Measure Theorey and Functional Analysis

DMTH505

Edited by:
Dr.Sachin Kaushal

LOVELY PROFESSIONAL UNIVERSITY

LOVELY PROFESSIONAL UNIVERSITY

MEASURE THEORY AND FUNCTIONAL ANALYSIS

Edited by
Dr. Sachin Kaushal

Printed by
EXCEL BOOKS PRIVATE LIMITED
A-45, Naraina, Phase-I,
New Delhi-110028
for
Lovely Professional University
Phagwara

SYLLABUS

Measure Theory and Functional Analysis

Objectives: This course is designed for the of analysis of various types of spaces like Banach Spaces, Hilbert Space, etc. and also

Sr. No	Description
1	Differentiation and Integration: Differentiation of monotone functions, Functions of bounded variation,
2	Differentiation of an integral, Absolute continuity
3	Spaces, Holder, Minkowski inequalities, Convergence and Completeness
4	Bounded linear functional on the *Lp* spaces, Measure spaces, Measurable Functions, Integration
5	General Convergence Theorems, Signed Measures, Radon-Nikodym theorem.
6	Banach spaces: Definition and some examples, Continuous linear transformations, The Hahn-Banach theorem
7	The natural imbedding of N in N**, The open mapping theorem, The closed graph theorem,
8	The conjugate of an operator, The uniform boundedness theorem, The uniform boundedness theorem, Hilbert spaces : The definition and some simple properties
9	Orthogonal complements, Orthonormal Sets, The conjugate space H*, The Adjoint of an Operator, Self Adjoint Operators
10	Normal and Unitary Operators, Projections, Finite dimensional spectral theory : the spectrum of an operator on a finite dimensional Hilbert space, the Spectral theorem

CONTENT

Notes

Unit 1: Differentiation and Integration: Differentiation of Monotone Functions

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand differentiation and integration
- Describe Lipschitz condition and Lebesgue point of a function
- State Vitali's Lemma and understand its proof.
- Explain four Dini's derivatives and its properties
- Describe Lebesgue differentiation theorem.

Introduction

Differentiation and integration are closely connected. The fundamental theorem of the integral calculus is that differentiation and integration are inverse processes. The general principle may be interpreted in two different ways:

1. If f is a Riemann integrable function over [a, b], then its indefinite integral i.e.

   F : [a, b] $\to$ R defined by F (x) = $\int_a^x f(t)\,dt$ is continuous on [a, b]. Furthermore if f is continuous at a point $x_o \in$ [a, b], then F is differentiable thereat and F′(xo) = f (x_o).

2. If f is Riemann integrable over [a, b] and if there is a differentiable function F on [a, b] such that $F' = f(x)$ for x $\in$ [a, b], then

$$\int_a^x f(t)\,dt = F(x) - F(a) \quad [a \le x \le b].$$

Notes

1.1 Differentiation and Integration

1.1.1 Lipschitz Condition

Definition: A function f defined on [a, b] is said to satisfy Lipschitz condition (or Lipschitzian function), if ∃ a constant M > 0 s.t.

$$|f(x) - f(y)| \le M\,|x - y|,\ \forall\, x, y \in [a, b].$$

1.1.2 Lebesgue Point of a Function

Definition: A point x is said to be a Lebesgue point of the function f (t), if

$$\lim_{h\to 0} \frac{1}{h}\int_x^{x+n} |f(t) - f(x)|\,dt = 0.$$

1.1.3 Covering in the Sense of Vitali

Definition: A set E is said to be covered in the sense of Vitali by a family of intervals (may be open, closed or half open), M in which none is a singleton set, if every point of the set E is contained in some small interval of M i.e., for each x ∈ E, ∃ and ε > 0, an interval I ∈ M s.t. x ∈ I and $\ell(I) < \varepsilon$.

The family M is called the *Vitali Cover* of set E.

Example: If E = {q : q is a rational number in the interval [a, b]}, then the family $\left[I_{qi}\right]$ where $\left[I_{qi}\right] = \left[q - \frac{1}{i}, q + \frac{1}{i}\right]$, i ∈ N is a vitali cover of [a, b].

Vitali's Lemma

Let E be a set of finite outer measure and M be a family of intervals which cover E in the sense of Vitali; then for a given ε > 0, it is possible to find a finite family of disjoint intervals $\{I_k, k = 1, 2, \ldots n\}$ of M, such that

$$m^*\left[E - \bigcup_{k=1}^{n} I_k\right] < \varepsilon.$$

Proof: Without any loss of generality, we assume that every interval of family M is a closed interval, because if not we replace each interval by its closure and observe that the set of end points of $I_1, I_2, \ldots\ldots I_n$ has measure zero.

[Due to this property some authors take family M of closed intervals in the definition of Vitali's covering].

Suppose 0 is an open set containing E s.t. m* (0) < m* (E) + 1 < ∞ we assume that each interval in M is contained in 0, if this can be achieved by discarding the intervals of M extending beyond 0 and still the family M will cover the set E in the sense of Vitali.

Now we shall use the induction method to determine the sequence $\langle I_k : k = 1, 2, \ldots n\rangle$ of disjoint intervals of M as follows:

Notes

Let I_1 be any interval in M and let ℓ_1 be the supremum (least upper bound of the lengths of the intervals in M disjoint from I_1 (i.e., which do not have any point common with I_1).

Obviously $\ell_1 < \infty$ as $\ell_1 \le m(0) < \infty$.

Now we choose an interval I_2 from M, disjoint from I_1, such that $\ell(I_2) > \frac{1}{2}\ell_1$. Let ℓ_2 be the supremums of lengths of all those intervals of M which do not have any point common with I_2 or I_2 obviously $\ell_2 < \infty$.

In general, suppose we have already chosen r intervals $I_1, I_2, \ldots I_r$ (mutually disjoint). Let ℓ_r be the supremums of the length of those intervals of M which do not have any point in common with $\bigcup_{i=1} I_i$ (i.e., which do not meet any of the intervals $I_1, I_2, \ldots I_r$. Then $\ell_r \le m(0) < \infty$.

Now if E is contained in $\bigcup_{i=1}^{r} I_i$, then Lemma established. Suppose $\bigcup_{i=1}^{r} I_i \subset E$. Then we can find interval I_{r+1} s.t. $\ell(I_{r+1}) > \frac{1}{2}\ell_r$ which is disjoint from $I_1, I_2, \ldots I_r$.

Thus at some finite iteration either the Lemma will be established or we shall get an infinite sequence $\langle I_r \rangle$ of disjoint intervals of M s.t. $\ell(I_{r+1}) > \frac{1}{2}\ell_r$ and $\ell_r < \infty$, n = 1, 2, 3 ….

Note that $\langle \ell_r \rangle$ is a monotonically decreasing sequence of non-negative real numbers.

Obviously, we have that $\bigcup_{i=1}^{\infty} I_r \subset 0 \Rightarrow \sum_{r=1}^{\infty} \ell(\ell_r) \le m(0) < \infty$ hence for any arbitrary $\varepsilon > 0$, we can find an integer N s.t.

$$\sum_{r=N+1}^{\infty} \ell(I_r) < \frac{1}{5}\varepsilon .$$

Let a set $F = \bigcup_{r=1}^{N} I_r$.

The lemma will be established if we show that $m^*(F) < \varepsilon$. For, let $x \in F$, then $x \notin \bigcup_{r=1}^{N} I_r \Rightarrow x$ is an element of E not belonging to the closed set $\bigcup_{r=1}^{N} I_r \Rightarrow \exists$ an interval I in M s.t. $x \in I$ and $\ell(I)$ is so small that I does not meet the $\bigcup_{r=1}^{N} I_r$, i.e.

$$I \cap I_r = \phi, \ \forall\, r = 1, 2, \ldots N.$$

Therefore we shall have $\ell(I) \le \ell_N < 2\ell(I_{n+1})$ as by the method of construction we take $\ell(I_{n+1}) \le \frac{1}{2}\ell_N$.

Notes

It also $I \cap I_{N+1} = \phi$, we should have $\ell(I) \le \ell_{N+1}$. Further if the interval I does not meet any of the intervals in the sequence $<I_r>$, we must have

$$\ell(I) \le \ell_r, \forall r$$

which is not true as $\ell_r < 2\ell(I_{r+1}) \to 0$ as $r \to \infty$.

$\Rightarrow$ I must meet at least one of the intervals of the sequence $<I_r>$. Let p be the least integer s.t. I meets I_p. Then $p > N$ and $\ell(I) \le \ell_{p-1} < 2\ell(I_p)$. Further let $x \in I$ as well $x \in I_p$, then the distance of x from the mid point of I_p is at most

$$\ell(I) + \frac{1}{2}\ell(\ell_p) < 2\ell(I_p) + \frac{1}{2}\ell(I_p) = \frac{5}{2}\ell(I_p)$$

Thus if I_p is an interval having the same mid point as I_p but length 5 times the length of I_p, i.e. $\ell(J_p) = 5\ell(I_p)$. Then $x \in J_p$ also.

Thus for every $x \in F$, $\exists$ an integer $p > N$ s.t. $x \in J_p$

and $\ell(J_p) = 5\ell(I_p)$. Also

$$F \subset \bigcup_{p=N+1}^{\infty} J_p$$

$$\Rightarrow \qquad m^*(F) \le \sum_{p=N+1}^{\infty} \ell(J_p) = 5 \sum_{p=N+1}^{\infty} \ell(I_p) < 5 \cdot \frac{\varepsilon}{5} = \varepsilon$$

and hence the Lemma holds good.

1.1.4 Four Dini's Derivatives

The usual condition of differentiability of a function f (x) is too strong. Here we are studying the functions under slightly weaker condition (measurability). So why we define four quantities, called as Dini's Derivatives, which may be defined even at the points where the function is not differentiable.

1. $D^+ f(x) = \overline{\lim_{n \to 0_+}} \dfrac{f(x+h) - f(x)}{h}$, called *upper right derivative*

2. $D_+ f(x) = \underline{\lim_{h \to 0_+}} \dfrac{f(x+h) - f(x)}{h}$, called *lower right derivative*

3. $D^- f(x) = \overline{\lim_{h \to 0_-}} \dfrac{f(x+h) - f(x)}{h}$,

   or $\overline{\lim_{h \to 0_+}} \dfrac{f(x-h) - f(x)}{-h}$, called *upper left derivative*

4. $D_- f(x) = \underline{\lim_{h \to 0_-}} \dfrac{f(x+h) - f(x)}{h}$

   or $\underline{\lim_{h \to 0_+}} \dfrac{f(x-h) - f(x)}{-h}$, called *lower left derivative*

Notes

Notes

1. $D^+f(x) \geq D_+ f(x)$ and $\overline{D}f(x) \geq \underline{D}f(x)$

   If $D^+ f(x) = D_+ f(x)$, then we conclude that right hand derivative of f (x) exists at the point x and denoted by $f'(x^+)$. Similarly if $D^- f(x) = D_- f(x)$, we say that f (x) is left differentiable at x and denote this common value by $f'(x^-)$.

2. The function is said to be differentiable at x if all the four Dini's derivatives are equal but different than $\pm\infty$, i.e. if

   $D^+ f(x) = D_+ f(x) = D^- f(x) = D_- f(x) \neq \pm\infty$

   and their common value is denoted by f′ (x).

Properties of Dini's Derivatives

1. Dini's derivatives always exist, may be finite or infinite for every function f.
2. $D^+ (f + g) \leq D^+ f + D^+ g$ with similar properties for the other derivatives.
3. If f and g are continuous at a point 'x', then

   $D^+ (f . g)(x) \leq f(x) D^+ g(x) + g(x) D^+ f(x)$.
4. $D_+ f(x) = -D^+(-f(x))$

   and $D_- f(x) = -D^-(-f(x))$.
5. If f is a continuous function on [a, b] and one of its derivatives (say D^+) is non-negative on (a, b). Then f is non-decreasing on [a, b] i.e.

   $$f(x) \leq f(y) \text{ whenever } x \leq y,\ y \in [a, b].$$
6. If f is any function on an interval [a, b], then the four derivatives if exist are measurable.

1.1.5 Lebesgue Differentiation Theorem

Statement: Let $f : [a, b] \to R$ be a finite valued monotonically increasing function, then f is differentiable. Also $f : [a, b] \to R$ is L-integrable and

$$\int_a^b f'(x)\,dx \leq f(b) - f(a).$$

Proof: Define a sequence $\langle f_n \rangle$ of non-negative functions, where $f_n : [a, b] \to R$ such that,

$$f_n(x) = n\left[f\left(x + \frac{1}{n}\right) - f(x)\right], \forall x \in [a, b] \qquad \ldots (1)$$

and set f (x) = f (b), for $x \geq b$.

By hypothesis, $f : [a, b] \to R$ is an increasing function, therefore $f_n : [a, b] \to R$ is also an increasing function and hence integrable in the Lebesgue sense.

Again from (i) we have

$$\lim_{n\to\infty} f_n(x) = \lim_{1/n\to 0} \frac{f\{x + (1/n)\} - f(x)}{(1/n)}, \forall x \in [a, b],$$

$$= f'(x), \text{ a.e.}$$

Notes

Thus, the sequence $<f_n>$ converges to f′ (x), a.e.

Using Fatou's Lemma, we have

$$\int_a^b f'(x)\,dx \le \lim_{n\to\infty} \inf \left\{ \int_a^b f_n(x)\,dx \right\} \qquad \ldots \text{(ii)}$$

Again
$$\lim_{n\to\infty} \inf \int_a^b f_n(x)\,dx = \lim_{n\to\infty} \inf n \int_a^b \left[f\left(x - \frac{1}{n}\right) - f(x) \right] dx$$

$$= \lim_{n\to\infty} \inf n \left[\int_a^b f\left(x + \frac{1}{n}\right) dx - \int_a^b f(x)\,dx \right]$$

Putting t = x + (1/n), we get

$$\int_a^b f\left(x + \frac{1}{n}\right) dx = \int_{a+(1/n)}^{b+(1/n)} f(t)\,dt = \int_{a+(1/n)}^{b+(1/n)} f(x)\,dx$$

[By the first property of definite integrals]

$$\therefore \quad \lim_{n\to\infty} \inf \int_a^b f_n(x)\,dx = \lim_{n\to\infty} \inf n \left[\int_{a+(1/n)}^{b+(1/n)} f(x)\,dx - \int_a^b f(x)\,dx \right]$$

$$= \lim_{n\to\infty} \inf n \left[\int_b^{b+(1/n)} f(x)\,dx - \int_a^{a+(1/n)} f(x)\,dx \right] \qquad \ldots \text{(iii)}$$

Now extend the definition of f by assuming

$$f(x) = f(b), \ \forall\, x \in [b, b + 1/n].$$

$$\Rightarrow \quad \int_b^{b+(1/n)} f(x)\,dx = \int_b^{b+(1/n)} f(b)\,dx = \frac{1}{n} f(b)$$

Also f (a) ≤ f (x), for $x \in \left(a, a + \frac{1}{n}\right)$, therefore

$$\int_a^{a+(1/n)} f(x)\,dx \ge \int_a^{a+(1/n)} f(a)\,dx = \frac{1}{n} f(a)$$

$$\Rightarrow \quad -\int_a^{a+(1/n)} f(x)\,dx \le -\frac{1}{n} f(a)$$

$$\text{(iii)} \Rightarrow \quad \lim_{n\to\infty} \inf \int_a^b f_n(x)\,dx = \lim_{n\to\infty} \inf n \left[\int_b^{b+(1/n)} f(b)\,dx + \left(-\int_a^{a+(1/n)} f(x)\,dx \right) \right]$$

$$\le \lim_{n\to\infty} \inf n \left[f(b) \cdot \frac{1}{n} + \left(-\frac{1}{n} \right) f(a) \right] \le f(b) - f(a)$$

Thus from (ii), we get

$$\int_a^b f'(x)\,dx \le f(b) - f(a)$$

⇒ f′(x) is integrable and hence finite a.e. thus f is differentiable a.e.

Notes

Example: Let f be a function defined by f (0) = 0 and f (x) = x sin (1/x) for x ≠ 0. Find $D^+ f(0)$, $D_+ f(0)$, $D^- f(0)$, $D_- f(0)$.

$$D^+ f(0) = \overline{\lim_{h\to o^+}} \frac{f(0+h)-f(o)}{h} = \overline{\lim_{h\to o}} \frac{h\sin\frac{1}{h}-0}{h}$$

$$= \overline{\lim_{h\to o}} \sin\frac{1}{h} = 1, \text{ as } -1 \le \sin\frac{1}{h} \le 1$$

Also
$$D_+ f(0) = \overline{\lim_{h\to o^+}} \frac{f(0+h)-f(0)}{h} = \overline{\lim_{h\to o^+}} \sin\frac{1}{h} = -1$$

$$D^- f(0) = \overline{\lim_{h\to o^-}} \frac{f(0-h)-f(0)}{0-h} = \overline{\lim_{h\to o}} \frac{(-h)\sin\left(\frac{-1}{h}\right)-0}{-h}$$

$$= \overline{\lim_{h\to o}} -\sin\frac{1}{h} = 1$$

and
$$D_- f(0) = \underline{\lim_{h\to o}} \frac{f(0-h)-f(0)}{-h} = \underline{\lim_{h\to o}} \left(-\sin\frac{1}{h}\right) = -1$$

Theorem: Let x be a Lebesgue point of a function f (t); then the indefinite integral

$$F(x) = F(a) + \int_a^x f(t)\,dt$$

is differentiable at each point x and F′ (x) = f (x).

Proof: Given that x is a Lebesgue point of f (t), so that

$$\lim_{h\leftarrow o} \frac{1}{h}\int_x^{x+h} |f(t)-f(x)|\,dt = 0 \qquad \ldots \text{(i)}$$

Now
$$\frac{1}{h}\int_x^{x+h} f(x)\,dt = \frac{1}{h} f(x)\int_x^{x+h} 1\,dt = \frac{1}{h} f(x)\,[t]_x^{x+h}$$

$$= \frac{1}{h} f(x).h = f(x)$$

Thus
$$f(x) = \frac{1}{h}\int_x^{x+h} f(x)\,dt \qquad \ldots \text{(ii)}$$

Also
$$F(x+h) - F(x) = \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt$$

$$= \int_a^x f(t)\,dt + \int_x^{x+h} f(t)\,dt - \int_a^x f(t)\,dt = \int_x^{x+h} f(t)\,dt$$

$\Rightarrow$
$$\frac{F(x+h)-F(x)}{h} = \frac{1}{h}\int_x^{x+h} f(t)\,dt \qquad \ldots \text{(iii)}$$

Notes

From (ii) and (iii) we have

$$\left|\frac{F(x+h)-F(x)}{h}-f(x)\right| = \left|\frac{1}{h}\int_x^{x+h} f(t)dt-\frac{1}{h}\int_x^{x+h} f(x)dt\right|$$

$$= \left|\frac{1}{h}\int_x^{x+h}\left[f(t)-f(x)\right]dt\right| \le \frac{1}{h}\int_x^{x+h}\left|f(t)-f(x)\right|dt$$

$$\therefore \quad \lim_{h\leftarrow o}\left|\frac{F(x+h)-F(x)}{h}-f(x)\right| \le \lim_{h\to o}\frac{1}{h}\int_x^{x+h}\left|f(t)-f(x)\right|dt \le 0 \qquad \text{[Using (i)]}$$

$$\text{or} \quad \lim_{h\to o}\left|\frac{F(x+h)-F(x)}{h}-f(x)\right| \le 0 \qquad \ldots \text{(iv)}$$

Since modulus of any quantity is always positive, therefore

$$\lim_{h\to o}\left|\frac{F(x+h)-F(x)}{h}-f(x)\right| \ge 0 \qquad \ldots \text{(v)}$$

Combining (iv) and (v), we obtain

$$\lim_{h\to o}\left|\frac{F(x+h)-F(x)}{h}-f(x)\right| = 0$$

$$\Rightarrow \quad \lim_{h\to o}\frac{F(x+h)-F(x)}{h} = f(x)$$

$$\Rightarrow \quad F'(x) = f(x).$$

Theorem: Every point of continuity of an integrable function f (t) is a Lebesgue point of f (t).

Proof: Let f (t) be integrable over the closed interval [a, b] and let f (t) be continuous at the point x_o.

f (t) is continuous at $t = x_o$ implies that $\forall\, \varepsilon > 0, \exists$ a $\delta > 0$ such that,

$|f(t) - f(x_o)| < \varepsilon$, whenever $|t - x_o| < \delta$.

$$\Rightarrow \quad \int_{x_o}^{x_o+h}|f(t)-f(x_o)|\,dt < \varepsilon\int_{x_o}^{x_o+h}dt + \varepsilon h \text{ whenever } |h| < \delta.$$

$$\therefore \quad \frac{1}{h}\int_{x_o}^{x_o+h}|f(t)-f(x_o)|\,dt < \varepsilon \qquad \ldots \text{(i)}$$

Now $h \to 0 \Rightarrow \varepsilon \to 0$. So from (i), we have

$$\lim_{h\to o}\frac{1}{h}\int_{x_o}^{x_o+h}|f(t)-f(x_o)|\,dt < 0 \qquad \ldots \text{(ii)}$$

$$\text{Now} \quad \lim_{h\to o}\left|\frac{1}{h}\int_{x_o}^{x_o+h}|f(t)-f(x_o)|\,dt\right|$$

$$\le \lim_{h\to o}\frac{1}{h}\int_{x_o}^{x_o+h}|f(t)-f(x_o)|\,dt \le 0 \qquad \text{[Using (ii)]}$$

Notes

or $\lim_{h \to o} \left| \frac{1}{h} \int_{x_o}^{x_o+h} |f(t) - f(x_o)| \, dt \right| = 0$ [$\because$ Modulus of any quantity is always non-negative]

i.e. $\lim_{h \to o} \frac{1}{h} \left| \int_{x_o}^{x_o+h} |f(t) - f(x_o)| \, dt \right| = 0$.

This shows that x_o is a Lebesgue point of f (t).

1.2 Summary

- A function f defined on [a, b] is said to satisfy Lipschitz condition if $\exists$ a constant M > 0 such that

$$| f(x) - f(y) \le M |x - y|, \ \forall \ x, y \in [a, b].$$

- A point x is said to be a Lebesgue point of the function f (t), if

$$\lim_{h \to o} \frac{1}{h} \int_{x}^{x+h} |f(t) - f(x)| dt \le 0 \quad = 0$$

- Let E be a set of finite outer measure and M be a family of intervals which cover E in the sense of Vitali; then for a given $\in$ > 0, it is possible to find a finite family of disjoint intervals

  $\{I_k, k = 1, 2, \ldots, n\}$ of M, such that

$$m^* \left[E - \bigcup_{k=1}^{n} I_k \right] < \in .$$

- Lebesgue differentiation theorem: Let f : [a, b] $\to$ R be a finite valued monotonically increasing function, then f is differentiable. Also

  f : [a, b] $\to$ R is L-integrable and

$$\int_a^b f'(x) \, dx \le f(b) - f(a) .$$

1.3 Keywords

Dinni's Derivatives: These are the ways to define the quantities to judge the measurability of the functions even at the points where it is not differentiable.

Fundamental Theorem of the Integral: The fundamental theorem of the integral calculus is that differentiation and integration are inverse processes.

Measurable functions: An extended real valued function f defined over a measurable set E is said to be measurable in the sense of Lebesgue if set

E (f > a) = {x $\in$ E : f (x) > a} is measurable for all extended real numbers a.

Vitali's Lemma: Let E be a set of finite outer measure and M be a family of intervals which cover E in the sense of Vitali; then for a given $\varepsilon > 0$, it is possible to find a finite family of disjoint intervals $\{I_k, k = 1, 2, \ldots n\}$ of M, such that

$$m^* \left[E - \bigcup_{k=1}^{n} I_k \right] < \varepsilon.$$

Notes

1.4 Review Questions

1. If the function f assumes its maximum at c, show that D+ f (c) ≤ 0 and D_ f (c) ≥ 0.
2. Give an example of functions such that $D^+ (f + g) \neq D^+ f + D^+ g$.
3. Find the four Dini's derivatives of function f : [0, 1] → R

   such that f (x) = 0, if x ∈ 0, if x ∈ Q and f (x) = 1, if x ≠ Q.
4. Evaluate the four Dini's derivative at x = 0 of the function f (x) given below:

$$f(x) = \begin{cases} ax \sin^2 \frac{1}{x} + bx \cos^2 \frac{1}{x}, x > 0 \\ px \sin^2 \frac{1}{x} + qx \cos^2 \frac{1}{x}, x < 0 \end{cases}$$

   and f (0) = 0, given that a < b, p < q.
5. Every point of continuity of an integrable function f (t) is a Lebesgue point of f (t). Elucidate.

1.5 Further Readings

Books

J. Yeh, *Real Analysis: Theory of Measure and Integration*

Bartle, Robert G. (1976). *The Elements of Real Analysis* (second edition ed.)

Online links

www.solitaryroad.com/c756.html

www.public.iastate.edu/.../Royden_Real_Analysis_Solutions.pdf

Notes

Unit 2: Functions of Bounded Variation

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define absolute continuous function.
- Define monotonic function.
- Understand functions of bounded variation.
- Solve problems on functions of bounded variation.

Introduction

Functions of bounded variation is a special class of functions with finite variation over an interval. In Mathematical analysis, a function of bounded variation, also known as a BV function, is a real-valued function whose total variation is bounded: the graph of a function having this property is well behaved in a precise sense. Functions of bounded variation are precisely those with respect to which one may find Riemann – Stieltjes integrals of all continuous functions.

In this unit, we will study about absolute continuous function, Monotonic function and functions of bounded variation.

2.1 Functions of Bounded Variation

2.1.1 Absolute Continuous Function

A real-valued function f defined on [a,b] is said to be absolutely continuous on [a,b], if for an arbitrary $\in > 0$, however small, $\exists\, a, \delta > 0$, such that

$$\sum_{r=1}^{n} \left| f(b_r) - f(a_r) \right| < \in, \text{wherever} \sum_{r=1}^{n} (b_r - a_r) < \delta,$$

Notes

where $a_1 < b_1 \le a_2 < b_2 \le ... \le a_n < b_n$ i.e., a_i's and b_i's are forming finite collection $\{(a_i, b_i) : i = 1,2,...,n\}$ of pair-wise disjoint (non-overlapping) intervals (or of disjoint closed intervals).

Obviously, every absolutely continuous function is continuous.

Notes

- If a function satisfies $\left|\sum \{f(b_r) - f(a_r)\}\right| < \varepsilon$, even then it is absolutely continuous.
- The condition $\sum_{r=1}^{n} (b_r - a_r) < \delta$, means that total length of all the intervals must be less than δ.

2.1.2 Monotonic Function

Recall that a function f defined on an interval I is said to be monotonically non-increasing, iff

$$x > y \Rightarrow f(x) \le f(y), \forall x, y \in I$$

and monotonically non-decreasing, iff

$$x > y \Rightarrow f(x) \ge f(y), \forall x, y \in I$$

Also f is said to be strictly decreasing, iff

$$x > y \Rightarrow f(x) < f(y)$$

and strictly increasing, iff

$$x > y \Rightarrow f(x) > f(y)$$

2.1.3 Functions of Bounded Variation – Definition

Let f be a real-valued function defined on [a,b] which is divided by means of points

$$a = x_0 < x_1 < x_2 < ... < x_n = b.$$

Then the set $P = \{x_0, x_1, x_2, ..., x_n\}$ is termed as subdivision or partition of [a,b].

Let us take $\overset{b}{\underset{a}{V}}(f,P) = \sum_{r=0}^{n-1} \left|f(x_{r+1}) - f(x_r)\right|$, and $\overset{b}{\underset{a}{V}}(f,P) = \sup \overset{b}{\underset{a}{V}}(f,P)$ for all possible subdivisions P of [a,b]. (This is called total variation of f over [a,b] and also denoted by $\overset{b}{\underset{a}{T}}(f)$).

Notes

If $\overset{b}{\underset{a}{V}}(f)$ is finite, then f is called a function of bounded variation or function of finite variation over [a,b].

Set of all the functions of bounded variation on [a,b] is denoted by BV [a,b].

Notes

If f is defined on R, then we define

$$\overset{\infty}{\underset{-\infty}{V}}(f) = \lim_{a\to\infty} \overset{a}{\underset{-a}{V}}(f).$$

Some important observations about the functions of bounded variations.

Let f: $[a,b] \to R$ and P be any subdivision of [a,b]. Then:

(i) $|f(x) - f(a)| \le \overset{x}{\underset{a}{V}}(f), x \in [a,b]$

(ii) $\overset{a}{\underset{a}{V}}(f) = 0$

(iii) $P_1 \subset P_2 \Rightarrow \overset{b}{\underset{a}{V}}(f,P_1) \le \overset{b}{\underset{a}{V}}(f,P_2)$, where P_1 and P_2 are any two subdivisions of [a,b].

(iv) $\overset{b}{\underset{a}{V}}(f,P) \le \overset{b}{\underset{a}{V}}(f)$, for all subdivisions P of [a,b].

(v) For each $\varepsilon > 0$, however small, $\exists$ at least one subdivision P′ of [a,b] such that

$$\overset{b}{\underset{a}{V}}(f,P') + \varepsilon > \overset{b}{\underset{a}{V}}(f).$$

(vi) $\overset{b}{\underset{a}{V}}(f) \ge 0.$

(vii) $a < b < c \Rightarrow \overset{b}{\underset{a}{V}}(f) \le \overset{c}{\underset{a}{V}}(f).$

2.1.4 Theorems and Solved Examples

Theorem 1: A monotonic function on [a,b] is of bounded variation.

Proof: Divide the interval [a,b] by means of points

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

without any loss of generality, we can take f(x) as increasing function on [a,b]. Since if f is a decreasing function, –f is an increasing function and so by taking –f = g, we see that g is an increasing function and so we are allowed to consider only increasing functions. Thus

$$x_r < x_{r+1} \Rightarrow f(x_r) \le f(x_{r+1})$$

$$\Rightarrow f(x_{r+1}) - f(x_r) \ge 0$$

$$\Rightarrow |f(x_{r+1}) - f(x_r)| = f(x_{r+1}) - f(x_r) \qquad \ldots(i)$$

Notes

Now $V = \sum_{r=0}^{n-1} \left| f(x_{r+1}) - f(x_r) \right| = \sum_{r=0}^{n-1} \left\{ f(x_{r+1}) - f(x_r) \right\}$ [using (i)]

$\therefore \quad V = f(x_n) - f(x_0) = f(b) - f(a).$

Now f is monotonic $\Rightarrow f(b)$ and $f(a)$ are finite quantities.

$\Rightarrow V =$ a finite quantity independent of the mode subdivision. Hence f is of bounded variation.

> *Notes*
>
> If f is a monotonic function on [a,b], then
>
> $$\overset{b}{\underset{a}{T}}(f) = \left| f(b) - f(a) \right|$$

Theorem 2: Let V, P, N denote total, positive and negative variations of a bounded function f on [a,b]; then prove that

V = P+N, and P–N= f(b)–f(a).

Proof: Let the interval [a,b] be divided by means of points

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

$$v = \sum_{r=0}^{n-1} \left| f(x_{r+1}) - f(x_r) \right|$$

If P denotes the sum of those differences $f(x_{r+1}) - f(x_r)$ which are +n for positive and –n for negative, then obviously,

v = p + n, f(b) – f(a) = p – n ...(i)

Let $P = \sup p, V = \sup v, N = \sup n,$ ...(ii)

where suprema are taken over all subdivisions of [a,b]. From (i), we have

v = 2p + f(a) – f(b), ...(iii)

v = 2n + f(b) – f(a). ...(iv)

Taking supremum in (iii) and (iv) and using (ii), we get

V = 2P + f(a) – f(b), ...(v)

V = 2N + f(b) – f(a). ...(vi)

By adding and subtracting, (v) and (vi) give

V = P+N and f(b) – f(a) = P–N.

Theorem 3: If f_1 and f_2 are non-decreasing functions on [a,b], then f_1–f_2 is of bounded variation on [a,b].

Proof: Let f = f_1 – f_2 defined on [a,b].

Then for any partition $P = \{a = x_0, x_1, \ldots, x_n = b\}$, we have

Notes

$$\sum\left|f(x_i)-f(x_{i-1})\right| \le \sum\left|f_1(x_i)-f_1(x_{i-1})\right| + \sum\left|f_2(x_i)-f_2(x_{i-1})\right|$$

$$\le \left[f_1(b)-f_1(a)\right]+\left[f_2(b)-f_2(a)\right]$$

as f_1 and f_2 are monotonically increasing.

$\Rightarrow \quad \overset{b}{\underset{a}{V}}(f) \le f_1(b)+f_2(b)-f_2(a)-f_1(a)$, which is a finite quantity.

$\Rightarrow \quad \overset{b}{\underset{a}{V}}(f) < \infty$ and hence f is of bounded variation.

Theorem 4: If $f \in BV\ [a,b]$ and $c \in (a,b)$, then $f \in BV\ [a,c]$ and $f \in BV\ [c,b]$. Also

$$\overset{b}{\underset{a}{V}}(f) = \overset{c}{\underset{a}{V}}(f) + \overset{b}{\underset{c}{V}}(f)$$

Proof: Since $f \in BV[a,b]$ and $[a,c] \subset [a,b]$ we get

$$\overset{c}{\underset{a}{V}}(f) < \overset{b}{\underset{a}{V}}(f) < \infty$$

$\Rightarrow f \in BV\ [a,c]$ and similarly $f \in BV\ [c,b]$.

Now if P_1 and P_2 are any subdivisions of [a,c] and [c,b] respectively, then $P = P_1 \cup P_2$ is a subdivision of [a,b].

$$\Rightarrow \overset{c}{\underset{a}{V}}(f,P_1) + \overset{b}{\underset{c}{V}}(f,P_2) = \overset{b}{\underset{a}{V}}(f,P) \le \overset{b}{\underset{a}{V}}(f).$$

But P_1 and P_2 are any subdivisions. So taking supremums on P_1 and P_2, we get

$$\overset{c}{\underset{a}{V}}(f) + \overset{b}{\underset{c}{V}}(f) \le \overset{b}{\underset{a}{V}}(f). \qquad ...(1)$$

Now let $P = \{a = x_0, x_1, x_2, ..., x_n = b\}$ be a subdivision of [a,b] and $c \in [x_{r-1}, x_r]$

$\Rightarrow P_1 = \{x_0, x_1, x_2, ..., x_{r-1}, c\}$ and

$P_2 = \{c, x_r, x_{r+1}, x_{r+2}, ..., x_n\}$ are the subdivisions of [a,c] and [c,b] respectively.

Now $\overset{b}{\underset{a}{V}}(f,P) = \sum_{i=1}^{r-1}\left|f(x_i)-f(x_{i-1})\right| + \left|f(x_r)-f(x_{r-1})\right| + \sum_{i=r+1}^{n}\left|f(x_i)-f(x_{i-1})\right|$

$$= \sum_{i=1}^{r-1}\left|f(x_i)-f(x_{i-1})\right| + \left|f(x_r)-f(c)+f(c)-f(x_{r-1})\right| + \sum_{i=r+1}^{n}\left|f(x_i)-f(x_{i-1})\right|$$

$$\le \left[\sum_{i=1}^{r-1}\left|f(x_i)-f(x_{i-1})\right| + \left|f(c)-f(x_{r-1})\right|\right] + \left[\left|f(x_r)-f(c)\right| + \sum_{i=r+1}^{n}\left|f(x_i)-f(x_{i-1})\right|\right]$$

$$\le \overset{c}{\underset{a}{V}}(f,P_1) + \overset{b}{\underset{c}{V}}(f,P_2) \le \overset{c}{\underset{a}{V}}(f) + \overset{b}{\underset{c}{V}}(f)$$

(i) and (ii) $\Rightarrow \overset{b}{\underset{a}{V}}(f) = \overset{c}{\underset{a}{V}}(f) + \overset{b}{\underset{c}{V}}(f)$. ...(ii)

Notes

Notes

- This theorem enables us to define a new function (called variation function) say

  $V(x) = \overset{x}{\underset{a}{V}}(f), \forall\ x \in [a,b]$.

- If x > y in [a,b], then $\overset{y}{\underset{a}{V}}(f) = \overset{x}{\underset{a}{V}}(f) + \overset{y}{\underset{x}{V}}(f)$.

  i.e. v(y) = v(x) + $\overset{y}{\underset{x}{V}}(f)$.

  $\Rightarrow v(x)$ is an increasing function.

- If $a < c_1 < c_2 < \ldots < c_n < b$, then

  $\overset{b}{\underset{a}{V}}(f) = \overset{c_1}{\underset{a}{V}}(f) + \overset{c_2}{\underset{c_1}{V}}(f) + \ldots + \overset{b}{\underset{c_n}{V}}(f)$

Corollary:

$f \in BV[a,b] \Leftrightarrow f \in BV[a,c]$,

$f \in BV[c,b]$ for each $c \in [a,b]$.

Theorem 5: If a function f of bounded variation in [a,b] is continuous at $c \in [a,b]$, then the function defined by v(x) = $\overset{x}{\underset{a}{V}}(f)$, is also continuous at x = c and vice versa.

Proof: Suppose f is continuous at x = c. Hence for arbitrary $\in/2 > 0$, we can find a δ_1 such that

$a \le c - \delta_1 < x < c$ or $|x - c| < \delta_1 \Rightarrow |f(x) - f(c)| < \in/2$ ...(i)

Also we know by remark (v) after the definition (2.1.3), for above $\in$, we can get a subdivision

P = $\{a = x_0, x_1, x_2, \ldots, x_n = c\}$ of [a,c]

s.t. $\overset{c}{\underset{a}{V}}(f) < \overset{c}{\underset{a}{V}}(f,P) + \frac{\in}{2}$ ...(ii)

Now choosing positive $\delta > \min[\delta_1, c - x_{n-1}]$, we get that for any x such that $c - \delta < x < c$, we also have $x_{n-1} < x < x_n$.

$$\text{(ii)} \Rightarrow \overset{c}{\underset{a}{V}}(f) < \sum_{r=1}^{n-1} |f(x_r) - f(x_{r-1})| + |f(x_n) - f(x_{n-1})| + \frac{\in}{2}$$

$$< \sum_{r=1}^{n-1} |f(x_r) - f(x_{r-1})| + |f(x_n) - f(x) + f(x) - f(x_{n-1})| + \frac{\in}{2}$$

Notes

$$< \sum_{r=1}^{n-1} \left| f(x_r) - f(x_{r-1}) \right| + \left| f(x) - f(x_{n-1}) \right| + \left| f(x_n) - f(x) \right| + \frac{\in}{2}$$

$$< \overset{x}{\underset{a}{V}}(f) + \left| f(c) - f(x) \right| + \frac{\in}{2}$$

$$\Rightarrow \overset{c}{\underset{a}{V}}(f) < \overset{x}{\underset{a}{V}}(f) + \in, \qquad \text{[by (i)]}$$

$$\Rightarrow 0 = \overset{c}{\underset{a}{V}}(f) - \overset{x}{\underset{a}{V}}(f) < \in,$$

$$\Rightarrow \forall x \text{ s.t. } c - \delta < x < c, \text{we have } v(c) - v(x) < \in$$

$$\Rightarrow \underset{x \to c-0}{\text{lt}}\, v(x) = v(c).$$

$\Rightarrow v(x)$ is continuous on the left at x = c.

Similarly considering the partition of [c,b], one can show that v(x) is right continuous also at x = c and hence v(x) is also continuous at x = c.

Converse of the above Theorem

If v(x) is continuous at $x = c = \in [a,b]$ so is f also at x – c.

Proof: Since v(x) is continuous at x = c, for arbitrary small $\in > 0, \exists$ a $\delta > 0$ such that

$$\left| v(x) - v(c) \right| < \in, x \in (c - \delta, c + \delta) \qquad \text{...(i)}$$

Now let $c < x < c + \delta$. Then by Note (ii) of Theorem 4, we get

$$\overset{x}{\underset{a}{V}}(f) = \overset{c}{\underset{a}{V}}(f) + \overset{x}{\underset{c}{V}}(f)$$

$$\Rightarrow \quad v(x) = v(c) + \overset{x}{\underset{c}{V}}(f)$$

$$\Rightarrow \quad v(x) - v(c) = \overset{x}{\underset{c}{V}}(f) \geq \left| f(x) - f(c) \right|$$

$$\Rightarrow \quad \left| f(x) - f(c) \right| \leq \left| v(x) - v(c) \right| < \in \text{ [by (i)]} \qquad \text{...(ii)}$$

Similarly, we can show that $\left| f(c) - f(x) \right| < \in$, if $c - \delta < x < c$. ...(iii)

(ii) and (iii) show that f(x) is also continuous at x = c.

Theorem 6: Let f and g be functions of bounded variation on [a,b] ; then prove that f+g, f-g, fg and f/g $\left(\left| g(x) \right| \geq \sigma > 0, \forall x \right)$ and cf are functions of bounded variation, c being constant.

Proof:

(i) Set f + g = h, then

$$\left| h(x_{r+1}) - h(x_r) \right| = \left| \left[f(x_{r+1}) + g(x_{r+1}) \right] - \left[f(x_r) + g(x_r) \right] \right|$$

where $a = x_0 < x_1 < x_2 < \ldots < x_n = b$

Notes

$$= \left|\left[f(x_{r+1}) - f(x_r)\right] + \left[g(x_{r+1}) - g(x_r)\right]\right|$$

$$\le \left|\left[f(x_{r+1}) - f(x_r)\right] + \left[g(x_{r+1}) - g(x_r)\right]\right| .$$

$$\therefore \sum_{r=0}^{n-1}\left|h(x_{r+1}) - h(x_r)\right| \le \sum_{r=0}^{n-1}\left|h(x_{r+1}) - h(x_r)\right| + \sum_{r=0}^{n-1}\left|g(x_{r+1}) - g(x_r)\right| .$$

or $\overset{b}{\underset{a}{V}}(h) \le \overset{b}{\underset{a}{V}}(f) + \overset{b}{\underset{a}{V}}(g).$

Now by hypothesis, f, g are functions of bounded variations.

$\Rightarrow \quad \overset{b}{\underset{a}{V}}(f)$ and $\overset{b}{\underset{a}{V}}(g)$ are finite.

$\Rightarrow \quad \overset{b}{\underset{a}{V}}(h) =$ a finite quantity.

Hence h = f + g is of bounded variation in [a,b].

(ii) Let h = f – g. Then as above,

$$\left|h(x_{r+1}) - h(x_r)\right| \le \left|\left[f(x_{r+1}) - f(x_r)\right] + \left[g(x_{r+1}) - g(x_r)\right]\right|.$$

$\Rightarrow \quad \overset{b}{\underset{a}{V}}(h) \le \overset{b}{\underset{a}{V}}(f) + \overset{b}{\underset{a}{V}}(g)$

$\Rightarrow \quad \overset{b}{\underset{a}{V}}(h) =$ a finite quantity.

Hence h = f – g is of bounded variation in [a,b].

(iii) Let h(x) = f(x).g(x). Then

$$\left|h(x_{r+1}) - h(x_r)\right| = \left|f(x_{r+1}).g(x_{r+1}) - f(x_r).g(x_r)\right|$$

$$= \left|f(x_{r+1}).g(x_{r+1}) - f(x_r)g(x_{r+1}) + f(x_r)g(x_{r+1}) - f(x_r)g(x_r)\right|$$

$$\le \left|g(x_{r+1})\left[f(x_{r-1}) - f(x_r)\right]\right| + \left|f(x_r)\left[g(x_{r+1}) - g(x_r)\right]\right| .$$

$$\le \left|g(x_{r+1})\right|\left|f(x_{r-1}) - f(x_r)\right| + \left|f(x_r)\right|.\left|g(x_{r+1}) - g(x_r)\right| .$$

Let A = $\sup\left\{\left|f(x)\right| : x \in [a,b]\right\}$,

B = $\sup\left\{\left|g(x)\right| : x \in [a,b]\right\}$,

$$\therefore \left|h(x_{r+1}) - h(x_r)\right| \le B.\left|f(x_{r+1}) - f(x_r)\right| + A.\left|g(x_{r+1}) - g(x_r)\right|.$$

$$\therefore \sum_{r=0}^{n-1}\left|h(x_{r+1}) - h(x_r)\right| \le B.\sum_{r=0}^{n-1}\left|h(x_{r+1}) - h(x_r)\right| + A \cdot \sum_{r=0}^{n-1}\left|g(x_{r+1}) - g(x_r)\right| .$$

Notes

i.e. $\overset{b}{\underset{a}{V}}(h) \le B.\overset{b}{\underset{a}{V}}(f) + A.\overset{b}{\underset{a}{V}}(g).$

= a finite quantity.

Hence h(x) = f(x).g(x) is of bounded variation in [a,b].

(iv) First, we shall show that 1/g is of bounded variation, where $g(x) \ge \sigma > 0, \forall x \in [a,b]$.

Now, $g(x) \ge \sigma > 0, \forall x \in [a,b]$

$$\Rightarrow \frac{1}{g(x)} \le \frac{1}{\sigma} > 0, \forall x \in [a,b].$$

Again, we observe that

$$\left|\frac{1}{g(x_{r+1})} - \frac{1}{g(x_r)}\right| = \left|\frac{g(x_r) - g(x_{r+1})}{g(x_r).g(x_{r+1})}\right| \le \frac{1}{\sigma^2}\left|g(x_r) - g(x_{r+1})\right|$$

$$\therefore \sum_{r=0}^{n-1}\left|\frac{1}{g(x_{r+1})} - \frac{1}{g(x_r)}\right| \le \frac{1}{\sigma^2}\sum_{r=0}^{n-1}\left|g(x_r) - g(x_{r+1})\right|$$

$$\Rightarrow \overset{b}{\underset{a}{V}}\left(\frac{1}{g}\right) \le \frac{1}{\sigma^2}\overset{b}{\underset{a}{V}}(g) = \text{a finite quantity.}$$

Hence $\frac{1}{g}$ is of bounded variation in [a,b].

Now f and $\frac{1}{g}$ are of bounded variation in [a,b].

$\Rightarrow f.\frac{1}{g}$ is of bounded variation in [a,b] [by case (iii)]

$\Rightarrow \frac{f}{g}$ is of bounded variation in [a,b].

(v) Do yourself. Note that $\overset{b}{\underset{a}{V}}(cf) = |c|\overset{b}{\underset{a}{V}}(f)$.

Notes

Since BV [a,b] is closed for all four algebraic operations, it is a linear space.

Theorem 7: Every absolutely continuous function f defined on [a,b] is of bounded variation.

Proof: Since f is absolutely continuous on [a,b]; for $\in = 1, \exists$ a $\delta > 0$

Notes

s.t. $\sum_{i=1}^{n}\left|f(b_i)-f(a_i)\right|<1,$

whenever $\sum_{i=1}^{n}(b_i-a_i)<\delta,$

and $a=a_1<b_1\le a_2<b_2\le\ldots\le a_n<b_n=b.$

Now consider another subdivision of [a,b] or say refinement of P by adjoining some additional points to P in such a way that all the intervals can be divided into r parts each of total length less than δ.

Let the r sub-intervals be $[c_0,c_1]$, $[c_1,c_2],\ldots,[c_{r-1},c_r]$ such that

$$a = c_0,\ c_r = b \text{ and } (c_{k+1}-c_k) < \delta, \forall K=0,1,2,\ldots,(r-1)$$

Obviously, $\sum_i\left|f(x_{i+1})-f(x_{i+1})\right|<1,$ where $x_{i+1},x_i.\in[c_k,c_{k+1}]$

or $\overset{c_{k+1}}{\underset{c_k}{V}}(f)<1,$ [Using (i)]

Hence $\overset{b}{\underset{a}{V}}(f)=\overset{c_1}{\underset{c_0}{V}}(f)+\overset{c_2}{\underset{c_1}{V}}(f)+\ldots+\overset{c_r}{\underset{c_{r-1}}{V}}(f)<1+1+1+\ldots+1=r=\text{finite quantity}.$

Hence, f is of bounded variation.

> *Notes*
>
> Converse of above theorem is not necessarily true. These exists functions of bounded variation but not absolutely continuous.

Theorem 8: Jordan Decomposition Theorem

A function f is of bounded variation, if and only if it can be expressed as a difference of two monotonic functions both non-decreasing.

Proof: Let f be the function of $f:[a,b]\to R.$

Case I. $f\in BV[a,b].$ Then we can write

$$f = v - (v - f), \qquad \ldots(i)$$

so that $f(x)=v(x)-(v(x)-f(x)), x\in[a,b].$

Now if x, y $\in$ [a, b] such that x < y, then by the remark (ii) of theorem 4, we get

$$\overset{y}{\underset{a}{V}}(f)=\overset{x}{\underset{a}{V}}(f)+\overset{y}{\underset{x}{V}}(f).$$

$$\Rightarrow v(y)-v(x)=\overset{y}{\underset{x}{V}}(f)\ge 0$$

$\Rightarrow v(x)\le v(y)$ and hence v is a non-decreasing function on [a,b].

Again, if x<y in [a,b], then as above **Notes**

$$v(y)-v(x)=\overset{y}{\underset{x}{V}}(f)\geq\left|f(y)-f(x)\right|\geq f(y)-f(x)$$

$$\Rightarrow v(y)-f(y)\geq v(x)-f(x)\Rightarrow(v-f)y\geq(v-f)x$$

$\Rightarrow v-f$ is also a non-decreasing function on [a,b].

Thus (i) shows that f is expressible as a difference of two monotonically non-decreasing functions.

Case II. Set g (x) and h (x) be increasing functions such that f(x) = g(x) – h(x).

Divide the closed interval [a,b] by means of points

$a = x_0 < x_1 < x_2 < ... < x_n = b$.

Let $V=\sum_{r=0}^{n-1}\left|f(x_{r+1})-f(x_r)\right|$

Now, we have that

$$\begin{aligned}\left|f(x_{r+1})-f(x_r)\right|&=\left|g(x_{r+1})-h(x_{r+1})-\{g(x_r)-h(x_r)\}\right|\\&=\left|\left[g(x_{r+1})-g(x_r)\right]+\left[h(x_r)-h(x_{r+1})\right]\right|\\&\leq\left|g(x_{r+1})-g(x_r)\right|+\left|h(x_r)-h(x_{r+1})\right|\\&\leq\left|g(x_{r+1})-g(x_r)\right|+\left|h(x_{r+1})-h(x_r)\right|\end{aligned}$$

Now, g(x) and h(x) are monotonically increasing functions, so that $g(x_{r+1})-g(x_r)\geq 0$ and $h(x_{r+1})-h(x_r)\geq 0$

$$\Rightarrow\left|g(x_{r+1})-g(x_r)\right|=g(x_{r+1})-g(x_r)$$

and $\left|h(x_{r+1})-h(x_r)\right|=h(x_{r+1})-h(x_r)$.

Hence $\left|f(x_{r+1})-f(x_r)\right|\leq\left[g(x_{r+1})-g(x_r)\right]+\left[h(x_{r+1})-h(x_r)\right]$

$$\therefore\sum_{r=0}^{n-1}\left|f(x_{r+1})-f(x_r)\right|\leq\sum_{r=0}^{n-1}\left[g(x_{r+1})-g(x_r)\right]+\sum_{r=0}^{n-1}\left[h(x_{r+1})-h(x_r)\right]$$

Now $\sum_{r=0}^{n-1}\left[g(x_{r+1})-g(x_r)\right]=\left[g(x_1)-g(x_0)\right]+\left[g(x_2)-g(x_1)\right]+.....+....+\left[g(x_n)-g(x_{n-1})\right]$

$$\begin{aligned}&=g(x_n)-g(x_0)\\&=g(b)-g(a)\qquad(\because\ x_n=b,x_0=a)\end{aligned}$$

Similarly, $\sum_{r=0}^{n-1}\left[h(x_{r+1})-h(x_r)\right]=h(b)-h(a)$.

Hence $\sum_{r=0}^{n-1}\left[f(x_{r+1})-f(x_r)\right]\leq g(b)-g(a)+h(b)-h(a)$.

Notes

since f is finite in [a,b] Now $\Rightarrow g(b), g(a)h(b), h(a)$ are finite numbers.

$$\therefore \sum_{r=0}^{n-1}\left[f(x_{r+1})-h(x_r)\right]<\infty$$

$$\Rightarrow \overset{b}{\underset{a}{V}}(f)<\infty.$$

$\Rightarrow$ f is a function of bounded variation. Alternatively, since g (x) and h(x) are both non-decreasing, so by theorem 3, g(x) – h(x) and hence f(x) is of bounded variation.

Corollary: A continuous function is of bounded variation iff it can be expressed are as a difference of two continuous monotonically increasing functions. It follows from the results of Theorems 5 and 8.

Theorem 9: An indefinite integral is a function of bounded variation, i.e. if $f \in L[a,b]$ and $F(x)$ is indefinite integral of $f(x)$ i.e. F (x) = $\int_a^x f(t)dt$, then $F \in BV[a,b]$. Also show that

$$\overset{b}{\underset{a}{V}}(f) \le \int_a^x |f|.$$

Proof: Since $f \in L[a,b]$, also $|f| \in L[a,b]$.

Let $P = \{x_i : i = 0,1,2,...,n\}$ be a subdivision of the interval [a,b]. Then

$$\sum_{r=0}^{n}\left|F(x_i)-F(x_{i-1})\right| = \sum_{i=1}^{n}\left|\int_a^{x_i} f - \int_a^{x_{i-1}} f\right|$$

$$= \sum_{i=1}^{n}\left|\int_{x_{i-1}}^{x_i} f\right| \le \sum_{i=1}^{n}\int_{x_{i-1}}^{x_i}|f|$$

$$= \int_a^b |f| < \infty.$$

$\Rightarrow f \in BV[a,b]$ and $\overset{b}{\underset{a}{V}}(f,p) \le \int_a^b |f|.$

Further above result is true for any subdivision of P of [a,b]. Therefore taking supremum, we get

$$\int_a^b (f) \le \int_a^b |f|.$$

Notes

Example: A function f of bounded variation on [a,b] is necessarily bounded on [a,b] but not conversely.

Solution: If $x \in [a,b]$, then $|f(x)-f(a)| \geq \overset{x}{\underset{a}{V}}(f) \leq \overset{b}{\underset{a}{V}}(f) < \infty$

$$\Rightarrow -\overset{b}{\underset{a}{V}}(f) \leq f(x)-f(a) \leq \overset{b}{\underset{a}{V}}(f)$$

$$\Rightarrow f(a)-\overset{b}{\underset{a}{V}}(f) \leq f(x) \leq \overset{b}{\underset{a}{V}}(f)+f(a)$$

$\Rightarrow f(x)$ is bounded on [a,b]

For the converse, define the function f on [0,1] by

$$f(x)=\begin{cases} 0, \text{if } x=0 \\ x.\sin\left(\dfrac{\pi}{x}\right), \text{if } 0<x\leq 1 \end{cases}$$

since $0 \leq x \leq 1$ and $-1 \leq \sin\left(\dfrac{\pi}{x}\right) \leq 1$, the function f is obviously bounded. Now consider the partition

$$P=\left\{0, \frac{2}{2n+1}, \frac{2}{2n-1}, \ldots, \frac{2}{7}, \frac{2}{5}, \frac{2}{3}, 1\right\} \text{of } [0,1]$$

Where $n \in N$. Then we get

$$\overset{1}{\underset{0}{V}}(f,P)=\left|f\left(\frac{2}{2n+1}\right)-f(0)\right|+\ldots+\left|f\left(\frac{2}{3}\right)-f\left(\frac{2}{5}\right)\right|+\left|f(1)-f\left(\frac{2}{3}\right)\right|$$

$$=\left|\frac{2}{2n+1}(-1)^n-0\right|+\ldots+\left|\frac{2}{3}(-1)-\frac{2}{5}.1\right|+\left|0-\frac{2}{3}(-1)\right|$$

$$=\frac{2}{2n+1}+\ldots+\left(\frac{2}{3}+\frac{2}{5}\right)+\frac{2}{3}$$

$$=4.\left(\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{2n+1}\right).$$

But we know that series $\sum \frac{1}{2n+1}$ is divergent. Therefore letting $n \to \infty$ we get that

$$\overset{1}{\underset{0}{V}}(f)=\underset{n\to\infty}{\text{lt}}\overset{1}{\underset{0}{V}}(f,P)=\infty$$

$\Rightarrow$ f is not of bounded variation.

Notes

Example: Show that the function

$$f(x)=\begin{cases} x\sin\frac{\pi}{x} \text{ if } (0<x\le 2) \\ 0, \text{if } x=0 \end{cases} \text{is continuous}$$

without being of bounded variation.

or

show that there exists a continuous function without being of bounded variation.

Solution: We know that $\underset{x\to 0}{\text{lt}}\, f(x)=0=f(0)$

$\Rightarrow f(x)$ is continuous but not of bounded variation (see converse of above example.)

Hence the result.

Problem: Show that if f′ exists and is bounded on [a, b], then f ∈ BV [a, b].

Solution: According to given, let $|f'|\le M$ on [a, b].

Then for any X_{i-1}, $x_i \in$ [a, b], we get

$$\left|\frac{f(x_i)-f(x_{i-1})}{x_i-x_{i-1}}\right|\le M \Rightarrow |f(x_i)-f(x_{i-1})|\le M(x_i-x_{i-1})$$

$\Rightarrow$ for any partition P of [a, b],

$$\overset{b}{\underset{a}{V}}(f)\le M\sum(x_i-x_{i-1})=M(b-a)$$

$\Rightarrow$ $f \in$ B V [a, b].

Problem: Show that the function f defined as

$$f(x)=x^p\sin\frac{1}{x} \text{ for } 0<x\le 1,\ f(o)=0,\ p\ge 2.$$

is of bounded variation [0, 1].

Solution: Note that $RF'(0) = \underset{h\to o}{\text{Lim}}\dfrac{(0+h)^p\sin\frac{1}{h}-0}{h}$

$$= \underset{h\to o}{\text{Lim}}\, h^{(p-1)}\sin\frac{1}{h}=0$$

and $Lf'(0)=\lim\limits_{h\to o}\dfrac{(-h)^p\sin\left(-\frac{1}{h}\right)-0}{-h}=0$

Notes

$\Rightarrow$ $f'(0) = 0$ and $f'(x) = x^p \cos\frac{1}{x}\left(-\frac{1}{x^2}\right) + px^{p-1}\sin\frac{1}{x}$

$\Rightarrow$ $f'(x) = x^{p-2}\left[px\sin\frac{1}{x} - \cos\frac{1}{x}\right]$, for $0 < x \le 1$

$\Rightarrow$ $f'(x)$ is bounded for $0 \le x \le 1$.

According to above problem, $f \in BV[0, 1]$.

2.2 Summary

- A real-valued function f defined on [a,b] is said to be absolutely continuous on [a,b], if for an arbitrary $\in > 0$, however small, $\exists\, a, \delta > 0$, s.t.

  $$\sum_{r=1}^{n}\left|f(b_r) - f(a_r)\right| < \in \text{ whenever } \sum_{r=1}^{n}(b_r - a_r) < \delta,$$

  where $a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n$

- A function f defined on an interval I is said to be monotonically non-increasing, iff

  $x > y \Rightarrow f(x) \le f(y), \forall x, y \in I.$

  and monotonically non-decreasing, iff $x > y \Rightarrow f(x), \ge f(y)\ \forall\ x, g \in I$.

- Let $\overset{b}{\underset{a}{V}}(f,P) = \sum_{r=0}^{n-1}\left|f(x_{r+1}) - f(x_r)\right|$, and $\overset{b}{\underset{a}{V}}(f) = \text{Sup}\,\overset{b}{\underset{a}{V}}(f,P)$ for all possible subdivisions P of [a,b]. If $\overset{b}{\underset{a}{V}}(f)$ is finite, then f is called a function of bounded variation over [a,b].

2.3 Keywords

Absolute Continuous Function: A real valued function f defined on [a, b] is said to be absolutely continuous on [a, b], if for an arbitrary $\varepsilon > 0$, however small, $\exists\, a, \delta > 0$, such that

$$\sum_{r=1}^{n}|f(b_r) - f(a_r)| < \varepsilon, \text{ wherever } \sum_{r=1}^{n}(b_r - a_r) < \delta$$

where $a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n$ i.e. a_1's and b_1's are forming finite collection $\{(a_i, b_i) : i = 1, 2, \ldots, n\}$ of pair-wise disjoint intervals.

Continuous: A continuous function is a function $f : X \to Y$ where the pre-image of every open set in Y is open in X.

Disjoint: Two sets A and B are said to be disjoint if they have no common element, i.e. $A \cap B = \phi$.

Notes

Monotonic Decreasing Function: A monotonic decreasing function is a function that either decreases or remains the same, never increases i.e. a function f(x) such that $f(x_2) \le f(x_1)$ for $x_2 > x_1$.

Monotonic Function: A monotonic function is a function that is either a monotonic increasing or monotonic decreasing.

Monotonic Increasing Function: A monotonic increasing function is a function that either increases or remains the same, never decreases i.e. a function f(x) such that $f(x_2) \ge f(x_1)$ for $x_2 > x_1$.

2.4 Review Questions

1. Show that sum and product of two functions of bounded variation is again a function of bounded variation.
2. Show that the function f defined on [0,1] by

$$f(x) = \begin{cases} x\cos\left(\frac{\pi x}{2}\right) \text{for } 0 < x \le 1 \\ 0 \text{ for } x = 0 \end{cases}$$

   is continuous but not of bounded variation on [0,1].

3. Show that the function f defined on [0,1] as f(x) = $x\sin\left(\frac{\pi}{x}\right)$ for $x > 0$, f(0)=0 is continuous but is not of bounded variation on [0,1].
4. Define a function of bounded variation on [a,b]. Show that every increasing function on [a,b] is of bounded variation and every function of bounded variation on [a,b] is differentiable on [a,b].
5. Show that a continuous function may not be of bounded variation.
6. Show that a function of bounded variation may not be continuous.
7. If f is a function such that its derivative f′ exists and is bounded. Then prove that the function f is of bounded variation.

2.5 Further Readings

Books

Halmos, Paul (1950), *Measure Theory*, Van Nostrand and Co.

Kolmogorov, Andrej N.; Fomin, Sergej V. (1969). *Introductory Real Analysis*, New York: Dovers Publications.

Online links

www.ams.org

www.whitman.edu/mathematics/SeniorProjectArchive/.../grady.pdf

Notes

Unit 3: Differentiation of an Integral

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define differentiation of an integral.
- Solve problems related to it.

Introduction

If f is an integrable function on [a, b], we define its indefinite integral to be the function F defined on [a, b] by

$$F(x) = \int_a^x f(t)\,dt$$

Here, it is shown that the derivative of the indefinite integral of an integrable function is equal to the integrand almost everywhere. We begin by establishing some lemmas.

3.1 Differentiation of an Integral

If f is an integrable function on [a, b] then f is integrable on any interval $[a, x] \subset [a, b]$. The function F given by

$$F(x) = \int_a^x f(t)\,dt + c\,,$$

where c is a constant, called the indefinite integral of f.

Lemma 1: If f is integrable on [a, b] then the indefinite integral of f namely the function F on [a, b] given by $F(x) = \int_a^x f(t)$ is a continuous function of bounded variation on [a, b].

Proof: Let x_o be any point of [a, b].

Notes

Then $$|F(x)-F(x_o)| = \left|\int_a^x f(t)\,dt - \int_a^{x_o} f(t)\,dt\right|$$

$$= \left|\int_a^x f(t)\,dt + \int_{x_o}^a f(t)\,dt\right|$$

$$= \left|\int_{x_o}^a f(t)\,dt + \int_a^x f(t)\,dt\right|$$

$$= \left|\int_{x_o}^x f(t)\,dt\right|$$

$$\le \left|\int_{x_o}^x |f(t)|\,dt\right|$$

But f is integrable on [a, b]

$\Rightarrow$ $|f|$ is integrable on [a, b]

[Since we know that measurable function f is integrable over E iff $|f|$ is integrable over E]

$\therefore$ Given $\varepsilon > 0$, $\exists\ \delta > 0$ such that for every measurable set $A \subset [a, b]$ with $m(A) < \delta$, we have $\int_A |f| < \varepsilon$ by theorem, "if f is a non-negative function which is integrable over a set E, then given $\varepsilon > 0$, there is a $\delta > 0$ such that for every set $A \subset E$ with $m(A) < \delta$, $\int_A f < \varepsilon$."

$$\Rightarrow \quad \left|\int_{x_o}^x |f(t)|\,dt\right| < \varepsilon, \text{ for } |x - x_o| < \delta.$$

$$\Rightarrow \quad |F(x) - F(x_o)| = \left|\int_{x_o}^x f(t)\,dt\right| \le \left|\int_{x_o}^x |f(t)|\,dt\right| < \in$$

whenever $|x - x_o| < \delta$.

$\Rightarrow$ $|F(x) - F(x_o)| < \varepsilon$ wherever $|x - x_o| < \delta$

$\Rightarrow$ F is continuous at x_o and hence in [a, b].

Now we shall show that F is a function of bounded variation.

Let $P = \{a = x_o < x_1 < x_2 < \ldots < x_n = b\}$ be a partition of [a, b].

Then

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| = \sum_{i=1}^n \left|\int_{x_{i-1}}^{x_i} f(t)\,dt\right|$$

Notes

$$\leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |f(t)|\,dt$$

$$= \int_a^b |f(t)|\,dt$$

$$\Rightarrow \qquad T_a^b(F) \leq \int_a^b |f(t)|\,dt$$

But $|f|$ is integrable therefore.

$$\int_a^b |f|\,dt < \infty$$

$$\Rightarrow \qquad T_a^b(F) < \infty$$

$$\Rightarrow \qquad F \in BV[a, b]$$

Hence the Proof.

Theorem 1: Let f be an integrable on [a, b].

If $\int_a^x f(t)\,dt = 0 \;\forall\; x \in [a, b]$ then f = 0 a.e. in [a, b].

Proof: Let if possible, $f \neq 0$ a.e. in [a, b].

Let f (t) > 0 on a set E of positive measure, then there exists a closed set $F \subset E$ with m (F) > 0.

Let A = (a, b) – F.

Then A is an open set.

Now $\quad \int_a^b f(t)\,dt = \int_{A \cup F} f(t)\,dt$

But $\quad \int_a^b f(t)\,dt = 0$

$$\Rightarrow \qquad \int_{A \cup F}^{b} f(t)\,dt = 0$$

$$\Rightarrow \qquad \int_A f(t)\,dt + \int_F f(t)\,dt = 0$$

$$\Rightarrow \qquad \int_A f(t)\,dt + \int_F f(t)\,dt = 0 \;\Rightarrow\; \int_A f(t)\,dt = -\int_F f(t)\,dt$$

Notes

But f (t) > 0 on F with m (F) > 0 implies

$$\int_F f(t)\,dt \neq 0$$

Therefore $\int_A f(t)\,dt \neq 0$

Now, A being as open set, it can be expressed as a union of countable collection $\{(a_n, b_n)\}$ of disjoint open intervals as we know that an open set can be expressed as a union of countable collection of disjoint open intervals.

Thus $\int_A f(t)\,dt = \sum_n \int_{a_n}^{b_n} f(t)\,dt$

But $\int_A f(t)\,dt \neq 0$

$\Rightarrow \quad \sum_n \int_{a_n}^{b_n} f(t)\,dt \neq 0$

$\Rightarrow \quad \int_{a_n}^{b_n} f(t)\,dt \neq 0$ for some n

$\Rightarrow \quad$ either $\int_a^{a_n} f(t)\,dt \neq 0$

Or $\int_a^{b_n} f(t)\,dt \neq 0$

In either case, we see that if f is positive on a set of positive measure, then for some $x \in [a, b]$ we have

$$\int_a^x f(t)\,dt \neq 0\,.$$

Similarly if f is negative on a set of positive measure we have

$$\int_a^x f(t)\,dt \neq 0\,.$$

But it leads to the contradiction of the given hypothesis. Hence our supposition is wrong.

f = 0 a.e. in [a, b].

Hence the proof.

Notes

Theorem 2: First fundamental theorem of calculus statement: If f is bounded and measurable on [a, b] and F (x) = $\int_a^x f(t)\,dt$ + F (a), then F′(x) = f (x) a.e. in [a, b].

Proof: Since every indefinite integral is a function of bounded variation, therefore F (x) is a function of bounded variation over [a, b]. Thus F (x) can be expressed as a difference of two monotonic functions and since every monotonic function has a finite differential coefficient at every point of a set of non-zero measure, therefore F (x) has a finite differential coefficient a.e. in [a, b]. Now F is given to be bounded;

$$\therefore \qquad |f| \le M \text{ (say)} \qquad \ldots (1)$$

Let
$$f_n(x) = \frac{F(x+h)-F(x)}{h}.$$

with $h = \frac{1}{x}$.

Then
$$|f_n(x)| = \left|\frac{1}{h}(F(x+h)-F(x)\right|$$

$$= \left|\frac{1}{h}\left(\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt\right)\right|$$

$$= \left|\frac{1}{h}\left(\int_a^{x+h} f(t)\,dt + \int_x^a f(t)\,dt\right)\right|$$

$$= \left|\frac{1}{h}\left(\int_x^a f(t)\,dt + \int_a^{x+h} f(t)\,dt\right)\right|$$

$$= \left|\frac{1}{h}\int_x^{x+h} f(t)\,dt\right|$$

But
$$|f| \le M$$

$$\therefore \qquad |f_n(x)| \le \frac{M}{h}\int_x^{x+h} dt = \frac{M}{h}(x+h-x)$$

$$\Rightarrow \qquad |f_n(x)| \le \frac{M}{h}(h)$$

$$\Rightarrow \qquad |f_n(x)| \le M$$

Since $f_n(x) \to F'(x)$ a.e.,

Notes

then the bounded convergence theorem implies that

$$\int_a^x F'(x)\,dx = \lim_{h\to\infty}\int_a^x F_x(x)\,dx$$

$$= \lim_{h\to 0}\frac{1}{h}\int_a^x [F(x+h)-F(x)]\,dx$$

$$= \lim_{h\to 0}\left[\frac{1}{h}\int_x^{x+h}[F(x)dx - \frac{1}{h}\int_a^{a+h}F(x)\,dx\right]$$

$$= F(x) - F(a)$$

$$= \int_a^x f(t)\,dt, \text{ by hypothesis}$$

$$\left[\because F(x) = \int_a^x f(t)\,dt + F(a) \Rightarrow F(x) - F(a) = \int_a^x f(t)\,dt\right]$$

or $\int_a^x [F'(t) - f(t)]\,dt = 0, \forall x$

$\Rightarrow$ $F'(x) - f(x) = 0$ a.e. in [a, b]

Hence $F'(x) = f(x)$ a.e. in [a, b] by the theorem, "If f is integrable on [a, b] and $\int_a^x f(t)\,dt = 0, \forall x \in [a, b]$ then f = 0 a.e. in [a, b]".

Hence $F'(x) = f(x)$ a.e. in [a, b].

Hence the proof.

Theorem 3: If f is an integrable function on [a, b] and if $F(x) = \int_a^x f(t)\,dt + F(a)$ then $F'(x) = f(x)$ a.e. in [a, b].

Proof: Without loss of generality, we may assume that $f(x) \geq 0 \ \forall x$

Let us define a sequence $\{f_n\}$ of functions

$f_n : [a, b] \to R$, where

$$f_n(x) = \begin{cases} f(x) \text{ if } f(x) \leq n, \\ n \text{ if } f(x) > n \end{cases}$$

Clearly, each f_n is bounded and measurable function and so, by the theorem,

Let f be a bounded and measurable function defined on [a, b]. If $F(x) = \int_a^x f(t)\,dt + F(a)$, then $F'(x) = f(x)$ a.e. in [a, b]", we have

$$\frac{d}{dx}\int_a^x f_n = f_n(x) \text{ a.e.}$$

Notes

Also, $f - f_n > 0 \ \forall$ n, and therefore, the function G_n defined by

$$G_n(x) = \int_a^x (f - f_n)$$

is an increasing function of x, which must have a derivative almost everywhere by Lebesgue theorem and clearly, this derivative must be non-negative.

Since $$G_n(x) = \int_a^x (f - f_n)$$

$$= \int_a^x f(t)\,dt - \int_a^x f_n(t)\,dt$$

$$\Rightarrow \quad \int_a^x f(t)\,dt = G_n(x) + \int_a^x f_n(t)\,dt$$

Now the relation

$$F(x) = \int_a^x f(t)\,dt + F(a) \text{ becomes}$$

$$F(x) = G_n(x) = \int_a^x f_n(t)\,dt + F(a),$$

$$\Rightarrow \quad F'(x) = G_n'(x) + f_n(x) \text{ a.e.}$$

$$\geq f_n(x) \text{ a.e. } \forall \text{ n.}$$

since n is arbitrary, we have

$$F'(x) \geq f(x) \text{ a.e.}$$

$$\Rightarrow \quad \int_a^b F'(x)\,dx \geq \int_a^b f(x)\,dx \qquad \ldots (1)$$

Also by the Lebesgue's theorem, i.e. "Let f be an increasing real-valued function defined on [a, b].

Then f is differentiable a.e. and the derivative f′ is measurable.

and $$\int_a^b f'(x)\,dx \leq f(b) - f(a)\text{", we have}$$

$$\int_a^b F'(x)\,dx \leq F(b) - F(a) \qquad \ldots (2)$$

Notes

But $$F(x) = \int_a^b f(t)\,dt + F(a)$$

$$\Rightarrow \qquad F(b) - F(a) = \int_a^b f(x)\,dx$$

Therefore (2) becomes

$$\int_a^b F'(x)\,dx \le \int_a^b f(x)\,dx \qquad \ldots (3)$$

From (1) and (3), we get

$$\int_a^b F'(x)\,dx = \int_a^b f(x)\,dx$$

$$\Rightarrow \qquad \int_a^b F'(x)\,dx - \int_a^b f(x)\,dx = 0$$

$$\Rightarrow \qquad \int_a^b \left[F'(x) - f(x)\right] dx = 0$$

since $F'(x) - f(x) \ge 0$ a.e., which gives that

$F'(x) - f(x) = 0$ a.e. and

so $F'(x) = f(x)$ a.e.

3.2 Summary

- If f is an integrable function on [a, b] then f is integrable on any interval [a, x] $\subset$ [a, b]. The function F given by

  $$F(x) = \int_a^x f(t)\,dt + c,$$

  where c is a constant, called the indefinite integral of F.

- Let f be an integrable on [a, b]. If $\int_a^x f(t)\,dt = 0 \;\forall\, x \in [a,b]$ then f = 0 a.e. in [a, b].

3.3 Keyword

Differentiation of an Integral: If f is an integrable function on [a, b] then f is integrable on any interval [a, x] $\subset$ [a, b]. The function F given by

$$F(x) = \int_a^x f(t)\,dt + c,$$

where c is a constant, called the indefinite integral of f.

Notes

3.4 Review Questions

1. If f is an integrable function on [a, b] and if F (x) = $\int_a^x f(t)\,dt + F(a)$ then check whether F′(x) = f (x) is absolute continuous function in [a, b] or not.

2. If F is an absolutely continuous function on [a, b], then prove that F (x) = $\int_a^x f(t)\,dt + C$ where f = F′ a.e. on [a, b] and C is constant.

3.5 Further Readings

Books

Flanders, Harley. *Differentiation under the Integral Sign*

Frederick S. Woods, *Advanced Calculus,* Ginn and Company

David V. Widder, *Advanced Calculus*, Dover Publications Inc., New Edition (Jul 1990).

Online links

www.physicsforums.com > Mathematics > Calculus & Analysis

www.sp.phy.cam.ac.uk/~ alt 36/partial diff.pdf

Notes

Unit 4: Absolute Continuity

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define Absolute Continuous function.
- Solve problems on absolute continuity
- Understand the proofs of related theorems.

Introduction

It may happen that a continuous function f is differentiable almost everywhere on [0,1], its derivative f′ is Lebesgue integrable, and nevertheless the integral of f′ differs from the increment of f. For example, this happens for the Cantor function, which means that this function is not absolutely continuous. Absolute continuity of functions is a smoothness property which is stricter than continuity and uniform continuity.

4.1 Absolute Continuity

4.1.1 Absolute Continuous Function

A real-valued function f defined on [a,b] is said to be absolutely continuous on [a,b], if for an arbitrary $\in > 0$, however small, $\exists\, a, \delta > 0$, such that

$$\sum_{r=1}^{n} \left| f(b_r) - f(a_r) \right| < \in \text{ whenever } \sum_{r=1}^{n} (b_r - a_r) < \delta,$$

where $a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n$ i.e. a_i′s and b_i′s are forming finite collection $\{(a_i, b_i) : i = 1, 2, \ldots, n\}$ of pair-wise disjoint intervals.

Obviously, every absolutely continuous function is continuous.

Notes

Notes

1. If a function satisfied $\left|\sum f(b_r) - f(a_r)\right| < \in$, even then it is absolutely continuous.

2. The condition $\sum_{r=1}^{n}(b_r - a_r) < \delta$ means that total length of all the intervals must be less than δ.

4.1.2 Theorems and Solved Examples

Theorem 1: Every absolutely continuous function f defined on [a,b] is of bounded variation.

Proof: Since f is absolutely continuous on [a,b]; for $\in = 1$, $\exists$ a $\delta > 0$ such that

$$\sum_{r=1}^{n}\left|f(b_i) - f(a_i)\right| < 1,$$

whenever $\sum_{r=1}^{n}(b_i - a_i) < \delta$,

and $a = a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n = b$.

Now consider another subdivision of [a,b] or say refinement of P by adjoining some additional points to P in such a way that all the intervals can be divided into r parts each of total length less than δ.

Let the r-sub-intervals be $[c_0, c_1], [c_1, c_2], \ldots, [c_{r-1}, c_r]$ such that

$a = c_0, c_r = b$ and $(c_{k+1} - c_k) < \delta, \forall k = 0, 1, 2, \ldots, (r-1)$

Obviously, $\sum_{i}\left|f(x_{i+1}) - f(x_i)\right| < 1$,

where $x_{i+1}, x_i \in [c_k, c_{k+1}]$

or $\overset{c_{k+1}}{\underset{c_k}{V}}(f) < 1$,

Hence $\overset{b}{\underset{a}{V}}(f) = \overset{c_1}{\underset{c_0}{V}}(f) + \overset{c_2}{\underset{c_1}{V}}(f) + \ldots + \overset{c_r}{\underset{c_{r-1}}{V}}(f) < 1 + 1 + \ldots + 1 = r =$ finite quantity.

Hence f is of bounded variation.

Notes

Notes

Converse of above theorem is not necessarily true. There exists functions of bounded variation but not absolutely continuous.

Theorem 2: Let f(x) and g(x) be absolutely continuous functions, then prove that $f(x) \pm g(x)$ and $f(x).g(x)$ are also absolutely continuous functions. Hence show that $\frac{f(x)}{g(x)}\left(\text{if}\left|g(x) > 0, \forall x\right|\right)$ is also absolutely continuous function.

Proof: Given $f(x)$ and $g(x)$ are absolutely continuous functions on the closed interval [a,b], therefore for each $\in > 0$, there exists $\delta > 0$ such that

$$\sum_{r=1}^{n} \left|f(b_r) - f(a_r)\right| < \in \text{ and}$$

$$\sum_{r=1}^{n} \left|g(b_r) - g(a_r)\right| < \in,$$

whenever $\sum_{r=1}^{n} (b_r - a_r) < \delta$, for all the points $a_1, b_1, a_2, b_2, ..., a_n, b_n$ such that

$a_1 < b_1 \le a_2 < b_2 \le ... \le a_n < b_n$.

(i) We have, $\sum_{r=1}^{n} \left|\left[f(b_r) \pm g(b_r)\right] - \left[f(a_r) \pm g(a_r)\right]\right| \le \sum_{r=1}^{n} \left|f(b_r) - f(a_r)\right| + \sum_{r=1}^{n} \left|g(b_r) - g(a_r).\right|$

Now if $\sum_{r=1}^{n} (b_r - a_r) < \delta$, then

$$\sum_{r=1}^{n} \left|f(b_r) - f(a_r)\right| < \frac{\in}{2} \text{ and} \sum_{r=1}^{n} \left|g(b_r) - g(a_r)\right| < \frac{\in}{2}.$$

$$\therefore \sum_{r=1}^{n} \left|\left[f(b_r) \pm g(b_r)\right] - \left[f(a_r) \pm g(a_r)\right]\right| < \frac{\in}{2} + \frac{\in}{2} = \in,$$

whenever $\sum_{r=1}^{n} (b_r - a_r) < \delta$.

This show that $\left[f(x) \pm g(x)\right]$ are also absolutely continuous functions over [a,b].

Notes

(ii) We have $\sum_{r=1}^{n}\left|f(b_r)g(b_r)-f(a_r)g(a_r)\right|$

$$=\sum_{r=1}^{n}\left|f(b_r)g(b_r)-f(b_r)g(a_r)+f(b_r)g(a_r)-f(a_r)g(a_r)\right|$$

$$=\sum_{r=1}^{n}\left|f(b_r)\left[g(b_r)-g(a_r)\right]+g(a_r)\left[f(b_r)-f(a_r)\right]\right|$$

$$\le\sum_{r=1}^{n}\left|f(b_r)\left[g(b_r)-g(a_r)\right]\right|+\sum_{r=1}^{n}\left|g(a_r)\left[f(b_r)-f(a_r)\right]\right|$$

$$\le\sum_{r=1}^{n}\left|f(b_r)\right|\left|g(b_r)-g(a_r)\right|+\sum_{r=1}^{n}\left|g(a_r)\right|\left|f(b_r)-f(a_r)\right|.$$

Now every absolutely continuous function is bounded therefore f(x) and g(x) are bounded in the closed interval [a,b].

Let $|f(x)|\le K_1, |g(x)|\le K_2, \forall x\in[a,b]$.

Then we have

$$\sum_{r=1}^{n}\left|f(b_r)g(b_r)-f(a_r)g(a_r)\right|\le|K_1|\in+|K_2|\in=\in\left(|K_1|+|K_2|\right),$$

Whenever $\sum_{r=1}^{n}\left|(b_r-a_r)\right|<\delta.$

Setting $\in\left(|K_1|+|K_2|\right)= \in^*,$

We have $\sum_{r=1}^{n}\left|f(b_r)g(b_r)-f(a_r)g(a_r)\right|< =\in^*,$

Whenever $\sum_{r=1}^{n}(b_r-a_r)<\delta,$

where $a_1<b_1\le a_2<b_2\le\ldots\le a_n<b_n$;

$\Rightarrow$ Product of two absolutely continuous functions is also absolutely continuous.

(iii) We have $|g(x)>0\forall x\in[a,b]|$; therefore

$$|g(x)|\ge\rho, \text{where } \rho>0, \forall x\in[a,b].$$

Notes

Now, $\sum_{r=1}^{n}\left|\frac{1}{g(b_r)}-\frac{1}{g(a_r)}\right| = \sum_{r=1}^{n}\left|\frac{g(a_r)-g(b_r)}{g(b_r)g(a_r)}\right| < \frac{\in}{\rho^2}$,

Whenever $\sum_{r=1}^{n}(b_r - a_r) < \delta$. Setting $\frac{\in}{\rho^2} = \in^*$, we get

$$\sum_{r=1}^{n}\left|\frac{1}{g(b_r)}-\frac{1}{g(a_r)}\right| < \in^*.$$

This show that $\frac{1}{g(x)}$ is absolutely continuous function over [a,b].

Now f(x), $\frac{1}{g(x)}$ are absolutely continuous.

$\Rightarrow f(x).\frac{1}{g(x)}$ is absolutely continuous.

$\Rightarrow \frac{f(x)}{g(x)}$ is also absolutely continuous over [a,b].

Hence the theorem is true.

Note

By Theorem 1, its remark and above theorem it follows that set of all absolutely continuous functions on [a,b] is a proper subspace of the space BV [a,b] of all functions of bounded variation on [a,b].

Theorem 3: If $\in BV[a,b]$, then f is absolutely continuous on [a,b], iff the variation function $v(x) = \overset{x}{\underset{a}{V}}(f)$ is absolutely continuous on [a,b].

Proof: Case I: Given v(x) is absolutely continuous.

$\Rightarrow$ For arbitrary $\in > 0, \exists\, \delta > 0$ s.t.

$\sum_{r=1}^{n}\left|v(b_r) - v(a_r)\right| < \in$, whenever $\sum_{r=1}^{n}(b_r - a_r) < \delta$.

Also, we know that $\left|f(x) - f(a)\right| \le \overset{x}{\underset{a}{V}}(f) = v(x)$

Notes

$$\Rightarrow \sum_{r=1}^{n} \left| f(b_r) - f(a_r) \right| = \sum_{r=1}^{n} \left| f(b_r) - f(a) + f(a) - f(a_r) \right|$$

$$= \sum_{r=1}^{n} \left[\left| f(b_r) - f(a) \right| + \left| f(a) - f(a_r) \right| \right]$$

Now taking supremum over all collections of P_i of $[a_i, b_i]$ for i = 2,...,n, we get

$$\sum_{r=1}^{n} \overset{b_i}{\underset{a_i}{V}}(f) < \in .$$

But $\overset{b_i}{\underset{a}{V}}(f) = \overset{a_i}{\underset{a}{V}}(f) + \overset{b_i}{\underset{a_i}{V}}(f)$

$$\Rightarrow \overset{b_i}{\underset{a_i}{V}}(f) = \overset{b_i}{\underset{a}{V}}(f) + \overset{a_i}{\underset{a}{V}}(f)$$

$$\Rightarrow \overset{b_i}{\underset{a_i}{V}}(f) = v(b_i) - v(a_i)$$

$$\Rightarrow \sum_{i=1}^{n} \left| v(b_i) - v(a_i) \right| < \in$$

$\Rightarrow v(x)$ is absolutely continuous.

Theorem 4: A necessary and sufficient condition that a function should be an indefinite integral is that it should be absolutely continuous.

Proof: Condition is sufficient.

Let f(x) be an absolutely continuous function over the closed interval [a,b].

Therefore f is of bounded variation and hence we can express f(x) as

$$f(x) = f_1(x) - f_2(x)$$

where $f_1(x)$ and $f_2(x)$ are monotonically increasing functions and hence both are differentiable.

$$\leq \sum_{r=1}^{n} \left| v(b_r) - v(a_r) \right| < \in \text{ whenever} \sum_{r=1}^{n} (b_r - a_r) < \delta$$

$\Rightarrow f$ is also absolutely continuous on [a,b].

Case II: Given f is absolutely continuous on [a,b].

$\Rightarrow$ for a given $\in > 0, \exists$ a $\delta > 0$ s.t.

$$\sum_{i=1}^{n} \left| f(b_i) - f(a_i) \right| < \in, \qquad \text{...(i)}$$

Notes

for every finite collection $P\{]a_i, b_i[, i = 1,2,...,n\}$ of pairwise disjoint sub-intervals of [a,b] such that $\sum_{i=1}^{n}(b_i - a_i) < \delta.$

Now, let $P_i\{]x_{k-1}^i, bx_{ki}^i[, k = 1,...,m_i\}$ be a finite collection of non-overlapping intervals of the interval $[a_i, b_i]$.

Then the collection $\{]x_{k-1}^i, bx_{ki}^i[: i = 1,2,...,n, k = 1,...,m_i\}$ is a finite collection of non-overlapping sub-intervals of [a,b] such that

$$\sum_{i=1}^{n}\left[\sum_{k=1}^{m_i}(x_k^i - x_{k-1}^i)\right] = \sum_{i=1}^{n}(b_i - a_i) < \delta.$$

and hence by (i), $\sum_{i=1}^{n}\sum_{k=1}^{m_i}\left|f(x_k^i) - f(x_{k-1}^i)\right| = \in.$

Hence f′(x) exists and $|f'(x)| \le f_1'(x) + f_2'(x)$

$\Rightarrow \int_a^b |f'(x)| \le f_1(b) + f_2(b) - f_1(a) - f_2(a) < \infty,$

$\Rightarrow f'(x)$ is integrable also.

Now let F(x) be an definite integral of f′(x) i.e.

$$F(x) = F(a) + \int_a^x f'(t)dt,\ x \in [a,b] \qquad ...(ii)$$

Using fundamental theorem of integral calculus,

We get

F′(x) = f′(x)

or F(x) = f(x) + constant (say c) ...(iii)

From (ii), we have F(a) = f(a),

Using this in (iii), we get c = 0 and hence F (x) = f(x).

Thus every absolutely continuous function f(x) is an indefinite integral of its own derivative.

Condition is necessary: Let f(x) be an indefinite integral of f(x) defined on the closed interval [a,b], so that

$F(x) = \int_a^x f(t)dt + f(a), \forall x \in [a,b]$ and f(x) is integrable over [a,b].

Corresponding to arbitrary small $\in > 0$, let δ >0 be such that if $m(A) < \delta$, then $\int_A |f| < \in,$

Notes

Now select 2n real numbers such that

$a_i < b_i \le a_2 < b_2 \le a_3 < ... \le a_n < b_n$

such that A = $\bigcup_{i=1}^{n}[a_i, b_i]$ and $\sum_{i=1}^{n}(b_i - a_i) < \delta.$

Then $\sum_{i=1}^{n}|F(b_i) - F(a_i)| = \sum_{i=1}^{n}\left|\int_a^{b_i} f - \int_a^{a_i} f\right|$

$$= \sum_{i=1}^{n}\left|\int_{a_i}^{b_i} F\right| \le \sum_{i=1}^{n}\left|\int_{a_i}^{b_i} f\right| = \int_A |f| < \in.$$

Thus, we have shown that for arbitrary small $\in > 0, \exists$ a $\delta > 0$ s.t. $\sum_{i=1}^{n}(b_i - a_i) < \delta.$

$$\Rightarrow \sum_{i=1}^{n}|F(b_i) - F(a_i)| < \in.$$

$\Rightarrow$ F is absolutely continuous.

Thus every indefinite integral is absolutely continuous.

Theorem 5: If a function f is absolutely continuous in an interval [a,b] and if f′(x) = 0. a.e. in [a,b], then f is constant.

Proof: Let $c \in [a, b]$ be arbitrary. If we show that f(c) = f(a), then the theorem will be proved.

Let E = $\{x \in]a, c[: f'(x) = 0\}$.

since c is arbitrary, therefore set $E \subset]a, c[$. This implies any $x \in E \Rightarrow f'(x) = 0.$

Let $\in$, η > 0 arbitrary. Now $f'(x) = 0, \forall x \in E \Rightarrow \exists$ an arbitrary small interval $[x, x+h] \subset [a, c]$

such that $\frac{|f(x+h) - f(x)|}{h} < \eta \Rightarrow |f(x+h) - f(x)| < \eta h.$

This implies that corresponding to every $x \in E, \exists$ an arbitrary small closed interval $[x, x+h]$ contained in [a,c] s.t.

$$|f(x+h) - f(x)| < \eta h.$$

Thus the interval $[x, x+h], \forall x \in E,$ over E in Vitali's sense. Thus by Vitali's Lemma, we can determine a finite number of non-overlapping intervals I_k, where

$$I_k = [x_k, y_k] \forall\ k = 1, 2, 3, ..., n$$

such that this collection covers all of E except for a set of measure less than $\delta > 0$ where δ is pre-assigned number which corresponds to $\in$ occurring in the definition of absolute continuity of f.

Notes

Suppose $x_k < x_{k+1}$; then adjoining the points y_0, x_{n+1}.

We have $a = y_0 \le x_1 < y_1 \le x_2 < y_2 \le \ldots \le x_n < y_n \le x_{n+1} = c$.

Now since f is absolutely continuous, therefore for above subdivision of [a,c], we have

$$\sum_{k=0}^{n} \left| f(x_{k+1}) - f(y_k) \right| < \in, \text{ whenever} \sum_{k=0}^{n} (x_{k+1} - y_k) < \delta.$$

(i) $$\Rightarrow \sum_{k=1}^{n} \left| f(y_k) - f(x_k) \right| \le n \sum_{k=1}^{n} (y_k - x_k) < n(c-a).$$

Now $$\left| f(c) - f(a) \right| = \left| \sum_{k=0}^{n} \left[f(x_{k+1}) - f(y_k) \right] + \sum_{k=1}^{n} \left[f(y_k) - f(x_k) \right] \right|$$

$$\le \sum_{k=0}^{n} \left| f(x_{k+1}) - f(y_k) \right| + \le \sum_{k=1}^{n} \left| f(y_k) - f(x_k) \right|$$

$$< \in + n(c-a)$$

But $\in, n$ and hence $\in + n(c-a)$ are arbitrary small positive numbers. So letting $\in \to 0, n \to 0$

We get f(c) = f(a)

$\Rightarrow f(x)$ is a constant function.

Corollary: If the derivatives of two absolutely continuous functions are equivalent, then the functions differ by a constant.

Proof: Let f and g be two absolutely continuous functions and f′ = g′ $\Rightarrow (f-g)' = 0 \Rightarrow$ by above theorem f – g = constant and hence the result.

Example: If f is an absolutely continuous monotone function on [a,b] and E a set of measure zero, then show that f (E) has measure zero.

Proof: Let the function f be monotonically increasing. By the definition of absolute continuity of f, for $\in > 0, \exists\, \delta > 0$ and non-overlapping intervals $\{I_n = [a_n, b_n]\}$ such that

$$\sum (b_n - a_n) < \delta \Rightarrow \sum \left| f(b_n) - f(a_n) \right| < \in$$

or $$\sum \left[f(b_n) - f(a_n) \right] < \in$$

Now, $$E \subseteq [a,b] \Rightarrow E \subseteq \cup I_n$$

$$\Rightarrow \quad f(E) \subseteq f(\cup I_n) = \cup f(I_n)$$

$$\Rightarrow \quad m^*(f(E)) \le \sum m^*(f(I_n)) \le \sum \left[\bar{f}(x_n) - \underline{f}(x_n) \right] < \in,$$

Notes

where $\bar{f}(x_n)$ and $\underline{f}(x_n)$ are the maximum and maximum values of f(x) in the interval $[a_n, b_n]$.

Also note that $\sum |\bar{x}_n - \underline{x}_n| \le \sum (b_n - a_n) < \delta$

$\Rightarrow m^*(f(E)) \le \in, \in$ being arbitrary.

$\Rightarrow m^*(f(E)) = 0 \Rightarrow m(f(E)) = 0.$

Example: Give an example which is continuous but not absolutely continuous.

Solution: Consider the function $f : F \to R$, where F is the Cantor's ternary set.

Let $x \in F \Rightarrow x = x_1 x_2 x_3 ... = \sum_{K=1}^{\infty} \frac{x_k}{3^k}, x_k = 0 \text{ or } 2$

Define $f(x) = \sum_{K=1}^{\infty} \frac{r_k}{2^k}$, where $r_k = \frac{1}{2} x_k$.

$= 0.\, r_1, r_2, r_3.....$

This function is continuous but not absolutely continuous.

(i) Note that this function is constant on each interval contained in the complement of the Cantor's ternary set.

For, let (a,b) be one of the countable open intervals contained in F^c. Then in ternary notation,

$a = 0.a_1 a_2 ... a_{n-1}\, 0\, 2\, 2\, 2$

and $b = 0.a_1 a_2 ... a_{n-1}\, 2\, 0\, 0\, 0$,

where $a_i = 0$ or 2, for $i \le n-1$.

$\Rightarrow f(a) = 0.r_1, r_2, ..., r_{n-1}\ 0\ 1\ 1\ 1\ 1\ ...,$ where $r_i = \left(\frac{a_i}{2}\right)$,

$f(b) = 0.r_1, r_2, ..., r_{n-1}\ 1\ 0\ 0\ 0\ 0\ ...$

But in binary notation

$0.r_1, r_2, ..., r_{n-1}\ 0\ 1\ 1\ 1\ 1\ ... = 0.r_1, r_2, ..., r_{n-1}\ 1\ 0\ 0\ 0\ 0\ ...$

$\Rightarrow f(a) = f(b).$

Thus, we extend the function f overall of the set [0,1] instead of F by defining $f(x) = f(b), \forall x \in (a,b) \subset F^c$. Thus, the Cantor's function is defined over [0,1] and maps it onto [0,1].

It is clearly a non-decreasing function.

Notes

(ii) To show that f(x) is a continuous function. Note that if $c', c'' \in F$, then we have

$$\left.\begin{array}{l} c' = 0.(2p_1)(2p_2)(2p_3)\ldots \\ c'' = 0.(2q_1)(2q_2)(2q_3)\ldots \end{array}\right\} \text{each } p_i, q_i = 0 \text{ or } =1$$

If $|c' - c''| < \left(\frac{1}{3^n}\right)$, then $p_i = q_i$, for $1 \le i \le n+1$ and hence

$$|f(c') - f(c'')| < \left(\frac{1}{2^n}\right) \qquad \ldots(i)$$

$\Rightarrow$ as $n \to \infty, c' \to c'', f(c') \to f(c'')$,

Hence if $c_0 \in F$ and $\langle c_n \rangle$ is a sequence in F such that $c_n \to c_0$, when $n \to \infty$, then $f(c_n) \to f(c_0)$, when $n \to \infty$.

Now let $x_0 \in [0,1]$ and let $\langle x_n \rangle$ be a sequence in [0,1] such that $x_n \to x_0$ as $n \to \infty$.

Case I: Let $x_0 \notin F \Rightarrow x_0 \in I$, say $(a,b) \subset F^c$

$$\Rightarrow x_n \in I \text{ and hence } f(x_n) = f(x) = f(a)$$

and hence $f(x_n) \to f(x_0)$ as $n \to \infty$.

Case II: Let $x_0 \in F$. Now for each n such that $x_n \in F$, set $x_n = c_n$ and hence $f(x_n) \to f(x_0)$.

If $x_n \notin F$, then $\exists$ an open interval $I \supseteq F^c$.

(i) if $x_n < x_0$, then set c_n as the upper end point of I.

(ii) If $x_0 < x_n$, then set c_n as the lower end point of I.

$\Rightarrow$ in any case $f(x_n) \to f(x_0)$ as $n \to \infty$.

But the sequence $\langle x_n \rangle$ was any sequence satisfying the stated conditions.

$\Rightarrow$ f is a continuous function.

(iii) To show f(x) is not absolutely continuous. Note that f′(x) = 0 at each $x \in F^c$.

$\Rightarrow f'(x)$ exists and is zero on [0,1] and is summable on [0,1].

We know that for f(x) to be absolutely continuous, we must have

$$f(x) = \int_0^x f'(x)dx + f(0).$$

Particularly, we must have

$$f(1) - f(0) = \int_0^1 f'(x)dx.$$

Notes

But f(1) – f(0) = 1 and $\int_0^x f'(1)dx = 0$ as $f'(x) = 0$

$\Rightarrow f(1) = f(0) \neq \int_0^x f'(1)dx = 0$ as $f'(x) = 0$

$\Rightarrow f(x)$ is not absolutely continuous.

Theorem 6: Prove that an absolutely continuous function on [a, b] is an indefinite integral.

Proof: Let f(x) be an absolutely continuous function in a closed interval [a, b] so that f′(x) &

$\int_a^x f'(t)\,dt$ exists finitely $\forall\, x \in$ [a, b].

Let F(x) be an indefinite integral of f′(x), so that

$$F(x) = f(a) + \int_a^x f'(t)\,dt,\ x \in [a, b] \qquad \ldots (1)$$

We shall prove that F(x) = f(x).

Since an indefinite integral is an absolutely continuous function.

Therefore F(x) is absolutely continuous in [a, b].

Then from (1),

$$F'(x) = f'(x) \text{ a.e.}$$

$$\Rightarrow \qquad \frac{d}{dx}[F(x) - f(x)] = 0.$$

Integrating, we get

$$F(x) - f(x) = c \text{ (constant)} \qquad \ldots (2)$$

Taking x = a in (1), we get

$$F(a) = f(a) + \int_a^a f'(t)\,dt$$

$$\Rightarrow \qquad F(a) - f(a) = 0$$

or $\qquad F(x) - f(x) = 0$ for x = a

Then from (2), we get c = 0.

Thus (2) reduces to

$$F(x) - f(x) = 0 \text{ a.e.}$$

$$\Rightarrow \qquad F(x) = f(x) \text{ a.e.}$$

which shows that f(x) is indefinite integral of its own derivative.

Notes

4.2 Summary

- A real-valued function f defined on [a,b] is said to be absolutely continuous on [a,b], if for an arbitrary $\in > 0$, however small, $\exists\, a, \delta > 0$, such that

$$\sum_{r=1}^{n} \left| f(b_r) - f(a_r) \right| < \in, \text{ whenever } \sum_{r=1}^{n} (b_r - a_r) < \delta.$$

- Every absolutely continuous function is continuous.
- Every absolutely continuous function f defined on [a,b] is of bounded variation.

4.3 Keywords

Absolute Continuity of Functions: Absolute continuity of functions is a smoothness property which is stricter than continuity and uniform continuity.

Absolute Continuous Function: A real-valued function f defined on [a,b] is said to be absolutely continuous on [a,b], if for an arbitrary $\in > 0$, however small, $\exists\, a, \delta > 0$, such that

$$\sum_{r=1}^{n} \left| f(b_r) - f(a_r) \right| < \in \text{ whenever } \sum_{r=1}^{n} (b_r - a_r) < \delta,$$

where $a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n$ i.e. a_i's and b_i's are forming finite collection $\{(a_i, b_i) : i = 1, 2, \ldots, n\}$ of pair-wise disjoint intervals.

4.4 Review Questions

1. Define absolute continuity for a real variable. Show that f(x) is an indefinite integral, if F is absolutely continuous.
2. If f,g: [0,1] $\to R$ are absolutely continuous, prove that f + g and fg are also absolutely continuous.
3. Show that the set of all absolutely continuous functions on an interval I is a linear space.
4. If g is a non-decreasing absolutely continuous function on [a,b] and f is absolutely continuous on [g(a), g(b)], show that fog is also absolutely continuous on [a,b].
5. If f is absolutely continuous on [a,b] and $f'(x) \ge 0$ for almost all $x \in [a, b]$, show that f is non-decreasing on [a,b].

4.5 Further Readings

Books

Krishna B Athreya, N Soumendra Lahiri, *Measure Theory and Probability Theory*, Springer (2006).

Notes

Aole Nielsen, *An Introduction to Integration and Measure Theory*, Wiley-Interscience, (1997).

N.L. Royden, *Real Analysis* (third ed.), Collier MacMillan, (1988).

Online links

dl.acm.org

mrich.maths.org

Notes

Unit 5: Spaces, Hölder

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand L^p-spaces, conjugate numbers and norm of an element of L^p-space
- Understand the proof of Hölder's inequality.

Introduction

In this unit, we discuss an important construction, which is extremely useful in virtually all branches of analysis. We shall study about L^p-spaces and Hölder's inequality.

5.1 Spaces, Hölder

5.1.1 L^p-Spaces

The class of all measurable functions f (x) is known as L^p-spaces over [a, b], if Lebesgue – integrable over [a, b] for each p exists, $0 < p < \infty$, i.e.

$$\int_a^b |f|^p \, dx < \infty, (p > 0)$$

and is denoted by L^p [a, b].

Notes

Note The symbol L^p is used for such classes when limits of integration are known and mentioning of interval is not necessary.

5.1.2 Conjugate Numbers

Let p, q be any two n on-negative extended real numbers s.t. $\frac{1}{p}+\frac{1}{q}=1$, then p, q are called (mutually) conjugate numbers.

Obviously, 2 is self-conjugate number.

Also if $p \neq 2$, then $q \neq 2$. Further, if $p = \infty$, then $q = 1 \Rightarrow 1, \infty$ are conjugate numbers.

Note Non-negativity $\Rightarrow p \geq 1, q \geq 1$.

5.1.3 Norm of an Element of L^p-space

The p-norm of any $f \in L^p [a, b]$, denoted by $\| f \|_p$, is defined as

$$\| f \|_p = \left[\int_a^b |f|^p\right]^{\frac{1}{p}}, 0 < p < \infty.$$

Theorem 1: If $f \in L^p [a, b]$ and $g \leq f$, then $g \in L^p [a, b]$.

Proof: Let α be any positive real number.

$\therefore \quad \{x \in [a, b] : g(x) > \alpha\} = \{x \in [a, b] : \alpha < g(x) \leq f(x)\} \quad (\because g \leq f)$

$= \{x \in [a, b] : f(x) > \alpha\}$

Again $f \in L^p [a, b]$

$\Rightarrow$ f is measurable over [a, b].

$\Rightarrow$ $\{x \in [a, b] : f(x) > \alpha\}$ is a measurable set.

$\Rightarrow$ $\{x \in [a, b] : g(x) > \alpha\}$ is a measurable set.

$\Rightarrow$ g is a measurable function over [a, b]

Again since $g(x) \leq f(x), \forall x \in [a, b]$

$$\Rightarrow \int_a^b |g|^p dx \leq \int_a^b |f|^p dx < \infty \quad (\because |f|^p \in L[a, b])$$

or $$\int_a^b |g|^p dx < \infty$$

Thus $|g|^p \in L[a, b]$.

Notes

Thus we have proved that g is a measurable function over [a, b] such that

$$|g|^p \in L[a, b]$$

Hence $\quad g \in L^p[a, b]$

Theorem 2: If $f \in L^p[a, b]$, $p > 1$, then $f \in L[a, b]$

Proof: $f \in L^p[a, b] \Rightarrow f$ is measurable over [a, b]

Let $\quad A_1 = \{x \in [a, b] : |f(x) \geq 1\}$

and $\quad A_2 = \{x \in [a, b] : |f(x) < 1\}$

Then $\quad [a, b] = A_1 \cup A_2$ and $A_1 \cap A_2 = \phi$

Using countable additive property of the integrals, we have

$$\int_a^b |f|\,dx = \int_{A_1} |f|\,dx + \int_{A_2} |f|\,dx \qquad \ldots \text{(i)}$$

Now $\because \quad |f(x)| \geq 1, x \in A_1$

$\therefore \quad |f| \leq |f|^p$ on A_1 as $p > 1$

$$\therefore \quad \int_{A_1} |f|\,dx \leq \int_{A_1} |f|^p\,dx < \infty \text{ as } f \in L^p[a, b] \qquad \ldots \text{(ii)}$$

Now $|f(x)| < |, \forall\ x \in A_2$

Using first mean value theorem, we get

$$\int_{A_2} |f|\,dx < m(A_2) = \text{A finite quantity} \qquad \ldots \text{(iii)}$$

Combining (ii) and (iii) and making use of (i), we get

$$\int_a^b |f|\,dx < \infty$$

Thus f is a measurable function over [a, b], such that

$$\int_a^b |f|\,dx < \infty$$

$\Rightarrow \quad |f| \subset L[a, b]$ and hence $f \in L[a, b]$.

Theorem 3: If $f \in L^p[a, b]$, $g \in L^p[a, b]$; then $f + g \in L^p[a, b]$

Proof: Since $f, g \in L^p[a, b] \Rightarrow f, g$ are measurable over [a, b]

$\Rightarrow f + g$ is measurable over [a, b]

Let $\quad A_1 = \{x \in [a, b] : |f(x)| \geq |g(x)|\}$

and $\quad A_2 = \{x \in [a, b] : |f(x)| < |g(x)|\}$

Then $[a, b] = A_1 \cup A_2$ and $A_1 \cap A_2 = \phi$

Notes

Therefore $\int |f+g|^p \, dx = \int_{A_1} |f+g|^p + \int_{A_2} |f+g|^p \, dx$.

Again, $|f + g|^p \le (|f| + |g|)^p \le (|g| + |g|)^p$ on A_2 and $\le (|f| + |f|)^p$ on A_1

$\le 2P \, |g|^p$ on A_2 and $\le 2^p \, |f|^p$ on A_1

Integrating, we have

$$\int_{A_1} |f+g|^p \le 2^p \int_{A_1} |f|^p$$

and

$$\int_{A_2} |f+g|^p \le 2^p \int_{A_2} |g|^p$$

Since f, g $\in L^p$ [a, b] $\Rightarrow \int_{A_1} |f|^p < \infty$ and $\int_{A_2} |g|^p < \infty$

$\because$ $\int_{A_1} |f+g|^p < \infty$ and $\int_{A_1} |f+g|^p < \infty$ [by (i) and (ii)]

$\Rightarrow$ $\int_a^b |f+g|^p \, dx < \infty \Rightarrow f + g \in L^p$ [a, b]

5.1.4 Simple Version of Hölder's Inequality

Lemma 1: Let p, q > 1 be such that $\frac{1}{p} + \frac{1}{q} = 1$, and let u and v be two non-negative numbers, at least one being non-zero. Then the function f : [0, 1] $\to$ R defined by

$$f(t) = ut + v(1-t^q)^{\frac{1}{q}}, \; t \in [0, 1],$$

has a unique maximum point at

$$s = \left[\frac{u^p}{u^p + v^p}\right]^{\frac{1}{q}} \qquad \ldots (1)$$

The maximum value of f is

$$\max_{t \in [0,1]} f(t) = (u^p + v^p)^{\frac{1}{p}} \qquad \ldots (2)$$

Proof: If v = 0, then f (t) = tu, $\forall$ t $\in$ [0, 1] (with u > 0), and in this case, the Lemma is trivial. Likewise, if u = 0, then

$$f(t) = v(1-t^q)^{\frac{1}{q}}, \; \forall \; t \in [0, 1] \text{ (with } v > 0), \qquad \ldots (3)$$

Notes

and using the inequality

$$(1-t^q)^{\frac{1}{q}} < 1, \forall\, t \in (0, 1],$$

We immediately get f (t) < f (0), $\forall$ t $\in$ (0, 1],

and the Lemma again follows.

For the remainder of the proof we are going to assume that u, v > 0.

Obviously f is differentiable on (0, 1) and the solutions of the equation (3)

$$f'(t) = 0$$

Let s be defined as in (1), so under the assumption that u, v > 0, we clearly have 0 < s < 1.

We are going to prove first that s is the unique solution in (0, 1) of the equation (3).

We have

$$f'(t) = u + v \cdot \frac{1}{q}(1-t^q)^{\frac{1}{q}-1} \cdot q \cdot t^{q-1} \qquad \ldots (4)$$

$$= u - v\left(\frac{t^q}{1-t^q}\right), \; t \in (0, 1)$$

so the equation (3) reads

$$u - v\left(\frac{t^q}{1-t^q}\right) = 0.$$

Equivalently, we have

$$\left(\frac{t^q}{1-t^q}\right)^{\frac{1}{p}} = u/v,$$

$$\Rightarrow \qquad \frac{t^q}{1-t^q} = (u/v)^p,$$

$$t^2 = \frac{(u/v)^p}{1-(u/v)^p} = \frac{u^p}{u^p + v^p},$$

Having shown that the "candidates" for the maximum point are 0, 1 and s let us show that s is the only maximum point.

For this purpose, we go back to (4) and we observe that f′ is also continuous on (0, 1).

Since $\lim_{t \to 0^+} f'(t) = u > 0$ and

$$\lim_{t \to 1^-} f'(t) = -\infty$$

and the equation (3) has exactly one solution in (0, 1), namely s, this forces

$$f'(t) > 0 \;\; \forall \;\; t \in (0, s)$$

Notes

and $f'(t) < 0 \ \forall \ t \in (s, 1)$.

This means that, f is increasing on [0, s] and decreasing on [s, 1], and we are done.

The maximum value of f is then given by

$$\max_{t \in [0,1]} f(t) = f(s),$$

and the fact that f (s) equals the value in (2) follows from an easy computation.

5.1.5 Hölder's Inequality

Statement: Let $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ be non-negative numbers. Let p, q > 1 be real number with the property $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\sum_{j=1}^{n} a_j b_j \le \left(\sum_{j=1}^{n} a_j^p \right)^{\frac{1}{p}} \cdot \left(\sum_{j=1}^{n} b_j^q \right)^{\frac{1}{q}} \qquad \ldots (5)$$

Moreover, one has equality only when the sequences $\left(a_1^p, \ldots, a_n^p\right)$ and $\left(b^q, \ldots, b_n^q\right)$ are proportional.

Proof: The proof will be carried on by induction on n. The case n = 1 is trivial.

Case n = 2.

Assume $(b_1, b_2) \neq (0, 0)$. (otherwise everything is trivial).

Define the number

$$r = \frac{b_1}{\left(b_1^q + b_2^q\right)^{\frac{1}{q}}}.$$

Notice that $r \in [0, 1]$ and we have

$$\frac{b_2}{\left(b_1^q + b_2^q\right)^{\frac{1}{q}}} = \left(1 - r^q\right)^{\frac{1}{q}}$$

Notice also that, upon dividing by $\left(b_1^q + b_2^q\right)^{\frac{1}{q}}$, the desired inequality

$$a_1 b_1 + a_2 b_2 \le \left(a_1^q + a_2^q\right)^{\frac{1}{p}} \left(b_1^q + b_2^q\right)^{\frac{1}{q}} \qquad \ldots (6)$$

reads

$$a_1 r + a_2 \left(1 - r^q\right)^{\frac{1}{q}} \le \left(a_1^q + a_2^q\right)^{\frac{1}{p}} \qquad \ldots (7)$$

It is obvious that this is an equality when $a_1 = a_2 = 0$. Assume $(a_1, a_2) \neq (0, 0)$, and set up the function.

$$f(t) = a_1 t + a_2 \left(1 - t^q\right)^{\frac{1}{q}}, \ t \in [0, 1].$$

Notes

We now apply Lemma (1) stated above, which immediately gives us (7).

Let us examine when equality holds.

If $a_1 = a_2 = 0$, the equality obviously holds, and in this case (a_1, a_2) is clearly proportional to (b_1, b_2). Assume $(a_1, a_2) \neq (0, 0)$.

Again by Lemma (1), we know that equality holds in (7), exactly when

$$r = \left[\frac{a_1^p}{a_1^p + a_2^p}\right]^{\frac{1}{q}}$$

that is
$$\frac{b_1}{\left(b_1^q + b_2^q\right)^{\frac{1}{q}}} = \left[\frac{a_1^p}{a_1^p + a_2^p}\right]^{\frac{1}{q}},$$

or equivalently

$$\frac{b_1^q}{b_1^q + b_2^q} = \frac{a_1^p}{a_1^p + a_2^p}.$$

Obviously this forces

$$\frac{b_1^q}{b_1^q + b_2^q} = \frac{a_1^q}{a_1^p + a_2^p},$$

so indeed $\left(a_1^p, a_2^p\right)$ and $\left(b_1^q, b_2^q\right)$ are proportional.

Having proven the case n = 2, we now proceed with the proof of:

The implication: Case n = k ⇒ case n = k + 1, start with two sequences $(a_1, a_2, \ldots, a_k, a_{k+1})$ and $(b_1, b_2, \ldots, a_k, b_{k+1})$.

Define the numbers

$$a = \left(\sum_{j=1}^{k} a_j^p\right)^{\frac{1}{p}} \text{ and } b = \left(\sum_{j=1}^{k} b_j^q\right)^{\frac{1}{q}}.$$

Using the assumption that the case n = k holds, we have

$$\sum_{j=1}^{k+1} a_j b_j \leq \left(\sum_{j=1}^{k} a_j^p\right)^{\frac{1}{p}} \cdot \left(\sum_{j=1}^{k} b_j^q\right)^{\frac{1}{q}} + a_{k+1}\, b_{k+1}$$

$$= ab + a_{k+1}\, b_{k+1} \qquad \ldots (8)$$

Using the case n = 2, we also have

$$ab + a_{k+1}\, b_{k+1} \leq \left(a^p + a_{k+1}^p\right)^{\frac{1}{p}} \cdot \left(b^q + b_{k+1}^q\right)^{\frac{1}{q}}$$

$$= \left(\sum_{j=1}^{k+1} a_j^p\right)^{\frac{1}{p}} \cdot \left(\sum_{j=1}^{k+1} b_j^q\right)^{\frac{1}{q}}, \qquad \ldots (9)$$

so combining with (8) we see that the desired inequality (5) holds for n = k + 1.

Notes

Assume now we have equality. Then we must have equality in both (8) and in (9).

On one hand, the equality in (8) forces $\left(a_1^p, a_2^p, \ldots, a_k^p\right)$ and $\left(b_1^q, b_2^q, \ldots, b_k^q\right)$ to be proportional (since we assume the case n = k). On the other hand, the equality in (9) forces $\left(a^p, a_{k+1}^p\right)$ and $\left(b^q, b_{k+1}^q\right)$ to be proportional (by the case m = 2). Since

$$a^p = \sum_{j=1}^{k} a_j^p \text{ and } b^q = \sum_{j=1}^{k} b_j^q ,$$

it is clear that $\left(a_1^p, a_2^p, \ldots, a_k^p, a_{k+1}^p\right)$ and $\left(b_1^q, b_2^q, \ldots, b_k^q, b_{k+1}^q\right)$ are proportional.

5.1.6 Riesz-Hölder's Inequality

Statement: Let p and q be conjugate indices or exponents (numbers) and $f \in L^p[a, b]$, $g \in L^q[a, b]$; then show that

(i) $f \cdot g \in L[a, b]$

(ii) $\|fg\| \le \|f\|_p \|g\|_q$ i.e.

$$\int |fg| \le \left(\int |f|^p\right)^{\frac{1}{p}} \left(\int |g|^q\right)^{\frac{1}{q}}$$

with equality only when $\alpha\ |f|^p = \beta\ |g|^q$ a.e. for some non-zero constants α and β.

Lemma: If A and B are any two non-negative real numbers and $0 < \lambda < 1$, then

$A^\lambda B1^{-\lambda} \le \lambda A + (1 - \lambda) B$, with equality when A = B.

Proof: If either A = 0 or B = 0, then the result is trivial.

Let A > 0, B > 0

Consider the function

$$\phi(x) = x^\lambda - \lambda x, \text{ where } 0 \le x < \infty \text{ and } 0 < \lambda < 1$$

$$\Rightarrow \quad \frac{d\phi}{dx} = \lambda x^{\lambda-1} - \lambda \text{ and } \frac{d^2\phi}{dx^2} = \lambda(\lambda - 1)x^{\lambda-1} .$$

Now solving $\frac{d\phi}{dx} = 0$, we get x = 1.

Also at x = 1, $\frac{d^2\phi}{dx^2} < 0$ as $0 < \lambda < 1$.

By calculus, $\phi(x)$ is maximum at x = 1, so

$$\phi(x) \le \phi(1) \text{ i.e. } x^\lambda - \lambda x \le 1^\lambda - \lambda. \qquad \ldots (1)$$

Now, putting $x = \frac{A}{B}$, we get

$$\left(\frac{A}{B}\right)^\lambda - \lambda\left(\frac{A}{B}\right) \le 1 - \lambda \text{ or } A^\lambda B^{-\lambda} - \lambda\frac{A}{B} \le 1 - \lambda$$

Notes

Or $$A^{\lambda} B^{1-\lambda} - \lambda A \le (1-\lambda) B \text{ or}$$

$$A^{\lambda} B^{1-\lambda} \le \lambda A + B(1-\lambda) \qquad \dots (2)$$

Obviously equality holds good only for x = 1, i.e. only when A = B.

Proof of Theorem

Note that when p = 1, q = ∞, the proof of theorem is obvious. Let us assume that 1 < p < ∞ and 1 < q < ∞.

Now set $$\lambda = \frac{1}{p}; p > 1 \quad \Rightarrow \quad \lambda < 1$$

Therefore $$\frac{1}{q} = 1 - \lambda \qquad \left(\because \frac{1}{p} + \frac{1}{q} = 1\right)$$

Putting these values of λ and 1 – λ in (2), we get

$$A^{\frac{1}{p}} B^{\frac{1}{q}} \le \frac{A}{p} + \frac{B}{q} \qquad \dots (3)$$

If one of the functions f (x) and g (x) is zero a.e. then the theorem is trivial. Thus, we assume that f ≠ 0, g ≠ 0 a.e. and hence the integrals

$$\int_a^b |f|^p \, dx \quad \text{and} \quad \int_a^b |g|^q \, dx$$

are strictly positive and hence $\| f \|_p > 0$, $\| g \|_q > 0$.

Set $$f(x) = \frac{f(x)}{\|f\|_p}, \quad g(x) = \frac{g(x)}{\|g\|_q}$$

and $$A^{\frac{1}{p}} = |f(x)|, \; B^{\frac{1}{q}} = |g(x)|.$$

Then (3) gives

$$|f(x)\, g(x)| \le \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}$$

Integrating, we get

$$\int_a^b |f(x)\, g(x)|\, dx \le \frac{1}{p}\int_a^b |f(x)|^p \, dx + \frac{1}{q}\int_a^b |g(x)|^p \, dx$$

$$= \frac{1}{p}\int_a^b \frac{|f(x)|^p}{\int_a^b |f|^p \, dx} dx + \frac{1}{q}\int_a^b \frac{|g(x)|^q}{\int_a^b |g|^q \, dx} dx$$

Notes

$$= \frac{1}{p}\frac{\int_a^b |f|^p\,dx}{\int_a^b |f|^p\,dx} + \frac{1}{q}\frac{\int_a^b |g|^q\,dx}{\int_a^b |g|^q\,dx}$$

$$= \frac{1}{p} + \frac{1}{q} = 1\,.$$

Hence $\int_a^b |f(x)\,g(x)|\,dx \le 1.$

Putting the values of f (x) and g (x), we get

$$\int_a^b \frac{|f(x)\,g(x)|\,dx}{\|f\|_p\,\|g\|_q} \le 1 \text{ or } \|fg\| \le \|f\|_p\,\|g\|_q \qquad \ldots (4)$$

Now $f \in L^p\,[a, b]$, $g \in L^q\,[a, b]$

$$\Rightarrow \quad \int_a^b |f|^p\,dx < \infty \text{ and } \int_a^b |g|^q\,dx < \infty$$

$$\Rightarrow \quad \|f\|_p < \infty \text{ and } \|g\|_q < \infty$$

Therefore, from (4), we have

$$\|fg\|_1 < \infty \Rightarrow f_g \in L'\,[a, b]$$

Also the equality will hold when A = B

i.e. $|f(x)|^p = |g(x)|^q$, a.e.

i.e. if $\frac{|f|^p}{\|f\|_p^p} = \frac{|g|^q}{\|g\|_q^q}$, a.e.

or if $\|g\|_q^q\,|f|^p = \|f\|_p^p\,|g|^q$, a.e.

or if we have got some non-zero constants α, β

$$\alpha\,|f|^p = \beta\,|g|^q, \text{ a.e.}$$

Hence the theorem.

5.1.7 Riesz-Hölder's Inequality for 0 < p < 1

If $0 < p < 1$ and p and q are conjugate exponents, and $f \in L^p$ and $g \in L^q$, then

$$\int |fg| \ge \|f\|_p\,\|g\|_q, \text{ provided } \int |g|^q \ne 0\,.$$

(In this case, the inequality is reversed than that of the case for $1 \le p < \infty$.)

Proof: Conjugacy of p, q $\Rightarrow \frac{1}{p} + \frac{1}{q} = 1$

Notes

$$\Rightarrow \frac{1}{q} = 1 - \frac{1}{p}$$

$$\Rightarrow \frac{-p}{q} = 1 - p$$

$$\Rightarrow p + \left(-\frac{p}{q}\right) = 1$$

If we take $p = \frac{1}{P}$ and $\frac{-p}{P} = \frac{1}{Q}$, then $\frac{1}{p} + \frac{1}{Q} = 1$ and since $0 < p < 1 \Rightarrow \frac{1}{0} < \frac{1}{P} < 1 \Rightarrow P > 1$,

i.e. $1 < P < \infty < \infty$ and also $\frac{1}{Q} = \frac{-p}{q} = 1 - p \Rightarrow 0 < \frac{1}{Q} < 1$ as $0 < p < 1 \Rightarrow Q > 1$.

P, Q are conjugate numbers with $1 < P < \infty$.

If we take $|fg| = F^P$ and $|g|^q = G^Q$.

Then
$$fg = |f_g|^{\frac{1}{P}} \cdot |g|^{\frac{q}{Q}} = |f|^P |g|^{\left(\frac{1}{P}\right)+\left(\frac{q}{Q}\right)\left(\frac{-p}{q}\right)}$$
$$= |f|^P.$$

f, g are non-negative measurable functions s.t.

Also $f \in L^P$ and $g \in L^q$.

Applying the Hölder's inequality for P, Q to the functions f and g, we get

$$\int |FG| \le \|F\|_P \, \|G\|_Q$$

$$\Rightarrow \int |f|^P \le \left(\int |F|^P\right)^{\frac{1}{P}} \left(\int |G|^Q\right)^{\frac{1}{Q}} \text{ as } |fg| = fg = |f|^P$$

$$\Rightarrow \int |f|^P \le \left(\int |fg|\right)^P \left(\int |g|^q\right)^{\frac{-p}{q}}$$

$$\Rightarrow \left(\int |f|^P\right)^{\frac{1}{P}} \le \left(\int |fg|\right)\left(\int |g|^q\right)^{\frac{-1}{q}}$$

$$\Rightarrow \left(\int |f|^P\right)^{\frac{1}{P}} \le \frac{\int |fg|}{\left(\int |g|^q\right)^{\frac{1}{q}}}, \text{ provided } \int |g|^q \ne 0$$

$$\Rightarrow \int |fg| \ge \left(\int |f|^P\right)^{\frac{1}{P}} \left(\int |q|^q\right)^{\frac{1}{q}} \ge \|f\|_P \, \|g\|_q$$

Notes

Theorem 4: SCHWARZ or CAUCHY-SCHWARZ INEQUALITY statement: Let f and g be square integrable, i.e.

$$f, g \in L^2[a, b]; \text{ then } fg \le L\,[a, b] \text{ and } \| fg \| \le \| f \|_2 \, \| g \|_2.$$

Proof: Let $x \in [a, b]$ be arbitrary, then

$$[\,|f(x)| - |g(x)|\,]^2 \ge 0$$

or $\quad 2\,|f(x)|\,.\,|g(x)| \le |f(x)^2 + |g(x)|^2.$

On integrating, we get

$$2\int_a^b |f(x)g(x)|\,dx \le \int_a^b |f(x)^2 dx + \int_a^b |g(x)|^2\,dx \qquad \ldots \text{(i)}$$

Now $f, g \in L^2[a, b] \Rightarrow$ f and g are measurable over [a, b] and $\int_a^b |f(x)|^2\,dx < \infty, \int_a^b |g(x)|^2\,dx < \infty$.

Using in (i), we get $\int_a^b |f(x)\,g(x)|\,dx < \infty$

Thus $fg \in L\,[a, b]$.

Let $a \in R$ be arbitrary. Then

$$(\alpha\,|f| + |g|)^2 \ge 0$$

$$\therefore \qquad \int_a^b (\alpha\,|f| + |g|)^2 \ge 0$$

or $\alpha^2 \int_a^b |f|^2\,dx + 2\alpha \int_a^b |fg|\,dx + \int_a^b |g|^2\,dx \ge 0$

Write $A = \int_a^b |f|^2\,dx,\ B = 2\int_a^b |fg|\,dx,\ C = \int_a^b |g|^2\,dx$

Then we have $\alpha^2 A + \alpha B + C \ge 0 \qquad \ldots \text{(ii)}$

Now, if A = 0, then f(x) = 0 a.e. in [a, b] and hence B = 0 and both sides of the inequality to be proved are zero. Thus when A = 0, the inequality is trivial.

Again, let $A \ne 0$. Writing $\alpha = -\dfrac{B}{2A}$ in (ii), we get

$$A\left(-\frac{B}{2A}\right)^2 + B\left(-\frac{B}{2A}\right) + C \ge 0\,.$$

which gives $B^2 \le 4AC$.

Notes

Now putting the values of A, B, C in last inequality, we have

$$4\left[\int_a^b |fg|\,dx\right]^2 \le 4\left[\int_a^b |f|^2\right]\left[\int_a^b |g|^2\right]$$

or
$$\int_a^b |f(x)\,(g(x)|\,dx \le \left(\int_a^b |f(x)|^2\right)^{1/2}\left(\int_a^b |g(x)|^2\right)^{1/2}$$

or
$$\|fg\| \le \|f\|_2\ \|g\|_2.$$

Note: The above theorem is a particular case of Hölder's inequality.

Example: Let f, g be square integrable in the Lebesgue sense then prove f + g is also square integrable in the Lebesgue sense, and $\|f+g\|_2 \le \|f\|_2 + \|g\|_2$.

Solution: By hypothesis $f^2 \in L\,[a, b]$, $g^2 \in L\,[a, b]$.

$$f^2, g^2 \in L\,[a, b] \Rightarrow fg \in L\,[a, b]. \qquad \text{[by Schwarz inequality]}$$

Again
$$(f+g)^2 = f^2 + g^2 + 2fg \in L\,[a, b].$$

Hence (f + g) is square integrable, again, we have

$$\int_a^b (f+g)^2 = \int_a^b f^2 + \int_a^b g^2 + 2\int_a^b fg$$

$$\le \int_a^b f^2 + \int_a^b g^2 + 2\left(\int_a^b f^2\right)^{1/2}\left(\int_a^b g^2\right)^{1/2} \qquad \text{(by Schwarz inequality)}$$

$$= \left[\left(\int_a^b f^2\right)^{1/2} + \left(\int_a^b g^2\right)^{1/2}\right]^2$$

$\therefore$
$$\left(\int_a^b (f+g)^2\right)^{1/2} \le \left(\int_a^b f^2\right)^{1/2} + \left(\int_a^b g^2\right)^{1/2}$$

or
$$\|f+g\|_2 \le \|f\|_2 + \|g\|_2.$$

Example: Prove that $\|f+g\|_1 \le \|f\|_1 + \|g\|_1$.

Solution: We know that $|f+g| \le |f| + |g|$.

Integrating both the sides.

$$\int |f+g| \le \int |f| + \int |g|$$

$\Rightarrow$
$$\|f+g\|_1 \le \|f\|_1 + \|g\|_1.$$

Notes

5.2 Summary

- The class of all measurable functions f (x) is known as L^p – space over [a, b], if Lebesgue-integrable over [a, b] for each p exists, $0 < p < \infty$, i.e.

$$\int_a^b |f|^p \, dx < \infty, \ (p > 0)$$

- The p-norm of any $f \in L^p$ [a, b], denoted by $\| f \|_p$, is defined as

$$\| f \|_p = \left[\int_a^b |f|^p \right]^{\frac{1}{p}}, \ 0 < p < \infty$$

- Let p, q > 1 be such that $\frac{1}{p} + \frac{1}{q} = 1$, and let u and v be two non-negative numbers, at least one being non-zero. Then the function f : [0, 1] → R defined by

$$f(t) = ut + v\left(1 - t^q\right)^{\frac{1}{q}}, \ t \in [0, 1],$$

has a unique maximum point at

$$s = \left[\frac{u^p}{u^p + v^p} \right]^{\frac{1}{q}}$$

- Let p and q be conjugate indices or exponents and $f \in L^p$ [a, b], $g \in L^q$ [a, b], then it is evident that

  (i) $f, g \in L$ [a, b]

  (ii) $\| fg \| \le \| f \|_p \ \| g \|_q$ i.e.

$$\int |fg| \le \left(\int |f|^p \right)^{\frac{1}{p}} \left(\int |g|^q \right)^{\frac{1}{q}}$$

5.3 Keywords

Conjugate Numbers: Let p, q be any two n on-negative extended real numbers s.t. $\frac{1}{p} + \frac{1}{q} = 1$, then p, q are called (mutually) conjugate numbers.

Hölder's Inequality: Let $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ be non-negative numbers. Let p, q > 1 be real number with the property $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\sum_{j=1}^{n} a_j b_j \le \left(\sum_{j=1}^{n} a_j^p \right)^{\frac{1}{p}} \cdot \left(\sum_{j=1}^{n} b_j^q \right)^{\frac{1}{q}}$$

Moreover, one has equality only when the sequences $\left(a_1^p, \ldots, a_n^p\right)$ and $\left(b^q, \ldots, b_n^q\right)$ are proportional.

Notes

L^p-Spaces: The class of all measurable functions f (x) is known as L^p-spaces over [a, b], if Lebesgue – integrable over [a, b] for each p exists, $0 < p < \infty$, i.e.

$$\int_a^b |f|^p \, dx < \infty, (p > 0)$$

and is denoted by L^p [a, b].

p-norm: The p-norm of any $f \in L^p$ [a, b], denoted by $\| f \|_p$, is defined as

$$\| f \|_p = \left[\int_a^b |f|^p \right]^{\frac{1}{p}}, 0 < p < \infty.$$

5.4 Review Questions

1. If f and g are non-negative measurable functions, then show that in Hölder's inequality, equality occurs iff $\exists$ some constants s and t (not both zero) such that $sf^p + tg^q = 0$.
2. State and prove Hölder's Inequality.

5.5 Further Readings

Books

G.H. Hardy, J.E. Littlewood, G. Polya, *Inequalities*, Cambridge University Press, (1934)

L.P. Kuptsov, *Hölder inequality*, Springer (2001)

Kenneth Kuttler, *An Introduction of Linear Algebra*, BRIGHAM Young University, 2007

Online links

www.m–hiKari.com

www.math.Ksu.edu

www.tandfonline.com

Notes

Unit 6: Minkowski Inequalities

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define L^p-space, conjugate numbers and norm of an element of L^p-space.
- Understand Minkowski inequality.
- Solve problems on Minkowski inequality.

Introduction

In mathematical analysis, the Minkowski inequality establishes that the L^p spaces are normed vector spaces. Let S be a measure space, let $1 \le p \le \infty$ and let f and g be elements of L^p (s). Then f + g is in L^p (s), we have the triangle inequality

$$\| f + g \|_p \le \| f \|_p + \| g \|_p$$

with equality for $1 < p < \infty$ if and only if f and g are positively linearly dependent, i.e. $f = \lambda_g$ for some $\lambda \ge 0$. In this unit, we shall study Minkowski's inequality for $1 \le p < \infty$ and for $0 < p < 1$. We shall also study almost Minkowski's inequality in integral form.

6.1 Minkowski Inequalities

Here, the norm is given by:

$$\| f \|_p = \left(\int | f |^p \, d\mu \right)^{1/p}$$

if $p < \infty$, or in the case $p = \infty$ by the essential supremum

$$\| f \|_\infty = \text{ess sup}_{x \in S} \, | f (x) | .$$

The Minkowski inequality is the triangle inequality in $L^p(S)$. In fact, it is a special case of the more general fact

Notes

$$\| f \|_p = \sup_{\|g\|_q=1} \int |fg| \, d\mu \, , \quad 1/p + 1/q = 1$$

where it is easy to see that the right-hand side satisfies the triangular inequality.

Like Holder's inequality, the Minkowski inequality can be specialized to sequences and vectors by using the counting measure:

$$\left(\sum_{k=1}^{n} | x_k + y_k |^p \right)^{1/p} \le \left(\sum_{k=1}^{n} | x_k |^p \right)^{1/p} + \left(\sum_{k=1}^{n} | y_k |^p \right)^{1/p}$$

for all real (or complex) numbers $x_1, ..., x_n, y_1, ..., y_n$ and where n is the cardinality of S (the number of elements in S).

Thus, we may conclude that

If p > 1, then Minkowski's integral inequality states that

$$\left| \int_a^b | f(x) + g(x) |^p \, dx \right|^{1/p} \le \left| \int_a^b | f(x) |^p \, dx \right|^{1/p} + \left| \int_a^b | g(x) |^p \, dx \right|^{1/p}$$

Similarly, if p >1 and $a_k, b_k > 0$, then Minkowski's sum inequality states that

$$\left| \sum_{k=1}^{n} | a_k + b_k |^p \right|^{1/p} \le \left(\sum_{k=1}^{n} | a_k |^p \right)^{1/p} + \left(\sum_{k=1}^{n} | b_k |^p \right)^{1/p}$$

Equality holds iff the sequences $a_1, a_2, ...$ and $b_1, b_2, ...$ are proportional.

6.1.1 Proof of Minkowski Inequality Theorems

Theorem 1: State and prove Minkowski inequality. If f and $g \in L^p$ $(1 \le p < \infty)$, then $f + g \in L^p$ and $\| f + g \|_p \le \| f \|_p + \| g \|_p$.

or

Let $1 \le p \le \infty$. Prove that for every pair $f, g \in L^p \{0, 1\}$, the function $f + g \in L^p \{0, 1\}$ and that $\| f + g \|_p \le \| f \|_p + \| g \|_p$. When does equality occur?

Or

Suppose $1 \le p < \infty$. Prove that for any two functions f and g in L^p [a, b]

$$\left(\int_a^b | f + g |^p \right)^{\frac{1}{p}} \le \left(\int_a^b | f |^p \, dx \right)^{\frac{1}{p}} + \left(\int_a^b | g |^p \, dx \right)^{\frac{1}{p}}$$

Proof: When p = 1, the desired result is obvious.

If $p = \infty$, then

$$| f | \le \| f \|_\infty \text{ a.e.}$$

$$| g | \le \| g \|_\infty \text{ a.e.}$$

Notes

$$\Rightarrow \qquad |f+g| \le |f| + |g|$$

$$\le \|f\|_\infty + \|g\|_\infty \text{ a.e.}$$

$$\Rightarrow \qquad \|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$$

Hence the result follows in this case also. Thus, we now assume that $1 < p < \infty$.

Since L^p is a linear space, $f + g \in L^p$.

Let q be conjugate to p, then $\frac{1}{p}+\frac{1}{q}=1$.

Now $\quad (f + g) \in L^p$

$\Rightarrow \quad (f + g)^{p/q} \in L^q$

Since $\frac{1}{p}+\frac{1}{q}=1 \Rightarrow \frac{1}{q}=1-\frac{1}{p} \Rightarrow \frac{1}{q}=\frac{p-1}{p}$

$$\Rightarrow (p-1)^q = p, \int \left(|f+g|^{p-1}\right)^q = \int |f+g|^p$$

and therefore $|f + g|^{p-1} \in L^p \Rightarrow (f + g)^{p/q} \in L^p$ because $p - 1 = \frac{p}{q}$.

On applying Hölder's inequality for f and $(f + g)^{p/q}$, we get

$$\int |f|\cdot|f+g|^{p/q}\,dx \le \left(\int |f|^p)\,dx\right)^{1/p}\left(\int |f+g|^{p/q\cdot q}\,dx\right)^{1/p}$$

or

$$\int |f|\cdot|f+g|^{p/q}\,dx \le \left(\int |f|^p)\,dx\right)^{1/p}\left(\int |f+g|^p\,dx\right)^{1/q} \qquad \ldots (1)$$

Since $g \in L^p$, therefore interchanging f and g in (1), we get

$$\int |g|\cdot|f+g|^{p/q}\,dx \le \left(\int |g|^p)\,dx\right)^{1/p}\left(\int |f+g|^p\,dx\right)^{1/q} \qquad \ldots (2)$$

Adding, we get

$$\int |f|\cdot|f+g|^{p/q}\,dx + \int |g|\cdot|f+g|^{p/q}\,dx \le \left[\left(\int |f|^p\,dx\right)^{1/p} + \left(\int |g|^p\,dx\right)^{1/p}\right]\cdot\left(\int |f+g|^p\,dx\right)^{1/q} \qquad \ldots (3)$$

Now $\qquad |f + g|^p = |f+g|\cdot|f+g|^{p-1}$

But $\qquad \frac{1}{p}+\frac{1}{q}=1 \Rightarrow 1+\frac{p}{q}=p \Rightarrow p-1=\frac{p}{q}$

$\therefore \qquad |f + g|^p = |f+g|\cdot|f+g|^{p/q}$

$$\le (|f|+|g|)\cdot|f+g|^{p/q}$$

Notes

or $$|f+g|^p \le |f|\cdot|f+g|^{p/q} + |g|\cdot|f+g|^{p/q}$$

Integrating, we get

$$\int |f+g|^p\,dx \le \int |f|\cdot|f+g|^{p/q} + |g|\cdot|f+g|^{p/q}\,dx \qquad \ldots (4)$$

Using (3), relation (4) becomes

$$\int |f+g|^p\,dx \le \left[\left(\int |f|^p\,dx\right)^{1/p} + \left(\int |g|^p\,dx\right)^{1/p}\right]\left(|f+g|^p\,dx\right)^{1/q}$$

Dividing each term by $\int \left(|f+g|^p\,dx\right)^{1/q}$, we get

$$\left(\int |f+g|^p\,dx\right)^{1-1/q} \le \left(\int |f|^p\,dx\right)^{1/p} + \left(\int |g|^p\,dx\right)^{1/p}$$

But $$\frac{1}{p}+\frac{1}{q}=1 \Rightarrow 1-\frac{1}{q}=\frac{1}{p}$$

So $$\left(\int |f+g|^p\,dx\right)^{1/p} \le \left(\int |f|^p\,dx\right)^{1/p} + \left(\int |g|^p\,dx\right)^{1/p}$$

or $$\|f+g\|_p \le \|f\|_p + \|g\|_p$$

Hence the proof.

Note Equality hold in Minkowski's inequality if and only if one of the functions f and g is a multiple of the other.

Theorem 2: Minkowski's inequality for $0 < p < 1$. If $0 < p < 1$ and f, g are non-negative functions in L^p, then

$$\|f+g\|_p \ge \|f\|_p + \|g\|_p.$$

Proof: For this proceed as in theorem Minkowski's inequality and applying the Hölder's inequality for $0 < p < 1$ for the functions $f \in L^p$ and $(f+g)^{p/q} \in L^q$, we get

$$\int |f||f+g|^{p/q} \ge \left(\int |f|^p\right)^{1/p}\left(\int |f+g|^{p/q})^q\right)^{1/q}$$

$$\Rightarrow \quad \int |f||f+g|^{p/q} \ge \left(\int |f|^p\right)^{1/p}\left(\int |f+g|^p\right)^{1/q} \qquad \ldots (i)$$

Also $g \in L^p$, proceeding as above, we get

$$\int |g||f+g|^{p/q} \ge \left(\int |g|^p\,dx\right)^{1/p}\left(\int |f+g|^p\right)^{1/q} \qquad \ldots (ii)$$

Notes

Adding these two,

$$\int |f+g|^{p/q}\,(|f|+|g|) \;\geq\; \left[\left\{|f|^p\right\}^{\frac{1}{p}} + \left\{|g|^p\right\}^{\frac{1}{p}}\right] \times \left\{|f+g|^p\right\}^{\frac{1}{q}} \qquad \ldots \text{(iii)}$$

Also $\qquad \dfrac{1}{p}+\dfrac{1}{q}=1 \;\Rightarrow\; 1+\dfrac{p}{q}=p$

$$\Rightarrow \qquad |f+g|^p = |f+g|\cdot|f+g|^{p/q} = (|f|+|g|)\cdot|f+g|^{p/q}, \text{ as } f\geq 0,\, g\geq 0$$

$$\Rightarrow \qquad \int |f+g|^p = \left[\left\{|f|^p\right\}^{1/p} + \left\{|g|^p\right\}^{1/p}\right]\left(\int |f+g|^p\right)^{1/q}$$

Dividing by $\left(\int |f+g|^p\right)^{1/q}$, we get

$$\left(\int |f+g|^p\right)^{1-(1/q)} \geq \left[\|f\|_p + \|g\|_p\right]$$

$$\Rightarrow \qquad \left(\int |f+g|^p\right)^{1/q} \geq \|f\|_p + \|g\|_p$$

$$\Rightarrow \qquad \|f+g\|_p \geq \|f\|_p + \|g\|_p$$

6.1.2 Minkowski Inequality in Integral Form

Statement: Suppose $f : \mathcal{R}\times\mathcal{R} \to \mathcal{R}$ is Lebesgue measurable and $1 \leq p < \infty$. Then

$$\left(\int\left|\int h(x,y)\,dy\right|dx\right)^{1/p} \;\leq\; \int\left(\left|\int h(x,y)^p\,dx\right|\right)^{1/p} dy$$

Proof: By an approximation argument we need only consider h of the form

$$h(x,y) = \sum_{j=1}^{N} f_j(x)1\,f_j(y),\ (x,y)\in\mathcal{R}\times\mathcal{R}\ ,$$

where N is a positive integer, f_j is Lebesgue measurable, and $F_j \in L_n$, $j = 1, \ldots N_i$ and $F_i \cap F_j = \phi$ if $1 \leq i < j \leq N$. We use Minkowski's inequality to estimate

$$\left(\int\left|\int h(x,y)\,dy\right|dx\right)^{1/p} = \left(\left|\sum_{j=1}^{N}\|F_j\|\,f_j(x)\right|_{dx}^{p}\right)^{1/p} \leq \sum_{j=1}^{N}\|F_j\|\left(\int |f_i(x)^p dx\right)^{1/p}$$

But

$$\int\left(\left|\int h(x,y)\right|^p dx\right)^{1/p} dy = \sum_{j=1}^{N}\int_{F_i}\left(\int |h(x,y)|^p\,dx\right)^{1/p} = \sum_{j=1}^{N}\int_{F_j}\left(\int |f_j(x)|^p\,dx\right)^{1/p}$$

Notes

Example: If $<f_n>$ is a sequence of functions belonging to $L^2(a, b)$ and also $f \in L^2(a, b)$ and $\text{Lim} \| f_n - f \|_2 = 0$, then prove that

$$\int_a^b f^2 dx = \text{Lim} \int_a^b f_n^2 dx$$

Solution: By Minkowski's inequality, we get

$$\left| \| f_n \|_2 - \| f_2 \| \right| \le \| f_n - f \|_2$$

$$\Rightarrow \quad \text{Lim} \left| \| f_n \|_2 - \| f \|_2 \right| \le \text{Lim} \| f_n - f \|_2 = 0$$

$$\Rightarrow \quad \text{Lim} \left| \| f_n \|_2 - \| f \|_2 \right| = 0 \Rightarrow \text{Lim} \| f_n \|_2 = \| f \|_2$$

$$\Rightarrow \quad \text{Lim} \left(\int_a^b (f_n)^2 dx \right)^{1/2} = \left(\int_a^b f^2 dx \right)^{1/2} \Rightarrow \text{Lim} \int_a^b f_n^2 \, dx = \int_a^b f^2 dx .$$

6.2 Summary

- The class of all measurable function f (x) is known as L^p space over [a, b], if Lebesgue integrable over [a, b] for each p exists, $0 < p < \infty$.
- If f and $g \in L^p$ $(1 \le p \le \infty)$, then $f + g \in L^p$ and $\| f + g \|_p \le \| f \|_p + \| g \|_p$.

6.3 Keywords

L^p-space: The class of all measurable functions f (x) is known as L^p-space over [a, b], if Lebesgue-integrable over [a, b] for each exists, $0 < p < \infty$, i.e.,

$$\int_a^b | f |^p \, dx < \infty, (p > 0)$$

and is denoted by L^p [a, b].

Minkowski Inequality in Integral Form: Suppose $f : \mathcal{R} \times \mathcal{R} \to \mathcal{R}$ is Lebesgue measurable and $1 \le p < \infty$. Then

$$\left(\int \left| \int h(x, y) \, dy \right| dx \right)^{1/p} \le \int \left(\left| \int h(x, y)^p \, dx \right| \right)^{1/p} dy$$

Minkowski Inequality: Minkowski inequality establishes that the L^p spaces are normed vector spaces. Let S be a measure space, let $1 \le p \le \infty$ and let f and g be elements of L^p (s). Then f + g is in L^p (s), we have the triangle inequality

$$\| f + g \|_p \le \| f \|_p + \| g \|_p$$

with equality for $1 < p < \infty$ if and only if f and g are positively linearly dependent.

Notes

6.4 Review Questions

1. If f, g are square integrable in the Lebesgue sense, prove that f + g is also square integrable and

$$\| f + g \|_2 \leq \| f \|_2 + \| g \|_2.$$

2. If | < p < ∞, then show that equality can be true, iff there are non-negative constants α and β, such that βf = αg.

6.5 Further Readings

Books

Books: Stein, Elias (1970). *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press.

Hardy, G.H.; Littlewood, J.E.; Polya, G. (1952). *Inequalities,* Cambridge Mathematical Library (second ed.). Cambridge: Cambridge University Press.

Online links

Mathworld.wolfram.com>Calculus and Analysis>Inequalities

Planet math.org/Minkowski In-equality.html

Notes

Unit 7: Convergence and Completeness

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand convergence and completeness.
- Understand Riesz-Fischer theorem.
- Solve problems on convergence and completeness.

Introduction

Convergence of a sequence of functions can be defined in various ways, and there are situations in which each of these definitions is natural and useful. In this unit, we shall start with the definition of convergence and Cauchy sequence and proceed with the topic completeness of L^p.

7.1 Convergence and Completeness

7.1.1 Convergent Sequence

Definition: A sequence $<x_n>$ in a normal linear space X with norm $\| \cdot \|$ is said to converge to an element $x \in X$ if for arbitrary $\in > 0$, however small, $\exists\, n_0 \in N$ such that $\| x_n - x \| < \in$, $\forall\ n > n_0$.

Then we write $\lim_{n\to\infty} x_n = x$.

Notes

7.1.2 Cauchy Sequence

Definition: A sequence $<x_n>$ in a normal linear space $(X, \|\cdot\|)$ is said to be a Cauchy sequence if for arbitrary $\in > 0$, $\exists\, n_0 \in N$ s.t.

$$\| x_n - x_m \| < \in, \ \forall \ n, m \geq n_0.$$

7.1.3 Complete Normed Linear Space

Definition: A normed linear space $(X, \|\cdot\|)$ is said to be complete if every Cauchy sequence $<x_n>$ in it converges to an element $x \in X$.

7.1.4 Banach Space

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

7.1.5 Summable Series

Definition: A series $\sum_{n=1}^{\infty} u_n$ in N_1 is said to be summable to a sum u if $u \in N_1$ and $\lim_{n\to\infty} S_n = u$, where

$$S_n = u_1 + u_2 + \ldots + u_n$$

In this case, we write

$$u = \sum_{n=1}^{\infty} u_n .$$

Further, the series $\sum_{n=1}^{\infty} u_n$ is said to be absolutely summable if $\sum_{n=1}^{\infty} \|u_n\| < \infty$.

7.1.6 Riesz-Fischer Theorem

Theorem: The normed L^p-spaces are complete for $(p \geq 1)$.

Proof: In order to prove the theorem, we shall show that every Cauchy sequence in L^p [a, b] space converges to some element f in L^p-space. Let $<f_n>$ be one of such sequences in L^p-space. Then for given $\in > 0$, $\exists$ a natural number n_0, such that

$$m, n \geq n_0 \Rightarrow \| f_m - f_n \|_p < \in,$$

since $\in$ is arbitrary therefore taking $\in = \frac{1}{2}$, we can find a natural number n_1 such that

$$\text{for all } m, n \geq n_1 \Rightarrow \| f_m - f_n \|_p < \frac{1}{2}$$

Similarly, taking $\in = \frac{1}{2^k}$, $\forall\, k \in N$, we can find a natural number n_k, such that

$$\text{for all } m, n \geq n_k \Rightarrow \| f_m - f_n \|_p < \frac{1}{2^k}$$

Notes

In particular, $m > n_k \Rightarrow \| f_m - f_{n_k} \| < \frac{1}{2^k}$.

Obviously $n_1 < n_2 < n_3 \ldots < n_k < \ldots$

i.e. $<n_k>$ is a monotonic increasing sequence of natural numbers.

Set $g_k = f_{n_k}$, then from above, we have

$$\| g_2 - g_1 \|_p = \left\| f_{n_2} - f_{n_2} \right\|_p < \frac{1}{2},$$

$$\| g_3 - g_2 \|_p = \left\| f_{n_3} - f_{n_2} \right\|_p < \frac{1}{2^2},$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$\| g_{k+1} - g_k \|_p = \left\| f_{n_{k+1}} - f_{n_k} \right\|_p < \frac{1}{2^k}.$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

Adding these inequalities, we get

$$\sum_{k=1}^{\infty} \left\| g_{k+1} - g_k \right\|_p < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1 \qquad \ldots \text{(i)}$$

Thus $\sum_{k=1}^{\infty} \left\| g_{k+1} - g_k \right\|_p$ is convergent. Define g such that

$$g(x) = \left| g_1(x) \right| + \sum_{k=1}^{\infty} \left| g_{k+1} - g_k \right|_p \text{ if R.H.S. is convergent} \qquad \ldots \text{(ii)}$$

and $g(x) = \infty$, if right hand side is divergent.

Now, $$\left(\int_a^b |g(x)|^p \, dx \right) = \lim_{n \to \infty} \left\{ \int_a^b |g_1(x) + \sum_{k=1}^{n} \left| g_{k+1} - g_k \right|^p_{dx} \right\}^{\frac{1}{p}}$$

or $$\| g \|_p = \lim_{n \to \infty} \left(\left\| g_1 \right\|_p + \sum_{k=1}^{n} \left\| g_{k+1} - g_k \right\|_p \right) \text{ (By Minkowski's inequality)}$$

$$= \| g_1 \|_p + \sum_{k=1}^{\infty} \left\| g_{k+1} - g_k \right\|_p < \| g_1 \|_p + 1, \qquad \text{[by (i)]}$$

$\Rightarrow$ $\| g \|_p < \infty \Rightarrow g \in L^p [a, b]$.

Let $E = \{x \in [a, b] : g(x) = \infty\}$.

Now we define a function f such that **Notes**

$$f(x) = 0, \ \forall \ x \in E$$

and $$f(x) = g_1(x) + \sum_{k=1}^{\infty}\left(g_{k+1} - g_k\right), \text{ for } x \in [a, b] \text{ but } x \notin E,$$

or $$f(x) = \lim_{m\to\infty}\left(g_1 + \sum_{k=1}^{m-1}\left(g_{k+1} - g_k\right)\right), \text{ for } x \notin E$$

$$= \lim_{m\to\infty} g_m(x)$$

Thus $$f(x) = 0, \text{ for } x \in E \text{ and}$$

$$f(x) = \lim_{m\to\infty} g_m(x) \text{ for } x \notin E.$$

$\therefore$ $$f(x) = \lim_{m\to\infty} g_m(x) \text{ a.e. in } [a, b]$$

or $$\lim_{m\to\infty}\left|g_m - f\right| = 0 \text{ a.e. in } [a, b] \qquad \ldots \text{(iii)}$$

Also, $$g_m(x) = g_1 + \sum_{k=1}^{m-1}\left(g_{k+1} - g_k\right)$$

$\Rightarrow$ $$|g_m| \le |g_1| + \sum_{k=1}^{m-1}\left(g_{k+1} - g_k\right)$$

$$\le |g_1| + \sum_{k=1}^{\infty}\left(g_{k+1} - g_k\right) = g,$$

$\Rightarrow$ $$|g_m| \le g, \ \forall \ m \in N$$

$\Rightarrow$ $$\lim_{m\to\infty}\left|g_m(x)\right| \le g$$

$\therefore$ $$\text{(iii)} \Rightarrow |f| \le g.$$

Again, $$|g_m - f| \le |g_m| + |f| \le g + g = 2g.$$

$\therefore$ $$|g_m - f| \le 2\,g.$$

Thus there exists a function $g \in L^p[a, b]$ s.t.

$$|g_m - f| \le 2g, \ \forall \ m$$

and $$\lim_{m\to\infty}\left|g_m - f\right| = 0 \text{ a.e. in } [a, b] \qquad \ldots \text{(iv)}$$

Applying Lebesgue dominated convergence theorem,

$$\lim_{m\to\infty}\int_a^b \left|g_m - f\right|^p dx = \int_a^b \lim_{m\to\infty}\left|g_m - f\right|^p dx = \int_a^b 0 \cdot dx = 0 \qquad \text{[Using (iv)]}$$

Notes

$$\Rightarrow \qquad \left(\lim_{m\to\infty}\int_a^b |g_m - f|^p\, dx\right)^{\frac{1}{p}} = 0$$

$$\Rightarrow \qquad \lim_{m\to\infty} \| g_m - f \|_p = 0$$

$$\Rightarrow \qquad \lim_{m\to\infty} \| f_{n_m} - f \|_p = 0 \text{ as } g_m = f_{n_m}$$

$$\Rightarrow \qquad \| f_{n_m} - f \|_p < \in'.$$

Also $$\| f_m - f_{n_m} \|_p < \in.$$

$$\therefore \qquad \| f_m - f \|_p = \| f_m - f_{n_m} + f_{n_m} - f \|_p$$

$$\leq \| f_m - f_{n_m} \|_p + \| f_{n_m} - f \|_p$$

$$< (\in + \in') = \in''.$$

Hence $$\lim_{m\to\infty} \| f_m - f \|_p = 0$$

or $$\lim_{m\to\infty} f_m = f \in L^p\,[a, b].$$

This proves the theorem.

Alternative Statement of this Theorem

A convergent sequence $<f_n>$ in L^p-spaces has a limit in L^p-space.

Or

Every Cauchy sequence $<f_n>$ in the L^p-space converges to a function in L^p-space.

Theorem: Prove that a normed linear space is complete iff every absolutely summable sequence is summable.

Proof: *Necessary part*

Let X be a complete normed linear space with norm $\| \cdot \|$ and $<f_n>$ be an absolutely summable sequence of elements of X

$$\Rightarrow \qquad \sum_{n=1}^{\infty} \| f_n \| = M < \infty,$$

$\Rightarrow$ For arbitrary $\in > 0$, however small, $\exists n_o \in N$

s.t. $$\sum_{n=n_o}^{\infty} \| f_n \| < \in, \qquad \ldots \text{(i)}$$

Now, if $S_n = \sum_{i=1}^{n} f_i$, then $\forall\ n \geq m \geq n_o$, we get

Notes

$$\| S_n - S_m \|_p = \left\| \sum_{i=m+1}^{n} f_i \right\| \le \sum_{i=m+1}^{n} \| f_i \|$$

$$\le \sum_{i=n_0}^{\infty} \| f_i \| < \in .$$

$\Rightarrow$ Sequence $<S_n>$ of partial sums is a Cauchy sequence

$\Rightarrow$ $<S_n>$ converges.

$\Rightarrow$ Sequence $<f_n>$ is summable to some element $S \in X$.

But X is a complete space. Therefore $<S_n>$ will converge to some element $S \in X$.

Sufficient part: Given that every absolutely summable sequence in the space X is summable.

To show that X is complete.

Let $<f_n>$ be a Cauchy sequence in X.

For each positive integer k, we can choose a number $n_k \in N$ such that

$$\| f_n - f_m \| < \frac{1}{2^k}, \ \forall \ n, m \ge n_k \qquad \ldots \text{(ii)}$$

We can choose these n_k's such that $n_{k+1} > n_k$.

Then $\left\langle f_{n_{k=1}} \right\rangle^{\infty}$ is a subsequence of $<f_n>$.

Setting $g_1 = f_{n_1}$ and $g_k = f_{n_k} - f_{n_{k-1}}$, (k > 1), we get a sequence $<g_k>$ s.t. its kth partial.

Sum = $S_k = g_1 + g_2 + \ldots + g_k = f_{n_1} + (f_{n_2} - f_{n_1}) + \ldots + (f_{n_k} - f_{n_{k-1}}) = f_{n_k}$.

Now, $\| g_k \| = \left\| f_{n_k} - f_{n_{k-1}} \right\| < \frac{1}{2^{k-1}}$, [by (ii)], $\forall$ k > 1

$$\Rightarrow \quad \sum_{k=1}^{\infty} \| g_k \| \le \| g_1 \| + \sum_{k=2}^{\infty} \frac{1}{2^{k-1}} = \| g_1 \| + 1 \ \text{(a finite quantity)}$$

$\Rightarrow$ The sequence $<g_k>$ is absolutely summable and hence by the hypothesis, it is a summable sequence.

$\Rightarrow$ The sequence of partial sums of this sequence converges to some $S \in X$.

$\Rightarrow$ The sequence $<S_k>$ converges and hence $\left\langle f_{n_k} \right\rangle$ converges to some $f \in X$.

Now, we shall show that the limit $f_n = f$.

Again, since $<f_n>$ is a Cauchy sequence, we get that for each $\in > 0$, however small, $\exists n' \in N$ s.t. $\forall$ n, m > n′.

$$\| f_n - f_m \| < \frac{\varepsilon}{2}.$$

Also since $f_{n_k} \to f, \exists n'' \in N$ such that $\forall k > n''$,

Notes

$$\left\| f_{n_k} - f \right\| < \frac{\varepsilon}{2}.$$

Choosing a number k, as large that k > n′′ and n_k > n′, we get

$$\left\| f_n - f \right\| \le \left\| f_n - f_{n_k} \right\| + \left\| f_{n_k} - f \right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \in,$$

$\Rightarrow$ $\forall$ n > n′, we obtain $\| f_n - f \| < \varepsilon$, where ε is an arbitrary quantity.

$\Rightarrow$ $f_n \to f \in X$ and hence X is a complete space.

Theorem: Let $\{f_n\}$ be a sequence in L^p, $1 \le p < \infty$, such that $f_n \to f$ a.e. and that $f \in L^p$.

If $\lim_{n\to\infty} \left\| f_n \right\|_p = \| f \|_p$, then

$$\lim_{n\to\infty} \left\| f_n - f \right\|_p = 0.$$

Proof: Without any loss of generality, we may assume that each $f_n \ge 0$ a.e. so that f is also ≥ 0 a.e. since the result in general case follows by considering $f = f^+ - f^-$.

Now, let a and b be any pair of non-negative real numbers, we have

$$|a - b|^p \le 2^p (|a|^p + |b|^q),$$

$$1 \le p < \infty$$

So, we get

$$2^p (|f_n|^p + (|f|^p) - |f_n - f|^p \ge 0 \text{ a.e.}$$

Thus, by Fatou's Lemma and by the given hypothesis,

We get

$$\begin{aligned} 2^{p+1} \int |f|^p &= \int \lim_{n\to\infty} \left[2^p \left(|f_n|^p + |f|^p \right) - |f_n - f|^p \right] \\ &\le \liminf_{n\to\infty} \int \left[2^p \left(|f_n|^p + |f|^p \right) - |f_n - f|^p \right] \\ &= 2^{p-1} \lim_{n\to\infty} \int |f_n|^p + 2^p \int |f|^p + \liminf_{n\to\infty} \left(-\int |f_n - f|^p \right) \\ &= 2^{p+1} \int |f|^p - \limsup_{n\to\infty} \int |f_n - f|^p . \end{aligned}$$

Since $\int |f|^p < \infty$ it follows that

$$\limsup_{n\to\infty} \int |f_n - f|^p \le 0.$$

Therefore $\limsup_{n\to\infty} \int |f_n - f|^p = \liminf_{n\to\infty} \int |f_n - f|^p = 0$,

So that $\lim_{n\to\infty} \int |f_n - f|^p = 0$

Notes

$$\Rightarrow \qquad \lim_{n\to\infty}\left(\int |f_n - f|^p \, dt\right)^{\frac{1}{p}} = 0$$

$$\Rightarrow \qquad \lim_{n\to\infty}\| f_n - f \|_p = 0.$$

Theorem: In a normed linear space, every convergent sequence is a Cauchy sequence.

Proof: Let the sequence $<x_n>$ in a normed linear space N, converges to a point $x_o \in N_1$. We shall show that it is a Cauchy sequence.

Let $\in > 0$ be given. Since the sequence converges to x_o $\exists$ a positive integer m_o s.t.

$$n \geq m_o \Rightarrow \| x_n - x_o \| < \in/2 \qquad \ldots (1)$$

Hence for all m, n $\geq m_o$, we have

$$\begin{aligned} \| x_m - x_n \| &= \| x_m - x_o + x_o - x_n \| \\ &\leq \| x_m - x_o \| + \| x_o - x_n \| \\ &< \frac{\in}{2} + \frac{\in}{2} = \in \text{ by (1)} \end{aligned}$$

It follows that the convergent sequence $<x_n>$ is a Cauchy sequence.

Theorem: Prove that L^∞ [0, 1] is complete.

Proof: Let (f_n) be any Cauchy sequence in L^∞, and let

$$A_k = \{x : |f_k(x)| > \| f_k \|_\infty\},$$

$$B_{m,n} = \{x : |f_k(x) - f_m(x)| > \| f_k - f_m \|_\infty\}.$$

Then $m(A_k) = 0 = m(B_{m,n})$ (k, m, n = 1, 2, 3, …),

So that if E is the union of these sets, we have m (E) = 0.

Now, if $x \in F = [0, 1] - E$, then

$$|f_k(x) \leq \| f_k \|_\infty$$

$$|f_n(x) - f_m(x) \leq \| f_n - f_m \|_\infty \to 0 \text{ as } n, m \neq \infty.$$

Hence the sequence (f_n) converges uniformly to a bounded function on F.

Define $f : [0, 1] \to R$ by

$$f(x) = \begin{cases} \lim\limits_{n\to\infty} f_n(x) & \text{if } x \in F \\ 0, & \text{if } x \in E \end{cases}$$

Then $f \in L^\infty$ and $\| f_n - f \|_\infty \to 0$ as $n \to \infty$.

Thus L^∞ is

Hence proved.

7.2 Summary

- A sequence $<x_n>$ in a normal linear space X with norm $\| \cdot \|$ is said to converge to an element $x \in X$ if for arbitrary $\in > 0$, however small, $\exists$ n_o N s.t. $\| x_n - x \| < \in$, $\forall$ $n > n_o$. Then we write $\lim\limits_{n\to\infty} x_n = x$.

Notes

- A complete normed linear space is also called Banach space.
- The normed L^p-spaces are complete for ($p \geq 1$).
- A convergent sequence $<f_n>$ in L^p-spaces has a limit in L^p-space.
- A normed linear space is complete iff every absolutely summable sequence is summable.
- In a normed linear space, every convergent sequence is a Cauchy sequence.

7.3 Keywords

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

Cauchy Sequence: A sequence $<x_n>$ in a normal linear space $(X, \|\cdot\|)$ is said to be a Cauchy sequence if for arbitrary $\in > 0$, $\exists\, n_0 \in N$ s.t.

$$\| x_n - x_m \| < \in, \ \forall \ n, m \geq n_0.$$

Complete Normed Linear Space: A normed linear space $(X, \|\cdot\|)$ is said to be complete if every Cauchy sequence $<x_n>$ in it converges to an element $x \in X$.

Convergence almost Everywhere: Let $<f_n>$ be a sequence of measurable functions defined over a measurable set E. Then $<f_n>$ is said to converge almost everywhere in E if there exists a subset E_o of E s.t.

(i) $f_n(x) \to f(x), \ \forall \ x \in E - E_o$.

and (ii) $m(E_o) = 0$.

Convergent Sequence: A sequence $<x_n>$ in a normal linear space X with norm $\|\cdot\|$ is said to converge to an element $x \in X$ if for arbitrary $\in > 0$, however small, $\exists\, n_0 \in N$ such that $\| x_n - x \| < \in, \ \forall \ n > n_0$.

Then we write $\lim\limits_{n\to\infty} x_n = x$.

Normed Linear Space: A linear space N together with a norm defined on it, i.e., the pair $(N, \| \ \|)$ is called a normed linear space.

Summable Series: A series $\sum\limits_{n=1}^{\infty} u_n$ in N_1 is said to be summable to a sum u if $u \in N_1$ and $\lim\limits_{n\to\infty} S_n = u$, where

$$S_n = u_1 + u_2 + \ldots + u_n$$

7.4 Review Questions

1. Prove that ℓ_p^n is complete.

2. Prove that the vector space L^∞ equipped with $\|\cdot\|_{L^\infty}$ is a complete vector space.

3. Suppose $f \in L^\infty$ is supported on a set of finite measure.

   Then $f \in L^p$ for all $p < \infty$, and

   $$\| f \|_{L^\infty} \to \| f \|_{L^\infty} \text{ as } p \to \infty.$$

Notes

If $f \in L^p$ $(p > 0)$, $f \geq 0$ and $f_n = \min(f, n)$ $(n \in N)$, show that $f_n \in L^p$ and $\lim_{n\to\infty} \| f_n - f \|_p = 0$.

7.5 Further Readings

Books

H.L. Royden, *Real analysis*.

Walter Rudin, *Real and Complex Analysis*, Third, McGraw-Hill Book Co., New York, 1987.

Online links

www.public.iastate.edu

www.scribd.com/doc/49732162/103.Convergence-and-Completeness.

Notes

Unit 8: Bounded Linear Functional on the L^p-spaces

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand bounded linear functional on L^p-spaces
- Understand related theorems.
- Solve problems on bounded linear functionals.

Introduction

In this unit, we obtain the representation of bounded linear functionals on L^p-space. We shall also study about linear functional, continuous linear functionals and norm of $f \in \ell_p^*$. Further, we shall prove important theorems on bounded linear functionals.

8.1 Bounded Linear Functionals on L^p-spaces

8.1.1 Linear Functional

Definition: Let N_1 be a normed space over a field R (or C). A mapping $f : N_1 \to R$ (or C) is called a linear functional on N_1 if $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$, $\forall\ x, y \in N_1$ and $\alpha, \beta \in R$ (or C).

8.1.2 Bounded Linear Functional

Definition: A linear functional f on a normed space N_1 is said to be bounded if there is a constant $k > 0$ such that

$$|f(x)| \le k \|x\|, \ \forall\ x \in N_1 \qquad \dots (1)$$

Notes

The smallest constant k for which (1) holds is called the norm of f, written $\| f \|$.

Thus $\| f \| = \sup\left\{\frac{|f(x)|}{\|x\|} : x \neq 0 \text{ and } x \in N_1\right\}$ or equivalently

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

Also $$| f (x) | \leq \| f \| \, \| x \| \;\; \forall \; x \in N_1.$$

Definition: Let $p \in R$, $p > 0$. We define $L^p = L^p [0, 1]$ to be the set of all real-valued functions on [0, 1] such that

(i) f is measurable and (ii) $\| f \|_p = \left(\int_0^1 | f |^p\right)^{\frac{1}{p}} < \infty$.

Note L^1 or simply L denotes the class of measurable function f (x) which are also L-integrable.

8.1.3 Bounded Linear Functional on L^p-spaces

If $x \in \ell_p$ and f is bounded linear functional on ℓ_p, then f has the unique representation of the form as an infinite series

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

8.1.4 Norm

The norm of $f \in \ell_p^*$ is given by

$$\| f \| = \left\{\sum_{k=1}^{\infty} | f(e_k) |^q\right\}^{\frac{1}{q}}$$

Likewise in finite dimensional case, the bounded linear functionals are characterised by the values they assume on the set e_k, k = 1, 2, 3, … .

8.1.5 Continuous Linear Functional

A linear functional f is continuous if given $\in > 0$ there exists $\delta > 0$ so that

$$| f (x) - f (y) | \leq \in \text{ whenever } \| x - y \| \leq \delta.$$

8.1.6 Theorems

Theorem 1: Suppose $1 \leq p < \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$, then, with $\mathcal{B} = L^p$ we have

$$\mathcal{B}^* = L^q,$$

Notes

in the following sense: For every bounded linear functional ℓ on L^p there is a unique $g \in L^q$ so that

$$\ell(f) = \int_x f(x)g(x)\,d\mu(x), \text{ for all } f \in L^p$$

Moreover, $\|\ell\|_{B^*} = \|g\|_{L^q}$.

This theorem justifies the terminology where by q is usually called the dual exponent of p.

The proof of the theorem is based on two ideas. The first, as already seen, is Hölder's inequality; to which a converse is also needed. The second is the fact that a linear functional ℓ on L^p, $1 \le p < \infty$, leads naturally to a (signed) measure υ. Because of the continuity of ℓ the measure υ is absolutely continuous with respect to the underlying measure μ, and our desired function g is then the density function of υ in terms of μ.

We begin with:

Lemma: Suppose $1 \le p, q \le \infty$, are conjugate exponents.

(i) If $g \in L^q$, then $\|g\|_{L^q} = \sup_{\|f\|_{L^p} \le 1} \left|\int fg\right|$.

(ii) Suppose g is integrable on all sets of finite measure and

$$\sup_{\substack{\|f\|_{L^p} \le 1 \\ f \text{ simple}}} |fg| = M < \infty$$

Then $g \in L^q$, and $\|g\|_{L^q} = M$.

For the proof of the lemma, we recall the signum of a real number defined by

$$\text{sign}(x) = \begin{cases} 1 \text{ if } x > 0 \\ -1 \text{ if } x < 0 \\ 0 \text{ if } x = 0 \end{cases}$$

Proof: We start with (i). If g = 0, there is nothing to prove, so we may assume that g is not 0 a.e., and hence $\|g\|_{L^q} \ne 0$. By Hölder's inequality, we have that

$$\|g\|_{L^q} \ge \sup_{\|f\|_{L^p} \le 1} \left|\int fg\right|.$$

To prove the reverse inequality we consider several cases.

- First, if q = 1 and p = ∞, we may take f (x) = sign g (x). Then, we have $\|f\|_{L^\infty} = 1$ and clearly $\int fg = \|g\|_{L^1}$.
- If $1 < p, q < \infty$, then we set $f(x) = |g(x)|^{q-1} \text{ sign } g(x) / \|g\|_{L^q}^{q-1}$. We observe that $\|f\|_{L^p}^p = \int |g(x)|^{p(q-1)}\,d\mu / \|g\|_{L^q}^{p(q-1)} = 1$ since p (q – 1) = q, and that $\int fg = \|g\|_{L^q}$.

Notes

- Finally, if $q = \infty$ and $p = 1$, let $\in > 0$, and E a set of finite positive measure, where $|g(x)| \geq \|g\|_{L^\infty} - \in$. (Such a set exists by the definition of $\|g\|_{L^\infty}$ and the fact that the measure μ is σ-finite). Then, if we take $f(x) = \mathcal{X}E(x)$ sign $g(x)/\mu(E)$, where $\mathcal{X}E$ denotes the characteristic function of the set E, we see that $\|f\|_{L^1} = 1$, and also

$$\left|\int fg\right| = \frac{1}{\mu(E)}\int_E |g| \geq \|g\|_\infty - \in .$$

This completes the proof of part (i).

To prove (ii) we recall that we can find a sequence $\{g_n\}$ of simple functions so that $|g_n(x) \leq |g(x)|$ while $g_n(x) \to g(x)$ for each x. When $p > 1$ (so $q < \infty$), we take $f_n(x) = |g_n(x)|^{q-1}$ sign $g(x)\big/\|g_n\|_{L^q}^{q-1}$. As before, $\|f_n\|_{L^p} = 1$. However

$$\int f_n g = \frac{\int |g_n(x)|^q}{\|g_n\|_{L^q}^{q-1}} = \|g_n\|_{L^q},$$

and this does not exceed M. By Fatou's Lemmas if follows that $\int |g|^q \leq M^q$, so $g \in L^q$ with $\|g\|_{L^q} \leq M$. The direction $\|g\|_{L^q} \geq M$ is of course implied by Hölder's inequality. When $p = 1$ the argument is parallel with the above but simpler. Here we take $f_n(x) = (\text{sign } g(x))\, \mathcal{X}E_n(x)$, where E_n is an increasing sequence of sets of finite measure whose union is X. The details may be left to the reader.

With the lemma established we turn to the proof of the theorem. It is simpler to consider first the case when the underlying space has finite measure. In this case, with ℓ the given functional on L^p, we can then define a set function υ by

$$\upsilon(E) = \ell\,(\mathcal{X}E),$$

where E is any measurable set. This definition make sense because $\mathcal{X}_E$ is now automatically in L^p since the space has finite measure. We observe that

$$|\upsilon(E)| \leq C\,(\mu(E))^{1/p} \qquad \ldots (1)$$

where C is the norm of the linear functional, taking into account the fact that $\|\mathcal{X}_E\|_{L^p} = (\mu(E))^{1/p}$.

Now the linearity of ℓ clearly implies that υ is finitely-additive. Moreover, if $\{E_n\}$ is a countable collection of disjoint measurable sets, and we put $E = \bigcup_{n=1}^{\infty} E_n$, $E_N^* = \bigcup_{n=N+1}^{\infty} E_n$, then obviously

$$\mathcal{X}_E = \mathcal{X}_{E_N^*} + \bigcup_{n=1}^{N} \mathcal{X}E_n .$$

Then $\upsilon(E) = \upsilon(E_N^*) + \sum_{n=1}^{N} \upsilon(E_n)$. However $\upsilon(E_N^*) \to 0$, as $N \to \infty$ because of (1) and the assumption $p < \infty$. This shows that υ is countably additive and moreover (1) also shows us that υ is absolutely continuous with respect to μ.

We can now invoke the key result about absolutely continuous measures, the Lebesgue-Radon – Nykodin theorem. It guarantees the existence of an integrable function g so that $\upsilon(E) = \int_E g\, du$

Notes

for every measurable set E. Thus we have $\ell(\mathcal{X}E) = \int \mathcal{X}Eg\, d\mu$. The representation $\ell(f) = \int fg\, d\mu$ then extends immediately to simple function f, and by a passage to the limit, to all $f \in L^p$ since the simple functions are dense in L^p, $1 \le p < \infty$. Also by lemma, we see that $\|g\|_{L^q} = \|\ell\|$.

To pass from the situation where the measure of X is finite to the general case, we use an increasing sequence $\{E_n\}$ of sets of finite measure that exhaust X, that is, $X = \bigcup_{n=1}^{\infty} E_n$. According to what we have just proved, for each n there is an integrable function g_n on E_n (which we can set to be zero in E_n^c) so that

$$\ell(f) = \int fg_n\, d\mu \qquad \dots (2)$$

whenever f is supported in E_n and $f \in L^p$. Moreover by conclusion (ii) of the lemma $\|g_n\|_{L^p} \le \|\ell\|$.

Now it is easy to see because of (2) that $g_n = g_m$ a.e. on E_m, whenever $n \ge m$. Thus $\lim_{n \to \infty} g_n(x) = g(x)$ exists for almost every x, and by Fatou's lemma, $\| g \|_{L^q} \le \| \ell \|$. As a result we have that $\ell(f) = \int f\, g\, du$ for each $f \in L^p$ supported in E_n, and then by a simple limiting argument, for all $f \in L^p$ supported in E_n. The fact that $\|\ell\| \le \| g \|_{L^q}$ is already contained in Hölder's inequality and therefore the proof of the theorem is complete.

Theorem 2: 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

$$|f(x_n)| \ge n \| x_n \| \qquad \dots (1)$$

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

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

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

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

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

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

It now follows from (1) and (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.

Notes

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

$|f(x_n)| \le k \|x_n\| \ \forall\, n = 1, 2, \ldots$ and $k \ge 0$.

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.

Theorem 3: If L is a linear space of all n-tuples, then

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

(ii) $\left(\ell_1^n\right)^* = \ell_\infty^n$

(iii) $\left(\ell_\infty^n\right)^* = \ell_1^n$

Proof: Let $(e_1, e_2, \ldots, e_n)$ be a standard basis for L so that any $x = (x_1, x_2, \ldots, x_n) \in L$ can be written as $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$.

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

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

$\Rightarrow$ f determines and is determined by n scalars $y_i = f(e_i)$.

Then the mapping

$y = (y_1, y_2, \ldots y_n) \to f$ 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 $L = \ell_p^n$ $(1 \le p < \infty)$ with the pth 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 $y \to f$ as an isometric isomorphism we try to find the norm of y's.

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

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

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

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

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

By using Holder's inequality, we get

$$\sum_{i=1}^{n} |x_i y_i| \le \left\{\sum_{i=1}^{n} |x_i|^p\right\}^{\frac{1}{p}} \left\{\sum_{i=1}^{n} |y_i|^q\right\}^{\frac{1}{q}} \text{ so that}$$

Notes

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

Using the definition of norm for f, we get

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

Consider the vector, defined by

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

Then,

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

Since q = p (q – 1) we have from (4),

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

Now

$$|f(x)| = \left|\sum_{i=1}^{n}x_iy_i\right| = \left|\sum_{i=1}^{n}\frac{|y_i|^q}{y_i}y_i\right| \qquad \dots (4)$$

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

So that

$$\sum_{i=1}^{n}|y_i|^q = |f(x) \le \|f\| \; \|x\| \qquad \dots (6)$$

From (5) and (6) we get,

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

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

Also from (2) and (7) we have

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

Notes

$y \to f$ is an isometric isomorphism.

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

(ii) Let $L = \ell_1^n$ 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 $$\sum_{i=1}^{n} |x_i|\,|y_i| \leq \max.\{|y_i|\} \sum_{i=1}^{n} |x_i| \text{ so that}$$

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

From the definition of the 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 $|y_i| = \max\limits_{1 \leq i \leq n}\{|y_i|\}$, let us consider the vector x as

$$x_i = \frac{|y_i|}{y_i} \text{ when } |y_i| = \max_{1 \leq i \leq n}\{|y_i|\} \text{ and } x_i = 0 \qquad \ldots (9)$$

otherwise

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

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

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

Hence $|y_i| = |f(x)| \leq \|f\|\,\|x\|$

Notes

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

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

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 $\left(\ell_1^n\right)^* = \ell_\infty^n$.

(iii) Let L = ℓ_∞^n with the norm

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

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.

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

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

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

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

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

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

$$\Rightarrow \qquad \sum_{i=1}^{n} |y_i| \le \|f\| \qquad \ldots (13)$$

Notes

It follows now from (11) and (13) that

$$\|f\| = \sum_{i=1}^{n} |y_i| \text{ so that } y \to f \text{ is an isometric isomorphism.}$$

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

This completes the proof of the theorem.

Note We need the signum function for finding the conjugate spaces of some infinite dimensional space which we define as follows:

If γ is a complex number, then

$$\left.\begin{aligned} \operatorname{sgn}\gamma &= \frac{\gamma}{|\gamma|} \text{ if } \gamma \neq 0 \\ &= 0 \text{ if } \gamma = 0 \end{aligned}\right\}$$

$\therefore$ (i) $|\operatorname{sgn}\gamma| = 0$ if $\gamma = 0$ and $|\operatorname{sgn}\gamma| = 1$ if $\gamma \neq 0$

(ii) $\gamma \operatorname{sgn}\bar{\gamma} = 0$ if $\gamma = 0$ and $\gamma \operatorname{sgn}\bar{\gamma} = \dfrac{\gamma\bar{\gamma}}{|\gamma|} = |\gamma|$, if $\gamma \neq 0$.

Theorem 4: 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 $\sum_{n=1}^{\infty} |x_n|^p < \infty$ … (1)

Let $e_n = (0, 0, 0, \ldots, 1, 0, 0, \ldots)$ where 1 is in the nth place.

$$e_n \in \ell_p \text{ for } n = 1, 2, 3, \ldots .$$

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

$(x_1, x_2, \ldots x_n, 0, 0, 0, \ldots)$ in the form $\sum_{k=1}^{n} x_k e_k$ and

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

Notes

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) if follows that

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

Let $f \in \ell^*_p$ and $S_n = \sum_{k=1}^{n} x_k e_k$ then

$$S_n \to x \text{ as } n \to \infty \qquad \text{(Using (4))}$$

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

$$f(S_n) \to f(x) \text{ as } n \to \infty.$$

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

which gives the form of the functional on ℓ_p.

Now we establish the isomeric isomorphism of ℓ^*_p onto ℓ_q, for which proceed as follows:

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

$T : \ell^*_p \to \ell_q$ given by

$T(f) = (\alpha_1, \alpha_2, \ldots, \alpha_k, \ldots)$ is an isomeric 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

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

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

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

Notes

Now $\alpha_k\beta_k = \alpha_k |\alpha_k|^{q-1} \operatorname{sgn} \overline{\alpha}_k = |\alpha_k|^{q-1} \alpha_\kappa \operatorname{sgn} \overline{\alpha}_k$

$$\Rightarrow \qquad \alpha_k \beta_k = |\alpha_k|^q = |\beta_k|^p \qquad \ldots (7)$$

(Using property of sgn function)

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

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

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

Since we can write

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

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

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

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.

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

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

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

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

Notes

We next show that T is onto ℓ_q.

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

Let $x \in \ell_p$ so that

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

We shall show that

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

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

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

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

$$\Rightarrow \qquad |g(x)| \le \|x\| \left\{\sum_{k=1}^{n} |\beta_k|^q\right\}^{\frac{1}{q}}$$

$\Rightarrow$ 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_q.

We next show that

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

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

$$\left\{\sum_{k=1}^{\infty} |\alpha_k|^q\right\}^{\frac{1}{q}} = \| Tf \| \le \| f \|$$

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

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

Notes

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

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

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

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

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

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 (already shown). Hence T is an isometric isomorphism of ℓ^*_p onto ℓ_q, i.e.,

$$\ell^*_p = \ell_q .$$

This completes the proof of the theorem.

Theorem 5: Let p > 1 with $\frac{1}{p}+\frac{1}{q}=1$ and let $g \in L_p(X)$. Then the function defined by

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

is a bounded linear functional on $L_p(X)$ and

$$\|F\| = \|g\|_q \qquad \ldots (1)$$

Proof: We first note that

F is linear on $L_p(X)$. For if $f_1, f_2 \in L_p(X)$, then we get

$$F(f_1 + f_2) = \int_X (f_1+f_2)g\, d\mu = \int_X f_1 g\, d\mu + \int_X f_2 g\, d\mu$$

$$= F(f_1) + F(f_2)$$

So that

$$F(f_1 + f_2) = F(f_1) + F(f_2)$$

and
$$F(\alpha F) = \alpha \int_X fg\, d\mu = \alpha F(f).$$

Now
$$|F(f)| = \left|\int_X fg\, d\mu\right| \le \int_X |fg|\, d\mu \qquad \ldots (2)$$

Notes

Making use of Hölder's inequality, we get

$$\int_X |fg|\,d\mu \le \left\{\int_X |f|^p\,d\mu\right\}^{\frac{1}{p}} \left\{\int_X |g|^q\,d\mu\right\}^{\frac{1}{q}}$$

$$= \|f\|_p \, \|g\|_q \qquad \ldots (3)$$

From (2) and (3) it follows that

$$|F(f)| \le \|f\|_p \, \|g\|_q.$$

Hence $\sup\left\{\dfrac{|F|f||}{\|f\|_p} : f \in Lp(X) \text{ and } f \ne 0\right\} \le \|g\|_q$

$\Rightarrow$ $\|F\| \le \|g\|_q$ (Using definition of the norm) … (4)

Further, let $f = |g|^{q-1} \operatorname{sgn} \bar{g}$ … (5)

Since $|\operatorname{sgn} \bar{g}| = 1$, we get

$$|f|^p = |g|^{p(q-1)} = |g|^q \qquad (\because p(q-1) = q)$$

Thus, $f \in L_p(X)$ and $\left\{\int_X |f|^p\,d\mu\right\}^{\frac{1}{p}} = \left\{\int_X |g|^q\,d\mu\right\}^{\frac{1}{q}}$ … (6)

But $\left\{\int_X |g|^q\,d\mu\right\}^{\frac{1}{p}} \left\{\int_X |g|^p\,d\mu\right\} = \|g\|_q^{q/p}$

which implies on using (6) that

$$\|f\|_p = \|g\|_q^{q/p} \qquad \ldots (7)$$

Now $F(f) = \int_X fg\,d\mu = \int_X |g|^{q-1}\, g \operatorname{sgn} \bar{g}\,d\mu$

$$= \int_X |g|^q\,d\mu = \|g\|_q^q$$

Hence $\|g\|_q^q \|g\| = F(f) \le \|F\| \, \|f\|_p.$

and this on using (7) yields that

$$\|g\|_q^q = F(f) \le \|F\| \, \|g\|_q^{q/p}$$

$\Rightarrow$ $\|g\|_q^{q-q/p} = \|g\|_q \le \|F\|$ … (8)

$(\because g \ne 0)$

Notes

From (4) and (8) it finally follows that

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

This completes the proof of the theorem.

Approximation by Continuous Function

Theorem 6: If f is a bounded measurable function defined on [a, b], then for given $\varepsilon > 0$, $\exists$ a continuous function g on [a, b], such that

$$\| f - g \|_2 < \varepsilon$$

Proof: Let $$F(x) = \int_a^x f(t)\,dt \text{ where } x \in [a, b].$$

Then $$|F(x+h) - F(x)| = \left| \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt \right|$$

$$= \left| \int_x^{x+h} f(t)\,dt \right| \le \int_x^{x+h} f(t)\,dt$$

$$\le Mh, \text{ where } |f(x)| \le M, \ \forall x \in [a, b].$$

Taking $h < \delta$, and $Mh < \varepsilon_1$, we get

$$|x + h - x| < \delta \Rightarrow |F(x+h) - F(x)| < \varepsilon_1$$

$$\Rightarrow F(x) \text{ is continuous on } [a, b].$$

Let $$G_n(x) = n \int_x^{x+h} f(t)\,dt : x \in [a, b] \text{ and } n \in N;$$

then $$G_n(x) = n\left[F\left(x + \frac{1}{n}\right) - F(x)\right]$$ ($\because$ F (x) is continuous on [a, b] $\Rightarrow$ $G_n(x)$ is continuous on [a, b] $\forall$ n)

Again, since $$F(x) = \int_a^x f(t)\,dt, x \in [a, b].$$

$\therefore$ $$F'(x) = f(x) \text{ a.e. in } [a, b].$$

Now, $$\lim_{n\to\infty} G_n(x) = \lim_{n\to\infty} \frac{F(x + (1/n) - F(x)}{1/n}$$

$$= \lim_{h\to 0} \frac{F(x+h) - F(x)}{h}, h = \frac{1}{n}$$

$$= F'(x) = f(x) \text{ a.e. in } [a, b]$$

and hence $$\lim_{n\to\infty} [G_n(x) - f(x)]^2 = 0.$$

Notes

Also $$|G_n(x)| = \left| \int_x^{x+(1/n)} f(t)\,dt \right| \le n \int_x^{x+(1/n)} |f(t)|\,dt = M.$$

Hence $|G_n(x)| \le M$, $\forall\, n \in N$ and $\forall\, x \in [a, b]$.

$\therefore$ $$[G_n(x) - f(x)]^2 \le (M + M)^2 = 4M^2,\ x \in [a, b].$$

On applying Lebesgue bounded convergence theorem, we get

$$\lim_{n\to\infty} \int_a^b (G_n - f)^2 = \int_a^b \lim (G_n - f)^2 = 0$$

$\Rightarrow$ $$\lim_{n\to\infty} [G_n - f]_2^2 = 0$$

$\Rightarrow$ $$\lim_{n\to\infty} \|G_n - f\|_2 = 0$$

or $$\lim_{n\to\infty} \|f - G_n\| = 0$$

$\Rightarrow$ for given $\varepsilon > 0$, $\exists\, n_o \in N$, such that $n \ge n_o$

$\Rightarrow$ $$\|f - G_n\|_2 < \varepsilon$$

Particularly for $n = n_o$.

$\Rightarrow$ $$\|f - G_{n_o}\|_2 < \varepsilon$$

$\Rightarrow$ $$\|f - g\|_2 < \varepsilon \qquad \text{(Taking } G_{n_o} = g\text{)}$$

Thus there exists a continuous function $G_{n_o}(x) = g(x)$

$$= n_o \int_x^{x+(1/n_o)} f(t)\,dt,\ x \in [a, b],$$

which satisfies the given condition.

8.2 Summary

- A linear functional f on a normed space N_1 is said to be bounded if there is a constant $k > 0$ such that

$$|f(x)| \le k\,\|x\|,\ \forall\ x \in N_1$$

- If $x \in \ell_p$ and f is bounded linear functional on ℓ_p, then f has the unique representation of the form as an infinite series.

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

- The norm of $f \in \ell_p^*$ is given by

$$\|f\| = \left\{ \sum_{k=1}^{\infty} |f(e_k)|^q \right\}^{\frac{1}{q}}$$

Notes

8.3 Keywords

Bounded Linear Functional on L^p-spaces: If $x \in \ell_p$ and f is bounded linear functional on ℓ_p, then f has the unique representation of the form as an infinite series

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

Bounded Linear Functional: A linear functional f on a normed space N_1 is said to be bounded if there is a constant k > 0 such that

$$|f(x)| \le k \|x\|, \ \forall \ x \in N_1$$

Continuous Linear Functional: A linear functional f is continuous if given $\in > 0$ there exists $\delta > 0$ so that

$$|f(x) - f(y)| \le \in \text{ whenever } \|x - y\| \le \delta.$$

Linear Functional: Let N_1 be a normed space over a field R (or C). A mapping $f : N_1 \to R$ (or C) is called a linear functional on N_1 if $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y), \ \forall \ x, y \in N_1$ and $\alpha, \beta \in R$ (or C).

Norm: The norm of $f \in \ell_p^*$ is given by

$$\|f\| = \left\{ \sum_{k=1}^{\infty} |f(e_k)|^q \right\}^{\frac{1}{q}}$$

8.4 Review Questions

1. Account for bounded linear functionals on L^p-space.
2. State and prove different continuous linear functional theorems.
3. Describe approximation by continuous function.
4. How will you explain norms of bounded linear functional on L^p-space?
5. What is Isometric Isomorphism?

8.5 Further Readings

Books

Rudin, Walter (1991), *Functional Analysis*, Mc-Graw-Hill Science/Engineering/Math

Kreyszig, Erwin, *Introductory Functional Analysis with Applications*, WILEY 1989.

T.H. Hilderbrandt, *Transactions of the American Mathematical Society*. Vol. 36, No. = 4, 1934.

Online links

www.math.psu.edu/yzheng/m597k/m597kLIII4.pdf

www.public.iastate.edu/…/Royden_Real_Analysis_Solutions.pdf

Notes

Unit 9: Measure Spaces

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define measure space.
- Define null set in a measure space.
- Understand theorems based on measure spaces.
- Solve problems on measure spaces.

Introduction

A measurable space is a set S, together with a non-empty collection, S, of subsets of S, satisfying the following two conditions:

1. For any A, B in the collection S, the set[1] A – B is also in S.
2. For any $A_1, A_2, \ldots \in S, \cup A_i \in S$.

The elements of S are called measurable sets. These two conditions are summarised by saying that the measurable sets are closed under taking finite differences and countable unions.

9.1 Measure Space

Measurable Space: Let $\mathcal{U}$ be a σ-algebra of subsets of set X. The pair $(X, \mathcal{U})$ is called a measurable space. A subset E of X is said to be $\mathcal{U}$-measurable if $E \in \mathcal{U}$.

(a) If μ is a measure on a σ-algebra $\mathcal{U}$ of subsets of a set X, we call the triple $(X, \mathcal{U}, u)$ a measure space.

(b) A measure μ on a σ-algebra $\mathcal{U}$ of subsets of a set X is called a finite measure if $m(X) < \infty$. In this case $(X, \mathcal{U}, \mu)$ is called a finite measure space.

Notes

(c) A measure μ on a σ-algebra $\mathcal{U}$ of subsets of a set X is called a *σ-finite measure* if there exists a sequence $(E_n : n \in \mathcal{N})$ in $\mathcal{U}$ such that $\mathcal{U}_{n \in \mathcal{N}} E_n = X$ and $\mu(E_n) < \infty$ for every $n \in \mathcal{N}$. In this case $(X, \mathcal{U}, \mu)$ is called a σ-finite measure space.

(d) A set $D \in \mathcal{U}$ in an arbitrary measure space $(X, \mathcal{U}, \mu)$ is called a *σ-finite set* if there exists a sequence $(D_n : n \in \mathcal{N})$ in $\mathcal{U}$ such that $U_{n \in \mathcal{N}} D_n = D$ and $\mu(D_n) < \infty$ for every $n \in \mathcal{N}$.

Lemma 1: (a) Let $(X, \mathcal{U}, \mu)$ be a measure space. If $D \in \mathcal{U}$ is a σ-finite set, then there exists an increasing sequence $(F_n : n \in \mathcal{N})$ in $\mathcal{U}$ such that $\lim_{n\to\infty} F_n = D$ and $\mu(F_n) < \infty$ for every $n \in \mathcal{N}$ and there exists a disjoint sequence $(G_n : n \in \mathcal{N})$ in $\mathcal{U}$ such that $\mathcal{U}_{n\in\mathcal{N}} G_n = D$ and $\mu(G_n) < \infty$ for every $n \in \mathcal{N}$.

(b) If $(X, \mathcal{U}, \mu)$ is a σ-finite measure space then every $D \in \mathcal{U}$ is a σ-finite set.

Proof 1: Let $(X, \mathcal{U}, \mu)$ be a measure space. Suppose $D \in \mathcal{U}$ is a σ-finite set. Then there exists a sequence $(D_n : n \in \mathcal{N})$ in $\mathcal{U}$ such that $U_{n \in \mathcal{N}} D_n = D$ and $\mu(D_n) < \infty$ for every $n \in \mathcal{N}$. For each $n \in \mathcal{N}$, let $F_n = U_{k=1}^{n} D_k$. Then $(F_n : n \in \mathcal{N})$ is an increasing sequence in $\mathcal{U}$ such that $\lim_{n\to\infty} F_n = U_{n\in\mathcal{N}} F_n$

$$= U_{n\in\mathcal{N}} D_n = D \text{ and } \mu(F_n) = \mu\left(\bigcup_{k=1}^{n} D_k\right) \le \sum_{k=1}^{n} \mu(D_k) < \infty \text{ for every } n \in \mathcal{N}.$$

Let $G_1 = F_1$ and $G_n = F_n \setminus U_{k=1}^{n-1} F_k$ for $n \ge 2$. Then $(G_n : n \in \mathcal{N})$ is a disjoint sequence in $\mathcal{U}$ such that $U_{n\in\mathcal{N}} G_n = U_{n\in\mathcal{N}} F_n = D$ as in the proof of Lemma "let $(E_n : n \in \mathcal{N})$ be an arbitrary sequence in an algebra $\mathcal{U}$ of subsets of a set X. Then there exists a disjoint sequence $(F_n : n \in \mathcal{N})$ in $\mathcal{U}$ such that

(1) $\bigcup_{n=1}^{N} E_n = \bigcup_{n=1}^{N} F_n$ for every $N \in \mathcal{N}$,

and

(2) $\bigcup_{n\in\mathcal{N}} E_n = \bigcup_{n\in\mathcal{N}} F_n$ ".

$\mu(G_1) = \mu(F_1) < \infty$ and $\mu(G_n) = \mu\left(F_n \middle| \bigcup_{k=1}^{n-1} F_k\right) \le \mu(F_n) < \infty$ for $n \ge 2$. This proves (a).

2. Let $(X, \mathcal{U}, \mu)$ be a σ-finite measure space. Then there exists a sequence $(E_n : n \in \mathcal{N})$ in $\mathcal{U}$ such that $U_{n \in \mathcal{N}} E_n = X$ and $\mu(E_n) < \infty$ for every $n \in \mathcal{N}$. Let $D \in \mathcal{U}$. For each $n \in \mathcal{N}$, let $D_n = D \cap E_n$. Then $(D_n : n \in \mathcal{N})$ is a sequence in $\mathcal{U}$ such that $U_{n\in\mathcal{N}} D_n = D$ and $m(D_n) \le \mu(E_n) < \infty$ for every $n \in \mathcal{N}$. Thus D is a σ-finite set. This proves (b).

9.1.1 Null Set in a Measure Space

Definition: Given a measure μ on a σ-algebra $\mathcal{U}$ of subsets of a set X · A subset E of X is called a null set with respect to the measure μ if $E \in \mathcal{U}$ and $\mu(E) = 0$. In this case we say also that E is a null set in the measure space $(X, \mathcal{U}, \mu)$. (Note that ϕ is a null set in any measure space but a null set in a measure space need not be ϕ.)

Theorem 1: A countable union of null sets in a measure space is a null set of the measure space.

Proof: Let $(E_n : n \in \mathcal{N})$ be a sequence of null sets in a measure space $(X, \mathcal{U}, \mu)$. Let $E = U_{n\in\mathcal{N}} E_n$. Since $\mathcal{U}$ is closed under countable unions,

we have $E \in \mathcal{U}$.

Notes

By the countable subadditivity of μ on $\mathcal{U}$,

we have $\mu(E) \leq \Sigma_{n \in \mathcal{N}} \mu(E_n) = 0$.

Thus $\mu(E) = 0$.

This shows that E is a null set in $(X, \mathcal{U}, \mu)$.

9.1.2 Complete Measure Space

Definition: Given a measure μ on a σ-algebra $\mathcal{U}$ of subsets of a set X. We say that the σ -algebra $\mathcal{U}$ is complete with respect to the measure μ if an arbitrary subset E_0 of a null set E with respect to μ is a member of $\mathcal{U}$ (and consequently has $\mu(E_0) = 0$ by the Monotonicity of μ). When $\mathcal{U}$ is complete with respect to μ, we say that $(X, \mathcal{U}, \mu)$ is a complete measure space.

Example: Let X = {a, b, c}. Then $\mathcal{U}$ = {ϕ, {a}, {b, c}, X} is a σ-algebra of subsets of X. If we define a set function μ on $\mathcal{U}$ by setting $\mu(\phi) = 0$, $\mu(\{a\}) = 1$, $\mu(\{b, c\}) = 0$, and $\mu(X) = 1$, then μ is a measure on $\mathcal{U}$. The set {b, c} is a null set in the measure space $(X, \mathcal{U}, \mu)$, but its subset {b} is not a member of $\mathcal{U}$. Therefore, $(X, \mathcal{U}, \mu)$ is not a complete measure space.

9.1.3 Measurable Mapping

Let f be a mapping of a subset D of a set X into a set Y. We write D (f) and R (f) for the domain of definition and the range of f respectively. Thus

$$D(f) = D \subset X,$$

$$R(f) = \{y \in Y : y = f(x) \text{ for some } x \in D(f)\} \subset Y.$$

For the image of D (f) by f, we have f (D (f)) = R (f). For an arbitrary subset E of y we define the preimage of E under the mapping f by

$$F^{-1}(E) := \{x \in X : f(x) \in E\} = \{x \in D(f) : f(x) \in E\}.$$

Notes

1. E is an arbitrary subset of Y and need not be a subset of R (f). Indeed E may be disjoint from $\mathcal{R}$ (f), in which case $f^{-1}(E) = \phi$. In general, we have $f(f^{-1}(E)) \subset E$.
2. For an arbitrary collection C of subsets of Y, we let $f^{-1}(C) := \{f^{-1}(E) : E \in C\}$.

Theorem 2: Given sets X and Y. Let f be a mapping with $D(f) \subset X$ and $\mathcal{R}(f) \subset Y$. Let E and E_α be arbitrary subsets of Y. Then

1. $f^{-1}(Y) = D(f)$,
2. $f^{-1}(E^C) = f^{-1}(Y \backslash E) = f^{-1}(Y) \backslash f^{-1}(E) = D(f) \backslash f^{-1}(E)$,
3. $f^{-1}(U_{\alpha \in \mathcal{A}} E_\alpha) = U_{\alpha \in \mathcal{A}} f^{-1}(E_\alpha)$,
4. $f^{-1}(\cap_{\alpha \in \mathcal{A}} E_\alpha) = \cap_{\alpha \in \mathcal{A}} f^{-1}(E_\alpha)$.

Theorem 3: Given sets X and Y. Let f be a mapping with $D(f) \subset X$ and $R(f) \subset Y$. If $\mathcal{B}$ is a σ-algebra of subsets of Y then $f^{-1}(\mathcal{B})$ is a σ-algebra of subsets of the set D (f). In particular, if D (f) = X then $f^{-1}(\mathcal{B})$ is a σ-algebra of subsets of the set X.

Notes

Proof: Let $\mathcal{B}$ be a σ-algebra of subsets of the set Y. To show that $f^{-1}(\mathcal{B})$ is a σ-algebra of subsets of the set D (f) we show that D (f) ∈ $f^{-1}(\mathcal{B})$; if A ∈ $f^{-1}(\mathcal{B})$ then D (f)\A ∈ $f^{-1}(\mathcal{B})$; and for any sequence $(A_n : n \in \mathcal{N})$ in $f^{-1}(\mathcal{B})$ we have $U_{n \in \mathcal{N}} A_n \in f^{-1}(B)$.

1. By (1) of above theorem, we have D (f) = $f^{-1}(Y) \in f^{-1}(B)$ since Y ∈ $\mathcal{B}$.
2. Let A ∈ $f^{-1}(\mathcal{B})$. Then A = $f^{-1}(\mathcal{B})$ for some B ∈ $\mathcal{B}$. Since $B^C \in \mathcal{B}$ we have $f^{-1}(B^C) \in f^{-1}(\mathcal{B})$. On the other hand by (2) of above theorem, we have $f^{-1}(B^C)$ = D (f)\$f^{-1}$ (B) = D (f)\A. Thus D (f)\A ∈ $f^{-1}(\mathcal{B})$.
3. Let $(A_n : n \in \mathcal{N})$ be a sequence in $f^{-1}(\mathcal{B})$. Then $A_n = f^{-1}(B_n)$ for some $B_n \in \mathcal{B}$ for each n ∈ $\mathcal{N}$. Then by (3) of above theorem, we have

$$\bigcup_{n\in\mathcal{N}} A_n = \bigcup_{n\in\mathcal{N}} f^{-1}(B_n) = f^{-1}\left(\bigcup_{n\in\mathcal{N}} B_n\right) \in f^{-1}(\mathcal{B}),$$

since $\bigcup_{n\in\mathcal{N}} B_n \in (\mathcal{B})$.

Measurable Mapping

Definition: Given two measurable spaces (X, $\mathcal{U}$) and (Y, $\mathcal{B}$). Let f be a mapping with D (f) ⊂ X and $\mathcal{R}$ (f) ⊂ Y. We say that f is a $\mathcal{U}/\mathcal{B}$ measurable mapping if $f^{-1}(B) \in \mathcal{U}$ for every B ∈ $\mathcal{B}$, that is, $f^{-1}(\mathcal{B}) \subset \mathcal{U}$.

Theorem 4: Given two measurable spaces (X, $\mathcal{U}$) and (Y, $\mathcal{B}$). Let f be a $\mathcal{U}/\mathcal{B}$-measurable mapping.

(a) If $\mathcal{U}$, is a σ-algebra of subsets of X such that $\mathcal{U}, \supset \mathcal{U}$, then f is $\mathcal{U}_1/\mathcal{B}$-measurable.

(b) If $\mathcal{B}_0$ is a σ-algebra of subsets of Y such that $\mathcal{B}_0 \subset \mathcal{B}$, then f is $\mathcal{U}/\mathcal{B}_0$-measurable.

Proof: (a) Follows from $f^{-1}(\mathcal{B}) \subset \mathcal{U} \subset \mathcal{U}_1$ and (b) from $f^{-1}(B_0) \subset f^{-1}(\mathcal{B}) \subset \mathcal{U}$.

Composition of two measurable mappings is a measurable mapping provided that the two measurable mappings from a chain.

Theorem 5: Given two measurable spaces (X, $\mathcal{U}$) and (Y, $\mathcal{B}$), where $\mathcal{B}$ = σ ($\mathcal{C}$) and $\mathcal{C}$ is arbitrary collection of subsets of Y. Let f be a mapping with D (f) ∈ $\mathcal{U}$ and $\mathcal{R}$ (f) ⊂ Y. Then f is a $\mathcal{U}/\mathcal{B}$-measurable mapping of D (f) into Y if and only if $f^{-1}(\mathcal{C}) \subset \mathcal{U}$.

Proof: If f is a $\mathcal{U}/\mathcal{B}$-measurable mapping of D (f) into Y, then $f^{-1}(\mathcal{B}) \subset \mathcal{U}$ so that $f^{-1}(\mathcal{C}) \subset \mathcal{U}$. Conversely if $f^{-1}(\mathcal{C}) \subset \mathcal{U}$, then σ ($f^{-1}(\mathcal{C})$ ⊂ σ ($\mathcal{U}$) = $\mathcal{U}$. Now by theorem,

"Let f be a mapping of a set X into a set Y. Then for an arbitrary collection C of subsets of Y, we have σ ($f^{-1}(\mathcal{C})$) = f^{-1} (σ ($\mathcal{C}$)."

σ ($f^{-1}(\mathcal{C})$ = f^{-1} (σ ($\mathcal{C}$)) = $f^{-1}(\mathcal{B})$. Thus $f^{-1}(\mathcal{B}) \subset \mathcal{U}$ and f is a $\mathcal{U}/\mathcal{B}$- measurable mapping of D (f).

Theorem 6: If X_0 is a thick subset of a measure space (X, S, μ), if $S_0 = S \cap X_0$, and if, for E in S, μ_0 (E ∩ X_0) = μ (E), then (X_0, S_0, μ_0) is a measure space.

Proof: If two sets, E_1 and E_2, in S are such that $E_1 \cap X_0 = E_2 \cap X_0$, then $(E_1 \Delta E_2) \cap X_o = 0$, so that μ ($E_1 \cap E_2$) = 0 and therefore μ ($E_1$) = μ ($E_2$). In other words μ_0 is indeed unambiguously defined on S_0.

Suppose next that $\{F_n\}$ is a disjoint sequence of sets in S_0, and let E_n be a set in S such that

$$F_n = E_n \cap X_0, n = 1, 2, \ldots .$$

Notes

If $\tilde{E}_n = E_n - \cup\{E_i : 1 \leqq i < n\}, n = 1, 2, \ldots$, then

$$\left(\tilde{E}_n \,\Delta\, E_n\right) \cap X_0 = \left(F_n - \cup\left\{F_i : 1 \leqq i < n\right\}\right)\Delta F_n$$

$$= F_n \,\Delta\, F_n = 0,$$

So that $\mu\left(\tilde{E}_n \Delta E_n\right) = 0$, and therefore

$$\Sigma_{n=1}^{\infty} \mu_0(F_n) = \Sigma_{n=1}^{\infty} \mu_0(E_n) = \Sigma_{n=1}^{\infty} \mu_0(\tilde{E}_n) = \mu\left(\Sigma_{n=1}^{\infty} \tilde{E}_n\right)$$

$$= \mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \mu_0\left(\bigcup_{n=1}^{\infty} F_n\right)$$

In other word μ_0 is indeed a measure, and the proof of the theorem is complete.

9.2 Summary

- Let $\mathcal{U}$ be a σ-algebra of subsets of a set X. The pair (X, $\mathcal{U}$) is called a measurable space. A subset E of X is said to be $\mathcal{U}$-measurable if E ∈ $\mathcal{U}$.
- If μ is a measure on a σ-algebra $\mathcal{U}$ of subsets of a set X, we call the triple (X, $\mathcal{U}$, μ) a measure space.
- A subset E of X is called a null set with respect to the measure μ if E ∈ $\mathcal{U}$ and μ (E) = 0.
- Two measurable spaces (X, $\mathcal{U}$) and (Y, $\mathcal{B}$). Let f be a mapping with D (f) ⊂ X and $\mathcal{R}$ (f) ⊂ Y. We say that f is a $\mathcal{U}/\mathcal{B}$-measurable mapping if f^{-1} (B) ∈ $\mathcal{U}$ for every B ∈ $\mathcal{B}$, that is f^{-1} ($\mathcal{B}$) ⊂ $\mathcal{U}$.

9.3 Keywords

Complete Measure Space: Given a measure μ on a σ-algebra $\mathcal{U}$ of subsets of a set X. We say that the σ -algebra $\mathcal{U}$ is complete with respect to the measure μ if an arbitrary subset E_0 of a null set E with respect to μ is a member of $\mathcal{U}$ (and consequently has μ (E_0) = 0 by the Monotonicity of μ). When $\mathcal{U}$ is complete with respect to μ, we say that (X, $\mathcal{U}$, μ) is a complete measure space.

Measurable Mapping: Given two measurable spaces (X, $\mathcal{U}$) and (Y, $\mathcal{B}$). Let f be a mapping with D (f) ⊂ X and $\mathcal{R}$ (f) ⊂ Y. We say that f is a $\mathcal{U}/\mathcal{B}$ measurable mapping if f^{-1} (B) ∈ $\mathcal{U}$ for every B ∈ $\mathcal{B}$, that is, f^{-1} ($\mathcal{B}$) ⊂ $\mathcal{U}$.

Measurable Space: A measurable space is a set S, together with a non-empty collection, S, of subsets of S.

Null Set in a Measure Space: A subset E of X is called a null set with respect to the measure μ if E ∈ $\mathcal{U}$ and μ (E) = 0. In this case we say also that E is a null set in the measure space (X, $\mathcal{U}$, μ).

Sigma Algebra: $\mathcal{F}$ is sigma algebra which establishes following relations:

(i) $A_k \in \mathcal{F}$ for all k implies $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$

(ii) $A \in \mathcal{F}$ implies $A^C \in \mathcal{F}$

(iii) $\phi \in \mathcal{F}$

Notes

9.4 Review Questions

1. Let $\mathcal{U}$ be a σ-algebra of subsets of a set X and let Y be an arbitrary subset of X. Let $\mathcal{B} = \{A \cap Y : A \in \mathcal{U}\}$. Show that $\mathcal{B}$ is a σ-algebra of subsets of Y.
2. Let $(X, \mathcal{U}, \mu)$ be a measure space. Show that for any $E_1, E_2 \in \mathcal{U}$ we have the equality: $\mu(E_1 \cup E_2) + \mu(E_1 \cap E_2) = \mu(E_1) + \mu(E_2)$.

9.5 Further Readings

Books

Paul Halmos, (1950). *Measure Theory*. Van Nostrand and Co.

Bogachev, V.I. (2007), *Measure Theory*, Berlin : Springer

Online links

planetmath.org/measurable space.html

mathworld.wolfram.com > Calculus and Analysis > Measure Theory

Notes

Unit 10: Measurable Functions

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand measurable functions.
- Define equivalent functions and characteristic function.
- Describe Egoroff's theorem and Riesz theorem.
- Define simple function and step function.

Introduction

In this unit, we shall see that a real valued function may be Lebesgue integrable even if the function is not continuous. In fact, for the existence of a Lebesgue integral, a much less restrictive condition than continuity is needed to ensure integrability of f on [a, b]. This requirement gives rise to a new class of functions, known as measurable functions. The class of measurable functions plays an important role in Lebesgue theory of integration.

Notes

10.1 Measurable Functions

10.1.1 Lebesgue Measurable Function/Measurable Function

Definition: Let E be a measurable set and R* be a set of extended real numbers. A function $f : E \to R^*$ is said to be a Lebesgue measurable function on E or a measurable function on E iff the set

$E(f > \alpha) = \{x \in E : f(x) > \alpha\} = f^{-1}\{\alpha, \infty)\}$ is a measurable subset of E $\forall\ \alpha \in R$.

Notes

1. The definition states that f is a measurable function if for every real number α, the inverse image of (α, ∞) under f is a measurable set.
2. The measure of the set $E(f > \alpha)$ may be finite or infinite.
3. A function whose values are in the set of extended real numbers is called an extended real valued function.
4. If E = R, then the set $E(f > \alpha)$ becomes an open set.

Example: A constant function with measurable domain is measurable.

Sol: Let f be a constant function defined over a measurable set E so that $f(x) = \in\ \forall x \in E$.

Then for any real number α,

$$E(f > \alpha) = \begin{cases} E, \text{if } c > \alpha \\ \phi, \text{if } c \le \alpha \end{cases}$$

The sets E and ϕ are measurable and hence $E(f > \alpha)$ is measurable i.e. the function f is measurable.

Theorem 1: Let f and g be measurable real valued functions on E, and c is a constant. Then each of the following functions is measurable on E.

(a) $f \pm c$ (b) $c f$

(c) $f + g$ (d) $f - g$

(e) $|f|$ (f) f^2

(g) fg (h) f/g (g vanishes no where on E)

Proof: Let α be an arbitrary real number.

(a) Since f is measurable and

$E(f \pm c > \alpha) = E(f > \alpha \mp c)$,

the function $f \pm c$ is measurable.

(b) To prove c f is measurable over E.

If c = 0, then cf is constant and hence measurable because a constant function is measurable.

Notes

Consider the case in which $c \neq 0$, then

$$E\,(cf > \alpha) = \begin{cases} E\left(f > \dfrac{\alpha}{c}\right) \text{if } c > 0 \\ E\left(f < \dfrac{\alpha}{c}\right) \text{if } c < 0 \end{cases}$$

Both the sets on R.H.S. are measurable.

Hence $E\,(cf > \alpha)$ is measurable and so cf is measurable $\forall\, c \in R$.

(c) Before proving f + g is measurable, we first prove that if f and g are measurable over E then the set E (f > g) is also measurable.

Now $f > g \Rightarrow \exists$ a rational number r such that

$$f\,(x) > r > g\,(x)$$

Thus $$E\,(f > g) = \bigcup_{r \in Q} [(E\,(f > r) \cap (E(g < r))]$$

= an enumerable union of measurable sets.

= measurable set, since Q is an enumerable set.

Now, we shall prove that f + g is measurable over E. Let a be any real number.

Now $$E\,(f + g > a) = E\,(f > a - g) \qquad \ldots (1)$$

Again, g is measurable

$\Rightarrow$cg is measurable, c is constant.

($\because$ We know that if f is a measurable function and c is constant then cf is measurable)

$\Rightarrow$ a + cg is measurable $\forall$ a, $c \in R$

$\Rightarrow$ a – g is measurable by taking $c' = -1$,

since f and a – g are measurable

$\Rightarrow$ E (f > a – g) is measurable.

$\Rightarrow$ E (f + g > a) is a measurable set.

$\Rightarrow$ f + g is a measurable function.

(d) To prove that f – g is measurable. Before proving f – g is measurable, we first prove that if f and g are measurable over E then the set E (f > g) is also measurable.

Now $f > g \Rightarrow \exists$ a rational number r, such that $f\,(x) > r > g\,(x)$.

Thus $$E\,(f > g) = \bigcup_{r \in Q} [(E\,(f > r) \cap (E(g < r))]$$

= an enumerable union of measurable sets.

= measurable sets, since Q is an enumerable set.

Now we shall prove that f – g is measurable over E.

Let a be any real number.

Now $E(f - g > a) = E(f > a + g)$ **Notes**

since g is measurable.

$\Rightarrow$ cg is measurable, c is constant.

$\Rightarrow$ a + cg is measurable $\forall$ a, c $\in$ R

$\Rightarrow$ a + g is measurable by taking c = 1,

since f and a + g are measurable

$\Rightarrow$ E (f > a + g) is measurable.

$\Rightarrow$ E (f – g > a) is a measurable set.

$\Rightarrow$ f – g is a measurable function.

(e) To prove $|f|$ is measurable.

We have

$$E(|f| > \alpha) = \begin{cases} E \text{ if } \alpha < 0 \\ [E(f > \alpha)] \cup [E(f < -\alpha)] \text{ if } \alpha \geq 0 \end{cases}$$

[because we know that $|x| > a \Rightarrow x > a$ or $x < -a$]

since f is measurable therefore E (f > α) and E (f < – α) are measurable by definition.

Also we know that finite union of two measurable sets is measurable.

$\Rightarrow$ E (f > α) $\cup$ E (f < – α) is measurable.

$\Rightarrow$ E ($|f|$ > α) is measurable.

$\Rightarrow$ $|f|$ is measurable.

(f) To prove f^2 is measurable.

We have $E(f^2 > \alpha) = \begin{cases} E \text{ if } \alpha < 0 \\ E(|f| > \sqrt{\alpha})] \text{ if } \alpha \geq 0 \end{cases}$

But $E(|f| > \sqrt{\alpha}) = [E(f > \sqrt{\alpha})] \cup [E(f < -\sqrt{\alpha})]$, if $\alpha \geq 0$ ($\because$ $|x| > a \Rightarrow x > a$ or $x < -a$)

$\therefore$ $E(f^2 > \alpha) = \begin{cases} E \text{ if } \alpha < 0 \\ [E(f > \sqrt{\alpha})] \cup [E(f < -\sqrt{\alpha})] \text{ if } \alpha \geq 0 \end{cases}$

But f is measurable over E.

$\Rightarrow$ $E(f > \sqrt{\alpha})$ and $E(f < -\sqrt{\alpha})$ are measurable sets.

$\Rightarrow$ $[E(f > \sqrt{\alpha})] \cup [E(f < -\sqrt{\alpha})]$ is measurable.

($\because$ union of two measurable sets is measurable)

$\Rightarrow$ E (f^2 > α) is measurable because both the sets on RHS are measurable.

$\Rightarrow$ f^2 is measurable over E.

(g) To prove fg is measurable.

Clearly, f + g and f – g are measurable functions over E.

Notes

$\Rightarrow$ $(f + g)^2$, $(f - g)^2$ are measurable functions over E.

$\Rightarrow$ $(f + g)^2 - (f - g)^2$ is a measurable function over E.

$\Rightarrow$ $\frac{1}{4}\left[(f+g)^2 - (f-g)^2\right]$ is a measurable function over E.

$\Rightarrow$ fg is a measurable function over E.

(h) To prove f/g is measurable.

Let g vanish nowhere on E, so that

$g(x) \neq 0 \quad \forall\ x \in E.$

$\Rightarrow$ $\frac{1}{g}$ exists.

Now we shall show that $\frac{1}{g}$ is measurable.

We have

$$E\left(\frac{1}{g} > \alpha\right) = \begin{cases} E(g > 0) \text{ if } \alpha = 0 \\ [E(g < 0)] \cap \left[E\left(g < \frac{1}{\alpha}\right)\right] \text{ if } \alpha > 0 \\ [E(g < 0)] \cup [E(g < 0)] \cap \left[E\left(g < \frac{1}{\alpha}\right)\right] \end{cases}$$

Also finite union and intersection of measurable sets are measurable. Hence $E\left(\frac{1}{g} > \alpha\right)$ is measurable in every case.

Since f and $\frac{1}{g}$ are measurable.

$\Rightarrow$ $(f)\Big/\left(\frac{1}{g}\right)$ is measurable over E.

$\Rightarrow$ $\frac{f}{g}$ is measurable over E.

10.1.2 Almost Everywhere (a.e.)

Definition: A property is said to hold almost everywhere (a.e.) if the set of points where it fails to hold is a set of measure zero.

Example: Let f be a function defined on R by

$$f(x) = \begin{cases} 0, & \text{if x is irrational} \\ 1, & \text{if x is rational} \end{cases}$$

Then f (x) = 0 a.e.

Notes

10.1.3 Equivalent Functions

Definition: Two functions f and g defined on the same domain E are said to be equivalent on E, written as $f \sim g$ on E, if $f = g$ a.e. on E, i.e. $f(x) = g(x)$ for all $x \in E - E_1$, where $E_1 \subset E$ with $m(E_1) = 0$.

Theorem 2: If $f, g : E \to R$ $(E \in M)$ such that $g \in \Omega(E)$.

Proof: Let α be any real number and let $E_1 = E(f > \alpha)$ and $E_2 = E(g > \alpha)$

Then
$$E_1 \Delta E_2 = (E_1 - E_2) \cup (E_2 - E_1)$$
$$= \{x \in E : f(x) \neq g(x)\}$$

so that by given hypothesis we have

$$m(E_1 \Delta E_2) = 0.$$

This together with the fact that E_1 is measurable

$\Rightarrow$ E_2 is measurable.

Hence $g \in \Omega(E)$.

10.1.4 Non-negative Functions

Definition: Let f be a function, then its positive part, written f^+ and its negative part, written f^{-1}, are defined to be the non-negative functions given by

$$f^+ = \max(f, 0) \text{ and } f^{-1} = \max(-f, 0) \text{ respectively.}$$

Note

$$f = f^+ - f^{-1}$$
$$\text{and } |f| = f^+ + f^{-1}$$

Theorem 3: A function is measurable iff its positive and negative parts are measurable.

Proof: For every extended real valued function f, we may write

$$f^+ = \frac{1}{2}[f + |f|]$$

and
$$f^{-1} = \frac{1}{2}[|f| - f]$$

But f is measurable then $|f|$ is measurable and hence positive and negative parts of f i.e. f^+ and f^- are measurable.

Conversely, let f^+ and f^{-1} be measurable.

Since
$$f = f^+ - f^{-1}$$

Since we know that if f and g are measurable functions defined on a measurable set E then $f - g$ is measurable on E.

Here $f^+ - f^{-1}$ is measurable.

and hence f is measurable.

Notes

Theorem 4: If f is a measurable function and f = g a.e. then g is measurable.

Proof: Let E = {x : f (x) ≠ g (x)}.

Then m (E) = 0

Let α be a real number.

$$\{x : g(x) > \alpha\} = \{x : f(x) > \alpha\} \cup \{x \in E : g(x) > \alpha\} - \{x \in E : g(x) \le \alpha\}$$

since f is measurable, the first set on the right is measurable i.e. {x : f (x) > α} is measurable.

The last two sets on the right are measurable since they are subsets of E and m (E) = 0.

Thus, {x : g (x) > α} is measurable.

So, g is measurable.

Example: Give an example of function for which f is not measurable but |f| is measurable.

Sol: Let k be a non-measurable subset of E = [0, 1).

Define a function f : E → R by

$$f(x) = \begin{cases} 1 \text{ if } x \in k \\ -1 \text{ if } x \notin k \end{cases}$$

The function f is not measurable, since E (f > 0) (=k) is a non-measurable set. But |f| is measurable as the set

$$E(|f| > \alpha) = \begin{cases} E \text{ if } \alpha < 1 \\ \phi \text{ if } \alpha \ge 1 \end{cases} \text{ is measurable}$$

10.1.5 Characteristic Function

Definition: Let A be subset of real numbers. We define the characteristic function χ_A of the set A as follows:

$$\chi_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases}$$

Note The characteristic function χ_A of the set A is also called the indicator function of A.

Theorem 5: Show that the characteristic function χ_A is measurable iff A is measurable.

Proof: Let χ_A be measurable.

Since A = {x : χ_A (x) > 0} is measurable.

But χ_A is measurable, therefore the set {x : χ_A (x) > 0} is measurable.

⇒ A is measurable.

Conversely, let A be measurable and α be any real number.

$$\text{then } E(\chi_A > \alpha) = \begin{cases} \phi \text{ if } \alpha \ge 1 \\ A \text{ if } 0 \le \alpha < 1 \\ E \text{ if } \alpha < 0 \end{cases}$$

Notes

Every set on R.H.S. is measurable.

Therefore E ($\chi_A > \alpha$) is measurable.

Hence χ_A is measurable.

Note The above theorem asserts that the characteristic function of non-measurable sets are non-measurable even though the domain set is measurable.

10.1.6 Simple Function

A real valued function ϕ is called simple if it is measurable and assumes only a finite number of values.

If ϕ is simple and has the values $\alpha_1, \alpha_2, \ldots \alpha_n$, then

$$\phi = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$$

where $A_i = \{x : \phi(x) = \alpha_i\}$

and $A_i \cap A_j$ is a null set.

Thus we can always express a simple function as a linear combination of characteristic function.

Notes

(i) ϕ is simple $\Leftrightarrow$ A'_is are measurable.

(ii) sum, product and difference of simple functions are simple.

(iii) the representation of ϕ as given above is not unique.

But if ϕ is simple and $\{\alpha_1, \alpha_2, \ldots\ldots, \alpha_n\}$ is the set of non-zero values of f, then

$$\phi = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$$

where $A_i = \{x : \phi(x) = \alpha_i\}$

This representation of ϕ is called the *Canonical representation*. Here A'_is are disjoint and α'_is are distinct and non-zero.

(iv) Simple function is always measurable.

10.1.7 Step Function

A real valued function S defined on an interval [a, b] is said to be a step function if these is a partition $a = x_o < x_1 \ldots < x_n = b$ such that the function assumes one and only one value in each interval.

Notes

Notes

(i) Step function also assumes finite number of values like simple functions but the sets $\{x : S(x) = C_i\}$ are intervals for each i.

(ii) Every step function is also a simple function but the converse is not true.

e.g. $f : R \to R$ such that $f(x) = \begin{cases} 1, x \text{ is rational} \\ 0, x \text{ is irrational} \end{cases}$

is a simple function but not step as the sets of rational and irrational are not intervals.

Theorem 6: If f and g are two simple functions then $\alpha f + \beta g$ is also a simple function.

Proof: Since f and g are simple functions and we know that every simple function can be expressed as the linear combination of characteristic function.

$\therefore$ f and g can be expressed as the linear combination of characteristic function.

$$\therefore \quad f = \sum_{i=1}^{m} \alpha_i \chi_{A_i}$$

and $$g = \sum_{j=1}^{m'} \beta_j \chi_{B_j}$$

where A'_is and B'_js are disjoint.

$$A_i = \{x : f(x) = \alpha_i\}$$

$$B_j = \{x : g(x) = \beta_j\}$$

The set E_k obtained by taking all intersections $A_i \cap B_j$ from a finite disjoint collection of measurable sets and we may write

$$f = \sum_{k=1}^{n} a_k \chi_{E_k}$$

and $$g = \sum_{k=1}^{n} b_k \chi_{E_k}$$

where $$n = mm'.$$

$$\alpha f + \beta g = \alpha \sum_{k=1}^{n} a_k \chi_{E_k} + \beta \sum_{k=1}^{n} b_k \chi_{E_k}$$

$$= \sum_{k=1}^{n} (\alpha a_k + \beta_{b_k}) \chi_{E_k}$$

which is a *linear combination* of characteristic functions, therefore it is simple.

Notes

Similarly $fg = \sum_{k=1}^{n} a_k + b_k \chi_{E_k}$

which is again a linear combination of characteristic function, therefore fg is simple.

Theorem 7: Let E be a measurable set with m (E) < ∞ and $\{f_n\}$ a sequence of measurable functions converging a.e. to a real valued function defined on E. Then, given $\varepsilon > 0$ and $\delta > 0$, there is a set $A \subset E$ with m (A) < δ and an integer N such that $|f_n(x) - f(x)| < \varepsilon$ for all $x \in E - A$ and all $n \geq N$.

Proof: Let F be the set of points of E for which $f_n \to f$. Then m (F) = 0 and $f_n(x) \to f(x)$ for all $x \in E - F = E_1$ (say). Then by the previous theorem for the set E_1, we get $A_1 \subset E_1$ with $m(A_1) < \delta$ and an integer N such that

$$|f_n(x) - f(x)| < \varepsilon \text{ for all } n \geq N \text{ and } x \in E_1 - A_1.$$

We get the required result by taking

$$A = A_1 \cup F \text{ since } m(F) = 0 \text{ and } E - A = E_1 - A_1$$

Note Before proving this theorem first prove the previous theorem.

10.1.8 Convergent Sequence of Measurable Function

Definition: A sequence $\{f_n\}$ of measurable functions is said to converge almost uniformly to a measurable function f defined on a measurable set E if for each $\in > 0$ there exists a measurable set $A \subset E$ with $m(A) < \in$ such that $\{f_n\}$ converges to f uniformly an E – A.

10.1.9 Egoroff's Theorem

Statement: Let E be a measurable set with m (E) < ∞ and $\{f_n\}$ a sequence of measurable functions which converge to f a.e. on E. Then, given $\eta > 0$ there is a set $A \subset E$ with $m(A) < \eta$ with that the sequence $\{f_n\}$ converges to f uniformly on E – A.

Proof: Applying the theorem, "Let E be a measurable set with m (E) < ∞ and $\{f_n\}$ a sequence of measurable function converging a.e. to real valued function f defined on E. Then given $\varepsilon > 0$ and $\delta > 0$ there is a set $A \subset E$ with m (A) < δ and an integer N such that

$$|f_n(x) - f(x)| < \varepsilon \text{ for all } x \in E - A \text{ and all } n \geq N"$$

with $\varepsilon = 1$, $\delta = \eta/2$, we get a measurable set

$A_1 \subset E$ with $m(A_1) < \eta/2$ and a positive integer N_1, such that

$$|f_n(x) - f(x)| < 1 \text{ for all } x \geq N_1$$

and $x \in E_1$, where $E_1 = E - A_1$.

Again, taking $\varepsilon = 1/2$ and $\delta = n/2^2$,

we get a measurable set $A_2 \subset E_1$ with $m(A_2) < \eta/2^2$, and a positive integer N_2 such that

$|f_n(x) - f(x)| < \frac{1}{2}$ $\forall\, n \geq N_2$ and $x \in E_2$ where $E_2 = E_1 - A_2$, and so on.

Notes

At the pth stage, we get a measurable set

$A_p \subset E_{p-1}$ with $m(A_p) < \frac{n}{2^p}$ and a positive integer N_p such that

$|f_n(x) - f(x)| < \frac{1}{p}$ $\forall\, n \geq N_p$ and $x \in E_p$ where

$$E_p = E_p - 1 - A_p.$$

Let $$A = \sum_{p=1}^{\infty} A_p\,,$$

then $$m(A) \leq \sum_{p=1}^{\infty} m(A_p)$$

But $$m(A_p) < \frac{n}{2p}$$

$\therefore$ $$m(A) < \sum_{p=1}^{\infty} \frac{n}{2^p}$$

But $\sum_{p=1}^{\infty} \frac{1}{2^p}$ is a G.P. series so $S_\infty = \frac{a}{1-r} = \frac{\frac{1}{2}}{1-\frac{1}{2}} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1$

$\therefore$ $$m(A) < \eta$$

Also, $$E - A = E - \bigcup_p A_p$$

$$= \bigcap_p (E - A_p)$$

$$= \bigcap_p (E_{p-1} - A_p)$$

$$= \bigcap_p E_p$$

Let $x \in E - A$. Then $x \in E_p$ $\forall\, p$ and so

$$|f_n(x) - f(x)| < \frac{1}{p}\,,\ \forall\, n \geq N_p.$$

Let us choose p such that $\frac{1}{p} < \varepsilon$, we get

$$|f_n(x) - f(x)| < \varepsilon\ \forall\, x \in E - A \text{ and } n \geq N_p = N$$

Notes

Note Egoroff's theorem can be stated as: almost every where convergence implies almost uniform convergence.

10.1.10 Riesz Theorem

Let $\{f_n\}$ be a sequence of measurable functions which converges in measure to f. Then there is a subsequence $\{f_{n_k}\}$ which converges in measure to f a.e.

Proof: Let $\{\varepsilon_n\}$ and $\{\delta_n\}$ be two sequences of positive real numbers such that $\varepsilon_n \to 0$ as $n \to \infty$ and $\sum_{n=1}^{\infty} \delta_n < \infty$.

Let us now choose an increasing sequence $\{n_k\}$ of positive integers as follows.

Let n be a positive integer such that

$$m\left(\left\{x \in E \left| f_n(x) - f(x) \right| \geq \varepsilon\right\}\right) < \delta$$

Since $f_n \to f$ in measure for a given $\varepsilon_1 > 0$ and $\delta_1 > 0$, $\exists$ a positive integer n_1 such that

$$m\left(\left\{x \in E, \left|f_{n_1}(x) - f(x)\right| \geq \varepsilon_1\right\}\right) < \delta_1, \forall\, n_1 \geq n$$

Similarly, let n_2 be a positive number such that $m\left(\left\{x \in E, \left|f_{n_2}(x) - f(x)\right| \geq \varepsilon_2\right\}\right) < \delta_2, \forall\, n_2 \geq n_1$ and so on.

In general let n_k be a positive number such that

$$m\left(\left\{x : x \in E, \left|f_{n_k}(x) - f(x)\right| \geq \varepsilon_k\right\}\right) < \delta_k \text{ and that } n_k \geq n_{k-1}.$$

We shall now prove that the subsequence $\{f_{n_k}\}$ converges to f a.e.

Let $A_k = \bigcup_{i=k}^{\infty} \left\{x : x \in E, \left| f_{n_i}(x) - f(x)\right| \geq \varepsilon_i\right\}, k \in N$ and $A = \bigcap_{k=1}^{\infty} A_k$.

Clearly, $\{A_k\}$ is a decreasing sequence of measurable sets and $m(A_1) < \infty$.

Therefore, we have

$$m(A) = \lim_{k \to \infty} m(A_k)$$

But $m(A_k) \leq \sum_{i=k}^{\infty} \delta_i \to 0$ as $k \to \infty$.

Hence m (A) = 0.

Now it remains to show that $\{f_n\}$ converges to f on E – A. Let $x_o \in$ E – A.

Notes

Then $x_o \notin A_{k_o}$ for some positive integer k_o.

or $\quad x_o \notin \{x \in E : |f_n(x) - f(x)| \geq \varepsilon_k\}, k \geq k_o$

which gives $|f_n(x_o) - f(x_o)| < \varepsilon_k, k \geq k_o$

But $\quad \varepsilon_k \to 0$ as $k \to \infty$.

Hence $\quad \lim_{k\to\infty} f_{n_k}(x_o) = f(x_o)$.

10.2 Summary

- Let E be a measurable set and R* be a set of extended real numbers. A function $f : E \to R^*$ is said to be a Lebesgue measurable function on E or a measurable function on E iff the set E $(f > \alpha) = \{x \in E : f(x) > \alpha\} = f^{-1}\{(\alpha, \infty)\}$ is a measurable subset of E $\forall\, \alpha \in R$.
- A property is said to hold almost everywhere (a.e.) if the set of points where it fails to hold is a set of measure zero.
- Two functions f and g defined on the same domain E are said to be equivalent on E, written as $f \sim g$ on E, if $f = g$ a.e. on E, i.e. $f(x) = g(x)$ for all $x \in E - E_1$, where $E_1 \subset E$ with $m(E_1) = 0$.
- $f^+ = \max(f, 0)$ and $f^{-1} = \max(-f, 0)$

  $|f| = f^+ + f^{-1}$
- Let A be subset of real numbers. We define the characteristic function χ_A of the set A as follows:

  $$\chi_A(x) = \begin{cases} 1, \text{ if } x \in A \\ 0, \text{ if } x \notin A \end{cases}$$
- A real valued function ϕ is called simple if it is measurable and assumes only a finite number of values.

10.3 Keywords

Almost Everywhere (a.e.): A property is said to hold almost everywhere (a.e.) if the set of points where it fails to hold is a set of measure zero.

Characteristic Function: Let A be subset of real numbers. We define the characteristic function χ_A of the set A as follows:

$$\chi_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases}$$

Egoroff's Theorem: Let E be a measurable set with $m(E) < \infty$ and $\{f_n\}$ a sequence of measurable functions which converge to f a.e. on E. Then given $n > 0$ there is a set $A \subset E$ with $m(A) < n$ such that the sequence $\{f_n\}$ converges to f uniformly on $E - A$.

Equivalent Functions: Two functions f and g defined on the same domain E are said to be equivalent on E, written as $f \sim g$ on E, if $f = g$ a.e. on E, i.e. $f(x) = g(x)$ for all $x \in E - E_1$, where $E_1 \subset E$ with $m(E_1) = 0$.

Notes

Haracteristic Function: Let A be subset of real numbers. We define the characteristic function χ_A of the set A as follows:

$$\chi_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases}$$

Lebesgue Measurable Function: A function $f : E \to R^*$ is said to be a Lebesgue measurable function on E or a measurable function on E iff the set

$E(f > \alpha) = \{x \in E : f(x) > \alpha\} = f^{-1}\{\alpha, \infty)\}$ is a measurable subset of E $\forall\ \alpha \in R$.

Measurable Set: A set E is said to be measurable if for each set T, we have

$$m^*(T) = m^*(T \cap E) + m^*\{T \cap E^c)$$

Non-negative Functions: Let f be a function, then its positive part, written f^+ and its negative part, written f^{-1}, are defined to be the non-negative functions given by

$$f^+ = \max(f, 0) \text{ and } f^{-1} = \max(-f, 0) \text{ respectively.}$$

Riesz Theorem: Let $\{f_n\}$ be a sequence of measurable functions which converges in measure to f. Then there is a subsequence $\{f_{n_k}\}$ which converges to f a.e.

Simple Function: A real valued function ϕ is called simple if it is measurable and assumes only a finite number of values.

If ϕ is simple and has the values $\alpha_1, \alpha_2, \ldots \alpha_n$, then

$$\phi = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$$

where $A_i = \{x : \phi(x) = \alpha_i\}$

and $A_i \cap A_j$ is a null set.

Step Function: A real valued function S defined on an interval [a, b] is said to be a step function if these is a partition $a = x_o < x_1 \ldots < x_n = b$ such that the function assumes one and only one value in each interval.

Subsequence: If (x_n) is a given sequence in X and (n_k) is an strictly increasing sequence of positive integers, then $\{x_{n_k}\}$ is called a subsequence of (x_n).

10.4 Review Questions

1. If f is a measurable function and c is a real number, then is it true to say that cf is measurable?
2. A non-zero constant function is measurable if and only if X is measurable comment.
3. Let Q be the set of rational number and let f be an extended real-valued function such that $\{x : f(x) > \alpha\}$ is measurable for each $\alpha \in Q$. Then show that f is measurable.
4. Show that if f is measurable then the set $\{x : f(x) = \alpha\}$ is measurable for each extended real number α.
5. If f is a continuous function and g is a measurable function, then prove that the composite function fog is measurable.

Notes

6. Show that

(i) $\chi_{A\cap B} = \chi_A \cdot \chi_B$

(ii) $\chi_{A\cup B} = \chi_A + \chi_B - \chi_A \cdot \chi_B$

(iii) $\chi_{A^c} = 1 - \chi_A$

10.5 Further Readings

Books

Dudley, R.M. (2002). *Real Analysis and Probability (2 ed.)*. Cambridge University Press

Strichartz, Robert (2000). *The Way of Analysis*. Jones and Bortlett.

Online links

mathworld.wolfram.com>calculus and Analysis > Measure theory

planetmath.org/Measurable functions.html

zeta.math.utsa.edu/~ mqr 328/class/real2/mfunct.pdf

Notes

Unit 11: Integration

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define the Riemann integral and Lebesgue integral of bounded function over a set of finite measure.
- Understand the Lebesgue integral of a non-negative function.
- Solve problems on integration.

Introduction

We now come to the main use of measure theory: to define a general theory of integration. The particular case of the integral with respect to the Lebesgue measure is not, in any way, simpler the general case, which will give us a tool of much wider applicability.

11.1 Integration

11.1.1 The Riemann Integral

Let f be a bounded real valued function defined on the interval [a, b] and let $a = x_0 < x_1 < \ldots < x_n = b$ be a sub-division of [a, b].

Then for each sub-division we can define the sums

$$S = \sum_{i=1}^{n} (x_i - x_{i-1}) M_i$$

and

$$s = \sum_{i=1}^{n} (x_i - x_{i-1}) m_i$$

Notes

where $$M_i = \sup_{x_{i-1}<x\le x_i} f(x),$$

$$m_i = \inf_{x_{i-1}<x\le x_i} f(x)$$

We then define the *upper Riemann integral* of f by

$$R\overline{\int_a^b} f(x)\,dx = \inf S$$

where the infimum is taken over all possible sub-divisions of [a, b].

Similarly, we define the *lower Riemann integral*

$$R\underline{\int_a^b} f(x)\,dx = \sup S$$

The upper integral is always at least as large as the lower integral, and if the two are equal, we say that f is *Riemann integrable* and we call this common value the Riemann integral of f.

It will be denoted by

$$R\int_a^b f(x)\,dx$$

Note By a *step function* we mean function Ψ s.t.

$$\Psi(x) = \alpha_i \;\forall\; x \in [x_{i-1}, x_i]$$

for some sub-division of [a, b] and some set of constant α_i then

$$\int_a^b \Psi(x)\,dx = \int_{x_o}^{x_1} \Psi(x)\,dx + \ldots + \int_{x_{n-1}}^{x_n} \Psi(x)\,dx$$

$$= \int_{x_o}^{x_1} \alpha_1\,dx + \int_{x_1}^{x_2} \alpha_2\,dx + \ldots + \int_{x_{n-1}}^{x_n} \alpha_n\,dx$$

$$= \alpha_1(x_1 - x_0) + \alpha_2(x_2 - x_1) + \ldots + \alpha_n(x_{n-1} - x_n)$$

$$= \sum_{i=1}^{n} \alpha_i(x_i - x_{i-1}) \qquad \ldots (1)$$

with this in mind, we see that

$$R\overline{\int_a^b} f(x)\,dx = \inf \cup_p (f)$$

Notes

$$= \inf \sum_{i=1}^{n} M_i (x_i - x_{i-1})$$

$$= \inf \int_a^b \Psi(x)\, dx \quad \text{for all step functions}$$

$$\Psi(x) \geq f(x)$$

Similarly $$R\int_a^b f(x)\, dx = \sup L_p(f)$$

$$= \sup \sum_{i=1}^{n} m_i (x_i - x_{i-1})$$

$$= \sup \int_a^b \phi(x)\, dx \quad \text{for all step function}$$

$$\phi(x) \leq f(x).$$

11.1.2 Lebesgue Integral of a Bounded Function over a Set of Finite Measure

Characteristic Function

The function χ_E defined by

$$\chi_E(x) = \begin{cases} 1 \text{ if } x \in E \\ 0 \text{ if } x \notin E \end{cases}$$

is called the *characteristic function* of E.

Simple Function

A linear combination $\phi(x) = \sum_{i=1}^{n} \alpha_i \chi_{E_i}(x)$ is called a *simple function* if the sets E_i are measurable.

This representation of ϕ is not unique.

However, a function ϕ is simple if and only if it is measurable and assumes only a finite number of values.

Canonical Representation

If ϕ is simple function and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ the set of non-zero values of ϕ, then

$$\phi = \sum_{i=1}^{n} \alpha_i \chi_{E_i},$$

where $E_i = \{x : \phi(x) = \alpha_i\}$.

Notes

This representation of ϕ is called the *canonical representation*. Here E_i's are disjoint and α_i's are finite in number, distinct and non-zero.

Elementary Integral

Definition: If ϕ vanishes outside a set of finite measure, we define the elementary integral of ϕ by $\int \phi(x)\,dx = \sum_{i=1}^{n} \alpha_i\, mE_i$ when ϕ has the canonical representation.

$$\phi = \sum_{i=1}^{n} \alpha_i\, \chi_{E_i}\ .$$

We sometimes abbreviate the expression for this integral $\int \phi$. If E is any measurable set, we define the elementary integral of ϕ over E by $\int_E \phi = \int \phi \cdot \chi_E$.

If E = [a, b], then the integral $\int_{[a,b]} \phi$ will be denoted by $\int_a^b \phi$.

Theorem 1: Let ϕ and Ψ be simple functions which vanish outside a set of finite measure, then

$$\int (a\phi + b\Psi) = a\int \phi + b\int \Psi \text{ and if } \phi \geq \Psi \text{ a.e., then } \int \phi \geq \int \Psi$$

Proof: Since ϕ and Ψ are simple functions.

Therefore these can be written in the canonical form

$$\phi = \sum^{m} \alpha_i\, \chi_{A_i}$$

and

$$\Psi = \sum_{j=1}^{m'} B_j\, \chi_{B_j}$$

where $\{A_i\}$ and $\{B_j\}$ are disjoint sequences of measurable sets and

$$A_i = \{x : \phi(x) = \alpha_i\}$$

and

$$B_j = \{x : \Psi(x) = \beta_j\}$$

The set E_k obtained by taking all intersections $A_i \cap B_j$ form a finite disjoint collection of measurable sets. We may write

$$\phi = \sum_{k=1}^{N} a_k\, \chi_{E_k} \text{ and}$$

$$\Psi = \sum_{k=1}^{N} b_k\, \chi_{E_k} \text{ (where } N = mm')$$

Notes

Now $$a\phi + b\Psi = a\sum_{k=1}^{N} a_k\,\chi_{E_k} + b\sum_{k=1}^{N} b_k\,\chi_{E_k}$$

$$= \sum_{k=1}^{N}(aa_k + bb_k)\,\chi_{E_k}$$

which is again a simple function.

Since $$\int\phi = \sum_i a_i\, m\, E_i$$

$\therefore$ $$\int(a\phi + b\Psi) = \sum_{k=1}^{N}(aa_k + bb_k)\, m\, E_k \quad \text{(by definition)}$$

$$= a\sum_{k=1}^{N} a_k\, m E_k + b\sum_{k=1}^{N} b_k\, m E_k$$

$$= a\int\phi + b\int\Psi$$

Now since $\phi \geq \Psi$ a.e.

$\Rightarrow$ $\phi - \Psi \geq 0$ a.e.

We have proved that

$$\int(a\phi + b\Psi) = a\int\phi + b\int\Psi$$

Put a = 1, b = –1 in the first part, we get

$$\int(\phi - \Psi) = \int\phi - \int\Psi$$

Since $\phi - \Psi \geq 0$ a.e. is a simple function, by the definition of the elementary integral, we have

$$\int(\phi - \Psi) \geq 0$$

$\Rightarrow$ $$\int\phi - \int\Psi \geq 0$$

$\Rightarrow$ $$\int\phi \geq \int\Psi$$

Theorem 2: Riemann integrable is Lebesgue integrable.

Proof: Since f is Riemann integrable over [a, b], we have

$$\inf_{\Psi_1 \geq f}\int_a^b \Psi_1(x)\,dx = \sup_{\phi \leq f}\int_a^b \phi_1(x)\,dx = R\int_a^b f(x)\,dx$$

where ϕ_1 and Ψ_1 vary over all step functions defined on [a, b].

Notes

Since we know that every step function is a simple function,

$$\therefore \qquad \sup_{\phi_1 \le f} \int_a^b \phi_1(x)\,dx \le \sup_{\phi \le f} \int_a^b \phi(x)\,dx$$

and
$$\inf_{\Psi_1 \ge f} \int_a^b \Psi_1(x)\,dx \ge \inf_{\Psi \ge f} \int_a^b \Psi(x)\,dx$$

where ϕ and Ψ vary over all the simple functions defined on [a, b]. Thus from the above relation, we have

$$R\int_a^b f(x)\,dx \le \sup_{\phi \le f} \int_a^b \phi(x)\,dx \le \inf_{\Psi \ge f} \int_a^b \Psi(x)\,dx \le R\int_a^b f(x)\,dx$$

$$\Rightarrow \qquad \sup_{\phi \le f} \int_a^b \phi(x)\,dx = \inf_{\Psi \ge f} \int_a^b \Psi(x)\,dx$$

$$\Rightarrow \qquad \int_a^b f(x)\,dx = R\int_a^b f(x)\,dx$$

Note The converse of this theorem is not true i.e.

A Lebesgue integrable function may not be Riemann integrable

e.g. Let f be a function defined on the interval [0, 1] as follows:

$$f(x) = \begin{cases} 1, \text{ if } x \text{ is rational} \\ 0, \text{ if } x \text{ is irrational} \end{cases}$$

Let us consider a partition p of an interval [0, 1].

$$U(p, f) = \sum_{i=1}^{N} M_i\,\Delta x_i$$

$$= 1\Delta x_1 + 1\Delta x_2 + \ldots\ldots + 1\Delta x_n = 1 - 0$$

$$= 1.$$

$$\therefore \qquad \overline{\int_0^1} f\,dx = \inf U(p, f) = 1 - 0 = 1.$$

$$\underline{\int_0^1} f\,dx = \sup L(p, f)$$

$$= \sup\{0\Delta x_1 + 0\Delta x_2 + \ldots\ldots + 0\Delta x_n\}$$

$$= 0$$

Notes

Thus $$\overline{\int} f\,dx \neq \underline{\int} f\,dx$$

$\therefore$ The function is not Riemann integrable.

Now for Lebesgue Integrability

Let A_1 be the set of all irrational numbers and A_2 be the set of all rational numbers in [0, 1].

The partition $P = \{A_1, A_2\}$ is a measurable partition of [0, 1] and $mA_1 = 0$, $mA_2 = 1$.

$$L(p, f) = \inf_{A_1} f(x)\cdot mA_1 + \inf_{A_2} f(x)\cdot mA_2$$

$$= 0\cdot mA_1 + 1\cdot mA_2 = 1.$$

$$U(p, f) = \sup_{A_1} f(x)\cdot mA_1 + \sup_{A_2} f(x)\cdot mA_2$$

$$= 0\cdot mA_1 + 1\cdot mA_2 = 1.$$

$$\sup_p L(p, f) = 1 = \inf_p U(p, f)$$

$\Rightarrow$ f is Lebesgue integrable over [0, 1].

Theorem 3: If f and g are bounded measurable functions defined on the set E of finite measure, then

(1) $$\int_E (af + bg) = a\int_E f + b\int_E g$$

(2) If f = g a.e., then $\int_E f = \int_E g$

(3) If $f \le g$ a.e., then $\int_E f \le \int_E g$

Hence $\left|\int f\right| \le \int |f|$

(4) If A and B are disjoint measurable set of finite measure, then

$$\int_{A\cup B} f = \int_A f + \int_B f$$

Proof of 1: Result is true if a = 0

Let $a \neq 0$.

If Ψ is a simple function then so is $a\Psi$ and conversely.

Hence for a > 0

$$\int_E af = \inf_{a\Psi \ge af} \int_E a\Psi$$

Notes

$$= \inf_{\Psi \geq f} \int_E a\Psi \qquad (\because a > 0)$$

$$= \inf_{\Psi \geq f} a \int_E \Psi$$

$$= a \inf_{\Psi \geq f} \int_E \Psi$$

$$= a \int_E f$$

Again if $a < 0$,

$$\int_E af = \inf_{a\Psi \geq af} \int_E a\Psi$$

$$= \sup_{\Psi \leq f} \int_E a\Psi \qquad (\because a < 0)$$

$$= \sup_{\Psi \leq f} a \int_E \Psi$$

$$= a \sup_{\Psi \leq f} \int_E \Psi$$

$$= a \int_E f$$

Therefore in each case

$$\int_E af = a \int_E f \qquad \ldots \text{(i)}$$

Now we prove that

$$\int_E (f+g) = \int_E f + \int_E g$$

Let Ψ_1 and Ψ_2 be two simple functions such that $\Psi_1 > f$ and $\Psi_2 \geq g$, then $\Psi_1 + \Psi_2$ is a simple function and $\Psi_1 + \Psi_2 \geq f + g$.

or $\qquad f + g = \Psi_1 + \Psi_2$

$$\therefore \qquad \int_E (f+g) \leq \int_E (\Psi_1 + \Psi_2)$$

But $\qquad \int_E \Psi_1 + \Psi_2 = \int_E \Psi_1 + \int_E \Psi_2$

Notes

$$\Rightarrow \qquad \int_E (f+g) \le \int_E \Psi_1 + \int_E \Psi_2$$

Since $$\inf_{f \le \Psi_1} \int_E \Psi_1 = \int_E f$$

and $$\inf_{g \le \Psi_2} \int_E \Psi_2 = \int_E g$$

$$\therefore \qquad \int_E (f+g) \le \int_E f + \int_E g \qquad \ldots (2)$$

On the other hand if ϕ_1 and ϕ_2 are two simple functions such that $\phi_1 < f$ and $\phi_2 \le g$. Then $\phi_1 + \phi_2$ is simple function and

$$\phi_1 + \phi_2 \le f + g,$$

or $$f + g \ge \phi_1 + \phi_2$$

$$\therefore \qquad \int_E (f+g) \ge \int_E (\phi_1 + \phi_2)$$

But $$\int_E (\phi_1 + \phi_2) = \int_E \phi_1 + \int_E \phi_2$$

$$\Rightarrow \qquad \int_E (f+g) \ge \int_E \phi_1 + \int_E \phi_2$$

Since $$\sup_{f \ge \phi_1} \int_E \phi_1 = \int_E f$$

and $$\sup_{f \ge \phi_2} \int_E \phi_2 = \int_E g$$

$$\therefore \qquad \int_E (f+g) \ge \int_E f + \int_E g \qquad \ldots (3)$$

From (2) and (3), we get

$$\int_E (f+g) = \int_E f + \int_E g$$

$$\therefore \qquad \int_E (af+bg) = \int_E af + \int_E bg$$

$$= a\int_E f + b\int_E g \quad \text{from (i)}$$

Proof of 2: Since f = g a.e.

$$\Rightarrow \qquad f - g = 0 \text{ a.e.}$$

Notes

Let $F = \{x : f(x) \neq g(x)\}$

Then by definition of a.e., we have $mF = 0$ and $F \subset E$.

$$\therefore \quad \int_E (f-g) = \int_{F \cup (E-F)} (f-g) = \int_F (f-g) + \int_{E-F} (f-g)$$

$$= (f-g)\, mF + (f-g)\, m(E-F)$$

$$= (f-g) \,.\, 0 + 0 \,.\, m(E-F) \quad [\because mF = 0 \text{ and } f-g = 0 \text{ over } E-F]$$

$$= 0$$

$$\therefore \quad \int_E (f-g) = 0$$

$$\Rightarrow \quad \int_E f - \int_E g = 0 \Rightarrow \int_E f = \int_E g$$

Note Converse need not be true

e.g. Let the functions $f : [-1, 1] \to R$ and $g : [-1, 1] \to R$ be defined by

$$f(x) = \begin{cases} 2 \text{ if } x \leq 0 \\ 0 \text{ if } x > 0 \end{cases}$$

and $\quad g(x) = 1 \;\forall\; x.$

Then $\quad \int_{-1}^{1} f(x)\,dx = 2 = \int_{-1}^{1} g(x)\,dx$

But $\quad f \neq g$ a.e.

In other words, they are not equal even for a single point in $[-1, 1]$.

Proof of 3: $f \leq g$ a.e.

$$\Rightarrow \quad f - g \leq 0 \text{ a.e.}$$

Let ϕ be simple function,

$$\phi = f - g$$

$$\Rightarrow \quad \phi \leq 0 \qquad [\because f - g = 0 \text{ a.e.}]$$

$$\Rightarrow \quad \int_E \phi \leq 0$$

$$\Rightarrow \quad \int_E (f-g) \leq 0$$

$$\Rightarrow \quad \int_E f - \int_E g \leq 0$$

Notes

$$\Rightarrow \qquad \int_E f \le \int_E g$$

Since $\qquad f \le |f|$

$$\Rightarrow \qquad \int_E f \le \int_E |f| \qquad \ldots (1)$$

Again $-f \le |f|$

$$\Rightarrow \qquad \int_E -f \le \int_E |f|$$

or

$$-\int_E |f| \le \int_E f \qquad \ldots (2)$$

From (1) and (2) we get

$$-\int_E |f| \le \int_E f \le \int_E |f|$$

$$\Rightarrow \qquad \left|\int_E f\right| \le \int_E |f|.$$

Proof of 4: It follows from (3) and the fact that $\int_E 1 = mE$.

Proof of 5:

$$\int_{A\cup B} f = \int f \chi_{A\cup B}$$

Now $\qquad \chi_{A\cup B} = \chi_A + \chi_B - \chi_{A\cap B}$

where A and B are disjoint measurable sets i.e.

$$A \cap B = \phi$$

$$\therefore \qquad \int_{A\cup B} f = \int f(\chi_A + \chi_B) - \int f \chi_{A\cap B}$$

$$= \int f \chi_A + \int f \chi_B - 0 \qquad [\because A \cap B = \phi \text{ and } m(\phi) = 0]$$

$$= \int_A f + \int_B f$$

11.1.3 The Lebesgue Integral of a Non-negative Function

Definition: If f is a non-negative measurable function defined on a measurable set E, we define

$$\int_E f = \sup_{h \le f} \int_E h,$$

Notes

where h is a bounded measurable function such that

$$m\{x : h(x) \neq 0\} < \infty$$

Theorem 4: Let f be a non-negative measurable function. Show that $\int f = 0$ implies f = 0 a.e.

Proof: Let ϕ be any measurable simple function such that

$$\phi \leq f.$$

Since f = 0 a.e. on E

$$\Rightarrow \qquad \phi \leq 0 \text{ a.e.}$$

$$\therefore \qquad \int_E \phi(x)\,dx \leq 0$$

Taking supremum over all those measurable simple functions $\phi \leq f$, we get

$$\int_E f\,dx \leq 0 \qquad \ldots (1)$$

Similarly let Ψ be any measurable simple function such that $\Psi \geq f$

Since $\qquad f = 0$ a.e.

$$\therefore \qquad \Psi \geq 0 \text{ a.e.}$$

$$\Rightarrow \qquad \int_E \Psi(x)\,dx \geq 0$$

Taking infimum over all those measurable simple functions $\Psi \geq f$, we get

$$\int_E f\,dx \geq 0 \qquad \ldots (2)$$

From (1) and (2), we get

$$\int_E f\,dx = 0.$$

Conversely, let $\qquad \int_E f\,dx = 0$

If $\qquad E_n = \left\{x : f(x) > \frac{1}{n}\right\}$, then

$$\int_E f\,dx > \int_E \frac{1}{n}\chi_{E_n}(x)\,dx$$

But $\qquad \int_E \frac{1}{n}\chi_{E_n}(x)\,dx = \frac{1}{n} m E_n \qquad$ (By definition)

Notes

$\therefore \qquad \int_E f\,dx > \frac{1}{n} m E_n$

Or $\qquad \frac{1}{n} m E_n < \int_E f\,dx$

But $\qquad \int_E f\,dx < 0$

$\therefore \qquad \frac{1}{n} m E_n < 0$

$\Rightarrow \qquad m E_n < 0$

But $m E_n \geq 0$ is always true

$\therefore \qquad m E_n = 0$

But $\qquad \{x : f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n$

and $\qquad m E_n = 0$

$\Rightarrow \qquad m\left(\bigcup_{n=1}^{\infty} E_n\right) = 0$

$\Rightarrow \qquad m\{x : f(x) > 0\} = 0$

$\therefore \qquad f = 0$ a.e. on E

Theorem 5: Let f and g be two non-negative measurable functions. If f is integrable over E and g (x) < f (x) on E, then g is also integrable over E, and

$$\int_E (f-g) = \int_E f - \int_E g\,.$$

Proof: Since we know that if f and g are non-negative measurable functions defined on a set E, then

$$\int_E (f+g) = \int_E f + \int_E g$$

Since $\qquad f = (f - g) + g,$

therefore we have

$$\int_E f = \int_E (f-g+g) = \int_E (f-g) + \int_E g \qquad \ldots (1)$$

Since the functions f – g and g are non-negative and measurable. Further, f being integrable over E, $\int_E f < \infty$ (by definition)

Notes

Therefore, each integral on the right of (1) is finite.

In particular, $\int_E g < \infty$,

which shows that g is an integrable function over E.

Since $$\int_E f = \int_E (f-g) + \int_E g$$

$$\Rightarrow \quad \int_E f - \int_E g = \int_E (f-g).$$

11.1.4 The General Lebesgue Integral

For the positive part f^+ of a function f, we define

$$f^+ = \max(f, 0)$$

and negative part f^- by f^-

$$f^- = \max(-f, 0)$$

and that f is measurable if and only if both f^+ and f^- are measurable.

Note

$$f = f^+ - f^-$$

and $$|f| = f^+ + f^-$$

Definition: A measurable function f is said to be *Lebesgue integrable* over E if f^+ and f^- are both Lebesgue integrable over E. In this case, we define $\int_E f = \int_E f^+ - \int_E f^-$.

Theorem 6: Let f and g be integrable over E, then

(a) The function of f is integrable over E, and $\int_E cf = c\int_E f$.

(b) Sum of two integrable functions is integrable i.e. the function f + g is integral over E, and

$$\int_E (f+g) = \int_E f + \int_E g$$

(c) If $f \le g$ a.e., then $\int_E f \le \int_E g$.

(d) If A and B are disjoint measurable sets contained in E, then

$$\int_{A \cup B} f = \int_A f \le \int_B f.$$

Notes

Proof: (a) If $c \geq 0$, then

$$(cf)^+ = cf^+$$

$$(cf)^- = cf^-$$

and if $c < 0$, then

$$(cf)^* = (-c) \cdot f^-$$

$$(cf)^- = (-c) \cdot f^+$$

Since f is integrable so f^+ and f^- are also integrable and conversely. Hence the result follows.

(b) In order to prove the required result first of all we show that if f_1 and f_2 are non-negative integrable functions such that $f = f_1 - f_2$, then

$$\int_E f = \int_E f_1 - \int_E f_2 \qquad \ldots (1)$$

Since $f = f^+ - f^-$,

Also $f = f_1 - f_2$,

then $f^+ - f^- = f_1 - f_2$

$\Rightarrow$ $f^+ + f_2 = f_1 + f^-$ … (2)

Also we know that if f and g are non-negative measurable functions defined on a set E, then

$$\int_E (f+g) = \int_E f + \int_E g$$

Then from (2), we get

$$\int_E f^+ + \int_E f_2 = \int_E f_1 + \int_E f^-$$

$$\Rightarrow \quad \int_E f^+ - \int_E f^- = \int_E f_1 - \int_E f_2 \qquad \ldots (3)$$

But f is integrable so f^+ and f^- are integrable i.e.

$$\int f = \int f^+ - \int f^-$$

Therefore (3) becomes

Hence $$\int_E f = \int_E f_1 - \int_E f_2$$

which proves (1).

Now, if f and g are integrable functions over E, then

$f^+ + g^+$, $f^- + g^-$ and $f + g = (f^+ + g^+) - (f^- + g^-)$

and also integrable functions over E.

Notes

Furthermore f and g are integrable, it implies that $|f|$ and $|g|$ are integrable.

($\because$ A measurable function f is integrable over E if and only if $|f|$ is integrable over E.)

Thus $|f| + |g|$ is integrable over E.

($\because \int_E (f+g) = \int_E f + \int_E g$ and by the definition of integrable)

Since $|f + g| \leq |+|f| + |g|$

which shows that f + g is integrable.

Hence sum of two integrable functions is integrable.

Thus
$$\int_E (f+g) = \int_E (f^+ + g^+) - \int_E (f^- + g^+)$$
$$= \int_E f^+ + \int_E g^+ - \int_E f^- - \int_E g^-$$
$$= \left(\int_E f^+ - \int_E f^-\right) + \left(\int_E g^+ - \int_E g^-\right)$$
$$= \int_E f + \int_E g$$

(c)
$$f \leq g \text{ a.e.}$$
$$\Rightarrow \quad f - g \leq 0 \text{ a.e.}$$
$$\Rightarrow \quad g - f \geq 0 \text{ a.e.}$$
$$\therefore \quad \int_E (g-f) \geq 0$$

Since g = f + (g – f) and f, g – f are integrable over E.

Then by the given hypothesis $(g - f)^- = 0$ a.e.

then $\int_E (g-f)^- = 0$,

(Since we know that if f = 0 a.e. then $\int_E f = 0$)

$$\therefore \int_E g = \int_E f + \int_E (g-f) = \int_E f + \int_E (g-f)^+ - \int_E (g-f)^-$$

becomes

$$\int_E g = \int_E f + \int_E (g-f)^+ - 0 = \int_E f + \int_E (g-f)^+ \qquad (\because (g-f) \geq 0)$$

Notes

$$\geq \int_E f$$

$$\therefore \quad \int_E g \geq \int_E f \Rightarrow \int_E f \leq \int_E g$$

(d)
$$\int_{A\cup B} f = \int f\,\chi_{A\cup B}$$

$$= \int f\cdot\chi_A + \int f\cdot\chi_B$$

$$= \int_A f + \int_B f\,.$$

Example: Let f be a non-negative integrable function. Show that the function F defined by

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \text{ is continuous on R.}$$

Solution: Since f is a non-negative integrable function, then given $\varepsilon > 0$ there is a $\delta > 0$ such that for every set $A \subset R$ with $m(A) < \delta$, we have

$$\left|\int_A f\right| < \varepsilon$$

If $x_o \in R$, then $\forall\ x \in R$ with $|x - x_o| < \delta$, we have

$$\left|\int_{x_o}^{x} f(t)\right| dt < \varepsilon$$

$$\Rightarrow \quad \left|\int_{x_o}^{-\infty} f(t)\,dt\right| + \left|\int_{-\infty}^{x} f(t)\,dt\right| < \varepsilon$$

$$\Rightarrow \quad \left|\int_{-\infty}^{x} f(t)\,dt\right| + \left|\int_{x_o}^{-\infty} f(t)\,dt\right| < \varepsilon$$

$$\Rightarrow \quad \left|\int_{-\infty}^{x} f(t)\,dt\right| - \left|\int_{-\infty}^{x_o} f(t)\,dt\right| < \varepsilon$$

$$\Rightarrow \quad |F(x) - F(x_o)| < \varepsilon$$

Hence F is continuous at x_o. Since $x_o \in R$ is arbitrary, F is continuous on R.

Notes

11.2 Summary

- Let ϕ and Ψ be simple functions which vanish outside a set of finite measure, then

$$\int (a\phi + b\Psi) = a\int \phi + b\int \Psi \text{ and if } \phi \geq \Psi \text{ a.e., then } \int \phi \geq \int \Psi$$

- A Lebesgue integrable function may not be Riemann integrable.
- Let A_1 be the set of all irrational numbers and A_2 be the set of all rational numbers in [0, 1].
- If f is a non-negative measurable function defined on a measurable set E, we define

$$\int_E f = \sup_{h \leq f} \int_E h,$$

where h is a bounded measurable function such that

$$m\{x : h(x) \neq 0\} < \infty$$

- Let f and g be two non-negative measurable functions. If f is integrable over E and g (x) < f (x) on E, then g is also integrable over E, and

$$\int_E (f - g) = \int_E f - \int_E g.$$

11.3 Keywords

Canonical Representation: If ϕ is simple function and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ the set of non-zero values of ϕ, then

$$\phi = \sum_{i=1}^{n} \alpha_i \chi_{E_i},$$

where $E_i = \{x : \phi(x) = \alpha_i\}$.

Characteristic Function: The function χ_E defined by

$$\chi_E(x) = \begin{cases} 1 \text{ if } x \in E \\ 0 \text{ if } x \notin E \end{cases}$$

is called the *characteristic function* of E.

Elementary Integral: If ϕ vanishes outside a set of finite measure, we define the elementary integral of ϕ by $\int \phi(x)\,dx = \sum_{i=1}^{n} \alpha_i\, mE_i$ when ϕ has the canonical representation.

$$\phi = \sum_{i=1}^{n} \alpha_i \chi_{E_i}.$$

Lebesgue Integrable: A measurable function f is said to be Lebesgue integrable over E if f^+ and f^- are both Lebesgue integrable over E. In this case, we define $\int_E f = \int_E f^+ - \int_E f^-$.

Notes

Simple Function: A linear combination $\phi(x) = \sum_{i=1}^{n} \alpha_i \chi_{E_i}(x)$ is called a simple function if the sets E_i are measurable.

Simple Function: A linear combination $\phi(x) = \sum_{i=1}^{n} \alpha_i \chi_{E_i}(x)$ is called a *simple function* if the sets E_i are measurable.

This representation of ϕ is not unique.

However, a function ϕ is simple if and only if it is measurable and assumes only a finite number of values.

The Lebesgue Integral of a Non-negative Function: If f is a non-negative measurable function defined on a measurable set E, we define

$$\int_E f = \sup_{h \le f} \int_E h,$$

where h is a bounded measurable function such that

$$m\{x : h(x) \neq 0\} < \infty$$

The Riemann Integral: Let f be a bounded real valued function defined on the interval [a, b] and let $a = x_0 < x_1 < \ldots < x_n = b$ be a sub-division of [a, b].

Then for each sub-division we can define the sums

$$S = \sum_{i=1}^{n} (x_i - x_{i-1}) M_i$$

and

$$s = \sum_{i=1}^{n} (x_i - x_{i-1}) m_i$$

where

$$M_i = \sup_{x_{i-1} < x \le x_i} f(x),$$

$$m_i = \inf_{x_{i-1} < x \le x_i} f(x)$$

11.4 Review Questions

1. Prove that $\int_E af = a\int_E f \; \forall$ real number a.

2. If f is bounded real valued measurable function defined on a measurable set E of finite measure such that $a \le f(x) \le b$, then show that $amE \le \int_E f \le bmE$.

3. If f and g are non-negative measurable functions defined on $E \in \mathcal{M}$ then prove that

   (a) $\int_E cf = c\int_E f, c > 0$

Notes

(b) $\int_E (f+g) = \int_E f + \int_E g$

(c) $\int f = 0 \Rightarrow f = 0$ a.e.

(d) If $f \le g$ a.e. then $\int_E f \le \int_E g$

4. If f is integrable over E, then show that $|f|$ is integrable over E, and $\left|\int_E f\right| \le \int_E |f|$.

5. Show that if f is a non-negative measurable function then f = 0 a.e. on E iff $\int_E f = 0$.

6. If $\int_E f = 0$ and $f(x) \ge 0$ on E, then f = 0 a.e.

11.5 Further Readings

Books

Erwin Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons Inc., New York, 1989

Walter Rudin, *Real and Complex Analysis*, Third McGraw Hill Book Co., New York, 1987

R.G. Bartle, *The Elements of Integration and Lebesgue Measure,* Wiley Interscience, 1995

Online links

www.maths.manchester.ac.uk

www.uir.ac.za

Notes

Unit 12: General Convergence Theorems

CONTENTS

Objectives

After studying this unit, you will be able to:

- Understand bounded convergence theorem.
- State and prove monotone convergence theorem and Lebesgue dominated convergence theorem.
- Solve related problems on these theorems.

Introduction

Convergence of a sequence of functions can be defined in various ways and there are situations in which each of these definitions is natural and useful. In this unit, we shall study about convergence almost everywhere, pointwise and uniform convergence. We shall also prove bounded convergence theorem and monotone convergence theorem which are so useful in solving problems on convergence. The dominated convergence theorem is one of the most important results of Lebesgue's integration theory. It gives a general sufficient condition for the validity of proceeding to the limit of a sequence of functions under the integral sign. It is an invaluable tool to study functions defined by integrals.

12.1 General Convergence Theorems

12.1.1 Convergence almost Everywhere

Let $<f_n>$ be a sequence of measurable functions defined over a measurable set E. Then $<f_n>$ is said to converge almost everywhere in E if there exists a subset E_0 of E s.t.

Notes

(i) $f_n(x) \to f(x), \ \forall \ x \in E - E_0$,

and (ii) $m(E_0) = 0$.

12.1.2 Pointwise Convergence

Let $<f_n>$ be a sequence of measurable functions on a measurable set E. Then $<f_n>$ is said to converge "pointwise" in E, if $\exists$ a measurable function f on E such that

$$f_n(x) \to f(x) \ \forall \ x \in E \text{ or}$$

$$\underset{n\to\infty}{\text{lt}} f_n(x) = f(x)$$

12.1.3 Uniform Convergence, Almost Everywhere (a.e.)

Let $<f_n>$ be a sequence of measurable functions defined over a measurable set E. Then the sequence $<f_n>$ is said to converge uniformly a.e. to f, if $\exists$ a set $E_0 \subset E$ s.t.

(i) $m(E_0) = 0$ and

(ii) $<f_n>$ converges uniformly to f on the set $E - E_0$.

12.1.4 Bounded Convergence Theorem

Theorem 1: State and Prove: Bounded Convergence Theorem

Statement: Let $\{f_n\}$ be a sequence of measurable functions defined on a set E of finite measure, and suppose that there is a real number M such that $|f_n(x)| \le M \ \forall$ n and x. If $f(x) = \lim_{n\to\infty} f_n(x)$ for each x in E, then

$$\int_E f = \lim_{n\to\infty} \int_E f_n$$

Proof: Since $f(x) = \lim_{n\to\infty} \int_E f_n(x)$ and f_n is measurable on E

$\therefore$ f is also measurable on E

Let $\varepsilon > 0$ be given

Then $\exists$ measurable set $A \subset E$ with $mA < \dfrac{\varepsilon}{4M}$ and a positive integer N such that

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2mE} \ \forall \ n \ge N \text{ and } x \in E - A$$

$$\therefore \quad \left| \int_E f_n - \int_E f \right| = \left| \int_E (f_n - f) \right|$$

$$\le \int_E |(f_n - f)|$$

Notes

$$= \int_{E-A} |(f_n - f)| + \int_A |(f_n - f)| \text{ as } (E - A) \cap A = \phi$$

$$\leq \frac{\varepsilon}{2mE} \int_{E-A} 1 + \int_A (|f_n| + |f|)$$

$$\leq \frac{\varepsilon}{2mE} m(E - A) + 2M \int_A 1$$

$$\leq \frac{\varepsilon}{2mE} mE + 2M\, mA \text{ as } m(E - A) \leq mE$$

$$< \frac{\varepsilon}{2} + 2M \frac{\varepsilon}{4M}$$

$$= \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

$$= \varepsilon$$

Thus $$\left| \int_E f_n - \int_E f \right| < \varepsilon$$

But ε was arbitrary

$\therefore$ $$\lim_{n\to\infty} \int_E f_n = \int_E f$$

12.1.5 Fatou's Lemma

If $\{f_n\}$ is a sequence of non-negative measurable functions and $f_n(x) \to f(x)$ almost everywhere on a set E, then

$$\int_E f \leq \underline{\lim}_{n\to\infty} \int_E f_n$$

i.e. $$\int_E f \leq \lim_{n\to\infty} \inf \int_E f_n$$

Proof: Since integrals over sets of measure zero are zero.

$\therefore$ Without loss of generality, we may assume that the convergence is everywhere. Let h be a bounded measurable function with $h \leq f$ and $h(x) = 0$ outside a set $E' \subset E$ of finite measure.

Define a function h_n by

$h_n(x) = \text{Min.}\{h(x), f_n(x)\}$

then $h_n(x) \leq h(x)$ and $h_n(x) \leq f_n(x)$

$\therefore$ h_n is bounded by the boundedness of h and vanishes outside E′ as $x \in E - E' \Rightarrow h(x) = 0$ $\Rightarrow h_n(x) = 0$ because

Notes

Since $h_n = h$ or $h_n = f_n$

$\therefore$ h_n is measurable function on E'

If $h_n = h$, then $h_n \to h$

If $h_n = f_n < h \le f$

then $f_n \to h$ as $f_n \to f$

$\Rightarrow$ $h_n \to h$

Thus $h_n \to h$

Since $h_n(x) \to h(x)$ for each $x \in E'$ and $\{h_n\}$ is a sequence of bounded measurable functions on E'

$\therefore$ By Bounded Convergence Theorem

$$\int_E h = \int_{E'} h + \int_{E-E'} h = \int_E h = \lim_{n\to\infty} \int_{E''} h_n$$

as $E = (E - E') \cup E'$ & $(E - E') \cap E' = \phi$

$$= \underline{\lim}_{n\to\infty} \int_{E'} h_n$$

$$\le \underline{\lim}_{n\to\infty} \int_{E'} f_n \text{ as } h_n \le f_n$$

$$\le \underline{\lim}_{n\to\infty} \int_{E'} f_n \text{ as } E' \subset E$$

$$\Rightarrow \quad \le \int_E h \le \underline{\lim}_{n\to\infty} \int_{E'} f_n$$

Taking supremum over all $h \le f$, we get

$$\sup_{n\le f} \int_{E'} h \le \int_E h = \underline{\lim}_{n\to\infty} \int_{E'} f_n$$

$$\Rightarrow \quad \int_E f \le \int_E f \le \underline{\lim}_{n\to\infty} \int_E f_n$$

Remarks:

(1) If in Fatou's Lemma, we take

$$f_n(x) = \begin{cases} 1, n \le x < n+1 \\ 0, \text{otherwise} \end{cases}$$

with $E = R$

then $\int_E f \le \underline{\lim}_{n\to\infty} \int_E f_n$

Thus in Fatou's Lemma, strict inequality is possible.

Notes

(2) $f_n \geq 0 \ \forall \ x \in E$ is essential for Fatou's Lemma

However, if we take

$$f_n(x) = \begin{cases} -n, \frac{1}{n} \leq x < \frac{2}{n} \\ 0, \text{otherwise} \end{cases}$$

with E = [0, 2]

Then $\int_E f \leq \underline{\lim}_{n\to\infty} \int_E f_n$.

12.1.6 Monotone Convergence Theorem

Statement: Let $\{f_n\}$ be an increasing sequence of non-negative measurable functions and let $f = \lim_{n\to\infty} \int f_n$. Then

$$\int f = \lim_{n\to\infty} \int f_n$$

Proof: Let h be a bounded measurable function with $h \leq f$ and $h(x) = 0$ outside a set $E' \subset E$ of finite measure

Define a function h_n by

$$h_n(x) = \text{Min.}\{h(x), f_n(x)\}$$

then $h_n(x) \leq h(x)$ and $h_n(x) \leq f_n(x)$

$\therefore$ h_n is bounded by the boundedness of h and vanishes outside E′ as

$x \in E - E' \Rightarrow h(x) = 0 \Rightarrow h_n(x) = 0$ because $f_n(x) \geq 0$

Since $h_n = h$ or $h_n = f_n$

$\therefore$ h_n is measurable function on E′

If $h_n = h$, then $h_n \to h$

then $f_n \to h$ as $f_n \to f$

$\Rightarrow$ $h_n \to h$

Thus $h_n \to h$

Since $h_n(x) \to h(x)$ for each $x \in E'$ and $\{h_n\}$ is a sequence of measurable functions on E′

$\therefore$ By Bounded Convergence Theorem

$$\int_E h = \int_{E'} h + \int_{E-E'} h = \int_{E'} h = \lim_{n\to\infty} \int_{E'} h_n$$

as $E = (E - E') \cup E'$ & $(E - E') \cap E' = \phi$.

$$= \underline{\lim}_{n\to\infty} \int_{E'} h_n$$

Notes

$$= \underline{\lim}_{n\to\infty} \int_{E'} f_n$$

$$\leq \underline{\lim}_{n\to\infty} \int_{E} f_n \text{ as } E' \subset E$$

$$\Rightarrow \quad \int_E h \leq \underline{\lim}_{n\to\infty} \int_E f_n$$

Taking supremum over all $h \leq f$, we get

$$\sup \int_E h \leq \underline{\lim}_{n\to\infty} \int_E f_n$$

$$\Rightarrow \quad \int_E f \leq \underline{\lim}_{n\to\infty} \int_E f_n \qquad \ldots (1)$$

Since $\{f_n\}$ is monotonically increasing sequence and $f_n \to f$

$$\therefore \quad f_n \leq f$$

$$\Rightarrow \quad \int f_n \leq \int f$$

$$\Rightarrow \quad \overline{\lim}_{n\to\infty} \int f_n \leq \int f \qquad \ldots (2)$$

$\therefore$ From (1) and (2), we have

$$\int f \leq \underline{\lim}_{n\to\infty} \int f_n \leq \overline{\lim}_{n\to\infty} \int f_n \leq \int f$$

$$\int f \leq \lim_{n\to\infty} \int f_n$$

Theorem 2: Let $\{u_n\}$ be a sequence of non-negative measurable functions, and let $f = \sum_{h=1}^{\infty} \int u_n$.

Then $\int f = \sum_{h=1}^{\infty} \int u_n$

Proof: Let $f_n = u_1 + u_2 + \ldots + u_n = \sum_{j=1}^{n} \int u_j$

then $f_n \to f$

i.e. $\lim_{n\to\infty} f_n = f$

Let h be a bounded measurable function with $h \leq f$ and $h(x) = 0$ outside a set $E' \subset E$ of finite measure.

Define a function h_n by

$$h_n(x) = \text{Min. } \{h(x), f_n(x)\}$$

Notes

then $h_n(x) \le h(x)$ and $h_n(x) \le f_n(x)$

$\therefore$ h_n is bounded by the boundedness of h and vanishes outside E′ as

$x \in E - E' \Rightarrow h(x) = 0 \Rightarrow h_n(x) = 0$ because $f_n(x) \ge 0$.

Since $h_n = h$ or $h_n = f_n$

$\therefore$ h_n is measurable function on E′

If $h_n = h$, then $h_n \to h$

If $h_n = f_n < h < f$

then $f_n \to h$ as $f_n \to f$

$\Rightarrow$ $h_n \to h$

Thus $h_n \to h$

Since $h_n(x) \to h(x)$ for each $x \in E'$ and $\{h_n\}$ is a sequence of measurable function on E′

$\therefore$ By Bounded Convergence Theorem

$$\int_E h = \int_{E'} h + \int_{E-E'} h = \int_{E'} h = \lim_{n\to\infty} \int_{E'} h_n$$

as $E = (E - E') \cup E'$ & $(E - E') \cap E' = \phi$

$$= \underline{\lim}_{n\to\infty} \int_{E'} h_n$$

$$= \underline{\lim}_{n\to\infty} \int_{E'} f_n$$

$$\le \underline{\lim}_{n\to\infty} \int_{E} f_n \quad \text{as } E' \subset E$$

$\Rightarrow$
$$\int_E h \le \underline{\lim}_{n\to\infty} \int_E f_n$$

Taking supremum over all $h \le f$, we get

$$\sup_{h\le f} \int_E h \le \underline{\lim}_{n\to\infty} \int_E f_n \qquad \dots (1)$$

$\Rightarrow$
$$\int_E f \le \underline{\lim}_{n\to\infty} \int_E f_n$$

Since $\{f_n\}$ is monotonically increasing sequence and $f_n \to f$

$\therefore$
$$f_n \le f$$

$\Rightarrow$
$$\int f_n \le \int f$$

$\Rightarrow$
$$\overline{\lim}_{n\to\infty} \int f_n \le \int f \qquad \dots (2)$$

Notes

$\therefore$ From (1) and (2), we have

$$\int f \le \underline{\lim}_{n\to\infty} \int f_n \le \overline{\lim}_{n\to\infty} \int f_n \le \int f$$

$$\Rightarrow \quad \int f \le \lim_{n\to\infty} \int f_n$$

$$= \lim_{n\to\infty} \int f_n$$

$$= \lim_{n\to\infty} \int \sum_{j=1}^{n} u_j$$

$$= \lim_{n\to\infty} \sum_{j=1}^{n} \int u_j$$

$$= \sum_{j=1}^{n} \int u_j$$

Hence $$\int f = \sum_{n=1}^{\infty} \int u_n$$

Theorem 3: Let f be a non-negative function which is integrable over a set E. Then given $\varepsilon > 0$ there is a $\delta > 0$ such that for every set $A \subset E$ with $mA < \delta$, we have

$$\int_A f < \varepsilon$$

Proof: If f is bounded function on E

Then $\exists$ positive real number M such that

$$|f(x)| \le M \ \forall \ x \in E$$

$\therefore$ For given $\varepsilon > 0 \ \exists \ \delta = \dfrac{\varepsilon}{M}$ such that for every set $A \subset E$ with $mA < \delta$, we have

$$\int_A f \le \int_A M = M\, mA < M.\delta = M\frac{\varepsilon}{M} = \varepsilon$$

i.e. $$\int_A f < \varepsilon$$

Thus the result is true if f is a bounded function. So assume that f is not a bounded function on E.

Define a function f_n on E by

$$f_n(x) = \begin{cases} f(x) & \text{if } f(x) \le n \\ n & \text{otherwise} \end{cases}$$

Then each f_n is bounded and **Notes**

$f_n \to f$ at each point

Since $\{f_n\}$ is an increasing sequence of bounded functions such that $f_n \to f$ on E

$\therefore$ By the monotone convergence theorem

$$\lim_{n\to\infty}\int_E f_n = \int_E f$$

$\therefore$ For given $\varepsilon > 0$ $\exists$ a positive integer N such that

$$\left|\int_E f_n - \int_E f\right| < \frac{\varepsilon}{2} \text{ for } n \geq N$$

$$\Rightarrow \quad \left|\int_E f - \int_E f_N\right| < \frac{\varepsilon}{2}$$

$$\Rightarrow \quad -\frac{\varepsilon}{2} < \int_E f - \int_E f_N < \frac{\varepsilon}{2}$$

$$\Rightarrow \quad \int_E (f - f_N) < \frac{\varepsilon}{2}$$

Choose $\delta < \dfrac{\varepsilon}{2N}$

$\therefore$ If $mA < \delta$, then we have

$$\begin{aligned}
\int_A f &= \int_A \left[(f - f_N) + f_N\right] \\
&= \int_A (f - f_N) + \int_A f_N \\
&\leq \int_E (f - f_N) + \int_A N \text{ as } f_N \leq N \\
&< \frac{\varepsilon}{2} + N\,mA \\
&< \frac{\varepsilon}{2} + N \cdot \delta \\
&< \frac{\varepsilon}{2} + N \cdot \frac{\varepsilon}{2N} \\
&= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon
\end{aligned}$$

Notes

$$\Rightarrow \qquad \int_A f < \varepsilon$$

12.1.7 Lebesgue Dominated Convergence Theorem

Theorem 4: State and prove Lebesgue dominated convergence theorem

Statement: Let g be an integrable function on E and let $\{f_n\}$ be a sequence of measurable functions such that $|f_n| \le g$ on E and $\lim_{n\to\infty} f_n = f$ a.e. on E. Then

$$\int_E f = \lim_{n\to\infty} \int_E f_n .$$

Proof: Since we know that if f is a measurable function over a set E and there is an integrable function g such that $|f| \le g$, then f is integrable over E. So clearly, each f_n is integrable over E.

Also $\quad \lim_{n\to\infty} f_n = f$ a.e. on E.

and $\quad |f_n| \le g$ a.e. on E

$\Rightarrow \quad |f| \le g$ a.e. on E.

Hence f is integrable over E.

Let $\{\phi_n\}$ be a sequence of functions defined by $\phi_n = f_{n+g}$. Clearly, ϕ_n is a non-negative and integrable function for each n.

Therefore, by Fatou's Lemma, we have

$$\int_E (f+g) \le \underline{\lim}_{n\to\infty} \int_E (f_n + g)$$

$$\Rightarrow \qquad \int_E f \le \underline{\lim}_{n\to\infty} \int_E f_n \qquad \ldots (1)$$

Similarly, let $\{\Psi_n\}$ be a sequence of functions defined by $\Psi_n = g - f_n$. Clearly Ψ_n is a non-negative and integrable function for each n. So, given by Fatou's Lemma, we have

$$\int_E (g-f) \le \underline{\lim}_{n\to\infty} \int_E (g - f_n)$$

$$\Rightarrow \qquad \int_E g - \int_E f \le \int_E g - \underline{\lim}_{n\to\infty} \int_E f_n$$

$$\Rightarrow \qquad -\int_E f \le -\underline{\lim}_{n\to\infty} \int_E f_n$$

$$\Rightarrow \qquad \int_E f \ge \overline{\lim} \int_E f_n \qquad \ldots (2)$$

Hence from (1) & (2), we get

$$\int_E f = \underline{\lim} \int_E f_n = \overline{\lim} \int_E f_n$$

Notes

But $$\underline{\lim}\int_E f_n = \overline{\lim}\int_E f_n = \lim\int_E f_n$$

Hence $$\int_E f = \lim\int_E f_n .$$

Corollary: Let $\{u_n\}$ be a sequence of integrable functions on E such that $\sum_{n=1}^{\infty} u_n$ converges a.e. on E. Let g be a function which is integrable on E and satisfy $\left|\sum_{i=1}^{n} u_i\right| \le g$ a.e. on E for each n. Then $\left|\sum_{n=1}^{\infty} u_n\right|$ is integrable on E and $\int_E \sum_{n=1}^{\infty} u_n = \sum_{n=1}^{\infty}\int_E u_n$.

Proof: Let $\sum_{i=1}^{n} u_i = f_n$.

Applying Lebesgue Dominated Convergence Theorem for the sequence $\{f_n\}$, we get

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

Corollary: If f is integrable over E and $\{E_i\}$ is a sequence of disjoint measurable sets such that $\sum_{i=1}^{\infty} E_i = E$, then

$$\int_E f = \sum_{i=1}^{\infty}\int_{E_i} f$$

Proof: Since $\{E_i\}$ is a sequence of disjoint measurable sets, we may write.

$$f = \sum_{i=1}^{\infty} f\cdot\chi_{E_i}$$

The function f. χ_{E_i} is integrable over E since $\left|f\,\chi_{E_i}\right| \le |f|$ and $|f|$ is integrable over E. Moreover

$$\left|\sum_{i=1}^{n} f\cdot\chi_{E_i}\right| \le |f|, \forall n \in N$$

Thus the conditions of above corollary are satisfied and hence

$$\int_E f = \int_E \sum_{i=1}^{\infty} f\cdot\chi_{E_i}$$

Notes

$$= \sum_{i=1}^{\infty} \int_E f \cdot \chi_{E_i}$$

$$= \sum_{i=1}^{\infty} \int_{E_i} f$$

Example: Show that the theorem of bounded convergence applies to $f_n(x) = \dfrac{nx}{1+n^2x^2}$, $0 \le x \le 1$.

Sol:

$$f_n(x) = \frac{nx}{1+n^2x^2}$$

$$= \frac{1}{\frac{1}{nx}+nx}$$

$$= \frac{1}{\left(\frac{1}{\sqrt{nx}}-\sqrt{nx}\right)^2+2}$$

$$\le \frac{1}{2}$$

Thus $\exists$ a number $\frac{1}{2}$ such that $|f_n(x)| \le \frac{1}{2}$.

Hence it satisfies the conditions of bounded convergence theorem.

Now

$$\lim_{n\to\infty}\int_0^1 f_n(x)\,dx = \lim \int_0^1 \frac{nx}{1+n^2x^2}\,dx$$

$$= \lim_{n\to\infty}\frac{1}{2n}\log(1+n^2x^2) \qquad \left(\because \frac{\infty}{\infty}\text{ form}\right)$$

$$= \lim_{n\to\infty}\frac{[1/(1+n^2x^2)]\,2nx^2}{2} \qquad \text{[Using L'Hospital Rule]}$$

$$= \lim_{n\to\infty}\frac{nx^2}{1+n^2x^2}$$

$$= \lim_{n\to\infty}\frac{\frac{1}{n}x^2}{\frac{1}{n^2}+x^2} = 0$$

and

$$\int_0^1 \lim_{n\to\infty} f_n(x)\,dx = \int_0^1 \lim_{n\to\infty}\left(\frac{nx}{1+n^2x^2}\right)dx$$

Notes

$$= \int_0^1 (0)\,dx = 0$$

$$\Rightarrow \qquad \lim_{n\to\infty}\int_0^1 f_n(x)\,dx = \int_0^1 \lim_{n\to\infty} f_n(x)\,dx$$

This verifies the result of bounded convergence theorem.

Example: Use Lebesgue dominated convergence theorem to evaluate $\lim_{n\to\infty}\int_0^1 f_n(x)\,dx$,

where

$$f_n(x) = \frac{n^{3/2}x}{1+n^2x^2},\ n = 1, 2, 3, \ldots\ 0 \le x \le 1.$$

Solution:

$$f_n(x) = \frac{n^{3/2}x}{1+n^2x^2}$$

$$= \frac{1}{x}\cdot\frac{n^{3/2}x^2}{1+n^2x^2}$$

$$\le \frac{1}{x} = g(x), \text{(say)}$$

$\Rightarrow \qquad f_n(x) \le g(x)$

and $\qquad g(x) \in L(0, 1]$,

Hence by Lebesgue Dominated Convergence Theorem.

$$\lim_{n\to\infty}\int_0^1 f_n(x)\,dx = \int_0^1 \lim_{n\to\infty} f_n(x)\,dx$$

$$= \int_0^1 \lim_{n\to\infty}\left(\frac{n^{3/2}x}{1+n^2x^2}\right)dx$$

$$= \int_0^1 \lim_{n\to\infty}\left(\frac{1}{\sqrt{n}}\right)\left(\frac{x}{\frac{1}{n^2}+x^2}\right)dx$$

$$= \int_0^1 0\,dx = 0.$$

Example: If (f_n) is a sequence of non-negative function s.t. $f_n \to f$ and $f_n \le f$ for each n, show that

$$\int f = \lim \int f_n$$

Notes

Solution: From the given hypothesis it follows that

$$\overline{\lim}\int f_n \le \int f \qquad \dots (1)$$

Also by Fatou's Lemma, we have

$$\int f \le \underline{\lim}\int f_n \qquad \dots (2)$$

Then from (1) and (2), we get

$$\int f \le \underline{\lim}\int f_n \le \overline{\lim}\int f_n \le \int f .$$

Hence
$$\int f = \underline{\lim}\int f_n \le \overline{\lim}\int f_n = \lim\int f_n .$$

Example: If $\alpha > 0$, prove that $\lim\limits_{n\to\infty}\int_o^n \left(1-\frac{x}{n}\right)^n x^{\alpha-1}dx = \int_o^\infty e^{-x}.x^{\alpha-1}dx$, where the integrals are taken in the Lebesgue sense.

Solution: If $f_n(x) = \left(1-\frac{x}{n}\right)^n .x^{\alpha-1} > 0$, then $f_n(x) \le g(x)$, where $g(x) = e^{-x}.x^{\alpha-1}$ $\left[\text{recall } \lim\limits_{n\to\infty}\left(1-\frac{x}{n}\right)^n = e^{-x}\right]$

Also $g(x) \in L[0, \infty]$, hence by Lebesgue dominated convergence theorem, we get

$$\lim_{n\to\infty}\int_o^n f_n(x)\,dx = \int_o^\infty \lim_{n\to\infty} f_n(x)\,dx$$

$$= \int_o^\infty \lim_{n\to\infty}\left(1-\frac{x}{n}\right)^n .x^{\alpha-1}\,dx$$

$$= \int_o^\infty e^{-x}.x^{\alpha-1}\,dx$$

Example: Show that if $\alpha > 1$,

$$\int_o^1 \frac{x\sin x}{1+(nx)^\alpha}dx = 0(n^{-1}) \text{ as } n \to \infty .$$

Solution: Consider the sequence $\langle f_n(x)\rangle$ s.t.

$$f_n(x) = \frac{nx\sin x}{1+(nx)^\alpha},\ n = 1, 2, \dots\dots\dots$$

Obviously since $\alpha > 1$, and $x \in [0, 1]$

Notes

$$\left|\frac{nx\sin x}{1+(nx)^{\alpha}}\right| \le 1$$

If $\Psi(x) = 1,\ \forall\, x$, then $|\ fn\ (x)\ | \le \Psi\ (x),\ \forall\, x$.

Hence by dominated convergence theorem, we get

$$\lim_{n\to\infty}\int_0^1 \frac{nx\sin x}{1+(nx)^{\alpha}}dx = \int_0^1 \lim_{n\to\infty}\frac{nx\sin x}{1+(nx)^{\alpha}}dx = \int_0^1 (0)\,dx = 0$$

$$\Rightarrow \qquad \lim_{n\to\infty} n\times\left[\int_0^1 \frac{x\sin x}{1+(nx)^{\alpha}}dx\right] = 0$$

$$\Rightarrow \qquad \int_0^1 \frac{x\sin x}{1+(nx)^{\alpha}}dx = 0\ (n^{-1}).$$

Example: Show that $\lim_{n\to\infty}\int_a^{\infty}\frac{n^2xe^{-n^2x^2}}{1+x^2}dx = 0$, if a > 0, but not for a = 0.

Solution: If a > 0, $\left.\begin{array}{l}\text{putting } nx = u\\ \text{and } \ du = ndx\end{array}\right\}$, we get

$$\int_a^{\infty}\frac{n^2xe^{-n^2x^2}}{1+x^2}dx = \int_{na}^{\infty}\frac{ue^{-u^2}du}{1+u^2/n^2} = \int_o^{\infty}\phi_{(na,\infty)}\frac{ue^{-u^2}}{1+u^2/n^2}du$$

Also $\left|\frac{u.e^{-u^2}}{1+(u^2/n^2)}\cdot\phi_{(na,\infty)}\right| < u.e^{-u^2} \in L\,[0,\infty]$

and $\lim_{n\to\infty}\phi_{(ne,\infty)}\frac{u.e^{-u^2}}{1+u^2/n^2} = 0$ as $\phi_{(\infty,\infty)} = 0$.

Hence by Lebesgue dominated convergence theorem, we obtain

$$\lim_{n\to\infty}\int_a^{\infty}\frac{n^2xe^{-n^2x^2}}{1+x^2}dx = \lim_{n\to\infty}\int_a^{\infty}\phi_{(na,\infty)}\frac{u.e^{-u^2}}{1+u^2/n^2}du$$

$$= \int_a^{\infty}\lim_{n\to\infty}\phi_{(na,\infty)}\frac{u.e^{-u^2}}{1+u^2/n^2}du = \int_o^{\infty}0\,dx = 0\,.$$

Now when a = 0,

$$\int_o^{\infty}\frac{n^2xe^{-n^2x^2}}{1+x^2}dx > \int_0^1\frac{n^2xe^{-n^2x^2}}{1+x^2}dx$$

Notes

$$> \frac{1}{2}\int_0^1 n^2 x.e^{-n^2x^2}dx \text{ (putting 1 in place of } x^2)$$

$$= -\frac{1}{4}\left[e^{-n^2x^2}\right]_0^1 > \frac{1}{4}$$

12.2 Summary

- ***Bounded Convergence Theorem:*** Let $\{f_n\}$ be a sequence of measurable functions defined on a set E of finite measure, and suppose that there is a real number M such that $|f_n(x)| < M$ $\forall$ n and all x. If f (x) = $\lim_{n\to\infty} f_n(x)$ for each x in E, then

$$\int_E f = \lim_{n\to\infty}\int_E f_n$$

- ***Monotone Convergence Theorem:*** Let $\{f_n\}$ be an increasing sequence of non-negative measurable functions and let f = $\lim_{n\to\infty} f_n$. Then

$$\int f = \lim_{n\to\infty}\int f_n$$

- ***Lebesgue Dominated Convergence Theorem:*** Let g be an integrable function on E and let $\{f_n\}$ be a sequence of measurable functions such that $|f_n| \le g$ on E and $\lim_{n\to\infty} f_n$ = f a.e. on E. Then

$$\int_E f = \lim_{n\to\infty}\int_E f_n$$

12.3 Keywords

Convergence almost Everywhere: Let $<f_n>$ be a sequence of measurable functions defined over a measurable set E. Then $<f_n>$ is said to converge almost everywhere in E if there exists a subset E_0 of E s.t.

(i) $f_n(x) \to f(x)$, $\forall$ $x \in E - E_0$,

and (ii) m $(E_0) = 0$.

Convergence: Refers to the notion that some functions and sequence approach a limit under certain conditions.

Fatou's Lemma: If $\{f_n\}$ is a sequence of non-negative measurable functions and $f_n(x) \to f(x)$ almost everywhere on a set E, then

$$\int_E f \le \liminf_{n\to\infty}\int_E f_n$$

Pointwise Convergence: Let $<f_n>$ be a sequence of measurable functions on a measurable set E. Then $<f_n>$ is said to converge "pointwise" in E, if $\exists$ a measurable function f on E such that

$$f_n(x) \to f(x) \ \forall \ x \in E \text{ or}$$

$$\underset{n\to\infty}{\text{lt}}\ f_n(x) = f(x)$$

Notes

Uniform Convergence, Almost Everywhere (a.e.): Let $<f_n>$ be a sequence of measurable functions defined over a measurable set E. Then the sequence $<f_n>$ is said to converge uniformly a.e. to f, if $\exists$ a set $E_0 \subset E$ s.t.

(i) $m(E_0) = 0$ and

(ii) $<f_n>$ converges uniformly to f on the set $E - E_0$.

12.4 Review Questions

1. Show that we may have strict inequality in Fatou's Lemma.
2. Let $<f_n>$ be an increasing sequence of non-negative measurable functions, and let $f = \lim f_n$.

   Show that $\int f = \lim \int f_n$.

   Deduce that $\int f = \sum_{n=1}^{\infty} \int u_n$, if u_n is a sequence of non-negative measurable functions and

   $f = \sum_{n=1}^{\infty} u_n$.
3. State the Monotone Convergence theorem. Show that it need not hold for decreasing sequences of functions.
4. Let $\{g_n\}$ be a sequence of integrable functions which converge a.e. to an integrable function g. Let $\{f_n\}$ be a sequence of measurable functions such that $|f_n| \le g_n$ and $\{f_n\}$ converges to f a.e.

   If $\int g = \lim_{n\to\infty} \int g_n$

   then prove that $\int f = \lim_{n\to\infty} \int f_n$.
5. State and prove monotone convergence theorem.

12.5 Further Readings

Books

G.F. Simmons, *Introduction to Topology and Modern Analysis*, New York: McGraw Hill, 1963.

H.L. Royden, *Real analysis*, Prentice Hall, 1988.

Online links

dl.acm.org

math.stanford.edu

www.springerlink.com

Notes

Unit 13: Signed Measures

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define signed measure.
- Describe positive and negative and null sets.
- Solve problems on signed measure.

Introduction

We have seen that a measure is a non-negative set function. Now we shall assume that it takes both positive and negative values. Such assumption leads us to a new type of measure known as signed measure. In this unit, we shall start with definition of signed measure and we shall prove some important theorems on it.

13.1 Signed Measures

13.1.1 Signed Measure: Definition

Definition: Let the couple $(X, \mathcal{A})$ be a measurable space, where $\mathcal{A}$ represents a σ-algebra of subsets of X. An extended real valued set function

$$\gamma : \mathcal{A} \to [-\infty, \infty]$$

defined on $\mathcal{A}$ is called a signed measure, if it satisfies the following postulates:

(i) γ assumes at most one of the values $-\infty$ or $+\infty$.

(ii) $\gamma(\phi) = 0$.

Notes

(iii) If $<A_n>$ is any sequence of disjoint measurable sets, then

$$\gamma\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \gamma(A_n),$$

i.e., γ is countably additive.

From this definition, it follows that a measure is a special case of a signed measure. Thus, every measure on $\mathcal{A}$ is a signed measure but the converse is not true in general, i.e. every signed measure is not a measure in general.

If $-\infty < \gamma(A) < \infty$, for very $A \in \mathcal{A}$, then we say that signed measure γ is finite.

13.1.2 Positive Set, Negative Set and Null Set

Definition

(a) ***Positive Set:*** Let $(X, \mathcal{A})$ be a measurable space and let A be any subset of X. Then $A \subset X$ is said to be a positive set relative to a signed measure γ defined on $(X, \mathcal{A})$, if

(i) $A \in \mathcal{A}$, i.e. A is measurable.

(ii) $\gamma(E) \geq 0, \ \forall E \subset A$ s.t. E is measurable.

Obviously, it follows from the above definition that:

(i) every measurable subset of a positive set is a positive set,

(ii) ϕ is a positive set w.r.t. every signed measure.

Also for A to be positive. $\gamma(A) \geq 0$ is the necessary condition, but not in general sufficient for A to be positive.

(b) ***Negative Set:*** Let $(X, \mathcal{A})$ be a measurable space. Then a subset A of X is said to be a negative set relative to a signed measure γ defined on measurable space $(X, \mathcal{A})$ if

(i) $A \in \mathcal{A}$ i.e., A is measurable.

(ii) $\gamma(E) \leq 0, \ \forall E \subset A$ s.t. E is measurable.

$\Rightarrow$ set A is negative w.r.t. γ, provided it is positive w.r.t. $-\gamma$.

(c) ***Null Set:*** A set $A \subset X$ is said to be a null set relative to a signed measure γ defined on measurable space $(X, \mathcal{A})$ is, A is both positive and negative relative to γ.

Thus, measure of every null set is zero.

Now, we know that a measurable set is a set of measure zero, iff every measurable subset of it has γ measure zero. Thus, if $A \subset X$ is a null set relative to γ then $\gamma(E) = 0, \ \forall$ measurable subsets $E \subset A$. In other words.

$$A \text{ is a null set} \Leftrightarrow \gamma(E) = 0, \ \forall \text{ measurable subsets } E \subset A.$$

Theorem 1: Countable union of positive sets w.r.t. a signed measure is positive.

Proof: Let $(X, \mathcal{A})$ be a measurable space and let γ be a signed measure defined on $(X, \mathcal{A})$. Let $<A_n>$ be a sequence of positive subsets of X, let $A = \bigcup_{i=1}^{n} A_i$ and let B be any measurable subset of A.

Set $\mathcal{B}_n = B \cap A_n \cap A_{n-1}^C \cap \ldots\ldots \cap A_1^C, \forall n \in N.$

Notes

where $A_n^C(n = 1,2,3\ldots n-1)$ denotes complement of $A_n(n = 1, 2, 3, \ldots n - 1)$ with respect to X.

Now, we know that complement of a measurable set is also measurable so that each $A_n^C(n = 1,2,3\ldots n-1)$ is measurable relative to γ. Again, intersection of countable collection of measurable sets is also measurable. Hence B_n is a measurable subset of the positive set A_n. Thus

$$\gamma(B_n) \geq 0 \qquad \text{(by the definition of positive set)} \ldots \text{(i)}$$

Obviously, the set B_n are disjoint and

if $$B = \bigcup_{n=1}^{\infty} B_n \text{, we get} \qquad \ldots \text{(ii)}$$

$$\gamma(B) = \sum_{n=1}^{\infty} \gamma(B_n) \qquad \ldots \text{(iii)}$$

In view of (i).

$$\gamma(B) \geq 0.$$

Thus, we have

(1) A is measurable for

A_n is a positive set $\Rightarrow A_n$ is a measurable set

$\Rightarrow$ countable union $\bigcup_{n=1}^{\infty} A_n$ is measurable,

$\Rightarrow A = \bigcup_{n=1}^{\infty} A_n$ is measurable

(2) $\gamma(B) \geq 0, \ \forall \ B \subset A$ s.t. B is a measurable set.

Hence A is a positive set, by definition.

Theorem 2: Let (X, $\mathcal{A}$) be a measurable space and let γ be a signed measure defined on (X, A). If B is a measurable set with finite negative measure i.e., $-\infty < \gamma(B) < 0$, then prove that B contains a negative set $A \subset B$ with the property $\gamma(A) < 0$.

Proof: If B is itself a negative set, then we may take A = B and theorem is done. Therefore consider the case when B is not a negative set. Then there must exist a measurable subset $E_1 \subset B$ and a smallest positive integer n_1, s.t.

$$\gamma(E_1) > \frac{1}{n_1}$$

$\because$ $$B = (B - E_1) \cup E_1 \text{ and } (B - E_1) \cap E_1 = \phi,$$

$\therefore$ $$\gamma(B) = \gamma(B - E_1) + \gamma(E_1) \qquad \ldots \text{(i)}$$

or $$\gamma(B - E_1) = \gamma(B) - \gamma(E_1) \qquad \ldots \text{(ii)}$$

Since γ (B) is finite, (i) implies that $\gamma(B - E_1)$ and $\gamma(E_1)$ are finite. Again $\gamma(B) < 0$, (ii) implies that $\gamma(B - E_1) < 0$.

Notes

Now, the set $B - E_1$ is either negative or contains a subset of positive measure. If the set $B - E_1$ is a negative set, then we may take $A = B - E_1$ and the theorem is done. So, suppose that $B - E_1$ is not a negative set. Then there must exist a measurable subset E_2 of $B - E_1$ and a smallest positive number n_2 with a property

$$\gamma(E_2) > \frac{1}{n_2}.$$

Since $\quad B = (B - E_1 \cup E_2) \cup (E_2 \cup E_2),$

and $\quad (B - E_1 \cup E_2) \cap (E_1 \cup E_2) = \phi,$

we have $\quad \gamma(B) = \gamma(B - E_1 \cup E_2) + \gamma(E_2 \cup E_2)$

or $\quad \gamma(B - E_1 \cup E_2) = \gamma(B) - \gamma(E_2 \cup E_2)$

$$= \gamma(B) - \gamma(E_1) - \gamma(E_2).$$

As before, $\quad \gamma(B - E_1 \cup E_2) > 0 \; [\because \gamma(B) < 0, \gamma(E_r) > 0 \text{ for } r = 1, 2]$

Thus, $B - E_1 \cup E_2$ is a set of negative measure, which is either a negative set or contains a subset of positive measure. If $B - E_1 \cup E_2$ is a negative set, then the theorem is done by taking $B = A - E_1 \cup E_2$. Otherwise we repeat the above process.

On repeating this process, at some stage we shall get either a negative subset $A \subset B$ s.t. $\gamma(A) < 0$ or a sequence $\langle E_r \rangle$ of disjoint measurable sets and a sequence $\langle n_r : r \in N \rangle$ of positive integers s.t.

$$E_r \subset B - \left[\bigcup_{n=1}^{r-1} E_n\right] \text{ and } \frac{1}{n_r} < \gamma(E_r) < \infty$$

In first case, we have nothing to do. In the latter case, let

$$A = B - \left[\bigcup_{n=1}^{\infty} E_n\right] \text{ or } B = A \cup \left[\bigcup_{n=1}^{\infty} E_n\right] \qquad \ldots \text{(iii)}$$

Then as before, it follows that

$$\gamma(B) = \gamma(A) + \sum_{n=1}^{\infty} \gamma(E_n).$$

$$> \gamma(A) + \sum_{k=1}^{\infty} \frac{1}{n_k} \qquad \ldots \text{(iv)}$$

[$\because$ change of suffix is in material]

Since $\gamma(B)$ is finite and γ assumes at most one of the values $-\infty$ and ∞, it follows from (iv) that $\gamma(A)$ is finite and the series $\sum_{k=1}^{\infty} \frac{1}{n_k}$ is convergent.

Then $\quad \gamma(A) < \gamma(B) - \sum_{k=1}^{\infty} \frac{1}{n_k}$

$$= \text{a finite negative number}$$

($\because \gamma(B)$ is a finite negative number)

Notes

or $$\gamma(A) < 0.$$

Again, we know that difference of two measurable sets is measurable and enumerable union of measurable sets is measurable therefore it follows from (iii) that A is a measurable set.

Now we shall prove that A is a negative set. Let $E \subset A$ be an arbitrary measurable set.

Since $$A = B - \left[\bigcup_{n=1}^{\infty} E_n\right],$$

$$E = B - \bigcup_{n=1}^{\infty} E_n .$$

Since $n_k \to \infty$, we can choose k so large that

$$\gamma(E) \le \frac{1}{n_{k-1}}$$

Letting $n_k \to \infty$, we obtain

$$\gamma(E) \le 0.$$

Thus we have

(1) A is measurable.

(2) $\gamma(E) \le 0, \ \forall \ E \subset A$ s.t. E is measurable.

Hence A is a negative set.

13.1.3 Hahn Decomposition Theorem

Theorem 3: Let γ be a signed measure on a measurable space $(X, \mathcal{A})$. Then there exists a positive set P and a negative set Q s.t.

$$P \cap Q = \phi \text{ and } P \cup Q = X.$$

Proof: Let $(X, \mathcal{A})$ be a measurable space and let γ be a signed measure defined on a measurable space $(X, \mathcal{A})$. Since, by definition, γ assumes at most one of the values $+\infty$ or collection of all negative subsets of X w.r.t. γ and let $\mathcal{B}$ be a collection of all negative subsets of X w.r.t. γ and let

$$k = \inf\{\gamma(E) : E \in \mathcal{B})$$

(i) $\Rightarrow$ that there exists a sequence $\langle E_n \rangle$ in $\mathcal{B}$ such that

$$\lim_{n\to\infty} \gamma(E_n) = k.$$

Let $$Q = \bigcup_{n=1}^{\infty} E_n .$$

Since $\mathcal{B}$ is a family of negative sets, $\langle E_n \rangle$ is a sequence of negative sets. Again, we know by remark of theorem 1 that countable union of negative sets is negative, it follows that Q is a negative subset of X so that

$$\gamma(Q) \ge K.$$

Notes

Now, $Q - E_n$ is a subset of Q, it follows that

$$\gamma (Q - E_n) \leq 0.$$

Since $(Q - E_n) \cap E_n = \phi$

and $Q = (Q - E_n) \cup E_n$,

we have $\gamma (Q) = \gamma (Q - E_n) + \gamma (E_n)$

$\Rightarrow$ $\gamma (Q) \leq \gamma (E_n)$, $\forall n \in N$ and $E_n \in B$.

Therefore $\gamma (Q) \leq K$. … (iii)

(ii) and (iii) $\Rightarrow \gamma (Q) = K \Rightarrow -\infty < k$. … (iv)

Now we shall show that $p = Q^C$, the complement of Q w.r.t. γ is a positive subset of X. Suppose not, i.e. P is negative. Then $\forall E \subset P$ s.t. E is measurable and $\gamma (E) < 0$. Now we know that if $-\infty < \gamma (E) < 0$, we get a negative set $A \subset E$ s.t. $\gamma (A) < 0$.

A, Q are distinct negative subsets of X

$\Rightarrow$ $A \cup Q$ is negative set

$\Rightarrow$ $\gamma (A \cup Q) \geq K$ [using (i)]

$\Rightarrow$ $\gamma (A) + \gamma (Q) \geq K$,

$\Rightarrow$ $\gamma (A) + K \geq K$, [using (iv)]

$\Rightarrow$ $\gamma (A) \geq 0$,

$\Rightarrow$ a contradiction, for $\gamma (A) < 0$

$\Rightarrow$ $P = Q^C$ is a positive subset of X

$\Rightarrow$ Q is a negative subset of X.

Thus $X = P \cup Q$, $P \cap Q = \phi$.

13.1.4 Hahn Decomposition: Definition

A decomposition of a measurable space X into two subsets s.t. $X = P \cup Q$, $P \cap Q = \phi$,

where P and Q are positive and negative sets respectively relative to the signal measure γ, is called as Hahn decomposition for the signed measure γ. P and Q are respectively called positive and negative components of X.

Notice that Hahn decomposition is not unique.

13.2 Summary

- Let the couple $(X, \mathcal{A})$ be a measurable space, where $\mathcal{A}$ represents a σ-algebra of subsets of X. An extended real-valued set function

  $$\gamma : \mathcal{A} \to [-\infty, \infty]$$

  defined on $\mathcal{A}$ is called a signed measure, if it satisfies the following postulates:

  (i) γ assumes at most one of the values $-\infty$ or $+\infty$.

  (ii) $\gamma (\phi) = 0$.

  (iii) If $\langle A_n \rangle$ is any sequence of disjoint measurable sets, then γ is countably additive.

Notes

- Let $(X, \mathcal{A})$ be a measurable space and then $A \subset X$ is said to be a positive set relative to a signed measure γ defined on $(X, \mathcal{A})$ if

  (i) A is measurable

  (ii) $\gamma(E) \geq 0, \ \forall \ E \subset A$ s.t. E is measurable.

- Let $(X, \mathcal{A})$ be a measurable space. Then $A \subset X$ is said to be negative set relative to a signed measure γ if

  (i) A is measurable

  (ii) $\gamma(E) \leq 0, \ \forall \ E \subset A$ s.t. E is measurable.

- $A \subset X$ is said to be a null set relative to a signed measure γ defined on measurable space $(X, \mathcal{A})$ is: A is both positive and negative relative to γ.

13.3 Keywords

Hahn Decomposition: Definition: A decomposition of a measurable space X into two subsets s.t. $X = P \cup Q, P \cap Q = \phi$.

Negative Set: Let $(X, \mathcal{A})$ be a measurable space. Then a subset A of X is said to be a negative set relative to a signed measure γ defined on measurable space $(X, \mathcal{A})$ if

(i) $A \in \mathcal{A}$ i.e., A is measurable.

(ii) $\gamma(E) \leq 0, \ \forall E \subset A$ s.t. E is measurable.

Null Set: A set $A \subset X$ is said to be a null set relative to a signed measure γ defined on measurable space $(X, \mathcal{A})$ is, A is both positive and negative relative to γ.

Positive Set: Let $(X, \mathcal{A})$ be a measurable space and let A be any subset of X. Then $A \subset X$ is said to be a positive set relative to a signed measure γ defined on $(X, \mathcal{A})$, if

(i) $A \in \mathcal{A}$, i.e. A is measurable.

(ii) $\gamma(E) \geq 0, \ \forall E \subset A$ s.t. E is measurable.

Signed Measure: Let the couple $(X, \mathcal{A})$ be a measurable space, where $\mathcal{A}$ represents a σ-algebra of subsets of X. An extended real valued set function

$$\gamma : \mathcal{A} \to [-\infty, \infty]$$

defined on $\mathcal{A}$ is called a signed measure, if it satisfies the following postulates:

(i) γ assumes at most one of the values $-\infty$ or $+\infty$.

(ii) $\gamma(\phi) = 0$.

13.4 Review Questions

1. If $\gamma(E) = \int_E xe^{-x^2}dx$, then find positive, negative and null sets w.r.t. γ. Also give a Hahn decomposition of R w.r.t. γ.
2. State and prove Hahn decomposition theorem for signed measures.
3. If μ is a measure and γ_1, γ_2 are the signed measures given by $\gamma_1(E) = \mu(A \cap E)$, $\gamma_2(E) = \mu(B \cap E)$, where $\mu(A \cap B) = 0$, show that $\gamma_1 \perp \gamma_2$.

Notes

4. Show that if γ_1 and γ_2 are two finite signed measures, then so is $a\gamma_1 + b\gamma_2$ where a, b are real numbers.

13.5 Further Readings

Books

Bartle, Robert G., *The Elements of Integration*, New York – London – Sydney: John Wiley and Sons

Cohn, Donald L. (1997) [1980], *Measure Theory (reprint ed.)*, Boston – Based – Stuttgart: Birkhauser Verlag

Online links

www.maths.bris.ac.uk

www.planetmath.org/signedmeasure.html

Notes

Unit 14: Radon-Nikodym Theorem

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define Absolutely continuous measure function
- State Radon-Nikodym theorem
- Understand the proof of Radon-Nikodym theorem
- Solve problems on this theorem

Introduction

In mathematics, the Radon-Nikodym theorem is a result in measure theory that states that given a measurable space (X, Σ), if a σ-finite measure on (X, Σ) is absolutely continuous with respect to a σ-finite measure on (X, Σ), then there is a measurable function f on X and taking values in $[0, \infty]$, such that for any measurable set A.

The theorem is named after Johann Radon, who proved the theorem for the special case where the underlying space is R^N in 1913, and for Otto Nikodym who proved the general case in 1930. In 1936 Hans Freudenthal further generalised the Radon-Nikodym theorem by proving the Freudenthal spectral theorem, a result in Riesz space theory, which contains the Radon-Nikodym theorem as a special case.

If Y is a Banach space and the generalisation of the Radon-Nikodym theorem also holds for functions with values in Y, then Y is said to have the Radon-Nikodym property. All Hibert spaces have the Radon-Nikodym property.

14.1 Radon-Nikodym Theorem

14.1.1 Absolutely Continuous Measure Function

Let $(X, \mathcal{A})$ be a measurable space and let γ and μ be measure functions defined on the space (X, A). The measure γ is said to be absolutely continuous w.r.t. μ if

$$\mu(A) = 0 \text{ or } |\mu|(A) = 0,\ A \in \mathcal{A} \Rightarrow \gamma(A) = 0, \text{ and is denoted by } \gamma \ll \mu.$$

Notes

Notes

- If μ is σ-finite, the converse is also true.
- If γ and μ are signed measures on $(X, \mathcal{A})$, then $\gamma \ll \mu$ if $|\gamma| \ll |\mu|$.

Radon-Nikodym Theorem

Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. If γ be a measure defined on A s.t. γ is absolutely continuous w.r.t. μ, then there exists a non-negative measurable function f on s.t.

$$\gamma(A) = \int_A f \, d\mu, \forall A \in \mathcal{A}.$$

The function f is unique in the sense that if g is any measurable function with the property defined as above, then f = g almost everywhere with respect to μ.

Proof: To establish the existence of the function f, we shall use the following two Lemmas:

Lemma 1: Let E be a countable set of real numbers. Let for each $a \in E$ there is a set $F_a \in \mathcal{A}$ s.t. $F_a \subset F_b$, whenever $b < a$ i.e. $\langle F_n \rangle$ is a monotonic decreasing sequence of subsets of $\mathcal{A}$ corresponding to the sequence $\langle a_n \rangle$ of real numbers in E. Then $\exists$ a measurable extended real valued function f on X s.t.

$f(x) \le a, x \in F_a,$

and $f(x) \ge a, x \in (X - F_a).$

Proof: Let $f(x) = \inf\{a : x \in F_a\}$ $x \in X$ and let, conventionally

$\inf\{\text{empty collection of real numbers}\} = \infty$

Now, $x \in F_a \Rightarrow f(x) \le a$

$x \notin F_a \Rightarrow x \in F_a$ for every $b < a$

$\Rightarrow f(x) \ge a$

Now, $f(x) < a \Rightarrow x \in F_b$ for some $b < a$

or $\{x : f(x) < a\} = \bigcup_{b<a} [F_b]$.

Also $x \in F_b \Rightarrow f(x) \le b < a$ for some $b < a$.

Hence f is measurable.

Again, by definition of f, we observe that

$f(x) \le a, x \in F_a,$

and $f(x) \ge a, x \notin F_a.$

Thus f is the required function.

Lemma 2: Let E be a countable set of real numbers. Let corresponding to each $a \in E$, there is a set $F_a \in \mathcal{A}$ s.t.

$\gamma(F_a - F_b) = 0$ whenever $b > a$.

Then there exists a measurable function f with the property

$x \in F_a \Rightarrow f(x) \le a$ a.e.

and $x \in (X - F_a) \Rightarrow f(x) > a$ a.e.

Notes

Proof: Let $P = \bigcup_{b<a} \{F_a - F_b\}$.

Evidently $\gamma(P) = 0$.

Let $F'_a = F_a \cup P$.

This $\Rightarrow F'_a - F'_b = (F_a - F_b) - P = \phi$ for $a < b$.

In view of Lemma 1, it follows that $\exists$ a measurable function f s.t.

$f(x) \le a,\ x \in F'_a$

and $f(x) \le a,\ x \in X - F'_a$

Thus we have

$$\left.\begin{array}{ll} x \in F_a \Rightarrow f(x) \le a & \text{a.e.,} \\ x \in X - F_a \Rightarrow f(x) > a & \text{a.e.,} \end{array}\right\} \text{except for } x \in P.$$

Proof of the main theorem

At first, suppose that μ is finite.

$\Rightarrow (\gamma - a\mu)$ is a signed measure on $\mathcal{A}$ for each rational number a.

Let (P_a, Q_a) be a Hahn decomposition for the measure $(\gamma - a\mu)$.

Let $P_O = X$ and $Q_O = \phi$.

By the definition of Hahn decomposition theorem,

$P_a \cup Q_a = X,$

and $P_b \cup Q_b = X.$

Therefore, $Q_a - Q_b = Q_a - (X - P_b)$

$= Q_a \cap P_b.$

Thus, $$(\gamma - a\mu)(Q_a - Q_b) \le 0 \qquad \ldots \text{(i)}$$

Similarly, we can prove that

$$(\gamma - b\mu)(Q_a - Q_b) \ge 0 \qquad \ldots \text{(ii)}$$

Let $a < b$, then from (i) and (ii), we have

$$\mu(Q_a - Q_b) = 0.$$

Therefore, by Lemma (ii)

$$f(x) \ge a, \text{ a.e. } x \in P_a$$

and $$f(x) \le a, \text{ a.e. } x \in Q_a,$$

where f is measurable

Since $Q_O = \phi$, it follows that f is non-negative

Again, let $A \in \mathcal{A}$ be arbitrary.

Define $A_r = A \cap \left(\frac{Q_{r+1}}{n_o} - \frac{Q_r}{n_o}\right)$

Notes

$$A_\infty = A - \cup \left(\frac{Q_r}{n_o}\right).$$

Evidently, $A = A_\infty \cup \left(\bigcup_{r=o}^{\infty} A_r\right)$,

where A is disjoint union of measurable sets.

$$\therefore \qquad \gamma(A) = \gamma(A_\infty) + \sum_{r=o}^{\infty} \gamma(A_r).$$

Obviously $A_r \subset \left(\frac{Q_{r+1}}{n_o} - \frac{Q_r}{n_o}\right)$

$$\Rightarrow \qquad \frac{r}{n_o} \le f(x) \le \frac{r+1}{n_o}, \forall x \in A_r$$

$$\Rightarrow \qquad \frac{r}{n_o}\mu(A_r) \le \int_{A_r} f d\mu \le \frac{r+1}{n_o}\mu(A_r) \qquad \text{[by first mean value theorem]}$$

Again $\frac{r}{n_o}\mu(A_r) \le \gamma(A_r) \le \frac{r+1}{n_o}\mu(A_r)$, we have

$$\left[\gamma(A_r) - \frac{1}{n_o}\mu(A_r)\right] \le \int_{A_r} f d\mu \le \gamma(A_r) + \frac{1}{n_o}\mu(A_r) \qquad \ldots \text{(iii)}$$

Now, if $\mu(A_\infty) > 0$, then $g(A_\infty) = 0$, $[\because (\gamma - a\mu) A_\infty$ is positive, $\forall a]$

and $\gamma(A_\infty) = 0$ if $\mu(A_\infty) = 0$ $[\because \gamma \ll \mu)]$

In either case, $\gamma(A_\infty) = \int_{A_r} f\, d\mu$.

Adding the inequalities (iii) over r, we get

$$\gamma(A) - \frac{1}{n_o}\mu(A) \le \int_A f\, d\mu \le \gamma(A) + \frac{1}{n_o}\mu(A_r).$$

Since n_o is arbitrary and $\mu(A)$ is assumed to be finite, it follows that

$$\gamma(A) = \int_A f\, d\mu \forall A \in \mathcal{A}.$$

To show that the theorem is true for σ-finite measure μ, decompose X into a countable union of X_i of finite μ-measure. Applying the same argument as above for each X_i, we get the required function.

To show the second part, let g be any measurable function satisfying the condition,

$$\gamma(A) = \int_A f\, d\mu \forall A \in \mathcal{A}.$$

Notes

For each $n \in N$, define

$$A_n = \left\{x \in X : f(x) - g(x) \geq \frac{1}{n}\right\} \in \mathcal{A}$$

and $$B_n = \left\{x \in X : g(x) - f(x) \geq \frac{1}{n}\right\} \in \mathcal{A}.$$

Since $f(x) - g(x) \geq \frac{1}{n}, \forall \mathcal{A} x \in A_n$, we have by first mean value theorem

$$\int_{A_n} (f-g)d\mu \geq \frac{1}{n}\mu(A_n)$$

$$\int_{A_n} fd\mu - \int_{A_n} gd\mu \geq \frac{1}{n}\mu(A_n)$$

$$\gamma(A_n) - \gamma(A_n) \geq \frac{1}{n}\mu(A_n) \text{ or } 0 \geq \frac{1}{n}\mu(A_n)$$

$\Rightarrow$ $\mu(A_n) \leq 0.$

Since $\mu(A_n)$ is always greater than equal to zero, we have $\mu(A_n) = 0$.

Similarly, we can show that

$\Rightarrow$ $\mu(B_n) \leq 0.$

If $$C = \{x \in X : f(x) \neq g(x)\}$$

$$= \bigcup_{n=1}^{\infty}(A_n \cup B_n),$$

then $\mu(C) = 0 \Rightarrow f = g$ a.e. on X w.r.t. μ.

Hence the theorem.

Theorem 1: If γ_1, γ_2 are σ-finite signed measures on (X, A) and $\gamma_1 \ll \mu, \gamma_2 \mu$, then

$$\frac{d(\gamma_1 + \gamma_2)}{d\mu} = \frac{d\gamma_1}{d\mu} + \frac{d\gamma_2}{d\mu} \text{ and } \frac{d\gamma_1}{d\mu} = \frac{d(-\gamma_1)}{d\mu}$$

Proof: Since γ_1, γ_2 are σ-finite and $\gamma_1 \ll \mu, \gamma_2 \ll \mu$, we have that $\gamma_1 + \gamma_2$ is also σ-finite and $\gamma_1 + \gamma_2 \ll \mu$.

Now for any $A \in \mathcal{A}$,

$$(\gamma_1 + \gamma_2)(A) = \gamma_1(A) + \gamma_2(A)$$

$$= \int_A \frac{d\gamma_1}{d\mu}d\mu + \int_A \frac{d\gamma_2}{d\mu}d\mu = \int_A \left[\frac{d\gamma_1}{d\mu} + \frac{d\gamma_2}{d\mu}\right]d\mu$$

$$\Rightarrow \quad \int_A \left[\frac{d\gamma_1 + \gamma_2}{d\mu}\right]d\mu = \int_A \left[\frac{d\gamma_1}{d\mu} + \frac{d\gamma_2}{d\mu}\right]d\mu$$

Notes

$$\Rightarrow \qquad \frac{d(\gamma_1+\gamma_2)}{d\mu} = \frac{d\gamma_1}{d\mu} + \frac{d\gamma_2}{d\mu}$$

Prove the other result yourself.

Theorem 2: If γ is a σ-finite signed measures and μ is a σ-finite measure s.t. $\gamma \ll \mu$, show that

$$\frac{d|\gamma|}{d\mu} = \left|\frac{d\gamma}{d\mu}\right|.$$

Proof: Let $\gamma = \gamma^+ - \gamma^-$ with Hahn decomposition A, B.

Then on A, $\left|\frac{d\gamma}{d\mu}\right| = \frac{d\gamma^+}{d\mu}$ and on B, $\left|\frac{d\gamma}{d\mu}\right| = \frac{d\gamma^-}{d\mu}$

$$\Rightarrow \qquad \left|\frac{d\gamma}{d\mu}\right| = \frac{d\gamma^+}{d\mu} + \frac{d\gamma^-}{d\mu} = \frac{d(\gamma^+ + \gamma^-)}{d\mu} = \frac{d|\gamma|}{d\mu}.$$

Theorem 3: If γ be a σ-finite signed measure and μ, λ be σ-finite measures on (X, A) s.t. $\gamma \ll \mu$, $\mu \ll \lambda$: then show that

$$\frac{d\gamma}{d\lambda} = \frac{d\gamma}{d\mu} \cdot \frac{d\mu}{d\lambda}$$

Proof: Since we may write $\gamma = \gamma^+ - \gamma^-$ and

$$\frac{-d\gamma^-}{d\mu} = \frac{d(-\gamma^-)}{d\mu}, \frac{-d\gamma^-}{d\lambda} = \frac{d(-\gamma^-)}{d\lambda}.$$

we need to prove the above result for measures only.

If $\frac{d\gamma}{d\mu} = f$ and $\frac{d\mu}{d\lambda} = g$, (f, g are non-negative functions as obtained in Radon-Nikodym Theorem), then we need to prove that

$$\gamma(F) = \int_F fg\, d\lambda.$$

Let Ψ be a measurable simple function s.t.

$$\Psi = \sum_{i=1}^{n} a_i\, \phi_{E_i},$$

then $\int_F \Psi d\mu = \sum_{i=1}^{n} a_i\, \mu(E_i \cap F)$

$$= \sum_{i=1}^{n} a_i \int_{E_i \cap F} g d\lambda = \int_F \Psi_g d\lambda.$$

Let $<\Psi_n>$ be a sequence of measurable simple function which converges to f, then

$$\gamma(F) = \int_F f d\mu = \lim \int_F \Psi_n d\mu.$$

Notes

$$= \lim \int_F \Psi_n g d\lambda = \int_F f_g d\lambda \text{ as } \Psi_n g \to fg$$

$$\Rightarrow \qquad = \frac{d\gamma}{d\lambda} = fg = \frac{d\gamma}{d\mu} \cdot \frac{du}{d\lambda}.$$

14.2 Summary

- Let (X, $\mathcal{A}$) be a measurable space and let r and m be measure functions, defined on the space (X, A). The measure γ is said to be absolutely continuous w.r.t. μ if

  μ (A) = 0 or $|\mu|$ (A) = 0, A $\in \mathcal{A} \Rightarrow \gamma$ (A) = 0, and is denoted by $\gamma \ll \mu$.

- Let (X, $\mathcal{A}$, μ) be a σ-finite measure space. If Y be a measure defined on A s.t. is absolutely continuous w.r.t. μ, then there exists a non-negative measurable function f s.t.

  $$\gamma(A) = \int_A f d\mu, \forall A \in \mathcal{A}$$

  The function f is unique in the sense that if g is any measurable function with the property defined as above, then f = g almost everywhere with respect to μ.

14.3 Keywords

Absolutely Continuous Measure Function: Let (X, $\mathcal{A}$) be a measurable space and let γ and μ be measure functions defined on the space (X, A). The measure γ is said to be absolutely continuous w.r.t. μ if

$$\mu(A) = 0 \text{ or } |\mu|(A) = 0, \ A \in \mathcal{A} \Rightarrow \gamma(A) = 0, \text{ and is denoted by } \gamma \ll \mu.$$

Radon-Nikodym Theorem: Let (X, $\mathcal{A}$, μ) be a σ-finite measure space. If γ be a measure defined on A s.t. γ is absolutely continuous w.r.t. μ, then there exists a non-negative measurable function f on s.t.

$$\gamma(A) = \int_A f \, d\mu, \forall A \in \mathcal{A}.$$

The function f is unique in the sense that if g is any measurable function with the property defined as above, then f = g almost everywhere with respect to μ.

14.4 Review Questions

1. Show that $\frac{d\gamma}{d\mu} = \left(\frac{d\mu}{d\gamma}\right)^{-1}$,

   where μ and γ are σ-finite signed measures and $\mu \ll \gamma, \gamma \ll \mu$.

2. If $\gamma(E) = \int_E f d\mu$, , where $\int_E f d\mu$ exists, then find $|\gamma|$ (E).

3. State and prove Radon-Nikodym theorem.

14.5 Further Readings

Notes

Book

G.E. Shilov and B.L. Gurevich, *Integral, Measure and Derivative: A Unified Approach*, Richard A. Silverman, trans. Dover Publications, 1978.

Online links

www.math.ksu.edu/nnagy/real-an/4-04-rn.pdf

mathworld.woltram.com

www.csun.edu

pioneer.netserv.chula.ac.th

Notes

Unit 15: 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.

15.1 Banach Spaces

15.1.1 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

Notes

(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

15.1.2 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 $\quad x_n \to x$ 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

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

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

Therefore

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

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

From (1) and (2) we get

$$\big| \| x \| - \| y \| \big| \leq \| x - y \|$$

This completes the proof of the theorem.

15.1.3 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.

Notes

Since x, y $\in \overline{M}$, 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 $\alpha = 0 = \beta$, 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 $\overline{[S]} = M$.

15.1.4 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 $x \in N$ such that

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

15.1.5 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.

Notes

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

$$\begin{aligned} | z_1 + z_2 |^2 &= (z_1 + z_2)\ (\bar{z}_1 + \bar{z}_2) \\ &= z_1 \bar{z}_1 + z_1 \bar{z}_2 + z_2 \bar{z}_1 + z_2 \bar{z}_2 \\ &\le |z_1|^2 + 2 |z_1 \bar{z}_2| + |z_2|^2 \end{aligned}$$

Notes

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

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

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

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

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

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

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

Notes

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

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

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 kth 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

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

Notes

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

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

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.

15.2 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

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

- 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.

15.3 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.

15.4 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

15.5 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 16: 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.

16.1 Continuous Linear Transformation

16.1.1 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.

Notes

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

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

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

Hence proved.

16.1.2 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

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

Notes

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

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

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

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 $\quad x_n \to 0$ 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.

16.1.3 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:

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

Notes

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

$$|f(x)| \le M \, \| x \|, \quad M \ge 0.$$

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

$$\frac{|f(x)|}{\|x\|} \le \|f\| \quad \forall \; x \ne 0.$$

or

$$|f(x)| \le \|f\| \; \|x\|.$$

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

If f, g ∈ N*, then

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

$$\le \sup_{\|x\| \ne 0} \frac{|f(x)|}{\|x\|} + \sup_{\|x\| \ne 0} \frac{|g(x)|}{\|x\|}$$

$$\Rightarrow \qquad \|f+g\| \le \|f\| + \|g\|.$$

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

$$|f(x)| \le M \|x\|, \; M \ge 0.$$

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

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

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

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

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

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

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

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

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

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

Notes

So that

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

Now consider

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

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

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

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

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

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

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

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

Notes

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

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

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

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n.$$

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

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

$\Rightarrow$ f determines and is determined by n scalars

$$y_i = f(e_i).$$

Then the mapping

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

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

By using Hölder's inequality, we get

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

so that

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

Using the definition of norm for f, we get

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

Consider the vector, defined by

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

Then

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

Since q = p (q – 1) we have from (4),

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

Now

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

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

Notes

So that

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

From (5) and (6), we get

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

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

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,

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

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

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

From the definition of norm for f, we have

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

Now consider the vector defined as follows:

If $|y_i| = \max_{1 \le i \le n}\{|y_i|\}$, 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

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

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

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

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

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

From (8) and (10), we obtain

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

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

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

(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

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

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

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

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

Notes

Consider the vector x defined by

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

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

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

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 $\quad \ell_p^* = \ell_q$.

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

Let $\ell_n = (0, 0, 0, \ldots, 1, 0, 0, \ldots)$ where 1 is in the mth place.

$$e_n \in \ell_p \text{ for } n = 1, 2, 3, \ldots$$

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 $(x_1, x_2, \ldots, x_n, 0, 0, 0, \ldots)$ in the form $\sum_{k=1}^{n} x_k e_k$ and

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

Notes

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

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

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

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

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

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

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

$T(f) = (\alpha_1, \alpha_2, \ldots, \alpha_k, \ldots)$ 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 $\beta_k = \begin{cases} |\alpha_k|^{g-1} \operatorname{sgn} \bar{\alpha}_k, & 1 \le k \le n \\ 0 & \forall\, n > k \end{cases}$

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

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

Notes

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

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

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

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

Since we can write

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

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

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

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.

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

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

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

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

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

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

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

$\Rightarrow$ 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

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

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

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

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

Notes

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

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

Hence, we have

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

which upon using definition of norm yields.

$$\|f\| \leq \|Tf\| \qquad \ldots (13)$$

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

$$m\|x\| \leq \|T(x)\| \ \forall x \in N. \qquad \ldots (1)$$

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

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

Hence (1) is equivalent to

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

$\Rightarrow$ 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

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

$\Rightarrow m \| x \| \le \| T(x) \|$ 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,

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

Let B be a bounded subset of N. Then

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

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

From above we see that

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

$\Rightarrow$ 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

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

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

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

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

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

This completes the proof of the theorem.

16.2 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

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

Notes

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

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

16.3 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 .

16.4 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 $\| T' \| = \| T \|$.

Notes

16.5 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 17: 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.

17.1 The Hahn-Banach Theorem

17.1.1 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,

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

$$= f_o (x + y + (\alpha + \beta) x_o)$$

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

Since f is linear M, $f (x + y) = f (x + y)$ … (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 $f_o(ky) = k[f(x) + \alpha r_o] = k f_o(y)$ … (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, $\| f_o \| \geq \| f \|$ … (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 $- f(x_1) - \| f \| \| x_1 + x_o \| \leq - f(x_2) + \| f \| \| x_2 + x_o \|$ … (6)

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

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

Choose r_o to be any real number such that

$$\begin{aligned} \sup_{y \in M} \left\{ -f(y) - \| f \| \left(\| y + x_o \| \right) \right\} &\leq r_o \\ &\leq \inf_{y \in M} \left\{ -f(y) + \| f \| \left(\| y + x_o \| \right) \right\} \end{aligned}$$

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

Notes

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

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

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

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

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

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

$$|f_o(z)| \le \|f\|\|z\| \quad \ldots (8)$$

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

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

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

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

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

Replacing z by –z in (9) we get

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

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

$$|f_o(z)| \le \|f\|\|z\| \quad \ldots (11)$$

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

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

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

$$\|f_o\| \le \|f\|$$

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

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

Notes

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

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

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

Notes

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 $f_o(x) = re^{i\theta}$

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

Since $f_o(x)$ is linear,

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

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

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

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

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

We get $$|f_o(x)| \le \| g_o \| \| x \|$$

Since g_o is the extension of g, we get

$$\| g_o \| \| x \| = \| g \| \; \| x \| \le \| f \| \| x \|.$$

Therefore

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

$$\| f_o \| \le \| f \|$$

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

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

This completes the proof of the theorem.

17.1.2 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

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

and $$p(\alpha x) = |\alpha|\, p(x) \;\; \forall \;\; x \in L \text{ and scalar } \alpha.$$

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

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

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

Notes

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

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 $g_o(x) \le p(x)\ \forall\ x \in L$.

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

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

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

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

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

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

It follows from (1) and (2) that

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

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 $p(x) \ge 0$ for all $x \in N$.

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

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

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

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

$$= p(x) + p(y)$$

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

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 $|f_o(x)| \le p(x) = \|f\|\ \|x\|\ \forall\, x \in N$.

$\Rightarrow$ $$\|f_o\| \le \|f\| \qquad \ldots (3)$$

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

$$\|f\| \le \|f_o\| \qquad \ldots (4)$$

From (3) and (4) it follows that

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

Notes

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 $$f(x_1 + x_2) = f(x_1) + f(x_2) \qquad \dots (1)$$

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

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

Notes

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 $x = 0$.

17.2 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 $x \in L$.

17.3 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)

Notes

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

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

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

17.4 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.

17.5 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 18: The Natural Imbedding of N in N**

Notes

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**.

18.1 The Natural Imbedding of N into N**

18.1.1 Definition: Natural Imbedding of N into N**

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

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

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.

Notes

18.1.2 Definition: 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). 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.

18.1.3 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

$J : N \to N^{**}$ 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.

18.1.4 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 $J : N \to N^{**} : J(x) = F_x$ 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 $\in$ N* and α, β be scalars. Then

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

$\Rightarrow$ F_x is linear

F_x is bounded.

For any f $\in$ N*, we have

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

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 $\|F_x\| \leq \|x\| \qquad \ldots (2)$

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 $x \neq 0$ be a vector in N. Then by theorem (If N is a normal linear space and $x_o \in N$, $x_o \neq 0$, then there exists a functional $f_o \in N^*$ such that

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

$\exists$ a functional f $\in$ 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 $\|F_x\| \geq \|x\| \qquad \ldots (3)$

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

From (2) and (3); we have

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

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

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

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

$\Rightarrow$ $F_{x+y} = F_x + F_y \qquad \ldots (6)$

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

Notes

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

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

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

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

(8) and (9) $\Rightarrow$ J is linear and also (4) shows that J is norm preserving.

For any x and y in N, we have

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

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. $$J(x) = J(y) \Rightarrow x = y \text{ so that J is one-one.}$$

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 $\ell^*_q = \ell_p$

$\Rightarrow$ $\ell^{**}_q = \ell_p$.

$\Rightarrow$ 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 $\{x_1, x_2, \ldots, x_m\}$ is a basis for N, and if $(\alpha_1\ \alpha_2 \ldots \alpha_m)$ 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 $x = x_i$.

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

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

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

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

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

$\Rightarrow$ $f_j(x) = \beta_j$

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

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

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

$\Rightarrow$ N* is m-dimensional.

18.2 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 $\| F_x \| = \| x \|$.

18.3 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**.

Notes

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).

18.4 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.

18.5 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

Notes

Unit 19: The Open Mapping Theorem

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.

19.1 The Open Mapping Theorem

19.1.1 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 $j : B' \to B'$ and $g : B' \to B'$

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 $g : B' \to B' : g(x) = 2n_0 y$ 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) $\quad f\left(\overline{T(S_{n_0})}\right) = \overline{f\left(T(S_{n_0})\right)}$

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

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

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

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

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

Notes

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

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

Also for $n > m$, we have

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

Notes

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

$$\begin{aligned} T(x) &= T\left(\lim_{n\to\infty} s_n\right) \\ &= \lim_{n\to\infty} T(s_n) \\ &= \lim_{n\to\infty}(y_1 + y_2 + \ldots + y_n) \\ &= y \end{aligned}$$

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

$$\begin{aligned} y \in S(x; r) &\Rightarrow \|y - x\| < r \\ &\Rightarrow \|z\| < r \text{ and } y - x = z, z \in S_r \\ &\Rightarrow y = x + z \text{ and } \|z\| < r \\ &\Rightarrow y \in x + S_r \end{aligned}$$

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

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

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

$$= r\,S_1$$

Thus $\quad S_r = r\,S_1$

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

Notes

19.1.2 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 $\Rightarrow$ 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.

19.1.3 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

$$m \| x \| \le \| T(x) \| \le M \| x \|.$$

Notes

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

$$T : N \to B' \text{ defined as follows.}$$

For each $f \in N$, let $T(f) = \Sigma f(i) \alpha_i$.

Notes

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

$$\| y_{n_{k+1}} - y_{n_k} \| < \frac{1}{2^k}$$

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 $(y_{n_{k+1}} - y_n) \in N$, there is a $n_k \in B$ and a constant M such that

$T(x_k) = y_{n_{k+1}} - y_n$ and $\| x_k \| \leq \| y_{n_{k+1}} - y_n \|$.

By on choice $\sum_{k=1}^{\infty} \| y_{n_{k+1}} - y_{n_k} \|$ 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.

Notes

Also $\| x \|_1 \le a \| x \|_2 \Rightarrow$ 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 $(N, \| x \|_1) \to (N, \|x \|_2)$

Since $i^{-1}(x) = x$, $\exists a, b$ s.t. $\| x \|_2 \le b \| \ \|_1$ … (2)

$\therefore$ (1) & (2) imply that the norms are equivalent.

19.2 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.

19.3 Keywords

Banach Space: A normed space V is said to be Banach space if for every Cauchy sequence $\{\nu_n\}_{n=1}^{\infty} \subset V$ then there exists an element $\nu \in V$ such that $\lim_{n\to\infty} \nu_n = \nu$.

Homeomorphism: A map $f : (X, T) \to (Y, U)$ 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.

19.4 Review Questions

1. If X and Y are Banach spaces and $A : X \to Y$ 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 $T_1 = T_2$.

19.5 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

Notes

Unit 20: 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.

20.1 The Closed Graph Theorem

20.1.1 Graph of Linear Transformation

Definition: Let N and N′ be a normed 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 consists of all ordered pairs (x, T (x)). It is generally denoted by G_T.

Therefore the graph of $T : N \to N'$ 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 $\in > 0$, we can find a n_o such that

$$\| (x_n, y_n) - (x_m, y_m) \| < \in \ \forall \ m, n \geq n_o. \qquad \ldots (1)$$

$\Rightarrow \| (x_n - x_m) \| < \in$ and $\| y_n - y_m \| < \in \ \forall \ m, n \geq n_o$

⇒ (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. $$\| (x_n - x_o) \| < \in \text{ and } \| y_n - y_m \| < \in \ \forall \ m, n \geq n_o. \qquad \ldots (2)$$

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))

⇒ $(x_n, y_n) \to (x_o, y_o)$ 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

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)

20.1.2 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 $\overline{G_T} = G_T$ ($\because G_T$ 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.

20.1.3 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

$$x_n \to x \text{ in B and } T(x_n) \to T(x) \text{ in B}' \qquad \dots (1)$$

Notes

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 $\|x_n - x\|_1 = \|x_n - x\| + \|T(x_n - x)\|$

$= \|x_n - x\| + \|T(x_n) - T(x)\|$

$= \|x_n - x\| + \|T(x_n - y)\| \to 0$ as $n \to \infty$. (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 $(x, y) \in G_T \Leftrightarrow (y, x) \in G_{T'}$.

$\Rightarrow$ $G_{T'}$ is closed in $B' \times B$.

$\Rightarrow$ T′ is continuous (By closed graph theorem)

$\Rightarrow$ 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.

20.2 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.

Notes

- 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 $y = T(x)$.
- 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.

20.3 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 $y = T(x)$.

20.4 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 $B = M \oplus N$.

20.5 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

Notes

Unit 21: The Conjugate of an Operator

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 $T \to T^*$.

21.1 The Conjugate of an Operator

21.1.1 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

$$\begin{aligned} f(T(\alpha x + \beta y)) &= \alpha f(T(x) + \beta f(T(y)) \\ &= \alpha [T'(f)](x) + \beta [T'(f)]y \\ &\Rightarrow \text{part (a).} \end{aligned}$$

(b) Let $f, g \in N^+$ and α, β be any scalars. Then

$$\begin{aligned} [T'(\alpha f + \beta g)(x)] &= (\alpha f + \beta g)(T(x)) = \alpha [T'(f)](x) + \beta (T'(g)](x) \\ &\Rightarrow T' \text{ is linear on } N^+ \\ &\Rightarrow \text{part (b)} \end{aligned}$$

(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

$$\begin{aligned} N^* &= T'(f) \text{ is in } N^* \text{ for each f in } N^*. \\ &\Rightarrow T'(N) \subseteq N^* \\ &\Rightarrow \text{part (c)} \end{aligned}$$

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.

21.1.2 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.

$$\| T^* \| = \sup \{ \| T^* (f) \| : \| f \| \le 1\}$$

$$= \sup \{ \| [T^* (f)] (x) \| : \| f \| \le 1 \text{ and } \| x \| \le \}$$

$$= \sup \{ | f (T (x)) | : \| f \| \le 1 \text{ and } \| x \| \le \}$$

$$= \sup \{ \| f \| \| T \| x \| : \| x \| \le 1 \text{ and } \| x \| \le \} \qquad \dots (1)$$

$$\le T$$

$\Rightarrow$ 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

$$\| f \| = 1 \text{ and } f (T (x)) = \| T (x) \| \qquad \dots (2)$$

Hence
$$\| T \| = \sup \left\{ \frac{\| T(x) \|}{\| x \|} : x \ne 0 \right\}$$

$$= \sup \left\{ \frac{\| f(T(x)) \|}{\| x \|} : \| f \| = 1, \ x \ne 0 \right\} \qquad \text{(by (2))}$$

$$= \sup \left\{ \frac{\| T^* (f) (x) \|}{\| x \|} : \| f \| = 1, \ x \ne 0 \right\} \qquad \text{(by (1))}$$

$$\le \sup \left\{ \frac{\| T^* (f) \| \ \| x \|}{\| x \|} : \| f \| = 1, \ x \ne 0 \right\}$$

$$= \sup \{ \| T^* (f) \| : \| f \| = 1 \} = \| T^* \| \qquad \dots (3)$$

From (1) and (3) it follows that

$$\| T \| = \| T^* \|. \qquad \dots (4)$$

Notes

Now we show that

$$\left\{\begin{array}{l}\phi : \beta(N) \to \beta(N^*) \text{ given by} \\ \phi(T) = T^*\end{array}\right\} \qquad \ldots (5)$$

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

$$\phi(\alpha T + \beta T_1) = (\alpha T + \beta T)^* \qquad \text{by (3)}$$

But $\quad [(\beta T + \beta T_1)^* (f)] (x) = f(\alpha T + \beta T_1)(x)$

$$= f(\alpha T(x) + \beta T_1(x))$$

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 $\quad (\alpha T + \beta T_1)^* = \alpha T^* + \beta T_1^* \qquad \ldots (6)$

Therefore $\quad \phi(\alpha T + \beta T_1) = (\alpha T + \beta T_1)^* = \alpha T^* + \beta T^*_1 = \alpha \phi(T) + \beta \phi(T_1)$

$\Rightarrow \phi$ 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 $\beta(N)$ onto $\beta(N^*)$.

Finally we show that ϕ reverses the product and preserves the identity transformation.

Now $\quad [(T T_1)^* (f)] (x) = f((T T_1)(x))$

$$= f(T(T_1(x))$$

$$= [T^*(f)] [T_1(x)], \text{ since } T_1(x) \in N \text{ and } T^*(f) \in N^*.$$

Notes

$$= [\dot{T}^*_1 (T^* (f))] (x)$$

$$= [(\dot{T}^*_1 \; T^*) (f)] (x)$$

Hence, we get

$$(T\, T_1)^* = \dot{T}^*_1 \; T^* \text{ so that}$$

$$\phi (T\, T_1) = (T\, T_1)^* = \dot{T}^*_1 \; T.$$

$\Rightarrow \phi$ 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 $\quad T^{**} (J (x))x' = J (T (x))x'.$

$\Rightarrow \quad T^{**} . J = J\, T$

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 $\Leftrightarrow$ 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

$$\therefore \quad (TT^{-1})^* = (T^{-1} T)^* = I^*$$

$$(T^{-1})^*T = T^* (T^{-1})^* = I$$

$\Rightarrow (T^*)^{-1}$ exists and it is given by $(T^*)^{-1} = (T^{-1})^*$. This completes the proof of the theorem.

21.2 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.

21.3 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^+.

21.4 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.

Notes

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 $(T^*)^{-1} = (T^{-1})^*$.

21.5 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 22: 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.

22.1 The Uniform Boundedness Theorem

22.1.1 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

$$x_k \to y \text{ as } k \to \infty$$

Notes

$\Rightarrow \quad T_i x_k \to T_{iy}$ 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)

$\Rightarrow \quad \| T_i y \| = \underset{k\to\infty}{\text{lt}} \| T_i x_k \|$

$\leq n \ \forall \ i$ $(\because x_k \in F_n)$

$\Rightarrow \quad y \in F_n.$

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 $\exists$ a real number $k \geq 0$ 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 $\quad s_o \in S_o \subseteq F_{n_o}$.

$\therefore \quad \| y \| = \| T_i s_o \| \leq n_o.$

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

$$\| T_i (S) \| = \left\| T_i \left(\frac{S_o - x_o}{r_o} \right) \right\|$$

$$= \frac{1}{r_o} \| T_i (S_o) - T_i (x_o) \|$$

$$\leq \frac{1}{r_o} \left(\| T_i (S_o) \| + \| T_i (x_o) \| \right)$$

$$\leq \frac{2n_o}{r_o} \forall i .$$

Hence $\quad \| T_i \| \leq \dfrac{2n_o}{r_o} \forall i .$

This completes the proof of the theorem.

Notes

22.1.2 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 $x \in B$, 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 $x \in B$, consider the vector $\dfrac{r_o x}{\|x\|}$.

Then $|f_n(x)| = \left| \dfrac{\|x\|}{r_o} f_n\left(\dfrac{r_o x}{\|x\|}\right) \right| \le \dfrac{\|x\|}{r_o}(m+k)$ so that $\dfrac{|f_n(x)|}{\|x\|} \le \left(\dfrac{m+k}{r_o}\right)$.

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 $\| f_n \| \ge \dfrac{| f_n(x) |}{\| x \|} = n$.

$\Rightarrow (\| f_n \|)$ 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 $\exists$ 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

$$\| T(x) \| \le \lambda \| x \|$$

$\Rightarrow$ 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 $\sum_{n=1}^{\infty} | a_n | < \infty$.

Notes

Solution: For every $x \in c_o$, 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 $\Leftrightarrow$ f (X) is a bounded set of numbers for each f in N*.

Proof: Let X be a bounded subset of N so that $\exists$ a positive constant λ_1 such that

$$\| x \| \le \lambda_1 \qquad \forall\ x \in X \qquad \ldots (1)$$

To show that f (X) is bounded for each $f \in N^*$. Now $f \in N^* \Rightarrow$ f is bounded.

$$\Rightarrow \qquad \exists\ \lambda_2 > 0 \text{ such that } | f(x) | \le \lambda_2 \| x \| \quad \forall\ x \in N \qquad \ldots (2)$$

It follows from (1) & (2) that

$$| f(x) | \le \lambda_1 \lambda_2 \qquad \forall\ x \in X.$$

$\Rightarrow$ f (X) is a bounded set of real numbers for each $f \in N^*$.

Conversely, let us assume that f (X) is a bounded set of real numbers for each $f \in N^*$.

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 $f \in N^*$ is equivalent to the assumption that $\{F_{x_i}(f)\}$ is bounded set for each $f \in N^*$.

Since N* is complete $\Rightarrow (F_{x_i})$ is bounded subset of N** by uniform boundedness theorem.

That is, $(\| F_{x_i} \|)$ 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 $T(x) \in N'$ for every $x \in N$ and f is a functional on N′ so that f (T (x)) is well defined and $f \circ T \in N^*$. Since T is continuous and f is continuous, f o T is continuous on N.

Notes

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*.

22.2 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.

22.3 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.

22.4 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, 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 $<h_n, h> \to 0$ for every h in $\mathcal{H}$ if and only if sup $\{\| h_n \| : n \geq 1\} < \infty$ and $<h_n, e> \to 0$ for every e in $\mathcal{E}$.

22.5 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

Notes

Unit 23: 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.

23.1 Hilbert Spaces

23.1.1 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:

Notes

(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)

$$\begin{aligned}(\alpha x - \beta y, z) &= (\alpha x + (-\beta) y, z)\\ &= \alpha (x, z) + (-\beta) (y, z)\\ &= \alpha (x, z) - \beta (y, z).\end{aligned}$$

(b)

$$\begin{aligned}(x, \beta y + \gamma z) &= \overline{(\beta y + \gamma z, x)} = \overline{(\beta y, x) + (\gamma z, x)}\\ &= \overline{\beta (y, x) + \gamma (z, x)}\\ &= \bar{\beta} (x, y) + \bar{\gamma} (x, z)\end{aligned}$$

(c)

$$\begin{aligned}(x, \beta y - \gamma z) &= (x, \beta y + (-\gamma) z) = \bar{\beta} (x, y) + \overline{(-\gamma)} (x, z)\\ &= \bar{\beta} (x, y) - \bar{\gamma} (x, z)\end{aligned}$$

(d)

$$(0, x) = (0\theta, x) = 0 (\theta, x) = 0, \text{ where } \theta \text{ is the zero element of } x \text{ and } \overline{(x, 0)} = \overline{(0, x)} = \bar{0} = 0 .$$

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

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)
$$(\alpha x + \beta y, z) = \sum_{i=1}^{n} \left(\alpha x_i + \beta \bar{y}_i\right)\bar{z}_i$$
$$= \sum_{i=1}^{n} \alpha x_i \bar{z}_i + \sum_{i=1}^{n} \beta y_i \bar{z}_i$$
$$= \alpha (x, z) + \beta (y, z)$$

(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.

23.1.2 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.

Notes

Theorem 2: If x and y are any two vectors in an inner product space then

$$|(x, y)| \le \|x\| \, \|y\| \qquad \dots (1)$$

Proof: If y = 0, we get $\|y\| = 0$ and also theorem 1 implies that $|(x, y)| = 0$ so that (1) holds.

Now, let $y \neq 0$, then for any scalar $\lambda \varepsilon C$ we have

$$0 \le \|x - \lambda y\|^2 = (x - \lambda y, x - \lambda y)$$

But

$$(x - \lambda y, x - \lambda y) = (x, x) - (x, \lambda y) - (\lambda y, x) + (\lambda y, \lambda y)$$

$$= (x, x) - \bar{\lambda}\,(x, y) - \lambda\,(y, x) + \lambda\bar{\lambda}\,(y, y)$$

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

$$\Rightarrow \|x\|^2 - \lambda\,(y, x) - \bar{\lambda}\,(x, y) + |\lambda|^2 \|y\|^2 = (x, \lambda y, x - \lambda y)$$

$$= \|x - \lambda y\|^2 \ge 0 \qquad \dots (2)$$

Choose $\lambda = \dfrac{(x,y)}{\|y\|^2}, y \neq 0, \|y\| \neq 0$.

$\therefore$ 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) $\|x\| = \sqrt{(x, x)} \Rightarrow \|x\|^2 = (x, x) \ge 0$ and $\|x\| = 0 \Leftrightarrow x = 0$.

(ii) Let $x, y \in X$, then

$$\|x + y\|^2 = (x + y, x + y)$$

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

$$= \|x\|^2 + (x, y) + (\overline{x, y}) + \|y\|^2$$

$$= \|x\|^2 + 2\mathrm{Re}\,(x, y) + \|y\|^2 \qquad [\because (x,y) + (\overline{x,y}) = 2\,\mathrm{Re}(x,y)]$$

$$\le \|x\|^2 + 2\,|(x, y)| + \|y\|^2 \qquad [\because \mathrm{Re}\,(x, y) \le |(x, y)|]$$

Notes

$$\leq \| x \|^2 + 2 \| x \| \; \| y \| + \| y \|^2 \qquad \text{[using Schwarz inequality]}$$

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

Therefore $\| x + y \| \leq \| x \| + \| y \|$

(iii) $\| \alpha x \|^2 = (\alpha x, \alpha x)$

$$= \alpha \bar{\alpha} \; (x, x)$$

$$= | \alpha |^2 \; \| x \|^2$$

$\Rightarrow$ $\| \alpha, x \| = \| \alpha \| x \|$

(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.

23.1.3 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 $\alpha, \beta \in C$.

23.1.4 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 $(x, y) = \sum_{n=1}^{\infty} x_n \bar{y}_n$, 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 $\sum_{n=1}^{\infty} x_n \bar{y}_n$ is convergent and this defines a complex number.

By Cauchy inequality, we have

$$\sum_{i=1}^{n} |x_i \bar{y}_i| \le \left\{\sum_{i=1}^{n} |x_i|^2\right\}^{\frac{1}{2}} \left\{\sum_{i=1}^{n} |y_i|^2\right\}^{\frac{1}{2}}$$

$$\le \left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{\frac{1}{2}} \left\{\sum_{n=1}^{\infty} |y_n|^2\right\}^{\frac{1}{2}}.$$

Since $\sum_{n=1}^{\infty} |x_n|^2$ and $\sum_{n=1}^{\infty} |y_n|^2$ are convergent, the sequence of partial sum $\sum_{i=1}^{n} |x_i \bar{y}_i|$ is a monotonic increasing sequence bounded above. Therefore, the series $\sum_{n=1}^{\infty} |x_i \bar{y}_i|$ is convergent. Hence $\sum_{n=1}^{\infty} |x_n \bar{y}_n|$ is absolutely convergent having its sum as a complex number.

Therefore $(x, y) = \sum_{n=1}^{\infty} x_n \bar{y}_n$ 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

$$\| (x + y) \|^2 = (x + y, x + y) \qquad \text{(By def. of Hilbert space)}$$

$$= (x, x + y) + (y, x + y)$$

$$= (x, x) + (x, y) + (y, x) + (y, y)$$

$$= \| x \|^2 + (x, y) + (y, x) + \| y \|^2 \qquad \ldots (1)$$

$$\| x - y \|^2 = (x - y, x - y)$$

$$= (x, x - y) - (y, x - y)$$

$$= (x, x) - (x, y) - (y, x) + \| y \|^2 \qquad \ldots (2)$$

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

$$\begin{aligned} |(x_n, y_n) - (x, y)| &= |(x_n, y_n) - (x_n, y) + (x_n, y) - (x,, y)| \\ &= |(x_n, y_n - y) + (x_n - x, y)| \\ &\qquad \text{(by linearity property of inner product)} \\ &\le |(x_n, y_n - y)| + |(x_n - x, y)| \qquad [\because |\alpha + \beta| \le |\alpha| + |\beta|] \\ &\le \| x_n \| + \| y_n - y \| + \| x_n - x \| \| y \| \qquad \text{[By Schwarz inequality]} \end{aligned}$$

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 $\| x_n \| \le M \ \forall\, n$.

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

$$\left\| \frac{x_m + x_n}{2} \right\| \ge \delta \Rightarrow \| x_m + x_n \| \ge 2\delta \qquad \dots (2)$$

Using parallelogram law, we get

$$\| x_m + x_n \|^2 + \| x_m - x_n \|^2 = 2 \| x_n \|^2 + 2 \| x_n \|^2$$

or

$$\begin{aligned} \| x_m - x_n \|^2 &= 2 \| x_m \|^2 + 2 \| x_n \|^2 - \| x_m - x_n \|^2 \\ &\le 2 \| x_m \|^2 + 2 \| x_n \|^2 - d\delta^2 && \text{(Using (2))} \\ &\to 0 \text{ as } m, n \to \infty && \text{(Using (1))} \end{aligned}$$

$\Rightarrow$ $\| x_m - x_n \|^2 \to 0$ 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 $\Rightarrow \dfrac{x+y}{2} \in E$

$$\Rightarrow \quad \left\|\frac{x+y}{2}\right\| \geq \delta \qquad \ldots (3)$$

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) = $\dfrac{t-a}{b-a}$. Then $\|x\| = 1$, $\|y\| = 1$

Now x (t) + y (t) = 1 + $\dfrac{t-a}{b-a}$ so that $\|x + y\| = 2$

x (t) – y (t) = 1 – $\dfrac{t-a}{b-a}$ 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

23.2 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.

23.3 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) $(x, x) \geq 0$, $(x, x) = 0 \Leftrightarrow x = 0$

23.4 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.

23.5 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

Notes

Unit 24: Orthogonal Complements

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.

24.1 Orthogonal Complement

24.1.1 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

$$(\alpha x, y) = \alpha(x, y)$$
$$= \alpha \cdot 0$$
$$= 0$$

$\Rightarrow \quad x \perp y \Rightarrow \alpha x \perp y.$

(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.

24.1.2 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,
$$\| x - y \|^2 = (x - y, x - y)$$
$$= (x, x) - (x, y) - (y, x) + (y, y)$$
$$= \| x \|^2 - 0 - 0 - \| y \|^2$$
$$= \| x \|^2 + \| y \|^2$$

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

24.1.3 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.

Notes

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 $y \in S_2$.

24.1.4 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 $S \cap S^\perp = \{0\}$.

Notes

(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.

Notes

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

$$M^{\perp} \cap M^{\perp\perp} = \{0\} \quad \text{… (2)}$$

(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

$$\| z_m - z_n \|^2 = \| x_m - x_n \|^2 + \| y_m - y_n \|^2 \quad \dots (1)$$

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.

24.1.5 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),

$$N^{\perp\perp} = N \quad \dots (2)$$

From (1) and (2), we have

$$N = M + M^\perp = H.$$

Notes

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 $-2\beta \, |(z_o, y)|^2 + \beta^2 \, |(z_o, y)|^2 \| y \|^2 \geq 0$

Notes

or $\beta \mid (z_o, y) \mid^2 \{\beta \| y \|^2 - 2\} \geq 0$ … (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 $\beta \mid (z_o, y) \mid^2 \{\beta \| y \|^2 - 2\} < 0$.

This contradicts (2).

Hence we must have $(z_o, y_o) = 0 \Rightarrow z_o \perp y, \ \forall y \in M$.

$\therefore \quad z_o \perp M$.

This completes the proof of the theorem.

24.2 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 $H = M \oplus M^\perp$.

24.3 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

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

24.4 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 $H \Leftrightarrow S^\perp = \{0\}$.

Notes

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.

24.5 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 25: 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.

25.1 Orthonormal Sets

25.1.1 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.

25.1.2 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 $(e_i, e_j) = 1$ for every i.

Notes

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.

25.1.3 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 ith 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 $\sum_{i=1}^{n} x_i \bar{y}_i$ 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 nth place and O's elsewhere is an orthonormal set.

25.1.4 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

$$\sum_{i=1}^{n} \left| (x, e_i) \right|^2 \leq \| x \|^2; \qquad \ldots (1)$$

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

$$y = x - \sum_{i=1}^{n} (x, e_i)\, e_i$$

Notes

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 $\| x \|^2 - \sum_{i=1}^{n} |(x, e_i)|^2 \geq 0$

$$\Rightarrow \qquad \sum_{i=1}^{n} |(x, e_i)|^2 \leq \| x \|^2$$

$\Rightarrow$ 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

$$S_n = \left\{ e_i : \left|(x, e_i)\right|^2 > \frac{\|x\|^2}{n} \right\}.$$

If the set S_n contains n or more than n' vectors, then we must have

$$\sum_{e_i \in S_n} \left|(x, e_i)\right|^2 > n\frac{\|x\|^2}{n} = \|x\|^2 \qquad \ldots (1)$$

By theorem (1), we have

$$\sum_{e_i \in S_n} \left|(x, e_i)\right|^2 \leq \|x\|^2 \qquad \ldots (2)$$

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 $S = \{e_1, e_2, \ldots, e_n\}$ for some positive integer n.

In this case, we have

$$\Sigma|(x, e_i)|^2 = \sum_{i=1}^{n} |(x, e_i)|^2 \leq \|x\|^2 \qquad \ldots (1)$$

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 $\{e_1, e_2, \ldots, e_n, \ldots\}$.

In this case we can write

$$\Sigma |(x, e_i)|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2 \qquad \ldots (2)$$

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

$$\sum_{i=1}^{n} |(x, e_i)|^2 \leq \|x\|^2 \qquad \ldots (3)$$

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 $e_i \in S$,

$$\Sigma |(x, e_i)|^2 = \sum_{n=1}^{\infty} |(x, e_i)|^2 \leq \|x\|^2$$

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 $|(x, e_n)|^2 \to 0$ as $n \to \infty$.

$\Rightarrow$ $(x, e_n) \to 0$ 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

$$S = \{e_1, e_2, \ldots, e_n\} \text{ and we define}$$

$$\Sigma\,(x, e_i)\, e_i = \sum_{i=1}^{n} (x, e_i)\, e_i,$$

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, $$\| s_m - s_n \|^2 = \left\| \sum_{i=n+1}^{m} \left| (x, e_i) e_i \right| \right\|^2$$

$$= \sum_{i=n+1}^{m} \left| (x, e_i) \right|^2$$

By Bessel's inequality, the series $\sum_{n=1}^{\infty} \left| (x, e_n) \right|^2$ is convergent.

Hence $\sum_{i=n+1}^{\infty} \left| (x, e_i) \right|^2$ 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_0} also.

Hence $s'_{m_0} - s_{n_0}$ 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 $\| s'_{m_0} - s_{n_0} \| \leq \sum_{i=n_0+1}^{\infty} |x, e_i|^2 < \in^2$ so that we have

$$\| s'_{m_0} - s'_{n_0} \| < \in \qquad \ldots (2)$$

Now $$\| s' - s \| = \| (s' - s'_{m_0}) + (s'_{m_0} - s'_{n_0}) + (s_{n_0} - s) \|$$

$$\leq \| s' - s'_{m_0} \| + \| s'_{m_0} - s'_{n_0} \| + \| s_{n_0} - s \|$$

$$< \in + \in + \in = 3\in \qquad \text{(Using (1) and (2))}$$

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

$$(s_n, e_j) = \left(\sum_{i=1}^{n} (x, e_i)\, e_i, e_j \right) = 0$$

$$\Rightarrow \qquad \lim_{n \to \infty} (s_n, e_j) = 0$$

Notes

Hence $(x - \Sigma (x, e_i) e_i, e_j) = (x, e_j) = 0$ since $e_j \notin S$.

$$\text{If } e_j \in S, \text{ then } (s_n, e_j) = \left(\sum_{i=1}^{n} (x, e_i) e_i, e_j\right) \quad \ldots (5)$$

$$\text{If } n > j, \text{ we get } \left(\sum_{i=1}^{n} (x, e_i) e_i, e_j\right) = (x, e_j) \quad \ldots (6)$$

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.

Notes

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 $i \ne k$ 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 $S = \left\{e_i : (x, e_i)(\overline{y, e_i}) \ne 0\right\}$

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

$$\leq \left\{\sum_{i=1}^{n}|(x, e_i)|^2\right\}^{\frac{1}{2}} \left\{\sum_{i=1}^{n}|(y, e_i)|^2\right\}^{\frac{1}{2}} \quad \text{(By Cauchy inequality)}$$

$$\leq \|x\|^2 \; \|y\|^2 \quad \text{(by Bessel's inequality for finite case)}$$

$$\therefore \quad \sum_{i=1}^{n}|(x, e_i)|\left|\overline{(y, e_i)}\right| \leq \|x\| \; \|y\| \quad \ldots (1)$$

Finally let S is countably infinite. Let the vectors in S be arranged in a definite order as

$$S = \{e_1, e_2, \ldots, e_n, \ldots\}.$$

Let us define

$$\sum|(x, e_i)|\left|\overline{(y, e_i)}\right| = \sum_{i=1}^{\infty}|(x, e_n)|\left|\overline{(y, e_n)}\right|.$$

But this sum will be well defined only if we can show that the series $\sum_{n=1}^{\infty}|(x, e_n)|\left|\overline{(y, e_n)}\right|$ 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

$$\sum_{n=1}^{\infty}|(x, e_n)|\left|\overline{(y, e_n)}\right| \leq \|x\| \; \|y\| \quad \ldots (2)$$

From (2), we see that the series $\sum_{n=1}^{\infty}|(x, e_n)|\left|\overline{(y, e_n)}\right|$ 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

$$\sum|(x, e_i)|\left|\overline{(y, e_i)}\right| = \sum_{n=1}^{\infty}|(x, e_n)|\left|\overline{(y, e_n)}\right|$$

and from (2), we see that this sum is $\leq \|x\| \; \|y\|$.

25.2 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 $e_1, e_2, \ldots, e_n$ where e_i is the i-tuple with 1 in the ith 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.

Notes

(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 nth place and O's elsewhere is an orthonormal set.

25.3 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.

25.4 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.

25.5 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

Notes

Unit 26: The Conjugate Space H^*

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. $\exists$ has a natural correspondence between H and H^{**} and this natural correspondence is an isometric isomorphism of H onto H^{**}.

26.1 The Conjugate Space H^*

26.1.1 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.

Notes

26.1.2 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 $x \in H$

$$|fy(x)| = |(x, y)| \le \|x\|.\|y\| \qquad ...(1)$$

(Schwarz inequality)

Let $\|y\| \le M$. Then for M > 0

$|fy(x)| \le M\|x\|$ so that fy is bounded and hence fy is continuous.

Now let y = 0, $\|y\| = 0$ and from the definition fy = 0 so that $\|fy\| \le \|y\|$.

Further let $y \ne 0$. Then from (1) we have $\dfrac{\text{Sup } |fy(x)|}{\|x\|} \le \|y\|$.

Hence using the definition of the norm of a functional,

we get $\|fy\| \le \|y\|$ ...(2)

Further $\|fy\| = \sup\{|fy(x)| : \|x\| \le 1\}$ ...(3)

Since $y \ne 0, \left(\dfrac{y}{\|y\|}\right)$ is a unit vector.

From (3), we get

$$\|fy\| \ge \left|fy\left(\frac{y}{\|y\|}\right)\right| \qquad ...(4)$$

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∀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, $\exists\, ay_0\neq 0$ 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

$$0=(z,y_0)=(f(x)y_0-f(y_0)x,y_0)$$

$$=f(x)(y_0,y_0)-f(y_0)(x,y_0)$$

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),

$$f(x) = \left| \frac{f(y_0)}{\|y_0\|^2} \right| (x, y_0) \qquad ...(4)$$

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

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

$$\Rightarrow \qquad \sup_{\|x\| \neq 0} \frac{|f(x)|}{\|x\|} \leq \|y\|.$$

Using definition of norm of f, we get from above

$$\|f\| \leq \|y\| \qquad ...(5)$$

Now let us take x = y in f(x) = (x,y), we get

$$\|y\|^2 = (y, y) = f(y) \leq \|f\| \|y\|$$

$$\Rightarrow \qquad \|y\| \leq \|f\| \qquad ...(6)$$

(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 $(x, y_1) - (x, y_2) = 0$

$$\Rightarrow \qquad (x, y_1 - y_2) = 0 \ \forall x \in H.$$

Notes

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

$$|f(x)|^2 \le \left(\sum_{n=1}^{\infty} \frac{|x_n|}{n}\right)^2 \le \left(\sum_{n=1}^{\infty} \frac{1}{n^2}\right)\left(\sum_{n=1}^{\infty} |x_n|^2\right)$$

$$\le \frac{\pi^2}{6}(x, x) = \frac{\pi^2}{6}\|x\|^2,$$

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.

Notes

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

$$f(x) = (x, y) = \sum x_n \bar{y}_n = \bar{y}_n \text{ as } x = e_n.$$

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 $\phi : H \to H^*$ defined by $\phi : H \to H^*$ 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

$$f_{y_1+y_2}(x) = (x, y_1 + y_2) = (x, y_1) + (x, y_2)$$

$$= f_{y_1}(x) + f_{y_2}(x)$$

$$\Rightarrow \quad f_{y_1+y_2} = \phi(y_1 + y_2) = f_{y_1} + f_{y_2} = \phi(y_1) + \phi(y_2)$$

(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

$\phi(y_1) = \phi(y_2) \Rightarrow f_{y_1} = f_{y_2}$

$$\Rightarrow \quad f_{y_1}(x) = f_{y_2}(x) \forall x \in H. \qquad ...(1)$$

$$f_{y_1}(x) = (x, y_1) \text{ and } f_{y_2}(x) = (x, y_2)$$

$\therefore$ from (1), we get

$$(x, y_1) = (x, y_2) \Rightarrow (x, y_1) - (x, y_2) = 0$$

$$\Rightarrow \quad (x, y_1 - y_2) = 0 \forall x \in H \qquad ...(2)$$

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

$$\|\phi(y_1) - \phi(y_2)\| = \|f_{y_1} - f_{y_2}\|$$

$$= \|f_{y_1} - f_{(-y_2)}\|$$

But $\|f_{y_1} + f_{-y_2}\| = \|f_{y_1 - y_2}\| = \|y_1 - y_2\|$ (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

$$(f_x, f_y) = (y,x) \qquad \ldots (1)$$

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 $(\alpha f_x + \beta f_y, f_z) = (f_{\bar{\alpha} x} + f_{\bar{\beta} y}, f_z)$ ... (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 $\left(z, \bar{\alpha}x + \bar{\beta}y\right) = \bar{\bar{\alpha}}(z,x) + \bar{\bar{\beta}}(z,y)$

$$= \alpha\left(f_x, f_z\right) + \beta\left(f_y, f_z\right) \quad \ldots (3)$$

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

$T_2\left(f_x\right) = F_{f_x}, F_{f_x}(f) = \left(f, f_x\right)$ 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 $(T_2.T_1)x = T_2(T_1(x)) = T_2(f_x) = F_{f_x}$.

Notes

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)

   $\therefore (F_x, F_y) = (F_{f_y}, F_{f_y}) = (f_y, f_x) = (x, y)$ 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.

26.2 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 $\|f\| = \|y\|$.

26.3 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:

Notes

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 $X \times X \to C$ 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) $(x, x) \geq 0$, $(x, x) = 0 \Leftrightarrow x = 0$

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.

26.4 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 $(F_x, F_y) = (x, y)$.

26.5 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

Notes

Unit 27: The Adjoint of an Operator

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.

27.1 Adjoint of an Operator

Let T be an operator on Hilbert space H. Then there exists a unique operator T* on H such that

$$(Tx,y)= (x,T^*y) \text{ for all } x, y \in H$$

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

$$(Tx,y)= (x,T^*y) \text{ for all } x, y \in H \qquad ...(1)$$

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

$$(Tx,y)= (x,T^*y) \text{ for all } x, y \in H. \qquad ...(2)$$

Notes

Let y be a vector in H and f_y its corresponding functional in H*.

Let us define

$$T^*: H \xrightarrow{\text{into}} H^* \text{ by}$$

$$T^*: f_y = f_z \qquad ...(3)$$

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:

$$y \to f_y \to T^* f_y = f_z \to z, \qquad ...(4)$$

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,

$$\left(T^* f_y\right)x = f_y(Tx) \qquad ...(5)$$

By Riesz representation theorem,

$y \to f_y$ so that

$$f_y(Tx) = (Tx, y) \qquad ...(6)$$

Since T* is defined on H*, we get

$$\left(T^* f_y\right)x = f_z(x) = (x, z) \qquad ...(7)$$

But we have from our definition T*y = z ...(8)

From (5) and (6) it follows that

$$\left(T^* f_y\right)x = (Tx, y) \qquad ...(9)$$

From (7) and (8) it follows that

$$\left(T^* f_y\right)x = (x, T^* y) \qquad ...(10)$$

From (9) and (10), we thus obtain

$$(Tx, y) = (x, T^* y) \forall x, y \in H.$$

This completes the proof of the theorem.

Notes

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

$$(T^*x,y)=(x,Ty)=\sum_{n=1}^{\infty}x_{n+1}\overline{y}_n=(Sx,y),$$

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 $(Tx,\alpha y_1+\beta y_2)=\overline{\alpha}(Tx,y_1)+\overline{\beta}(Tx,y_2)$

$$=\overline{\alpha}(Tx,y_1)+\overline{\beta}(x,T^*y_2)$$

$$=(x,\alpha T^*y_1)+(x,\beta T^*y_2).$$

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

$$= (x, \alpha T^* y_1 + \beta T^* y_2).$$

$\Rightarrow$ 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

$$(Tx, y) = (x, T' y) \quad ...(2)$$

and $$(Tx, y) = (x, T^* y) \quad ...(3)$$

From (2) and (3) it follows that

$$(x, T' y) = (x, T^* y) \forall x, y \in H$$

$$\Rightarrow \quad (x, (T' y - T^* y)) = 0$$

Notes

$\Rightarrow \quad (x,(T'-T^*)y)=0 \forall x \in H$

$\Rightarrow \quad (T'-T^*)y = 0$ 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 nth place.

Then using (3) we obtain

$$(T_{e_n},e_1)=1.\sum\frac{e_n(j)}{j}=1.\frac{1}{n}.$$

Notes

Now we check whether there is T* which is adjoint of T. Now $(e_n, T^*(e_1)) = T^*(e_1).e_n$, where the R.H.S. gives the component wise inner product. Since $T^*(e_1) \in M, T^*(e_1).e_n$ cannot be equal to $\frac{1}{n} \forall\ n = 1, 2, ...$

$\Rightarrow$ there is no T* on M such that

$$(T(e_n), e_1) = (e_n, T^*(e_1))$$

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) $(T_1 + T_2)^* = T_1^* + T_2^*$ (preserve addition)

(ii) $(T_1T_2)^* = T_2^* T_1^*$ (reverses the product)

(iii) $(\alpha T)^* = \bar{\alpha} T^*$ (conjugate linear)

(iv) $\|T^*\| \geq \|T\|$

(v) $\|T^* T\| = \|T\|^2$

Proof: (i) For every x, y $\in$ H, we have

$$\begin{aligned}
(x, (T_1 + T_2)^*y &= ((T_1 + T_2)\, x, y) && \text{(By def. of adjoint)}\\
&= (T_1x + T_2x, y)\\
&= (T_1x, y) + (T_2x, y)\\
&= (x, T_1^*y) + (x, T_2^*y)\\
&= (x, T_1^*y + T_2^*y)\\
&= (x, (T_1^* + T_2^*)y)\\
&\Rightarrow (T_1 + T_2)^* = T_1^* + T_2^* && \text{by uniqueness of adjoint operator}
\end{aligned}$$

(ii) For every $x, y \in H$, we have

$$(x(T_1T_2)^* y) = ((T_1T_2)x, y)$$

$$= (T_1(T_2x), y)$$

Notes

$$= (T_2x, T_1 * y)$$

$$= (x, T_2^*, (T_1^* y))$$

$$= (x, (T_2^* T_1^*), y)$$

Therefore from the uniqueness of adjoint operator, we have $(T_1T_2)* = T_2^*T_1^*$.

(iii) For every $x, y \in H$, we have

$$(x, (\alpha T) * y) = ((\alpha T)x, y) = (\alpha(Tx), y)$$

$$= \alpha(Tx, y)$$

$$= \alpha(x, T * y) = (x, \bar{\alpha}(T * y))$$

$$= (x(\bar{\alpha}T *)y).$$

Therefore from the uniqueness of adjoint operator, we have

$$(\alpha T *) = \bar{\alpha}T *.$$

(iv) For every $y \in H$ we have

$$\|T * y\|^2 = (T * y, T * y)$$

$$= (TT * y, y)$$

$$= |(TT * y, y)| \qquad \left[\because \|T * y\|^2 = (TT * y, y) \text{is a real number} \geq 0\right]$$

$$\leq \|TT * y\| \; \|y\| \qquad (\text{By Schwarz inequality})$$

$$\leq \|T\| \; \|T * y\| \; \|y\| \qquad [\because \|Tx\| \leq \|T\| \; \|x\|]$$

Thus $\|T * y\|^2 \leq \|T\| \|T * y\| \|y\| \forall y \in H$

$$\Rightarrow \|T * y\| \leq \|T\| \|y\| \forall y \in H. \qquad ...(1)$$

Now $\|T *\| = \text{Sup}\{\|T * y\| : \|y\| \leq 1\}$

from (1), we see that if $\|y\| \leq 1$ then $\|T * y\| \leq \|T\|$

$$\Rightarrow \|T *\| \leq \|T\| \qquad ...(2)$$

Now applying (2) from the operator T* in place of operator T, we get

$$\|(T *)*\| \leq \|T *\|$$

Notes

$$\Rightarrow \|T^{**}\| \le \|T^*\|$$

$$\Rightarrow \|T\| \le \|T^*\| \quad [\because T^{**} = T] \qquad ...(3)$$

From (2) and (3) it follows that

$$\|T\| \le \|T^*\|.$$

(v) We have $\|T^*T\| \le \|T^*\| \; \|T\|$

$$= \|T\| \; \|T\| \quad [\because \|T^*\| = \|T\|]$$

$$= \|T\|^2 \qquad ...(4)$$

Further for every $x \in H$, we have

$$\|Tx\|^2 = (Tx, Tx)$$

$$= (T^*Tx, x)$$

$$= |((T^*T)x, x)|$$

$$\le \|T^*T\| \; \|x\| \|x\| \qquad \text{(By Schwarz inequality)}$$

$$= \|T^*T\| \; \|x\|^2$$

Then we have

$$\|Tx\|^2 \le \|T^*T\| \; \|x\|^2 \; \forall x \in H \qquad ...(5)$$

Now $\|T\| = \sup\{\|Tx\| : \|x\| \le 1\}$

$$\therefore \|T\|^2 = \left[\sup\{\|Tx\| : \|x\| \le 1\}\right]^2$$

$$= \sup\left\{\|Tx\|^2 : \|x\| \le 1\right\}$$

From (5) we see that

if $\|x\| \le 1$, then $\|Tx\|^2 \le \|T^*T\|$.

Therefore, $\text{Sup}\left\{\|Tx\|^2 : \|x\| \le 1\right\} \le \|T^*T\|$

$$\Rightarrow \|T\|^2 \le \|T^*T\|. \qquad ...(6)$$

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

$$TT^{-1} = T^{-1}T = I.$$

Notes

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 $\beta(H)$ 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 $\beta(H)$. Then $T^* \in \beta(H)$ and we have $\phi(T^*) = (T^*)^* = T^{**} = T$. Hence, the mapping ϕ is onto.

27.2 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.

Notes

- 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) $(\alpha T)^* = \overline{\alpha} T^*$

  (iv) $\|T^*\| = \|T\|$

  (v) $\|T^* T\| = \|T\|^2$

27.3 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) 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.

27.4 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 $(T^*)^{-1} = (T^{-1})^*$.

27.5 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 28: 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.

28.1 Self Adjoint Operator

28.1.1 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

$$(\alpha A_1 + \beta A_2)^* = (\alpha A_1)^* + (\beta A_2)^*$$

$$= \bar{\alpha} A_1^* + \bar{\beta} A_2^*$$

$$= \alpha A_1 + \beta A_2 \quad \left[\because \alpha, \beta \text{ are real numbers}, \therefore \bar{\alpha} = \alpha,\ \bar{\beta} = \beta\right]$$

$\Rightarrow \alpha A_1 + \beta A_2$ 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

$$\|A - A^*\| = \|A - A_n + A_n - A^*\|$$

$$\le \|A - A_n\| + \|A_n - A^*\|$$

$$= \|A - A_n\| + \|A_n - A_n^*\| + \|A_n^* - A^*\|$$

$$\le \|A - A_n\| + \|A_n - A_n^*\| + \|A_n^* - A^*\|$$

$$= \|-(A_n - A)\| + \|A_n - A_n\| + \|(A_n - A)^*\| \qquad [\because A_n \in S \Rightarrow A_n^* = A_n]$$

$$= \|A_n - A\| + \|0\| + \|A_n - A\|$$

Notes

$= 2\|A_n - A\|$ $\quad\quad \left[\because \|0\| = 0, \|T^*\| = \|T\| \text{ and } \|-T\| = \|T\|\right]$

$\to 0$ 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. $\quad$ ($\because$ 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, $(Tx, y) = O \forall \ x, y \in H$

$\Rightarrow \quad (Tx, Tx) = O \forall \ x, y \in H \quad$ (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

$$(Tx, x) = 0 \forall \text{ x in } H \Leftrightarrow T = O.$$

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

$$\left(T(\alpha x + \beta y), \alpha x + \beta y\right) = \left(\alpha Tx + \beta Ty, \alpha x + \beta y\right)$$

$$= \alpha\left(Tx, \alpha x + \beta y\right) + \beta\left(Ty, \alpha x + \beta y\right)$$

$$= \alpha\bar{\alpha}(Tx, x) + \alpha\bar{\beta}(Tx, y) + \beta\bar{\alpha}(Ty, x) + \beta\bar{\beta}(Ty, x)$$

$$= |\alpha|^2 (Tx, x) + \alpha\bar{\beta}(Tx, y) + \beta\bar{\alpha}(Ty, x) + |\beta|^2 (Ty, x)$$

$$\therefore \left(T(\alpha x + \beta y), \alpha x + \beta y\right) - |\alpha|^2 (Tx, x) - |\beta|^2 (Ty, y) = \alpha\bar{\beta}(Tx, y) + \beta\bar{\alpha}(Ty, x) \quad ...(1)$$

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

$$\alpha\bar{\beta}(Tx, y) + \beta\bar{\alpha}(Ty, x) = 0 \quad ...(2)$$

for all scalars α, β and $\forall x, y \in H$.

Putting $\alpha = 1, \beta = 1$ in (2) we get

$$(Tx, y) + (Ty, x) = 0 \quad ...(3)$$

Again putting $\alpha = i, \beta = 1$ in (2) we obtain

$$i(Tx, y) - i(Ty, x) = 0 \quad ...(4)$$

Multiply (3) by (i) and adding to (4) we get

$$2i(Tx, y) = 0 \forall x, y \in H$$

$\Rightarrow \quad (Tx, y) = 0 \forall x, y \in H$

$\Rightarrow \quad (Tx, Tx) = 0 \forall x, y \in H$ (Taking y = Tx)

$\Rightarrow \quad Tx = 0 \forall x, y \in H$

Notes

$\Rightarrow$ 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

$$(Tx, Tx) = (x, T^* x) = (x, Tx) = \left(\overline{Tx, x}\right)$$

$\Rightarrow (Tx, x)$ 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 $\forall x \in H$,

$\therefore$ $(Tx, x) = \left(\overline{Tx, x}\right) = \left(\overline{x, T^* x}\right) = (T^* x, x)$

$\Rightarrow$ $(Tx, x) - (T^* x, x) = 0 \forall x \in H$

$\Rightarrow$ $(Tx - T^* x, x) = 0 \forall x \in H$

$\Rightarrow$ $((T - T^*)x, x) = 0 \forall x \in H$

$\Rightarrow$ T – T* = 0 [$\because$ 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 $x, y \in H$.

$$(Ax, y) = (x, A^* y) = \left(\overline{A^* y, x}\right).$$

since H is real Hilbert space $\left(\overline{A^* y, x}\right) = (A^* y, x)$ so that $(Ax, y) = (Ay, x)$ [$\because A^* = A$]

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 $\le$ 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 $'\le'$ is a partial order relation on S. $'\le'$ 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 $A_1 \le A_2 \Rightarrow (A_1x,x) \le (A_2x,x) \forall x \in H.$

Notes

$\Rightarrow \quad (A_1x,x)+(Ax,x)\le(A_2x,x)+(Ax,x)\forall x\in H.$

$\Rightarrow \quad ((A_1+A_2)x,x)\le((A_1+A)x,x)\forall x\in H.$

$\Rightarrow \quad A_1+A_2\le A_2+A,$ 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.

28.1.2 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 $\quad (T^*Tx,x)=(Tx,T^{**}x)=(Tx,Tx)=\|Tx\|^2\ge 0$

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 $(I+T)x=0\Rightarrow x=0$

Notes

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

$$\|(I+T)x\|^2 = \|x+Tx\|^2$$

$$=(x+Tx,x+Tx)$$

$$=(x,x)+(x,Tx)+(Tx,x)+(Tx,Tx)$$

$$=\|x\|^2+\|Tx\|^2+\left(\overline{Tx,x}\right)+(Tx,x)$$

$$=\|x\|^2+\|Tx\|^2+2(Tx,x) \qquad [\because \text{T is positive} \Rightarrow \text{T is self-adjoint} \Rightarrow (Tx,x)\text{real}]$$

$$\geq \|x\|^2 \qquad [\because \text{T is positive} \Rightarrow (Tx,x)\geq 0]$$

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 $\text{Lim}\{(I+T)x_n\} = (I+T)(\lim\ x_n)$ $\quad [\because I+T \text{ is a continuous mapping}]$

$= (I+T)x \in M(\text{range of } I+T)$

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

$x_0 \perp M \Rightarrow (\{I+T\}x_0, x) = 0$

$\Rightarrow (x_0 + Tx_0, x_0) = 0$

$\Rightarrow (x_0, x_0) + (Tx_0, x_0) = 0$

$\Rightarrow \|x\|^2 + (Tx_0, x_0) = 0$

$\Rightarrow -\|x\|^2 = (Tx_0, x_0)$

$\Rightarrow -\|x\|^2 \geq 0 \qquad [\because T \text{ positive} \Rightarrow (Tx_0, x_0) \geq 0]$

$\Rightarrow \|x\|^2 \leq 0$

$\Rightarrow \|x\| = 0 \qquad [\because \|x\|^2 \geq 0]$

$\Rightarrow x = 0$

$\Rightarrow$ 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.

28.2 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 $(Ax, x) \geq 0 \forall x \in H$.

Notes

28.3 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 $(Ax, x) \geq 0 \forall x \in H.$

Self Adjoint: An operator T on a Hilbert space H is said to be self adjoint if T*=T.

28.4 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 $\sup\left\{|(Tx, x)|/\|x\|^2 : x \neq 0\right\} \leq \|T\|$. Prove that if T is self-adjoint, then equality holds here.

28.5 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

Notes

Unit 29: 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.

29.1 Normal and Unitary Operators

29.1.1 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.

Notes

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.

⇒ T is a normal operator on H.

But T = T* ⇒ 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 ⇒ TT*=T*T

Hence T*T**=T*T=TT*=T**T* so that T*T**=T**T

⇒ 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.

$$\|TT^* - T^*T\| = \|TT^* - T_kT_k^* + T_kT_k^* - T_k^*T_k + T_k^*T - TT^*\| \le \|TT^* - T_kT_k^*\| + \|T_kT_k^* - T_k^*T_k\| + \|T_k^*T_k - TT^*\|$$

$$\Rightarrow \quad \|TT^* - T^*T\| \le \|TT^* - T_kT_k^*\| + \|T_k^*T_k - TT^*\| \qquad ...(1)$$

$[\because T_k$ 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 $(1) \to 0$

$\Rightarrow \quad \|TT^* - T^*T\| = 0$

$\Rightarrow \quad TT^* = T^*T$

⇒ 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

$$\|T_n^* - T^*\| = \|T_n - T^*\| = \|T_n - T\| \to 0.$$

$$\Rightarrow \quad \|T_n^* - T^*\| \to 0 \Rightarrow T_n^* = T^*.$$

Now, $\|TT^* - T^*T\| = \|(TT^* - T_nT^*) + (T_nT_n^* - T^*T)\| \le \|TT^* - T_nT_n^*\| + \|T_nT_n^* - T^*T\|$

$$= \|(TT^* - T_nT^*)\| + \|(T_nT_n^* - T_n^*T_n) + (T_n^*T_n - T^*T)\|$$
$$\le \|TT^* - T_nT_n^*\| + \|T_nT_n^* - T_n^*T_n\| + \|T_n^*T_n - T^*T\|$$

$$= \|(TT^* - T_nT_n^*)\| + \|0\| + \|T_n^*T_n - T^*T\|$$

$$\left[\because T_n \in M \Rightarrow T_n \text{ is a normal operator on H i.e. } T_nT_n^* = T_n^*T_n \text{ and } |0| = 0\right]$$

$$= \|TT^* - T_nT_n^*\| + \|T_n^*T_n - T^*T\| \to 0 \text{ since } T_n \to T \text{ and } T_n^* \to T^*$$

Thus, $\|TT^* - T^*T\| = 0 \Rightarrow TT^* - T^*T = 0$

$\Rightarrow \quad TT^* = T^*T \Rightarrow T$ is normal operator on H.

$\Rightarrow \quad T \in M$ 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. $T \in M$

$\Rightarrow \quad \alpha T \in M$, 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 $(\alpha T)^*(\alpha T) = (\bar{\alpha} T)(\alpha T) = \bar{\alpha}\alpha(T^*T) = (\alpha\bar{\alpha})(TT^*)$

$\Rightarrow \quad (\alpha T)(\alpha T)^* = (\alpha T)^*(\alpha T)$

$\Rightarrow \quad \alpha T$ 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

$$(N_1 + N_2)(N_1 + N_2)^* = (N_1 + N_2)\left(N^*_1 + N^*_2\right)$$

$$\Rightarrow \quad N_1N^*_1 + N_1N_2 + N_2N^*_1 + N_2N^*_2 \qquad ...(4)$$

$$\Rightarrow \quad N_1N^*_1 + N^*_2N_1 + N^*_1N_2 + N^*_2N_2$$

$$= \quad \left(N^*_1 + N^*_2\right)\left(N_1 + N_2\right)$$

$$= \quad \left(N_1 + N_2\right)^*\left(N_1 + N_2\right) \qquad \text{(using (1) and (2))}$$

$$\Rightarrow \quad \left(N_1 + N_2\right)\left(N_1 + N_2\right)^* = \left(N_1 + N_2\right)^*\left(N_1 + N_2\right)$$

$$\Rightarrow \quad N_1 + N_2 \text{ is normal.}$$

Now we show that N_1N_2 is normal i.e.

$$\left(N_1N_2\right)\left(N_1N_2\right)^* = \left(N_1N_2\right)^*\left(N_1N_2\right).$$

$$\text{L.H.S.} = \left(N_1N_2\right)\left(N_1N_2\right)^* = N_1N_2N^*_2N^*_1$$

$$= N_1\left(N_2N^*_2\right)N^*_1$$

$$= N_1\left(N^*_2N_2\right)N^*_1$$

$$= \left(N_1N^*_2\right)\left(N_2N^*_1\right)$$

$$= \left(N^*_2N_1\right)\left(N^*_1N_2\right)$$

$$= N^*_2\left(N_1N^*_1\right)N_2$$

$$= \left(N^*_2N^*_1\right)\left(N_1N_2\right)$$

$$= \left(N_1N_2\right)^*\left(N_1N_2\right)$$

$$\Rightarrow \quad \left(N_1N_2\right)\left(N_1N_2\right)^* = \left(N_1N_2\right)^*\left(N_1N_2\right)$$

$$\Rightarrow \quad N_1N_2 \text{ is normal.}$$

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 $\Leftrightarrow TT^* = T^*T$

$\Leftrightarrow TT^* - T^*T = 0$

Notes

$\Leftrightarrow ((TT^* - T^*T)x, x) = 0 \forall x$

$\Leftrightarrow (TT^*x; x) = (T^*Tx, x) \forall x$

$\Leftrightarrow (T^*x, T^*x) = (Tx, T^{**}x) \forall x$

$\Leftrightarrow \|T^*x\|^2 = \|Tx\|^2 \forall x \qquad [\because T^{**} = T]$

$\Leftrightarrow \|T^*x\| = \|Tx\| \forall x.$

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 $T^*_1 = \left[\frac{1}{2}(T + T^*)\right]^*$

$= \frac{1}{2}(T + T^*)^*$

$= \frac{1}{2}(T^* + T^{**})$

$= \frac{1}{2}(T^* + T) = \frac{1}{2}(T + T^*) = T_1$

Notes

$\Rightarrow \quad T^*_1 = T_1$

$\Rightarrow \quad T_1$ is self-adjoint.

Also $T^*_2 = \left[\frac{1}{2i}(T - T^*)\right]^*$

$$= \left(\frac{\overline{1}}{2i}\right)(T - T^*)^*$$

$$= -\frac{1}{2i}(T^* - T^{**})$$

$$= -\frac{1}{2}(T^* - T) = \frac{1}{2i}(T - T^*) = T_2$$

$\Rightarrow \quad T^*_2 = T_2$

$\Rightarrow \quad T_2$ 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 $T^* = (U_1 + iU_2)^*$

$$= U^*_1 + (iU_2)^*$$

$$= U^*_1 + \bar{i}U^*_2$$

$$= U^*_1 - iU^*_2 = U_1 - iU_2$$

$\therefore T + T^* = (U_1 - iU_2) + (U_1 - iU_2) = 2U,$

$\Rightarrow \quad U_1 = \frac{1}{2}(T + T^*) = T_1$

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.

Notes

Theorem 7: If T is an operator on a Hilbert space H, then T is normal $\Leftrightarrow$ 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 $\quad T^* = (T_1 + iT_2)^* = T_1^* + (iT_2)^*$

$$= T_i^* + \bar{i}\ T_2^*$$

$$= T_i^* - iT_2^*$$

$$= T_1 - iT_2$$

Now $\quad TT^* = (T_1 + iT_2)\,(T_i - iT_2)$

$$= T_1^2 + T_2^2 + i\,(T_2T_1 - T_1T_2) \qquad \ldots (1)$$

and $\quad T^*T = (T_i - iT_2)\,(T_1 - iT_2)$

$$= T_1^2 + T_2^2 + i\,(T_1T_2 - T_2T_1) \qquad \ldots (2)$$

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 $\Rightarrow TT^* = T^*T$

Also $\quad (T - \lambda I)^* = T^* - (\lambda I)^*$

$$= T^* - \bar{\lambda}\, I^*$$

$$= T^* - \bar{\lambda}\, I.$$

Now $\quad (T - \lambda I)\,(T - \lambda I)^* = (T - \lambda I)\,(T^* - \bar{\lambda}\, I)$

$$= TT^* - \bar{\lambda}\, I - \lambda T^* + |\lambda|^2 I \qquad \ldots (1)$$

Also $\quad (T - \lambda I)^*\,(T - \lambda I) = (T^* - \bar{\lambda}\, I)\,(T - \lambda I)$

$$= T^*T - \lambda I^* - \bar{\lambda}\, T + |\lambda|^2 I \qquad \ldots (2)$$

Notes

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.

29.1.2 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) ∀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.

29.1.3 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.

$\Rightarrow \quad \|Tx\| = \|x\| \ \forall x$

$\Rightarrow \quad T^*T = I$ [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.

Notes

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.

29.2 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 $x \in H$.

29.3 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 $x \in H$.

29.4 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.

29.5 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/

Notes

Unit 30: 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.

30.1 Projections

30.1.1 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

$z = x + y, x \in M, y \in N.$

Notes

Thus $\quad Pz = P(x+y)$

$$= Px + Py$$

$$= Px = x \quad (y \in N)$$

$\therefore (Pz,z) = (x,z) \quad [\because Pz = P(x+y) = x, P \text{ being projection on } H]$

$$= (x, x+y)$$

$$= (x,x) + (x,y)$$

$$= \|x\|^2$$

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. $P = P^*$

$\Rightarrow \quad$ 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 $x \in M \Leftrightarrow Px = x$.

Notes

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

$$\|Px + (I-P)x\|^2 = \|Px\|^2 + \|(I-P)x\|^2 \qquad ...(2)$$

Notes

From (1) and (2), we have

$$\|x\|^2 = \|Px\|^2 + \|(I-P)x\|^2$$

$\Rightarrow \quad \|(I-P)x\|^2 = 0 \quad [\because \text{by hypothesis } \|Px\| = \|x\|]$

$\Rightarrow \quad (I-P)x = 0$

$\Rightarrow \quad x - Px = 0$

$\Rightarrow \quad Px = 0$

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

$(Px, x) = (PPx, x)$

$= (Px, P^*x) = (Px, Px) = \|Px\|^2 \geq 0$

$\Rightarrow \quad (Px, x) \geq 0 \forall x \in H.$

$\Rightarrow$ P is a positive operator i.e. $P \geq 0$.

(ii) P is a projection on H$\Rightarrow$ I – P is also a projection on H.

$\Rightarrow \quad I - P \geq 0.$ (by (i))

$\Rightarrow \quad P \leq I$

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

$$\|Px + (I-P)x\|^2 = \|Px\|^2 + \|(I-P)x\|^2$$

$\Rightarrow \quad \|x\|^2 = \|Px\|^2 + \|(I-P)x\|^2 \quad [\because Px + (I-P)x = x]$

Notes

$\Rightarrow \quad \|x\|^2 \geq \|Px\|^2$

$\Rightarrow \quad \|Px\| \geq \|x\|.$

(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 $\Leftrightarrow$ 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.

$\therefore \quad (PQ)^* = PQ$

$\Rightarrow \quad Q^* P^* = PQ$

$\Rightarrow \quad QP = PQ \qquad (\because Q^* = Q, P^* = P)$

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

$(PQ)(x) = P(Qx) = Px \qquad [\because$ N is range of Q and $x \in N \Rightarrow Qx = x]$

$= x \qquad [\because$ M is range of P and $x \in P]$

$\Rightarrow \quad (PQ)x = x$

$\therefore \quad x \in R(PQ)$

$\Rightarrow \quad x \in M \cap N \Rightarrow x \in R(PQ)$

Notes

$\therefore \quad M \cap N \subset R(PQ)$

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

$$PP^* = P^*P^* \qquad \text{[taking } P^* \text{ in place of P in L.H.S.]}$$

$$= P^*P \qquad [\because P^* = P]$$

$\Rightarrow$ P is normal.

Conversely, let $PP^* = P^*P$.

Then to prove that $P^* = P$.

For every vector $y \in H$, we have

$$(Py, Py) = (y, P^*Py) = (y, PP^*y) \qquad [\because P^*P = PP^*]$$

$$= (P^*y, P^*y) \qquad [\because (P^*)^* = P]$$

From this we conclude that

$$Py = 0 \Leftrightarrow P^*y = 0.$$

Now let x be any vector in H.

Let $y = x - Px$. Then

Notes

$$Py = P(x - Px) = Px - P^2x = Px - Px = 0$$

$$0 = P^*y = P^*(x - Px) = P^*x - P^*Px$$

$$\Rightarrow \quad P^*x = P^*Px \ \forall \ x \in H$$

$$\therefore \quad P^* = P^*P$$

Now $\quad P = (P^*)^* = (P^*P)^* = P^*P = P^*$

$\therefore$ P is a self adjoint operator.

Also $P^2 = P$.

Hence P is a projection on H.

30.1.2 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.

$$(Tx,y) = 0$$

$$\Rightarrow \quad (x,Ty) = 0$$

$\Rightarrow$ $\quad T^*y$ 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 $(M^\perp)^\perp$ is invariant under T.

But $(M^\perp)^\perp = M^{\perp\perp} = M$ 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 $PT = PTP$ $\qquad [\because P \text{ is projection} \Leftrightarrow P^* = P. \text{ Also} TT^* = T]$

Notes

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. $M \perp N$.

30.1.3 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

$$\|P_i x\|^2 = \sum_{j=1}^{n} \|P_j x\|^2$$

$\Rightarrow \quad \|P_j x\|^2 = O$ if $j \neq i$

$\Rightarrow \quad \|P_j x\| = O,\ j \neq i$

$\Rightarrow \quad P_j x = O,\ j \neq i$

$\Rightarrow \quad$ x is in the null space of $P_j, i \neq j$

$\Rightarrow \quad x \in M_j^{\perp}$, if $j = i$

$\Rightarrow \quad$ 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 $i = 1, 2, ..., n$

$\Rightarrow \quad x_1 + x_2 + ... + x_n \in R(P)$.

$\Rightarrow \quad x \in R(P)$.

Then $x \in M \Rightarrow x \in R(P)$

$\therefore M \subset R(P)$ ...(3)

Now suppose that $x \in R(P)$. Then

$Px = x$

$\Rightarrow \quad (P_1 + P_2 + ... + P_n)x = x$

$\Rightarrow \quad P_1x + P_2x + ... + P_nx = x$

But $P_1(x) \in M_1, P_2(x) \in M_2, ..., P_n(x) \in M_n$.

$\therefore x \in M_1 + M_2 + ... + M_n$ and so $R(P) \subset M$ ...(4)

Hence from (3) and (4), we get

M = R(P)

This completes the proof of the theorem.

30.2 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.

30.3 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

30.4 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 $N \cap M^{\perp}$.

30.5 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.

Notes

Unit 31: Finite Dimensional Spectral Theory

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.

Notes

31.1 Finite Dimensional Spectral Theory

31.1.1 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 $T \in \beta(H)$ (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 ith row and $(\alpha_{1j}, \alpha_{2j}, \ldots \alpha_{nj})$ as its jth 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.

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 $T \in \beta$ (H) with (T) = $[\alpha_{ij}]$, then T is singular $\Leftrightarrow$ $[\alpha_{ij}]$ 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 $\Leftrightarrow$ $[T^{-1} T] = [TT^{-1}] = [I]$

from (2) $[T^{-1}] [T] = [T] [T^{-1}] = [I] = [\delta_{ij}]$

so that $[T^{-1}] [\delta_{ij}] = [\delta_{ij}] [T^{-1}] = [\delta_{ij}]$, $[T] = [\alpha_{ij}]$.

$\Rightarrow$ $[\alpha_{ij}]$ is a non-singular and $[\alpha_{ij}]^{-1} = [T^{-1}]$.

This completes the proof of the theorem.

31.1.2 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

$$B = C^{-1} AC.$$

Notes

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.

31.1.3 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 ([δ_{ij}]) = 1.

(ii) det (T_1 T_2) = (det T_1) (det T_2)

(iii) det (T) $\neq$ 0 $\Leftrightarrow$ [T] is non-singular

$\Leftrightarrow$ det ([T]) $\neq$ 0.

Hence det (T) $\neq$ 0 $\Leftrightarrow$ [T] is non-singular.

31.1.4 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 $\in$ 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 $\neq$ 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...

Notes

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 $x \in M_\lambda \Leftrightarrow Tx = \lambda x \quad \ldots (1)$

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

$$T(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$$

Let λ be a eigenvalue of T. Then $\exists$ a non zero vector

$$y = (y_1, y_2, \ldots) \text{ in } \ell_2 \text{ such that } Ty = \lambda y.$$

Notes

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.

31.1.5 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

Notes

that $\sigma(T) \neq \phi$. 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 $\sigma(T)$ 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

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \qquad \ldots (1)$$

(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

$$\begin{bmatrix} \alpha_{11} - \lambda & \alpha_{12} \\ \alpha_{21} & \alpha_{22} - \lambda \end{bmatrix} = 0$$

$$\Rightarrow \quad \lambda^2 - (\alpha_{11} + \alpha_{22}) \lambda + (\alpha_{11} \alpha_{22} - \alpha_{12} \alpha_{21}) = 0 \qquad \ldots (2)$$

Since (2) is true for λ, we can take

$$T = \lambda I \qquad \ldots (3)$$

From (2) and (3) we get

$$T^2 - (\alpha_{11} + \alpha_{22}) T + (\alpha_{11} \alpha_{22} - \alpha_{12} \alpha_{21}) I = 0 \qquad \ldots (4)$$

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 $P(\lambda) \in \sigma(P(T))$.

Notes

Proof:

(a) T is singular $\Leftrightarrow \exists$ a non-zero vector $x \in H$ such that $Tx = 0$ or $Tx = 0$. Hence T is singular $\Leftrightarrow$ 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, $\exists$ a non-zero vector $x \in H$ s.t. $Tx = \lambda x$.

Premultiplying by T^{-1} we get

$T^{-1} Tx = T^{-1} (\lambda x)$

$\Rightarrow T^{-1}(x) = \frac{1}{\lambda} x$ 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 $S - \lambda I$.

Now $S - \lambda I = ATA^{-1} - A(\lambda I) A^{-1}$

$= A(T - \lambda I) A^{-1}$

$\therefore \quad \det(S - \lambda I) = \det(A(T - \lambda I) A^{-1})$

$= \det(T - \lambda I)$

$\Rightarrow \quad \lambda$ is an eigenvalue of T $\Leftrightarrow \det(T - \lambda I) = 0$.

Hence $\det(T - \lambda I) = 0 \Leftrightarrow \det(S - \lambda I) = 0$

$\Rightarrow$ 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 $\exists$ 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.

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.

31.1.6 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.

Notes

(a) The M_i's are pairwise orthogonal and span H:

(b) The P_i's are pairwise orthogonal and $P_1 + P_2 + \ldots + P_m = I$ and $T = \lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_m P_m$.

(c) T is normal operator on H.

Proof: We shall show that

(i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i)

(i) $\Rightarrow$ (ii)

Assume that M_i's are pairwise orthogonal and span H. Hence every $x \in H$ can be represented uniquely as

$$x = x_1 + x_2 + \ldots + x_m \qquad \ldots (1)$$

where $x_i \in M_i$ for i = 1, 2, …, m

by hypothesis M_i's are pairwise orthogonal. Since P_i's are projections in M_i's $\Rightarrow$ 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,

$$\begin{aligned} P_i(x) &= P_i(x_1 + x_2 + \ldots + x_m) \\ &= P_i x_1 + P_i x_2 + \ldots + P_i x_m \qquad \ldots (2) \end{aligned}$$

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)

$\Rightarrow$ $x_j \in M_i^{\perp} \Rightarrow P_i x_j = x_i$

$\therefore$ from (2) we get

$$P_i x = x_i \qquad \ldots (3)$$

Since I is the identity mapping on H, we get

$$\begin{aligned} Ix &= x_1 + x_2 + \ldots + x_m && \ldots \text{(by (1))} \\ &= P_1 x + P_2 x + \ldots + P_m x && \ldots \text{(by (3))} \\ &= (P_1 + P_2 + \ldots + P_m)\, x \ \forall x \in H. \end{aligned}$$

This show that $I = P_1 + P_2 + \ldots\ldots + P_m$.

For every $x \in H$, we have from (1)

$$\begin{aligned} T(x) &= T(x_1 + x_2 + \ldots + x_m) \\ &= Tx_1 + Tx_2 + \ldots + Tx_m \end{aligned}$$

Since $x_i \in M_i \Rightarrow Tx_i = \lambda x_i$

$$\therefore \quad T_x = \lambda_1 x_1 + \lambda_2 x_2 + \ldots\ldots + \lambda_m x_m \qquad \ldots (4)$$

$$= \lambda_1 P_1 x + \lambda_2 P_2 x + \ldots\ldots + \lambda_m P_m x \qquad \ldots (5)$$

$$\Rightarrow \quad T = \lambda_1 P_1 + \lambda_2 P_2 + \ldots\ldots + \lambda_m P_m$$

(ii) $\Rightarrow$ (iii)

Notes

Let $T = \lambda_1 P_1 + \lambda_2 P_2 + \ldots\ldots + \lambda_m P_m$, 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 $P_i P_j = 0$ for $i \neq j$

Since adjoint operation is conjugate linear, we get

$$T^* = (\lambda_1 P_1 + \lambda_2 P_2 + \ldots\ldots + \lambda_m P_m)^*$$

$$= \bar{\lambda}_1 P^*_1 + \bar{\lambda}_2 P^*_2 + \ldots\ldots + \bar{\lambda}_m P^*_m$$

$$= \bar{\lambda}_1 P_1 + \bar{\lambda}_2 P_2 + \ldots\ldots + \bar{\lambda}_m P_m .$$

Now
$$TT^* = (\lambda_1 P_1 + \lambda_2 P_2 + \ldots\ldots + \lambda_m P_m)(\bar{\lambda}_1 P_1 + \bar{\lambda}_2 P_2 + \ldots\ldots + \bar{\lambda}_m P_m)$$

$$= |\lambda_1|^2 P_1^2 + |\lambda_2|^2 P_2^2 + \ldots\ldots + |\lambda_1|^2 P_m^2 \qquad (\because P_i P_j = 0, i \neq j)$$

$$= |\lambda_1|^2 P_1 + |\lambda_2|^2 P_2 + \ldots\ldots + |\lambda_m|^2 P_m \qquad \ldots \text{(by (6))}$$

Similarly T*T can be found s.t.

$$T^*T = |\lambda_1|^2 P_1 + |\lambda_2|^2 P_2 + \ldots\ldots + |\lambda m|^2 P_m$$

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 $\Rightarrow P_i T = TP_i$ for each P_i.

$\therefore$
$$TP = T(P_1 + P_2 + \ldots\ldots + P_m)$$

$$= TP_1 + TP_2 + \ldots\ldots + TP_m$$

$$= P_1 T + P_2 T + \ldots\ldots + P_m T$$

$$= (P_1 + P_2 + \ldots\ldots + P_m) T$$

$$= PT$$

$\therefore$ 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.

Notes

$\therefore \quad M^{\perp} = \{0\}$.

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.

31.1.7 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

$$\begin{aligned} Tx &= (\lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_m p_m)x \\ &= (\lambda_1 p_1 p_i + \lambda_2 p_2 p_i + \ldots + \lambda_i p_i^2 + \ldots + \lambda_m p_m p_i)x \end{aligned}$$

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 $Tx = \lambda x = \lambda Ix$.

$\Rightarrow \quad (\lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_m p_m)\, x = \lambda\, (p_1 + p_2 + \ldots + p_m)\, x$

$\Rightarrow \quad \{(\lambda_1 - \lambda)\, p_1 + (\lambda_2 - \lambda)\, p_2 + \ldots + (\lambda_m - \lambda)\, p_m\}\, x = 0 \qquad \ldots (2)$

Since $p_i^2 = p_i$ and $p_i\, p_j = 0$ for $i \neq j$ operating with p_i throughout (2), we get

$(\lambda_i - \lambda)\, p_i x = 0$ for $i = 1, 2, \ldots, m$.

Notes

If $\lambda_i \neq \lambda$ for each i, $p_i x = 0$ for each i.

Hence $p_i x + p_2 x + \ldots + p_m x = 0$

$\Rightarrow \quad (p_1 + p_2 + \ldots + p_m)\, x = 0$

$\Rightarrow \quad Ix = 0$

$\Rightarrow \quad x = 0$, 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 $T = \mu_1 Q_1 + \mu_2 Q_2 + \ldots + \mu_n Q_n$ … (3)

be another revolution of T. Then $\mu_1, \mu_2, \ldots \mu_n$ 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

$$T = \lambda_1 Q_1 + \lambda_2 Q_2 + \ldots + \lambda_m Q_m \qquad \ldots (4)$$

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

$$T^0 = I = p_1 + p_2 + \ldots + p_m$$

$$T^1 = \lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_m p_m$$

$$T^2 = \lambda_1^2 p_1 + \ldots + \lambda_m^2 p_m$$

and $$T^n = \lambda_1^n p_1 + \ldots + \lambda_m^n p_m \text{ for any positive integer n.} \qquad \ldots (5)$$

Now if g (t) is a polynomial with complex coefficient in the complex variable t, we can write g (T) as

$$g(T) = g(\lambda_1)\, p_1 + g(\lambda_2)\, p_2 + \ldots + g(\lambda_m)\, p_m$$

$$= \sum_{j=1}^{m} g(\lambda_j)\, p_j \qquad \ldots \text{(by 5)}$$

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 $$\pi_i(t) = \frac{(t-\lambda_1)\ldots(t-\lambda_{i-1})\,(t-\lambda_{i+1})\ldots(t-\lambda_m)}{(\lambda_i-\lambda_1)\ldots(\lambda_i-\lambda_{i-1})\,(\lambda_i-\lambda_{i+1})\ldots(\lambda_i-\lambda_m)}$$

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.

$\therefore \quad p_i = Q_i$ for each i.

This completes the proof of the theorem.

Notes

31.1.8 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.

$T : N \to N'$ is compact of every bounded set $B \subset N$, $\overline{T(B)}$ is compact in N′.

31.1.9 Properties of Compact Operators

1. Let $T : N \to N'$ 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 $B \subset N$. 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 $I : N \to N$ is not compact operator.

   For consider a closed unit sphere.

   $S = \{x \in N : \| x \| \le 1\}$ then S is bounded.

   Since N is a infinite dimensional.

   $I(S) = S = \overline{S}$ is not necessarily compact.

   Hence $I : N \to N$ 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.

$$\| x_n - y_n \| \le \frac{1}{n} \qquad \ldots (1)$$

Notes

and $$x_{n_k} \to x \in \overline{A} \qquad \ldots (2)$$

we can find a (y_{n_k}) of (y_n) s.t.

$$\begin{aligned} \left\| y_{n_k} - x \right\| &= \left\| y_{n_k} - x_{n_k} + x_{n_k} - x \right\| \\ &\le \left\| y_{n_k} - x_{n_k} \right\| + \left\| x_{n_k} - x \right\| \\ &\to 0 \text{ as } n \to \infty. \\ &\Rightarrow \overline{A} \text{ is compact.} \end{aligned}$$

This completes the proof of the theorem.

31.2 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.

31.3 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.

Notes

31.4 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.

31.5 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

LOVELY PROFESSIONAL UNIVERSITY
Jalandhar-Delhi G.T. Road (NH-1)
Phagwara, Punjab (India)-144411
For Enquiry: +91-1824-521360
Fax.: +91-1824-506111
Email: odl@lpu.co.in

978-93-90164-64-6
9 789390 164646