already proved that T gives rise to an unique operator T* and H* such that (T*f) (x) = f(Tx) $\forall f \in H^*$ and $\forall x \in H$. The operator T* on H* is called the conjugate of the operator T on H.

In the definition of conjugate T* of T, we have never made use of the correspondence between H and H*. Now we make use of this correspondence to define the operator T* on H called the adjoint of T. Though we are using the same symbol for the conjugate and adjoint operator on H, one should note that the conjugate operator is defined on H*, while the adjoint is defined on H.

13.1 Adjoint of an Operator

Let T be an operator on Hilbert space H. Then there exists a unique operator T* on H such that

$(Tx,y)= (x,T^*y)$ for all $x, y \in H$

The operator T* is called the adjoint of the operator T.

Theorem 1: Let T be an operator on Hilbert space H. Then there exists a unique operator T* on H such that

$(Tx,y)= (x,T^*y)$ for all $x, y \in H$ ...(1)

The operator T* is called the adjoint of the operator T.

Proof: First we prove that if T is an operator on H, there exists a mapping T* on H onto itself satisfying

$(Tx,y)= (x,T^*y)$ for all $x, y \in H$. ...(2)

Notes

Let y be a vector in H and f_y its corresponding functional in H*.

Let us define

$$T^*: H \xrightarrow{\text{into}} H^* \text{ by}$$

$$T^*: f_y = f_z \quad \text{...(3)}$$

Under the natural correspondence between H and H*, let $z \in H$ corresponding to $f_z \in H^*$. Thus starting with a vector y in H, we arrive at a vector z in H in the following manner:

$$y \to f_y \to T^* f_y = f_z \to z, \quad \text{...(4)}$$

where $T^*: H^* \to H^*$ and $y \to f_y$ and $z \to f_z$ are on H to H*

under the natural correspondence. The product of the above three mappings exists and it is denoted by T*.

Then T* is a mapping on H into H such that

$$T^* y = z.$$

We define this T* to be the adjoint of T. We note that if we identify H and H* by the natural correspondence $y \to f_y$, then the conjugate of T and the adjoint of T are one and the same.

After establishing, the existence of T*, we now show (1). For $x \in H$, by the definition of the conjugate T* on an operator T,

$$\left(T^* f_y\right) x = f_y(Tx) \quad \text{...(5)}$$

By Riesz representation theorem,

$y \to f_y$ so that

$$f_y(Tx) = (Tx, y) \quad \text{...(6)}$$

Since T* is defined on H*, we get

$$\left(T^* f_y\right) x = f_z(x) = (x, z) \quad \text{...(7)}$$

But we have from our definition T*y = z ...(8)

From (5) and (6) it follows that

$$\left(T^* f_y\right) x = (Tx, y) \quad \text{...(9)}$$

From (7) and (8) it follows that

$$\left(T^* f_y\right) x = (x, T^* y) \quad \text{...(10)}$$

From (9) and (10), we thus obtain

$$(Tx, y) = (x, T^* y) \forall x, y \in H.$$

This completes the proof of the theorem.

Note The relation $(Tx, y) = (x, T^* y)$ can be equivalently written as

$(T^* x, y) = (x, Ty)$ since

$(T^* x, y) = (\overline{y, T^* x}) = (\overline{Ty, x}) = (\overline{\overline{x, Ty}}) = (x, Ty)$

$\Rightarrow$ $(T^* x, y) = (x, Ty)$.

Example: Find adjoint of T if T is defined on ℓ_2 as $Tx = (0, x_1, x_2, ...)$ for every $x = (x_n) \in \ell_2$.

Let T* be the adjoint of T. Using inner product in ℓ_2, we have

$$(T^* x, y) = (x, Ty)$$

since $Ty = (0, y_1, y_2, ...)$, we have

$$(T^* x, y) = (x, Ty) = \sum_{n=1}^{\infty} x_{n+1} \bar{y}_n = (Sx, y),$$

where $S(x) = (x_2, x_3, ...)$

Hence $(T^* x, y) = (Sx, y)$ for every x in ℓ_2.

Since T* is unique, T*=S so that we have

$$T^*(x) = (x_2, x_3, x_4, ...).$$

Theorem 2: Let H be the given Hilbert space and T* be adjoint of the operator T. Then T* is a bounded linear transformation and T determine T* uniquely.

Proof: T* is linear.

Let $y_1, y_2 \in H$ and α, β be scalars. Then for $x \in H$, we have

$$(x, T^*(\alpha y_1 + \beta y_2)) = (Tx, \alpha y_1 + \beta y_2)$$

But $(Tx, \alpha y_1 + \beta y_2) = \bar{\alpha}(Tx, y_1) + \bar{\beta}(Tx, y_2)$

$$= \bar{\alpha}(Tx, y_1) + \bar{\beta}(x, T^* y_2)$$

$$= (x, \alpha T^* y_1) + (x, \beta T^* y_2).$$

Hence for any $x \in H$,

$$(x, T^*(\alpha y_1 + \beta y_2)) = (x, \alpha T^* y_1) + (x, \beta T^* y_2)$$

Notes

$$= \left(x, \alpha T^* y_1 + \beta T^* y_2\right).$$

$\Rightarrow$ T* is linear.

T* is bounded

for any $y \in H$, let us consider

$$\|T^* y\|^2 = (T^* y, T^* y)$$

$$= |(TT^* yy)|$$

$$\leq \|TT^* y\| \; \|y\| \text{(using Schwarz inequality)}$$

$$\leq \|T\| \; \|T^* y\| \; \|y\|$$

Hence $\|T^* y\|^2 \leq \|T\| \; \|T^* y\| \; \|y\| > 0$ ...(1)

If $\|T^* y\| = 0$ then $\|T^* y\| \leq \|T\|\|y\|$ because $\|T\|\|y\| > 0$

Hence let $\|T^* y\| \neq 0.$

Then we get from (1)

$$\|T^* y\| \leq \|T\| \; \|y\|.$$

since T is bounded,

$\|T\| \leq M$ so that

$\|T^* y\| \leq M\|y\|$ for every $y \in H.$

$\Rightarrow$ T* is bounded.

$\Rightarrow$ T* is continuous.

Uniqueness of T*.

Let if T* is not unique, let T′ be another mapping of H into H with property $(Tx, y) = (x, T^* y) \; \forall \; x, y \in H.$

Then we have

$$(Tx, y) = (x, T'y) \quad \text{...(2)}$$

and $$(Tx, y) = (x, T^* y) \quad \text{...(3)}$$

From (2) and (3) it follows that

$$(x, T'y) = (x, T^* y) \forall x, y \in H$$

$$\Rightarrow \quad (x, (T'y - T^* y)) = 0$$

$\Rightarrow \quad (x,(T'-T^*)y)=0 \forall x \in H$

$\Rightarrow \quad (T'-T^*)y = 0$ for every $y \in H$

Hence $T'y = T^*y$ for every $y \in H$.

$\Rightarrow \quad T = T^*$.

This completes the proof of the theorem.

Notes

1. We note that the zero operator and the identity operator I are adjoint operators. For,

   (i) $(x, 0^*y) = (0x, y) = (0, y) = 0 = (x, 0) = (x, 0y)$

   so from uniqueness of adjoint 0* = 0.

   (ii) (x,Iy) = (Ix,y) = (x,y) = (x,Iy)

   so from uniqueness of adjoint I*=I.

2. If H is only an inner product space which is not complete, the existence of T* corresponding to T in the above theorem is not guaranteed as shown by the following example.

Example: Let M be a subspace of L_2 consisting of all real sequences, each one containing only finitely many non-zero terms. M is an incomplete inner product space with the same inner product for ℓ_2 given by

$$(x, y) = \sum_{n=1}^{\infty} x_n y_n \qquad ...(1)$$

For each $x \in M$, define

$$T(x) = \left(\sum_{n=1}^{\infty} \frac{x_n}{n}, 0, 0, ...\right) \qquad ...(2)$$

Then from the definition, for $x, y \in M$,

$$T(x, y) = y_1 \sum_{n=1}^{\infty} \frac{x_n}{n}.$$

Now let $e_n = (0,0,...,1,0,...)$ where 1 is in the n^{th} place.

Then using (3) we obtain

$$(T_{e_n}, e_1) = 1.\sum \frac{e_n(j)}{j} = 1.\frac{1}{n}.$$

Now we check whether there is T* which is adjoint of T. Now $\left(e_n, T^*\left(e_1\right)\right) = T^*\left(e_1\right).e_n$, where the R.H.S. gives the component wise inner product. Since $T^*\left(e_1\right) \in M, T^*\left(e_1\right).e_n$ cannot be equal to $\frac{1}{n} \forall\ n = 1,2,...$

$\Rightarrow$ there is no T* on M such that

$$\left(T\left(e_n\right), e_1\right) = \left(e_n, T^*\left(e_1\right)\right)$$

Hence completeness assumption cannot be ignored from the hypothesis.

Notes

1. The mapping $T \to T^*$ is called the adjoint operation on $\beta(H)$.
2. From Theorem (2), we see that the adjoint operation is mapping $T \to T^*$ on $\beta(H)$ into itself.

Theorem 3: The adjoint operation $T \to T^*$ on $\beta(H)$ has the following properties:

(i) $\left(T_1 + T_2\right)^* = T_1^* + T_2^*$ (preserve addition)

(ii) $\left(T_1 T_2\right)^* = T_2^* T_1^*$ (reverses the product)

(iii) $(\alpha T)^* = \overline{\alpha} T^*$ (conjugate linear)

(iv) $\|T^*\| \geq \|T\|$

(v) $\|T^* T\| = \|T\|^2$

Proof: (i) For every $x, y \in H$, we have

$$\begin{aligned}(x, (T_1 + T_2)^*y &= ((T_1 + T_2)\, x, y) \quad \text{(By def. of adjoint)}\\ &= (T_1x + T_2x, y)\\ &= (T_1x, y) + (T_2x, y)\\ &= (x, T_1^*y) + (x, T_2^*y)\\ &= (x, T_1^*y + T_2^*y)\\ &= (x, (T_1^* + T_2^*)y)\end{aligned}$$

$\Rightarrow (T_1 + T_2)^* = T_1^* + T_2^*$ by uniqueness of adjoint operator

(ii) For every $x, y \in H$, we have

$$\begin{aligned}\left(x\left(T_1 T_2\right)^* y\right) &= \left(\left(T_1 T_2\right)x, y\right)\\ &= \left(T_1\left(T_2 x\right), y\right)\end{aligned}$$

Notes

$= (T_2 x, T_1 * y)$

$= (x, T_2^*, (T_1^* y))$

$= (x, (T_2^* T_1^*), y)$

Therefore from the uniqueness of adjoint operator, we have $(T_1 T_2)^* = T_2^* T_1^*$.

(iii) For every $x, y \in H$, we have

$(x, (\alpha T)^* y) = ((\alpha T)x, y) = (\alpha(Tx), y)$

$= \alpha(Tx, y)$

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

$= (x(\overline{\alpha} T^*) y)$.

Therefore from the uniqueness of adjoint operator, we have

$(\alpha T^*) = \overline{\alpha} T^*$.

(iv) For every $y \in H$ we have

$\|T^* y\|^2 = (T^* y, T^* y)$

$= (TT^* y, y)$

$= |(TT^* y, y)|$ $\quad\left[\because \|T^* y\|^2 = (TT^* y, y) \text{is a real number} \geq 0\right]$

$\leq \|TT^* y\| \; \|y\|$ $\quad$ (By Schwarz inequality)

$\leq \|T\| \; \|T^* y\| \; \|y\|$ $\quad [\because \|Tx\| \leq \|T\| \; \|x\|]$

Thus $\|T^* y\|^2 \leq \|T\| \|T^* y\| \|y\| \forall y \in H$

$\Rightarrow \|T^* y\| \leq \|T\| \|y\| \forall y \in H.$ ...(1)

Now $\|T^*\| = \text{Sup}\{\|T^* y\| : \|y\| \leq 1\}$

from (1), we see that if $\|y\| \leq 1$ then $\|T^* y\| \leq \|T\|$

$\Rightarrow \|T^*\| \leq \|T\|$ ...(2)

Now applying (2) from the operator T* in place of operator T, we get

$\|(T^*)^*\| \leq \|T^*\|$

$\Rightarrow \|T^{**}\| \le \|T^*\|$

$\Rightarrow \|T\| \le \|T^*\| \quad [\because T^{**} = T]$ ...(3)

From (2) and (3) it follows that

$\|T\| \le \|T^*\|$.

(v) We have $\|T^*T\| \le \|T^*\| \; \|T\|$

$= \|T\| \; \|T\| \quad [\because \|T^*\| = \|T\|]$

$= \|T\|^2$ ...(4)

Further for every $x \in H$, we have

$\|Tx\|^2 = (Tx, Tx)$

$= (T^*Tx, x)$

$= |((T^*T)x, x)|$

$\le \|T^*T\| \; \|x\| \|x\|$ (By Schwarz inequality)

$= \|T^*T\| \; \|x\|^2$

Then we have

$\|Tx\|^2 \le \|T^*T\| \; \|x\|^2 \; \forall x \in H$ ...(5)

Now $\|T\| = \sup\{\|Tx\| : \|x\| \le 1\}$

$\therefore \|T\|^2 = [\sup\{\|Tx\| : \|x\| \le 1\}]^2$

$= \sup\{\|Tx\|^2 : \|x\| \le 1\}$

From (5) we see that

if $\|x\| \le 1$, then $\|Tx\|^2 \le \|T^*T\|$.

Therefore, $\text{Sup}\{\|Tx\|^2 : \|x\| \le 1\} \le \|T^*T\|$

$\Rightarrow \|T\|^2 \le \|T^*T\|$. ...(6)

From (5) and (6) it follows that

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

This completes the proof of the theorem.

Notes

Cor: If (T_n) is a sequence of bounded linear operators on a Hilbert space and $T_n \to T$, then $T_n^* \to T^*$.

We have

$$\|T_n^* - T^*\| = \|(T_n - T)^*\|$$

$$= \|T_n - T\| \quad \text{(By properties of T*)}$$

Since $T_n \to T$ as $n \to \infty$

$\Rightarrow T_n^* \to T^*$ as $n \to \infty$.

Theorem 4: The adjoint operation on $\beta(H)$ is one-to-one and onto. If T is a non-singular operator on H, then T* is also non-singular and

$$(T^*)^{-1} = (T^{-1})^*.$$

Proof: Let $\phi : \beta(H) \to \beta(H)$ is defined by

$$\phi(T) = T^* \text{ for every } T \in \beta(H).$$

To show ϕ is one-to-one, let $T_1, T_2 \in \beta(H)$. Then we shall show that $\phi(T_1) = \phi(T_2) \Rightarrow T_1 = T_2$.

Now $\phi(T_1) = \phi(T_2)$

$\Rightarrow T^*_1 = T^*_2$

$\Rightarrow (T^*_1)^* = (T^*_2)^*$ (using Theorem 4. prop (iv))

$\Rightarrow T_1 = T_2$

$\Rightarrow \phi$ is one-to-one.

ϕ is onto:

For $T^* \in \beta(H)$, we have on using Theorem 4 (iv),

$$\phi(T^*) = (T^*)^* = T.$$

Thus for every $T^* \in \beta(H)$, there is a $T^* \in \beta(H)$ such that

$$\phi(T^*) = T \Rightarrow \phi \text{ is onto.}$$

Next let T be non-singular operator on H. Then its inverse T^{-1} exists on H and

$$TT^{-1} = T^{-1}T = I.$$

Taking the adjoint on both sides of the above, we obtain

$$\left(TT^{-1}\right)^* = \left(T^{-1}T\right)^* = I^*.$$

By using Theorem 4 and note (2) under Theorem 2, we obtain

$$\left(T^{-1}\right)^* T^* = T^*\left(T^{-1}\right)^* = I.$$

$\Rightarrow$ T^* is invertible and hence non-singular.

Further from the above, we conclude

$$(T^*)^{-1} = \left(T^{-1}\right)^*.$$

This completes the proof of the theorem.

Note From the properties of the adjoint operation $T \to T^*$ on $\beta(H)$ discussed in Theorems (3) and (4), we conclude that the adjoint operation $T \to T^*$ is one-to-one conjugate linear mapping on $\beta(H)$ into itself.

Example: Show that the adjoint operation is one-to-one onto as a mapping of β(H) into itself.

Solution: Let $\phi : \beta(H) \to \beta(H)$ be defined

$$\phi(T) = T^* \; \forall \; T \in \beta(H)$$

We show ϕ is one-to-one and onto.

ϕ is one-one:

Let $T_1, T_2 \in \beta(H)$. Then

$$\begin{aligned} \phi(T_1) = \phi(T_2) &\Rightarrow T_1^* = T_2^* \\ &\Rightarrow (T_1^*)^* = (T_2^*)^* \\ &\Rightarrow T_1^{**} = T_2^{**} \\ &\Rightarrow T_1 = T_2 \\ &\Rightarrow \phi \text{ is one-to-one.} \end{aligned}$$

ϕ is onto:

Let T be any arbitrary member of β(H). Then $T^* \in \beta(H)$ and we have $\phi(T^*) = (T^*)^* = T^{**} = T$. Hence, the mapping ϕ is onto.

13.2 Summary

- Let T be an operator on Hilbert Space H. Then there exists a unique operator T* on H such that $(Tx, y) = (x, T^*y)$ for all $x, y \in H$. The operator T* is called the adjoint of the operator T.

- The adjoint operation $T \to T^*$ on $\beta(H)$ has the following properties:

  (i) $(T_1 + T_2)^* = T_1^* + T_2^*$

  (ii) $(T_1 T_2)^* = T_2^* T_1^*$

  (iii) $(\alpha T)^* = \bar{\alpha} T^*$

  (iv) $\|T^*\| = \|T\|$

  (v) $\|T^* T\| = \|T\|^2$

13.3 Keywords

Adjoint of the Operator T: Let T be an operator on Hilbert space H. Then there exists a unique operator T* on H such that

$$(Tx, y) = (x, T^*y) \text{ for all } x, y \in H$$

The operator T* is called the adjoint of the operator T.

Conjugate of the Operator T on H: T gives rise to an unique operator T* and H* such that (T*f) (x) = f(Tx) $\forall f \in H^*$ and $\forall x \in H$. The operator T* on H* is called the conjugate of the operator T on H.

13.4 Review Questions

1. Show that the adjoint operation is one-to-one onto as a mapping of $\beta(H)$ into itself.
2. Show that $\|TT^*\| = \|T\|^2$.
3. Show that O*=O and I*=I. Use the latter to show that if T is non-singular, then T* is also non-singular, and that in this case $(T^*)^{-1} = (T^{-1})^*$.

13.5 Further Readings

Books

N.I. Akhiezer and I.M. Glazman, *Theory of Linear Operators in Hilbert Space*, Vol. II, Pitman, 1981.

K. Yosida, *Functional Analysis*, Academic Press, 1965.

Online links

www.math.osu.edu/~ gerlach.1/math.BVtypset/node 78.html.

sepwww.standford.edu/sep/prof/pvi/conj/paper_html/node10.html.

Notes

Unit 14: Self Adjoint Operators

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define self adjoint operator.
- Define positive operator.
- Solve problems on self adjoint operator.

Introduction

The properties of complex number with conjugate mapping $z \to \bar{z}$ motivate for the introduction of the self-adjoint operators. The mapping $z \to \bar{z}$ of complex plane into itself behaves like the adjoint operation in $\beta(H)$ as defined earlier. The operation $z \to \bar{z}$ has all the properties of the adjoint operation. We know that the complex number is real iff $z = \bar{z}$. Analogue to this characterization in $\beta(H)$ leads to the motion of self-adjoint operators in the Hilbert space.

14.1 Self Adjoint Operator

14.1.1 Definition: Self Adjoint

An operator T on a Hilbert space H is said to be self adjoint if T*=T.

We observe from the definition the following properties:

(i) O and I are self adjoint $(\because O^* = O \text{ and } I^* = I)$

(ii) An operator T on H is self adjoint if

$(Tx, y) = (x, Ty) \forall\ x, y \in H$ and conversely.

Notes

If T* is an adjoint operator T on H then we know from the definition that

$(Tx, y) = (x, T^* y) \forall\ x, y \in H$.

If T is self-adjoint, then T = T*.

$\therefore (Tx, y) = (x, Ty) \forall\ x, y \in H.$

Conversely, if $(Tx, y) = (x, Ty) \forall\ x, y \in H$ then we show that T is self-adjoint.

If T* is adjoint of T then (Tx, y) = (x, T*y)

$\therefore$ We have $(x, Ty) = (x, T^* y)$

$\Rightarrow (x, (T - T^*)y) = 0 \forall\ x, y \in H$

But since $x \neq 0 \Rightarrow (T - T^*)y = 0 \forall\ y \in H$

$\Rightarrow$ T = T*

$\Rightarrow$ T is self adjoint.

(iii) For any $T \in \beta(H), T + T^*$ and $T^* T$ are self adjoint. By the property of self-adjoint operators, we have

$$(T + T^*)^* = T^* + T^{**}$$
$$= T^* + T$$
$$= T + T^*$$

$\Rightarrow (T + T^*)^* = T + T^*,$

and $(T^* T)^* = T^* T^{**} = T^* T$

$\Rightarrow (T^* T^*)^* = T^* T.$

Hence $T + T^*$ and $T^* T$ are self adjoint.

Theorem 1: If (A_n) is a sequence of self-adjoint operators on a Hilbert space H and if (A_n) converges to an operator A, then A is self adjoint.

Proof: Let (A_n) be a sequence of self adjoint operators and let $A_n \to A.$

A_n is self adjoint $\Rightarrow$ $A_n^* = A_n$ for $n = 1, 2, ...$

We claim that $A = A^*$

Now $A - A^* = A - A_n + A_n - A_n^* + A_n^* - A^*$

$\Rightarrow \|A - A^*\| \leq \|A - A_n\| + \|A_n - A_n^*\| + \|(A_n - A)^*\|$

$\leq \|-(A_n - A)\| + \|A_n - A_n\| + \|A_n - A\|$ $\qquad [\because A_n^* = A_n]$

$= \|A_n - A\| + 0 + \|A_n - A\|$

Notes

$= 2\| A_n - A \|$

$\to 0$ as $n \to \infty$

$\Rightarrow \|A - A^*\| = 0$ or $A - A^* = 0 \Rightarrow A - A^*$

$\Rightarrow$ A is self-adjoint operator.

This completes the proof of the theorem.

Theorem 2: Let S be the set of all self-adjoint operators in $\beta(H)$. Then S is a closed linear subspace of $\beta(H)$ and therefore S is a real Banach space containing the identity transformation.

Proof: Clearly S is a non-empty subset of $\beta(H)$, since O is self adjoint operator i.e. $O \in S$.

$\beta(H)$, since O is self adjoint operator i.e. $O \in S$.

Let $A_1, A_2 \in S$, We prove that $\alpha A_1 + \beta A_2 \in S$.

$A_1, A_2 \in S \Rightarrow A_1^* = A_1$ and $A_2^* = A_2$ ...(1)

For $\alpha, \beta \in R$, we have

$(\alpha A_1 + \beta A_2)^* = (\alpha A_1)^* + (\beta A_2)^*$

$= \bar{\alpha} A_1^* + \bar{\beta} A_2^*$

$= \alpha A_1 + \beta A_2 \quad [\because \alpha, \beta \text{ are real numbers}, \therefore \bar{\alpha} = \alpha,\ \bar{\beta} = \beta]$

$\Rightarrow \alpha A_1 + \beta A_2$ is also a self adjoint operator on H.

$\Rightarrow A_1, A_2 \in S \Rightarrow \alpha A_1 + \beta A_2 \in S$.

$\Rightarrow$ S is a real linear subspace of $\beta(H)$.

Now to show that S is a closed subset of the Banach space $\beta(H)$. Let A be any limit point of S. Then $\exists$ a sequence of operator A_n is such that $A_n \to A$. We shall show that $A \in S$ i.e. $A = A^*$.

Let us consider

$\|A - A^*\| = \|A - A_n + A_n - A^*\|$

$\leq \|A - A_n\| + \|A_n - A^*\|$

$= \|A - A_n\| + \|A_n - A_n^*\| + \|A_n^* - A^*\|$

$\leq \|A - A_n\| + \|A_n - A_n^*\| + \|A_n^* - A^*\|$

$= \|-(A_n - A)\| + \|A_n - A_n\| + \|(A_n - A)^*\| \quad [\because A_n \in S \Rightarrow A_n^* = A_n]$

$= \|A_n - A\| + \|0\| + \|A_n - A\|$

$= 2\|A_n - A\|$ $\quad [\because \|0\| = 0, \|T^*\| = \|T\| \text{ and } \|-T\| = \|T\|]$

$\to 0$ as $A_n \to A$

$\therefore \|A - A^*\| = 0 \Rightarrow A - A^* = 0$

$\Rightarrow \quad A = A^* \Rightarrow A$ is self adjoint

$\Rightarrow \quad A \in S$

$\Rightarrow \quad S$ is closed.

Now since S is a closed linear subspace of the Banach space $\beta(H)$, therefore S is a real Banach space. ($\because$ S is a complete linear space)

Also $I^* = I \Rightarrow$ the identity operator $I \in S$.

This completes the proof of the theorem.

Theorem 3: If A_1, A_2 are self-adjoint operators, then their product A_1, A_2 is self adjoint $\Leftrightarrow A_1, A_2 = A_2, A_1$(i.e. they commute)

Proof: Let A_1, A_2 be two self adjoint operators in H.

Then $A_1^* = A_1, A_2^* = A_2$.

Let A_1, A_2 commute, we claim that A_1, A_2 is self-adjoint.

$$(A_1, A_2)^* = A_2^* A_1^* = A_2 A_1 = A_1 A_2$$

$\Rightarrow \quad (A_1, A_2)^* = A_1 A_2$

$\Rightarrow \quad A_1 A_2$ is self adjoint.

Conversely, let $A_1 A_2$ is self adjoint, then

$$(A_1 A_2)^* = A_1 A_2$$

$\Rightarrow \quad A_2^* A_1^* = A_1 A_2$

$\Rightarrow \quad A_1, A_2$ commute

This completes the proof of the theorem.

Theorem 4: If T is an operator on a Hilbert space H, then T = $T = 0 \Leftrightarrow (Tx, y) = 0 \ \forall \ x, y \in H$.

Proof: Let T = 0 (i.e. zero operator). Then for all x and y we have

$$T(x, y) = (Ox, y) = (O, y) = O.$$

Conversely, $(Tx, y) = O \forall \ x, y \in H$

$\Rightarrow \quad (Tx, Tx) = O \forall \ x, y \in H$ (taking y = Tx)

$\Rightarrow \quad Tx = O \forall \ x, y \in H$

Notes

$\Rightarrow \quad T = O$ i.e. zero operator.

This completes the proof of the theorem.

Theorem 5: If T is an operator on a Hilbert space H, then

$$(Tx, x) = 0 \forall\ x \text{ in } H \Leftrightarrow T = O.$$

Proof: Let T = O. Then for all x in H, we have

$$(Tx, x) = (Ox, x) = (0, x) = 0.$$

Conversely, let $(Tx, x) = 0 \forall\ x, y \in H$. Then we show that T is the zero operator on H.

If α, β any two scalars and x, y are any vectors in H, then

$$(T(\alpha x + \beta y), \alpha x + \beta y) = (\alpha Tx + \beta Ty, \alpha x + \beta y)$$

$$= \alpha(Tx, \alpha x + \beta y) + \beta(Ty, \alpha x + \beta y)$$

$$= \alpha\bar{\alpha}(Tx, x) + \alpha\bar{\beta}(Tx, y) + \beta\bar{\alpha}(Ty, x) + \beta\bar{\beta}(Ty, x)$$

$$= |\alpha|^2 (Tx, x) + \alpha\bar{\beta}(Tx, y) + \beta\bar{\alpha}(Ty, x) + |\beta|^2 (Ty, x)$$

$$\therefore (T(\alpha x + \beta y), \alpha x + \beta y) - |\alpha|^2 (Tx, x) - |\beta|^2 (Ty, y) = \alpha\bar{\beta}(Tx, y) + \beta\bar{\alpha}(Ty, x) \quad \text{...(1)}$$

But by hypothesis $(Tx, x) = 0 \forall x \in H$.

$\Rightarrow$ L.H.S. of (1) is zero, consequently the R.H.S. of (1) is also zero. Thus we have

$$\alpha\bar{\beta}(Tx, y) + \beta\bar{\alpha}(Ty, x) = 0 \quad \text{...(2)}$$

for all scalars α, β and $\forall x, y \in H$.

Putting $\alpha = 1, \beta = 1$ in (2) we get

$$(Tx, y) + (Ty, x) = 0 \quad \text{...(3)}$$

Again putting $\alpha = i, \beta = 1$ in (2) we obtain

$$i(Tx, y) - i(Ty, x) = 0 \quad \text{...(4)}$$

Multiply (3) by (i) and adding to (4) we get

$$2i(Tx, y) = 0 \forall x, y \in H$$

$\Rightarrow \quad (Tx, y) = 0 \forall x, y \in H$

$\Rightarrow \quad (Tx, Tx) = 0 \forall x, y \in H \qquad$ (Taking y = Tx)

$\Rightarrow \quad Tx = 0 \forall x, y \in H$

$\Rightarrow$ T = 0 (zero operator)

This completes the proof of the theorem.

Theorem 6: An operator T on a Hilbert space H is self-adjoint.

$\Leftrightarrow (Tx, x)$ is real for all x.

Proof: Let T* =T (i.e. T is self adjoint operator)

Then for every $x \in H$, we have

$$(Tx, Tx) = (x, T^* x) = (x, Tx) = (\overline{Tx, x})$$

$\Rightarrow (Tx, x)$ equals its own conjugate and is therefore real.

Conversely, let (Tx, x) is real $\forall x \in H.$ We claim that T is self adjoint i.e. T*=T.

since (Tx, x) is real $\forall x \in H,$

$$\therefore \quad (Tx, x) = (\overline{Tx, x}) = (\overline{x, T^* x}) = (T^* x, x)$$

$$\Rightarrow \quad (Tx, x) - (T^* x, x) = 0 \forall x \in H$$

$$\Rightarrow \quad (Tx - T^* x, x) = 0 \forall x \in H$$

$$\Rightarrow \quad ((T - T^*)x, x) = 0 \forall x \in H$$

$\Rightarrow$ T – T* = 0 $\quad [\because$ if $(Tx, x) = 0 \Rightarrow T = 0]$

$\Rightarrow$ T = T*

$\Rightarrow$ T is self adjoint.

This completes the proof of the theorem.

Cor. If H is real Hilbert space, then A is self adjoint

$$\Leftrightarrow (Ax, y) = (Ay, x) \forall x, y \in H.$$

A is self adjoint $\Leftrightarrow$ for any $x, y \in H.$

$$(Ax, y) = (x, A^* y) = (\overline{A^* y, x}).$$

since H is real Hilbert space $(\overline{A^* y, x}) = (A^* y, x)$ so that $(Ax, y) = (Ay, x)$ $\quad [\because A^* = A]$

Theorem 7: The real Banach space of all self-adjoint operators on a Hilbert space H is a partially ordered set whose linear and order structures are related by the following properties:

(a) If $A_1 \le A_2$ then $A_1 + A \le A_2 + A$ for every $A \in S$;

(b) If $A_1 \le A_2$ and $\alpha \ge 0$, then $\alpha A_1 \le \alpha A_2$.

Proof: Let S represent the set of all self-adjoint operators on H. We define a relation ≤ on S as follows:

If $A_1 A_2 \in S$, we write $A_1 \le A_2$ if $(A_1 x, x) \le (A_2, x) \forall$ x in H.

We shall show that '≤' is a partial order relation on S. '≤' is reflexive.

Notes

Let $A \in S$. Then

$$(Ax,x)=(Ax,x)\forall x\in H$$

$\Rightarrow$ $(Ax,x)\le(Ax,x)\forall x\in H$

$\Rightarrow$ By definition $A\le A$.

$\Rightarrow$ $'\le'$ on S is reflexive.

$'\le'$ is transitive.

Let $A_1\le A_2$ and $A_2\le A_3$ then

$$(A_1x,x)\le(A_2x,x)\forall x\in H.$$

and $(A_2x,x)\le(A_3x,x)\forall x\in H.$

From these we get

$$(A_1x,x)\le(A_3x,x)\forall x\in H.$$

and $(A_2x,x)\le(A_3x,x)\forall x\in H.$

From these we get

$$(A_1x,x)\le(A_3x,x)\forall x\in H.$$

Therefore by definition $A_1\le A_3$ and so the relation is transitive.

$'\le'$ is anti-symmetric.

Let $A_1\le A_2$ and $A_2\le A_1$ then to show that $A_1=A_2$.

We have $A_1\le A_2\Rightarrow(A_1x,x)\le(A_2x,x)\forall x\in H.$

Also $A_2\le A_1\Rightarrow(A_2x,x)\le(A_1x,x)\forall x\in H.$

From these we get

$$(A_1x,x)=(A_2x,x)\forall x\in H.$$

$\Rightarrow$ $(A_1x-A_2x,x)=0\forall x\in H.$

$\Rightarrow$ $((A_1-A_2)x,x)=0\forall x\in H.$

$\Rightarrow$ $A_1-A_2=0$

$\Rightarrow$ $A_1=A_2$

$\Rightarrow$ $'\le'$ on anti-symmetric.

Hence $'\le'$ is a partial order relation on S.

Now we shall prove the next part of the theorem.

(a) We have $A_1\le A_2\Rightarrow(A_1x,x)\le(A_2x,x)\forall x\in H.$

$\Rightarrow \quad (A_1x,x)+(Ax,x)\le(A_2x,x)+(Ax,x)\forall x\in H.$

$\Rightarrow \quad ((A_1+A_2)x,x)\le((A_1+A)x,x)\forall x\in H.$

$\Rightarrow \quad A_1+A_2\le A_2+A$, by def. of $\le$.

(b) We have $A_1\le A_2\Rightarrow(A_1x,x)\le(A_2x,x)\forall x\in H$

$\Rightarrow \quad \alpha(A_1x,x)\le\alpha(A_2x,x)\forall x\in H \qquad [\because \alpha\ge 0]$

$\Rightarrow \quad (\alpha A_1x,x)\le(\alpha A_2x,x)\forall x\in H$

$\Rightarrow \quad ((\alpha A_1)x,x)\le((\alpha A_2)x,x)\forall$ x in H

$\Rightarrow \quad \alpha A_1\le\alpha A_2$, by def. of '$\le$'

This completes the proof of the theorem.

14.1.2 Definition – Positive Operator

A self adjoint operator on H is said to be positive if $A\ge 0$ in the order relation. That is

if $\quad (Ax,x)\ge 0\forall x\in H.$

We note the following properties from the above definition.

(i) Identity operator I and the zero operator O are positive operators.

Since I and O are self adjoint and

$(Ix,x)=(x,x)=\|x\|^2\ge 0$

also (Ox, x) = (0, x) = 0

$\Rightarrow I, O$ are positive operators.

(ii) For any arbitrary T on H, both TT* and T*T are positive operators. For, we have

$(TT^*)^*=(T^*)^*T^*=TT^*$

$\Rightarrow TT^*$ is self adjoint

Also $(T^*T)^*=T^*(T^*)^*=T^*T$

$\Rightarrow T^*T$ is self adjoint

Further we see that

$(TT^*x,x)=(T^*x,T^*x)=\|T^*x\|^2\ge 0$

and $(T^*Tx,x)=(Tx,T^{**}x)=(Tx,Tx)=\|Tx\|^2\ge 0$

Therefore by definition both TT* and T*T are positive operators.

Theorem 8: If T is a positive operator on a Hilbert space H, then I+T is non-singular.

Proof: To show I+T is non-singular, we are to show that I+T is one-one and onto as a mapping of H onto itself.

I+T is one-one.

First we show $(I+T)x=0\Rightarrow x=0$

We have $(I+T)x = 0 \Rightarrow Ix + Tx = 0 \Rightarrow x + Tx = 0$

$$\Rightarrow Tx = -x$$

$\Rightarrow \quad (Tx, x) = (-x, x) = -\|x\|^2$

$\Rightarrow \quad -\|x\|^2 \geq 0 \quad [\because (Tx, x) \geq 0]$

$\Rightarrow \quad \|x\|^2 \leq 0$

$\Rightarrow \quad \|x\|^2 = 0 \quad [\because \|x\|^2 \text{ is always} \geq 0]$

$\Rightarrow \quad x = 0$

$\therefore (I+T)x = 0 \Rightarrow x = 0.$

Now $(I+T)x = (I+T)y \Rightarrow (I+T)(x-y) = 0$

$\Rightarrow \quad x - y = 0 \Rightarrow x = y$

Hence I+T is one-one.

I+T is onto.

Let M = range of I+T. Then I+T will be onto if we prove that M=H.

We first show that M is closed.

For any $x \in H$, we have

$$\|(I+T)x\|^2 = \|x + Tx\|^2$$

$$= (x + Tx, x + Tx)$$

$$= (x, x) + (x, Tx) + (Tx, x) + (Tx, Tx)$$

$$= \|x\|^2 + \|Tx\|^2 + \left(\overline{Tx, x}\right) + (Tx, x)$$

$$= \|x\|^2 + \|Tx\|^2 + 2(Tx, x) \qquad [\because T \text{ is positive} \Rightarrow T \text{ is self-adjoint } \Rightarrow (Tx,x) \text{real}]$$

$$\geq \|x\|^2 \qquad [\because T \text{ is positive} \Rightarrow (Tx, x) \geq 0]$$

Thus $\|x\| \leq \|(I+T)x\| \forall x \in H$

Now let $\langle (I+T)x_n \rangle$ be a CAUCHY sequence in M. For any two positive integers m,n we have

$$\|x_m - x_n\| \leq \|(I+T)(x_m - x_n)\|$$

$$= \|(I+T)x_m - (I+T)x_n\| \to 0,$$

since $\langle (I+T)x \rangle$ is a CAUCHY sequence.

$\therefore \|x_m - x_n\| \to 0$

$\Rightarrow \langle x_n \rangle$ is a CAUCHY sequence in H. But H is complete. Therefore by CAUCHY sequence $\langle x_n \rangle$ in H converges to a vector, say x in H.

Notes

Now $\text{Lim}\{(I+T)x_n\} = (I+T)(\lim x_n)$ $[\because I+T \text{ is a continuous mapping}]$

$$= (I+T)x \in M(\text{range of } I+T)$$

Thus the CAUCHY sequence $\langle (I+T)x_n \rangle$ in M converges to a vector $(I+T)x$ in M.

$\Rightarrow$ every CAUCHY sequence in M is a convergent sequence in M.

$\Rightarrow$ M is complete subspace of a complete space is closed.

$\Rightarrow$ M is closed.

Now we show that M = H. Let if possible $M \neq H$.

Then M is a proper closed subspace of H.

Therefore, $\exists$ a non-zero vector x_0 in H s.t. x_0 is orthogonal in M.

Since $(I+T)x_0 \in M,$ therefore

$x_0 \perp M \Rightarrow (\{I+T\}x_0, x) = 0$

$\Rightarrow (x_0 + Tx_0, x_0) = 0$

$\Rightarrow (x_0, x_0) + (Tx_0, x_0) = 0$

$\Rightarrow \|x\|^2 + (Tx_0, x_0) = 0$

$\Rightarrow -\|x\|^2 = (Tx_0, x_0)$

$\Rightarrow -\|x\|^2 \geq 0$ $\quad [\because T \text{ positive} \Rightarrow (Tx_0, x_0) \geq 0]$

$\Rightarrow \|x\|^2 \leq 0$

$\Rightarrow \|x\| = 0$ $\quad [\because \|x\|^2 \geq 0]$

$\Rightarrow x = 0$

$\Rightarrow$ a contradiction to the fact that $x_0 \neq 0$.

Hence we must have M = H and consequently I+T is onto. Thus I+T is non-singular.

This completes the proof of the theorem.

Cor. If T is an arbitrary operator on H, then the operator I+TT* and I+T*T are non-singular.

Proof: We know that for an arbitrary T on H, T*T and TT* are both positive operators.

Hence by Theorem (8) both the operators I+TT* and I+T*T are non-singular.

14.2 Summary

- An operator T on a Hilbert space H is said to be self adjoint if T*=T.
- A self adjoint operator on H is said to be positive if $A \geq 0$ in the order relation. That is if $(Ax, x) \geq 0 \forall x \in H.$

14.3 Keywords

Positive Operator: A self adjoint operator on H is said to be positive if $A \geq 0$ in the order relation. That is

if $\quad (Ax, x) \geq 0 \forall x \in H.$

Self Adjoint: An operator T on a Hilbert space H is said to be self adjoint if T*=T.

14.4 Review Questions

1. Define a new operation of "Multiplication" for self-adjoint operators by $A_1 \circ A_2 = \frac{(A_1A_2 + A_2A_1)}{2}$, and note that $A_1 \circ A_2$ is always self-adjoint and that it equals A_1A_2 whenever A_1 and A_2 commute. Show that this operation has the following properties:

   $A_1 \circ A_2 = A_2 \circ A_1,$

   $A_1 \circ (A_2 + A_3) = A_1 \circ A_2 + A_1 \circ A_3,$

   $\alpha(A_1 \circ A_2) = (\alpha A_1) \circ A_2 = A_1 \circ (\alpha A_2),$

   and $A \circ I = I \circ A = A$. Show that

   $A_1 \circ (A_2 \circ A_3) = (A_1 \circ A_2) \circ A_3$ whenever A_1 and A_3 commute.

2. If T is any operator on H, it is clear that $|(Tx, x)| \leq \|Tx\|\|x\| \leq \|T\|\|x\|^2$; so if $H \neq \{0\}$, we have $\sup\{|(Tx, x)|/\|x\|^2 : x \neq 0\} \leq \|T\|$. Prove that if T is self-adjoint, then equality holds here.

14.5 Further Readings

Books

Akhiezer, N.I.; Glazman, I.M. (1981), *Theory of Linear Operators in Hilbert Space*

Yosida, K., *Functional Analysis*, Academic Press

Online links

www.ams.org/bookstore/pspdf/smfams-14-prev.pdf-UnitedStatesmath
world.wolfram.com

Unit 15: Normal and Unitary Operators

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand the concept of Normal and Unitary operators.
- Define the terms Normal, Unitary and Isometric operator.
- Solve problems on normal and unitary operators.

Introduction

An operator T on H is said to be normal if it commutes with its adjoint, that is, if TT*=T*T. We shall see that they are the most general operators on H for which a simple and revealing structure theory is possible. Our purpose in this unit is to present a few of their more elementary properties which are necessary for our later work. In this unit, we shall also study about Unitary operator and Isometric operator.

15.1 Normal and Unitary Operators

15.1.1 Normal Operator

Definition: An operator T on a Hilbert space H is said to be normal if it commutes with its adjoint i.e. if TT* = T*T

Conclusively every self-adjoint operator is normal. For if T is a self adjoint operator i.e. T*=T then TT* =T*T and so T is normal.

Note A normal operator need not be self adjoint as explained below by an example.

Example: Let H be any Hilbert space and $I : H \to H$ be the identity operator.

Define T = 2iI. Then T is normal operator, but not self-adjoint.

Solution: Since I is an adjoint operator and the adjoint operation is conjugate linear,

$T^* = -2iI^* = -2iI$ so that

$TT^* = T^*T = 4I.$

$\Rightarrow$ T is a normal operator on H.

But $T = T^* \Rightarrow$ T is not self-adjoint.

Note If $T \in \beta(H)$ is normal, then T* is normal.

since if T* is the adjoint of T; then T**= T.

T is normal $\Rightarrow$ TT*=T*T

Hence T*T**=T*T=TT*=T**T* so that T*T**=T**T

$\Rightarrow$ T* is normal if $T \in \beta(H)$.

Theorem 1: The limit T of any convergent sequence (T_k) of normal operators is normal.

Proof: Now $\|T_k^* - T^*\| = \|(T_k - T)^*\| = \|T_k - T\|$

$\Rightarrow T_k^* - T^*$ as $k \to \infty$ since $T_k \to T$ as $k \to \infty$.

Now we prove $TT^* = T^*T$ so that T is normal.

$$\|TT^* - T^*T\| = \|TT^* - T_kT_k^* + T_kT_k^* - T_k^*T_k + T_k^*T - TT^*\| \leq \|TT^* - T_kT_k^*\| + \|T_kT_k^* - T_k^*T_k\|$$

$$+ \|T_k^*T_k - TT^*\|$$

$\Rightarrow$ $\|TT^* - T^*T\| \leq \|TT^* - T_kT_k^*\| + \|T_k^*T_k - TT^*\|$ ...(1)

$[\because T_k$ is normal i.e. $T_kT_k^* = T_k^*T_k]$

since $T_k \to T$ as $T_k^* \to T^*$, R.H.S. of $(1) \to 0$

$\Rightarrow$ $\|TT^* - T^*T\| = 0$

$\Rightarrow$ $TT^* = T^*T$

$\Rightarrow$ T is normal.

This completes the proof of the theorem.

Theorem 2: The set of all normal operators on a Hilbert space H is a closed subspace of $\beta(H)$ which contains the set of all set-adjoint operators and is closed under scalar multiplication.

Proof: Let M be the set of all normal operators on a Hilbert space H. First we shall show that M is closed subset of $\beta(H)$.

Notes

Let T be any limited point of M. Then to show that $T \in M$ i.e. to show that T is a normal operator on H.

Since T is a limited point of M, therefore $\exists$ a sequence(T_n) of distinct point of M such that $T_n \to T$. We have

$$\|T_n^* - T^*\| = \|T_n - T^*\| = \|T_n - T\| \to 0.$$

$\Rightarrow$ $\|T_n^* - T^*\| \to 0 \Rightarrow T_n^* = T^*$.

Now, $\|TT^* - T^*T\| = \|(TT^* - T_nT^*) + (T_nT_n^* - T^*T)\| \le \|TT^* - T_nT_n^*\| + \|T_nT_n^* - T^*T\|$

$$= \|(TT^* - T_nT^*)\| + \|(T_nT_n^* - T_n^*T_n) + (T_n^*T_n - T^*T)\|$$
$$\le \|TT^* - T_nT_n^*\| + \|T_nT_n^* - T_n^*T_n\| + \|T_n^*T_n - T^*T\|$$

$$= \|(TT^* - T_nT_n^*)\| + \|0\| + \|T_n^*T_n - T^*T\|$$

$$[\because T_n \in M \Rightarrow T_n \text{ is a normal operator on H i.e. } T_nT_n^* = T_n^*\ T_n \text{ and } |0| = 0]$$

$$= \|TT^* - T_nT_n^*\| + \|T_n^*T_n - T^*T\| \to 0 \text{ since } T_n \to T \text{ and } T_n^* \to T^*$$

Thus, $\|TT^* - T^*T\| = 0 \Rightarrow TT^* - T^*T = 0$

$\Rightarrow$ $TT^* = T^*T \Rightarrow T$ is normal operator on H.

$\Rightarrow$ $T \in M$ and so M is closed.

Now every self adjoint operator is normal. Therefore the set M contains the set of all self-adjoint operators on H.

Finally, we show that M is closed with respect to scalar multiplication i.e. $T \in M$

$\Rightarrow$ $\alpha T \in M$, α is any scalar.

In other words, we are to show that if T is a normal operator on H and α is any scalar, then αT is normal operator on H. Since T is normal, therefore $TT^* = T^*T$.

We have $(\alpha T)^* = \bar{\alpha} T^*$.

Now $(\alpha T)(\alpha T)^* = (\alpha T)(\bar{\alpha} T^*) = \alpha\bar{\alpha}(TT^*)$.

Also $(\alpha T)^*(\alpha T) = (\bar{\alpha} T)(\alpha T) = \bar{\alpha}\alpha(T^*T) = (\alpha\bar{\alpha})(TT^*)$

$\Rightarrow$ $(\alpha T)(\alpha T)^* = (\alpha T)^*(\alpha T)$

$\Rightarrow$ αT is normal.

This completes the proof of the theorem.

Theorem 3: If N_1, N_2 are normal operators on a Hilbert space H with the property that either commutes with the adjoint of the other, then $N_1 + N_2$ and N_1N_2 are also normal operators.

Proof: Since N_1, N_2 are normal operators, therefore

$$N_1N_1^* = N_1^*N_1 \text{ and } N_2N_2^* = N_2^*\ N_2 \qquad ...(1)$$

Notes

Also by hypothesis, we have $N_1 \, N^*_2 = N^*_2 N_1$ and $N_2 N^*_1 = N^*_1 N_2$ … (2)

we claim that $N_1 + N_2$ is normal.

i.e. $(N_1 + N_2)(N_1 + N_2)^* = (N_1 + N_2)^*(N_1 + N_2)$ ...(3)

since adjoint operation preserves addition, we have

$$(N_1 + N_2)(N_1 + N_2)^* = (N_1 + N_2)\left(N^*_1 + N^*_2\right)$$

$\Rightarrow \quad N_1 N^*_1 + N_1 N_2 + N_2 N^*_1 + N_2 N^*_2$ ...(4)

$\Rightarrow \quad N_1 N^*_1 + N^*_2 N_1 + N^*_1 N_2 + N^*_2 N_2$

$= \quad \left(N^*_1 + N^*_2\right)\left(N_1 + N_2\right)$

$= \quad \left(N_1 + N_2\right)^*\left(N_1 + N_2\right)$ (using (1) and (2))

$\Rightarrow \quad \left(N_1 + N_2\right)\left(N_1 + N_2\right)^* = \left(N_1 + N_2\right)^*\left(N_1 + N_2\right)$

$\Rightarrow \quad N_1 + N_2$ is normal.

Now we show that $N_1 N_2$ is normal i.e.

$$\left(N_1 N_2\right)\left(N_1 N_2\right)^* = \left(N_1 N_2\right)^*\left(N_1 N_2\right).$$

$$\begin{aligned}
\text{L.H.S.} = \left(N_1 N_2\right)\left(N_1 N_2\right)^* &= N_1 N_2 N^*_2 \, N^*_1 \\
&= N_1\left(N_2 N^*_2\right)N^*_1 \\
&= N_1\left(N^*_2 \, N_2\right)N^*_1 \\
&= \left(N_1 N^*_2\right)\left(N_2 N^*_1\right) \\
&= \left(N^*_2 \, N_1\right)\left(N^*_1 \, N_2\right) \\
&= N^*_2\left(N_1 N^*_1\right)N_2 \\
&= \left(N^*_2 \, N^*_1\right)\left(N_1 N_2\right) \\
&= \left(N_1 N_2\right)^*\left(N_1 N_2\right)
\end{aligned}$$

$\Rightarrow \quad \left(N_1 N_2\right)\left(N_1 N_2\right)^* = \left(N_1 N_2\right)^*\left(N_1 N_2\right)$

$\Rightarrow \quad N_1 N_2$ is normal.

This completes the proof of the theorem.

Theorem 4: An operator T on a Hilbert space H is normal $\Leftrightarrow \|T^* x\| = \|Tx\|$ for every $x \in H$.

Proof: We have T is normal $\Leftrightarrow TT^* = T^*T$

$\Leftrightarrow TT^* - T^*T = 0$

$\Leftrightarrow ((TT^* - T^*T)x, x) = 0 \forall x$

$\Leftrightarrow (TT^*x; x) = (T^*Tx, x) \forall x$

$\Leftrightarrow (T^*x, T^*x) = (Tx, T^{**}x) \forall x$

$\Leftrightarrow \|T^*x\|^2 = \|Tx\|^2 \forall x \qquad [\because T^{**} = T]$

$\Leftrightarrow \|T^*x\| = \|Tx\| \forall x.$

This completes the proof of the theorem.

Theorem 5: If N is normal operator on a Hilbert space H, then $\|N\|^2 = \|N^2\|$.

Proof: We know that if T is a normal operator on H then

$$\|Tx\| = \|T^*x\| \forall x \qquad ...(1)$$

Replacing T by N, and x by Nx we get

$$\|NNx\| = \|N^*Nx\| \forall x$$

$$\Rightarrow \qquad \|N^2x\| = \|N^*Nx\| \forall x \qquad ...(2)$$

Now $\|N^2\| = \text{Sup}\{\|N^2x\| : \|x\| \le 1\}$

$= \text{Sup}\{\|N^*Nx\| : \|x\| \le 1\}$ (by (2))

$= \|N^*N\|$

$= \|N\|^2$

This completes the proof of the theorem.

Theorem 6: Any arbitrary operator T on a Hilbert space H can be uniquely expressed as $T = T_1 + iT_2$ where T_1, T_2 are self-adjoint operators on H.

Proof: Let $T_1 = \frac{T + T^*}{2}$ and $T_2 = \frac{1}{2i}(T - T^*)$

Then $T_1 + iT_2 = T$ ...(1)

Now $T^*_1 = \left[\frac{1}{2}(T + T^*)\right]^*$

$= \frac{1}{2}(T + T^*)^*$

$= \frac{1}{2}(T^* + T^{**})$

$= \frac{1}{2}(T^* + T) = \frac{1}{2}(T + T^*) = T_1$

$\Rightarrow \qquad T^*_1 = T_1$

$\Rightarrow \qquad T_1$ is self-adjoint.

Also $T^*_2 = \left[\frac{1}{2i}(T - T^*)\right]^*$

$$= \left(\overline{\frac{1}{2i}}\right)(T - T^*)^*$$

$$= -\frac{1}{2i}(T^* - T^{**})$$

$$= -\frac{1}{2}(T^* - T) = \frac{1}{2i}(T - T^*) = T_2$$

$\Rightarrow \qquad T^*_2 = T_2$

$\Rightarrow \qquad T_2$ is self-adjoint.

Thus T can be expressed in the form (1) where T_1, T_2 are self adjoint operators.

To show that (1) is unique.

Let $T = U_1 + iU_2$, U_1, U_2 are both self-adjoint

We have $T^* = (U_1 + iU_2)^*$

$$= U^*_1 + (iU_2)^*$$

$$= U^*_1 + \bar{i}U^*_2$$

$$= U^*_1 - iU^*_2 = U_1 - iU_2$$

$\therefore T + T^* = (U_1 - iU_2) + (U_1 - iU_2) = 2U,$

$\Rightarrow \qquad U_1 = \frac{1}{2}(T + T^*) = T_1$

and $T - T^* = (U_1 + iU_2) - (U_1 - iU_2) = 2iU_2$

$$U_2 = \frac{1}{2i}(T - T^*) = T_2$$

$\Rightarrow$ expression (1) for T is unique.

This completes the proof of the theorem.

Note The above result is analogous to the result on complex numbers that every complex number z can be uniquely expressed in the form z = x + iy where x, y are real. In the above theorem $T = T_1 + T_2$, T_1 is called real part of T and T_2 is called the imaginary part of T.

Theorem 7: If T is an operator on a Hilbert space H, then T is normal ⇔ its real and imaginary parts commute.

Proof: Let T_1 and T_2 be the real and imaginary parts of T. Then T_1, T_2 are self-adjoint operators and $T = T_1 + i\,T_2$.

We have
$$T^* = (T_1 + iT_2)^* = T_1^* + (iT_2)^*$$
$$= T_i^* + \bar{i}\ T_2^*$$
$$= T_i^* - iT_2^*$$
$$= T_1 - iT_2$$

Now
$$TT^* = (T_1 + iT_2)(T_i - iT_2)$$
$$= T_1^2 + T_2^2 + i(T_2T_1 - T_1T_2) \qquad \ldots (1)$$

and
$$T^*T = (T_i - iT_2)(T_1 - iT_2)$$
$$= T_1^2 + T_2^2 + i(T_1T_2 - T_2T_1) \qquad \ldots (2)$$

Since T is normal i.e. $TT^* = T^*T$.

Then from (1) and (2), we see that
$$T_1^2 + T_2^2 + i(T_2T_1 - T_1T_2) = T_1^2 + T_2^2 + i(T_1T_2 - T_2T_1)$$
$$\Rightarrow \quad T_2T_1 - T_1T_2 = T_1T_2 - T_2T_1$$
$$\Rightarrow \quad 2T_2T_1 = 2T_1T_2$$
$$\Rightarrow \quad T_2T_1 = T_1T_2 \Rightarrow T_1, T_2 \text{ commute.}$$

Conversely, let T_1, T_2 commute

i.e. $\quad T_1T_2 = T_2T_1$, then from (1) and (2)

We see that
$$TT^* = T^*T \Rightarrow T \text{ is normal.}$$

Example: If T is a normal operator on a Hilbert space H and λ is any scalar, then $T - \lambda I$ is also normal.

Solution: T is normal ⇒ $TT^* = T^*T$

Also
$$(T - \lambda I)^* = T^* - (\lambda I)^*$$
$$= T^* - \bar{\lambda} I^*$$
$$= T^* - \bar{\lambda} I.$$

Now
$$(T - \lambda I)(T - \lambda I)^* = (T - \lambda I)(T^* - \bar{\lambda} I)$$
$$= TT^* - \bar{\lambda} I - \lambda T^* + |\lambda|^2 I \qquad \ldots (1)$$

Also
$$(T - \lambda I)^*(T - \lambda I) = (T^* - \bar{\lambda} I)(T - \lambda I)$$
$$= T^*T - \lambda I^* - \bar{\lambda} T + |\lambda|^2 I \qquad \ldots (2)$$

Since TT* = T*T, therefore R.H.S. of (1) and (2) are equal.

Hence their L.H.S. are also equal.

$\therefore \qquad (T - \lambda I)(T - \lambda I)^* = (T - \lambda I)^*(T - \lambda I)$

$\Rightarrow$ T – λI is normal.

15.1.2 Unitary Operator

An operator U on a Hilbert space H is said to be unitary if UU* =U*U =I.

Notes

(i) Every unitary operator is normal.

(ii) $U^* = U^{-1}$ i.e. an operator is unitary iff it is invertible and its inverse is precisely equal to its adjoint.

Theorem 8: If T is an operator on a Hilbert space H, then the following conditions are all equivalent to one another.

(i) T*T = I.

(ii) (Tx,Ty) = (x,y) for all $x, y \in H$.

(iii) $\|Tx\| = \|x\| \forall x \in H$.

Proof: (i) $\Rightarrow$ (ii)

(Tx,Ty) = (x,T*Ty) = (x,Iy) = (x,y) $\forall$x and y.

(ii) $\Rightarrow$ (iii)

We are given that

$(Tx, Ty) = (x, y) \forall x, y \in H.$

Taking y = x, we get

$$(Tx,Tx) = (x,x) \Rightarrow \|Tx\|^2 = \|x\|^2$$

$$\Rightarrow \qquad \|Tx\| = \|x\| \forall x \in H.$$

(iii) $\Rightarrow$ (i)

Given $\quad \|Tx\| = \|x\| \forall x$

$$\Rightarrow \qquad \|Tx\|^2 = \|x\|^2$$

$$\Rightarrow \qquad (Tx, Tx) = (x, x)$$

$$\Rightarrow \qquad (T^*Tx, x) = (x, x)$$

Notes

$\Rightarrow \quad ((T^*T - I)x, x) = O \forall x \in H$

$\Rightarrow \quad T^*T - I = O$

$\Rightarrow \quad T^*T = I$

This completes the proof of the theorem.

15.1.3 Isometric Operator

Definition: An operator T on H is said to be isometric if $\|Tx - Ty\| = \|x - y\| \forall x, y \in H$.

Since T is linear, the condition is equivalent to $\|Tx\| = \|x\|$ for every $x \in H$.

For example: let $\{e_1, e_2, ..., e_n, ...\}$ be an orthonormal basis for a separable Hilbert space H and $T \in \beta(H)$ be defined as $T(x_1e_1 + x_2e_2 + ...) = x_1e_2 + x_2e_3 + ...$ where $x = (x_n)$.

Then $\|Tx\|^2 = \sum_{n=1}^{\infty} |x_n|^2 = \|x\|^2$

$\Rightarrow$ T is an isometric operator.

The operator T defined is called the right shift operator given by $Te_n = e_{n+1}$.

Theorem 9: If T is any arbitrary operator on a Hilbert space H then H is unitary $\Leftrightarrow$ it is an isometric isomorphism of H onto itself.

Proof: Let T is a unitary operator on H. Then T is invertible and therefore T is onto.

Further TT* = I.

Hence $\|Tx\| = \|x\|$ for every $x \in H$. [By Theorem (7)]

$\Rightarrow$ T preserves norms and so T is an isometric isomorphism of H onto itself.

Conversely, let T is an isometric isomorphism of H onto itself. Then T is one-one and onto. Therefore T^{-1} exists. Also T is an isometric isomorphism.

$\Rightarrow \quad \|Tx\| = \|x\| \ \forall x$

$\Rightarrow \quad T^*T = I \qquad$ [By Theorem (7)]

$\Rightarrow \quad (T^*T)T^{-1} = IT^{-1}$

$\Rightarrow \quad T^*(TT^{-1}) = T^{-1}$

$\Rightarrow \quad T^*I = T^{-1}$

$\Rightarrow \quad TT^* = I = T^*T$ and so T is unitary.

This completes the proof of the theorem.

Note If T is an operator on a Hilbert space H such that $\|Tx\| = \|x\| \ \forall x \in H$ and T is definitely an isometric isomorphism of H onto itself. But T need not be onto and so T need not be unitary. The following example will make the point more clear.

Example: Let T be an operator on l_2 defined by $T\{x_1, x_2, ...\} = \{0, x_1, x_2, ...\}$

$\Rightarrow$ $\|Tx\| = \|x\| \ \forall x \in I_2$.

$\Rightarrow$ T is an isometric isomorphism of l_2 into itself.

However T is not onto. If $(y_1, y_2, ...)$ is a sequence in l_2 such that $y_1 \neq 0$, then $\exists$ no sequence in l_2 whose T-image is $(y_1, y_2, ...)$. Therefore T is not onto and so T is not unitary.

15.2 Summary

- An operator T on a Hilbert space H is said to be normal if it commutes with its adjoint i.e. if $TT^* = T^*T$. Conclusively every self adjoint operator is normal.
- The set of all normal operators on a Hilbert space H is a closed subspace of $\beta(H)$ which contains the set of all set-adjoint operators and is closed under scalar multiplication.
- An operator U on a Hilbert space H is said to be unitary if UU* = U*U =I.
- An operator T on H is said to be isometric if $\|Tx - Ty\| = \|x - y\| \forall x, y \in H$, since T is linear, the condition is equivalent to $\|Tx\| = \|x\|$ for every $x \in H$.

15.3 Keywords

Normal Operator: An operator T on a Hilbert space H is said to be normal if it commutes with its adjoint i.e. if TT* = T*T.

Unitary Operator: An operator U on a Hilbert space H is said to be unitary if UU* = U*U = I.

Isometric Operator: An operator T on H is said to be isometric if $\|Tx - Ty\| = \|x - y\| \forall x, y \in H$. Since T is linear, the condition is equivalent to $\|Tx\| = \|x\|$ for every $x \in H$.

15.4 Review Questions

1. If T is an operator on a Hilbert space H, then T is normal $\Leftrightarrow$ its real and imaginary part commute.
2. An operator T on H is normal $\Leftrightarrow \|T^*x\| = \|Tx\|$ for every x.
3. The set of all normal operators on H is a closed subset of $\beta(H)$ which contains the set of all self-adjoint operators and is closed under scalar multiplication.
4. If H is finite-dimensional, show that every isometric isomorphism of H into itself is unitary.
5. Show that the unitary operators on H form a group.

15.5 Further Readings

Notes

Books

Arch W. Naylor, R Sell George, *Linear Operator Theory in Engineering and Sciences*, New York Springer, (1982).

Paul Garret, *Operators in Hilbert Spaces*, 2005.

Online link

www.maths.leeds.ac.uk/nkisilv/

Unit 16: Projections

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define perpendicular projections.
- Define invariance and orthogonal projections.
- Solve problems on projections.

Introduction

We have already defined projections both in Banach spaces and Hilbert spaces and explained how Hilbert spaces have plenty of projection as a consequence of orthogonal decomposition theorem or projection theorem. Now, the context of our present work is the Hilbert space H, and not a general Banach space, and the structure which H enjoys in addition to being a Banach space enables us to single out for special attention those projections whose range and null space are orthogonal. Our first theorem gives a convenient characterisation of these projections.

16.1 Projections

16.1.1 Perpendicular Projections

A projection P on a Hilbert space H is said to be a perpendicular projection on H if the range M and null space N of P are orthogonal.

Theorem 1: If P is a projection on a Hilbert space H with range M and null space N then $M \perp N \Leftrightarrow P$ is self-adjoint and in this case $N = M^{\perp}$.

Proof: Let $M \perp N$ and z be any vector in H. Then since $H = M \oplus N$, we can write z uniquely as

$z = x + y, x \in M, y \in N.$

Thus $Pz = P(x+y)$

$= Px + Py$

$= Px = x \quad (y \in N)$

$\therefore (Pz,z) = (x,z) \quad [\because Pz = P(x+y) = x, P \text{ being projection on } H]$

$= (x, x+y)$

$= (x,x) + (x,y)$

$= \|x\|^2$

and $(Pz^*,z) = (z,Pz)$

$= (x+y,x) = (x,x) + (x,y)$

$= \|x\|^2.$

Hence $(Pz,z) = (Pz^*,z) \forall z \in H$

$\Rightarrow \quad ((P - P^*)z,z) = 0 \forall z \in H$

$\Rightarrow \quad P - P^* = 0$ i.e. $P = P^*$

$\Rightarrow$ P is self adjoint.

Further, $M \perp N \Rightarrow N \subseteq M^{\perp}$

If $N \neq M^{\perp}$, then N is a proper closed linear subspace of the Hilbert space $M^{\perp}$ and therefore $\exists$ a vector $z_0 \neq 0 \in M^{\perp}$ s.t. $z_0 \perp N$.

Now $z_0 \perp M$ and $z_0 \perp N$ and $H = M \oplus N$.

$\Rightarrow z_0 \perp H \Rightarrow z_0 = 0,$ a contradiction.

Hence $N = M^{\perp}$

Conversely, let $P^* = P, x, y$ be any vectors in M and N respectively. Then

$(x,y) = (Px,y)$

$= (x,P^*y) = (x,Py)$

$= (x,0) = 0$

$\Rightarrow M \perp N.$

This completes the proof of the theorem.

Theorem 2: If P is the projection on the closed linear subspace M of H, then

$x \in M \Leftrightarrow Px = x \Leftrightarrow \|Px\| = \|x\|.$

Proof: We have, P is a projection on H with range M then, to show $x \in M \Leftrightarrow Px = x.$

Let Px = x. Then X is in the range of P because Px is in the range of P.

$Px = x \Rightarrow x \in M.$

Conversely, let $x \in M.$ Then to show Px = x.

Let Px = y. Then we must show that y = x.

We have

$$Px = y \Rightarrow P(Px) = Py \Rightarrow P^2x = Py$$

$$\Rightarrow Px = Py \quad [\because P^2 = P]$$

$$\Rightarrow P(x - y) = 0$$

$\Rightarrow x - y$ is a in null space of P.

$$\Rightarrow x - y \in M^{\perp}.$$

$$\Rightarrow x - y = z, z \in M^{\perp}.$$

$$\Rightarrow x = y + z.$$

Now $y = Px \Rightarrow y$ is in the range of P.

i.e. y is in M. Thus we have expressed

$x = y + z, y \in M, z \in M^{\perp}.$

But x is in M. So we can write x = x+0, $x \in M, 0 \in M^{\perp}$

But $H = M \oplus M^{\perp}.$

Therefore we must have y = x, z = 0

Hence $x \in M \Rightarrow Px = x.$

Now we shall show that Px = x $\Leftrightarrow \|Px\| = \|x\|.$

If Px = x then obviously $\|Px\| = \|x\|.$

Conversely, suppose that $\|Px\| = \|x\|.$

We claim that Px = x. We have

$$\|x\|^2 = \|Px + (I - P)x\|^2 \qquad ...(1)$$

Now Px is in M. Also P is the projection on M.

$\Rightarrow$ $I - P$ is the projection on $M^{\perp}.$

$\Rightarrow$ $(I + P)x$ in $M^{\perp}.$

$\Rightarrow$ Px and $(I + P)x$ are orthogonal vectors.

Then by Pythagorean theorem, we get

$$\|Px + (I - P)x\|^2 = \|Px\|^2 + \|(I - P)x\|^2 \qquad ...(2)$$

From (1) and (2), we have

$$\|x\|^2 = \|Px\|^2 + \|(I-P)x\|^2$$

$\Rightarrow$ $\|(I-P)x\|^2 = 0$ $[\because$ by hypothesis $\|Px\| = \|x\|]$

$\Rightarrow$ $(I-P)x = 0$

$\Rightarrow$ $x - Px = 0$

$\Rightarrow$ $Px = 0$

This completes the proof of the theorem.

Theorem 3: If P is a projection on a Hilbert space H, then

(i) P is a positive operator i.e. $P \geq 0$

(ii) $0 \leq P \leq 1$

(iii) $\|Px\| \leq \|x\| \forall x \in H$.

(iv) $\|P\| \leq 1$.

Proof: P, projection on H $\Rightarrow P^2 = P, P^* = P$.

Let M = range of P.

(i) Let $x \in H$. Then

$$(Px, x) = (PPx, x)$$

$$= (Px, P^*x) = (Px, Px) = \|Px\|^2 \geq 0$$

$\Rightarrow$ $(Px, x) \geq 0 \forall x \in H$.

$\Rightarrow$ P is a positive operator i.e. P ≥ 0.

(ii) P is a projection on H$\Rightarrow$ I – P is also a projection on H.

$\Rightarrow$ $I - P \geq 0$. (by (i))

$\Rightarrow$ $P \leq I$

But $P \geq 0$, consequently $0 \leq P \leq 1$.

(iii) Let $x \in H$. If M is the range of P, then $M^\perp$ is the range of (I – P).

Now Px is in M and (I – P)x is in $M^\perp$.

Therefore Px and (I – P)x are orthogonal vector. So by Pythagorean theorem, we have

$$\|Px + (I-P)x\|^2 = \|Px\|^2 + \|(I-P)x\|^2$$

$\Rightarrow$ $\|x\|^2 = \|Px\|^2 + \|(I-P)x\|^2$ $[\because Px + (I-P)x = x]$

$\Rightarrow \quad \|x\|^2 \geq \|Px\|^2$

$\Rightarrow \quad \|Px\| \geq \|x\|.$

(iv) We have $\|P\| = \sup\{\|Px\| : \|x\| \leq 1\}$

But $\|Px\| \leq \|x\| \forall x \in H$ (by (iii))

$\therefore \sup\{\|Px\| : \|x\| \leq 1\} \leq 1$

Hence $\|P\| \leq 1$

This completes the proof of the theorem.

Example: If P and Q are the projections on closed linear subspaces M and N of H. Show that PQ is a projection ⇔ PQ = QP. In this case, show that PQ is the projection on $M \cap N$.

Solution: Since P and Q are projections on H, therefore $P^2 = P$, $P^* = P$, $Q^2 = Q$, $Q^* = Q$. Also it is given that M is range of P and N is the range of Q.

Now suppose PQ is projection on H. Then to prove that PQ = QP.

Since PQ is a projection on H.

$\therefore \quad (PQ)^* = PQ$

$\Rightarrow \quad Q^* P^* = PQ$

$\Rightarrow \quad QP = PQ \quad (\because Q^* = Q, P^* = P)$

Conversely, let PQ = QP. We shall show that PQ is a projection on H.

We have $(PQ)^* = Q^*P^* = QP = PQ$.

Also $(PQ)^2 = (PQ)(PQ) = (PQ)(QP)$

$= PQ^2P = PQP$

$= QPP = QP^2$

$= QP = PQ$

Thus $(PQ)^* = PQ$ and $(PQ)^2 = PQ$.

$\Rightarrow$ PQ is a projection on H.

Finally we are to show that PQ is the projection on $M \cap N$, i.e. we are to show that range of PQ is $M \cap N$.

Let R (PQ) = range of PQ.

Let $x \in M \cap N \Rightarrow x \in M, x \in N$ we have

$(PQ)(x) = P(Qx) = Px \quad [\because$ N is range of Q and $x \in N \Rightarrow Qx = x]$

$= x \quad [\because$ M is range of P and $x \in P]$

$\Rightarrow \quad (PQ)x = x$

$\therefore \quad x \in R(PQ)$

$\Rightarrow \quad x \in M \cap N \Rightarrow x \in R(PQ)$

$\therefore \quad M \cap N \subset R(PQ)$

Now let $x \in R(PQ)$. Then $(PQ)x = x$

Now $\quad (PQ)x = x$

$\Rightarrow \quad P[(PQ)x] = Px$

$\Rightarrow \quad [P(PQ)]x = Px$

$\Rightarrow \quad (P^2Q)x = Px$

$\Rightarrow \quad (PQ)x = Px$

But $\quad (PQ)x = x.$

$\therefore$ We have $Px = x \Rightarrow x \in M$ i.e. the range of P.

Also $\quad PQ = QP$

$\therefore \quad x \in R(PQ) \Rightarrow (PQ)x = x$

$\Rightarrow \quad (QP)x = x \Rightarrow Q[(QP)x] = Qx$

$\Rightarrow \quad (Q^2P)x = Qx \Rightarrow (QP)x = Qx$

But $(QP)x = x$, $\therefore Qx = x \Rightarrow x \in N$.

Thus $\quad x \in R(PQ) \Rightarrow x \in M$ and $x \in N$

$\Rightarrow x \in M \cap N$

$\therefore \quad R(PQ) \subset M \cap N$

Hence $\quad R(PQ) = M \cap N.$

Example: Show that an idempotent operator on a Hilbert space H is a projection on H $\Leftrightarrow$ it is normal.

Solution: P is an idempotent operator on H i.e. $P^2 = P$.

Let P be a projection on H. Then $P^* = P$. We have

$$PP^* = P^* P^* \quad \text{[taking } P^* \text{ in place of P in L.H.S.]}$$

$$= P^* P \quad [\because P^* = P]$$

$\Rightarrow$ P is normal.

Conversely, let $PP^* = P^*P$.

Then to prove that $P^* = P$.

For every vector $y \in H$, we have

$$(Py, Py) = (y, P^* Py) = (y, PP^*y) \quad [\because P^*P = PP^*]$$

$$= (P^*y, P^*y) \quad [\because (P^*)^* = P]$$

From this we conclude that

$$Py = 0 \Leftrightarrow P^*y = 0.$$

Now let x be any vector in H.

Let $y = x - Px$. Then

$$Py = P(x - Px) = Px - P^2x = Px - Px = 0$$

$$0 = P^*y = P^*(x - Px) = P^*x - P^*Px$$

$\Rightarrow$ $P^*x = P^*Px \ \forall \ x \in H$

$\therefore$ $P^* = P^*P$

Now $P = (P^*)^* = (P^*P)^* = P^*P = P^*$

$\therefore$ P is a self adjoint operator.

Also $P^2 = P$.

Hence P is a projection on H.

16.1.2 Invariance

Definition: Let T be an operator on a Hilbert space H and M be a closed subspace of H. Then M is said to be invariant under T if $T(M) \subset M$. If we do not take into account the action of T on vectors outside M, then T can be regarded as an operator on M itself. The operator T on H induces on operator T_M on M such that $T_M(x) = T(x)$ for every $x \in M$. This operator T_M is called the restriction of T on M.

Further, let T be an operator on Hilbert space H. If M is a closed subspace of H and if M and $M^{\perp}$ are both invariant under T, then T is said to be reduced by M. If T is reduced by M, we also say that M reduces T.

Theorem 4: A closed linear subspace M of a Hilbert space H is invariant under the operation $T \Leftrightarrow M^{\perp}$ is invariant under T*.

Proof: Let M is invariant under T, we show $M^{\perp}$ is invariant under T*.

Let y be any arbitrary vector in $M^{\perp}$. Then to show that T^*y is also in $M^{\perp}$ i.e. T^*y is orthogonal to every vector in M.

Let x be any vector in M. Then $Tx \in M$ because M is invariant under T.

Also $y \in M^{\perp} \Rightarrow y$ is orthogonal to every vector in M.

Therefore y is orthogonal to Tx i.e.

$(Tx,y) = 0$

$\Rightarrow$ $(x,Ty) = 0$

$\Rightarrow$ T^*y is orthogonal to every vector x in M.

$\therefore T^*y$ is in $M^{\perp}$ and so $M^{\perp}$ is invariant under T*.

Conversely, let $M^{\perp}$ is invariant under T*. Thus to show that M is invariant under T. Since $M^{\perp}$ is a closed linear subspace of H invariant under T*, therefore by first case $\left(M^{\perp}\right)^{\perp}$ is invariant under T.

But $\left(M^{\perp}\right)^{\perp} = M^{\perp\perp} = M$ and $(T^*)^* = T^{**} = T$.

Notes

Hence M is invariant under T.

This completes the proof of the theorem.

Theorem 5: A closed linear subspace M of a Hilbert space H reduces on operator $\Leftrightarrow$ M is invariant under both T and T*.

Proof: Let M reduces T, then by definition both M and $M^{\perp}$ are invariant under T*. But by theorem 4, if $M^{\perp}$ is invariant under T then $(M^{\perp})^{\perp}$ i.e. M is invariant under T*. Thus M is invariant under T and T*.

Conversely, let M is invariant under both T and T*. Since M is invariant under T*, therefore $M^{\perp}$ is invariant under $(T^*)^* = T$ (by theorem 4). Thus both M and $M^{\perp}$ are invariant under T. Therefore M reduces T.

Theorem 6: If P is the projection on a closed linear subspace M of a Hilbert space H, then M is invariant under an operator T $\Leftrightarrow$ TP = PTP.

Proof: Let M is invariant under T.

Let $x \in H$. Then Px is in the range of T, $Px \in M \Rightarrow TPx \in M$.

Now P is projection and M is the range of P. Therefore $TPx \in M \Rightarrow TPx$ will remain unchanged under P. So, we have

$$PTPx = TPx$$

$$\Rightarrow \quad PTP = TP \qquad \text{(By equality of mappings)}$$

Conversely, let PTP = TP. Let $x \in M$. Since P is a projection with range M and $x \in M$, therefore

$$Px = x$$

$$\Rightarrow \quad TPx = Tx$$

$$\Rightarrow \quad PTPx = Tx \qquad [\because PTP = TP]$$

$$\Rightarrow \quad PTPx = TPx \qquad [\because TPx = Tx]$$

But P is a projection with range M.

$\therefore P(TPx) = TPx \Rightarrow TPx \in M \Rightarrow Tx \in M$

Since TPx = Tx.

Thus $x \in M \Rightarrow Tx \in M$

$\Rightarrow$ M is invariant under T.

Theorem 7: If P is the projection on a closed linear subspace of M of a Hilbert space H, then M reduces an operator $\Leftrightarrow TP = PT$.

Proof: M reduces $T \Leftrightarrow$ M is invariant under T and T*.

$\Leftrightarrow TP = PTP$ and $T^*P = PT^*P$

$\Leftrightarrow TP = PTP$ and $(T^*P)^* = (PT^*P)^*$

$\Leftrightarrow TP = PTP$ and $P^*T^{**} = P^*T^{**}P^*$

$\Leftrightarrow TP = PTP$ and $PT = PTP$ $\qquad [\because P \text{ is projection} \Leftrightarrow P^* = P. \text{ Also} T^{**} = T]$

Thus M reduces T.

$\Leftrightarrow$ TP = PTP and PT = PTP ...(1)

Now suppose M reduces T. Then from (1), TP = PTP and PT = PTP. This gives TP = PT.

Conversely, let TP = PT

$\Rightarrow$ PTP = P^2T (Multiplying both sides on left by P.)

or PTP = PT $\left[\because P^2 = P\right]$

similarly multiplying both sides of TP = PT on the right of P, we get

TP^2 = PTP or TP = PTP. Thus

TP = PT $\Rightarrow$ TP = PTP and PT = PTP.

Therefore from (1), we conclude that M reduces T.

Theorem 8: If M and N are closed linear subspace of a Hilbert space H and P and Q are the projections on M and N respectively, then

(i) $M \perp N \Leftrightarrow PQ = O$. and

(ii) $PQ = O \Leftrightarrow QP = O$.

Proof: Since P and Q are projections on a Hilbert space H, therefore P* = P, Q* = Q.

We first observe that

$$PQ = O \Leftrightarrow (PQ)^* = (O)^* \Leftrightarrow Q^*P^* = O^*$$

$$\Leftrightarrow QP = O.$$

Therefore to prove the theorem it suffices to prove that

$$M \perp N \Leftrightarrow PQ = O.$$

First suppose $M \perp N$. If y is any vector in N, then $M \perp N \Leftrightarrow$ y is orthogonal to every vector in M.

so $y \in M^{\perp}$.Consequently $N \subset M^{\perp}$.

Now, let z be any vector in H. Then Qz is the range of Q i.e. Qz is in N.

Consequently Qz is in $M^{\perp}$ which is null space of P.

Therefore P(Qz) = O.

Thus PQz = O $\forall z \in H$

PQ = O

Conversely, let PQ = O and $x \in M$ and $y \in N$.

since M is the range of P, therefore Px = x. Also N is the range of Q. Therefore

Qy = y

We have (x,y) = (Px, Qy) = (x,P*Qy)

Notes

$= (x, PQy) \quad [\because P^* = P]$

$= (x, Oy) \quad [\because PQ = O]$

$= (x, O) = O$

$\therefore$ M and N are orthogonal i.e. $M \perp N$.

16.1.3 Orthogonal Projections

Definition: Two projections P and Q on a Hilbert space H are said to be orthogonal if PQ = O.

Note: By theorem 8, P and Q are orthogonal iff their ranges M and N are orthogonal.

Theorem 9: If $P_1, P_2, \ldots P_n$ are projections on closed linear subspaces $M_1, M_2, \ldots M_n$ of a Hilbert space H, then $P = P_1 + P_2 + \ldots + P_n$ is a projection $\Leftrightarrow$ the P_i's are pair-wise orthogonal (in the sense that $P_iP_j = 0, i \neq j$).

Also then P is the projection on $M = M_1 + M_2 + \ldots + M_n$.

Proof: Given that $P_1, P_2, \ldots P_n$ are projections on H.

Therefore $P_i^2 = P_i = P_i^*$ for each i = 1, 2, ...,n.

Let $P = P_1 + P_2 + \ldots + P_n$. Then $P^* = (P_1 + P_2 + \ldots + P_n)^* = P_1^* + \ldots + P_n^*$

$= P_1 + P_2 + \ldots + P_n = P$.

Sufficient Condition:

Let $P_iP_j = O, i \neq j$. Then to prove that

P is a projection on H. We have

$P^2 = PP = (P_1 + P_2 + \ldots + P_n)(P_1 + P_2 + \ldots + P_n)$

$= P_1^2 + P_2^2 \ldots + P_n^2 \quad [P_iP_j = 0, i \neq j]$

$= P_1 + P_2 + \ldots + P_n$

$= P$

Thus, $P^* = P = P^*$.

Therefore P is a projection on H.

Necessary Condition:

Let P is a projection on H.

Then $P^2 = P = P^*$.

We are to prove that $P_iP_j = 0$ if $i \neq j$.

We first observe that if T is any projection on H and z is any vector in H, then

$(Tz, z) = (T\,Tz, z) = (Tz, T^*z)$

$= (Tz, Tz)$

$= \|Tz\|^2 \quad \ldots(1)$

Notes

Now let x belongs to the range of some P_i so that $P_i x = x$. Then

$$\|x\|^2 = \|P_i x\|^2$$

$$\leq \sum_{j=1}^{n} \|P_j x\|^2 = \|P_1 x\|^2 + \ldots + \|P_n x\|^2$$

$$= \sum_{j=1}^{n} (P_j x, x) \qquad \text{[Using (i)]}$$

$$= (P_1 x, x) + \ldots + (P_n x, x)$$

$$= ((P_1 + P_2 \ldots + P_n) x)$$

$$= (Px, x)$$

$$= \|Px\|^2 \qquad \text{[by (1)]}$$

$$\leq \|x\|^2 \qquad \ldots(2)$$

Thus we conclude that sign of equality must hold throughout the above computation. Therefore we have

$$\|P_i x\|^2 = \sum_{j=1}^{n} \|P_j x\|^2$$

$\Rightarrow$ $\|P_j x\|^2 = O$ if $j \neq i$

$\Rightarrow$ $\|P_j x\| = O,\ j \neq i$

$\Rightarrow$ $P_j x = O,\ j \neq i$

$\Rightarrow$ x is in the null space of $P_j, i \neq j$

$\Rightarrow$ $x \in M_j^{\perp}$, if $j = i$

$\Rightarrow$ x is orthogonal to the range M_j of every P_j with $j \neq i$.

Thus every vector x in the range P_i(i = 1,...,n) is orthogonal to the range of every P_j with $j \neq i$.

Therefore the range of P_i is orthogonal to the range of every P_j with $j \neq i$. Hence

$P_i P_j = O,\ i \neq j$ [By theorem (8)]

Finally in order to show that P is the projection on $M = M_1 + M_2 + \ldots + M_n$

We are to show that R(P) = M where R(P) is the range of P.

Let $x \in M$. Then $x = x_1 + x_2 + \ldots + x_n$

where $x_i \in M_i,\ 1 \leq i \leq n$. Now from (2), we observe that $\|x\|^2 = \|Px\|^2$ if x is the range of some P_i.

$\therefore x_i \in M$, i.e. the range of P_i.

Notes

$\Rightarrow \quad \|P_i x\|^2 = \|x_i\|^2 \Rightarrow \|P_i x\| = \|x_i\|$

$\Rightarrow \quad Px_i = x_i$

$\Rightarrow \quad x_i \in$ the range of P.

$\Rightarrow \quad x_i \in R(P)$, for each $i = 1, 2, \ldots, n$

$\Rightarrow \quad x_1 + x_2 + \ldots + x_n \in R(P)$.

$\Rightarrow \quad x \in R(P)$.

Then $x \in M \Rightarrow x \in R(P)$

$\therefore M \subset R(P)$ ...(3)

Now suppose that $x \in R(P)$. Then

$Px = x$

$\Rightarrow \quad (P_1 + P_2 + \ldots + P_n)x = x$

$\Rightarrow \quad P_1x + P_2x + \ldots + P_nx = x$

But $P_1(x) \in M_1, P_2(x) \in M_2, \ldots, P_n(x) \in M_n$.

$\therefore x \in M_1 + M_2 + \ldots + M_n$ and so $R(P) \subset M$ ...(4)

Hence from (3) and (4), we get

M = R(P)

This completes the proof of the theorem.

16.2 Summary

- A projection P on a Hilbert space H is said to be a perpendicular projection on H if the range M and null space N of P are orthogonal.
- Let T be an operator on a Hilbert space H and M be a closed subspace of H. Then M is said to be invariant under T if $T(M) \subset M$.
- Let T be an operator on Hilbert space H, if M is closed subspace of H and if M and $M^{\perp}$ are both invariant under T, then T is said to be reduced by M.
- Two projections P and Q on a Hilbert space H are said to be orthogonal if PQ = O.

16.3 Keywords

Invariance: Let T be an operator on a Hilbert space H and M be a closed subspace of H. Then M is said to be invariant under T if $T(M) \subset M$.

Orthogonal Projections: Two projections P and Q on a Hilbert space H are said to be orthogonal if PQ = O.

Perpendicular Projections: A projection P on a Hilbert space H is said to be a perpendicular projection on H if the range M and null space N of P are orthogonal.

Notes

16.4 Review Questions

1. If P and Q are the projections on closed linear subspaces M and N of H, prove that PQ is a projection $\Leftrightarrow PQ = QP$. In this case, show that PQ is the projection on $M \cap N$.

2. If P and Q are the projections on closed linear subspaces M and N of H, prove that the following statements are all equivalent to one another:

   (a) $P \leq Q$;

   (b) $Px\| \leq \|Qx\|$ for every x;

   (c) $M \subseteq N$;

   (d) $PQ = P$;

   (e) $QP = P$.

3. If P and Q are the projections on closed linear subspaces M and N of H, prove that $Q - P$ is a projection $\Leftrightarrow P \leq Q$. In this case, show that Q – P is the projection on $N \cap M^{\perp}$.

16.5 Further Readings

Books

Borbaki, Nicolas (1987), *Topological Vector Spaces, Elements of mathematics*, Berlin: Springer – Verlag

Rudin, Walter (1987), *Real and Complex Analysis*, McGraw-Hill.

Online links

www.math.Isu.edu/~sengupta/7330f02/7330f02proiops.pdf

Planetmath.org/....OrthogonalProjections OntoHilbertSubspaces.html.

Unit 17: Finite Dimensional Spectral Theory

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand finite dimensional spectral theory.
- Describe spectral analysis and spectral resolution of an operator.
- Define compact operators and understand properties of compact operators.
- Solve problems on spectral theory.

Introduction

The generalisation of the matrix eigenvalue theory leads to the spectral theory of operators on a Hilbert space. Since the linear operators on finite dimensional spaces are determined uniquely by matrices, we shall study to some extent in detail the relationship between linear operators in a finite dimension Hilbert spaces and matrices as a preliminary step towards the study of spectral theory of operators on finite dimensional Hilbert spaces.

17.1 Finite Dimensional Spectral Theory

17.1.1 Linear Operators and Matrices on a Finite Dimensional Hilbert Space

Let H be the given Hilbert space of dimension n with ordered basis B = $\{e_1, e_2, \ldots, e_n\}$ where the ordered of the vector is taken into consideration. Let T ∈ β(H) (the set of all bounded linear operators). Since each vector in H is uniquely expressed as linear combination of the basis, we can express Te_j as $Te_j = \sum_{i=1}^{n} \alpha_{ij} e_i$, where the n-scalars $\alpha_{1j}, \alpha_{2j}, \ldots \alpha_{nj}$ are uniquely determined by Te_j.

Then vectors $Te_1, Te_2, \ldots, Te_j$ determine uniquely the n^2 scalars α_{ij}, i, j = 1, 2, … n. These n^2 scalars determine matrix with $(\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})$ as the i^{th} row and $(\alpha_{1j}, \alpha_{2j}, \ldots \alpha_{nj})$ as its j^{th} column. We denote this matrix by {T} and call this matrix as the matrix of the operator T with respect to the ordered basis B.

$$\text{Hence } \mathcal{T} = [\alpha_{ij}] = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & & & \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{bmatrix}$$

We note that

(i) [0] = 0, which is the zero matrix.

(ii) [I] = I = $[\delta_{ij}]$, which is a unit matrix of order n. Here δ_{ij} is the Kronecker delta.

Definition: The set of all n × n matrices denoted by A_n is complex algebra with respect to addition, scalar multiplication and multiplication defined for matrices.

This algebra is called the total matrices algebra of order n.

Theorem 1: Let B be an ordered basis for a Hilbert space of dimension n. Let T ∈ β (H) with (T) = $[\alpha_{ij}]$, then T is singular ⇔ $[\alpha_{ij}]$ is non-singular and we have $[\alpha_{ij}]^{-1} = [T^{-1}]$.

Proof: T is non-singular iff there exists an operator T^{-1} on H such that

$$T^{-1}\,T = T\,T^{-1} = I \qquad \ldots (1)$$

Since there is one-to-one correspondence between T and $[T^{-1}]$,

(1) is true ⇔ $[T^{-1}\,T] = [TT^{-1}] = [I]$

from (2) $[T^{-1}]\,[T] = [T]\,[T^{-1}] = [I] = [\delta_{ij}]$

so that $[T^{-1}]\,[\delta_{ij}] = [\delta_{ij}]\,[T^{-1}] = [\delta_{ij}]$, $[T] = [\alpha_{ij}]$.

⇒ $[\alpha_{ij}]$ is a non-singular and $[\alpha_{ij}]^{-1} = [T^{-1}]$.

This completes the proof of the theorem.

17.1.2 Similar Matrices

Let A, B are square matrix of order n over the field of complex number. Then B is said to be similar to A if there exists a n × n non-singular matrix C over the field of complex numbers such that

$$B = C^{-1}\,AC.$$

This definition can be extended similarly for the case when A, B are operators on a Hilbert space.

Notes

1. The matrices in A_n are similar iff they are the matrices of a single operator on H relative to two different basis H.
2. Similar matrices have the same determinant.

17.1.3 Determinant of an Operator

Let T be an operator on an n-dimensional Hilbert space H. Then the determinant of the operator T is the determinant of the matrix of T, namely [T] with respect to any ordered basis for H.

Following we given properties of a the determinant of an operator on a finite dimensional Hilbert space H.

(i) det (I) = 1, I being identity operator.

Since det (I) = det ([I]) = det ($[\delta_{ij}]$) = 1.

(ii) det ($T_1 T_2$) = (det T_1) (det T_2)

(iii) det (T) ≠ 0 ⇔ [T] is non-singular

⇔ det ([T]) ≠ 0.

Hence det (T) ≠ 0 ⇔ [T] is non-singular.

17.1.4 Spectral Analysis

Definition: Eigenvalues

Let T be bounded linear operator on a Hilbert space H. Then a scalar λ is called an eigenvalue of T if there exists a non-zero vector x in H such that Tx = λx.

Eigenvalues are sometimes referred as characteristic values or proper values or spectral values.

Definition: Eigenvectors

If λ is an eigenvalue of T, then any non-zero vector x ∈ H such Tx = λx, is called a eigenvector (characteristic vector or proper vector or spectral vector) of T.

Properties of Eigenvalues and Eigenvectors

Notes If the Hilbert space has no non-zero vectors then T cannot have any eigenvectors and consequently the whole theory reduces to triviality. So we shall assume that H ≠ 0 throughout this unit.

1. If x is an eigenvector of T corresponding to the eigenvalue λ and α is a non-zero scalar, then αx is also an eigenvector of T corresponding to the same eigenvalue λ.

Contd...

Since x is an eigenvector of T, corresponding to the eigenvalue $\lambda \neq 0$ and $Tx = \lambda x$.

$\alpha \neq 0 \Rightarrow \alpha x \neq 0$

Hence $T(\alpha x) = \alpha T(x) = \alpha(\lambda x)$

Therefore corresponding to an eigenvalue λ there are more than one eigenvectors.

2. If x is an eigenvector of T, then x cannot correspond to more than one eigenvalue of T.

   If possible let λ_1, λ_2 be two eigenvalues of T, $(\lambda_1 \neq \lambda_2)$ for eigenvector x. Then

$$Tx = \lambda_1 x \text{ and } Tx = \lambda_2 x$$

$$\Rightarrow \quad \lambda_1 x = \lambda_2 x$$

$$\Rightarrow \quad (\lambda_1 - \lambda_2)x = 0$$

$$\Rightarrow \quad \lambda_1 - \lambda_2 = 0 \qquad (\because x \neq 0)$$

$$\Rightarrow \quad \lambda_1 = \lambda_2$$

3. Let λ be an eigenvalue of an operator T on H. If M_λ is the set consisting of all eigenvectors of T corresponding to λ together with the vector 0, then M_λ is a non-zero closed linear subspace of H invariant under T.

   By definition $x \in M_\lambda \Leftrightarrow Tx = \lambda x \quad \dots (1)$

   By hypothesis $0 \in M_\lambda$ and 0 vector satisfies (1).

$\therefore \quad M_\lambda = \{x \in H : Tx = \lambda x\} = \{x \in H : (T - \lambda I)\, x = 0\}$

Since T and I are continuous, M_λ is the null space of continuous transformation $T - \lambda I$. Hence M_λ is closed.

Next we show that if $x \in M_\lambda$, then $Tx \in M_\lambda$. If $x \in M_\lambda$ then $Tx = \lambda x$.

Since M_λ is a linear subspace of H, $x \in M_\lambda \Rightarrow \lambda x = Tx \in M_\lambda$.

$\Rightarrow M_\lambda$ is invariant under T.

Definition: Eigenspace

The closed subspace M_λ is called the eigenspace of T corresponding to the eigenvalue λ.

From property (3), we have proved that each eigenspace of T is a non-zero linear subspace of H invariant under T.

Note It is not necessary for an operator T on a Hilbert space H to possess an eigenvalue.

Example: Consider the Hilbert space ℓ_2 and the operator T on ℓ_2 defined by

$$T(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$$

Let λ be a eigenvalue of T. Then $\exists$ a non zero vector

$$y = (y_1, y_2, \dots) \text{ in } \ell_2 \text{ such that } Ty = \lambda y.$$

Now $Ty = \lambda y \quad \Rightarrow T(y_1, y_2, \ldots) = \lambda(y_1, y_2, \ldots)$

$\Rightarrow (0, y_1, y_2, \ldots) = (\lambda y_1, \lambda y_2 \ldots)$

$\Rightarrow \lambda y_1 = 0, \lambda y_2 = y_1, \ldots\ldots$

Now y is a non-zero vector $\Rightarrow y_1 \neq 0$

$\therefore \quad \lambda y_1 = 0 \Rightarrow \lambda = 0.$

Then $\lambda y_2 = y_1 \Rightarrow y_1 = 0$ and this contradicts the fact that y is a non-zero vector. Therefore T cannot have an eigenvalue.

17.1.5 Spectrum of an Operator

The set of all eigenvalues of an operator is called the spectrum of T and is denoted by σ (T).

Theorem 1: If T is an arbitrary operator on a finite dimensional Hilbert space H, then the spectrum of T namely σ (T) is a finite subset of the complex plane and the number of points in σ (T) does not exceed the dimension n of H.

Proof: First we shall show that an operator T on a finite dimensional Hilbert space h is singular if and only if there exists a non-zero vector $x \in H$ such that Tx = 0.

Let $\exists$ a non-zero vector $x \in H$ s.t. Tx = 0. We can write Tx = 0 as Tx = T0. Since $x \neq 0$, the two distinct elements x, $0 \in H$ have the same image under T. Therefore T is not one-to-one. Hence T^{-1} does not exist. Hence it is singular.

Conversely, let T is singular. Let $\exists$ no non-zero vector such that Tx = 0. This means $Tx = 0 \Rightarrow x = 0$. Then T must be one-to-one. Since H is finite dimensional and T is one-to-one, T is onto, so that T is a non-singular, contradicting the hypothesis that T is singular. Hence there must be non-zero vector x s.t. Tx = 0.

Now if T is an operator on a finite dimensional Hilbert space H of dimension n. Then A scalar $\lambda \in \sigma$ (T), if there exists a non-zero vector x such that $(T - \lambda I)x = 0$.

Now $(T - \lambda I)x = 0 \Leftrightarrow (T - \lambda I)$ is a singular.

$(T - \lambda I)$ is singular $\Leftrightarrow \det(T - \lambda I) = 0$. Thus $\lambda \in \sigma(T) \Leftrightarrow \lambda$ satisfies the equation $\det(T - \lambda I) = 0$.

Let B be an ordered basis for H. Thus $\det(T - \lambda I)$

$$= \det([T - \lambda I]_B)$$

But $\det([T - \lambda I]_B) = \det([T]_B - \lambda[I]_B)$

Thus $\det(T - \lambda I) = \det([T]_B - \lambda[\delta_{ij}])$.

So $\det(T - \lambda I) = 0 \Rightarrow \det([T]_B - \lambda[\delta_{ij}]) = 0 \qquad \ldots (1)$

If $[T]_B = [\alpha_{ij}]$ is a matrix of T then (1) gives

$$\begin{bmatrix} \alpha_{11} - \lambda & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{22} & \alpha_{22} - \lambda & \ldots & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \ldots & \ldots & \alpha_{nn} - \lambda \end{bmatrix} \qquad \ldots (2)$$

The expression of determinant of (2) gives a polynomial equation of degree n in λ with complex coefficients in the variable λ. This equation must have at least one root in the field of complex number (by fundamental theorem of algebra). Hence every operator T on H has eigenvalue so that σ (T) $\neq \phi$. Further, this equation in λ has exactly n roots in complex field. If the equation has repeated roots, then the number of distinct roots are less than n. So that T has an eigenvalue and the number of distinct eigenvalue of T is less than or equal to n. Hence the number of elements of σ(T) is less than or equal to n. This completes the proof of the theorem.

Example: For a two dimensional Hilbert space H, let B = $\{e_1, e_2\}$ be a basis and T be an operator on H given by the matrix

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \quad \dots (1)$$

(i) If T is given by $Te_1 = e_2$ and $Te_2 = -e_2$, find the spectrum T.

(ii) If T is an arbitrary operator on H with the same matrix representation, then

$$T^2 (\alpha_{11} + \alpha_{22}) T + (\alpha_{11} \alpha_{22} - \alpha_{12} \alpha_{21}) I = 0$$

Sol:

(i) Using the matrix A of the operator T, we have

$Te_1 = \alpha_{11} e_1 + \alpha_{21} e_2 = e_2$ so that $\alpha_{11} = 0$ and $\alpha_{21} = 1$

$Te_2 = \alpha_{12} e_1 + \alpha_{22} e_2 = -e_1$ so that $\alpha_{12} = -1$ and $\alpha_{22} = 0$

Hence $[T] = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

For this matrix, the eigenvalue are given by the characteristic equation

$$\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = 0$$

$\Rightarrow \quad \lambda^2 + 1 = 0 \Rightarrow \lambda = \pm i$ so $\sigma(T) = \{\pm i\}$.

(ii) Let us consider the eigenvalues of A, which are given by

$$\begin{bmatrix} \alpha_{11} - \lambda & \alpha_{12} \\ \alpha_{21} & \alpha_{22} - \lambda \end{bmatrix} = 0$$

$$\Rightarrow \quad \lambda^2 - (\alpha_{11} + \alpha_{22}) \lambda + (\alpha_{11} \alpha_{22} - \alpha_{12} \alpha_{21}) = 0 \quad \dots (2)$$

Since (2) is true for λ, we can take

$$T = \lambda I \quad \dots (3)$$

From (2) and (3) we get

$$T^2 - (\alpha_{11} + \alpha_{22}) T + (\alpha_{11} \alpha_{22} - \alpha_{12} \alpha_{21}) I = 0 \quad \dots (4)$$

The operator T on H having λ as an eigenvalue satisfies equation (4).

Theorem 2: If T is an operator on a finite dimension Hilbert space, then the following statements are true.

(a) T is singular $\Leftrightarrow 0 \in \sigma$ (T)

(b) If T is non-singular, then $\lambda \in \sigma (T) \Leftrightarrow \lambda^{-1} \in \sigma (T^{-1})$

(c) If A is non-singular, then $\sigma (ATA^{-1}) = \sigma (T)$

(d) If $\lambda \in \sigma$ (T) and if P is polynomial then $P(\lambda) \in \sigma (P(T))$.

Proof: **Notes**

(a) T is singular ⇔ ∃ a non-zero vector $x \in H$ such that $Tx = 0$ or $Tx = 0$. Hence T is singular ⇔ 0 is the eigenvalue of T i.e. $0 \in \sigma(T)$.

(b) Let T be non-singular and $\lambda \in \sigma(T)$.

Hence $\lambda \neq 0$ by (a) so that λ^{-1} exists. Since λ is an eigenvalue of T, ∃ a non-zero vector $x \in H$ s.t. $Tx = \lambda x$.

Premultiplying by T^{-1} we get

$$T^{-1}Tx = T^{-1}(\lambda x)$$

$$\Rightarrow T^{-1}(x) = \frac{1}{\lambda}x \text{ for } x \neq 0$$

Hence $\lambda^{-1} \in (\sigma(T^{-1}))$

Conversely, if λ^{-1} is an eigenvalue of T^{-1} then $(\lambda^{-1})^{-1} = \lambda$ is an eigenvalue of $(T^{-1})^{-1} = T$.

Hence $\lambda \in \sigma(T)$.

(c) Let $S = ATA^{-1}$. Then we find $S - \lambda I$.

$$\text{Now } S - \lambda I = ATA^{-1} - A(\lambda I)A^{-1}$$

$$= A(T - \lambda I)A^{-1}$$

$$\therefore \quad \det(S - \lambda I) = \det(A(T - \lambda I)A^{-1})$$

$$= \det(T - \lambda I)$$

⇒ λ is an eigenvalue of T ⇔ $\det(T - \lambda I) = 0$.

Hence $\det(T - \lambda I) = 0 \Leftrightarrow \det(S - \lambda I) = 0$

⇒ S and T have the some eigenvalues so that

$\sigma(ATA^{-1}) = \sigma(T)$.

(d) If $\lambda = \sigma(T)$, λ is an eigenvalue of T. Then ∃ a non-zero vector x such that $Tx = \lambda x$.

Hence $T(Tx) = T(\lambda x) = \lambda Tx = \lambda^2 x$.

Hence if λ is an eigenvalue of T, then λ^2 is an eigenvalue of T^2. Repeating this we get that if λ is an eigenvalue of T, then λ^n is an eigenvalue of T^n for any positive integer n.

Let $P(t) = a_0 + a_1 t + \ldots + a_m t^m$, $a_0, a_1, \ldots\ldots, a_m$ are scalars.

$$\text{Then } [P(T)]x = (a_0 I + a_2 T + \ldots\ldots + a_m T^m)x$$

$$= a_0 x + a_1(\lambda x) + \ldots\ldots + a_m(\lambda^m x)$$

$$= [a_0 + a_1(\lambda) + \ldots\ldots + a_m \lambda^m]x$$

Hence $P(\lambda) = a_0 + a_1\lambda + \ldots\ldots + a_m\lambda^m$ is an eigenvalue of $P(T)$.

This if $\lambda \in \sigma(T)$, then $P(\lambda) \in \sigma(P(T))$.

This completes the proof of the theorem.

17.1.6 Spectral Theorem

Statement: Let T be an operator on a finite dimensional Hilbert space H with $\lambda_1, \lambda_2, \ldots, \lambda_m$ as the eigenvalues of T and with $M_1, M_2, \ldots, M_m$ be then corresponding eigenspaces. If $P_1, P_2, \ldots, P_m$ are the projections on the spaces, then the following statements are equivalent.

(a) The M_i's are pairwise orthogonal and span H:

(b) The P_i's are pairwise orthogonal and $P_1 + P_2 + \dots + P_m = I$ and $T = \lambda_1 P_1 + \lambda_2 P_2 + \dots + \lambda_m P_m$.

(c) T is normal operator on H.

Proof: We shall show that

(i) ⇒ (ii) ⇒ (iii) ⇒ (i)

(i) ⇒ (ii)

Assume that M_i's are pairwise orthogonal and span H. Hence every $x \in H$ can be represented uniquely as

$$x = x_1 + x_2 + \dots + x_m \quad \dots (1)$$

where $x_i \in M_i$ for $i = 1, 2, \dots, m$

by hypothesis M_i's are pairwise orthogonal. Since P_i's are projections in M_i's ⇒ P_i's are pairwise orthogonal, i.e. $i \neq j \Rightarrow P_i P_j = 0$.

If x is any vector in H, then from (1) for each i,

$$P_i(x) = P_i (x_1 + x_2 + \dots + x_m)$$
$$= P_i x_1 + P_i x_2 + \dots + P_i x_m \quad \dots (2)$$

Since P_i is the range of M_i, $P_i\, x_i = x_i$.

For $i \neq j$ $M_j \perp M_i$ since $x_j \in M_j$ for each j we have

$x_j \perp M_i$ for $j \neq i$.

Hence $x_j \in M_i^{\perp}$ (null space of P_i)

⇒ $x_j \in M_i^{\perp} \Rightarrow P_i x_j = x_i$

∴ from (2) we get

$$P_i x = x_i \quad \dots (3)$$

Since I is the identity mapping on H, we get

$$Ix = x_1 + x_2 + \dots + x_m \quad \dots \text{(by (1))}$$
$$= P_1 x + P_2 x + \dots + P_m x \quad \dots \text{(by (3))}$$
$$= (P_1 + P_2 + \dots + P_m)\, x \ \ \forall x \in H.$$

This show that $I = P_1 + P_2 + \dots\dots + P_m$.

For every $x \in H$, we have from (1)

$$T(x) = T(x_1 + x_2 + \dots + x_m)$$
$$= Tx_1 + Tx_2 + \dots + Tx_m$$

Since $x_i \in M_i \Rightarrow Tx_i = \lambda x_i$

$$\therefore \quad T_x = \lambda_1 x_1 + \lambda_2 x_2 + \dots\dots + \lambda_m x_m \quad \dots (4)$$
$$= \lambda_1 P_1 x + \lambda_2 P_2 x + \dots\dots + \lambda_m P_m x \quad \dots (5)$$

$$\Rightarrow \quad T = \lambda_1 P_1 + \lambda_2 P_2 + \dots\dots + \lambda_m P_m$$

(ii) ⇒ (iii)

Notes

Let $T = \lambda_1P_1 + \lambda_2P_2 + \ldots\ldots + \lambda_mP_m$, where P_i's are pairwise orthogonal projections and to show that T is normal.

Since P_i's are projection and $\Rightarrow P^*_i = P_i$ and $P_i^2 = P_i$ … (6)

Further we have $P_iP_j = 0$ for $i \neq j$

Since adjoint operation is conjugate linear, we get

$$T^* = (\lambda_1P_1 + \lambda_2P_2 + \ldots\ldots + \lambda_mP_m)^*$$

$$= \bar{\lambda}_1P^*_1 + \bar{\lambda}_2P^*_2 + \ldots\ldots + \bar{\lambda}_mP^*_m$$

$$= \bar{\lambda}_1P_1 + \bar{\lambda}_2P_2 + \ldots\ldots + \bar{\lambda}_mP_m .$$

Now

$$TT^* = (\lambda_1P_1 + \lambda_2P_2 + \ldots\ldots + \lambda_mP_m)(\bar{\lambda}_1P_1 + \bar{\lambda}_2P_2 + \ldots\ldots + \bar{\lambda}_mP_m)$$

$$= |\lambda_1|^2 P_1^2 + |\lambda_2|^2 P_2^2 + \ldots\ldots + |\lambda_1|^2 P_m^2 \qquad (\because P_iP_j = 0, i \neq j)$$

$$= |\lambda_1|^2P_1 + |\lambda_2|^2P_2 + \ldots\ldots + |\lambda_m|^2P_m \qquad \ldots \text{(by (6))}$$

Similarly T*T can be found s.t.

$$T^*T = |\lambda_1|^2P_1 + |\lambda_2|^2P_2 + \ldots\ldots + |\lambda m|^2 P_m$$

Hence $T^*T = TT^* \Rightarrow$ T is normal.

(iii) $\Rightarrow$ (i)

Let T is normal operator on H and prove that M_i's are pairwise orthogonal and M_i's span H.

We know that if

T is normal on H $\Rightarrow$ its eigenspaces M_i's are pairwise orthogonal.

So it suffices to show that M_i's span H.

Let

$$M = M_1 + M_2 + \ldots\ldots + M_m$$

and

$$P = P_1 + P_2 + \ldots\ldots + P_m$$

Since T is normal on H, each eigenspace M_i reduces T. Also M_i reduces $T \Rightarrow P_iT = TP_i$ for each P_i.

$\therefore$

$$TP = T(P_1 + P_2 + \ldots\ldots + P_m)$$

$$= TP_1 + TP_2 + \ldots\ldots + TP_m$$

$$= P_1T + P_2T + \ldots\ldots + P_mT$$

$$= (P_1 + P_2 + \ldots\ldots + P_m)\,T$$

$$= PT$$

$\therefore$ Since P is projection on M and TP = PT, M reduces T and so $M^\perp$ is invariant under T. Let U be the restriction of T to $M^\perp$. Then U is an operator on a finite dimensional Hilbert space $M^\perp$ and $Ux = Tx \ \forall \ x \in M^\perp$. If x is an eigenvector for U corresponding to eigenvalue λ then $x \in M^\perp$ and $Ux = \lambda x$.

$\therefore$ $Tx = \lambda x$ and so x is also eigenvector for T.

Hence each eigenvector of U is also an eigenvector for T. But T has no eigenvector in $M^\perp$. Hence $M \cap M^\perp = \{0\}$. So U is an operator on a finite dimensional Hilbert space $M^\perp$ and U has no eigenvector and so it has no eigenvalue.

$\therefore \quad M^{\perp} = \{0\}$.

For if $M^{\perp} \neq \{0\}$, then every operator on a non-zero finite dimensional Hilbert space must have an eigenvalue.

Now $M^{\perp} = \{0\} \Rightarrow M = H$.

Thus $M = M_1 + M_2 + \ldots\ldots + M_m = H$ and so M_i's span H.

This complete the proof of the theorem.

17.1.7 Spectral Resolution of an Operator

Let T be an operator on a Hilbert space H. If there exist distinct complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_m$ and non-zero pairwise orthogonal projections $p_1, p_2, \ldots, p_m$ such that

$$T = \lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_m p_m$$

and $\quad p = p_1 + p_2 + \ldots + p_m$, then the expression

$T = \lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_m p_m$ for T is called the spectral resolution for T.

> *Note* We note that the spectral theorem coincides with the spectral resolution for a normal operator on a finite dimensional Hilbert space.

Theorem: The spectral resolution of the normal operator on a finite dimensional non-zero Hilbert space is unique.

Proof: Let $T = \lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_m p_m$

be a spectral resolution of a normal operator on a non-zero finite dimensional Hilbert space H. Then $\lambda_1, \lambda_2, \ldots, \lambda_m$ are distinct complex numbers and p_i's are non-zero pairwise orthogonal projections such that $p_1 + p_2 + \ldots + p_m = 1$. We establish that $\lambda_1 + \lambda_2, \ldots, \lambda_m$ are precisely the distinct eigenvalues of T.

To this end we show first that for each i, λ_i is an eigenvalue of T. Since $p_i \neq 0$ is a projection, $\exists$ a non-zero x in the range of p_1 such that $p_i x = x$

Let us consider

$$Tx = (\lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_m p_m)x$$

$$= (\lambda_1 p_1 p_i + \lambda_2 p_2 p_i + \ldots + \lambda_i p_i^2 + \ldots + \lambda_m p_m p_i)x$$

So p_i's are pairwise orthogonal $p_i p_j = 0$ for $i \neq j$ and $p_i^2 = p_i$, we have $Tx = \lambda_i p_i x = \lambda_i x$ by $p_i x = x$.

$\Rightarrow \quad \lambda_i$ is an eigenvalue of T.

Next we show that each eigenvalue of T is an element of the set $(\lambda_1, \lambda_2, \ldots, \lambda_m)$. Since T is an operator on a finite dimensional Hilbert space, T must have an eigenvalue.

If λ is an eigenvalue of T then $Tx = \lambda x = \lambda I x$.

$$\Rightarrow \quad (\lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_m p_m)\, x = \lambda\, (p_1 + p_2 + \ldots + p_m)\, x$$

$$\Rightarrow \quad \{(\lambda_1 - \lambda)\, p_1 + (\lambda_2 - \lambda)\, p_2 + \ldots + (\lambda_m - \lambda)\, p_m\}\, x = 0 \qquad \ldots (2)$$

Since $p_i^2 = p_i$ and $p_i\, p_j = 0$ for $i \neq j$ operating with p_i throughout (2), we get

$$(\lambda_i - \lambda)\, p_i x = 0 \text{ for } i = 1, 2, \ldots, m.$$

Notes

If $\lambda_i \neq \lambda$ for each i, $p_i x = 0$ for each i.

Hence $p_i x + p_2 x + \dots + p_m x = 0$

$\Rightarrow$ $(p_1 + p_2 + \dots + p_m)\, x = 0$

$\Rightarrow$ $Ix = 0$

$\Rightarrow$ $x = 0$, a contradiction to the fact that $x \neq 0$. Hence λ must be equal to λ_i for some i. This in the spectral resolution (1) of T, the scalar λ_i are the precisely the eigenvalue of T.

If the spectral resolution is not unique.

Let $T = \mu_1 Q_1 + \mu_2 Q_2 + \dots + \mu_n Q_n$ … (3)

be another revolution of T. Then $\mu_1, \mu_2, \dots \mu_n$ is the same set of eigenvalues of T written in different order. Hence writing the eigenvalues in the same order as in (1) and renaming the projections, we can write (3) as

$$T = \lambda_1 Q_1 + \lambda_2 Q_2 + \dots + \lambda_m Q_m \quad \dots (4)$$

To prove uniqueness, we shall show that p_i in (1) and Q_i in (4) are some.

Using the fact $p_i^2 = p_i$, $p_i p_j = 0 \;\; \forall\, i \neq j$, we can have

$$T^0 = I = p_1 + p_2 + \dots + p_m$$

$$T^1 = \lambda_1 p_1 + \lambda_2 p_2 + \dots + \lambda_m p_m$$

$$T^2 = \lambda_1^2 p_1 + \dots + \lambda_m^2 p_m$$

and $T^n = \lambda_1^n p_1 + \dots + \lambda_m^n p_m$ for any positive integer n. … (5)

Now if g (t) is a polynomial with complex coefficient in the complex variable t, we can write g (T) as

$$g(T) = g(\lambda_1)\, p_1 + g(\lambda_2)\, p_2 + \dots + g(\lambda_m)\, p_m$$

$$= \sum_{j=1}^{m} g(\lambda_j)\, p_j \quad \dots \text{(by 5)}$$

Let π_i be a polynomial such that $\pi_i(\lambda_i) = 1$ and $\pi_i(\lambda_j) = 0$

if $i \neq j$

Taking π_i in place of g, we get

$$\pi_i(T) = \sum_{j=1}^{m} \pi_i(\lambda_j)\, p_j = \sum_{j=1}^{m} \delta_{ij}\, p_j = p_i$$

Hence for each i, let $p_i = \pi_i(T)$ which is a polynomial in T. The proof is complete if we show the existence of π_i over the field of complex number.

Now $$\pi_i(t) = \frac{(t-\lambda_1)\dots(t-\lambda_{i-1})(t-\lambda_{i+1})\dots(t-\lambda_m)}{(\lambda_i-\lambda_1)\dots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\dots(\lambda_i-\lambda_m)}$$

satisfies our requirements i.e. $\pi_i(\lambda_i) = 1$ and $\pi_i(\lambda_j) = 0$ if $i \neq j$

Repeating the above discussion for Q_i's we get in a similar manner $Q_i = \pi_i(T)$ for each i.

$\therefore$ $p_i = Q_i$ for each i.

This completes the proof of the theorem.

17.1.8 Compact Operators

Definition: A subset A in a normed linear space N is said to be relatively compact if its closure $\overline{A}$ is compact.

A linear transformation T of a normed linear space N into a normed linear space N′ is said to be a compact operator if it maps a bounded set of N into a relatively compact set in N′, i.e.

T : N → N′ is compact of every bounded set B ⊂ N, $\overline{T(B)}$ is compact in N′.

17.1.9 Properties of Compact Operators

1. Let T : N → N′ be a compact operator. Then T is bounded (continuous) linear operator. For, let B be a bounded set in N. Since T is compact, $\overline{T(B)}$ is compact in N′. So $\overline{T(B)}$ is complete and totally bounded in N′. Since a totally bounded set is always bounded, $\overline{T(B)}$ is bounded and consequently T(B) is bounded, since a subset of a bounded set is bounded.

∴ T is a bounded linear transformation and it is continuous.

2. Let T be a linear transformation on a finite dimensional space N. Then T is compact operator. For, N is finite dimensional and T is linear T(N) is finite dimensional. Since any linear transformation on a finite dimensional space is bounded. T(B) is bounded subset of T(N) for every bounded set B ⊂ N. Now if T(B) is bounded so is $\overline{T(B)}$ and is closed. T(N) is finite dimensional, any closed and bounded subset of T(N) is compact, so that $\overline{T(B)}$ is compact, being closed and bounded subset of T(N).

3. The operator O on any normed linear space N is compact.

4. If the dimension of N is infinite, then identity operator I : N → N is not compact operator.

   For consider a closed unit sphere.

   $S = \{x \in N : \|x\| \leq 1\}$ then S is bounded.

   Since N is a infinite dimensional.

   $I(S) = S = \overline{S}$ is not necessarily compact.

   Hence I : N → N is not compact operator. But I is a bounded (continuous) operator.

Theorem: A set A in a normed linear space N is relatively compact ⇔ every sequence of points in A contains a convergent sub sequence.

Proof: Let A is relatively compact.

Since $A \subset \overline{A}$, every sequence in A is also sequence in $\overline{A}$. Since $\overline{A}$ is compact, such a sequence in $\overline{A}$ contains a convergent subsequence. Hence every sequence in A has a convergent subsequence.

Conversely, let every sequence in A has a convergent subsequence.

Let (y_n) be a sequence of points in $\overline{A}$. Since A is dense in $\overline{A}$, ∃ a sequence (x_n) of points of A s.t.

$$\|x_n - y_n\| \leq \frac{1}{n} \qquad \dots (1)$$

and $$x_{n_k} \to x \in \overline{A} \qquad \dots (2)$$

we can find a (y_{n_k}) of (y_n) s.t.

$$\left\| y_{n_k} - x \right\| = \left\| y_{n_k} - x_{n_k} + x_{n_k} - x \right\|$$

$$\leq \left\| y_{n_k} - x_{n_k} \right\| + \left\| x_{n_k} - x \right\|$$

$$\to 0 \text{ as } n \to \infty.$$

$$\Rightarrow \overline{A} \text{ is compact.}$$

This completes the proof of the theorem.

17.2 Summary

- If T is an arbitrary operator on a finite dimensional Hilbert space H, then the spectrum of T namely σ (T) is a finite subset of the complex plane and the number of points in σ (T) does not exceed the dimension n of H.
- Let T be bounded linear operator on a Hilbert space H. Then a scalar λ is called an eigenvalue of T if there exists a non-zero vector x in H such that Tx = λx.
- The closed subspace M_λ is called the eigenspace of T corresponding to the eigenvalue λ.
- The set of all eigenvalues of an operator is called the spectrum of T. It is denoted by σ (T).
- The spectral resolution of the normal operator on a finite dimensional non-zero Hilbert space is unique.
- A subset A in a normed linear space N is said to be relatively compact if its closure $\overline{A}$ is compact.

17.3 Keywords

Eigenspace: The closed subspace M_λ is called the eigenspace of T corresponding to the eigenvalue λ.

Eigenvalues: Let T be bounded linear operator on a Hilbert space H. Then a scalar λ is called an eigenvalue of T if there exists a non-zero vector x in H such that Tx = λx.

Eigenvalues are sometimes referred as characteristic values or proper values or spectral values.

Eigenvectors: If λ is an eigenvalue of T, then any non-zero vector x ∈ H such Tx = λx, is called a eigenvector.

Similar Matrices: Let A, B are square matrix of order n over the field of complex number. Then B is said to be similar to A if there exists a n × n non-singular matrix C over the field of complex numbers such that

$$B = C^{-1} AC.$$

Spectrum of an Operator: The set of all eigenvalues of an operator is called the spectrum of T and is denoted by σ (T).

Total Matrices Algebra: The set of all n × n matrices denoted by A_n is complex algebra with respect to addition, scalar multiplication and multiplication defined for matrices.

This algebra is called the total matrices algebra of order n.

17.4 Review Questions

1. If $T \in \beta$ (H) is a self-adjoint operator, then σ (T) = {m, M} where m, M are spectral values.
2. If T is self-adjoint operator then σ (T) is the subset of the real line $[- \|T\|, \|T\|]$.
3. Let $\| R_\lambda (T) \| = (T - \lambda I)^{-1}$ for a $T \in B(X, X)$. Prove that $\| R_\lambda T \| \to 0$ as $\lambda \to \infty$.
4. Prove that the projection of a Hilbert space H onto a finite dimensional subspace of H is compact.

17.5 Further Readings

Books

Walter Rudin, *Real and complex analysis*, Third, McGraw-Hill Book Co., New York, 1987.

Erwin Kreyszig, *Introductory functional analysis with applications*, John Wiley & Sons Inc., New York, 1989.

Online links

www.math.washington.edu

chicago.academia.edu

www.math.ethz.ch/~ Kowalski/spectral-theory.pdf