M.Sc. Final Year

Mathematics, MM-06

INTEGRATION THEORY AND FUNCTIONAL ANALYSIS

मध्यप्रदेश भोज (मुक्त) विश्वविद्यालय – भोपाल

MADHYA PRADESH BHOJ (OPEN) UNIVERSITY - BHOPAL

COURSE WRITERS

Ekta Punia, Former Assistant Professor (Mathematics), Daulat Ram College, University of Delhi, Delhi
Units (1, 2, 3.0-3.1, 3.3-3.10, 4, 5.0-5.1, 5.3.2- 5.10)

V. K. Khanna, Retd. Professor, Department of Mathematics, Kirori Mal College, University of Delhi, Delhi
S. K. Bhambri, Retd. Professor, Department of Mathematics, Kirori Mal College, University of Delhi, Delhi
Units (3.2, 5.2, 5.3, 5.3.1)

Published by Registrar, MP Bhoj (Open) University, Bhopal in 2020

VIKAS®

Vikas® is the registered trademark of Vikas® Publishing House Pvt. Ltd.

VIKAS® PUBLISHING HOUSE PVT. LTD.
E-28, Sector-8, Noida - 201301 (UP)
Phone: 0120-4078900 • Fax: 0120-4078999
Regd. Office: A-27, 2nd Floor, Mohan Co-operative Industrial Estate, New Delhi 1100 44
• Website: www.vikaspublishing.com • Email: helpline@vikaspublishing.com

SYLLABI-BOOK MAPPING TABLE

Integration Theory and Functional Analysis

Syllabi	Mapping in Book
Integration Theory	
I. Signed measure. Hahn decomposition theorem, mutually singular measures. Rodon-Nikodym theorem. Lebesgue decomposition. Riesz representation theorem. Extension theorem (Caratheodory), Lebesgue-Stieltijes integral, product measures, Fubini's theorem.	**Unit-1:** Signed Measures **(Pages 3-43)**
II. Baire sets. Baire measure, continuous functions with compact support. Regularity of measures on locally compact spaces. Integration of continuous functions with compact support, Riesz-Markoff theorem.	**Unit-2:** Baire Sets and Baire Measure **(Pages 45-65)**
Functional Analysis	
III. Normed Linear Spaces, Banach Spaces with examples, Quotient space of normed linear space and its completeness, bounded linear transformations, normed linear space of bounded linear transformations, dual (conjugate) spaces with examples, natural imbedding of a normed linear space in its second dual, open mapping theorem, closed graph theorem, uniform boundedness principle and its consequences.	**Unit-3:** Normed Linear Spaces **(Pages 67-141)**
IV. Finite dimensional normed spaces and subspaces, Equivalent norms, finite dimensional normed linear spaces and compactness, Riesz lemma, Hahn Banach theorem for real llnear space, complex linear space, and normed linear space, Adjoint operators, Reflexive spaces, Weak convergence, weak* Convergence.	**Unit-4:** Finite Dimensional Normed Spaces and Subspaces **(Pages 143-184)**
V. Inner product space, Hilben space, Orthogonal complements, Orthonormal sets, Bessel's inequality, complete orthonormal sets and Parseval's identity, coniugate space H* and reflexivity of Hilbert space, hljoin1 of on operator on a Hilbert space, self - adjoint operators, positive, projection, normal and unitary operators.	**Unit-5:** Inner Product Space and Hilbert Space **(Pages 185-220)**

CONTENTS

INTRODUCTION

At the close of the 19th century Stieltjes showed that the integral is a linear functional and hence was able to generalize Riemann's theory to introduce Riemann-Stieltjes integrals. However, there were many shortcomings in Riemann's theory and it was felt that an integration theory which could encompass a larger class of functions was needed. This paved the way for sthe concept of measure, which was introduced by Borel and Lebesgue at the beginning of 20th century.

Functional analysis is a branch of mathematical analysis, the core of which is formed by the study of vector spaces endowed with some kind of limit-related structure such as inner product, norm, topology, etc., and the linear operators acting upon these spaces and respecting these structures in a suitable sense. It is concerned with infinite-dimensional vector spaces and mappings between them. The historical roots of functional analysis lie in the study of spaces of functions and the formulation of properties of transformations of functions such as the Fourier transform as transformations defining certain operators between function spaces. This turned out to be particularly useful for the study of differential and integral equations. The usage of the word functional goes back to the calculus of variations, implying a function whose argument is a function.

In modern introductory texts to functional analysis, the subject is seen as the study of vector spaces endowed with a topology, in particular infinite dimensional spaces. A significant part of functional analysis is the extension of the theory of measure, integration and probability to infinite dimensional spaces, also known as infinite dimensional analysis.

This book, *Integration Theory and Functional Analysis* is divided into five units that follow the self-instruction mode with each unit beginning with an Introduction to the unit, followed by an outline of the Objectives. The detailed content is then presented in a simple but structured manner interspersed with Check Your Progress Questions to test the student's understanding of the topic. A Summary along with a list of Key Terms and a set of Self-Assessment Questions and Exercises is also provided at the end of each unit for recapitulation.

UNIT 1 SIGNED MEASURES

NOTES

Structure

1.0 INTRODUCTION

In mathematics, signed measures is a generalization of the concept of measures that permits it to have negative values. There are two somewhat different concepts of a signed measures which depends on the fact that one should permit or not the infinite values. There are two types of measures, extended signed measures and finite signed measures. For a specified or given measurable space (X, Σ), that is for a set X with an σ algebra or sigma algebra Σ on it, an extended signed measures is considered a function. A finite signed measures can be defined except that it only takes the real values, i.e., it cannot take $+\infty$ or $-\infty$. Finite signed measures form a vector space. The sum of two finite signed measures is a finite signed measures because it is the product of a finite signed measures by a real number which is considered closed under linear combination. It follows the assumption that the set of finite signed measures on a measures space (X, Σ) is a real vector space. The Hahn decomposition theorem is named after the Austrian Mathematician Hans Hahn. The theorem states that given a measurable space (X, Σ) and a signed measures μ defined on the σ algebra Σ then there exist two sets P and N in Σ and the pair (P, N) is termed as a Hahn decomposition of the signed measures μ. In measures theory, Lebesgue's decomposition theorem states that for given μ and ν two σ-finite signed measures on a measurable space (Ω, Σ).

In this unit, you will learn about the basics of signed measures, Hahn decomposition theorem, mutually singular measures, Radon-Nikodym theorem, extension theorem (Caratheodory), Lebesgue-Stieltjes integral, product measures and Fubini's theorem.

1.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Explain the significance of signed measures
- State Hahn and Jordan decomposition theorems
- Discuss the meaning of mutually singular measures is
- Explain Radon-Nikodym theorem
- Elaborate on the Caratheodory extension theorem
- Describe Lebesgue decomposition
- State the significance of Lebesgue-Stieltjes integral
- Discuss the product measures and Fubini's theorem

1.2 BASICS OF SIGNED MEASURES

In mathematics, **signed measures** is referred as a simplification of the concept of measure that allows it to have negative values. There are two different notions of a signed measures that depends on the condition that how the infinite values are taken. In advanced concept typically the signed measures take finite values, while sometimes generally the infinite values are taken. Hence the former concept is termed as 'Finite Signed Measures' while the later is termed as 'Extended Signed Measures'.

For a specified (given) measurable space (X, Σ), that is for a set X with a σ algebra or sigma algebra Σ on it, an **extended signed measures** is a function as follows:

$$\mu : \Sigma \to \mathbb{R} \cup \{\infty, -\infty\}$$

This implies that $\mu(\tilde{A}) = 0$ where μ is sigma additive. It satisfies the following equality for any sequence $A_1, A_2, ..., A_n$ of disjoint sets in Σ.

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)\Bbbk$$

One possibility is that any extended signed measures can take the value as $+\infty$ or it can take the value as $-\infty$. But both the values are not available.

Similarly, a **finite signed measures** can be defined except that it only takes the real values, i.e., it cannot take $+\infty$ or $-\infty$. Finite signed measures form a vector space.

The measures can be extended signed measures but may not be the general finite signed measures. For example, consider a nonnegative measure ν on the space (X, Σ) and a measurable function $f: X \to \mathbf{R}$ such that,

$$\int_X |f(x)|\, d\nu(x) < \infty.$$

Subsequently, a finite signed measures is given by,

$$\mu(A) = \int_A f(x)\, d\nu(x)$$

This is for all A in $\sum$.

This signed measures will take only finite values. To permit it to take $+\infty$ as a value, substitute the assumption regarding f being absolutely integrable with the more relaxed condition,

$$\int_X f^-(x)dv(x) < \infty,$$

Here $f^-(x) = \max(-f(x), 0)$ is the negative part of f.

Properties

The two results follow which implies that an extended signed measures is the difference of two nonnegative measures and a finite signed measures is the difference of two finite nonnegative measures.

The Hahn decomposition theorem states that for a given signed measures μ, there exist two measurable sets P and N such that,

1. $P \cup N = X$ and $P \cap N = \varnothing$.
2. $\mu(E) \geq 0$ for each E in Σ such that $E \subseteq P$ or in other words P is a positive set.
3. $\mu(E) \leq 0$ for each E in Σ such that $E \subseteq N$ that is N is a negative set.

Furthermore, this decomposition is unique for adding/subtracting μ null sets from P and N.

Now consider the two nonnegative measures μ^+ and μ^- defined by,

$$\mu^+(E) = \mu(P \cap E)$$

And

$$\mu^-(E) = -\mu(N \cap E)$$

This is for all measurable sets E, that is E in $\sum$.

Here both μ^+ and μ^- are nonnegative measures. The measures take only finite values and are termed as the *positive part* and *negative part* of μ, respectively. Thus we have $\mu = \mu^+ - \mu^-$. The measure $|\mu| = \mu^+ + \mu^-$ is termed as the *variation* of μ and its maximum possible value specified as $||\mu|| = |\mu|(X)$ is termed as the *total variation* of μ.

This possibility of the Hahn decomposition theorem is termed as the Jordan decomposition. The measures μ^+, μ^- and $|\mu|$ are independent of the option of P and N in the Hahn decomposition theorem.

Space of Signed Measures

The sum of two finite signed measures is a finite signed measures because it is the product of a finite signed measure by a real number which is considered closed under linear combination. It follows the assumption that the set of finite signed measures on a measure space $(X, \sum)$ is a real vector space. This is quite opposite to the positive measures which are only closed under conical combination and thus form a convex cone but not a vector space. In addition, the total variation defines a norm for which the space of finite signed measures becomes a Banach space. As per the Riesz representation theorem, if X is a compact separable space

NOTES

then the space of finite signed Baire measures is considered the dual of the real Banach space of all continuous real valued functions on X.

Let (X, S) be a measurable space. A function $v : S \to [-\infty, \infty]$ is considered as a signed measures if the following conditions are true:

- If $v(\tilde{A}) = 0$.
- If $\{-\infty, \infty\} \cap$ (range v) is a singleton set or empty.
- If $\{E_i\}_{i=1}^{\infty}$ are a pairwise disjoint collection of measurable sets then we have,

$$v(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} v(E_i)$$

Here the sum converges absolutely if $\sum_1^{\infty} v(E_i) < \infty$.

For example, let $g \in L^1(X, S, \mu)$ where μ is a measure and defines,

$$v(E) = \int_E g d\mu = \int_E g^+ d\mu - \int_E g^- d\mu$$

Definition: A set $E \in S$ is a *positive* set for the signed measures ν if $v(F) \geq 0$ for all $F \in S$ with $F \subset E$. Similarly, a *negative* set can also be defined.

Lemma 1: Suppose ν is a signed measures and E is a positive set. If $F \subset E$ is measurable then F is considered as a positive set. Additionally, if in a countable family $E_i \in S$ are all positive then $\bigcup_1^{\infty} E_i$.

Proof: The first assertion follows immediately from the definition.

Assume that $\{E_i\}_1^{\infty}$ are positive sets then we can write $\bigcup_1^{\infty} E_i$ as a disjoint union of the form $\bigcup_1^{\infty} F_i$ and consequently $F_i \subset E_i$.

In addition, if $B \subset \bigcup_1^{\infty} E_i = \bigcup_1^{\infty} F_i$ then,

$$B = \bigcup_1^{\infty} (B \cap F_i)$$

Subsequently,

$$v(B) = \sum_{i=1}^{\infty} v(B \cap F_i) \geq 0$$

Since for all i, $B \cap F_i \subset E_i$.

Lemma 2: Consider that (X, S) be a measurable space and ν a signed measures and assume a subset $E \in S$ and $0 < v(E) < \infty$. Then there exists a positive set $P \subset E$ such that, $v(P) > 0$.

Proof: For the condition, if E is positive then take $E = P$. Consequently assume that there exists a subset $N \subset E$ such that, $v(N) < 0$. Let $\mathcal{N}(E) = \{N \in S : N \subset E, v(N) < 0\}$. Since E is not positive hence $\mathcal{N}$ is nonempty.

Let $n \in \mathbb{N}$ be the smallest natural number such that there exists $N_1 \in \mathcal{N}(E)$ with $v(N_1) < -\dfrac{1}{n_1}$. If $E \setminus N_1$ is positive then it follows the definition and is proved. If it is negative then continue inductively.

Let n_2 be the smallest positive integer such that there exists $N_2 \in \mathcal{N}(E)$ such that, $v(N_2) < -\frac{1}{n_2}$. Reproduce to obtain $N_1, \ldots, N_k$ such that $v(N_j) < -\frac{1}{n_j}$, $1 \le j \le k$ and n_{k+1} is selected as the smallest positive integer and there exists $N_{k+1} \in \mathcal{N}(E \setminus \cup_1^k N_j)$ such that, $v(N_{k+1}) < -\frac{1}{n_{k+1}}$.

Observe that $n_{k+1} \ge n_k$ and $\{N_k\}$ is a disjoint family of sets.

Let,

$$P = E \setminus \left(\bigcup_{k=1}^{\infty} N_k \right)$$

We have,

$$E = P \bigcup \left(\bigcup_{k=1}^{\infty} N_k \right)$$

Therefore,

$$v(E) = v(P) + \sum_{1}^{\infty} v(N_k)$$

Because v(E) < ∞ we obtain that the sum on the right converges absolutely.

Subsequently,

$$\sum_{k=1}^{\infty} \frac{1}{n_k} < \sum_{k=1}^{\infty} | v(N_k) | < \infty$$

Specifically, $n_k \to \infty$ since $k \to \infty$. We have,

$$v(\bigcup_{k=1}^{\infty} N_k) < 0 \text{ and } v(E) > 0$$

As a result we have, $v(P) > 0$

Let $\varepsilon > 0$ and consider that K is very large such that we have,

$$\frac{1}{n_{k+1} - 1} < \varepsilon, \ \forall k \ge K$$

We already know that,

$$P \subset E \setminus \left(\bigcup_{k=1}^{K} N_k \right),$$

Subsequently by construction, P contains no measurable set F with $v(F) < \frac{-1}{n_{k+1} - 1}$.

Hence, P contains no subset $F \in S$ with $v(F) < -\varepsilon$. This is true for all $\varepsilon > 0$ and hence P must be positive. Any measure is thus a signed measures and sometimes termed as a positive measure.

NOTES

Signed Measures and Complex Measures

Now we will explain a generalization of the notion of a measure where the values are permitted to be outside $[0, \infty]$.

Definition: Suppose $\mathcal{A}$ is a σ algebra on a non-empty set X. A function $\mu : \mathcal{A} \to [-\infty, \infty]$ is termed as a signed measures on $\mathcal{A}$, if it has the properties explained below.

1. Either one of the following is true:
   - $\mu(A) < \infty, \forall A \in \mathcal{A}$
   - $\mu(A) > -\infty, \forall A \in \mathcal{A}$
2. If $\mu(\varnothing) = 0$.
3. For any pairwise disjoint sequence $(A_n)_{n=1}^{\infty} \subset \mathcal{A}$ there is the equality,

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n) \qquad ...(1.1)$$

The convention that if one term on the right hand side of Equation (1.1) is equal to $\pm\infty$ then the entire sum is equal to $\pm\infty$. It is significant to use Condition (1) because it avoids conditions when one term is ∞ and another term is $-\infty$.

A complex measures simplifies the concept of measure using the complex values or we can say that these are sets whose size (length, area and volume) is a complex number.

A complex measure μ on a measurable space (X, Σ) is a function defined on Σ which takes complex values that is sigma additive. We can write,

$$\mu : \Sigma \to \mathbb{C}$$

This specifies that for any sequence $(A_n)_n$ of disjoint sets in Σ there is,

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

This is possible provided that the sum on the right hand side converges completely or diverges accurately in analogy with the real valued signed measures.

Integration with Respect to a Complex Measures

The *integral* of a complex valued measurable function can be defined with respect to a complex measures by approximating a measurable function with simple functions. The already existing integral of a real valued function can be used with respect to a nonnegative measure. The real and imaginary parts μ_1 and μ_2 of a complex measures μ are considered finite valued signed measures. Using the Hahn-Jordan decomposition theorem these measures can be split as follows:

$$\mu_1 = \mu_1^+ - \mu_1^- \text{ and } \mu_2 = \mu_2^+ - \mu_2^-$$

Here $\mu_1^+, \mu_1^-, \mu_2^+, \mu_2^-$ are the unique finite valued nonnegative measures. Subsequently, for a measurable function f which is real valued for the moment, we can define:

$$\int_X f\, d\mu = \left(\int_X f\, d\mu_1^+ - \int_X f\, d\mu_1^-\right) + i\left(\int_X f\, d\mu_2^+ - \int_X f\, d\mu_2^-\right)$$

This expression holds provided that the expression on the right hand side is defined such that all four integrals exist. At the time of addition of these integrals the indeterminate $\infty - \infty$ is not encountered.

For a given complex valued measurable function, its real and imaginary components can be integrated independently as already discussed and we can denote this as follows:

$$\int_X f\,d\mu = \int_X \Re(f)\,d\mu + i\int_X \Im(f)\,d\mu.$$

Variation of a Complex Measure and Polar Decomposition

For a complex measure μ its variation or absolute value $|\mu|$ can be defined using the formula,

$$|\mu|(A) = \sup \sum_{n=1}^{\infty} |\mu(A_n)|$$

Here A is in Σ and the supremum flows over all sequences of disjoint sets $(A_n)_n$ whose union is A. Considering only the finite partitions of the set A into measurable subsets, we can obtain an equivalent definition.

This implies that $|\mu|$ is a nonnegative finite measure. Correspondingly since a complex number can be represented in a polar form we can have the polar decomposition for a complex measures. There exists a measurable function θ with real values such that,

$$d\mu = e^{i\theta} d|\mu|$$

This implies that,

$$\int_X f\ d\mu = \int_X f\, e^{i\theta} d|\mu|$$

This holds for any absolutely integrable measurable function f, i.e., f satisfying the condition,

$$\int_X |f|\,d|\mu| < \infty.$$

The Radon-Nikodym theorem can be used to prove that the variation is a measure and the polar decomposition exists.

The Space of Complex Measures

The sum of two complex measures is also referred as a complex measures and similarly the product of a complex measures by a complex number. To be precise, the set of all complex measures on a measure space (X, Σ) forms a vector space. Additionally, the total variation $\|\mu\|$ is defined as follows,

$$\|\mu\| = |\mu|(X)$$

This is considered as the norm with respect to which the space of complex measures is termed as a Banach space.

1.2.1 Hahn Decomposition Theorem

The Hahn decomposition theorem is named after the Austrian mathematician Hans Hahn. The theorem states that given a measurable space (X, Σ) and a signed

NOTES

measures μ defined on the σ algebra Σ then there exist two sets P and N in Σ such that,

1. $P \cup N = X$ and $P \cap N = \varnothing$.
2. For each E in Σ such that, $E \subseteq P$ one has $\mu(E) \geq 0$, i.e., P is a positive set for μ.
3. For each E in Σ such that, $E \subseteq N$ one has $\mu(E) \leq 0$, i.e., N is a negative set for μ.

In addition, basically this decomposition is unique because for any additional pair (P', N') of measurable sets which fulfils the above three conditions, the symmetric differences $P \Delta P'$ and $N \Delta N'$ are μ null sets based on the logic that every measurable subset of them has zero measure. The pair (P, N) is termed as a *Hahn decomposition* of the signed measures μ.

Hahn-Jordan Decomposition

A consequence of this theorem is the Jordan decomposition theorem, which states that every signed measure μ can be expressed as a difference of two positive measures μ^+ and μ^-, of which at least one is finite. Here μ^+ and μ^- are the positive and negative part of μ, respectively. These two measures can be characterized as follows:

$$\mu^+(E) := \mu(E \cap P) \text{ and } \mu^-(E) := -\mu(E \cap N)$$

This holds for every E in Σ and both μ^+ and μ^- can be verified as positive measures on the space (X, Σ) where at least one of them is finite, since μ cannot take both $+\infty$ and $-\infty$ as values and satisfy $\mu = \mu^+ - \mu^-$. The pair (μ^+, μ^-) is termed as *Jordan decomposition* and also sometimes Hahn-Jordan decomposition of μ.

Proof of the Hahn Decomposition Theorem

Preparation: Consider that μ does not take the value $-\infty$ or else decompose according to $-\mu$. As already explained, a negative set is a set A in Σ such that $\mu(B)\ d \leq 0$ for every B in Σ which is a subset of A.

Claim: Assume that a set D in Σ satisfies $\mu(D) \leq 0$. Then there is a negative set $A \subseteq D$ such that $\mu(A) \leq \mu(D)$.

Proof of the Claim: Define $A_0 = D$. Further presume for a natural number n that $A_n \subseteq D$ has been constructed. Let,

$$t_n = \sup\{\mu(B) : B \in \Sigma, B \subset A_n\}$$

This denotes the supremum of $\mu(B)$ for all the measurable subsets B of A_n. This supremum may be infinite. Since the empty set $\varnothing$ is a feasible B in the definition of t_n and $\mu(\varnothing) = 0$ hence we obtain $t_n \geq 0$. By definition of t_n there exists $B_n \subseteq A_n$ in Σ which satisfies that,

$$\mu(B_n) \geq \min\{1, t_n / 2\}.$$

Set $A_{n+1} = A_n \backslash B_n$ concludes the induction step. Define,

$$A = D \setminus \bigcup_{n=0}^{\infty} B_n .$$

NOTES

Since the sets $(B_n)_{n\geq 0}$ are disjoint subsets of D hence it follows from the sigma (σ) additivity of the signed measures μ that,

$$\mu(A) = \mu(D) - \sum_{n=0}^{\infty} \mu(B_n) \leq \mu(D) - \sum_{n=0}^{\infty} \min\{1, t_n / 2\}.$$

This shows that $\mu(A) \leq \mu(D)$. Presume that A is not a negative set, i.e., there exists a B in Σ which is a subset of A and satisfies $\mu(B) > 0$. Then $t_n \geq \mu(B)$ for every n hence the series on the right must diverge to $+\infty$ which signifies that $\mu(A) = -\infty$, which is not permitted. Consequently, A must be a negative set.

Construction of the Decomposition

Consider that set $N_0 = \varnothing$. Initiating that N_n is given then define s_n such that,

$$s_n := \inf\{\mu(D) : D \in \Sigma, D \subset X \setminus N_n\}.$$

This is the infimum of $\mu(D)$ for all the measurable subsets D of $X \setminus N_n$. This infimum may be $-\infty$. Since the empty set is a feasible D and $\mu(\varnothing) = 0$ hence we have $s_n \leq 0$. Therefore there exists a D_n in Σ with $D_n \subseteq X \setminus N_n$ and $\mu(D_n) \leq \max\{S_{n/2} - 1\} \leq 0$.

As per the above claim we can define that there is a negative set $A_n \subseteq D_n$ such that, $\mu(A_n) \leq \mu(D_n)$. We can define $N_{n+1} = N_n \cup A_n$ to conclude the initiation step.

Define,

$$N = \bigcup_{n=0}^{\infty} A_n .$$

Since the sets $(A_n)_{n\geq 0}$ are disjoint hence we have for every $B \subseteq N$ in Σ that,

$$\mu(B) = \sum_{n=0}^{\infty} \mu(B \cap A_n)$$

This is true by the sigma or σ additivity of μ. Specifically, this proves that N is a negative set. Define $P = X \setminus N$. If P is not a positive set then there exists $D \subseteq P$ in Σ with $\mu(D) < 0$. Subsequently $s_n \leq \mu(D)$ for all n and,

$$\mu(N) = \sum_{n=0}^{\infty} \mu(A_n) \leq \sum_{n=0}^{\infty} \max\{s_n / 2, -1\} = -\infty,$$

This is not possible for μ consequently P is a positive set.

Proof of the Uniqueness Statement: Consider that (N', P') is an additional Hahn decomposition of X. Subsequently $P \cap N'$ is considered as a positive set and also a negative set. As a result, every measurable subset of it has measure zero. The similar applies to $N \cap P^1$. Because

$$(P\Delta P')\bigcup(N\Delta N') = (P\bigcap N')\bigcup(N\bigcap P').$$

Hence the proof is complete.

NOTES

Signed Measures: Hahn and Jordan Decomposition

The theoretical measure is termed as a nonnegative extended real valued function defined on a measurable space (X, $\mathcal{B}$). On the contrary, a signed measures can also use negative values. By definition, a signed measures ν must satisfy the following properties:

1. The ν assumes at most one of the values $-\infty$ and ∞.
2. $v(\text{Ã}) = 0$.
3. $v\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} v(E_i)$ for any sequence $\langle E_i \rangle$ of the disjoint measurable sets. Consequently, if $v\left(\bigcup_{i=1}^{\infty} E_i\right)$ is finite then the series converges absolutely.

Property (1) is sufficient to avoid $\infty - \infty$ whereas Property (2) defines that there is at least one positive set. Property (3) is essential for the existence of a positive set having positive measures that are included in a measurable set with finite positive measure.

We can state that a set A is positive with respect to a signed measure ν if A is measurable and all measurable subsets of A have nonnegative measure. In the same way, B is termed as negative if it is measurable and every measurable subset of it has non positive measure. The set C is termed as a null set if it is both positive and negative.

Evidently, every measurable subset of a positive set is termed positive and a union of a countable collection of disjoint positive sets is also positive, by Property (3). For a countable collection of positive sets $\langle P_i \rangle$ which may not be necessarily disjoint, we can define that their union is also positive. If E be an arbitrary measurable set of $\bigcup P_i$ then we can define,

$$E_i = E \bigcup P_i \setminus \bigcup_{n=1}^{i-1} P_n.$$

Here $\langle E_i \rangle$ is a sequence of disjoint measurable sets whose union is E and each E_i is contained in positive P_i. Consequently, $v(E) = v\left(\bigcup E_i\right) = \sum v(E_i) \geq 0$ as per Property (3).

Eventually it can be proved that given a signed measures space (X, B, ν) there is a positive set A and a negative set B such that A and B partition X. This is termed as the Hahn decomposition theorem. The following is an important lemma.

Lemma: Let E be a measurable set having finite positive measure. Then there is a positive set P having positive measure contained in E.

Proof: If E is positive then the theorem holds. Assume that E is not positive then it has a measurable subset of negative measure. Let $n_1 \in \mathbb{Z}_+$ be the smallest such that there is a measurable set E_1 with $v(E_1) < -\frac{1}{n}$. If we consider n as small then there is no E_1 fulfilling the form. Next we take $E \backslash E_1$.

NOTES

Take $v(E \backslash E_1) = v(E) - v(E_1) > v(E) > 0$. If E/E_1 is positive then we exit the proof.

Initiating, if $E \setminus \bigcup_{i=1}^{k-1} E_i$ is not positive then let $n_k \in \mathbb{Z}_+$ be the smallest such that there is a measurable set E_k with $v(E_k) < -\frac{1}{n_k}$. Continuing further, let $P = E \setminus \bigcup_{i=1}^{\infty} E_i$. Subsequently, $E = P \cup (\bigcup E_i)$. Here P and $\langle E_i \rangle$ are disjoint. Consequently by Property (3), $v(E) = v(P) + \sum v(E_i)$. But by definition $v(E) < \infty$ hence $\sum v(E_i)$ is absolutely convergent. Therefore, $\lim_{i \to \infty} n_i = 0$. Now consider that $\in > 0$, accordingly there is k such that, $\frac{1}{n_k - 1} < \in$.

Consequently, $-\in < -\frac{1}{n_k - 1}$.

Here P is contained in $E \setminus \bigcup_{i=1}^{k-1} E_i$. If A contains a measurable set which is henceforth contained in $E \setminus \bigcup_{i=1}^{k-1} E_i$ with measure less than $-\in$ and consequently less than $-\frac{1}{n_k - 1}$ then n_k is considered no longer the smallest positive integer that makes the existence of a measurable set having measure less than the negative of its reciprocal.

As a result, P contains no measurable set having measure less than $-\in$. Since $\in$ is arbitrary, hence P contains no measurable set having negative measure. This shows that P is positive.

It is obvious that $v(P) = v(E) - \Sigma\, v(E_i) > v(E) > 0$.

Hence proved.

Theorem 1.1 (Hahn Decomposition Theorem): Let $(x, \mathcal{B}, \nu)$ be a signed measures space. Then there is a positive set P and a negative set N such that, $X = P \cup N$ and $P \cap N = \tilde{A}$.

Proof: Assume that ν never takes ∞. Let $p = \sup_P$ is positive $\nu(P)$. Since $\tilde{A}$ is positive, consequently $p \geq 0$. By the definition of sup, there exists is a sequence of positive sets $\langle P_i \rangle$ such that, $p = \lim_{i \to \infty} v(P_i)$. Let $P = \bigcup P_i$ then P is positive. Therefore, $p \geq v(P)$. But for any i, $v(A) = v(A_i) + v(A \setminus A_i) \geq v(A_i)$.

Hence, $v(A) \geq p$. As a consequence, $0 \leq v(A) = p < \infty$.

Let, $N = X \setminus P$.

NOTES

Here if we can assert that N is negative then no further proof required. Presume that E is a positive subset of N. Subsequently E and P are disjoint and $E \bigcup P$ is positive.

Consequently, $p \geq v(E \bigcup P) = v(E) + v(p) = v(E) + p$. This implies that $v(E) = 0$. Therefore, N does not contain any positive subsets of positive measure. As a result, by the contrapositive of the lemma, N does not contain any subsets of positive measure.

Thus the theorem proves that the Hahn decomposition of a measurable space associated with a signed measures exists. Though, it is not exceptional. Consider that m be the Lebesgue measure on $\mathbb{R}$.

Now define $v_1(E) = m(E \bigcap [-1, 0])$ and $v_2(E) = m(E \bigcap [0, 1])$.

Let $v = v_1 - v_2$. Now consider that $(-\infty, 0)$ and $(0, \infty)$ is a Hahn decomposition since for any measurable $E \subset (-\infty, 0]$, $v(E) = v_1(E \bigcap [-1, 0])$ $-v_2(E \bigcap [0, 1) = v_1(E \bigcap [-1, 0]) \geq 0$ and any $F \subset (0, \infty), v(F) = v_1(\text{Ã}) - v_2$ $(F \bigcap [0, 1]) \leq 0$.

Similarly, it can be verified that $[-1, 0]$ and $(-\infty, -1) \bigcup (0, \infty)$ is an additional Hahn decomposition. Hahn decomposition is considered unique except for null sets.

If ν is a signed measures and P and N is Hahn decomposition then we can define v^+ through $v^+(E) = v(E \bigcap P)$ and v^- through $v^-(E) = -v(E \bigcap N)$.

Consequently, $v = v^+ - v^-$. Here v^+ and v^- are measures that are mutually singular because there is a binary partition $\{A, B\}$ of X such that $v^+(A) = v^-(B) = 0$ when $A = N$ and $B = P$.

A decomposition of the measure ν as a difference of two mutually singular measures v^+ and v^- is termed as **Jordon decomposition**. The Jordan decomposition is in fact independent of the Hahn decomposition and there is only one pair of such decomposition and is unique in nature.

Let $v = v^+ - v^-$ be a Jordan decomposition. By definition, there exists a partition A and B of X such that $v^+(A) = v^-(B) = 0$. Here (B, A) is a Hahn decomposition.

Now consider that $E \subset B$, then $v^-(E) \leq v^-(B) = 0$, i.e., $v^-(E) = 0$.

Consequently, $v(E) = v^+(E) - v^-(E) \geq 0$. This implies that B is positive. Similarly A is negative. Subsequently we can deduce that Hahn decomposition is up to null sets where as the Jordan decomposition is unique.

Theorem 1.2 (Jordan Decomposition Theorem): Let v be a signed measures on (X, S). Then there exists two measures v^+ and v^- on (X, S) such that, $v = v^+ + v^-$ and $v^+ \perp v^-$.

NOTES

Furthermore, if μ^+ and μ^- on (X, S) such that, $v = \mu^+ - \mu^-$ and $\mu^+ \perp \mu^-$.

Then $\mu^+ = v^+$ and $v^- = \mu^-$.

Proof: Assume that μ^+ and μ^- satisfy the last two conditions. Then, there exists measurable sets A, B such that, $X = A \bigcup B$ and $A \bigcap B = \tilde{A}$. In addition, $\mu^+(B) = \mu^-(A) = 0$.

If P, N be the corresponding sets for v^+ and v^- then,

- A is positive for v.
- P is positive for v.
- B is negative for v.
- N is negative for v.

Observe that, $P \setminus A \subset P$ and $P \setminus A \subset B$

Subsequently, $v(P \setminus A) = \mu^+(P \setminus A) - \mu^-(P \setminus A) = 0 - \mu^-(P \setminus A) \leq 0$

Consequently, we obtain that $\mu^-(P \setminus A) = 0$ and $v(P \setminus A) = 0$. In the same way we deduce that $v(A \setminus P) = 0$.

For $E \in S$ we establish that,

$$\begin{aligned} \mu^+(E) &= \mu^+(E \bigcap A) = v(E \bigcap A) = v(E \bigcap (P \bigcap A)) + v(E \bigcap (P \setminus A)) \\ &= v((E \bigcap P \bigcap A) = v(E \bigcap A) \bigcap P) = v^+(E \bigcap A) = v^+(E) \end{aligned}$$

Similarly, we obtain $\mu^-(E) = v^-(E)$ for all $E \subset S$. The following definitions will prove the assumptions.

Definition 1: v^+ and v^- are termed as the *positive* and *negative variation* of v. $v^+ + v^-$ is termed as the total variation of v, $|v|$. This is precisely defined in the theorem.

Note: $E \subset S$ is null for a signed measures v if $v(F) = 0$ for all $F \subset E$, $F \in S$. This is true if and only if $|v|(E) = 0$.

For signed measures v_1, v_2 we can state that $v_1 \perp v_2$ if and only if $|v_1| \perp |v_2|$.

Definition 2: For a signed measures and a measurable function f, we state that f is integrable with respect to v if f is integrable with respect to v^+ and f is also integrable with respect to v^-. Furthermore, we can define that,

$$\int_X f\,dv = \int_X f\,dv^+ - \int_X f\,dv^-$$

The right hand side of the above equation homogeneously obtain the integration over X and replace it with integration over P and N, respectively.

Note: If v is a signed measures and μ is a (honest) measure on a measurable space (X, S) then we can state that v is absolutely continuous with respect to μ if for $E \in S$ and $\mu(E) = 0$ we have $v(E)$.

NOTES

We can define that $v << \mu$ when v is absolutely continuous.

Note: $v << \mu \Leftrightarrow |v| << \mu$.

Absolutely continuous and mutually singular are at times the opposite terms. If v and μ are measures such that both are $\mu \perp v$ and $v << \mu$ then $v \equiv 0$. The similar condition holds for v a signed measures.

Proof: To state that $v \perp \mu$ means that $X = A \bigcup B$ where $A, B \in S$, $A \bigcap B = \tilde{A}$, and $v(B) = \mu(A) = 0$. Then because $v << \mu$, we obtain $v(A) = v(B) = 0$ and that $v(X) = v(A \bigcup B) = v(A) + v(B) = 0$. For signed measures, just replace v and μ in the statement by the total variance $|v|$ and $|\mu|$.

Proposition: Assume that v is a finite signed measures, i.e., $|v|(X) < \infty$ and μ is a positive measure, both on the measurable space (X, S). Then, $v << \mu$ if and only if for all $\in > 0$, there exists a $\delta > 0$ such that $|v(E)| < \in$ whenever $\mu(E) < \delta$.

Proof: We include $v << \mu \Leftrightarrow |v| << \mu$ consequently without loss of generality assume that v is a measure, i.e., $v \geq 0$. The condition $\in - \delta$ holds, if $\mu(E) = 0$ then $\mu(E) < \delta$ for all $\delta > 0$. Specifically, we obtain $v(E) < \in$ for all $\in > 0$. Subsequently, $v(E) = 0$. Accordingly, $v << \mu$.

Conversely, assume that the $\in - \delta$ condition fails. Therefore there exists $\in > 0$ with no δ. Fix such an $\in > 0$. Hence, $\forall \delta > 0$, there exists a $E_\delta \in S$ with $\mu(E_\delta) < \delta$ and $v(E_\delta) \geq \in$. Let $E_n \in S$ satisfy $\mu(E_n) < \dfrac{1}{2^n}$ and $v(E_n) \geq \in$. Take $F_k = \bigcup_{n=k}^{\infty} E_n$. Then, $\mu(F_k) \leq \dfrac{1}{2^{k-1}}$ and $v(F_k) \geq \in$. Using the continuity theorem, we obtain,

$$\mu(\bigcap_{k=1}^{\infty} F_k) = 0,$$

However,

$$v(\bigcap_{k=1}^{\infty} F_k) = \lim_{k \to \infty} v(F_k) \geq \in$$

Hence, v is not absolutely continuous with respect to μ.

Corollary: If $f \in L^1(X, \mu)$ where μ is a measure, then for all $\in > 0$ there exists a $\delta > 0$ such that, $|\int_E f d\mu| < \in$ whenever $\mu(E) < \delta$.

Proof: Consider that $v(E) = \int_E f d\mu$ then $v << \mu$ and apply the proposition.

1.2.2 Mutually Singular Measures

Two complex measures μ and ν on a measure space X are considered **mutually singular** if they are provided on different subsets. More specifically, $X = A \bigcup B$

where A and B are two disjoint sets such that the following properties hold for any measurable set E,

1. The sets $A \cap E$ and $B \cap E$ are measurable.
2. The total variation of μ is supported on A and that of ν on B, i.e.,

   $\|\mu\|(B \cap E) = 0 = \|v\|(A \cap E)$.

The relation of two measures being singular is defined as $\mu \perp v$ which is evidently symmetric. However, it is sometimes stated that ν is singular with respect to μ.

A discrete singular measures with respect to Lebesgue measures on the real integrals is a measure λ defined at 0. We can state that $\lambda(E) = 1$ iff $0 \in E$. Generally, a measure λ is concentrated on a subset A if $\lambda(E) = \lambda\ (E \cap A)$. In this case the measure is concentrated at 0.

Two positive or signed or complex measures μ and ν defined on a measurable space (Ω, Σ) are called singular if there exist two disjoint sets A and B in Σ whose union is Ω such that μ is zero on all measurable subsets of B while ν is zero on all measurable subsets of A. This is denoted by $\mu \perp v$

A polished form of Lebesgue's decomposition theorem decomposes a singular measure into a singular continuous measure and a discrete measure.

As a special case, a measure defined on the Euclidean space $\mathbf{R}^n$ is called singular if it is singular in respect to the Lebesgue measure on this space. For example, the Dirac delta function is a singular measure.

Consider the following discrete measure function on the real line,

$$H(x) \overset{\text{def}}{=\!=} \begin{cases} 0, & x < 0; \\ 1, & x \geq 0; \end{cases}$$

This has the Dirac delta distribution δ_0 as its distributional derivative. This is a measure on the real line and a point mass at 0. Though, the Dirac measure δ_0 is neither absolutely continuous with respect to Lebesgue measure λ nor is λ absolutely continuous with respect to δ_0: $\lambda(\{0\}) = 0$. But $\delta_0(\{0\}) = 1$, if U is any open set not containing 0 then $\lambda(U) > 0$ but $\delta_0(U) = 0$.

Check Your Progress

1. What are signed measures?
2. What is finite signed measures?
3. State Hahn decomposition theorem.
4. How can the space of signed measures be defined?
5. Define the integral of a complex valued measurable function.
6. What is Jordon decomposition?
7. When do complex measures become mutually singular?
8. When are the two positive or signed or complex measures called singular?

1.3 RADON -NIKODYM THEOREM

NOTES

The theorem is named after Johann Radon, who proved the theorem for the special case where the underlying space is $\mathbf{R}^N$ and Otto–Nikodym who proved the general case. The Radon-Nikodym theorem is a consequence in measure theory that states that given a measurable space (X, Σ), if a σ-finite measure ν on (X, Σ) is absolutely continuous with respect to a σ-finite measure μ on (X, Σ) then there is a measurable function f on X which takes values in $[0, \infty)$ such that,

$$v(A) = \int_A f \, d\mu$$

This holds for any measurable set A.

Radon-Nikodym Derivative

The function f satisfies the above stated equality is uniquely defined up to a μ-null set. If g is an additional function which satisfies the same property then $f = g$, μ-almost everywhere (μ-αε). f is generally described as $dv/d\mu$ and is termed as the Radon-Nikodym derivative. The option of notation and the name of the function reflect the fact that the function is analogous to a derivative in calculus and describes the rate of change of density of one measure with respect to another. A similar theorem can be proved for signed and complex measures if μ is a nonnegative σ-finite measure and v is a finite valued signed or complex measures such that $|v| << \mu$ then there is μ-integrable real or complex valued function g on X such that,

$$v(A) = \int_A g \, d\mu,$$

This holds for any measurable set A.

Properties

If Y is a Banach space and the generalization of the Radon-Nikodym theorem also holds for functions with values in Y then Y is said to contain the Radon-Nikodym property. All Hilbert spaces have the Radon-Nikodym property. The following properties hold for Radon-Nikodym:

- Consider that ν, μ and λ are the σ-finite measures on the same measure space. If ν << λ and μ << λ, i.e., ν and μ are absolutely continuous with respect to λ then we can state that,

  $$\frac{d(v+\mu)}{d\lambda} = \frac{dv}{d\lambda} + \frac{d\mu}{d\lambda} \ \lambda\text{-almost everywhere.}$$

- If ν << μ << λ then we can state that,

  $$\frac{dv}{d\lambda} = \frac{dv}{d\mu}\frac{d\mu}{d\lambda} \ \lambda\text{-almost everywhere.}$$

- Particularly, if μ << ν and ν << μ then we can state that,

  $$\frac{d\mu}{dv} = \left(\frac{dv}{d\mu}\right)^{-1} \ v\text{-almost everywhere.}$$

- If $\mu << \lambda$ and g is a μ-integrable function then we can state that,

$$\int_X g\, d\mu = \int_X g \frac{d\mu}{d\lambda} d\lambda.$$

- If ν is a finite signed or complex measures then we can state that,

$$\frac{d\,|\nu|}{d\mu} = \left|\frac{d\nu}{d\mu}\right|.$$

NOTES

Divergences

If μ and ν are measures over X and $\nu << \mu$ then,

- The Kullback-Leibler divergence from μ to ν is stated as follows:

$$D_{KL}(\mu \,||\, \nu) = \int_X \log\left(\frac{d\mu}{d\nu}\right) d\mu$$

- For $\alpha > 0$, $\alpha \neq 1$ the Rényi divergence of order α from μ to ν is stated as follows:

$$D_\alpha(\mu \,||\, \nu) = \frac{1}{\alpha - 1} \log\left(\int_X \log\left(\frac{d\mu}{d\nu}\right)^{\alpha-1} d\mu\right).$$

The Assumption of σ-Finiteness

The Radon-Nikodym theorem holds the assumption that the measure μ with respect to which one computes the rate of change of ν is sigma finite. When μ is not sigma finite then the Radon-Nikodym theorem fails to hold.

Consider the Borel σ-algebra on the real line. Let the counting measure μ of a Borel set A be defined as the number of elements of A, if A is finite and $+\infty$ otherwise. It can be checked that μ is certainly a measure. It is not sigma finite, because not every Borel set is atmost a countable union of finite sets. If ν be the usual Lebesgue measures on this Borel algebra then ν is absolutely continuous with respect to μ, since for a set A we can state $\mu(A) = 0$ only if A is the empty set and then $\nu(A)$ is also zero.

Assume that the Radon-Nikodym theorem holds for some measurable function f then we can state that,

$$\nu(A) = \int_A f\, d\mu$$

This holds for all Borel sets. Taking A to be a singleton set, $A = \{a\}$ and using the above equality we obtain,

$$0 = f(a)$$

This holds for all real numbers a. This implies that the function f and consequently the Lebesgue measure ν is zero, which is a contradiction.

Proof

For finite measures μ and ν, consider functions f by $f\,d\mu\, d\leq d\nu$. The supremum of all such functions together with the monotone convergence theorem provides the

Radon-Nikodym derivative. The truth that the remaining part of μ is singular with respect to ν follows from a procedural fact regarding finite measures. After establishing the result for finite measures, extending to σ-finite, signed and complex measures can be obtained logically.

For Finite Measures

Assume that μ and ν are both finite valued nonnegative measures. Let F be the set of those measurable functions $f : X \rightarrow [0, +\infty]$ which satisfy the given notation as follows:

$$\int_A f\, d\mu \leq v(A)$$

This is for every $A \in \Sigma$. This set is not empty contains at least the zero function. Now consider that $f_1, f_2 \in F$ where let A be an arbitrary measurable set, $A_1 = \{x \in A \mid f_1(x) > f_2(x)\}$ and $A_2 = \{x \in A \mid f_2(x)\ \text{e}\geq f_1(x)\}$. Then we obtain the following expression,

$$\int_A \max\{f_1, f_2\}\, d\mu = \int_{A_1} f_1\, d\mu + \int_{A_2} f_2 d\mu \leq v(A_1) + v(A_2) = v(A),$$

Consequently, max $\{f_1, f_2\} \in F$.

Now, consider that $\{f_n\}_n$ be a sequence of functions in F such that,

$$\lim_{n\rightarrow\infty} \int_X f_n\, d\mu = \sup_{f \in F} \int_X f\, d\mu.$$

By substituting f_n with the maximum of the first n functions, it can be assumed that the sequence $\{f_n\}$ is increasing. Let g be a function defined as,

$$g(x) := \lim_{n\rightarrow\infty} f_n(x).$$

By Lebesgue's monotone convergence theorem we obtain,

$$\int_A g\, d\mu = \lim_{n\rightarrow\infty} \int_A f_n\, d\mu \leq v(A)$$

This is for each $A \in \Sigma$ and consequently $g \in F$. Moreover, by the construction of g we obtain,

$$\int_X g\, d\mu = \sup_{f\in F} \int_X f\, d\mu.$$

Now, because $g \in F$ we have,

$$v_0(A) := v(A) - \int_A g\, d\mu$$

This defines a nonnegative measure on Σ. Assume that $\nu_0 \neq 0$, then because μ is finite there is a $\varepsilon > 0$ such that, $\nu_0(X) > \varepsilon\, \mu(X)$. Let (P, N) be a Hahn decomposition for the signed measures $\nu_0 - \varepsilon\, \mu$ then for every $A \in \Sigma$ we have $\nu_0(A \cap P) \geq \varepsilon\, \mu(A \cap P)$ and consequently we obtain the following expression:

$$v(A) = \int_A g\, d\mu + v_0 A \geq \int_A g\, d\mu + v_0(A \cap P)$$

$$\geq \int_A g\,d\mu + \varepsilon\mu(A \cap P) = \int_A (g + \varepsilon 1_P) d\mu.$$

In addition, $\mu(P) > 0$ if $\mu(P) = 0$. Then because ν is absolutely continuous in relation to μ, we have $\nu_0(P) \leq \nu(P) = 0$, subsequently $\nu_0(P) = 0$ and we obtain the expression,

NOTES

$$\nu_0(X) - \varepsilon\mu(X) = (\nu_0 - \varepsilon\mu)(N) \leq 0.$$

Contradicting the fact that $\nu_0(X) > \varepsilon\,\mu(X)$.

Subsequently, because

$$\int_X g d\mu \leq \nu(X) < +\infty\,.$$

$g + \varepsilon\,1_P \in F$ and satisfies the expression of the form,

$$\int_X (g + \varepsilon 1_P) d\mu > \int_X g d\mu = \sup_{f \in F} \int_X f d\mu.$$

This is not possible. As a result the initial assumption that $\nu_0 \neq 0$ must be false. Accordingly $\nu_0 = 0$ as required.

Now, because g is μ-integrable, the set $\{x \in X \,|\, g(x) = +\infty\}$ is considered as μ-null. Consequently, if a f is defined as follows,

$$f(x) = \begin{cases} g(x) & \text{if } g(x) < \infty \\ 0 & \text{otherwise,} \end{cases}$$

Then f contains the required properties.

For the uniqueness, consider that $f, g : X \to [0, +\infty)$ be measurable functions which satisfy the expression,

$$\nu(A) = \int_A f\ d\mu = \int_A g\ d\mu$$

This is for every measurable set A. Then, $g - f$ is μ-integrable and we obtain,

$$\int_A (g - f)\ d\mu = 0.$$

Specifically it holds for $A = \{x \in X \,|\, f(x) > g(x)\}$ or $\{x \in X \,|\, f(x) < g(x)\}$. It follows that,

$$\int_X (g - f)^+\ d\mu = 0 = \int_X (g - f)^-\ d\mu.$$

Subsequently we can state that $(g - f)^+ = 0$ μ-almost everywhere. The same is true for $(g - f)^-$ and hence $f = g$ μ-almost everywhere, as required.

For σ-Finite Positive Measures

If μ and ν are σ-finite, then X can be defined as the union of a sequence $\{B_n\}_n$ of disjoint sets in Σ, each of which has finite measure for both μ and ν. For each n, there is a Σ-measurable function $f_n : B_n \to [0, +\infty)$ such that,

$$\nu(A) = \int_A f_n\ d\mu$$

This is for each $\sum$-measurable subset A of B_n. The union f of those functions is then termed as the required function. For the uniqueness, because each of the f_n is μ-almost everywhere unique then consequently is f.

NOTES

For Signed and Complex Measures

If ν is a σ-finite signed measures, then it can be termed as Hahn-Jordan decomposed because $\nu = \nu^+ - \nu^-$ where one of the measures is finite. Applying the earlier obtained result to those two measures, we can obtain two functions, $g, h : X \to [0, +\infty)$ which satisfy the Radon-Nikodym theorem for ν^+ and ν^-, respectively, where at least one is μ-integrable, i.e., its integral with respect to μ is finite. Then it is obvious that $f = g - h$ satisfies the required properties including uniqueness because both g and h are unique up to μ-almost everywhere equality.

If ν is a complex measure, then it can be decomposed as $\nu = \nu_1 + i\nu_2$ where both ν_1 and ν_2 are considered as the finite valued signed measures. As a result, we obtain two functions of the form $g, h : X \to [0, +\infty)$ which satisfy the required properties for ν_1 and ν_2, respectively. Evidently, $f = g + i\,h$ is the required function.

Theorem 1.3: Johann Radon-Otton Nikodym Marcin: Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space and let v be a measure defined on $\mathcal{B}$ such that $v << \mu$. Then there is a unique nonnegative measurable function f up to sets of μ-measure zero such that

$$v(E) = \int_E f\, d\mu,$$

for every $E \in \mathcal{B}$. f is called the Radon-Nikodym derivative of v with respect to μ and it is often denoted by $\left[\frac{dv}{d\mu}\right]$.

Proof: Consider the following examples to prove the theorem:

Let $(\mathbb{R}, \mathcal{M}, v)$ be the Lebesgue measures space. Let μ be the counting measure on $\mathcal{M}$. So μ is not σ-finite. For any $E \in \mathcal{M}$, if $\mu(E) = 0$, then $E = \emptyset$ and hence $v(E) = 0$. This defines $v << \mu$. Suppose that f is Radon-Nikodym derivation, then for each $x \in \mathbb{R}$, $0 = v(\{x\}) = \int_{\{x\}} f d\mu = \int_{\{x\}} f \chi(x) d\mu = f(x)\mu(\{x\})$. Hence, $f = 0$. This means for every $E \in \mathcal{M}$, $v\{E\} = \int_E 0 d\mu = 0$, which contradicts that v is the Lebesgue measures.

If v is still the same measure as explained above but we let – to be defined by $\mu(\emptyset) = 0$ and $\mu\{A\} = \infty$ if $A \neq \emptyset$. Clearly, μ, is not σ-finite and $v << \mu$. The Radon-Nikodym derivative does not exist. Suppose f is one. Then for any $0 = v(\{x\}) = \int_{\{x\}} f\, d\mu = f(x)\mu(\{x\}) = f(x)\infty$. Thus, $f = 0$. Hence for any $E \in \mathcal{M}$, $v(E) = \inf_E 0 d\mu = 0$.

Obviously if $\mu(X) = 0$, then the measure of every set in M. with respect to μ and v is zero.

Assume that $(X, \mathcal{B}, \mu)$ is a finite measure space and defines the existence and uniqueness of the Radon-Nikodym derivative. Then the case when μ is σ-finite follows by the pasting of the derivatives on each set of finite measure.

Let $\mathcal{F} = \{f$ is measurable $| \, v(E) \geq \int_E f \, d\mu$, for all $E \in \mathcal{M}\}$. $\mathcal{F}$ is nonempty because the zero function is in it. Let $s = \sup_{f \in \mathcal{F}} \int_X f \, d\mu$. Then there is a sequence $\langle h_n \rangle$ in $\mathcal{F}$ such that, $\lim_{n\to\infty} \int_X f_n d\mu = s.$

Let f_1, f_2 $\mathcal{F}$, then for any $E \in \mathcal{M}$, $\int_E f_1 \vee f_2 d\mu = \int_{\{x \in E | f_1(x) \geq f_2(x)\}} f_1 \vee f_2 d\mu$ $+ \int_{\{x \in E | f_1(x) < f_2(x)\}} f_1 \vee f_2 d\mu = \int_{\{x \in E | f_1(x) \geq f_2(x)\}} f_1 d\mu + \int_{\{x \in E | f_1(x) < f_2(x)\}} f_2 d\mu \leq v(\{x \in E$ $| f_1(x) \geq f_2(x)\})$ $+ v(\{x \in E \,|\, f_1(x) < f_2(x)\}) = v(E)$

Therefore, $f_1 \vee f_2 \in \mathcal{F}$.

Let $f_n = V_{k=1}^n h_k$. Then $\langle f_n \rangle$ is a nonnegative increasing sequence in $\mathcal{F}$ and $\lim_{n\to\infty} \int_X f_n d\mu = s$. Define g by $g(x) = \lim_{n\to\infty} f_n(x)$ for $x \in X$. Then by the motonone convergence theorem, for any $E \in \mathcal{M}$, $\int_E g d\mu = \lim_{n\to\infty} \int_E f_n d\mu \leq v(E)$. This shows $g \in \mathcal{F}$ and $\int_X g d\mu = \lim_{n\to\infty} \int_X f_n d\mu = s.$

Therefore, the function v_0 defined on $\mathcal{M}$ by $v_0(E) = v(E) - \int_E g d\mu$ is a measure. We can define that $v_0 = 0$ and then g is the required function. Suppose v_0 is not zero. Since $v_0(X) > 0$ and $\mu(X) < \infty$, there is $\varepsilon > 0$ such that, $v_0(X) - \in \mu(X) > 0$. Let $\{A, B\}$ be a Hahn decomposition for the signed measures $v_0 - \in \mu$. Then for every $E \in \mathcal{M}$, $v_0(A \cap E) - \in \mu(A \cap E) \geq 0$. Subsequently, $v(E) = v_0(E) + \int_E \; g d\mu \geq v_0(E$ $\cap A) + \int_E \; g d\mu \geq \in \mu \,(A \cap E) + \int_E \; g d\mu, = \int_E (g + \in \chi_A)$. Therefore, $g + \in \chi_A$ is also in $\mathcal{F}$. However, if $\mu(A) > 0$, then $\int_X (g + \in \chi_A) d\mu$ $= \int_X g d\mu + \mu(A) > \int_X g d\mu = s$, which is a contradiction. Obviously, if $\mu(A) = 0$, since $v << \mu$, $v(A) = 0$. Therefore $v_0(A) = v(A) - \int_A g d\mu \leq v(A) = 0$. Hence $v_0(A) = 0$. Consequently, $v_0(X) - \in \mu(X) = v_0(B) - \in \mu(B) \leq 0$, contradicting that $v_0(X) - \in \mu(X) > 0$.

Accordingly, $v_0 = 0$, which means $v(E) = \int_E \; g\mu$ for every $E \in \mathcal{M}$.

To show uniqueness, let $v(E) = \int_E f d\mu = \int_E g d\mu$. Then $\int_E (f - g) d\mu = 0$. Since E is arbitrary, $\int_{\{f-g \geq 0\}} f - g d\mu = 0$. This shows $f = g$ *a.e.* on $\{x \in X) | f(x) \geq g(x)\}$. Similarly, $f = g$ *a.e.* on $\{x \in X | f(x) < g(x)\}$. Hence $f = g$ *a.e.* on X.

Note: If we add the condition that v is finite, then the function g in our proof is integrable.

1.3.1 Lebesgue Decomposition

NOTES

In measure theory, Lebesgue's decomposition theorem states that for given μ and ν two σ-finite signed measures on a measurable space (Ω, Σ), there exist two σ-finite signed measures ν_0 and ν_1 such that:

- $v = v_0 + v_1$
- $v_0 << \mu$, i.e., v_0 is absolutely continuous with respect to μ.
- $v_1 \perp \mu$, i.e., ν_1 and μ are singular.

These two measures can be uniquely determined. The Lebesgue's decomposition theorem can be defined and improved in various ways. Primarily the decomposition of the singular part can be improved as follows:

$v = v_{cont} + v_{sing} + v_{pp}$

Where,

- ν_{cont} is the absolutely continuous part.
- ν_{sing} is the singular continuous part.
- ν_{pp} is the pure point part or a discrete measure.

Subsequently, absolutely continuous measures can be classified using the Radon-Nikodym theorem. Consequently, Lebesgue decomposition provides an extremely explicit description of measures. The Lebesgue decomposition theorem states that if (X, Ω) is a measurable space and μ is a finite measure on X, then for every measure v, there is a unique decomposition $\nu = v_1 + \nu_2$ such that, $\nu_1 << \mu$ and $\nu_2 \perp \mu$. Generalization of Lebesgue decomposition theorem can be defined as follows.

Assume that the space of all finite measures on (X, Ω) is denoted by M. Then the above is equivalent to the statement that $M = S \oplus T$, where S is the space of all measures that are absolutely continuous with respect to μ, while T is the space of all measures that are singular with respect to μ. We can characterize T in terms of S as follows,

$T = S^{\perp} = \{\nu \in M \mid \nu \perp m \text{ for all } m \in S\}$

Consequently state that a subspace $S \subset M$ has property D (for decomposition) if $M = S \oplus S^{\perp}$. Then the Lebesgue decomposition theorem defines that $\{\nu | \nu << \mu\}$ has property D for any fixed μ.

Theorem 1.4 (Lebesgue Decomposition): Let v be a σ-finite signed measures on (X, S) and μ be a σ-finite positive measure on (X, S). Then there exists unique σ-finite signed measures v_0 and v_1 on (X, S) such that $v = v_0 + v_1$, $\nu \perp \mu$ and $v_1 << \mu$.

This can be proved with the help of Lemma 1.

Lemma 1: Assume that v and μ are positive measures on (X, S) such that they do not take the value in $\{\infty, -\infty\}$. Then either $v \perp \mu$ or there exists a $\varepsilon > 0$ and $E \in S$ such that $\mu(E) > 0$ and $v(F) \geq \varepsilon\, \mu(F)$ for all $R \in S, F \subset E$.

Note: The conclusion is, that is a positive set for $v-\varepsilon\mu$.

Proof of Lemma: For each $n\in\mathbb{N}$, consider that (P_n, N_n) be the Hahn decomposition for $v-\frac{1}{n}\mu$. Put $P=\bigcup_1^\infty P_n$ and $N=\bigcap_1^\infty N_n$. Here N is a negative set for $v-\frac{1}{n}\mu, \forall n\in\mathbb{N}$. Therefore, $0\le v(N)\le\frac{\mu(N)}{n}$ for all n. Consequently, $v(N) = 0$.

NOTES

(a) If $\mu(P) = 0$ then because $P\bigcup N = X$ and $P\bigcap N=\emptyset$ we can state that $\mu \perp v$.

(b) If $\mu(P) > 0$ then $\mu(P_n) > 0$ for some $n\in\mathbb{N}$. Because P_n is a positive set for $v-\frac{1}{n}\mu$, we can obtain that $v(F)\ge\frac{1}{n}\mu(F)$ for all $F\in P_n$. Now take $P_n = E$ and $\varepsilon=\frac{1}{n}$. Hence the lemma is proved.

Proof: Assume that μ and v are finite positive measures.

Consider that $\mathcal{F} = \{f : X \to [0, \infty] : f$ is measurable and $\int_E fd\mu\le v(E), \forall E\in S\}$.

Then $\mathcal{F}\neq\emptyset$ because $0 \in \mathcal{F}$. If $f, g\in\mathcal{F}$, then we can state that $h = \max f, g \in \mathcal{F}$

If $A = \{x \in X : f(x) > g(x)\}$, then for given any $E\in S$,

$$hd\mu=\int_{E\cap A} hd\mu+\int_{E\setminus A} hd\mu=\int_{E\cap A} fd\mu+\int_{E\setminus A} gd\mu$$

$$\le v(E\cap A)+v(E\setminus A)=v(E)$$

If $a=\sup_{\mathcal{F}}\{\int_X f\,d\mu : f\in\mathcal{F}\}$ then we can state that $a\le v(V)<\infty$. Select $f_n\in\mathcal{F}$ such that, $\int_X f_n d\mu\to a$. Use $g_n=\max\{f_1,....,f_n\}$

Then, $g_n\le g_{n+1}$ and because $g_n\in\mathcal{F}$ we obtain, $a\ge\int_X g_n d\mu\ge\int_X f_n d\mu$,

Consequently, $\int_X g_n d\mu\to a$.

Now set $f(x)=\lim_{n\to\infty} g_n(x)$.

By the monotone convergence theorem, we obtain the following expression:

$$\int_X f\,d\mu=a<\infty$$

This implies that f is integrable and is finite almost everywhere.

NOTES

Set $v_0(E) = v(E) - \int_E f\, d\mu$ for $E \in S$. Here v_0 is a positive measure because $f \in \mathcal{F}$.

We can state that $\mu \perp v_0$. If it is not, then by Lemma 1 there exists $\varepsilon > 0$ and $E_0 \in S$ which is positive for $v_0 - \varepsilon\mu$ such that, $\mu(E_0) > 0$. Subsequently,

$$\int_E (F + \varepsilon\chi E_0) d\mu = \int_E fd\,\mu + \varepsilon\mu(E \cap E_0)$$

$$\leq \int_E f\, d\mu + v_0(E \cap E_0) \leq \int_E f\, d\mu + v_0(E) = v(E).$$

Therefore, $f + \varepsilon\chi E_0 \in \mathcal{F}$.

But, it is obvious that $\int_X f + \varepsilon\chi E_0 d\mu = a + \varepsilon\mu(E_0) > a$, which contra-dicts the truth that $\int_X f = a$ is a supremum in $\mathcal{F}$. As a result, we have $\mu \perp v_0$.

Hence we can state that $v_1(E) = \int_E f\, d\mu$.

Uniqueness: Assume that $v = \tau_0 + \tau_1$, where τ_i are signed measures as $\tau_0 \perp \mu$ and $\tau_1 << \mu$. We include $v = \tau_1 + \tau_0 = \tau_0 + \tau_1$. Consequently, $\tau_0 - v_0 = v_1 - \tau_1$.

By uniqueness of the Lebesgue decomposition theorem, $v = \int_E f\, d\mu$.

For the σ-finite case we have the following Lemma 2:

Lemma 2: Let λ be a σ-finite positive measure on (X, S) and assume that $\lambda(X) = \infty$. Then there exists an integrable function $w : X \to (0, 1)$ such that if $\lambda_w(E) = \int_E w\, d\lambda$ we have the following properties:

(a) λ_w is a finite measure.

(b) $\forall f \geq 0$ measurable and $\int_X f\, d\lambda_w = \int_X f\, wd\lambda$.

(c) $\lambda(E) = \int_E \frac{1}{w} d\lambda_w$.

(d) $\forall f \geq 0$ measurable and $\int_X F\, d\lambda = \int_X \frac{f}{w} d\lambda_w$.

(e) $E \in S$, $\lambda(E) = 0 \equiv \lambda_w(E) = 0$.

Proof: Let A_i be disjoint measurable sets such that $\bigcup_{i=1}^{\infty} A_i = X$ and $1 < \lambda(A_1) < \infty$.

Set $w = \sum_{i=1}^{\infty} \frac{1}{2^j \lambda(A_j)} \chi A_j$. Here $w(x) \in (0, 1)$ for all $x \in X$. Furthermore, by the monotone convergence theorem we have,

NOTES

$$\int_X w d\lambda = \sum_{j=1}^{\infty} \frac{1}{2^j \lambda(A_j)} \int_X \chi A_j d\lambda$$

Therefore, w is integrable and λ_w is a finite measure.

Proof of σ-Finite Positive Measures

Consider that v, μ be σ-finite positive measures. To obtain measurable functions $v, w : X \rightarrow (0, 1)$ with uv-integrable and w μ-integrable apply the lemma.

Set, $v_v(E) = \int_E v dv$ and $\mu_w(E) = \int_E w d\mu.$

Obtain unique measures v_v^3 and v_v^a such that $v_v^3 \perp \mu_w$, $v_v^a \ll \mu_w$ and $v_v^3 + v_v^a = v_v$.

Define $v_0(E) = \int_E \frac{1}{v} dv_v^3$ and $v_1(E) = \int_E \frac{1}{v} dv_v^a$.

Subsequently, we have $v_0(E) + v_1(E) = \int_E \frac{1}{v} d(v_v^3 + v_v^a) = v(E)$

Because $v_v^3 \perp \mu_w$ hence we have $v_0 \perp \mu$ and because $v_v^a \ll \mu_w$ we have $v_1 \ll \mu$.

Given that $v_v^a \ll \mu_w$ there exists an $h^{\mu w}$-integrable such that,

$$v_v^a(E) = \int_E h d\mu_w$$

Define, $v_1(E) \int_E \frac{1}{v} dv_v^a = \int_E \frac{1}{v} h d\mu_w = \int_E \frac{1}{v} h w d\mu$

To establish the theorem for a signed measures v relate the above to v^+ and v^- obtain the difference $v = v^+ - v^-$.

Notation: If $v \ll \mu$, we can state the $v(E) = \int_E f\, d\mu$. The function f is termed as the Radon-Nikodym derivative of v with respect to μ. We denote f by $\left[\frac{dv}{d\mu}\right]$.

When $v \ll \mu$ we can state that $dv = f d\mu$. This is stimulated by the truth that,

$$\int h dv = \int h\, f d\mu$$

Proposition: Assume that μ be a σ-finite signed measure and λ, μ are positive σ-finite measures such that $v \ll \mu$ and $\mu \ll \lambda$. Then the following conditions hold:

NOTES

(a) If $g \in L^1(v)$, then $\int g dv = \int_X g \left[\frac{dv}{d\mu}\right] d\mu$

(b) If $v \ll \lambda$, then $\left[\frac{dv}{d\lambda}\right] = \left[\frac{dv}{d\mu}\right] \cdot \left[\frac{d\mu}{d\lambda}\right]$

Proof: To prove (a), establish it for v^+ and v^- using standard methods and subtract.

To prove (b), we obtain:

$$v(E) = \int_E \left[\frac{dv}{d\mu}\right] d\mu = \int_E \left[\frac{dv}{d\mu}\right] \cdot \left[\frac{d\mu}{d\lambda}\right] d\lambda$$

Consequently, $v \ll \mu$ and subsequently

$$\left[\frac{dv}{d\mu}\right] = v(E) = \int_E \left[\frac{dv}{d\mu}\right] d\mu = \int_E \left[\frac{dv}{d\mu}\right] \cdot \left[\frac{d\mu}{d\lambda}\right] d\lambda$$

Hence proved.

1.3.2 Riesz Representation Theorem

Let H be a Hilbert space over $\mathbb{R}$ or $\mathbb{C}$, and T a bounded linear functional on (a bounded operator from to the field, $\mathbb{R}$ or $\mathbb{C}$, over which is defined), then there exists some $g \in H$ such that for every $f \in H$ we have

$$T(f) = \langle f, g \rangle$$

Moreover, $\| T \| = \| g \|$ (here $\| T \|$ (here denotes the operator norm of T, while $\| g \|$ is the Hilbert space norm of g.

Proof: Let us assume that H is separable and consider the case on $\mathbb{R}$.

Since H is separable wse can choose an orthonormal basis $\phi_j, j \geq 1$, for . Let be a bounded linear functional and set .

Choose $f \in H$, let $c_j = \langle f, \phi_j \rangle$, and define $f_n = \sum_{j=1}^{n} c_j \phi_j$. Since the ϕ_j form a basis we know that $\|f - f_n\| \to 0$ as $n \to \infty$.

$$T(f_n) = \sum_{j=1}^{n} a_j c_j. \quad \ldots(1.2)$$

Since T is bounded, say with norm $\| T \| < \infty$, we have

$$|T(f) - T(f_n)| \leq \| T \| \, \|f - f_n\|. \quad \ldots(1.3)$$

Since $\|f - f_n\| \to 0$ as $n \to \infty$ from (1.2) and (1.3) we can say that

$$T(f) = \lim_{n\to\infty} T(f_n) = \sum_{j=1}^{\infty} a_j c_j. \quad \ldots(1.4)$$

NOTES

Also the sequence a_j must itself be square-summable. To check this, first note that since $|T(f)| \leq \| T \| \| f \|$ we have

$$\left| \sum_{j=1}^{\infty} c_j a_j \right| \leq \| T \| \left(\sum_{j=1}^{\infty} c_j^2 \right)^{1/2}. \qquad \ldots(1.5)$$

The above inequality must hold for any square-summable sequence c_j (since any such c_j corresponds to some element in *H*). Fix a positive integer *N* and define a sequence *cj* = *aj* for $j \leq N, c_j = 0$ for . Clearly such a sequence is square-summable and equation (1.5) then gives

$$\left| \sum_{j=1}^{N} a_j^2 \right| \leq \| T \| \left(\sum_{j=1}^{N} a_j^2 \right)^{1/2}$$

or

$$\left(\sum_{j=1}^{N} a_j^2 \right)^{1/2} \leq \| T \|. \qquad \ldots(1.6)$$

Thus a_j is square-summable, since the sequence of partial sums is bounded above.

Since a_j is square-summable the function $g = \sum_i a_i \phi_i$ is well-defined as an element of H, and $T(f) = \sum_i a_i c_i = < f, g >$. Then from equation (1.6), $\| g \| \leq \| T \|$. But from Cauchy-Schwarz we have $|T(f)| = | < f, g > | \leq \| f \| \| g \|$ or $\frac{|T(f)|}{\|f\|} \leq \| g \|$, implying $\| T \| \leq \| g \|$, so $\| T \| = \| g \|$.

Application

The example below shows how functional analytic methods are used in ODE. Consider the equation below:

$$-f''(x) + b(x)f(x) = q(x) \qquad \ldots(1.7)$$

on the interval $0 < x < 1$, where $b(x) \geq \delta > 0$ for some δ. Assume that the functions b and q are continuous on $[0,1]$. Let us find a solution to equation (1.7) with $f'(0) = f'(1) = 0$ (considering arbitrary boundary conditions). If we multiply (1.7) by a C^1 function ϕ and integrate the first term, $-f''\phi$ by parts from $x = 0$ to $x = 1$, we get

$$\int_0^1 (f'(x)\phi'(x) + b(x)f(x)\phi(x))\, dx = \int_0^1 q(x)\phi(x)dx. \qquad \ldots(1.8)$$

The above equation must hold for any $\phi \in C^1([0,1])$, if f is a $C^2(0,1)$ solution to equation (1.7) whsich is continuous on $[0,1]$. Conversely, if for a given C^2 function we find that (1.8) holds for all , then must be a solution to equation (1.7), for if we "undo" the integration by parts in (1.8) we obtain

$$\phi(1)f'(1) - \phi(0)f'(0) + \phi(x)(-f''(x) + b(x)f(x)) = \phi(x)q(x)$$

for all ϕ. A familiar PDE argument then shows that $f'(0) = f'(1) = 0$ and equation (1.7) must hold.

We can show that there is a unique solution to equation (1.8). Such a solution will not necessarily be twice-differentiable as required by equation (1.7), but it will satisfy equation (1.8). Equation (1.8) is often called the "weak" formulation of the problem.

NOTES

Define an inner product

$$< g,h >= \int_0^1 (g'(x)h'(x) + b(x)g(x)h(x))\,dx$$

onsa the space $C^1([0,1])$, and let H denote the completion of the this space. (We must use $b \geq \delta > 0$ to assure that $\leq r$ anglerally is an inner product, so that, $\| g \| = \sqrt{\langle g,g \rangle} = 0$ iff $g \equiv 0$). The space H is a Hilbert space, and can be interpreted (if need be) as a subspace of $C([0,1])$

Define a functional $T: H \to \mathbb{R}$ by

$$T(\phi) = \int_0^1 q\,(x)\phi(x)dx$$

One can easily check that T is bounded on H using Cauchy-Schwarz theorem. From the Riesz Representation Theorem it then follows that there must exist some function $f \in H$ such that

$$T(\phi) = \langle f,\phi \rangle$$

for all $\phi \in H$. This is exactly equation (1.8), the weak form of the ODE, , the function f that satisfies equation 1.8 lies in H and F is a continuous function.

1.4 EXTENSION THEOREM (CARATHÉODORY)

We fix a topological space Ω. The power set of Ω is denoted by $\mathcal{P}(\Omega)$ and consists of all subsets of Ω.

Definition: A ring on Ω is a subset $\mathcal{R}$ of $\mathcal{P}(\Omega)$, such that

1. $\theta \in \mathcal{R}$.
2. $A, B \in \mathcal{R} \quad \Rightarrow \quad A \cup B \in \mathcal{R}$.
3. $A, B \in \mathcal{R} \quad \Rightarrow \quad A \setminus B \in \mathcal{R}$.

Definition: A σ-algebra on Ω is a subset Σ of $\mathcal{P}(\Omega)$, such that

1. $\theta \in \Sigma$.
2. $(A_n)_{n \in \mathbb{N}} \in \Sigma \quad \Rightarrow \quad \bigcup_n A_n \in \Sigma$.
3. $A \in \Sigma \quad \Rightarrow \quad A^C \in \Sigma$.

Since, $A \cap B = A \setminus (A \setminus B)$ it follows that any ring on Ω is closed under finite intersections; hence any ring is also a semi-ring. Since, $\bigcap_n A_n = (\bigcap_n A^C)^C$ it follows that any σ-algebra is closed under arbitrary intersections. And from $A \setminus B = A \cap B^c$, we deduce that any σ-algebra is also a ring.

If $(R_i)_{i\in I}$ is a set of rings on Ω then it is clear that $\bigcap_I R_i$ is also a ring on Ω. Let S be any subset of $P(\Omega)$, then we call the intersection of all rings on Ω containing S the ring generated by S.

Definition: Let $\mathcal{A}$ be a subset of $\mathcal{P}(\Omega)$. A measure on $\mathcal{A}$ is a map $\mu : \mathcal{A} \to [0, +\inf]$ such that

1. $\mu(\theta) = 0$.
2. If $A_n \in \mathcal{A}$ are disjoint and $A = \uplus_n A_n \in \mathcal{A}$ then $\mu(A) = \sum_n \mu(A_n)$.

If $\mathcal{A}$ is a σ-algebra, we do not need to assume that in addition $\uplus_n A_n \in \mathcal{A}$. By taking all but finitely many A_n to be the empty sets one sees that $\mu(A_1 \uplus ... \uplus A_N) = \mu(A_1) + ... + \mu(A_n)$. If $A \subset B$ then $A \uplus (B \setminus A) = B$ and hence $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.

Definition: We call an outer measure on Ω a map $\lambda : \mathcal{P}(\Omega) \to [0, +\infty]$ with,

1. $\lambda(\theta) = 0$.
2. $A \subset B \quad \Rightarrow \quad \lambda(A) \leq \lambda(B)$.
3. $(A_n)_{n\in\mathbb{N}} \mathcal{P}(\Omega),\ \lambda(\cup_n A_n) \leq \sum_n \lambda(A_n)$.

By taking all but finitely many A_n to be the empty set one sees that an outer measure is subadditive; $\lambda(A \cup B) \leq \lambda(A) + \lambda(B)$.

Let λ be an outer measure on Ω. We define $\sum_\lambda$ to be the set of all subsets $A \subset \Omega$ such that for any $X \subset \Omega$, we have $\lambda(X) = \lambda(X \cap A) + \lambda(X \cap A^C)$.

In other words, $\sum_\lambda$ consists of all subsets $A \subset \Omega$ that cut Ω in two in a good way. Clearly $\Omega \in \Sigma_\lambda$ and by the very form of the definition of Σ_λ, we have $A \in \Sigma_\lambda \quad \Leftrightarrow \quad A^C \in \Sigma_\lambda$.

Theorem 1.5: Let, λ be an outer measure on Ω and let $\sum_\lambda$ be as defined above. Then $\sum_\lambda$ is a σ-algebra on Ω.

Proof: After the preliminary remarks preceding the Theorem, it only remains to prove that $\sum_\lambda$ is closed under countable unions. We will first prove that $\sum_\lambda$ is closed under finite intersections and unions.

Let, A, B $\in \sum_\lambda$ and let, X be any subset of Ω. We have $X \cap A^C = X \cap (A \cap B^C) \cap A^C$ since $(A \cap B)^C \supset A^C$. On the other hand we have $(A \cap B)^C = A^C \cup B^C$ and hence $X \cap (A \cap B)^C \cap A = (X \cap A \cap B^C) \cup (X \cap A \cap A^C) = X \cap A \cap B^C$. Therefore we have, $\lambda(X \cap (A \cap B)^C) = \lambda(X \cap (A \cap B)^C \cap A) + \lambda(X \cap (A \cap B)^C \cap A^C = \lambda(X \cap A^C) + \lambda(X \cap A \cap B^C)$

NOTES

Now adding $\lambda(X \cap A \cap B)$ and using that $\lambda(X \cap A) = \lambda(X \cap A \cap B)$ $+ \lambda(X \cap A \cap B^C)$, one obtains $\lambda(X \cap A \cap B) + \lambda(X \cap (A \cap B^C) = \lambda(X)$. Hence $A \cap B \in \Sigma_\lambda$.

Since $A \cup B = (A^C \cap B^C)^C$ and $A \setminus B = A \cap B^C$ we see that Σ_λ is closed under finite unions and the set-theoretic difference. Thus Σ_λ is a ring on Ω.

If A, $B \in \Sigma_\lambda$ are disjoint and $X \subset \Omega$ then, $\lambda(X \cap (A \uplus B) = \lambda(X) - \lambda(X \cap A^C \cap B^C) = \lambda(X) - \lambda(X \cap A^C) + \lambda\ (X \cap A^C \cap B) = \lambda(X \cap A) + \lambda$ $(X \cap B)$ as $A^C \cap B = B$

Using induction we obtain $\lambda(X \cap \uplus_{n=1}^N A_n) = \sum_{n=1}^N \lambda(X \cap A_n)$ whenever A_n are in Σ_λ and pairwise disjoint. Now we fix a sequence A_n in Σ_λ which are pairwise disjoint and we denote the union $\bigcup_n A_n$ by A. Furthermore, we fix an arbitrary $X \in \Omega$ and an arbitrary large integer N.

Since $X \cap A^C \subset X \cap (\uplus_{n=1}^N A_n)^C$ and Σ_λ is closed under finite unions, we have $\lambda(X \cap A^C) + \lambda(X \cap (\uplus_{n=1}^N A_n)) \leq \lambda(X \cap (\uplus_{n=1}^N A_n)^C) + \sum_n \lambda(X \cap A_n) = \lambda(X)$

But N is arbitrary in this equation and so we can obtain,

$$\lambda(X \cap A^C) + \sum_n \lambda(X \cap A_n) \leq \lambda(X) \qquad \ldots(1.9)$$

On the other hand we have $\lambda(X) \leq \lambda(X \cap A^C) + \lambda(X \cap A)$, which again by the definition of an outer measure is less than or equal to $\lambda(X \cap A^C) + \sum_n \lambda(X \cap A_n)$. Hence again using Equation (1.9) we obtain,

$$\lambda(X) \leq \lambda(X \cap A^C) + \lambda(X \cap A) \leq \lambda(X \cap A^C) + \sum_n \lambda(X \cap A_n) \leq \lambda(X)$$

From this we conclude that Σ_λ is indeed closed under countable unions and , by taking $X = A$ that $\lambda(A) = \sum_n \lambda(A_n)$. Therefore the restriction of λ to Σ_λ is a measure on Σ_λ. Hence proved.

We will call Σ_λ the σ-algebra related to λ.

Now we come to a critical step; we want to associate an outer measure λ_μ to a given measure λ on some ring $\mathcal{R}$. Of course, we want the restriction of the outer measure λ_μ to the ring to coincide with the measure μ.

Let $\mathcal{R}$ be a ring on Ω and let μ be a measure on $\mathcal{R}$. If $X \subset \Omega$ is any subset we can cover X with sets from $\mathcal{R}$ to approximate X inside $\mathcal{R}$—we call an $\mathcal{R}$-cover of X a countable subset $(A_n)_n$ of $\mathcal{R}$ with $X \subset \bigcup_n A_n$. This leads to the following definition; for any $X \subset \Omega$ we define $\lambda_\mu(X)$ to be the infimum of all sums $\sum_n \mu(A_n)$ where $(A_n)_{n\in\mathbb{N}}$ is any countable cover of X with A_n in $\mathcal{R}$. We need to check that this is an outer measure.

Theorem 1.6: The map $\lambda_\mu : \mathcal{P}(\Omega) \to [0, +\infty]$ defined above defines an outer measure on Ω.

Proof: Since $\phi \in \mathcal{R}$, we have $\lambda_\mu(\phi) = 0$. If $X \subset Y$ are two subsets of Ω, then any cover of Y with sets from $\mathcal{R}$ also covers X and hence $\lambda_\mu(X) \le \lambda_\mu(Y)$.

Now let X_n be any sequence of subsets. By the definition of the infimum we can find $\varepsilon > 0$ and for each n an $\mathcal{R}$-cover $(A_{n,m})_m$ of X_n such that, $\sum_m \mu(A_{n,m}) < \lambda_\mu(X) + \frac{\varepsilon}{2^n}$. The sets $A_{n,m}$ form a countable cover of $X = \bigcup_n X_n$; we can for example set $B_1 = A_{1,1}, B_2 = A_{2,1}, B_3 = A_{1,2}, B_4 = A_{3,1}$ and so on, similar to Cantor's proof of the countability of $\mathbb{Q}$. But then, $\lambda_\mu(X) = \lambda_\mu(\bigcup_n X_n) \le$

$$\sum_{n,m} \mu(A_{n,m}) < \sum_n (\lambda_\mu(X_n) + \frac{\varepsilon}{2^n} = \sum_n \lambda_\mu(X_n) + \varepsilon$$

But ε was arbitrary and hence $\lambda_\mu(\bigcup_n X_n) \le \sum_n \lambda_\mu(X_n)$.

Theorem 1.7: The restriction of λ_μ to $\mathcal{R}$ is μ.

Proof: For any $A \in \mathcal{R}$ the set A itself forms a cover and hence $\lambda_\mu(A) \le \mu(A)$. On the other hand, let (A_n) be an $\mathcal{R}$-cover of A. We define $B_1 = A_1 \cap A$ and $B_{n+1} = (A_n \cap A) \setminus \cup_{k \ge n} (A_k \cap A)$. Then clearly $B_n \in \mathcal{R}$, the B_n are disjoint, $\uplus_n B_n = A$ and $\mu(B_n) \le \mu(A_n)$. Since μ is a measure on $\mathcal{R}$ we have $\mu(A) = \sum_n \mu(B_n)$ which is less than or equal to $\sum_n \mu(A_n)$. Since this holds for any $\mathcal{R}$-cover of A we have $\mu(A) \le \lambda_\mu(A)$. Therefore equality holds and the Theorem is proved.

We will call λ_μ the outer measure associated to μ.

So, we now have two constructions; given a ring and a measure on it we can construct an outer measure. Given an outer measure, we can construct a σ-algebra such that the restriction of the outer measure to the σ-algebra is a measure on the σ-algebra. So it seems feasible that we can construct a measure on a σ-algebra starting from a ring with a measure on it.

Lemma: Let $\mathcal{R}$ be a ring on Ω and let μ be a measure on $\mathcal{R}$. Let λ be the outer measure associated to μ. Let Σ be the σ-algebra related to λ. Then $\mathcal{R} \in \Sigma$.

Proof: Let A be an element of $\mathcal{R}$ and let X be any subset of Ω. Since, λ is an outer measure on Ω we have

$$\lambda(X) = \lambda((X \cap A) \cup (X \cap A^C)) \le \lambda(X \cap A) + \lambda(X \cap A^C)$$

Now let $(A_n)_{n \in \mathbb{N}}$ be any $\mathcal{R}$-cover of X. Then the $A_n \cap A$ form an $\mathcal{R}$-cover of $X \cap A$ and the $A_n \cap A^C$ form an $\mathcal{R}$-cover of $X \cap A^C$. Hence we

NOTES

have that $\lambda(X \cap A) + \lambda(X \cap A^C) \leq \sum_n \mu(A_n \cap A) + \sum_n (A_n \cap A^C) = \sum_n \mu(A_n)$, where the last step follows from the fact that μ is a measure and hence $\mu(C \uplus D) = \mu(C) + \mu(D)$. Since the inequality holds for any $\mathcal{R}$-cover of X we need $\lambda(X \cap A) + \lambda(X \cap A^C) \leq \lambda(X)$. We thus need equality; for any $X \subset \Omega$ we have $\lambda(X) = \lambda(X \cap A) + \lambda(X \cap A^C)$, or in other words $A \in \Sigma$ and since A was an arbitrary element of $\mathcal{R}$ the lemma is proved.

Note: Any ring generates a σ-algebra; one simply enlarges the ring with countable unions. Or, the σ-algebra generated by the ring $\mathcal{R}$ is the intersection of all σ-algebras that contain $\mathcal{R}$. Therefore, the above lemma shows that the σ-algebra generated by $\mathcal{R}$ is contained in Σ.

Theorem 1.8 (Caratheodory): Let $\mathcal{R}$ be a ring on Ω and let μ be a measure on $\mathcal{R}$. Then there exists a measure μ′ on the σ-algebra generated by $\mathcal{R}$ such that the restriction of μ′ to $\mathcal{R}$ coincides with μ.

Proof: Let λ be the outer measure on Ω associated to μ. Let Σ be the σ-algebra associated to λ. Then by the above lemma the σ-algebra generated by $\mathcal{R}$ is contained in Σ. Hence, λ restricts to a measure on the σ-algebra generated by $\mathcal{R}$. By Theorem 1.7 this restriction of λ to $\mathcal{R}$ coincides with μ. Hence the theorem is proved.

For example, consider Ω to be the real line. Then the open intervals generate a σ-algebra Σ. For any open interval (a, b) with $a < b$ we can put $\mu((a, b)) = b - a$. Then there exists a measure μ′ on Σ such that $\mu'((a, b)) = b - a$. Indeed, for countable unions of disjoint intervals we can define $\mu(\cup_n (a_n, b_n)) = \sum_n (b_n - a_n)$. Hence μ does give rise to a measure on the ring generated by all intervals.

1.4.1 Lebesgue-Stieltjes Integral

Lebesgue-Stieltjes integrals are named after Henri Leon Lebesgue and Thomas Joannes Stieltjes. The Lebesgue-Stieltjes integration generalizes the Riemann-Stieltjes and Lebesgue integration. The Lebesgue-Stieltjes integral is the ordinary Lebesgue integral with respect to a measure known as the Lebesgue-Stieltjes measure which can be combined to any function of bounded variation on the real line. The Lebesgue-Stieltjes measure is a regular Borel measure and conversely we can state that every regular Borel measure on the real line is of this type.

The Lebesgue-Stieltjes integral is,

$$\int_a^b f(x)\, dg(x)$$

This is defined as $f : [a, b] \to \mathbf{R}$ which is Borel measurable and bounded, and $g : [a, b] \to \mathbf{R}$ which is of bounded variation in $[a, b]$ and right continuous or when f is nonnegative and g is monotone and right continuous. To establish, presume that f is nonnegative and g is monotone non-decreasing and right continuous. Describe $w((s, t]) := g(t) - g(s)$ and $w(\{a\}) := 0$. There is a unique Borel measure μ_g on $[a, b]$ which agrees with w on every interval I. The measure μ_g occurs from an outer measure given by,

$$\mu_g(E) = \inf\left\{\sum_i \mu_g(I_i) \mid E \subset \bigcup_i I_i\right\}$$

The infimum is taken over all E by countably several semi-open intervals. This measure is termed as the Lebesgue-Stieltjes measure associated with g.

NOTES

We can define the Lebesgue-Stieltjes integral as the Lebesgue integral of f with respect to the measure μ_g in the standard method. If g is non-increasing, then define,

$$\int_a^b f(x)dg(x) := -\int_a^b f(x)d(-g)(x),$$

If g is of bounded variation and f is bounded, then we can write as follows:

$g(x) = g_1(x) - g_2(x)$

Here $g_1(x) := V_a^x g$ is the total variation of g in the interval $[a,x]$ and $g_2(x) = g_1(x) - g(x)$. Both g_1 and g_2 are monotone non-decreasing. Now the Lebesgue-Stieltjes integral with respect to g is defined as follows,

$$\int_a^b f(x)dg(x) = \int_a^b f(x)dg_1(x) - \int_a^b f(x)dg_2(x),$$

Integration by Parts

A function f is said to be regular at a point a if the right and left hand limits $f(a+)$ and $f(a-)$ exist and the function takes the average value at the limiting point as follows,

$$f(a) = \frac{1}{2}(f(a-) + f(a+)),$$

Given two functions U and V of finite variation, if at each point either U or V is continuous or if both U and V are regular then there is an integration by parts for the Lebesgue-Stieltjes integral. It can be expressed as follows:

$$\int_a^b U\,dV + \int_a^b V\,dU = U(b+)V(b+) - U(a-)\,V(a-),$$

Here $b > a$.

When $g(x) = x$ for all real x, then μ_g is the Lebesgue measure and the Lebesgue-Stieltjes integral of f with respect to g is equivalent to the Lebesgue integral of f.

1.4.2 Product Measures and Fubini's Theorem

Given two measurable spaces and measures on them we can obtain the product measurable space and the product measure on that specific space. Theoretically, this can be defined using the Cartesian product of sets. Consider that (X_1, Σ_1) and (X_2, Σ_2) be two measurable spaces, i.e., Σ_1 and Σ_2 are sigma algebras on X_1 and X_2, respectively. Now let μ_1 and μ_2 be measures on these spaces. We can denote this by $\Sigma_1 \times \Sigma_2$ the sigma algebra on the Cartesian product $X_1 \times X_2$ produced by subsets of the form $B_1 \times B_2$, where $B_1 \in \Sigma_1$ and $B_2 \in \Sigma_2$.

The product measure $\mu_1 \times \mu_2$ is defined to be the unique measure on the measurable space $(X_1 \times X_2, \times \Sigma_1 \times \Sigma_2)$ satisfying the property,

$(\mu_1 \times \mu_2)(B_1 \times B_2) = \mu_1(B_1)\, \mu_2(B_2)$

This holds for all $B_1 \in \Sigma_1, B_2 \in \Sigma_2$.

Actually, when the spaces are σ-finite then for every measurable set E we can define,

$$(\mu_1 \times \mu_2)(E) = \int_{X_2} \mu_1(E^y)\, d\mu_2(y) = \int_{X_1} \mu_2(E_x)\, d\mu_1(x)$$

Here $E_x = \{y \in X_2 | (x,y) \in E\}$ and $E^y = \{x \in X_1 | (x,y) \in E\}$, which are both measurable sets.

The uniqueness of product measure is guaranteed only in the case that both (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) are σ-finite. The Borel measure on the Euclidean space $\mathbf{R}^n$ can be obtained as the product of n of the Borel measure on the real line $\mathbf{R}$.

Let $(E_1, \varepsilon_1, \mu_1)$ and $(E_2, \varepsilon_2, \mu_2)$ be finite measure spaces. Then the set,

$\mathcal{A} = \{A_1 \times A_2 : A_1 \in \varepsilon_1, A_2 \in \varepsilon_2\}$ is a π-system of subsets of $E = E_1 \times E_2$. Define the product σ-algebra $\varepsilon_1 \otimes \varepsilon_2 = \sigma(\mathcal{A})$. Set $\varepsilon = \varepsilon_1 \otimes \varepsilon_2$.

Lemma 1: Let $f : E \to \mathbb{R}$ be ε-measurable. Then for all $x_1 \in E_1$ the function $x_2 \mapsto f(x_1, x_2) : E_2 \to \mathbb{R}$ is ε_2-measurable.

Proof: The set of ε-measurable functions can be denoted by v to evaluate the result and for which the condition holds. Subsequently v is considered as a vector space which contains the indicator function of every set $A \in \mathcal{A}$. Further if $f_n \in v$ for all n and if f is bounded with $0 \le f_n \uparrow f$ subsequently $f \in v$ also. Thus, according to monotone class theorem v contains all bounded ε-measurable functions.

Lemma 2: For all bounded ε-measurable functions f the function is of the form,

$$x_1 \mapsto f_1(x_1) = \int_{E_2} f(x_1, x_2)\mu_2(dx_2) : E_1 \to \mathbb{R}$$

This function is considered bounded and ε_1-measurable. This lemma can be easily solved by using the monotone class theorem as we have done in Lemma 1. For this μ_1 and μ_2 must be finite.

Theorem 1.9 (Product Measure): There exists a unique measure $\mu = \mu_1 \otimes \mu_2$ on ε such that, $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$. This holds for all $A_1 \in \varepsilon_1$ and $A_2 \in \varepsilon_2$.

Proof: Uniqueness holds since $\mathcal{A}$ is a π-system generating ε. Using the already defined lemmas we can define the existence as follows,

$$\mu(A) = \int_{E_1}\left(\int_{E_2} 1_A(x_1, x_2)\mu_2(dx_2)\right)\mu_1(dx_1)$$

Use monotone convergence to define that μ is countably additive.

Proposition: Let $\hat{\varepsilon} = \varepsilon_2 \otimes \varepsilon_1$ and $\hat{\mu} = \mu_2 \otimes \mu_1$. For a function f on $E_1 \times E_2$ we can define $\hat{f}$ for the function on $E_2 \times E_1$ specified by $\hat{f}(x_2, x_1) = f(x_1, x_2)$. If f is ε-measurable then $\hat{f}$ is $\hat{\varepsilon}$-measurable. Further if f is non-negative then $\hat{\mu}(\hat{f}) = \mu(f)$.

Fubini's Theorem

Fubini's theorem is named after Guido Fubini. It is a consequence which provides conditions for which it is possible to compute a double integral using iterated integrals. As a consequence it also permits the order of integration to be changed in iterated integrals.

Fubini's theorem is also sometimes termed as Tonelli's theorem. It is used to establish a connection between a multiple integral and a repeated one. If $f(x, y)$ is continuous on the rectangular region $R : a \le x \le b, c \le y \le d$, then the equality of the form $\iint_R f(x,y)d(x,y) = \int_a^b \int_c^d f(x,y)dy\,dx$ holds.

Theorem Statement

Let A and B are complete measure spaces. Assume that $f(x,y)$ is $A \times B$ measurable if,

$$\int_{A\times B} |f(x,y)|\,d(x,y) < \infty,$$

Here the integral is taken with respect to a product measure on the space over $A \times B$. Subsequently,

$$\int_A \left(\int_B f(x,y)\,dy\right)dx = \int_B \left(\int_A f(x,y)\,dx\right)dy = \int_{A\times B} f(x,y)\,d(x,y),$$

Here the first two integrals are considered as the iterated integrals with respect to two measures and the third integral is with respect to a product of these two measures. If the integral of the absolute value is not finite then the two iterated integrals may essentially have different values.

Corollary: If $f(x,y) = g(x)h(y)$ for some functions g and h, then we can state that:

$$\int_A g(x)\,dx \int_B h(y)\,dy = \int_{A\times B} f(x,y)\,d(x,y),$$

Here the integral on the right side is defined with respect to a product measure.

Another alternate statement of Fubini's theorem states that if A and B are σ-finite measure spaces and not essentially absolute, and if either $\int_A \left(\int_B |f(x,y)|\,dy\right)dx < \infty$ or $\int_B \left(\int_A |f(x,y)|\,dx\right)dy < \infty$ holds then $\int_{A\times B} |f(x,y)|\,d(x,y) < \infty$.

And, $\int_A \left(\int_B f(x, y)\, dy \right) dx = \int_B \left(\int_A f(x, y)\, dx \right) dy = \int_{A \times B} f(x, y)\, d(x, y)$.

For this the essential condition is that the measures must be σ-finite.

NOTES

Check Your Progress

9. Define Radon-Nikodym theorem.
10. State the Caratheodory extension theorem.
11. What does Lebesgue's decomposition theorem state?
12. What is Lebesgue-Stieltjes integral?
13. Define Fubini's theorem.

1.5 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Signed measures is referred as a simplification of the concept of measures that allows it to have negative values. There are two different notions of a signed measures that depends on the condition that how the infinite values are taken.
2. A finite signed measures can be defined except that it only takes the real values, i.e., it cannot take $+\infty$ or $-\infty$. Finite signed measures form a vector space.
3. Hahn decomposition theorem states that for a given signed measures μ, there exist two measurable sets P and N. This decomposition is unique for adding/subtracting μ null sets from P and N.
4. The sum of two finite signed measures is a finite signed measures because it is the product of a finite signed measures by a real number which is considered closed under linear combination. It follows the assumption that the set of finite signed measures on a measures space (X, Σ) is a real vector space. The total variation defines a norm for which the space of finite signed measures becomes a Banach space. As per the Riesz representation theorem, if X is a compact separable space then the space of finite signed Baire measures is considered the dual of the real Banach space of all continuous real valued functions on X.
5. The integral of a complex valued measurable function can be defined with respect to a complex measures by approximating a measurable function with simple functions. The already existing integral of a real valued function can be used with respect to a nonnegative measure. The real and imaginary parts μ_1 and μ_2 of a complex measures μ are considered finite valued signed measures.
6. A decomposition of the measures ν as a difference of two mutually singular measures v^+ and v^- is termed as Jordon decomposition.
7. Two complex measures μ and ν on a measures space X are considered mutually singular if they are provided on different subsets.

8. Two positive or signed or complex measures μ and ν defined on a measurable space (Ω, Σ) are called singular if there exist two disjoint sets A and B in Σ whose union is Ω such that μ is zero on all measurable subsets of B while ν is zero on all measurable subsets of A.
9. The Radon-Nikodym theorem is a consequence in measure theory that states that given a measurable space (X, Σ), if a σ-finite measure ν on (X, Σ) is absolutely continuous with respect to a σ-finite measure μ on (X, Σ) then there is a measurable function f on X which takes values in $[0, \infty)$ such that,

   $$v(A) = \int_A f \, d\mu$$

   This holds for any measurable set A.
10. Let $\mathcal{R}$ be a ring on Ω and let μ be a measures on $\mathcal{R}$. Then there exists a measures μ′ on the σ-algebra generated by $\mathcal{R}$ such that the restriction of μ′ to $\mathcal{R}$ coincides with μ.
11. In measures theory, Lebesgue's decomposition theorem states that for given μ and ν two σ-finite signed measures on a measurable space (Ω, Σ), there exist two σ-finite signed measures ν_0 and ν_1.
12. The Lebesgue-Stieltjes integral is the ordinary Lebesgue integral with respect to a measures known as the Lebesgue-Stieltjes measures which can be combined to any function of bounded variation on the real line.
13. Fubini's theorem is also sometimes termed as Tonelli's theorem. It is used to establish a connection between a multiple integral and a repeated one. If $f(x, y)$ is continuous on the rectangular region $R : a £ x £ b, c £ y £ d$, then the equality of the form $\iint_R f(x, y)d(x, y) = \int_a^b \int_c^d f(x, y)dy\,dx$ holds.

1.6 SUMMARY

- Signed measures is referred as a simplification of the concept of measures that allows it to have negative values. There are two different notions of a signed measures that depends on the condition that how the infinite values are taken.
- A finite signed measures can be defined except that it only takes the real values, i.e., it cannot take +∞ or –∞. Finite signed measures form a vector space.
- An extended signed measures is the difference of two nonnegative measures and a finite signed measures is the difference of two finite nonnegative measures.
- Hahn decomposition theorem states that for a given signed measures μ, there exist two measurable sets P and N. This decomposition is unique for adding/subtracting μ null sets from P and N.
- The sum of two finite signed measures is a finite signed measures because it is the product of a finite signed measures by a real number which is

NOTES

considered closed under linear combination. It follows the assumption that the set of finite signed measures on a measures space (X, Σ) is a real vector space.

- The total variation defines a norm for which the space of finite signed measures becomes a Banach space. As per the Riesz representation theorem, if X is a compact separable space then the space of finite signed Baire measures is considered the dual of the real Banach space of all continuous real valued functions on X.
- A complex measures μ on a measurable space (X, Σ) is a function defined on Σ which takes complex values that is sigma additive.
- The integral of a complex valued measurable function can be defined with respect to a complex measures by approximating a measurable function with simple functions.
- The already existing integral of a real valued function can be used with respect to a nonnegative measure. The real and imaginary parts μ_1 and μ_2 of a complex measures μ are considered finite valued signed measures.
- The sum of two complex measures is also referred as a complex measures and similarly the product of a complex measures by a complex number.
- The pair (P, N) is termed as a Hahn decomposition of the signed measures μ.
- Jordan decomposition theorem states that every signed measures μ can be expressed as a difference of two positive measures μ^+ and μ^-, of which at least one is finite.
- Every measurable subset of a positive set is termed positive and a union of a countable collection of disjoint positive sets is also positive.
- Given a signed Measures space (X, B, ν) there is a positive set A and a negative set B such that A and B partition X.
- A decomposition of the measures ν as a difference of two mutually singular measures v^+ and v^- is termed as Jordon decomposition.
- For a signed measures and a measurable function f, we state that f is integrable with respect to v if f is integrable with respect to $v+$ and f is also integrable with respect to v^-.
- If v is a signed measures and μ is a (honest) measures on a measurable space (X, S) then we can state that v is absolutely continuous with respect to μ if for and $\mu(E) = 0$ we have $v(E)$.
- Two complex measures μ and ν on a measures space X are considered mutually singular if they are provided on different subsets.
- A discrete singular measures with respect to Lebesgue measures on the real integrals is a measures λ defined at 0.
- Two positive or signed or complex measures μ and ν defined on a measurable space (Ω, Σ) are called singular if there exist two disjoint sets A and B in Σ whose union is Ω such that μ is zero on all measurable subsets of B while ν is zero on all measurable subsets of A.

- Let $\mathcal{R}$ be a ring on Ω and let μ be a measures on $\mathcal{R}$. Then there exists a measures μ' on the σ-algebra generated by $\mathcal{R}$ such that the restriction of μ' to $\mathcal{R}$ coincides with μ.
- A polished form of Lebesgue's decomposition theorem decomposes a singular measures into a singular continuous measures and a discrete measure.
- As a special case, a measures defined on the Euclidean space R_n is called singular if it is singular in respect to the Lebesgue measures on this space. For example, the Dirac delta function is a singular measure.
- The Radon-Nikodym theorem is a consequence in measures theory that states that given a measurable space (X, Σ), if a σ-finite measures ν on (X, Σ) is absolutely continuous with respect to a σ-finite measures μ on (X, Σ) then there is a measurable function f on X which takes values in $(0, \infty)$.
- The function f satisfies the above stated equality is uniquely defined up to a μ-null set. If g is an additional function which satisfies the same property then $f = g$ μ-almost everywhere (μ-ae). f is generally described as $dv/d\mu$ and is termed as the Radon-Nikodym derivative.
- In measures theory, Lebesgue's decomposition theorem states that for given μ and ν two σ-finite signed measures on a measurable space (Ω, Σ), there exist two σ-finite signed measures ν_0 and ν_1.
- The Lebesgue-Stieltjes integral is the ordinary Lebesgue integral with respect to a measures known as the Lebesgue-Stieltjes measures which can be combined to any function of bounded variation on the real line.
- Given two measurable spaces and measures on them we can obtain the product measurable space and the product measures on that specific space.
- Fubini's theorem is also sometimes termed as Tonelli's theorem. It is used to establish a connection between a multiple integral and a repeated one. If $f(x, y)$ is continuous on the rectangular region $R : a \le x \le b, c \le y \le d$, then the equality of the form $\iint_R f(x,y)d(x,y) = \int_a^b \int_c^d f(x,y)dy\,dx$ holds.

1.7 KEY TERMS

- **Signed measure:** Signed measure is referred to simplification of the concept of measures that allows it to have negative values. There are two different notions of a signed measures that depends on the condition that how the infinite values are taken.
- **Hahn decomposition theorem:** Hahn decomposition theorem states that for a given signed measures μ, there exist two measurable sets P and N. This decomposition is unique for adding/subtracting μ null sets from P and N.
- **Jordan decomposition theorem**: Jordan decomposition theorem states that every signed measures μ can be expressed as a difference of two positive measures μ^+ and μ^- of which at least one is finite.

NOTES

- **Jordon decomposition:** A decomposition of the measures ν as a difference of two mutually singular measures v^+ and v^- is termed as Jordon decomposition.
- **Mutually singular:** Two complex measures μ and ν on a measures space *X* are considered mutually singular if they are provided on different subsets.
- **Fubini's theorem:** Fubini's theorem is named after Guido Fubini. It is a consequence which provides conditions for which it is possible to compute a double integral using iterated integrals. As a consequence it also permits the order of integration to be changed in iterated integrals.

1.8 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is signed measure?
2. Why are Hahn and Jordan decomposition theorems used?
3. What is mutually singular measure?
4. Define Radon-Nikodym theorem.
5. State the drawbacks of Caratheodory extension theorem?
6. What is the significance of Lebesgue decomposition?
7. Define Lebesgue-Stieltjes integrals.
8. Specify the term product measure.
9. What does Fubini's theorem state?

Long-Answer Questions

1. Explain in detail the signed measures with the help of examples.
2. Discuss Hahn and Jordan decomposition theorems with the help of proof and examples.
3. Explain mutually singular measures with the help of appropriate examples.
4. Discuss Radon-Nikodym theorem with the help of proof.
5. Explain Lebesgue decomposition and its importance with reference to signed measures and decomposition.
6. Describe Riesz representation theorem with application.
7. Explain Lebesgue-Stieltjes integral with the help of examples.
8. Discuss product measures and Fubini's theorem with reference to signed measures and decomposition.
9. Brief a note on the importance of Caratheodary extension theorem with examples.

NOTES

1.9 FURTHER READING

Yang, Bicheng and Qiang Chen. 2016. *Integration Theory and Functional Analysis*. New York: Magnum Publishing LLC.

Kadets, Vladimir. 2018. *A Course in Functional Analysis and Measure Theory*. New York: Springer International Publishing.

Pundir, R. and S. K. Pundir. 2015. *Integration Theory and Functional Analysis*. Meerut (Uttar Pradesh): Pragati Prakashan.

Rudin, Walter. 1986. *Real and Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Ahlfors, Lars V. 1978. *Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Lang, Serge. 1998. *Complex Analysis*, 4th Edition. New York: Springer-Verlag New York Inc.

Shastri, Anant R. 2000. *An Introduction to Complex Analysis*. India: Macmillan Publishers India Ltd.

Gupta, S. L. and Nisha Rani. 2003. *Fundamental Real Analysis*, 4th Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, 1st Edition. United Kingdom: Cambridge University Press.

Basener, William F. *Topology and Its Applications*. New Jersey: John Wiley & Sons, Inc.

Bourbaki, Nicolas. 1966. *Elements of Mathematics: General Topology*, Part 1. United States: Addison-Wesley.

UNIT 2 BAIRE SETS AND BAIRE MEASURE

NOTES

Structure

2.0 INTRODUCTION

In mathematics, Baire set describes the specific relations between measure theory and topology. Particularly, Baire sets help to evaluate measures on non-metrizable topological spaces. The Baire sets form a subclass of the Borel sets. A subset of a compact Hausdorff topological space is termed as a Baire set if it is a member of the smallest σ–algebra which contains all compact G_δ sets.

Functions with compact support in X are those with support that is a compact subset of X. For example, if X is the real line, they are functions of bounded support and therefore vanish at infinity and negative infinity. Real valued compactly supported smooth functions on a Euclidean space are termed as bump functions.

The Riesz–Markov–Kakutani representation theorem relates linear functionals on spaces of continuous functions on a locally compact space to measures in measure theory. The theorem is named for Frigyes Riesz (1909) who introduced it for continuous functions on the unit interval, Andrey Markov (1938) who extended the result to some non-compact spaces, and Shizuo Kakutani (1941) who extended the result to compact Hausdorff spaces. There are many closely related variations of the theorem, as the linear functionals can be complex, real, or positive, the space they are defined on may be the unit interval or a compact space or a locally compact space, the continuous functions may be vanishing at infinity or have compact support, and the measures can be Baire measures or regular Borel measures or Radon measures or signed measures or complex measures.

In this unit, you will learn about the Baire sets and Baire measure, continuous functions with compact support, regularity of measures on locally compact spaces and Riesz-Markov theorem.

2.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Explain Baire sets and Baire measure
- Describe continuous functions with compact support
- State the regularity of measures on locally compact spaces
- Discuss Riesz-Markov theorem

2.2 INTRODUCTION TO BAIRE SETS AND BAIRE MEASURE

The Baire set describes the specific relations between measure theory and topology. Particularly, Baire sets help to evaluate measures on non-metrizable topological spaces. The Baire sets form a subclass of the Borel sets. A subset of a compact Hausdorff topological space is termed as a Baire set if it is a member of the smallest σ–algebra which contains all compact G_δ sets.

As per Dudley, a subset of a topological space X is termed as a Baire set if it belongs to the smallest σ–algebra for which all continuous functions defined on X into the real line are measurable. A discrete topological space is locally compact and Hausdorff. Therefore any function defined on a discrete space is continuous and as per Dudley all subsets of a discrete space are Baire.

Properties

The following properties hold for Baire sets:

- Baire sets correspond with Borel sets in every metric or metrizable space. Specifically they correspond in Euclidean spaces and all their subsets considered as topological spaces.
- For every compact Hausdorff space, every finite Baire measure, i.e., a measure on the σ-algebra of all Baire sets is regular.
- For every compact Hausdorff space, every finite Baire measure has a unique extension to a regular Borel measure.

In the descriptive set theory, a set of reals or subset of the Baire space or Cantor space is termed as universally Baire if it has a definite strong regularity property. Universally Baire sets are used in Ω-logic. A subset A of the Baire space is universally Baire if it has one of the following equivalent properties:

1. For every notion of forcing, there are trees T and U such that A is the projection of the set of all branches through T and it is forced that the projections of the branches through T and the branches through U are complements of each other.
2. For every compact Hausdorff space Ω and every continuous function f from Ω to the Baire space, the preimage of A under f has the property of Baire in Ω.
3. For every cardinal λ and every continuous function f from λ^ω to the Baire space, the preimage of A under f has the property of Baire.

Baire Measures

A Baire measure is a measure on the σ-algebra of Baire sets of a topological space. In spaces that are not metric spaces, the Borel sets and the Baire sets may differ. Baire measures can be used because they connect to the properties of continuous functions directly.

Definition: A measure μ on $(\mathbb{R}, \mathbb{B}_{\mathbb{R}})$ is a Baire measure (on $\mathbb{R}$), if μ (E) $< \infty$ whenever E is a bounded (Borel) set.

Assume that μ is a finite Baire measure. Then, define

$F(x) := \mu((-\infty, x])$

F is termed as the cumulative distribution function of μ. Now examine:

$\mu(a, b] = \mu(-\infty, b] - \mu(-\infty, a] = F(b) - F(a)$

Furthermore,

$$(a,b] = \bigcap_{1}^{\infty}(a,b,+\frac{1}{n}]$$

Consequently,

$$\mu(a, b] = F(b) - F(a) = \lim_{n\to\infty}\mu(a,b+\frac{1}{n}] = \lim_{n\to\infty}F(b+\frac{1}{n}) - F(a)$$

Accordingly,

$$F(b) = \lim_{n\to\infty}F(b+\frac{1}{n})$$

Hence, and we can conclude that F is right continuous.

$$\mu(\{b\}) = F(b) - F(\frac{1}{n})$$

$\mu\{b\} = F(b) = F(a)$

As a result, F is continuous at b if and only if $\mu(b) = 0$.

Proposition: Let μ be a finite Baire measure and let F be its cumulative distribution function. Then F is monotone increasing, bounded, right continuous and $\lim\limits_{x\to\infty} F(x) = 0$.

For a given cumulative distribution function that is increasing and right continuous we can construct a Baire measure. The following is the notation for this:

$$F(-\infty) = \lim_{x\to-\infty}F(x)$$

$$F(\infty) = \lim_{x\to\infty}F(x)$$

Theorem 2.1: If F is a monotone increasing function which is right continuous, then there exists a unique Baire measure μ such that $\mu(a, b] = F(b) - F(a)$.

Corollary: Every such F which is also bounded is the cumulative distribution function of a finite Baire measure, provided that $F(-\infty) = 0$. The Lebesgue-Stiltjes integral with respect to F can be defined with μ the appropriate measure as follows:

$$\int \phi dF = \int \phi d\mu$$

NOTES

Let μ be the counting measure on $\mathbb{N}$ σ-algebra is $\mathcal{P}(\mathbb{N})$ and all functions are measurable.

If $f : \mathbb{N} \to \mathbb{R}$ is non-negative then how the following will be evaluated,

$$\int_N f d\mu = \sum_{n=1}^{\infty} f(n)$$

A function $f : \mathbb{N} \to \mathbb{F}$ is integrable if and only if $|f|$ is,

$$\sum_{n=1}^{\infty} | f(n) | < \infty$$

By the discrete cosine transform, if a sum converges absolutely in n then we have the expression,

$$\lim_{n\to\infty} \sum_{m=1}^{\infty} x_{m,n} = \sum_{m=1}^{\infty} \lim_{n\to\infty} x_{m,n}$$

Check Your Progress

1. What is Baire set?
2. Give the properties of Baire set.
3. Define Baire measure.

2.3 CONTINUOUS FUNCTIONS WITH COMPACT SUPPORT

Functions with compact support in X are those with support that is a compact subset of X. For example, if X is the real line, they are functions of bounded support and therefore vanish at infinity and negative infinity. Real valued compactly supported smooth functions on a Euclidean space are termed as bump functions. The specific functions with compact support are considered dense in the space of functions that vanish at infinity. For more complex examples and in limits, for any $\varepsilon > 0$, any function f on the real line **R** that vanishes at infinity can be approximated by selecting an appropriate compact subset C of **R** such that,

$$| f(x) - I_C(x) f(x) | < \varepsilon$$

This holds for all $x \in X$, where I_C is termed as the indicator function of C. Every continuous function on a compact topological space has compact support since every closed subset of a compact space is indeed compact.

Let X be a separable and locally compact metric space, then for each compact set $K \subset X$ there is a continuous function with compact support and such that $f|K = 1$.

Definitely, $X = \bigcup_{n=1}^{\infty} U_n$, where $\{U_n\}$ is a increasing sequence of open and pre-compact subset of X (this follows from the Lindelöf theorem). Consequently there is an $m \in \mathbb{N}$, such that $K \subset U_m$.

2.3.1 Regularity of Measures on Locally Compact Spaces

Definition: A Borel measure on $(-a, b]$ is defined as the σ-additive measure defined on the Borel algebra of $(a, b]$ and is the smallest σ-algebra containing all the open subsets of $(a, b]$.

Theorem 2.2: Let μ be a finite Borel measure, μ $((a, t])$, less than ∞. Then, $F(t)$= μ $((a, t])$ is an increasing function such that, $\lim_{s \to t} F(s) = F(t)$.

On the other hand a Borel measure on $(a, b]$ is uniquely determined by every such function.

Proof: Let μ be a finite Borel measure. By σ-continuity we have, $\lim_{n\to\infty} \mu(\bigcup_n (a, b+s_n)\mu(a,b)$ for every decreasing sequence $(s_n) = 0$. For the reverse we will use the extension procedure. Let F be as defined above. Consider the algebra $A_{\mathbb{R}}$ generated by the intervals $(s, t]$. For such an interval define,

$$\mu F(s, t) = F(t) - F(s)$$

Now, we need to prove that,

$$\mu F(s, t) \le \sum_j \mu(B_j)$$

whenever $B_j \in A$ and $\cup_j B_j = (s, t]$ is a disjoint union. As every B_j is a finite union of intervals, assume as disjoint partition

$$(s, t) = \bigcup_j (s_j, t_j)$$

Let, ε > 0. For fixed j we can choose $\delta j > 0$, such that

$$F(t_j + \delta_j) - F(s_j) < F(t_j) - F(s_j) + 2^{-j}\varepsilon$$

Let, δ > 0. Then,

$$[s+\delta, t] \subset \bigcup_j (s_j, t_j + \delta_j)$$

By compactness we can find n_0, such that

$$[s+\delta, t] \subset \bigcup_{j=1}^{n} (s_j, t_j + \delta_j)$$

Now, reorganize the intervals such that $s_1 < s + \delta < t_1 + \delta_1$ and $s_2 < t_1 + \delta_1 < t_2 + \delta_2$, etc., so that $t < t_m + \delta_m$. This gives,

$$F(t) - F(s+\delta) \le \sum_{i=1}^{m} (F(t_i + \delta_i) - F(s_i)) \le \sum_j \mu F((s_j, t_j]) + 2^{-j}\varepsilon$$

$$\le \varepsilon + \sum_j \mu F((s_j, t_j])$$

NOTES

Since $\varepsilon > 0$ is arbitrary, we have

$$F(t) - F(s+\delta) \le \sum_j \mu F((s_j, t_j])$$

Using $\lim_{\delta \to 0} F(s+\delta) = F(s)$, we deduce

$$F(t) - F(s) \le \sum_j \mu F((s_j, t_j])$$

Find a σ-additive measure $\tilde{\mu} F$, by applying the Caratheodory extension procedure which satisfies,

$$\tilde{\mu} F((s,t]) = F(t) - F(s)$$

Here, the sets of meausre 0 are defined by the outer measure μ_F^* which is uniquely determined by the values $\tilde{\mu} F((s, t])$. Since the measure is finite, for every measurable set $E \subset (a, b]$ we can find $B \in A$ such that,

$$\mu_F^*(E \Delta B) < \varepsilon$$

This gives,

$$\mu_F(B) - \varepsilon < \tilde{\mu}_F(E) < \mu_F(B) + \varepsilon$$

Hence, the extension is uniquely determined from its values on intervals $(s, t]$.

In the following equations we suppose that $F : [a, b] \to \mathbb{R}$ is a monotone increasing function, continuous from the right. Here, we need the following four derivatives:

$$D^+ F(x) = \limsup_{h \downarrow 0} \frac{F(x+h) - F(x)}{h}$$

$$D^- F(x) = \limsup_{h \downarrow 0} \frac{F(x) - F(x-h)}{h}$$

$$D_+ F(x) = \liminf_{h \downarrow 0} \frac{F(x+h) - F(x)}{h}$$

$$D_- F(x) = \liminf_{h \downarrow 0} \frac{F(x) - F(x-h)}{h}$$

Note that

$$D^+ F(x) \ge D_+ F(x) \text{ and } D^- F(x) \ge D_- F(x)$$

We know that F is differentiable at x if all the values coincide and are finite.

Theorem 2.3: Let F be a monotone increasing function. Then F is differentiable almost everywhere. Furthermore,

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

Proof: Here, define $F(x) = F(b)$ for $x \ge b$. For example, consider

$$E_{u,v} = \{x : D^+F(x) > u > v > D\_F(x)\}$$

for rational u, v. We have to prove that $m*(E_{uv}) = 0$. Then we can conclude that,

NOTES

$D_+F(x) \le D^+F(x) \le D\_F(x) \le D^-F(x)$ almost everywhere. In the same way, $D^-F(x) \le D_+F$ and so, F is differentiable almost everywhere.

Now, we consider $s = m^*(E_{uv})$ and $\varepsilon > 0$. We can find an open set O such that $E_{u,v} \subset O$ and $m(O) < s + \varepsilon$. For each point $x \in E_{uv}$ there is a small interval $[x-h,\ x] \subset O$ such that,

$$F(x) - F(x-h) < vh$$

Now, we can find a finite disjoint collection $I_1, \ldots, I_N$ of such intervals such that,

$$m\left(\bigcup_{n=1}^{N} \text{Interior}(I_n)\right) > s - \varepsilon$$

Now, $I_n = [x_n - h_n, x_n]$. So, from disjointness

$$\sum_{n=1}^{N}(F(x_n) - F(x_n - h_n)) < v\sum_n h_n < vm(O) < v(s+\varepsilon)$$

Now define $A = \bigcup_{n=1}^{N} \text{Interior}(I_n) \cap E_{u,v}$. We know that every $y \in A$ is a left endpoint of an interval $(y, y+k)$ contained on some I_n such that,

$$F(y+k) - F(k) > uk$$

By applying the covering theorem, we obtain disjoint intervals $J_1, \ldots, J_M$ such that $\bigcup_i J_i$ contains a subset of A of measure greater than $s - 2\varepsilon$. Then,

$$\sum_{i=1}^{M}[F(y_k + k_i) - F(k_i)] > u\sum_i k_i > u(s - 2\varepsilon)$$

Note here that, every J_i is contained in some I_n. Hence, by monotonicity and disjointness we get,

$$\sum_{J_i \subset I_n}[F(y_k + k_i) - F(k_i)] \le F(x_n) - F(x_n - h_n)$$

Thus we have,

$$u(s-2\varepsilon) < \sum_{i=1}^{M}[F(y_k + k_i) - F(k_i)] \le \sum_{n=1}^{N}(F(x_n) - F(x_n - h_n)) \le v(x+\varepsilon)$$

Passing to the limit $\varepsilon \to 0$ we get $us \le vs$. Since $u > v$ we must have $s = 0$. In the following we may assume that,

$$g(x) = \lim_{h\to 0}\frac{F(x+h) - F(x)}{h} = F'(x)$$

exists almost everywhere. We define,

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

NOTES

Since, F is increasing we know that $f_n(x) \geq 0$. Since $\lim_n f_n(x) = F'(x)$ almost everywhere, we know that F' is measurable. By Fatou's lemma we find,

$$\int_a^b f \leq \liminf_n n\int_a^b [F(x+\frac{1}{n}) - F(x)]dx$$

$$= \liminf_n \left(n\int_b^{b+\frac{1}{n}} F - n\int_b^{a+\frac{1}{n}} F \right)$$

$$= \liminf_n \left(F(b) - n\int_b^{a+\frac{1}{n}} F \right)$$

$$= F(b) - \lim_n n\int_b^{a+\frac{1}{n}} F = F(b) - F(a)$$

Corollary 1: A function of bounded variation is differentiable almost everywhere.

Corollary 2: Let μ be a finite Borel measure on $(a, b]$. Then there exists an absolute continuous measure μ_n and singular measure μ_s such that,

$\mu = \mu_n + \mu_s$.

Proof: Consider $F(x) = \mu((a, x])$. Then, F is differentiable almost everywhere and we may define the absolute continuous measure

$$\mu_n((a,x]) = \int_a^x F'(x)dm$$

Then, $G(x) = F(x) - \int_a^x F'(x)dm$ is again an increasing function. The singular measure is determined by,

$$\mu_g((a,x]) = G(x)$$

Obviously, $G' = 0$ almost everywhere.

2.3.2 Measure and Outer Measure

Definition: An outer measure $\mu*$ is an extended real valued set function defined on all subsets of a space X having the following properties:

(a) $\mu*\phi = 0$

(b) $A \subset B \Rightarrow \mu*A \leq \mu*B$ (Monotonicity)

(c) $E \subset \sum_{i=1}^{\infty} E_i \Rightarrow \mu * E \leq \sum_{i=1}^{\infty} \mu * E_i$ (Subadditivity)

The outer measure $\mu*$ is said to be finite if $\mu*X < \infty$.

NOTES

By correspondence with the case of Lebesgue measure we say that a set E is measurable with respect to $\mu*$ if for every set A we have,

$$\mu * A = \mu * (A \cap E) + \mu * (A \cap E^C)$$

Since, $\mu*$ is subadditive, in order to show that E is measurable, we only need to prove that,

$$\mu * A \geq \mu * (A \cap E) + \mu * (A \cap E^C)$$

for every A.

When $\mu * A = \infty$, this inequality holds trivially. So, we only have to prove it for sets A with $\mu * A$ finite.

Theorem 2.4: The class β of $\mu*$-measurable sets are a σ-algebra. If $\bar{\mu}$ is restricted to β, then $\bar{\mu}$ is a complete measure on β.

Proof: It is clear that the empty set is measurable. From the symmetry of the definition of measurability in E and E^c, we have that E^c is measurable whenever E is measurable. Now, let E_1 and E_2 be measurable sets. From the measurability of E_2,

$$\mu * A = \mu * (A \cap E_2) + \mu * (A \cap E_2{}^C)$$

and by the measurability of E_1,

$$\mu * A = \mu * (A \cap E_2) + \mu * (A \cap E_2{}^C \cap E_1) + \mu * (A \cap E_2{}^C \cap E_2{}^C)$$

Now, since

$$A \cap [E_1 \cup E_2] = [A \cap E_2] \cup [A \cap E_1 \cap E_2{}^C]$$

we have,

$$\mu * (A \cap [E_1 \cup E_2]) \leq \mu * (A \cap E_2) + \mu * (A \cap E_2{}^C \cap E_1)$$

by subadditivity, and so

$$\mu * A \geq \mu * (A \cap [E_1 \cup E_2]) + \mu * (A \cap E_2{}^C \cap E_2{}^C)$$

This implies that $E_1 \cup E_1$ is measurable. So we get that the union of two measurable sets is measurable. But by induction, the union of any finite number of measurable sets is measurable. Hence, β is an algebra of sets. Suppose, $E = \cup E_i$, where $<E_i>$ is a disjoint sequence of measurable sets, and fix

$$G_n = \bigcup_{i=1}^{n} E_i$$

Then G_n is measurable, and

$$\mu * A = \mu * (A \cap G_n) + \mu * (A \cap G_n{}^C)$$

$$\geq \mu * (A \cap G_n) + \mu * (A \cap E^C)$$

because $E^c \subset G_n{}^c$.

NOTES

Now, $G_n \cap E_n = E_n$ and $G_n \cap E_n^{\,c} = G_{n-1}$, and by the measurability of E_n, we have

$$\mu^*(A \cap G_n) = \mu^*(A \cap E_n) + \mu^*(A \cap G_{n-1})$$

By induction (as above, $\mu^*(A \cap G_{n-1}) = \mu^*(A \cap E_{n+1}) + \mu^*(A \cap E_{n-2})$ and so on),

$$\mu^*(A \cap G_n) = \sum_{i=1}^{n} \mu^*(A \cap E_i)$$

and so,

$$\mu^* A \geq \mu^*(A \cap E^C) + \sum_{i=1}^{\infty} \mu^*(A \cap E_i)$$

$$\geq \mu^*(A \cap E^C) + \mu^*(A \cap E)$$

Since, $A \cap E \subset \bigcup_{i=1}^{\infty} (A \cap E_i)$

Thus, E is measurable.

Since, the union of any sequence of sets in an algebra can be replaced by a disjoint union of sets in an algebra, it follows that B is a σ-algebra.

Let us now prove that $\bar{\mu}$ is finitely additive. Let E_1 and E_2 be disjoint measurable sets.

Then, the measurability of E_2 implies that,

$$\bar{\mu}(E_1 \cup E_2) = \mu^*(E_1 \cup E_2)$$

$$= \bar{\mu}^*([E_1 \cup E_2] \cap E_2 + \mu^*([E_1 \cup E_2] \cap E_2^{\,C})$$

$$= \mu^* E_2 + \mu^* E_1$$

Finite additivity now follows by induction.

If E is the disjoint union of the measurable sets $\{E_i\}$, then

$$\bar{\mu}_i E \geq \bar{\mu}\left(\bigcup_{i=1}^{n} E_i \right) = \sum_{i=1}^{n} \bar{\mu} E_i$$

and so,

$$\bar{\mu}\, E \geq \sum_{i=1}^{\infty} \bar{\mu}\, E_i$$

But,

$\bar{\mu}\, E \leq \sum_{i=1}^{\infty} \bar{\mu}\, E_i$ by the subadditivity of μ*. Hence, $\bar{\mu}$ is countably additive.

So $\bar{\mu}$ is a measure since it is non negative and $\bar{\mu}\,\phi = \mu^*\phi = 0$.

2.3.3 Extension of a Measure

A measure on an algebra is a non negative extended real valued set function μ defined on an algebra ***A*** of sets such that,

(a) $\mu\phi = 0$

(b) If $<A_i>$ is a disjoint sequence of sets in ***A*** whose union is also in ***A***, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu A_i$$

Thus, a measure on an algebra ***A*** is a measure iff ***A*** is a σ-algebra.

We construct an outer measure μ* and show that the measure $\bar{\mu}$ is an extension of measure μ defined on an algebra. Define, $\mu{*}E = \inf \sum_{i=1}^{\infty} \mu A_i$, where $<A_i>$ ranges over all sequence from ***A*** such that

$$E \subset \bigcup_{i=1}^{\infty} A_i$$

Lemma 1: If $A \in$ ***A*** and if $<A_i>$ is any sequence of sets in ***A*** such that $A \subset \bigcup_{i=1}^{\infty} A_i$, then $\mu A \leq \sum_{i=1}^{\infty} \mu A_i$.

Proof: Fix, $B_n = A \cap A_n \cap A^C_{n-1} \cap ... \cap A_i^C$. Then $B_n \in$ ***A*** and $B_n \subset A_n$. But since A is the disjoint union of the sequence $<B_n>$, by countable additivity

$$\mu A = \sum_{n=1}^{\infty} \mu B_n \leq \sum_{n=1}^{\infty} \mu A_n$$

Corollary: If $A \in$ ***A***, $\mu{*}A = \mu A$.

Actually, from above, we have

$$\mu A \leq \sum_{n=1}^{\infty} \mu A_n < \mu{*}A + \varepsilon$$

or,

$$\mu A \leq \mu{*}A + \varepsilon$$

Now, as since ε is arbitrary, we have

$$\mu A \leq \mu{*}A$$

Also, by definition

$$\mu{*}A \leq \mu A$$

Therefore,

$$\mu{*}A = \mu A$$

NOTES

Lemma 2: The set function $\mu*$ is an outer measure.

Proof: We know that from the definition, $\mu*$ is a monotone non negative set function defined for all sets and $\mu*\phi = O$. Now it is only remained to prove that it is countably subadditive. Let $E \subseteq \bigcup_{i=1}^{\infty} E_i$. If $\mu*E_i = \infty$ for any i, then we have $\mu*E \leq \sum\mu*E_i = \infty$. If $\mu*E_i \neq \infty$, then given $\varepsilon > 0$, there exists for each i a sequence $< A_{ij} >_{j=1}^{\infty}$ of sets in $\boldsymbol{A}$ such that $E_i \subset \bigcup_{j=1}^{\infty} A_{ij}$ and

$$\sum_{j=1}^{\infty} \mu A_{ij} < \mu * E_i + \frac{\varepsilon}{2^i}$$

Then,

$$\mu * E \leq \sum_{i,j} \mu A_{ij} < \sum_{i=1}^{\infty} \mu * E_i + \varepsilon$$

Since ε is an arbitrary positive number, we have

$$\mu * E \leq \sum_{i=1}^{\infty} \mu * E_i$$

which proves that $\mu*$ is subadditive.

Lemma 3: If $A \in \boldsymbol{A}$, then A is measurable with respect to $\mu*$.

Proof: Suppose E be an arbitrary set of finite outer measure and ε be a positive number. Then there is a sequence $<A_i>$ from $\boldsymbol{A}$ such that $E \subset \cup A_i$ and $\Sigma\mu\, A_i < \mu * E + \varepsilon$.

By the additivity of μ on $\boldsymbol{A}$, we have

$$\mu(A_i) = \mu(A_i \cap A) + \mu(A_i \cap A^C)$$

Hence,

$$\mu * E + \varepsilon > \sum_{i=1}^{\infty} \mu(A_i \cap A) + \sum_{i=1}^{\infty} \mu(A_i \cap A^C)$$

$$> \mu * (E \cap A) + \mu * (E \cap A^C)$$

because

$$E \cap A \subset \cup(A_i \cap A)$$

and

$$E \cap A^C \subset \cup(A_i \cap A^C)$$

Since ε is an arbitrary positive number, we have

$$\mu * E \geq \mu * (\varepsilon \cap A) + \mu * (E \cap A^C)$$

and thus A is $\mu*$-measurable.

NOTES

Note: The outer measure $\mu*$ which we have defined above is known as the outer measure induced by μ.

Notation: For a given algebra $\boldsymbol{A}$ of sets we use $\boldsymbol{A}_\sigma$ to denote those sets which are countable unions of sets of $\boldsymbol{A}$ and use $\boldsymbol{A}_{\sigma\delta}$ to denote those sets which are countable intersection of sets in $\boldsymbol{A}_\sigma$.

Theorem 2.5: Let μ be a measure on an algebra $\boldsymbol{A}$, $\mu*$ be the outer measure induced by μ, and E be any set. Then for $\varepsilon > 0$, there exists a set $A \in \boldsymbol{A}_\sigma$ with $E \subset A$ and $\mu * A \le \mu * E + \varepsilon$.

There is also a set $B \in \boldsymbol{A}_{\sigma\delta}$ with $E \subset B$ and $\mu*E = \mu*B$.

Proof: From the definition of $\mu*$ there is a sequence $<A_i>$ from $\boldsymbol{A}$ such that $E \subset \cup A_i$ and

$$\sum_{i=1}^{\infty} \mu A_i \le \mu * E + \varepsilon \qquad \ldots (2.1)$$

Fix $A = \cup A_i$

Then, $\mu * A \le \Sigma \mu * A_i$

$$= \Sigma \mu A_i \qquad \ldots(2.2)$$

because $\mu*$and μ agree on members of $\boldsymbol{A}$ by the above mentioned corollary.

Hence, Equations (2.1) and (2.2) imply

$\mu * A \le \mu * E + \varepsilon$

which proves the first part.

To prove the second statement, we note that for each positive integer n there is a set A_n in $\boldsymbol{A}_\sigma$, such that, $E \subset A_n$ and

$\mu * A_n < \mu * E + \dfrac{1}{n}$ (From first part proved above)

Let $B = \cap A_n$. Then, $B \in \boldsymbol{A}_{\sigma\delta}$ and $E \subset B$. Since $B \subset A_n$,

$$\mu * B < \mu * A_n \le \mu * E + \frac{1}{n}$$

Since n is arbitrary, by monotonicity, $\mu*B \le \mu*E$. Hence $\mu*B = \mu*E$.

2.3.4 Riesz-Markov Theorem

Consider that X denotes a locally compact Hausdorff space. Let f be a real valued continuous function on X. The support of f is the subset of the form,

Supp (f) : = Closure of $\{x \in \mathrm{X} : f(x) \ne 0\}$

Here f has compact support if the support of f is a compact subset on X. f is zero outside a compact set. A linear functional or linear form, also termed as a one-form or covector, is a linear map from a vector space to its field of scalars. In $\mathbf{R}^n$, if vectors are represented as column vectors then linear functionals are represented as row vectors and their action on vectors is specified by the dot product or the matrix product with the row vector on the left and the column vector on the right. Generally, if V is a vector space over a field k then a linear functional f is a function from V to k, which is linear,

NOTES

$$f(\vec{v}+\vec{w}) = f(\vec{v}) + f(\vec{w}) \text{ for all } \vec{v}, \vec{w} \in V$$

$$f(a\vec{v}) = af(\vec{v}) \text{ for all } \vec{v} \in V, a \in k.$$

The set of all linear functionals from V to k, $\mathrm{Hom}_k(V,k)$, is itself a vector space over k. This space is called the dual space of V or sometimes the algebraic dual space. It is written as V^* or V' when the field k is implicit. Every non-degenerate bilinear form on a finite dimensional vector space V gives an isomorphism from V to V^*. Specifically, the bilinear form on V is denoted by $<, >$ and there is a natural isomorphism $V \to V^* : v \mapsto v^*$ given by the following expression:

$$v^*(w) := \langle v, w \rangle.$$

The inverse isomorphism is given by $V^* \to V : f \mapsto f^*$ where f^* is the unique element of V for which for all $w \in V$. The vector is then,

$$\langle f^*, w \rangle = f(w).$$

The above defined vector $v^* \in V^*$ is said to be the dual vector of $v \in V$.

In an infinite dimensional Hilbert space, equivalent results hold by the Riesz representation theorem. There is a mapping $V \to V^*$ into the continuous dual space V^*.

Riesz Representation Theorem for Linear Functionals on $C_c(X)$

The following theorem represents positive linear functionals on $C_c(X)$, i.e., the space of continuous compactly supported complex valued functions on a locally compact Hausdorff space X. The Borel sets in the given statement refer to the σ-algebra produced by the open sets.

A non-negative countably additive Borel measure μ on a locally compact Hausdorff space X is regular if and only if,

- $\mu(K) < \infty$ for every compact K.
- For every Borel set E, $\mu(E) = \inf \{\mu(U) : E \subseteq U, U \text{ open}\}$
- The relation $\mu(E) = \sup (\mu(K) : K \subseteq E, K \text{ compact})$ holds whenever E is open or when E is Borel and $\mu(E) < \infty$.

Theorem 2.6: Let X be a locally compact Hausdorff space. For any positive linear functional ψ on $C_c(X)$, there is a unique Borel regular measure μ on X such that,

$$\psi(f) = \int_X f(x)\, d\mu(x)$$

This holds for all f in $C_c(X)$.

One method to measure theory is to initiate with a Radon measure defined as a positive linear functional on $C(X)$. In its original form by F. Riesz (1909) the

NOTES

theorem states that every continuous linear functional $A[f]$ over the space $C[0,1]$ of continuous functions in the interval [0,1] can be represented in the form,

$$A(f) = \int_0^1 f(x)\, d\alpha(x).$$

Here $\alpha(x)$ is a function of bounded variation on the interval [0,1] and the integral is a Riemann-Stieltjes integral. Since there is a one-to-one association between Borel regular measures in the interval and functions of bounded variation consequently the above stated theorem generalizes the original statement of F. Riesz. He assigns to each function of bounded variation the consequent Lebesgue-Stieltjes measure and the integral with respect to the Lebesgue-Stieltjes measure agrees with the Riemann-Stieltjes integral for continuous functions.

Riesz-Markov Representation Theorem for the Dual of $C_0(X)$

The Riesz-Markov theorem gives a concrete realization of the dual space of $C_0(X)$, i.e., the set of continuous functions on X which vanish at infinity. The Borel sets in the statement of the theorem also refer to the σ-algebra generated by the open sets.

If μ is a complex valued countably additive Borel measure, then μ is regular iff the non-negative countably additive measure $|\mu|$ is regular as defined above.

Theorem 2.7: Let X be a locally compact Hausdorff space. For any continuous linear functional ψ on $C_0(X)$, then there is a unique regular countably additive complex Borel measure μ on X such that,

$$\psi(f)\int_X f(x)\, d\mu(x)$$

This holds for all f in $C_0(X)$. The norm of ψ as a linear functional is the total variation of μ is given as follows,

$$\|\psi\| = |\mu|(X).$$

To conclude, ψ is positive iff the measure μ is non-negative.

2.4 Integration of Continuous Functions with Compact Support

Let ν be a k-form on $\mathbb{R}^n$. We define the support of to be the closure of the set

$$\{x \in \mathbb{R}^n, \nu_x \neq 0\}$$

and we say that ν is compactly supported if $\operatorname{supp}\nu$ is compact. We will denote by $\Omega_c^k(\mathbb{R}^n)$ the set of all C^∞ k-forms which are compactly supported, and if U is an open subset of $\mathbb{R}^n$, we will denote by $\Omega_c^k(U)$ the set of all compactly supported k-forms whose support is contained in .

Let $\omega = f dx_1 \wedge \cdots \wedge dx_n$ be a compactly supported n-form with $f \in C_0^\infty(\mathbb{R}^n)$. We will define the integral of ω over $\mathbb{R}^n$:

$$\int_{\mathbb{R}^n} \omega$$

to be the usual integral of f over $\mathbb{R}^n$

$$\int_{\mathbb{R}^n} f\, dx.$$

NOTES

(Since f is C^∞ and compactly supported this integral is well-defined.) Now let Q be the rectangle

$$[a_1, b_1] \times \cdots \times [a_n, b_n].$$

Theorem 2.8

Let ω be a compactly supported *n*-form, with supp $\omega \subseteq$ Int *Q*. Then the following assertions are equivalent:

1. $\int \omega = 0$.
2. There exists a compactly supported ($n-1$)-form, μ, with $\text{supp}\mu \subseteq$ Int Q satisfying $d\mu = \omega$.

   We will first prove that (2.4) ⇒ (2.3). Let

$$\mu = \sum_{i=1}^{n} f_i\, dx_1 \wedge \ldots \wedge \hat{dx_i} \wedge \ldots \wedge dx_n$$

(the "hat" over the dx_i meaning that has to be omitted from the wedge product). Then

$$d\mu = \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial f_i}{\partial x_i} dx_1 \wedge \ldots \wedge dx_n,$$

and to show that the integral of $d\mu$ is zero it suffices to show that each of the integrals

$$\int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i} dx \qquad \text{...(2.3)}$$

is zero. By Fubini we can compute (2.3) by first integrating with respect to the variable, x_i , and then with respect to the remaining variables. But

$$\int \frac{\partial f}{\partial x_i} dx_i = f(x)\Big|_{x_i=a_i}^{x_i=b_i} = 0$$

since f_i is supported on *U*.

We will prove that (2.3) ⇒ (2.4) by proving a somewhat stronger result. Let *U* be an open subset of $\mathbb{R}^m$. We'll say that *U* has property *p* if every form, $\omega \in \Omega_c^m(U)$ whose integral is zero in $d\Omega_c^{m-1}(U)$.

Theorem 2.9

Let *U* be an opsen subset of $\mathbb{R}^{n-1}$ and $A \subseteq \mathbb{R}$ an open interval. Then if has property does as well.

Remark

It is simple to see that the open interval A itself has property p. Hence it follows by induction from Theorem 2.8 that

$$\mathrm{Int}Q = A_1 \times \cdots \times A_n, \quad A_i = (a_i, b_i)$$

NOTES

has property p, and this proves "$(2.3) \Rightarrow (2.4)$".

To prove Theorem 2.9 let $(x,t) = (x_1, \ldots, x_{n-1}, t)$ be product coordinates on $U \times A$. Given $\omega \in \Omega_c^n(U \times A)$ we can express ω as a wedge product, $dt \wedge \alpha$ with $\alpha = f(x,t)dx_1 \wedge \cdots \wedge dx_{n-1}$ and $f \in C_0^\infty(U \times A)$. Let $\theta \in \Omega_c^{n-1}(U)$ be the form

$$\theta = \left(\int_A f(x,t)dt\right)dx_1 \wedge \cdots \qquad \text{...(2.4)}$$

Then

$$\int_{\mathbb{R}^{n-1}} \theta = \int_{\mathbb{R}^n} f(x,t)dxdt = \int_{\mathbb{R}^n} \omega$$

so if the integral of ω is zero, the integral of θ is zero. Hence since U has property $P, \beta = dv$ for some $v \in \Omega_c^{n-1}(U)$. Let $\rho \in C^\infty(\mathbb{R})$ be a bump function which is supported on A and whose integral over is one. Setting

$$\kappa = -\rho(t)dt \wedge v$$

we have

$$d\kappa = \rho(t)dt \wedge dv = \rho(t)dt \wedge \theta,$$

and hence

$$\omega - d\kappa = dt \wedge (\alpha - \rho(t)\theta) = dt \wedge u(x,t)dx_1 \wedge \cdots \wedge dx_{n-1}$$

where

$$u(x,t) = f(x,t) - \rho(t)\int_A f(x,t)dt$$

by (2.3). Thus

$$\int u(x,t)dt = 0.$$

Let a and b be the end points of A and let

$$v(x,t) = \int_a^t i(x,s)ds. \qquad \text{…(2.5)}$$

By (2.4) $v(a,x) = v(b,x) = 0$, so v is in $C_0^\infty(U \times A)$ and by (2.5), $\partial v/\partial t = u$. Hence if we let γ be the form, $v(x,t)dx_1 \wedge \cdots \wedge dx_{n-1}$, we have:

$$d\gamma = u(x,t)dx \wedge \cdots \wedge dx_{n-1} = \omega - d\kappa$$

and

$$\omega = d(\gamma + \kappa).$$

Since γ and k are both in $\Omega_c^{n-1}(U \times A)$ this proves that ωis in $d\Omega_c^{n-1}(U \times A)$ and hence that $U \times A$ has property P.

NOTES

Check Your Progress

4. What are continuous functions with compact support?
5. What is Borel measure?
6. State a condition for an outer measure to be finite.
7. Why is Riesz-Markov theorem used?
8. What does Riesz-Markov representation theorem for the dual of $C_0(X)$ state?

2.5 ANSWERS TO 'CHECK YOUR PROGRESS'

1. The Baire sets form a subclass of the Borel sets. A subset of a compact Hausdorff topological space is termed as a Baire set if it is a member of the smallest σ–algebra which contains all compact G_δ sets.
2. Baire sets correspond with Borel sets in every metric or metrizable space. Specifically they correspond in Euclidean spaces and all their subsets considered as topological spaces. For every compact Hausdorff space, every finite Baire measure, i.e., a measure on the σ-algebra of all Baire sets is regular.
3. A measure μ on $(\mathbb{R}, \mathbb{B}_{\mathbb{R}})$ is a Baire measure (on $\mathbb{R}$), if μ (E) $< \infty$ whenever E is a bounded (Borel) set.
4. Functions with compact support in X are those with support that is a compact subset of X. For example, if X is the real line, they are functions of bounded support and therefore vanish at infinity and negative infinity. Real valued compactly supported smooth functions on a Euclidean space are termed as bump functions.
5. A Borel measure on $(-a, b]$ is defined as the σ-additive measure defined on the Borel algebra of $(a, b]$ and is the smallest σ-algebra containing all the open subsets of $(a, b]$.
6. The outer measure $\mu*$ is said to be finite if $\mu*X < \infty$.
7. The Riesz-Markov theorem gives a concrete realization of the dual space of $C_0(X)$, i.e., the set of continuous functions on X which vanish at infinity. The Borel sets in the statement of the theorem also refer to the σ-algebra generated by the open sets.
8. The Riesz-Markov theorem gives a concrete realization of the dual space of $C_0(X)$, i.e., the set of continuous functions on X which vanish at infinity. The Borel sets in the statement of the theorem also refer to the σ-algebra generated by the open sets.

2.6 SUMMARY

NOTES

- The Baire set describes the specific relations between measure theory and topology. Particularly, Baire sets help to evaluate measures on non-metrizable topological spaces.
- The Baire sets form a subclass of the Borel sets. A subset of a compact Hausdorff topological space is termed as a Baire set if it is a member of the smallest σ–algebra which contains all compact G_δ sets.
- Baire sets correspond with Borel sets in every metric or metrizable space. Specifically they correspond in Euclidean spaces and all their subsets considered as topological spaces.
- For every compact Hausdorff space, every finite Baire measure, i.e., a measure on the σ-algebra of all Baire sets is regular.
- For every compact Hausdorff space, every finite Baire measure has a unique extension to a regular Borel measure.
- A measure μ on $(\mathbb{R}, \mathbb{B}_{\mathbb{R}})$ is a Baire measure (on $\mathbb{R}$), if $\mu\,(E) < \infty$ whenever E is a bounded (Borel) set.
- Let μ be a finite Baire measure and let F be its cumulative distribution function. Then F is monotone increasing, bounded, right continuous and $\lim\limits_{x \to \infty} F(x) = 0$.
- If F is a monotone increasing function which is right continuous, then there exists a unique Baire measure μ such that $\mu(a, b] = F(b) - F(a)$.
- Functions with compact support in X are those with support that is a compact subset of X. For example, if X is the real line, they are functions of bounded support and therefore vanish at infinity and negative infinity. Real valued compactly supported smooth functions on a Euclidean space are termed as bump functions.
- A Borel measure on $(-a, b]$ is defined as the σ-additive measure defined on the Borel algebra of $(a, b]$ and is the smallest σ-algebra containing all the open subsets of $(a, b]$.
- The class β of $\mu*$-measurable sets are a σ-algebra. If $\bar{\mu}$ is restricted to β, then $\bar{\mu}$ is a complete measure on β.
- In R^n, if vectors are represented as column vectors then linear functionals are represented as row vectors and their action on vectors is specified by the dot product or the matrix product with the row vector on the left and the column vector on the right.
- The Riesz-Markov theorem gives a concrete realization of the dual space of $C_0(X)$, i.e., the set of continuous functions on X which vanish at infinity. The Borel sets in the statement of the theorem also refer to the σ-algebra generated by the open sets.

2.7 KEY TERMS

- **Baire set:** The Baire sets form a subclass of the Borel sets. A subset of a compact Hausdorff topological space is termed as a Baire set if it is a member of the smallest σ–algebra which contains all compact G_δ sets.
- **Borel measure:** A Borel measure on $(-a, b)$ is defined as the σ-additive measure defined on the Borel algebra of (a, b) and is the smallest σ-algebra containing all the open subsets of (a, b).
- **Riesz representation theorem:** The following theorem represents positive linear functionals on $C_c(X)$, i.e., the space of continuous compactly supported complex valued functions on a locally compact Hausdorff space X. The Borel sets in the given statement refer to the σ-algebra produced by the open sets.

2.8 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define the terms Baire sets and Baire measures.
2. What do you mean by continuous functions with compact support?
3. Define regularity of measures on locally compact spaces.
4. How will you define the regularity of measure on locally compact spaces?
5. What is the difference between measure and outer measure?
6. State extension of a measure.
7. What does Riesz-Markov theorem state?

Long-Answer Questions

1. Explain the importance and applications of Baire sets and Baire measure in signed measures and decomposition.
2. Describe and prove the uniqueness of continuous functions and compact support.
3. Describe the concept of regularity of measure on locally compact spaces.
4. Discuss the differences among measure, outer measure, extension of a measure and measure space.
5. Discuss Riesz-Markov theorem in signed measures and decomposition.
6. Illustrate the applications of signed measures and decomposition.

2.9 FURTHER READING

NOTES

Yang, Bicheng and Qiang Chen. 2016. *Integration Theory and Functional Analysis*. New York: Magnum Publishing LLC.

Kadets, Vladimir. 2018. *A Course in Functional Analysis and Measure Theory*. New York: Springer International Publishing.

Pundir, R. and S. K. Pundir. 2015. *Integration Theory and Functional Analysis*. Meerut (Uttar Pradesh): Pragati Prakashan.

Rudin, Walter. 1986. *Real and Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Ahlfors, Lars V. 1978. *Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Lang, Serge. 1998. *Complex Analysis*, 4th Edition. New York: Springer-Verlag New York Inc.

Shastri, Anant R. 2000. *An Introduction to Complex Analysis*. India: Macmillan Publishers India Ltd.

Gupta, S. L. and Nisha Rani. 2003. *Fundamental Real Analysis*, 4th Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, 1st Edition. United Kingdom: Cambridge University Press.

Basener, William F. *Topology and Its Applications*. New Jersey: John Wiley & Sons, Inc.

Bourbaki, Nicolas. 1966. *Elements of Mathematics: General Topology*, Part 1. United States: Addison-Wesley.

UNIT 3 NORMED LINEAR SPACES

NOTES

Structure

3.0 INTRODUCTION

In mathematics, a linear space is a basic structure in incidence geometry. It consists of a family of subsets of a set such that the intersection of two subsets contains at most one element of the set. The elements of the set are called points and the subsets are called lines. Linear spaces can be seen as a generalization of projective and affine planes. The term linear space was coined by Libois in 1964. Linear transformations are the transformations that can be represented by matrices. Vector spaces stem from affine geometry, through the introduction of coordinates in the plane or three-dimensional space. The foundation of the definition of vectors was Bellavitis notion of the bipoint, an oriented segment one of whose ends is the origin and the other one a target. Vectors are elements in $\mathbf{R}^2$, $\mathbf{R}^4$, etc., and are used in systems of linear equations. An important development of vector spaces is due to the construction of function spaces by Lebesgue. This was later formalized by Banach and Hilbert, around 1920 and was used for evaluating algebra and the field of functional analysis. The 2- or 3-dimensional vectors are defined through real valued entries and the 'Length' of a vector can easily be extended to any real vector space $\mathbf{R}^n$.

The Hahn-Banach theorem is an essential tool in functional analysis. It permits the extension of bounded linear functionals defined on a subspace of some vector space to the complete space and also illustrates that there are 'Enough' continuous linear functionals defined on every normed vector space for studying the dual space. It is named for Hans Hahn and Stefan Banach who proved this theorem independently and a general extension theorem from which the Hahn-Banach

NOTES

theorem can be derived, which was proved in 1923 by Marcel Riesz. A Banach space is a complete normed vector space or a Banach space is a vector space which is equipped with a norm and which is complete with respect to that norm. Two common types of Banach spaces are real Banach spaces and complex Banach spaces, which are Banach spaces whose underlying vector spaces are defined over the field of real numbers or complex numbers, respectively. Various infinite dimensional function spaces evaluated in analysis are Banach spaces, including spaces of continuous functions (continuous functions on a compact Hausdorff space), spaces of Lebesgue integrable functions known as L^p spaces and spaces of holomorphic functions known as Hardy spaces. These are the most commonly used topological vector spaces and their topology is based on a norm.

In this unit, you will learn about the normed linear spaces, Banach spaces, conjugate spaces, nature imbedding of a normal linear space in its second dual and uniform boundedness principle and its consequences.

3.1 OBJECTIVES

After going through this unit, you will be able to:

- Explain the meaning of normed linear spaces
- Discuss the significance of linear transformation
- Describe linear transformation and dual spaces
- Explain quotient spaces
- Elaborate on the Banach spaces and completeness of l_p, L^p, R^n, C^n and C [a, b]
- Discuss about the conjugate spaces
- Know about the natural imbedding of a normed linear space in its second dual
- State uniform boundedness principle, open mapping theorem and closed graph theorem

3.2 NORMED LINEAR SPACES

A normed linear space is a vector space X and a non-negative valued mapping $\|\cdot\|$ on X termed as the **norm**, which satisfies the following properties:

1. $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$.
2. $\|\mathbf{a}\ \mathbf{x}\| = |\mathbf{a}|\ \|\mathbf{x}\|$, for all scalars $\mathbf{a}$.
3. $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

Here $\|\mathbf{x}\|$ is considered as the length of **x** or the distance from **x** to **0**. For a given vector **x**, if **y** is defined as $(1/\|\mathbf{x}\|)\ \mathbf{x}$, then **y** has unit length and is called the normalized vector for **x**.

The 2- or 3-dimensional vectors are defined through real valued entries and the 'Length' of a vector can easily be extended to any real vector space $\mathbf{R}^n$. The following properties of vector length are essential:

1. The zero vector '**0**' has zero length whereas every other vector has a positive length.

$$\| x \| > 0 \text{ if } x \neq 0$$

2. Multiplying a vector by a positive number changes its length without changing its direction. Moreover,

$\| \alpha x \| = |\alpha| \ \| x \|$ for any scalar α.

3. The triangle inequality holds, i.e., taking norms as distances, the distance from point A through B to C is never shorter than going directly from A to C or the shortest distance between any two points is a straight line.

$\|x+y\| \leq \|x\| + \|y\|$ for any vectors **x** and **y**. (By triangle inequality)

NOTES

The generalization of these three properties shows the ways to the notion of norm. A vector space on which a norm is defined is then called a normed vector space. Normed vector spaces are essential to study linear algebra and functional analysis.

A seminormed vector space is a pair (V, p) where V is a vector space and p a seminorm on V.

A normed vector space is a pair (V, $\cdot$) where V is a vector space and $\cdot$ a norm on V.

A vector norm can be taken as any real valued function that satisfies all the three properties. Properties 1 and 2 together imply that,

$\| x \| = 0$ if and only if $x = 0$.

A functional variation of the triangle inequality is given as,

$\| x - y \| \geq | \, \| x \| - \| y \| \, |$ for any vectors x and y.

This also illustrates that a vector norm is a continuous function.

3.2.1 Linear Transformation

By a Linear Transformation ($L.T.$) we mean a map $T : V \to W$, such that or such that, $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ where $x, y \in V$, $\alpha, \beta \in F$ and V, W are vector spaces over the field F. Also, we will be dealing with vector spaces that are finite dimensional, unless mentioned otherwise.

Theorem 3.1: *A L.T.* $T : V \to V$ is one-one iff T is onto.

Proof: Let $T : V \to V$ be one-one. Let dim $V = n$.

Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of V, then $\{T(v_1), \ldots, T(v_n)\}$ will also be a basis of V as

$$\alpha_1 T(v_1) + \alpha_2 T(v_2) + \ldots + \alpha_n T(v_n) = 0$$

$$\Rightarrow \quad T(\alpha_1 v_1 + \ldots + \alpha_n v_n) = T(0) \quad (T \text{ a } L.T.)$$

$$\Rightarrow \quad \alpha_1 v_1 + \ldots + \alpha_n v_n = 0 \quad (T \text{ is } 1\text{-}1)$$

$$\Rightarrow \quad \alpha_i = 0 \text{ for all } i$$

NOTES

thus $T(v_1)$, ... $T(v_n)$ are *L.I.* and as dim $V = n$ the result follows.

Let now $v \in V$ be any element

$$\text{then} \quad v = a_1T(v_1) + a_2T(v_2) + ... + a_nT(v_n) \quad a_i \in F$$
$$= T(a_1v_1 + ... + a_nv_n)$$
$$= T(v') \text{ for some } v'$$

Hence T is onto.

Conversely, let T be onto.

Here again we show that if $\{v_1, v_2, ... v_n\}$ is a basis of V then so also is $\{T(v_1), T(v_2), ..., T(v_n)\}$

For any $v \in V$, since T is onto, $\exists$ some $v' \in V$ such that,

$$T(v') = v$$

Again $v' \in V$ means $v' = \sum \alpha_i v_i \quad \alpha_i \in F$

$$\therefore \quad v = T(v') = T(\sum \alpha_i v_i) = \sum \alpha_i T(v_i)$$

$$\Rightarrow \quad T(v_1), T(v_2), ..., T(v_n) \text{ span } V$$

and as dim $V = n$, $\{T(v_1), ..., T(v_n)\}$ forms a basis of V.

Now if $v \in \text{Ker } T$ be any element

$$\text{then} \quad T(v) = 0$$
$$\Rightarrow \quad T(\Sigma\, \alpha_i v_i) = 0$$
$$\Rightarrow \quad \Sigma\, \alpha_i T(v_i) = 0$$
$$\Rightarrow \quad \alpha_i = 0 \text{ for all } i \text{ as } T(v_1), ..., T(v_n) \text{ are } L.I.$$
$$\Rightarrow \quad v = \Sigma \alpha_i v_i = 0$$
$$\Rightarrow \quad \text{Ker } T = \{0\} \Rightarrow T \text{ is 1-1.}$$

Theorem 3.2: Let V and W be two vector spaces over F. Let $\{v_1, v_2, ..., v_n\}$ be a basis of V and $w_1, w_2, ..., w_n$ be any vectors in W (not essentially distinct). Then there exists a unique *L.T*

$$T : V \rightarrow W \text{ such that, } T(v_i) = w_i \quad i = 1, 2, ..., n.$$

Proof: Let $v \in V$ be any element, then $v = \sum_{i=1}^{n} \alpha_i v_i$, $\quad \alpha_i \in F$

as $\{v_1, v_2, ..., v_n\}$ is a basis of V.

$$\text{Define} \quad T : V \rightarrow W \text{ s.t.,}$$
$$T(v) = \Sigma\, \alpha_i w_i$$

Then T is a linear transformation (verify!).

Clearly here, $T(v_i) = T(ov_1 + ... + 1 \,.\, v_i + ... + ov_n) = 1w_i$ for all i

To show uniqueness let T' be any other *L.T.* from $V \rightarrow W$ such that,

$$T'(v_i) = w_i$$

Let $v \in V$ be any element, then $v = \Sigma\, \alpha_i v_i$

$$T'(v) = T'(\Sigma\, \alpha_i v_i) = \Sigma\, \alpha_i T'(v_i) = \Sigma\, \alpha_i w_i = T(v)$$

Hence $\qquad T' = T.$

Thus we notice that a linear transformation is completely determined by its values on the elements of a basis.

Definition: Let $T : V \to W$ be a $L.T.$

then we define $\quad$ Rank of T = dim Range $T = r(T)$

Nullity of T = dim Ker $T = \nu(T)$.

Theorem 3.3: (Sylvester's Law) : Let $T : V \to W$ be a $L.T.$, then

$$Rank\ T + Nullity\ T = dim\ V.$$

Proof: Let $\{x_1, x_2, ..., x_m\}$ be a basis of Ker T then $\{x_1, x_2, ..., x_m\}$ being $L.I.$ in Ker T will be $L.I.$ in V. Thus it can be extended to form a basis of V.

Let $\{x_1, x_2, ..., x_m, v_1, v_2, ..., v_n\}$ be the extended basis of V.

Then $\quad$ dim Ker T = nullity of $T = m$

$\quad$ dim $V = m + n$

we show $\{T(v_1), T(v_2), ..., T(v_n)\}$ is a basis of Range T

Now $\quad \alpha_1 T(v_1) + \alpha_2 T(v_2) + ... + \alpha_n T(v_n) = 0$

$\Rightarrow \quad T(\alpha_1 v_1 + ... + \alpha_n v_n) = 0$

$\Rightarrow \quad \alpha_1 v_1 + \alpha_2 v_2 ... + \alpha_n v_n \in$ Ker T

$\Rightarrow \quad \alpha_1 v_1 + ... + \alpha_n v_n = \beta_1 x_1 + ... + \beta_m x_m$

or $\quad \alpha_1 v_1 + ... + \alpha_n v_n + (-\beta_1)x_1 + ... + (-\beta_m)x_m = 0$

$\Rightarrow \quad \alpha_1 = \alpha_2 = ... = \beta_1 = ... = \beta_m = 0$

$\Rightarrow \quad \alpha_i = 0$ for all i

i.e., $\quad \{T(v_1), T(v_2), ..., T(v_n)\}$ is $L.I.$

Now if $T(v) \in$ Range T be any element then as $v \in V$

$$v = a_1 x_1 + ... + a_m x_m + b_1 v_1 + ... + b_n v_n \quad a_i, b_j \in F$$

$$\therefore \quad T(v) = a_1 T(x_1) + ... + a_m T(x_m) + b_1 T(v_1) + ... + b_n T(v_n)$$

$$= 0 + ... + 0 + b_1 T(v_1) + ... + b_n T(v_n) \text{ [as } x_i \in \text{Ker } T]$$

or that $T(v)$ is a linear combination of $T(v_1), ..., T(v_n)$

which, therefore, form a basis of Range T.

$\therefore$ dim Range $T = n$ = rank T

which proves the theorem.

Theorem 3.4: If $T : V \to V$ be a $L.T.$ Show that the following statements are equivalent.

(*i*) Range $T \cap$ Ker $T = \{0\}$

(*ii*) If $T(T(v)) = 0$ then $T(v) = 0$, $v \in V$

Proof: $(i) \Rightarrow (ii)$

$$T(T(v)) = 0 \Rightarrow T(v) \in \text{Ker } T$$

Also $\quad T(v) \in$ Range $T \quad$ (by definition)

$\Rightarrow \quad T(v) = 0$

NOTES

$(ii) \Rightarrow (i)$

Let $x \in$ Range $T \cap$ Ker T

$\Rightarrow$ $x \in$ Range T and $x \in$ Ker T

$\Rightarrow$ $x = T(v)$ for some $v \in V$

and $T(x) = 0$

$$x = T(v) \Rightarrow T(x) = T(T(v))$$
$$\Rightarrow 0 = T(T(v))$$
$$\Rightarrow T(v) = 0 \quad \text{(given condition)}$$
$$\Rightarrow v = 0.$$

Algebra of Linear Transformations

Let V and W be two vector spaces over the same field F. Let $T : V \to W$ and $S : V \to W$ be two linear transformations. We define $T + S$, the sum of T and S by

$$T + S : V \to W, \text{ such that,}$$
$$(T + S)v = T(v) + S(v), \quad v \in V$$

Then $T + S$ is also a $L.T.$ from $V \to W$ as

$$(T + S)(\alpha x + \beta y) = T(\alpha x + \beta y) + S(\alpha x + \beta y)$$
$$= \alpha T(x) + \beta T(y) + \alpha S(x) + \beta S(y)$$
$$= \alpha(T + S)x + \beta(T + S)y$$

Again for $\alpha \in F$, we define the product of a $L.T.$ $T : V \to W$ with α, by $(\alpha T) : V \to W$ such that, $(\alpha T)v = \alpha(T(v))$.

It is easy to see that αT is a also a $L.T.$ from $V \to W$. Let Hom (V, W) be the set of all linear transformations from $V \to W$. Then we show Hom (V, W) forms a vector space over F under the addition and scalar multiplication as defined above.

We have already seen that when $T, S \in$ Hom (V, W), $\alpha \in F$ then $T + S$, $\alpha T \in$ Hom (V, W), thus closure holds for these operations. We verify some of the other conditions in the definition.

$$(T + S)v = T(v) + S(v) = S(v) + T(v) = (S + T)v \text{ for all } v \in V$$
$$\Rightarrow \quad T + S = S + T \text{ for all } S, T \in \text{Hom}(V, W)$$

The map $O : V \to W$, such that, $O(v) = 0$ is a $L.T.$ and

$$(T + O)v = T(v) + O(v) = T(v) = (O + T)v \text{ for all } v$$

thus O is zero of Hom (V, W)

For any $T \in$ Hom (V, W), the map $(-T) : V \to W$, such that,

$$(-T)v = -T(v)$$

will be additive inverse of T.

Again, $[\alpha(T + S)]v = \alpha[(T + S)v] = \alpha[T(v) + S(v)] = \alpha T(v) + \alpha S(v)$

$$= (\alpha T)v + (\alpha S)v = (\alpha T + \alpha S)v \quad \text{for all } v \in V$$
$$\Rightarrow \quad \alpha(T + S) = \alpha T + \alpha S$$

$$[(\alpha\beta)T]v = (\alpha\beta)T(v) = \alpha[\beta T(v)] = [\alpha(\beta T)]v \quad \text{for all } v$$

$$\Rightarrow \qquad (\alpha\beta)T = \alpha(\beta T)$$

$$(1T)v = 1 \,.\, T(v) = T(v) \quad \text{for all } v$$

$$\Rightarrow \qquad 1 \,.\, T = T$$

Hence one notices that Hom (V, W) forms a vector space over F.

Note: The notation $L(V, W)$ is also used for denoting Hom (V, W).

Definition: Product (composition) of two linear transformations

Let V, W, Z be three vector spaces over a field F

Let $T : V \to W$, $S : W \to Z$ be $L.T.$

We define $ST : V \to Z$, such that,

$$(ST)v = S(T(v))$$

then ST is a linear transformation (verify!), called product of S and T.

Note: TS may not be defined and even if it is defined it may not equal ST.

Definition: A $L.T.$ $T : V \to V$ is called a *linear operator* on V, whereas a $L.T.$ $T : V \to F$ is called a *linear functional*. We use notation T^2 for $T.T$ and $T^n = T^{n-1}T$, etc.

Theorem 3.5: Let T, T_1, T_2 be linear operators on V, and let $I : V \to V$ be the identity map $I(v) = v$ for all v (which is clearly a $L.T.$) then

(*i*) $IT = TI = T$

(*ii*) $T(T_1 + T_2) = TT_1 + TT_2$

$(T_1 + T_2)T = T_1T + T_2T$

(*iii*) $\alpha(T_1T_2) = (\alpha T_1)T_2 = T_1(\alpha T_2) \quad \alpha \in F$

(*iv*) $T_1(T_2T_3) = (T_1T_2)T_3$.

Proof: (*i*) Obvious.

(*ii*)
$$\begin{aligned} [T(T_1 + T_2)]x &= T[(T_1 + T_2)x] = T[T_1(x) + T_2(x)] \\ &= T(T_1(x)) + T(T_2(x)) = TT_1(x) + TT_2(x) \\ &= (TT_1 + TT_2)x \end{aligned}$$

$$\Rightarrow \quad T(T_1 + T_2) = TT_1 + TT_2$$

Other result follows similarly.

(*iii*) $[\alpha(T_1T_2)]x = \alpha[(T_1T_2)x] = \alpha[T_1(T_2(x))]$

$[(\alpha T_1)T_2]x = (\alpha T_1)\ [T_2(x)] = \alpha[T_1(T_2(x)]$

$[T_1(\alpha T_2)]x = T_1(\alpha T_2)x = T_1(\alpha T_2(x)) = \alpha T_1(T_2(x))]$

Hence the result follows.

(*iv*) Follows easily by definition.

Refer exercises for the generalised version of above theorem.

Theorem 3.6: Let V and W be two vector spaces (over F) of *dim* m and n respectively. Then Hom (V, W) has *dim* mn.

NOTES

Proof: Let $\{v_1, v_2, ..., v_m\}$ and $\{w_1, w_2, ..., w_n\}$ be basis of V and W respectively.

Define mappings $T_{ij} : V \to W$, such that,

$$T_{ij}(v) = \alpha_i w_j \; 1 \leq i \leq m$$
$$1 \leq j \leq n$$

where $v \in V$ is any element and therefore,

$v = \alpha_1 v_1 + \alpha_2 v_2 + ... \alpha_m v_m$ for some $\alpha_i \in F$

Note also that $T_{ij}(v_k) = 0$ if $k \neq i$

$= w_j$ if $k = i$

We show T_{ij} are $L.T.$

Let $x, y \in V$ then $x = \sum_1^m \alpha_i v_i, \quad y = \sum_1^m \beta_i v_i \quad \alpha_i, \beta_i \in F$

$$\begin{aligned}
\text{Now} \quad T_{ij}(x + y) &= T_{ij}[(\alpha_1 v_1 + ... + \alpha_m v_m) + (\beta_1 v_1 + ... + \beta_m v_m)] \\
&= T_{ij}[(\alpha_1 + \beta_1)v_1 + ... + (\alpha_m + \beta_m)v_m] \\
&= T_{ij}(\gamma_1 v_1 + ... + \gamma_m v_m) \\
&= \gamma_i w_j \\
&= (\alpha_i + \beta_i)w_j = \alpha_i w_j + \beta_i w_j = T_{ij}(x) + T_{ij}(y)
\end{aligned}$$

$$\begin{aligned}
\text{Also,} \quad T_{ij}(\lambda x) &= T_{ij}(\lambda(\alpha_1 v_1 + ... + \alpha_m v_m)) \\
&= T_{ij}(\lambda\alpha_1 v_1 + ... + \lambda\alpha_m v_m) \\
&= (\lambda\alpha_i)w_j = \lambda(\alpha_i w_j) = \lambda T_{ij}(\Sigma\alpha_i v_i) \\
&= \lambda T_{ij}(x)
\end{aligned}$$

Hence $T_{ij} \in \text{Hom}(V, W)$. We claim $S = \{T_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ forms a basis of Hom (V, W)

Suppose,

$$\beta_{11}T_{11} + \beta_{12}T_{12} + ... + \beta_{1n}T_{1n} + \beta_{21}T_{21} + \beta_{22}T_{22} + ... + \beta_{2n}T_{2n} + ... + \beta_{m1}T_{m1} + \beta_{m2}T_{m2} + ... + \beta_{mn}T_{mn} = 0, \; \beta_{ij} \in F$$

[where 0 is, of course, zero of Hom (V, W)]

By operating on v_1, we get

$$\beta_{11}T_{11}(v_1) + \beta_{12}T_{12}(v_1) + ... + \beta_{1n}T_{1n}(v_1) + \beta_{21}T_{21}(v_1) + ... = 0$$

$\Rightarrow \beta_{11}w_1 + \beta_{12}w_2 + ... + \beta_{1n}w_n + 0 + ... + 0 + ... = 0$

But $w_1, w_2 ..., w_n$ are $L.I.$

$\Rightarrow \beta_{11} = \beta_{12} = ... \beta_{1n} = 0$

Similarly, by operating on v_2 we'll get $\beta_{21} = \beta_{22} = ... \beta_{2n} = 0$

Thus by operating on v_3, v_4 ... we find that all the coefficients are zero and hence S is $L.I.$. So, $o(S) = mn$.

Let Now $T \in \text{Hom}(V, W)$ be any element, then

$T : V \to W$ is a $L.T.$

We show T is a linear combination of T_{ij}

Consider v_1, then $T(v_1) \in W$ and thus is a linear combination of $w_1, w_2, ... w_n$

Let $T(v_1) = \alpha_{11}w_1 + \alpha_{12}w_2 + ... + \alpha_{1n}w_n$

Put $T_0 = \alpha_{11}T_{11} + \alpha_{12}T_{12} + ... + \alpha_{1n}T_{1n} + \alpha_{21}T_{21} + \alpha_{22}T_{22} + ... + a_{mn}T_{mn}$

(where α_{11}, α_{12} ... are, of course, the same as before)

Then $$T_0(v_1) = \alpha_{11}T_{11}(v_1) + \alpha_{12}T_{12}(v_1) + ...$$
$$= \alpha_{11}w_1 + \alpha_{12}w_2 + \alpha_{2n}w_n + 0 + 0 + ... + 0$$
$$\Rightarrow T_0(v_1) = T(v_1)$$

Similarly proceeding with $v_2, v_3, ... v_m$ we get

$$T_0(v_2) = T(v_2)$$
$$..........$$
$$T_0(v_m) = T(v_m)$$

Thus T_0 and T agree on all elements of the basis of V.

$\Rightarrow T_0$ and T agree on all elements of $V \Rightarrow T_0 = T$

But T_0 is a linear combination of members of S

$\Rightarrow T$ is a linear combination of members of S

$\Rightarrow S$ spans Hom (V, W)

or that S forms a basis of Hom (V, W)

Hence dim Hom $(V, W) = mn$.

Corollary : Obviously dim Hom $(V, V) = m^2$ where dim $V = m$ and

dim Hom $(V, F) = m\ .1 = m$ as dim $F(F) = 1$ as F is generated by 1 and thus $\{1\}$ is a basis of $F(F)$.

Example 3.1: Find the range, Rank, Ker and nullity of the linear transformation

$T : \mathbf{R}^3 \to \mathbf{R}^3$, such that,

$T(x, y, z) = (x + z, \quad x + y + 2z, \quad 2x + y + 3z)$

Solution: Let $(x, y, z) \in$ Ker T be any element, then

$$T(x, y, z) = (0, 0, 0)$$
$$\Rightarrow (x + z,\ x + y + 2z,\ 2x + y + 3z) = (0, 0, 0)$$
$$\Rightarrow x + 0 + z = 0$$
$$x + y + 2z = 0$$
$$2x + y + 3z = 0$$

Giving $x = -z$, $-z + y + 2z = 0$ i.e., $y = -z$

Thus Ker T consists of all elements of the type $(x, x, -x) = x(1, 1, -1)$ where x is any real number, i.e., Ker T is spanned by $(1, 1, -1)$ which is *L.I.* Note $(1, 1, -1) \in$ Ker T

Hence dim (Ker T) = 1 = nullity of T

NOTES

Again, from definition of T, we notice elements of the types $(x + z,\ x + y + 2z,\ 2x + y + 3z)$ are in Range T

Now $(x + z,\ x + y + 2z,\ 2x + y + 3z) = (x + 0 + z,\ x + y + 2z,\ 2x + y + 3z)$

$= (x, x, 2x) + (0, y, y) + (z, 2z, 3z)$

$= x(1, 1, 2) + y(0, 1, 1) + z(1, 2, 3)$

Thus Range T is spanned by $\{(1, 1, 2), (0, 1, 1), (1, 2, 3)\}$

Since $(1, 1, 2) + (0, 1, 1) = (1, 2, 3)$ we find these vectors are *L.D.* So dim Range $T \leq 2$

Again as $(1, 1, 2)$ and $(0, 1, 1)$ are *L.I.* we find

$$\text{dim Range } T = 2 = \text{Rank } T.$$

Example 3.2: Find the range, rank, Ker and nullity of the following linear transformations

(*a*) $T : \mathbf{R}^2 \to \mathbf{R}^3$ such that, $T(x_1, x_2) = (x_1, x_1 + x_2, x_2)$

(*b*) $T : \mathbf{R}^4 \to \mathbf{R}^3$ such that, $T(x_1, x_2, x_3, x_4) = (x_1 - x_4, x_2 + x_3, x_3 - x_4)$

Solution: (*a*) From definition of T, we notice elements of the type $(x_1, x_1 + x_2, x_2)$ will have pre images in $\mathbf{R}^2$, i.e., elements of this type are in Range T.

$$\begin{aligned} \text{Now,}\quad (x_1, x_1 + x_2, x_2) &= (x_1 + 0, x_1 + x_2, 0 + x_2) \\ &= (x_1, x_1, 0) + (0, x_2, x_2) \\ &= x_1(1, 1, 0) + x_2(0, 1, 1) \end{aligned}$$

or that Range T is spanned by $\{(1, 1, 0), (0, 1, 1)\}$ and since

$$\alpha_1(1, 1, 0) + \alpha_2(0, 1, 1) = (0, 0, 0)$$

$$\Rightarrow \alpha_1 = \alpha_2 = 0$$

these are *L.I.* and thus form a basis of Range T

$$\Rightarrow \text{Rank } T = \text{dim Range } T = 2.$$

$$\begin{aligned} \text{Again,}\quad (x_1, x_2) \in \text{Ker } T \quad &\Rightarrow T(x_1, x_2) = (0, 0, 0) \\ &\Rightarrow (x_1, x_1 + x_2, x_2) = (0, 0, 0) \\ &\Rightarrow x_1 = 0, x_1 + x_2 = 0, x_2 = 0 \\ &\Rightarrow x_1 = x_2 = 0 \\ &\Rightarrow \text{Ker } T = \{(0, 0)\} \end{aligned}$$

Also then nullity T = dim Ker $T = 0$.

(*b*) From defintion of T, we find elements of the type $(x_1 - x_4, x_2 + x_3, x_3 - x_4)$ have pre image in $\mathbf{R}^4$.

Now,

$(x_1 - x_4,\ x_2 + x_3,\ x_3 - x_4) = (x_1 + 0 + 0 - x_4,\ 0 + x_2 + x_3 + 0,\ 0 + 0 + x_3 - x_4)$

$= x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 1, 1) + x_4(-1, 0, -1)$

or that Range T is spanned by

$\{(1, 0, 0), (0, 1, 0), (0, 1, 1), (-1, 0, -1)\}$

Since Range T is a subspace of $\mathbf{R}^3$ which has dim 3 these four elements cannot form basis of Range T.

In fact these are *L.D.*, elements as

$$(-1, 0, -1) + (1, 0, 0) + (0, 1, 0) + (0, 1, 1) = (0, 0, 0)$$

If we consider three members

$$(1, 0, 0), (0, 1, 0), (0, 1, 1)$$

we notice $\quad \alpha_1(1, 0, 0) + \alpha_2(0, 1, 0) + \alpha_3(0, 1, 1) = (0, 0, 0)$

$\Rightarrow \quad \alpha_i = 0$ for all i

or that (1, 0, 0), (0, 1, 0) (0, 1, 1) are *L.I.* and hence form basis of Range T

$\Rightarrow$ dim Range $T = 3$ = rank of T

one might notice here that as

$$(-1, 0, -1) = -1(1, 0, 0) -1(0, 1, 0) -1(0, 1, 1)$$

the elements (1, 0, 0), (0, 1, 0), (0, 1, 1) span Range T

Also then $\qquad$ Range $T = \mathbf{R}^3$

Again $\quad (x_1, x_2, x_3, x_4) \in \text{Ker } T \Rightarrow T(x_1, x_2, x_3, x_4) = (0, 0, 0)$

$$\Rightarrow x_1 - x_4 = 0$$
$$x_2 + x_3 = 0$$
$$x_3 - x_4 = 0$$

if we fix x_4, we get $x_1 = x_4,\ x_2 = -x_3 = -x_4,\ x_3 = x_4$

or that elements of the type $(x_4, -x_4, x_4, x_4)$ are in the Ker T

i.e., Ker T is spanned by $(1, -1, 1, 1)$ (Note $(1, -1, 1, 1) \in$ Ker T)

this being *L.I.* forms basis of Ker T

$\Rightarrow$ dim Ker $T = 1$

$\Rightarrow$ nullity of $T = 1$.

Example 3.3: Let F be a subfield of complex numbers and T a function from $F^3 \to F^3$ defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3,\ 2x_1 + x_2,\ -x_1 - 2x_2 + 2x_3)$$

(*i*) Show that T is a *L.T.*

(*ii*) What are the conditions on *a, b, c* such that (*a, b, c*) be in the null space of *T*? Find nullity of *T*.

Solution: $T[(x_1, x_2, x_3) + (y_1, y_2, y_3)] = T(x_1 + y_1,\ x_2 + y_2,\ x_3 + y_3)$

$= (x_1 + y_1 - x_2 - y_2 + 2x_3 + 2y_3,\ 2x_1 + 2y_1 + x_2 + y_2,$
$- x_1 - y_1 - 2x_2 - y_2 + 2x_3 + 2y_3)$

Also $T(x_1, x_2, x_3) + T(y_1, y_2, y_3) = (x_1 - x_2 + 2x_3,\ 2x_1 + x_2,\ -x_1 - 2x_2 + 2x_3)$

$+ (y_1 - y_2 + 2y_3,\ 2y_1 + y_2,\ -y_1 - 2y_2 + 2y_3)$

$= (x_1 - x_2 + 2x_3 + y_1 - y_2 + 2y_3,\ 2x_1 + x_2 + 2y_1 + y_2$

$-x_1 - 2x_2 + 2x_3 - y_1 - 2y_2 + 2y_3)$

$= (x_1 + y_1 - x_2 - y_2 + 2x_3 + 2y_3,\ 2x_1 + 2y_1 + x_2 + y_2,$

$-x_1 - y_1 - 2x_2 - y_2 + 2y_3 + 2y_3)$

NOTES

Thus $T((x_1, x_2, x_3) + (y_1, y_2, y_3)) = T(x_1, x_2, x_3) + T(y_1, y_2, y_3)$

It is easy to see that for any α

$$T(\alpha(x_1, x_2, x_3)) = \alpha T(x_1, x_2, x_3)$$

Thus T is a $L.T.$

Now if $(a, b, c) \in \text{Ker } T$ then $T(a, b, c) = (0, 0, 0)$

$$\Rightarrow (a - b + 2c,\ 2a + b,\ -a - 2b + 2c) = (0, 0, 0)$$

$$\Rightarrow a - b + 2c = 0$$

$$2a + b = 0$$

$$-a - 2b + 2c = 0$$

Since $\begin{vmatrix} 1 & -1 & 2 \\ 2 & 1 & 0 \\ -1 & -2 & 2 \end{vmatrix} = 0$

The above equations have a non zero solution.

Solving the equiations, we find

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 0 \\ -1 & -2 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \to R_2 - 2R_1,\ R_3 \to R_3 + R_1$

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -4 \\ 0 & -3 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 \to R_3 + R_2$

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -4 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow a - b + 2c = 0$$

$$3b - 4c = 0$$

Since rank of coeficient matrix is 2, the number of $L.I.$ solutions is $3 - 2 = 1$.

If we take $c = k$, we get $a = -\frac{2k}{3}$, $b = \frac{4k}{3}$, $c = k$ as solution of the given equations. In other words a, b, c should satisfy the relation $\frac{a}{-2} = \frac{b}{4} = \frac{c}{4}$ for (a, b, c) to be in Ker T.

Now $(-2, 4, 3)$ is one member of Ker T and all other members would be multiples of this, i.e., $\{(-2, 4, 3)\}$ generates ker T. Since $(-2, 4, 3)$ being non zero is *L.I.* $\{(-2, 4, 3)\}$ forms a basis of Ker T or that dim Ker T = nullity $T = 1$.

In fact, the result dim V = dim Range T + dim Ker T will then give us dim Range T = Rank $T = 2$ as dim $V = \dim F^3 = 3$.

Example 3.4: If $T_1, T_2 \in Hom\,(V, W)$ then show that

(*i*) $r(\alpha T_1) = r(T_1)$ for all $\alpha \in F$, $\alpha \neq 0$

(*ii*) $| r(T_1) - r(T_2) | \leq r(T_1 + T_2) \leq r(T_1) + r(T_2)$

where $r(T)$ means rank of T.

Solution: (*i*) $T_1 : V \to W$

thus $T_1(V)$ = range T_1, is a subspace of W

Now, $\quad (\alpha T_1)v = \alpha(T_1(v)) \in T_1(V) \qquad$ for all $v \in V$

$$\Rightarrow (\alpha T_1)V \subseteq T_1\,(V) \qquad ...(1)$$

Again as $\alpha \neq 0$, α^{-1} exists and thus,

$$(\alpha^{-1}T_1)\ V \subseteq T_1(V)$$

$$\alpha(\alpha^{-1}T_1)V \subseteq \alpha T_1(V)$$

$$\Rightarrow T_1(V) \subseteq \alpha T_1(V) \Rightarrow T_1(V) = \alpha T_1(V) \quad \text{by Equation (1)}$$

$$\Rightarrow \dim T_1(V) = \dim \alpha T_1(V)$$

or $\qquad r(T_1) = r(\alpha T_1)$.

(*ii*) Since, $\quad (T_1 + T_2)x = T_1(x) + T_2(x) \quad$ for all $x \in V$

$$(T_1 + T_2)V \subseteq T_1(V) + T_2(V)$$

$$\Rightarrow \dim\,[(T_1 + T_2)V] \quad \leq \dim\,[T_1(V) + T_2(V)]$$

$$\leq \dim T_1(V) + \dim T_2(V)$$

$$\Rightarrow r(T_1 + T_2) \leq r(T_1) + r(T_2)$$

Again, $\quad T_1 = (T_1 + T_2) - T_2 = (T_1 + T_2) + (-T_2)$

$$\Rightarrow r(T_1) = r[(T_1 + T_2) + (-T_2)]$$

$$\leq r(T_1 + T_2) + r(-T_2) = r(T_1 + T_2) + r(T_2)$$

(using Equation (1) $\alpha = -1$)

$$\Rightarrow r(T_1) - r(T_2) \leq r(T_1 + T_2)$$

Similarly, $\quad r(T_2) - r(T_1) \leq r(T_1 + T_2)$

$$\Rightarrow \quad | r(T_1) - r(T_2) | \leq r(T_1 + T_2) \leq r(T_1) + r(T_2).$$

Example 3.5: Let T be a linear operator on V. If $T^2 = 0$, what can you say about the relation of the range of T to the null space of T? Give an example of linear operator T of $\mathbf{R}^2$ such that $T^2 = 0$, but $T \neq 0$.

Solution: $T^2 = 0 \Rightarrow T^2\,(v) = 0$ for all $v \in V$

$$\Rightarrow T(T(V)) = 0$$

$$\Rightarrow T(v) \in \text{Ker } T \qquad \text{for all } v \in V$$

$$\Rightarrow \text{range } T \subseteq \text{Ker } T.$$

NOTES

Define $T : \mathbf{R}^2 \to \mathbf{R}^2$, such that

$$T(x_1, x_2) = (x_2, 0)$$

then T is a linear operator (Verify!)

Since $T(2, 2) = (2, 0) \neq (0, 0)$

$$T \neq 0$$

But $T^2(x_1, x_2) = T(T(x_1, x_2)) = T(x_2, 0) = (0, 0)$

$$\Rightarrow T^2 = 0.$$

Example 3.6: Let T be a linear operator on V and let Rank T^2 = Rank T then show that Range $T \cap Ker\ T = \{0\}$.

Solution: $T : V \to V$, $T^2 : V \to V$

$$\text{Rank } T^2 = \dim V - \dim \text{Ker } T^2$$

$$\Rightarrow \dim \text{Ker } T = \dim \text{Ker } T^2$$

We claim Ker T = Ker T^2

$$x \in \text{Ker } T \Rightarrow T(x) = 0 \Rightarrow T^2(x) = T(0) = 0$$

$$\Rightarrow x \in \text{Ker } T^2 \Rightarrow \text{Ker } T \subseteq \text{Ker } T^2$$

$$\Rightarrow \text{Ker } T = \text{Ker } T^2 \text{ (as they have same dim)}$$

Now, $x \in$ Range $T \cap$ Ker $T \Rightarrow x \in$ Range T and $x \in$ Ker T

$$\Rightarrow T(x) = 0,\ x = T(y) \text{ for some } y \in V$$

$$\Rightarrow T(T(y)) = 0$$

$$\Rightarrow T^2(y) = 0$$

$$\Rightarrow y \in \text{Ker } T^2 = \text{Ker } T$$

$$\Rightarrow T(y) = 0 \Rightarrow x = 0$$

$$\Rightarrow \text{Ker } T \cap \text{Range } T = \{0\}.$$

Invertible Linear Transformations

We recall that a map $T : V \to W$ is invertible iff it is 1-1 onto, and inverse of T is the map $T^{-1}: W \to V$ such that

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

We show that inverse of a (1–1 onto) $L.T.$ is also a $L.T.$ Let $T : V \to W$ be a 1-1 onto $L.T.$ and $T^{-1}: W \to V$ be its inverse.

We have to prove

$$T^{-1}(\alpha w_1 + \beta w_2) = \alpha T^{-1}(w_1) + \beta T^{-1}(w_2) \quad \alpha, \beta \in F,\ w_1, w_2 \in W$$

Since T is onto, for $w_1, w_2 \in W$, $\exists\ v_1, v_2 \in V$

such that, $T(v_1) = w_1$, $T(v_2) = w_2$

$$\Leftrightarrow v_1 = T^{-1}(w_1),\ v_2 = T^{-1}(w_2)$$

Now, $T^{-1}(\alpha w_1 + \beta w_2) = T^{-1}(\alpha T(v_1) + \beta T(v_2))$

$$= T^{-1}(T(\alpha v_1) + T(\beta v_2))$$

$$= T^{-1}(T(\alpha v_1 + \beta v_2))$$

$= \alpha v_1 + \beta v_2$

$= \alpha T^{-1}(w_1) + \beta T^{-1}(w_2).$

NOTES

Definition: A *L.T.* $T : V \to W$ is called *non-singular* if Ker $T = \{0\}$, i.e., if T is 1-1.

Theorem 3.7: A linear transformation $T : V \to W$ is non singular iff T carries each *L.I.* subset of V onto a *L.I.* subset of W.

Proof: Let T be non-singular and $\{v_1, v_2 ..., v_n\}$ be a *L.I.* subset of V. we show $\{T(v_1), T(v_2) ..., T(v_n)\}$ is *L.I.* subset of W.

Now $\quad \alpha_1 T(v_1) + \alpha_2 T(v_2) + ... + \alpha_n T(v_n) = 0 \quad \alpha_i \in F$

$\Rightarrow \; T(\alpha_1 v_1 + ... + \alpha_n v_n) = 0$

$\Rightarrow \; \alpha_1 v_1 + ... \; \alpha_n v_n \in \text{Ker } T = \{0\}$

$\Rightarrow \; \alpha_1 v_1 + ... \; \alpha_n v_n = 0$

$\Rightarrow \; \alpha_i = 0$ for all i as $v_1, v_2 ..., v_n$ are *L.I.*

Conversely, let $v \in$ Ker T be any element

Then, $\quad T(v) = 0$

$\Rightarrow \; \{T(v)\}$ is not *L.I.* in W

$\Rightarrow \; v$ is not *L.I.* in V. (by hypothesis)

$\Rightarrow \; v = 0 \Rightarrow$ Ker $T = \{0\}$

$\Rightarrow \; T$ is non singular.

Theorem 3.8: Let $T : V \to W$ be a *L.T.* where V and W are two *F.D.V.S.* with same dimension. Then the following are equivalent

(*i*) T is invertible

(*ii*) T is non singular (i.e., T is 1-1)

(*iii*) T is onto (i.e., Range $T = W$)

(*iv*) If $\{v_1, v_2, ..., v_n\}$ is a basis of V then $\{T(v_1), T(v_2), ..., T(v_n)\}$ is a basis of W.

Proof: (*i*) $\Rightarrow$ (*ii*) F follows by definition.

(*ii*) $\Rightarrow$ (*iii*) T is non-singular

$\Rightarrow$ Ker $T = \{0\}$

$\Rightarrow$ dim Ker $T = 0$

Since $\quad$ dim Range T + dim Ker T = dim V, we get

dim Range T = dim V

$\Rightarrow$ dim Range T = dim W (given condition)

But Range T being a subspace of W, we find

Range $T = W$

(*iii*) $\Rightarrow$ (*i*) T onto means Range $T = W$

$\Rightarrow$ dim Range T = dim W = dim V

and as dim Range T + dim Ker T = dim V, we get

NOTES

dim Ker $T = 0$

$\Rightarrow$ Ker $T = \{0\}$

or that T is 1-1 and as it is onto T will be invertible.

$(i) \Rightarrow (iv)$ T is invertible $\Rightarrow T$ is 1-1 onto

i.e., T is an isomorphism.

$(iv) \Rightarrow (i)$

Let $\{T(v_1),..., T(v_n)\}$ be basis of W where $\{v_1,...v_n\}$ is basis of V. Any $w \in W$ can be put as

$$w = \alpha_1 T(v_1) +...+ \alpha_n T(v_n)$$
$$= T(\alpha_1 v_1 +...+ \alpha_n v_n) = T(v) \text{ for some } v \in V$$

$\therefore$ T is onto. Thus (iii) holds.

Hence (i) holds.

Example 3.7: Let T be a linear operator on $\mathbf{R}^3$, defined by

$$T(x_1, x_2, x_3) = (3x_1, x_1 - x_2, \ 2x_1 + x_2 + x_3)$$

show that T is invertible and find the rule by which T^{-1} is defined.

Solution: $T : \mathbf{R}^3 \to \mathbf{R}^3$

Let $(x_1, x_2, x_3) \in$ Ker T be any element

Then $\quad T(x_1, x_2, x_3) = (0, 0, 0)$

$$\Rightarrow (3x_1, \ x_1 - x_2, \ 2x_1 + x_2 + x_3) = (0, 0, 0)$$
$$\Rightarrow 3x_1 = 0, \ x_1 - x_2 = 0, \ 2x_1 + x_2 + x_3 = 0$$
$$\Rightarrow x_1 = x_2 = x_3 = 0 \text{ or that Ker } T = \{(0, 0, 0)\}$$

$\Rightarrow$ T is non singular and thus invertible (Refer Theorem 3.8)

Now if (z_1, z_2, z_3) be any element of $\mathbf{R}^3$, then (x_1, x_2, x_3) will be its image under T if,

$$T(x_1, x_2, x_3) = (z_1, z_2, z_3)$$
$$\Rightarrow 2x_1 = z_1$$
$$x_1 - x_2 = z_2$$
$$2x_1 + x_2 + x_3 = z_3$$

which give $x_1 = \dfrac{z_1}{3}$, $x_2 = \dfrac{z_1}{3} - z_2$, $z_3 = z_3 - z_1 + z_2$

Hence $T^{-1} : \mathbf{R}^3 \to \mathbf{R}^3$ is defined by

$$T^{-1}(z_1, z_2, z_3) = \left(\frac{z_1}{3}, \frac{z_1}{3} - z_2, \ z_3 - z_1 + z_2\right)$$

Example 3.8: If $T : V \to V$ is a *L.T.*, such that T is not onto, then show that there exists some $0 \neq v$ in V such that, $T(v) = 0$.

Solution: Since T is not onto, it is not 1-1 (theorem done)

Suppose $\exists$ no $0 \neq v \in V$ s.t. $T(v) = 0$

Then $T(v) = 0$ only when $v = 0$

$\Rightarrow$ Ker $T = \{0\} \Rightarrow T$ is 1-1, a contradiction.

Theorem 3.9: Let $T : V \to W$ and $S : W \to U$ be two linear transformations. Then

NOTES

(*i*) If S and T are one-one onto then ST is one-one onto and $(ST)^{-1} = T^{-1} S^{-1}$.

(*ii*) If ST is one-one then T is one-one

(*iii*) If ST is onto then S is onto.

Proof: (*i*) Since S and T are 1-1 onto, S^{-1} and T^{-1} exist.

Let $\quad ST(x) = ST(y)$

Then $\quad S(T(x)) = S(T(y))$

$\Rightarrow T(x) = T(y)$ as S is 1-1

$\Rightarrow x = y$ as T is 1-1

$\Rightarrow ST$ is 1-1.

Again $ST : V \to U$, let $u \in$ U be any element then as S in onto, $\exists\, w \in W$ such that, $S(w) = u$ and as $T : V \to W$ is onto $\exists\, v \in V$ such that, $T(v) = w$

Now $\quad T(v) = w \Rightarrow S(T(v)) = S(w) \Rightarrow ST(v) = u$

or that ST is onto.

Also $(ST)(T^{-1}S^{-1}) = S(T(T^{-1}S^{-1})) = S(TT^{-1})S^{-1} = S(IS^{-1}) = SS^{-1} = I$

Similarly $(T^{-1}S^{-1})(ST) = T^{-1}(S^{-1}(ST)) = T^{-1}(S^{-1}S)T = T^{-1}(IT) = T^{-1}T = I$

Showing that, $\quad (ST)^{-1} = T^{-1}S^{-1}$.

(*ii*) Let $v \in$ Ker T be any element

Then $\quad T(v) = 0$

$\Rightarrow S(T(v)) = S(0)$

$\Rightarrow ST(v) = 0$

$\Rightarrow v \in$ Ker ST and Ker $ST = (0)$ as ST is 1-1

$\Rightarrow v = 0 \Rightarrow$ Ker $T = (0) \Rightarrow T$ is 1-1.

(*iii*) Let $u \in U$ be any element. Since $ST : V \to U$ is onto, $\exists$ some $v \in V$ such that, $ST(v) = u$

i.e., $\quad S(T(v)) = u$

Let $T(v) = w$ and $w \in W$ such that,

$\quad S(w) = u$

Then, $\quad S$ is onto.

Example 3.9: In the above theorem show that if ST is 1-1 onto then T is 1-1 and S is onto. Again, if *V, W, U* are of same dimension and ST is one-one onto then so are S and T.

Solution: First part of the problem follows by (*ii*) and (*iii*) of Theorem 3.9.

Let now $\quad$ dim V = dim W = dim U

NOTES

The result then follows by using Theorem 3.8 we proved earlier that if $T : V \to W$ is a $L.T.$ where dim V = dim W then T is 1-1 iff T is onto.

Example 3.10: Let T be a linear operator on $F.D.V.S.$ Suppose there is a linear operator U on V such that $TU = I$. Show that T is invertible and $T^{-1} = U$.

Solution: We have $T : V \to V$, $U : V \to V$ such that, $TU = I$ we claim U is 1-1.

Let $\quad U(x) = U(y)$

Then $\quad T(U(x)) = T(U(y))$

$\Rightarrow I(x) = I(y)\ (TU = I)$

$\Rightarrow x = y$

or that U is 1-1 and, therefore, onto also.

Hence U is invertible.

Now $U^{-1} : V \to V$ such that, $UU^{-1} = 1$

Thus $UT = (UT)I = UT(UU^{-1}) = U(TU)U^{-1} = UU^{-1} = I$

$\Rightarrow UT = I = TU$

$\Rightarrow T$ is invertible and $T^{-1} = U$.

Example 3.11: Show that the conclusion of the previous problem fails if V is not finite dimensional.

Solution: Let V be the vector space of all polynomials in x over a filed F.

Let T = Differential operator on V.

i.e., $\quad T : V \to V$, such that,

$$T(f(x)) = \frac{d}{dx} f(x)$$

Notice this T is a linear transformation.

Let $\quad U : V \to V$ such that,

$$U(f) = \int_0^x f(t)\,dt$$

Then U is a linear transformation.

Again $TU(f) = T\int_0^x f(t)\,dt = f = I(f)$

$\Rightarrow TU = I$

Now $T(2x) = 2,\ T(2x + 3) = 2$

and as $2x \neq 2x + 3$, T is not 1-1 and hence T is not invertible.

Thus $UT \neq I$.

Example 3.12: Let V_1 and V_2 be vector spaces over F. Show that $V_1 \times V_2$ is $F.D.V.S.$ if and only if V_1 and V_2 are $F.D.V.S.$

Solution: Let $\quad V_1' = \{(v_1, 0) \mid v_1 \in V_1\}$

$V_2' = \{(0, v_2) \mid v_2 \in V_2\}$

then $\quad V_1'$ and V_2' are subspaces of $V_1 \times V_2$

Define $\theta_1 : V_1 \to V_1'$ such that,

$$\theta_1(v_1) = (v_1, 0)$$

Then θ_1 is an isomorphism (Prove!)

Similarly $\theta_2 : V_2 \to V_2'$ such that,

$$\theta_2(v_2) = (0, v_2)$$

will be an isomorphism.

So $\quad V_1 \cong V_1', \quad V_2 \cong V_2'$

Suppose $V_1 \times V_2$ is *F.D.V.S.*, then V_1' and V_2' are *F.D.V.S.* (being subspaces of $V_1 \times V_2$)

$$\Rightarrow \; V_1 \text{ and } V_2 \text{ are } F.D.V.S.$$

Conversely, if V_1 and V_2 are *F.D.V.S.* then $V_1 \times V_2$ is *F.D.V.S.* and dim $(V_1 \times V_2) = \dim V_1 + \dim V_2$. (Note: If $\{e_1, e_2, ..., e_m\}$ and $\{f_1, f_2, ..., f_n\}$ are basis of V_1 and V_2 respectively, then $\{(e_1, 0), ..., (e_m, 0), (0, f_1), ..., (0, f_n)\}$ is a basis of $V_1 \times V_2$.)

Example 3.13: Let W_1 and W_2 be subspaces of V such that $\frac{V}{W_1}$ and $\frac{V}{W_2}$ are *F.D.V.S.* Show that $\frac{V}{W_1 \cap W_2}$ is also a *F.D.V.S.*

Solution: Define $\theta : V \to \frac{V}{W_1} \times \frac{V}{W_2}$ such that,

$$\theta(v) = (W_1 + v, W_2 + v)$$

It is easy to see that θ is a linear transformation where Ker $\theta = W_1 \cap W_2$.

Hence $\quad \frac{V}{\text{Ker } \theta} \cong \theta(V)$

Again, since $\frac{V}{W_1}$ and $\frac{V}{W_2}$ are *F.D.V.S.*, so will be $\frac{V}{W_1} \times \frac{V}{W_2}$. In fact $\dim\left(\frac{V}{W_1} \times \frac{V}{W_2}\right) = \dim \frac{V}{W_1} + \dim \frac{V}{W_2}$.

Also $\theta(V)$ is a subspace of $\frac{V}{W_1} \times \frac{V}{W_2}$ and is therefore, finite dimensional.

Hence $\frac{V}{W_1 \cap W_2}$ is *F.D.V.S.*

3.2.2 Metric on Normed Linear Spaces

Let $U(F)$, $V(F)$ be vector spaces of dimension n and m respectively. Let $\beta = \{u_1, ..., u_n\}$, $\beta' = \{v_1, ..., v_m\}$ be their ordered basis respectively. Suppose $T : U \to V$ is a linear transformation. Since $T(u_1), ..., T(u_n) \in V$ and $\{v_1, ..., v_m\}$ spans V, each $T(u_i)$ is a linear combination of vectors $v_1, ..., v_m$.

Let $\quad T(u_1) = \alpha_{11}v_1 + ... \alpha_{m1}v_m$

$\quad\quad T(u_2) = \alpha_{12}v_1 + ... + \alpha_{m2}v_m$

....

$$T(u_n) = \alpha_{1n}v_1 + ... + \alpha_{mn}v_m$$

where each $\alpha_{ij} \in F$. Then the $m \times n$ matrix

NOTES

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & ... & ... & \alpha_{1n} \\ : & : & ... & ... & : \\ : & : & ... & ... & : \\ : & : & ... & ... & : \\ \alpha_{m1} & \alpha_{m2} & ... & ... & \alpha_{mn} \end{bmatrix}$$

is called matrix of T with repsect to ordered basis β, β' respectively. A is uniquely determined by T as each $\alpha_{ij} \in F$ is uniquely determined. We write

$$A = [T]_{\beta, \beta'}$$

The word *ordered basis is very significant, for as the order of basis is changed, the entries* α_{ij} *will change their positions and so the corresponding matrix will be different.*

In particular if $U = V$, $\beta = \beta'$, then instead of writing $[T]_{\beta, \beta'}$, we write $[T]_{\beta}$.

Let $M_{m\times n}(F)$ denote the vector space of all $m \times n$ matrices over F. Let Hom (U, V) denote the vector space of all linear transformations from $U(F)$ into $V(F)$. We prove

Theorem 3.10: *Hom* $(U, V) \cong M_{m\times n}(F)$.

Proof: Define θ : Hom $(U, V) \to M_{m\times n}(F)$, such that,

$$\theta(T) = [T]_{\beta, \beta'}$$

Where $\beta = \{u_1, ... u_n\}$, $\beta' = \{v_1, ... v_m\}$ are ordered basis of U, V respectively. θ is well defined as $[T]_{\beta, \beta'}$ is uniquely determined by T.

It is not difficult to verify that θ is a linear transformation.

Let $\quad \theta(S) = \theta(T), \quad S, T \in$ Hom (U, V)

Then, $[S]_{\beta, \beta'} = [T]_{\beta, \beta'}$

$\Rightarrow \quad (a_{ij}) = (b_{ij})$

$\Rightarrow \quad a_{ij} = b_{ij}$ for all i, j

$$\Rightarrow \quad S(u_j) = \sum_{i=1}^{m} a_{ij}v_i = \sum_{i=1}^{m} b_{ij}v_i = T(u_j) \text{ for all } j = 1, ... n$$

$\Rightarrow \quad S = T \Rightarrow \theta$ is 1-1.

Let $A = (a_{ij})_{m\times n} \in M_{m\times n}(F)$. Then $\exists$ a linear transformation $T \in$ Hom (U, V) such that,

$$T(u_j) = \sum_{i=1}^{m} a_{ij}v_i \quad \text{for } j = 1, ..., n$$

$\therefore \quad A = [T]_{\beta, \beta'} = \theta(T) \Rightarrow \theta$ is onto.

Hence θ is an isomorphism and so Hom $(U, V) \cong M_{m\times n}(F)$.

NOTES

Corollary : dim Hom $(U, V) = mn$.

Proof: S = set of all $m \times n$ matrices with only one entry 1 and all other entries zero, is a basis of $M_{m\times n}(F)$.

Clearly, $o(S) = mn \Rightarrow \dim M_{m\times n}(F) = mn$

$$\dim \text{Hom}(U, V) = mn.$$

Theorem 3.11: Let S, T be two linear transformations from $V(F)$ into $V(F)$. Let β be an ordered basis of V. Then

$$[ST]_\beta = [S]_\beta[T]_\beta.$$

Proof: Let $\beta = \{v_1, \ldots v_n\}$

Let $S(v_1) = a_{11}v_1 + \ldots a_{n1}v_1$

....

$S(v_n) = a_{1n}v_1 + \ldots + a_{nn}v_n$

where $a_{ij} \in F$

In general, $S(v_j) = \sum_{i=1}^{n} a_{ij}v_i$ for all $j = 1, \ldots, n$

$\therefore \quad [S]_\beta = (a_{ij})$

Similarly,

$T(v_1) = b_{11}v_1 + \ldots + b_{n1}v_n$

....

$T(v_n) = b_{1n}v_1 + \ldots + b_{nn}v_n$ where $b_{ij} \in F$

In general, $T(v_k) = \sum_{j=1}^{n} b_{jk}\, v_j$, for all $k = 1, \ldots, n$

$\therefore \quad [T]_\beta = (b_{jk})$

$$\therefore \quad ST(v_k) = S\left(\sum_{j=1}^{n} b_{jk}\, v_j\right) = \sum_{j=1}^{n} b_{jk}\, S(v_j)$$

$$= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)v_i$$

$$[ST]_\beta = (c_{ik}), \text{ where } c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$$

Also, (i, k)th entry in $[S]_\beta\,[T]_\beta$

$$= \sum_{j=1}^{n} a_{ij}b_{jk} = c_{ik} = (i, k)\text{th entry in } [ST]_\beta$$

$$\therefore \quad [ST]_\beta = [S]_\beta\,[T]_\beta$$

NOTES

Corollary : If S is an invertible linear transformation from $V(F)$ into $V(F)$, then so is $[S]_\beta$ with respect to any basis β of V and conversely.

Proof: Since S is invertible, $\exists\ T : V \to V$ such that, $ST = I = TS$. Let β be an ordered basis of V. Then by above theorem,

$$[ST]_\beta = [I]_\beta = I, \text{ where } T = S^{-1}$$

$$\Rightarrow\ [S]_\beta[T]_\beta = I$$

$$\Rightarrow\ [S]_\beta[S^{-1}]_\beta = I$$

$$\Rightarrow\ [S^{-1}]_\beta = [S]^{-1}{}_\beta \text{ for any basis } \beta \text{ of } V$$

Conversely, let $[S]_\beta$ be invertible. Then $\exists$ a matrix $A = (a_{ij})$ over F such that, $[S]_\beta\ A = I$

Let $T : V \to V$ be a linear transformation such that,

$$T(v_j) = \sum_{i=1}^{n} a_{ij} v_i \quad \text{for all } j = 1, \ldots n$$

$$\therefore \quad [T]_\beta = A$$

$$\therefore \quad [S]_\beta\ [T]_\beta = I$$

$$\Rightarrow\ [ST]_\beta = I$$

$$\Rightarrow\ (ST)(v_j) = v_j \quad \text{for all } j = 1, \ldots, n$$

$$\begin{aligned}\Rightarrow\ (ST)(x) &= (ST)(\alpha_1 v_1 + \ldots + \alpha_n v_n)\\ &= \alpha_1 v_1 + \ldots + \alpha_n v_n\\ &= x \quad \text{for all } x \in V\end{aligned}$$

$$\Rightarrow\ ST = I \Rightarrow S \text{ is invertible.}$$

We now give a relation between matrices of a linear transformation with respect to two different basis of a vector space.

Theorem 3.12: Let $T : V(F) \to V(F)$ be a linear transformation. Let $\beta = \{u_1, \ldots, u_n\}$, $\beta' = \{v_1, \ldots, v_n\}$ be two ordered basis of V. Then $\exists$ a non singular matrix P over F such that

$$[T]_{\beta'} = P^{-1}[T]_\beta P.$$

Proof: Let $S : V \to V$ be a linear transformation such that $S(u_i) = v_i$ for all $i = 1, \ldots n$.

Now $x \in \text{Ker } S \Rightarrow S(x) = 0$, $x = \alpha_1 u_1 + \ldots + \alpha_n u_n$, $\alpha_i \in F$

$$\Rightarrow\ S(\alpha_1 u_1 + \ldots + \alpha_n u_n) = 0$$

$$\Rightarrow\ \alpha_1 S(u_1) + \ldots + \alpha_n S(u_n) = 0$$

$$\Rightarrow\ \alpha_1 v_1 + \ldots + \alpha_n v_n = 0$$

$$\Rightarrow\ \alpha_i = 0 \quad \text{for all } i$$

$$\Rightarrow\ x = 0$$

$$\Rightarrow\ \text{Ker } S = \{0\}$$

$$\Rightarrow\ S \text{ is 1-1 and so onto.}$$

$\therefore$ S is an isomorphism. Let $[T]_\beta = (a_{ij})$

Then $\quad T(u_j) = \sum_{i=1}^{n} a_{ij} u_i$

$\therefore \quad (STS^{-1})(v_j) = ST(u_j)$

$$= S\left(\sum_{i=1}^{n} a_{ij} u_i\right) = \sum_{i=1}^{n} a_{ij} v_i$$

$\therefore \quad [STS^{-1}]_{\beta'} = (a_{ij}) = [T]_\beta$

$\Rightarrow \quad [S]_{\beta'}[T]_{\beta'}[S^{-1}]_{\beta'} = [T]_\beta$

$\Rightarrow \quad [S]_{\beta'}[T]_{\beta'}[S]_{\beta}^{-1} = [T]_\beta$

$\Rightarrow \quad [T]_{\beta'} = [S]_{\beta'}^{-1}[T]_\beta[S]_{\beta'}$

$= P^{-1}[T]_\beta P$, where $P = [S]_{\beta'}$.

Example 3.14: Let T be a linear operator on $\mathbf{C}^2$ defined by $T(x_1, x_2) = (x_1, 0)$ Let $\beta = \{\in_1 = (1, 0), \in_2 = (0, 1)\}$, $\beta' = \{\alpha_1 = (1, i), \alpha_2 = (-i, 2)\}$ be ordered basis for $\mathbf{C}^2$. What is the matrix of T relative to the pair β, β'?

Solution: Now $T(\in_1) = T(1, 0)$

$= (1, 0)$

$= a(1, i) + b(-i, 2)$

$\Rightarrow \quad a - bi = 1$ where $a, b \in \mathbf{C}$

$ai + 2b = 0$

$\Rightarrow \quad a = 2, b = -i$

$\Rightarrow \quad T(\in_1) = 2\alpha_1 - i\alpha_2$

Also $T(\in_2) = T(0, 1) = (0, 0) = 0\alpha_1 + 0\alpha_2$

$$\therefore \quad [T]_{\beta\,\beta'} = \begin{bmatrix} 2 & 0 \\ -i & 0 \end{bmatrix}.$$

Example 3.15: Let T be the linear operator on $\mathbf{R}^2$ defined by $T(x_1, x_2) = (-x_2, x_1)$

(*i*) Prove that for all real numbers c, the operator $(T - cI)$ is invertible.

(*ii*) Prove that if β is any ordered basis for $\mathbf{R}^2$ and $[T]_\beta = A$, then $a_{12}a_{21} \neq 0$, where $A = (a_{ij})$.

Solution: (*i*) Let $\beta = \{\in_1 = (1, 0), \in_2 = (0, 1)\}$ be an ordered basis for $\mathbf{R}^2$.

Then, $\quad T(\in_1) = T(1, 0) = (0, 1) = 0\in_1 + 1\in_2$

$T(\in_2) = T(0, 1) = (-1, 0) = -1\in_1 + 0\in_2$

$$\therefore \quad [T]_\beta = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},\ [cI]_\beta = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}$$

$$\therefore \quad [T - cI]_\beta = \begin{bmatrix} -c & -1 \\ 1 & -c \end{bmatrix}$$

$\det [T - cI]_\beta = c^2 + 1 \neq 0$ for all real numbers c

$\therefore \quad [T - cI]_\beta$ is invertible.

$\Rightarrow \quad T - cI$ is invertible for all real numbers c.

(*ii*) Let β be any ordered basis for $\mathbf{R}^2$ such that,

NOTES

$$[T]_\beta = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = A, \quad a_{ij} \in \mathbf{R}$$

By (*i*) $T - a_{11}I$ is Invertible

$\Rightarrow \quad [T - a_{11}I]_\beta$ is invertible

$\Rightarrow \quad \begin{bmatrix} 0 & a_{12} \\ a_{21} & a_{22} - a_{11} \end{bmatrix}$ is invertible

$\Rightarrow \quad -a_{12}a_{21} \neq 0$ as det of abve matrix $\neq 0$

$\Rightarrow \quad a_{12}a_{21} \neq 0.$

Example 3.16: Let T be the linear operator on $\mathbf{R}^3$ defined by

$$T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3)$$

Show that T is invertible.

Solution: Let $\beta = \{\in_1 = (1, 0, 0), \in_2 = (0, 1, 0), \in_3 = (0, 0, 1)\}$ be an ordered basis of $\mathbf{R}^3$. Then $[T]_\beta = \begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix} = A.$

$$\det A = 3(4) + 1\,(-4 + 1) = 12 - 3 = 9 \neq 0$$

So, A is invertible

$\Rightarrow T$ is invertible.

Example 3.17: Let A be an $n \times n$ matrix over F. Show that A is invertible if and only if columns of A are linearly independent over F.

Solution: Let $V(F)$ be a vector space of dimension n. Let $\beta = \{v_1,..., v_n\}$ be an ordered basis of V. Let $A = (a_{ij})$. Then $\exists$ a linear transformation $T : V \to V$ such that,

$$T(v_j) = \sum_{i=1}^{n} a_{ij} v_i$$

$\therefore \quad [T]_\beta = A.$

Let $M_n(F)$ denote the vector space of all $n \times n$ matrices over F.

Let $A \in M_n(F)$ be invertible. Then T is also invertible (by Corollary to Theorem 3.11) and so T is 1-1, onto.

Let, $\quad \alpha_1 \begin{bmatrix} a_{11} \\ : \\ a_{n1} \end{bmatrix} + ... + \alpha_n \begin{bmatrix} a_{1n} \\ : \\ a_{nn} \end{bmatrix} = 0, \; \alpha_i \in F$

NOTES

$$\Rightarrow \begin{bmatrix} \alpha_1 a_{11} & \cdots & +\alpha_n a_{1n} \\ \cdots & \cdots & \cdots \\ \alpha_1 a_{n1} + & \cdots & +\alpha_n a_{nn} \end{bmatrix} = 0$$

$$\Rightarrow \quad \alpha_1 a_{11} + ... + \alpha_n a_{1n} = 0$$
$$... \qquad ...$$
$$\alpha_1 a_{n1} + ... + \alpha_n a_{nn} = 0$$
$$\Rightarrow \quad \alpha_1 a_{11}\ v_1 + ... + \alpha_n a_{1n}\ v_1 = 0$$
$$... \qquad ...$$
$$\alpha_1 a_{n1}\ v_n + ... + \alpha_n a_{nn}\ v_n = 0$$
$$\Rightarrow \quad \alpha_1(a_{11}v_1 + ... + a_{n1}v_n) + ... + \alpha_n(a_{1n}v_1 + ... + \alpha_{nn}v_n) = 0$$
$$\Rightarrow \quad \alpha_1 T(v_1) + ... + \alpha_n T(v_n) = 0$$
$$\Rightarrow \quad T(\alpha_1 v_1 + ... + \alpha_n v_n) = 0$$

$\Rightarrow$ $\alpha_1 v_1 + ... + \alpha_n v_n = 0$ as T is 1-1

$\Rightarrow$ $\alpha_i = 0$ for all i

$\Rightarrow$ Columns of A are linearly independent.

Conversely, let columns of A be linearly independent over F.

Now, $x \in \text{Ker } T$

$$\Rightarrow \quad T(x) = 0,\ x \in V$$
$$\Rightarrow \quad T(\alpha_1 v_1 + ... + \alpha_n v_n, = 0$$
$$\Rightarrow \quad \alpha_1 T(v_1) + ... + \alpha_n T(v_n) = 0$$
$$\Rightarrow \quad \sum_{j=1}^{n} \alpha_j T(v_j) = 0 \quad \Rightarrow \quad \sum_{j=1}^{n} \alpha_j \left(\sum_{i=1}^{n} a_{ij} v_i \right) = 0$$
$$\Rightarrow \quad \sum_{i=1}^{n} \left(\sum_{j=1}^{n} (\alpha_j a_{ij}) \right) v_i = 0$$
$$\Rightarrow \quad \sum_{j=1}^{n} \alpha_j a_{ij} = 0 \text{ for all } i = 1, ..., n$$
$$\Rightarrow \quad \alpha_1 \begin{bmatrix} a_{11} \\ : \\ a_{n1} \end{bmatrix} + ... + \alpha_n \begin{bmatrix} a_{1n} \\ : \\ a_{nn} \end{bmatrix} = 0$$

$\Rightarrow$ each $\alpha_i = 0$ as columns are linearly independent.

$\Rightarrow$ $x = 0$ $\Rightarrow$ Ker $T = \{0\}$

$\Rightarrow$ T is 1-1 and so onto.

$\Rightarrow$ T is invertible.

NOTES

Example 3.18: Let T be the linear operator on $\mathbf{R}^2$ defined by $T(x_1, x_2) = (-x_2, x_1)$.

Let $\beta = \{\in_1 = (1, 0), \in_2 = (0, 1)\}$

$\beta' = \{\alpha_1 = (1, 2), \alpha_2 = (1, -1)\}$

be ordered basis for $\mathbf{R}^2$. Find a matrix P such that,

$$[T]_{\beta'} = P^{-1}[T]_\beta P.$$

Proof: $T(\in_1) = T(1, 0) = (0, 1) = 0\in_1 + 1\in_2$

$T(\in_2) = T(0, 1) = (-1, 0) = -1\in_1 + 0\in_2$

$$\therefore \qquad [T]_\beta = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Define $S : \mathbf{R}^2 \to \mathbf{R}^2$ such that,

$S(\in_i) = \alpha_i \quad i = 1, 2$

Now, $\alpha_1 = (1, 2) = 1\in_1 + 2\in_2$

$\alpha_2 = (1, -1) = 1\in_1 + (-1)\in_2$

$\Rightarrow$ $S(\alpha_1) = 1\alpha_1 + 2\alpha_2$

$S(\alpha_2) = 1\alpha_1 + (-1)\alpha_2$

$$\Rightarrow \qquad [S]_{\beta'} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

$$\Rightarrow \qquad P = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \text{ and } P^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

$$\Rightarrow \qquad P^{-1}\,[T]_\beta\,P = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{2}{3} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} \\ -\frac{5}{3} & -\frac{1}{3} \end{bmatrix}$$

$$= [T]_{\beta'}$$

Example 3.19: Let T be linear operator on $\mathbf{R}^3$, the matrix of which in the standard ordered basis is

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$$

Find a basis for the range of T and a basis for the null space of T.

Solution: Det $A = 1(4 - 3) - 2\,(1) + 1(1)$

$= 1 - 2 + 1 = 0$

$\therefore$ A is not invertible and so T is not invertible.

Let $\{\in_1 = (1, 0, 0), \in_2 = (0, 1, 0), \in_3 = (0, 0, 1)\}$

be standard ordered basis of $\mathbf{R}^3$.

Let $(x_1, x_2, x_3) \in \text{Ker } T$

Then, $T(x_1, x_2, x_3) = 0$

$\Rightarrow$ $T(x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1)) = 0$

$\Rightarrow$ $T(x_1 \in_1 + x_2\in_2 + x_3\in_3) = 0$

$\Rightarrow$ $x_1T(\in_1) + x_2T(\in_2) + x_3T(\in_3) = 0$

$\Rightarrow$ $x_1(1, 0, -1) + x_2(2, 1, 3) + x_3(1, 1, 4) = 0$

$\Rightarrow$ $(x_1 + 2x_2 + x_3, x_2 + x_3, -x_1 + 3x_2 + 4x_3) = 0$

$\Rightarrow$ $x_1 + 2x_2 + x_3 = 0, x_2 + x_3 = 0, -x_1 + 3x_2 + 4x_3 = 0$

$\Rightarrow$ $x_1 + x_2 = 0, x_2 + x_3 = 0$

$\Rightarrow$ $(x_1, x_2, x_3) = (-x_2, x_2, -x_3)$

$= x_2(-1, 1, -1)$

$\Rightarrow$ Every element in Ker T is multiple of $(-1, 1, -1)$

$\Rightarrow$ Ker T is spanned by $(-1, 1, -1)$

Since $(-1, 1, -1) \neq 0$, $\{(-1, 1, -1)\}$ is a basis of Ker T.

$\therefore$ dim Ker $T = 1 \Rightarrow$ dim Range $T = 2$

Since $T\in_1 = (1, 0, -1)$

$T\in_2 = (2, 1, 3)$

belong to Range T and $aT\in_1 + bT\in_2 = 0$

we find $a(1, 0, -1) + b(2, 1, 3) = 0$

$\Rightarrow$ $b = 0, a = 0$

$\Rightarrow \{T\in_1, T\in_2\}$ is a linearly independent set in Range T. As dim Range $T = 2$, $\{(1, 0\,, -1), (2, 1, 3)\}$ is a basis of Range T.

Example 3.20: Let T be a linear operator on F^n and let A be the matrix of T in the standard ordered basis for F^n. Let W be the subspace of F^n spanned by the column vectors of A. Find a relation between W and T.

Solution: $T : F^n \to F^n$

Let $\beta = \{e_1 = (1, 0, 0, ... 0), ..., e_n = (0, 0, ..., 1)\}$ be the standard ordered basis of F^n and let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ : & : & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

NOTES

thus, $T(e_1) = a_{11}e_1 + a_{21}e_2 + ... + a_{n1}e_n$

$T(e_2) = a_{12}e_1 + a_{22}e_2 + ... + a_{n2}e_2$

... ...

$T(e_n) = a_{1n}e_1 + a_{2n}e_2 + ... + a_{nn}e_n$

and also W is spanned by

$$\{(a_{11}, a_{21}, ..., a_{n1}), (a_{12}, a_{22}, ..., a_{n2}), ..., (a_{1n}, a_{2n}, ..., a_{nn})\}$$

We claim $T : F^n \to W$ is onto $L.T.$

For any $x \in F^n$, $x = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n$

$\Rightarrow T(x) = \alpha_1 T(e_1) + \alpha_2 T(e_2) + ... + \alpha_n T(e_n)$

$\Rightarrow T(x) \in W$ as $T(e_1), T(e_2), ..., T(e_n) \in W$

Again, for any $w \in W$, $w = \beta_1 T(e_1) + \beta_2 T(e_2) + ... + \beta_n T(e_n)$

$= T(\beta_1 e_1 + \beta_2 e_2 + ... + \beta_n e_n)$

showing that T is onto.

$\Rightarrow$ Range $T = W \Rightarrow$ dim Range T = dim W

or that rank of T = dim W

which is the required relation between T and W.

3.2.3 Linear Transformation and Dual Spaces

The set of all linear transformations from vector space V over F into vector space W over F, is also a vector space over F. Further, if dim $V = m$, dim $W = n$, then dim How $(V, W) = mn$. In particular, if $W = F$, then,

Hom (V, F) is called dual space of V over F. It is denoted by $\hat{V}$ and read as V dual. In this section we study these dual spaces.

First we will construct a basis of $\hat{V}$, from a given basis of V.

Theorem 3.13: Let $\{v_1, ..., v_n\}$ be a basis of V.

Define $\hat{v}_i : V \to F$ such that,

$\hat{v}_i (\alpha_1 v_1 + ... + \alpha_n v_n) = \alpha_i \quad i = 1, 2, ..., n$

Then $\hat{v}_i$ is a linear transformation for all $i = 1, ..., n$ and $\{\hat{v}_1, ..., \hat{v}_n\}$ is a basis of $\hat{V}$. Hence $dim\ V = dim\ \hat{V}$.

Proof: Let $v, v' \in V$

Suppose $v = \alpha_1 v_1 + ... + \alpha_n v_n$

$v' = \beta_1 v_1 + ... + \beta_n v_n, \quad \alpha_i, \beta_i \in F$

If $v = v'$, then $\alpha_j = \beta_j$ for all $j = 1, ..., n$

$\therefore \quad \hat{v}_i(v) = \alpha_i = \hat{v}_i(v')$

$\therefore \hat{v}_i$ is well defined for all $i = 1, ..., n$

Also $\hat{v}_i(v + v') = \hat{v}_i(\overline{\alpha_1 + \beta_1}v_1 + ... + \overline{\alpha_n + \beta_n}v_n)$

$= \alpha_i + \beta_i$

$= \hat{v}_i(v) + \hat{v}_i(v')$

and $\hat{v}_i(\alpha v) = \hat{v}_i(\alpha\alpha_1 v_1 + ... + \alpha\alpha_n v_n)$

$= \alpha\alpha_i = \alpha\hat{v}_i(v)$

$\therefore$ $\hat{v}_i$ is a *L.T.* for all $i = 1, ..., n$

By definition, $\hat{v}_i(v_j) = (0v_1 + ... + 1v_j + ... + 0v_n) = 0$ if $j \neq i$

$= 1$ if $j = i$

$\therefore$ $\hat{v}_i(v_j) = \delta_{ij}$ for all $i, j = 1, ..., n$

Let $\alpha_1\hat{v}_1 + ... + \alpha_n\hat{v}_n = 0 \quad \alpha_i \in F$

Then, $(\alpha_1\hat{v}_1 + ... + \alpha_n\hat{v}_n)(v_j) = 0(v_j) = 0$

$\Rightarrow$ $\alpha_j\hat{v}_j(v_j) = 0$

$\Rightarrow$ $\alpha_j = 0$ for all $j = 1, ..., n$

$\therefore$ $\{\hat{v}_1, ..., \hat{v}_n\}$ is *L.I.* over F.

Let $f \in \hat{V}$. Let $f(v_i) = \alpha_i \quad i = 1, ..., n$

Then $(\alpha_1\hat{v}_1 + ... + \alpha_n\hat{v}_n)(v_i)$

$= \alpha_i\hat{v}_i(v_i)$

$= \alpha_i \quad i = 1, ..., n$

$\therefore$ f and $\alpha_1\hat{v}_1 + ... + \alpha_n\hat{v}_n$ agree on all bases elements of V.

So, $f = \alpha_1\hat{v}_1 + ... + \alpha_n\hat{v}_n$

$\therefore$ $\{\hat{v}_1, ..., \hat{v}_n\}$ spans $\hat{V}$.

Hence, $\{\hat{v}_1, ..., \hat{v}_n\}$ is a basis of $\hat{V}$, called *dual basis* of $\{v_1, ..., v_n\}$ such that, $\hat{v}_i(v_j) = \delta_{ij}$.

Corollary : Let V be a finite dimensional vector space over F. Let $0 \neq v \in V$. Then $\exists f \in \hat{V}$ such that, $f(v) \neq 0$.

Proof: Since $v \neq 0$, $\{v\}$ is *L.I.* set. So, it can be extended to form a basis of V.

Let $\{v = v_1, v_2, ..., v_n\}$ be a basis of V.

Let $\{\hat{v}_1, ..., \hat{v}_n\}$ be corresponding dual basis. Then $\hat{v}_i(v_j) = \delta_{ij}$

$\therefore$ $\hat{v}_1(v_1) = 1$

Let $f = \hat{v}_1 \in \hat{V}$

Then $f(v) = f(v_1) = \hat{v}_1(v_1) = 1 \neq 0$.

Theorem 3.14: Let V be a finite dimensional vector space over F.

Define $\theta : V \to \hat{\hat{V}}$ such that,

$\theta(v) = T_v$ *for all* $v \in V$

where $T_v : \hat{V} \to F$ such that,

$T_v(f) = f(v)$ for all $f \in \hat{V}$

Then θ is an isomorphism from V onto $\hat{\hat{V}}$. (Here $\hat{\hat{V}}$ = dual of $\hat{V}$, called double dual of V).

Proof: Let $f, g \in \hat{V}$

Then $T_v(f + g) = (f + g)(v)$

$= f(v) + g(v)$

$= T_v(f) + T_v(g)$

Let $\alpha \in F$

Then $T_v(\alpha f) = (\alpha f)\ (v)$

NOTES

$= \alpha f(v)$

$= \alpha T_v(f)$

$\therefore \quad T_v \in \hat{\hat{V}}$

θ is well defined as $v = v' \Rightarrow T_v(f) = f(v)$

$= f(v') = T_{v'}(f)$ for all $f \in \hat{V} \Rightarrow T_v = T_{v'}$

θ is a $L.T.$ as

$$\theta(v + v') = T_{v+v'} = T_v + T_{v'} = \theta(v) + \theta(v')$$

as $\quad T_{v+v'}(f) = f(v + v')$

$= f(v) + f(v')$

$= T_v(f) + T_{v'}(f)$

$= (T_v + T_{v'})\ (f)$ for all $f \in \hat{V}$

$\mathrm{T}_{v+v'} = T_v + T_{v'}$

Also $\quad \theta(\alpha v) = T_{\alpha v} = \alpha T_v = \alpha\theta(v)$

as $\quad T_{\alpha v}(f) = f(\alpha v)$

$= \alpha f(v)$

$= \alpha T_v(f)$ for all $f \in \hat{V}$

$\therefore \quad T_{\alpha v} = \alpha T_v$

Let $0 \neq v \in \operatorname{Ker} \theta \Rightarrow \theta(v) = 0 \Rightarrow T_v = 0$

By Corollary to Theorem 3.13 $\exists\, f \in \hat{V}$ such that, $f(v) \neq 0$

$\therefore \quad T_v(f) \neq 0$

a contradiction as $T_v = 0 \Rightarrow T_v(f) = 0$

$\therefore \quad \operatorname{Ker} \theta = \{0\} \Rightarrow \theta$ is 1-1

$\therefore \quad V \cong \theta(V) \subseteq \hat{\hat{V}}$

$\Rightarrow \quad \dim \theta(V) = \dim V = \dim \hat{V} = \dim \hat{\hat{V}}$ (by Theorem 3.13)

$\therefore \quad \theta(V) = \hat{\hat{V}}$ as $\theta(V)$ is a subspace of $\hat{\hat{V}}$

$\therefore\ \theta$ is onto from V to $\hat{\hat{V}}$.

Thus θ is an isomorphism.

Corollary 1: Let V be a finite dimensional vector space over F. If L is a linear functional on $\hat{V}$, then $\exists$ a unique $v \in V$ such that, $L(f) = f(v)$ for all $f \in \hat{V}$.

Proof: L is a linear functional on $\hat{V}$

$\Rightarrow L \in \hat{\hat{V}} \Rightarrow \exists$ unique $v \in V$ such that,

$\theta\ (v) = L$ as θ is 1-1 onto

$\therefore \quad T_v = L$

$\Rightarrow L(f) = T_v(f) = f(v)$ for all $f \in \hat{V}$.

Corollary 2: Let V be a finite dimensional vector space over the field F. Then each basis for $\hat{V}$ is the dual of some basis for V.

Proof: Let $\{f_1, ..., f_n\}$ be a basis for $\hat{V}$.

NOTES

By Theorem 3.13, $\exists$ a basis $\{L_1, ..., L_n\}$ for $\hat{\hat{V}}$ s.t., $L_i(f_j) = \delta_{ij}$. As in Corollary 1 $\exists$ unique $v_i \in V$ for each i,

s.t., $L_i = T_{v_i} = \theta(v_i)$

Since $\{L_1, L_2, ..., L_n\}$ is a basis for $\hat{\hat{V}}$, $\{\theta^{-1} L_1, ..., \theta^{-1} L_n\} = \{v_1, ..., v_n\}$ is basis for V as θ is an isomorphism.

Also $\delta_{ij} = L_i(f_j) = T_{v_i}(f_j) = f_j(v_i)$

$\{f_1, ..., f_n\}$ is dual of basis $\{v_1, ..., v_n\}$ for V.

Example 3.21: Let V be the vector space of all polynomial functions from **R** to **R** which have degree less than or equal to 2, Let t_1, t_2, t_3 be three distinct real numbers and let $L_i : V \to F$ be such that, $L_i(p(x)) = p(t_i)$, $i = 1, 2, 3$. Show that $\{L_1, L_2, L_3\}$ is a basis of $\hat{V}$. Determine a basis for V such that, $\{L_1, L_2, L_3\}$ is its dual.

Solution: $L_i\,(p(x) + q(x))$

$= L_i(r(x)),\quad r(x) = p(x) + q(x)$

$= r(t_i) = p(t_i) + q(t_i)$

$= L_i(p(x)) + L_i(q(x))$

Also $L_i(\alpha p(x)),\quad \alpha \in F$

$= L_i(q(x)),\ q(x) = \alpha p(x)$

$= q(t_i)$

$= \alpha p(t_i) = \alpha\, L_i(p(x))\quad$ for all $i = 1, 2, 3$

$L_i \in \hat{V}\quad$ for all $i = 1, 2, 3$

Let $\quad \alpha_1 L_1 + \alpha_2 L_2 + \alpha_3 L_3 = 0$

Apply it on polynomials $1, x, x^2$ to get

$\alpha_1 + \alpha_2 + \alpha_3 = 0$

$\alpha_1 t_1 + \alpha_2 t_2 + \alpha_3 t_3 = 0$

$a_1 t_1^2 + \alpha_2 t_2^2 + \alpha_3 t_3^2 = 0$

$$\begin{bmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$A \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0,\ A = \begin{bmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{bmatrix}$$

$\det A = (t_1 - t_2)\,(t_2 - t_3)\,(t_3 - t_1)$

$\neq 0$ as t_1, t_2, t_3 are distinct

Thus A^{-1} exists.

NOTES

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0 \Rightarrow \alpha_1 = \alpha_2 = \alpha_3 = 0$$

Hence, $\{L_1, L_2, L_3\}$ is a *L.I.* set.

Since dim $V = 3$, $\{L_1, L_2, L_3\}$ is a basis of $\hat{V}$.

Let $\{p_1(x), p_2(x), p_3(x)\}$ be a basis of V such that, $\{L_1, L_2, L_3\}$ is its dual basis.

Then, $L_1(p_1) = 1,\ L_2(p_1) = 0,\ L_3(p_1) = 0$

$L_2(p_1) = 0 \quad \Rightarrow p_1(t_2) = 0$

$\Rightarrow t_2$ is a root of $p_1(x)$

$L_3(p_1) = 0 \quad \Rightarrow p_1(t_3) = 0$

$\Rightarrow t_3$ is a root of $p_1(x)$

Since, deg $p_1(x) \le 2$,

$p_1(x) = \alpha(x - t_2)\ (x - t_3), \quad \alpha = \text{Constant}$

$L_1\ (p_1) = 1 \quad \Rightarrow p_1(t_1) = 1$

$\Rightarrow \alpha(t_1 - t_2)\ (t_1 - t_3) = 1$

$$\Rightarrow \alpha = \frac{1}{(t_1 - t_2)\,(t_1 - t_3)}$$

$$\therefore \quad p_1(x) = \frac{(x - t_2)\,(x - t_3)}{(t_1 - t_2)\,(t_1 - t_3)}$$

Similarly, $p_2(x) = \dfrac{(x - t_1)\,(x - t_3)}{(t_2 - t_1)\,(t_2 - t_3)}$, $p_3(x) = \dfrac{(x - t_1)\,(x - t_2)}{(t_3 - t_1)\,(t_3 - t_2)}$.

Example 3.22: Let V be the vector space of all polynomial functions p from **R** into **R** which have degree 2 or less. Define three linear functionals on V by

$$f_1(p) = \int_0^1 p(x)\,dx, \quad f_2(p) = \int_0^1 p(x)\,dx,$$

$$f_3(p) = \int_0^{-1} p(x)\,dx$$

Show that $\{f_1, f_2, f_3\}$ is basis of $\hat{V}$. Determine a basis for V such that, $\{f_1, f_2, f_3\}$ is its dual basis.

Solution: It can be easily seen that $f_1, f_2, f_3 \in \hat{V}$.

Let $\alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3 = 0, \quad \alpha_i \in \mathbf{R}$

Apply it on $1, x, x^2$ to get

$$\alpha_1 + 2\alpha_2 - \alpha_3 = 0$$

$$\frac{\alpha_1}{2} + \frac{4}{2}\alpha_2 + \frac{\alpha_3}{2} = 0$$

$$\frac{\alpha_1}{3} + \frac{8}{3}\alpha_2 - \frac{\alpha_3}{3} = 0$$

NOTES

Let, $A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 4 & 1 \\ 1 & 8 & -1 \end{bmatrix}$

Then, $A\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0$, $\det A \neq 0$

$\therefore$ $A^{-1}A\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = 0 \Rightarrow \alpha_1 = \alpha_2 = \alpha_3 = 0$

$\therefore$ $\{f_1, f_2, f_3\}$ is a *L.I.* set.

Since $\dim V = 3$, $\{f_1, f_2, f_3\}$ is a basis of $\hat{V}$.

Let $\{p_1(x), p_2(x), p_3(x)\}$, be a basis of V such that, $\{f_1, f_2, f_3\}$ is its dual basis.

$\therefore$ $f_1(p_1) = 1, f_2(p_1) = 0, f_3(p_1) = 0$

Let $p_1(x) = c_o + c_1x + c_2x^2$

$$f_2(p_1) = 0 \Rightarrow c_o x + c_1\frac{x^2}{2} + c_2\frac{x^3}{3}\Bigg|_0^2 = 0$$

$$\Rightarrow c_o x + \frac{c_1}{2}x^2 + \frac{c_2}{3}x^3 = 0 \text{ when } x = 2$$

$$f_3(p_1) = 0 \quad \Rightarrow c_o x + c_1\frac{x^2}{2} + c_2\frac{x^3}{3}\Bigg|_0^{-1} = 0$$

$$\Rightarrow c_o x + \frac{c_1}{2}x^2 + \frac{c_2}{3}x^3 = 0 \text{ when } x = -1$$

$$\therefore \quad c_o x + \frac{c_1}{2}x^2 + \frac{c_2}{3}x^3 = \alpha x(x-2)(x+1)$$

$$f_1(p_1) = 1 \quad \Rightarrow c_o x + \frac{c_1}{2}x^2 + \frac{c_2}{3}x^3 = 1 \text{ when } x = 1$$

$$\Rightarrow \alpha \,.\, 1(-1)(2) = 1 \Rightarrow \alpha = -\frac{1}{2}$$

$$c_o x + \frac{c_1}{2}x^2 + \frac{c_2}{3}x^3 = -\frac{1}{2}x(x-2)(x+1)$$

$$= -\frac{1}{2}x^3 + \frac{1}{2}x^2 + x$$

NOTES

$$\therefore \quad \frac{c_2}{3} = -\frac{1}{2}, \frac{c_1}{2} = \frac{1}{2}, \quad c_o = 1$$

$$\therefore \quad c_o = 1, c_1 = 1, c_2 = -\frac{3}{2}$$

$$\therefore \quad p_1(x) = 1 + x - \frac{3}{2}x^2$$

Similarly, we can find $p_2(x), p_3(x)$.

Definition: Let W be a sub-set of V.

Define $\quad A(W) = \{f \in \hat{V} \mid f(w) = 0 \quad \text{for all } w \in W\}$

Then $A(W)$ is a sub-space of $\hat{V}$ as $\alpha, \beta \in F$,

$$\begin{aligned} f, g \in A(W) &\Rightarrow f(w) = 0 = g(w) \text{ for all } w \in W \\ &\Rightarrow \alpha f(w) + \beta g(w) = 0 \quad \text{for all } w \in W \\ &\Rightarrow (\alpha f + \beta g)\ (w) = 0 \quad \text{for all } w \in W \\ &\Rightarrow \alpha f + \beta g \in A(W) \end{aligned}$$

$A(W)$ is called *annihilator* of W.

Example 3.23: Let U, W be sub-sets of V. If $U \subseteq W$, show that $A(U) \subseteq A(W)$.

Solution: Let $f \in A(W)$ then,

$$\begin{aligned} & f(w) = 0 \quad \text{for all } w \in W \\ \Rightarrow\ & f(u) = 0 \quad \text{for all } u \in U \text{ as } U \subseteq W \\ \Rightarrow\ & f \in A(U). \end{aligned}$$

Theorem 3.15: Let V be a finite dimensional vector space and W, a subspace of V. Then $dim\ A(W) = dim\ V - dim\ W$.

Proof: Let $\{w_1, ..., w_m\}$ be a basis of W.

It can be extended to form a basis of V.

Let $\{w_1, ..., w_m, v_{m+1}, ..., v_n\}$ be a basis of V.

Let $\{f_1, ..., f_m, f_{m+1}, ..., f_n\}$ be corresponding dual basis.

Then $\quad f_i(w_j) = 0 \quad i = m + 1, ..., n$

$\qquad\qquad\qquad\qquad j = 1, ..., m$

$\therefore \quad f_i \in A(W) \quad \text{for all } i = m + 1, ..., n$

We show $\{f_{m+1}, ..., f_n\}$ is a basis of $A(W)$.

Let $\quad \alpha_{m+1} f_{m+1} + ... + \alpha_n f_n = 0$

$\therefore \quad (\alpha_{m+1} f_{m+1} + ... + \alpha_n f_n)\ (v_k) = 0 \quad \text{for all } k = m + 1,.., n$

$\therefore \quad \alpha_k f_k\ (v_k) = 0$

$\therefore \quad \alpha_k = 0 \quad \text{for all } k = m + 1, ..., n$

So, $\{f_{m+1}, ..., f_n\}$ is a $L.I.$ set.

Let $f \in A(W)$ then $f(w) = 0 \quad \text{for all } w \in W, f \in \hat{V}$

$$\begin{aligned} f \in \hat{V} &\Rightarrow f = \beta_1 f_1 + ... + \beta_m f_m + ... + \beta_n f_n \\ &\Rightarrow 0 = f(w_j) = \beta_j f_j(w_j) = \beta_j \quad \text{for all } j = 1, ..., m \\ &\Rightarrow f = \beta_{m+1} f_{m+1} ... + \beta_n f_n \end{aligned}$$

$$\Rightarrow \{f_{m+1}, ..., f_n\} \text{ spans } A(W)$$

$\therefore \{f_{m+1}, ..., f_n\}$ is a basis of $A(W)$.

Hence $\dim A(W) = n - m = \dim V - \dim W$.

Corollary 1: $\dfrac{\hat{V}}{A(W)} \cong \hat{W}$

Proof: Since $\dim \dfrac{\hat{V}}{A(W)} = \dim \hat{V} - \dim A(W)$

$$= \dim V - \dim V + \dim W$$

$$= \dim W = \dim \hat{W}$$

Hence, $\dfrac{\hat{V}}{A(W)} \cong \hat{W}$.

Corollary 2: If V is a finite dimensional vector space and W, a subspace of V, then

$$A(A(W)) \cong W.$$

Proof: Define $\theta : W \to A(A(W))$ such that,

$$\theta(w) = T_w$$

where $T_w : \hat{W} \to F$ such that,

$$T_w(f) \to f(w)$$

$$T_w \in A(A(W)) \text{ as } T_w(f) = f(w) = 0 \quad \text{for all } f \in A(W)$$

Then as in Theorem 3.14, θ is well defined 1-1 linear transformation.

$\therefore \quad W \cong \theta(W) \subseteq A(A(W))$

Since $\dim A(A(W)) = \dim \hat{V} - \dim A(W)$

$$= \dim V - \dim A(W)$$

$$= \dim W$$

(by Theorem 3.15)

and $\dim \theta(W) = \dim W$

$$A(A(W)) = \theta(W)$$

$\therefore$ θ is onto from W to $A(A(W))$

Hence $W \cong A(A(W))$.

For sake of convenience, we shall write $A(A(W)) = W$.

Consider for example, $V = \mathbf{R}^2$, $W = \{(x, 0) \mid x \in R\}$

Then $A(W)$ is a subspace of $\hat{V}$ spanned by f

where $f(x_1, x_2) = x_2$

In fact, $\{f\}$ is a basis of $A(W)$ as $\dim W = 1$.

Also, $A(A(W))$ is spanned by T_w where $w = (1, 0)$

Since $\dim A(A(W)) = 1$, $\{T_w\}$ is a basis of $A(A(W))$

NOTES

Then $\theta : W \to A(A(W))$ such that,

$$\theta(w) = T_w$$

is an isomorphism as basis of W is mapped to basis of $A(A(W))$.

Example 3.24: Let W_1, W_2 be subspaces of finite dimensional vector space V. Determine $A(W_1 + W_2)$.

Solution: $f \in A\,(W_1 + W_2)$

$$\Leftrightarrow \quad f(x) = 0 \quad \text{for all } x \in W_1 + W_2$$

$$\Leftrightarrow f(w_1) = 0 = f(w_2) \quad \text{for all } w_1 \in W_1, w_2 \in W_2$$

$$\Leftrightarrow \quad f \in A(W_1) \cap A(W_2)$$

$$\therefore \qquad A(W_1 + W_2) = A(W_1) \cap A(W_2).$$

Example 3.25: Let f_1, f_2, f_3 be three linear functionals on $\mathbf{R}^4$ defined as follows:

$$f_1(x_1, x_2, x_3, x_4) = x_1 + 2x_2 + 2x_3 + x_4$$

$$f_2(x_1, x_2, x_3, x_4) = 2x_2 + x_4$$

$$f_3(x_1, x_2, x_3, x_4) = -2x_1 - 4x_3 + 3x_4$$

Determine the subspace W *of* $\mathbf{R}^4$ such that,

$$f_i(w) = 0,\ w \in W \quad i = 1, 2, 3.$$

Solution: Let $(x_1, x_2, x_3, x_4) \in W$

Then $\quad f_i(x_1, x_2, x_3, x_4) = 0 \qquad i = 1, 2, 3$

$$\therefore \qquad x_1 + 2x_2 + 2x_3 + x_4 = 0$$

$$2x_2 = x_4 = 0$$

$$-2x_1 - 4x_3 + 3x_4 = 0$$

$$\therefore \qquad \begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 2 & 0 & 1 \\ -2 & 0 & -4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0$$

By elementary row transformations, we get

$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0$$

$$\therefore \qquad x_1 + 2x_3 = 0,\ x_2 = 0,\ x_4 = 0$$

$$\therefore \qquad (x_1, x_2, x_3, x_4) = (-2x_3, 0, x_3, 0) = x_3(-2, 0, 1, 0)$$

$\therefore$ W is spanned by $(-2, 0, 1, 0)$.

Example 3.26: Let W be the subspace of $\mathbf{R}^5$ spanned by the vectors

$$\alpha_1 = (2, -2, 3, 4, -1),\ \alpha_3 = (0, 0, -1, -2, 3)$$

$$\alpha_2 = (-1, 1, 2, 5, 2),\ \alpha_4 = (1, -1, 2, 3, 0)$$

Describe $A(W)$.

Solution: Let $f \in A(W)$

Then $\quad f(w) = 0 \quad$ for all $w \in W$

$\Rightarrow f(\alpha_i) = 0$ for all $i = 1, 2, 3, 4$

Let $\quad f(x_1, x_2, x_3, x_4, x_5) = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5$

(Note $v_1 = (1, 0, 0, 0, 0)$, $v_2 = (0, 1, 0, 0, 0)$, $v_3 = (0, 0, 1, 0, 0)$, $v_4 = (0, 0, 0, 1, 0)$, $v_5 = (0, 0, 0, 0, 1)$ form a basis of $\mathbf{R}^5$).

Let $\{\hat{v}_1, \hat{v}_2, \hat{v}_3, \hat{v}_4, \hat{v}_5\}$ be its dual basis.

Then $\quad f = c_1\hat{v}_1 + c_2\hat{v}_2 + c_3\hat{v}_3 + c_4\hat{v}_4 + c_5\hat{v}_5$

$$\Rightarrow f(x_1, x_2, x_3, x_4, x_5) = \sum_1^5 c_i\hat{v}_i\,(x_1, x_2, x_3, x_4, x_5)$$

$$= \sum_1^5 c_i\hat{v}_i\,(x_1v_1 + x_2v_2 + x_3v_3 + x_4v_4 + x_5v_5)$$

$$= c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5$$

as $\quad \hat{v}_i\,(v_j) = \delta_{ij}$

$\therefore \quad f(\alpha_i) = 0 \quad$ for all $i = 1, 2, 3, 4$

$$\Rightarrow \begin{bmatrix} 2 & -2 & 3 & 4 & -1 \\ -1 & 1 & 2 & 5 & 2 \\ 0 & 0 & 1 & -2 & 3 \\ 1 & -1 & 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix} = 0$$

By elementary row transformations, we get

$$\begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix} = 0$$

$\Rightarrow c_1 - c_2 - c_4 = 0, \; c_3 + 2c_4 = 0, \; c_5 = 0$

$\Rightarrow 2c_1 - 2c_2 + c_3 = 0, \; c_5 = 0, \; c_3 = -2c_4$

Let $\quad c_2 = a, \; c_4 = b$

Then $\quad c_3 = -2b$

$\quad 2c_1 - 2a - 2b = 0 \Rightarrow c_1 = a + b$

$\Rightarrow f(x_1, x_2, x_3, x_4, x_5) = (a + b)\,x_1 + ax_2 - 2bx_3 + bx_4$

Take $\quad a = 1, \; b = 0$

Then $\quad f_1(x_1, x_2, x_3, x_4, x_5) = x_1 + x_2$

Take $\quad a = 0, \; b = 1$

Then $\quad f_2(x_1, x_2, x_3, x_4, x_5) = x_1 - 2x_3 + x_4$

$\therefore \quad f = af_1 + bf_2$

$\therefore \{f_1, f_2\}$ spans $A(W)$

Let $\alpha f_1 + \beta f_2 = 0$. Apply it on v_1, v_2 respectively. We get $\alpha + \beta = 0$, $\alpha = 0 \Rightarrow \beta = 0$.

$\therefore$ $\{f_1, f_2\}$ is *L.I.* So, $\{f_1, f_2\}$ is a basis of $A(W)$

NOTES

Hence dim $A(W) = 2$.

Example 3.27: Let V be a finite dimensional vector space. Suppose $V = W_1 \oplus W_2$, where W_1, W_2 are subspaces of V. Show that $\hat{V} = A(W_1) \oplus A(W_2)$.

Solution: $\dim V = \dim (W_1 \oplus W_2)$

$= \dim W_1 + \dim W_2$

Also $\dim (A(W_1) \oplus A(W_2))$

$= \dim A(W_1) + \dim A(W_2)$

$= \dim V - \dim W_1 + \dim V - \dim W_2$

$= 2 \dim V - (\dim W_1 + \dim W_2)$

$= 2 \dim V - \dim V = \dim V = \dim \hat{V}$

Since $A(W_1) \oplus A(W_2)$ is a subspace of $\hat{V}$

and $\dim \hat{V} = \dim (A(W_1) \oplus A(W_2))$,

$\hat{V} = A(W_1) \oplus A(W_2)$.

Example 3.28: If f and g are in $\hat{V}$ such that, $f(v) = 0$ implies $g(v) = 0$, prove that $g = cf$ for some $c \in F$.

Solution: If $f = 0$, then $g = 0 = cf$ where $c = 0 \in F$.

Let $f \neq 0$ then $\exists\ v \neq 0$ in V such that, $f(v) \neq 0$

Let $\quad c = \dfrac{g(v)}{f(v)}$

$h = g - cf$ and $x \in V$

and $\quad \alpha = \dfrac{f(x)}{f(v)}$.

Then $\quad f(x - \alpha v) = f(x) - \alpha f(v) = 0$

$\Rightarrow x - \alpha v \in \text{Ker } f$

$\Rightarrow x - \alpha v = y \in \text{Ker } f$

$\Rightarrow x = y + \alpha v$

$\therefore \quad h(x) = g(x) - cf(x)$

$= g(y) + \alpha g(v) - cf(y) - c\alpha f(v)$

$= \alpha g(v) - c\alpha f(v)$ as $y \in \text{Ker } f \Rightarrow y \in \text{Ker } g$

$= \alpha g(v) - \alpha g(v) = 0$ for all $x \in V$

$\therefore \quad h = 0 \Rightarrow g = cf$

Hence the result follows.

Definition: Consider the system of m equations

$a_{11}x_1 + ... + a_{1n}x_n = 0$

$\ldots \quad \ldots \quad \ldots$

$$a_{m1}x_1 + ... + a_{mn}x_n = 0, \text{ where } a_{ij} \in F$$

in n unknowns.

Let U be the subspace of $F^{(n)}$ generated by m vectors

$u_1 = (a_{11}, ..., a_{1n}), ..., u_m = (a_{m1}, ..., a_{mn})$

If dim $U = r$, we say the system of equations has rank r.

NOTES

We determine the number of linearly independent solutions to the system of equations in $F^{(n)}$. Consider

Theorem 3.16: If the system of homogeneous linear equations

$$a_{11}x_1 + ... + a_{1n}x_n = 0$$

$$... \quad .. \quad ...$$

$$a_{m1}x_1 + ... + a_{mn}x_n = 0,$$

where $a_{ij} \in F$ is of rank r; then there are $n - r$ linearly independent solutions in $F^{(n)}$.

Proof: Let S be the set of solutions of the given system of equations

$$S = \{(\alpha_1, \alpha_2, ..., \alpha_n) \in F^n \mid \sum a_{ij}\,\alpha_j = 0, \quad i = 1, 2, .., m\}$$

Then S is a subspace of $F^n = V$

Let $\{v_1, v_2, ..., v_n\}$ be the standard basis of V

and $\{f_1, f_2, ..., f_n\}$ be its dual basis

Let U be the subspace of V as described above

Define $\theta : S \to A(U)$, such that,

$$\theta((\alpha_1, \alpha_2, ..., \alpha_n)) = \alpha_1 f_1 + \alpha_2 f_2 + ... + \alpha_n f_n$$

Let $\quad f = \alpha_1 f_1 + \alpha_2 f_2 + ... + \alpha_n f_n$

Then
$$\begin{aligned} f(u_1) &= (\alpha_1 f_1 + ... + \alpha_n f_n)\,(a_{11}v_1 + ... + a_{1n}v_n) \\ &= \alpha_1 a_{11} + ... \alpha_n a_{1n} \\ &= 0 \;\text{ as } (\alpha_1, ..., \alpha_n) \in S \end{aligned}$$

Similarly $f(u_2) = ... = f(u_m) = 0$

So $f \in A(U)$

It can be easily shown that θ is a linear transformation.

If $(\alpha_1, \alpha_2, ..., \alpha_n) \in \text{Ker } \theta$ then $\sum_1^n \alpha_i f_i = 0$

$\Rightarrow \alpha_i = 0 \quad \forall\, i$

$\Rightarrow \text{Ker } \theta = \{0\}$ or that θ is 1–1.

Let now $f \in A(U) \subseteq \hat{V}$

and suppose $f = \alpha_1 f_1 + \alpha_2 f_2 + ... + \alpha_n f_n$

Then, $0 = f(u_1) = \alpha_1 a_{11} + ... + \alpha_n a_{1n}$

$$... \qquad ... \qquad ...$$

$$0 = f(u_m) = \alpha_1 a_{m1} + ... + \alpha_n a_{mn}$$

$\therefore\ (\alpha_1, \alpha_2, ..., \alpha_n) \in S$

and $\theta\ ((\alpha_1, \alpha_2, ..., \alpha_n)) = \alpha_1 f_1 + ... + \alpha_n f_n = f$

or that θ is onto.

NOTES

Hence $S \cong A(U)$

$\Rightarrow \dim S = \dim A(U) = \dim V - \dim U$

$= n - r$

Hence there are $n - r$ linearly independent solutions of the given system of equations.

Corollary : If $n > m$, that is, if the number of unknowns exceed the number of equations, then the system of equations has a non zero solution.

Proof: Since U is generated by m vectors, $r = \dim U \leq m < n \Rightarrow n - r > 0 \Rightarrow$ system of equations has a linearly independent solution, which is non zero (as zero vector is not linearly independent).

Example 3.29: Let m and n be positive integers. Let $f_1, ..., f_m$ be linear functionals on $F^{(n)}$. For α in $F^{(n)}$ define $T(\alpha) = (f_1\ (\alpha), ..., f_m(\alpha))$.

Show that T is a linear transformation from $F^{(n)}$ into $F^{(m)}$. Then show that every linear transformation from $F^{(n)}$ into $F^{(m)}$ is of the above form, for some $f_1, ..., f_m$.

Solution: Since $f_1, ..., f_m$ are linear transformations, so is T. Let $\{e_1, ..., e_n\}$ be the standard basis of $F^{(n)}$.

Then $T(e_i) \in F^{(m)} \quad \forall\ i = 1, ..., n.$

So, $\quad T(e_i) = (\beta_{i1}, ..., \beta_{im}) \quad \forall\ i = 1, ..., n.$

$$\begin{aligned}
\therefore \quad T(\alpha) &= T(\alpha_1 e_1 + ... + \alpha_n e_n),\ \alpha = \alpha_1 e_1 + ... + \alpha_n e_n \\
&= \alpha_1 T(e_1) + ... + \alpha_n T\ (e_n) \\
&= \alpha_1(\beta_{11}, ..., \beta_{1m}) + ... + \alpha_n(\beta_{n1}, ..., \beta_{nm}) \\
&= (\alpha_1\beta_{11} + ... + \alpha_n\beta_{n1}, ..., \alpha_1\beta_{1m} + ... + \alpha_n\beta_{nm})
\end{aligned}$$

For each $i (1 \leq i \leq m)$, $\exists$ a linear transformation

$$f_i : F^{(n)} \to F \text{ such that,}$$

$$f_i(e_1) = \beta_{1i}, ..., f_i(e_n) = \beta_{ni}$$

$$\begin{aligned}
\therefore \quad f_1(\alpha) &= f_1(\alpha_1 e_1 + ... + \alpha_n e_n) \\
&= \alpha_1\beta_{11} + ... + \alpha_n\beta_{n1}
\end{aligned}$$

....

$$\begin{aligned}
f_m(\alpha) &= f_m\ (\alpha_1 e_1 + ... + \alpha_n e_n) \\
&= \alpha_1\beta_{1m} + ... + \alpha_n\beta_{nm}
\end{aligned}$$

So, $\quad T(\alpha) = (f_1(\alpha), ..., f_m(\alpha)).$

Example 3.30: Let V be the vector space of all 2×2 matrices over the field of real numbers and let

$$B = \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix}$$

Let W be the subspace of V consisting of all A such that $AB = 0$. Let f be a linear functional on V which is in the annihilator of W. Suppose that $f(I) = 0$ and $f(C) = 3$, where I is the 2×2 identity matrix and

$$C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Find $f(B)$.

Solution: Now $\quad W = \{A \mid AB = 0\}$

Let $\qquad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in V$

Then $\qquad A = a_{11}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_{21}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

$\therefore \qquad f(A) = a_{11}f\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12}f\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_{21}f\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22}f\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

$\qquad\qquad = a_{11}\alpha + a_{12}\beta + a_{21}\gamma + a_{22}\delta$ (say).

$\therefore \qquad 0 = f(I) = \alpha + \delta$

$\qquad\qquad 3 = f(C) = \delta$

So, $\qquad \alpha = -3,\ \delta = 3$

Let $\qquad D = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$. Then $DB = 0$

So, $D \in W$

$\Rightarrow f(D) = 0$ as $f \in A(W)$.

$\therefore \qquad 0 = \alpha + 2\beta \Rightarrow \beta = \dfrac{3}{2}$

Also, let $E = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}$.

Then $EB = 0$.

So, $E \in W$.

$\Rightarrow f(E) = 0$ as $f \in A(W)$

$\therefore\ \gamma + 2\delta = 0 \Rightarrow \gamma = -6$

So, $\quad f(B) = 2 \times (-3) + (-2)\left(\dfrac{3}{2}\right) + (-1)(-6) + (3)(1)$

$\qquad\qquad = -6 - 3 + 6 + 3 = 0.$

Example 3.31: Let F be a subfield of complex numbers. We define n linear functionals on $F^{(n)}$ $(n \geq 2)$ by

$$f_k(x_1, ..., x_n) = \sum_{j=1}^{n} (k - j)x_j,\ 1 \leq k \leq m.$$

NOTES

What is the dimension of the subspace annihilated by $f_1, ..., f_n$?

Solution: Now $f_1(x_1, ..., x_n) = ox_1 - x_2 - 2x_3 \; ... - (n-1)\, x_n$

$f_2(x_1, ..., x_n) = x_1 + ox_2 - x_3 \; ... - (n-2)\, x_n$

$f_3(x_1, ..., x_n) = 2x_1 + x_2 + ox_3 \; ... - (n-3)\, x_n$

....

$f_n(x_1, ..., x_n) = (n-1)\, x_1 + (n-2)\, x_2 + (n-3)\, x_3 + ... + 1\, x_{n-1} + ox_n$

Let W be the subspace of $F^{(n)}$ annihilated by $f_1, ..., f_n$.

Then $\quad (x_1, ..., x_n) \in W$

$\Rightarrow f_k(x_1, ..., x_n) = 0 \qquad \forall\; k = 1, 2, ..., n.$

$$\begin{bmatrix} 0 & -1 & -2 & ... & ... & -(n-1) \\ 1 & 0 & -1 & ... & ... & -(n-2) \\ 2 & 1 & 0 & ... & ... & -(n-3) \\ ... & ... & ... & ... & ... & ... \\ n-1 & n-2 & n-3 & ... & ... & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = 0$$

i.e., $AX = 0$, where A is the matrix on the left and $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

It can be easily seen that Rank $A = 2$.

$\therefore$ number of linear independent solutions in W is $n-2$.

$\therefore$ dim $W = n - 2$.

Transpose of a Linear Transformation

Let V, W be vector spaces over F.

Let T be a linear transformation from V into W.

Define $\quad T^t : \hat{W} \to \hat{V}$ such that,

$T^t(g) = gT$

Then T^t is a linear transformation called the *transpose* of T.

It can be easily shown that

(*i*) $(T_1 + T_2)^t = T_1^t + T_2^t$, where T_1, T_2 are linear transformations from V into W.

(*ii*) $(T_1T_2)^t = T_2^t\, T_1^t$, where $T_1 : W \to V$ and $T_2 : V \to W$ are linear transformations

(*iii*) $(\alpha T)^t = \alpha T^t$, $\alpha \in F$, $T : V \to W$ is a linear trasnsformation

(*iv*) $I^t = I$, $I : V \to V$ is the identity map.

Theorem 3.17: Let $T : V \to W$ be a linear transformation. Then

(*a*) The null space of T^t = the annihilator of range of T.

(*b*) If V, W are finite dimensional, then

(*i*) Rank of T = rank of T^t

(*ii*) Range of T^t = annihilator of the null space of T.

Proof: (*a*) Now $g \in$ Null space of T^t

$\Leftrightarrow T^t(g) = 0$

$\Leftrightarrow gT = 0 \Leftrightarrow gTV = 0 \quad \Leftrightarrow g(\text{Range } T) = 0$

$\Leftrightarrow g \in A(R_T)$

Where $A(R_T)$ denotes the annihilator of range T.

(*b*) Let dim $V = n$, dim $W = m$,

Let r = rank of T = dim R_T = dim $T(V)$

where R_T denotes the range of T.

Now dim$A(R_T)$ = dim $A(TV)$

= dim W – dim $T(V) = m - r$

Nullity of T^t = dimension of the null space of T^t

= dim $A(R_T) = m - r$

But nullity of T^t = dim W – rank T^t

dim W – rank T^t

$\Rightarrow \quad m - r = m -$ rank T^t

$\Rightarrow \quad$ rank $T^t = r =$ rank T

This proves (*i*).

Let N denote the null space of T.

Then $A(N) = \{f \in \hat{V} \mid f(n) = 0 \ \forall\ n \in N\}$ = Annihilator of the null space of T.

Now $f \in$ Range T^t

$\Rightarrow \quad f = T^t g, \quad g \in W$

$= gT$

$\Rightarrow \quad f(n) = gT(n) = g(0) = 0 \quad \forall\ n \in N$

$\Rightarrow \quad f \in A(N)$

$\Rightarrow$ Range$T^t \subseteq A(N)$

So, dim $A(N)$ = dim V – dim N

= dim V – nullity T = rank T

= rank T^t = dim Range T^t

Therefore, $A(N)$ = Range T^t

This proves (*ii*).

Lemma: Let $T : V \to W$ be a linear transformation. Let $\beta = \{v_1, ..., v_n\}$, $\beta' = \{w_1, ..., w_m\}$ be ordered basis of V, W respectively. Let $\hat{\beta} = \{f_1, ..., f_n\}$ be the dual basis of V such that $f_i(v_j) = \delta_{ij}$. Let $F \in \hat{V}$.

Then

$$f = \sum_1^n f(v_i)\, f_i$$

Proof: Suppose $f = \sum_{1}^{n} c_i f_i,\ c_i \in F$

NOTES

Then $\quad f(v_j) = \sum c_i f_i(v_j) = \sum c_i \delta_{ij} = c_j$

So, $\quad f = \sum_{1}^{n} f(v_i) f_i.$

Theorem 3.18: Let $T : V \to W$ be a linear transformation. Let $\beta = \{v_1, ..., v_n\}$, $\beta' = \{w_1, ..., w_m\}$ be ordered basis of V, W respectively. Let $\hat{\beta} = \{f_1, ..., f_n\}$, $\hat{\beta}' = \{g_1, .., g_m\}$ be the dual basis of V, W respectively.

Let $A = (a_{ij})$ be the matrix of T with respec to β, β' and $B = (b_{ij})$ be the matrix of T^t with respecy to, $\hat{\beta}'\ \hat{\beta}$.

Then $a_{ij} = b_{ji}\ \ \forall\ i, j.$

This shows that the matrix of T^t is the transpose of the matrix of T. For this reason T^t is called the transpose of T.

Proof: Now $T^t : \hat{W} \to \hat{V}$ such that,

$$T^t(g_j) = g_j T = f \text{ (say)}$$

Then $\quad f(v_i) = (T^t g_j)\ (v_i)$

$$= (g_j T)\ (v_i)$$

$$= (g_j T)\ (v_i) = g_j\left(\sum_{1}^{m} a_{ki}\ w_k\right)$$

$$= \sum a_{ki}\, g_j(w_k) = \sum a_{ki}\, \delta_{jk} = a_{ji}$$

By above lemma,

$$f = \sum_{1}^{n} f(v_j) f_i = \sum_{1}^{n} a_{ji} f_i$$

But $\qquad f \quad = T^t g_j = \sum_{1}^{n} b_{ij} f_i$

So, $\qquad \sum_{1}^{n} b_{ij} f_i = \sum_{1}^{n} a_{ji} f_i$

$\Rightarrow \qquad \sum_{1}^{n}\left(b_{ij} - a_{ji}\right) f_i = 0$

$\Rightarrow \qquad b_{ij} = a_{ji}\ \forall\ i, j$. This proves the theorem.

Let $A = (a_{ij})$ be the $m \times n$ matrix over F. Then *row rank* of A is defined as the dimension of the subspace of $F^{(n)}$ spanned by $(a_{11}, ..., a_{1n}), ..., (a_{m1}, ..., a_{mn})$.

Similarly, *column rank* of A is defined as the dimension of the subspace of $F^{(m)}$ spanned by $(a_{11}, a_{21}, ..., a_{m1}), ..., (a_{1n}, ..., a_{mn})$.

NOTES

Theorem 3.19: Let A be an $m \times n$ matrix over F. Then

Row rank of A = column rank of A.

Proof: Define $T : F^{(n)} \to F^{(m)}$ such that,

$$T((x_1, ..., x_n)) = (y_1, ..., y_m)$$

where $y_i = \sum_{j=1}^{n} a_{ij} x_j$

Then T is a linear transformation.

$$\begin{aligned} \text{Range } T &= \{T(x_1, ..., x_n) \mid x_i \in F\} \\ &= \{T(x_1 (1, ..., 0) + ... + x_n(0, -0, 1)) \mid x_i \in F\} \\ &= \{x_1 T(e_1) + ... + x_n T(e_n) \mid x_i \in F\} \end{aligned}$$

e_i = nth-tuple with ith co-ordinate 1 and zero elesewhere

= {linear combination of columns of A}

$\subseteq$ subspace generated by columns of A and vice-versa

Thus, Range T = subspace of $F^{(n)}$ generated by columns of A

So, Rank T = column rank of A

Also, Rank T^t = column rank of A^t

= Dimension of subspace of $F^{(m)}$ generated by columns of A^t

= Dmension of subspace generated by rows of A

= Row rank of A

Thus, column rank of A

= Row rank of A (as Rank T^t = Rank T)

= Rank T.

Example 3.32: Let V be a finite dimensional vector space over F. Let T be a linear operator on V. Let $c \in F$. Suppose $\exists\, 0 \neq v \in V$ such that $T(v) = cv$. Prove that there is a non zero linear functional f on V such that, $T^t f = cf$.

Solution: Now $(T - cI)v = 0$, $v \neq 0$

$\Rightarrow v \in \text{Ker } (T - cI)$

$\Rightarrow \text{Ker } (T - cI) \neq \{0\}$

$\Rightarrow \dim \text{Ker } (T - cI) \geq 1$

$\Rightarrow$ nullity of $(T - cI) \geq 1$

$\Rightarrow$ rank of $(T - cI) < n$

$\Rightarrow$ rank of $(T - cI)^t < n$

$\Rightarrow$ nulity of $(T - cI)^t \geq 1$

$\Rightarrow \exists f \in \hat{V}$ such that $f \neq 0$ and $(T - cI)^t f = 0$

$\Rightarrow T^t f = cf, f \neq 0$.

Example 3.33: Let A be $m \times n$ matrix with real entries. Prove that $A = 0 \Leftrightarrow$ *Trace* $(A^t A) = 0$.

NOTES

Solution: Let $A^t = B = (b_{ij})_{n \times m}$

$$A = (a_{jk})_{m \times n}$$

$$A^tA = BA = C = (c_{ik}), \quad c_{ik} = \sum_{j=1}^{m} b_{ij}\, a_{jk}$$

$$\text{Trace } (A^tA) = 0$$

$$\Rightarrow \sum_{1}^{n} c_{ii} = 0$$

$$\Rightarrow c_{11} + ... + c_{nn} = 0$$

$$\Rightarrow \sum_{1}^{m} b_{ij} a_{ji} + ... + \sum_{1}^{m} b_{nj} a_{jn} = 0$$

$$\Rightarrow \sum (a_{ji})^2 + ... + \sum (a_{jn})^2 = 0$$

$$\Rightarrow a_{ji} = 0 \quad \forall\ i, j$$

$$\Rightarrow A = 0.$$

Converse is obvious.

Quotient Spaces

If W be a subspace of a vector space $V(F)$ then since $< W, + >$ forms an abelian group of $< V, + >$, we can talk of cosets of W in V. Let $\frac{V}{W}$ be the set of all cosets $W + v$, $v \in V$, then we show that $\frac{V}{W}$ also forms a vector space over F, under the operations defined by

$$(W + x) + (W + y) = W + (x + y) \qquad x, y \in V$$

$$\alpha(W + x) = W + \alpha x \qquad \alpha \in F$$

Addition is well defined, since,

$$W + x = W + x'$$

$$W + y = W + y'$$

$$\Rightarrow x - x' \in W, y - y' \in W$$

$$\Rightarrow (x - x') + (y - y') \in W$$

$$\Rightarrow (x + y) - (x' + y') \in W$$

$$\Rightarrow W + (x + y) = W + (x' + y')$$

Again, $\quad W + x = W + x'$

$$\Rightarrow x - x' \in W,$$

$$\Rightarrow \alpha(x - x') \in W \qquad \alpha \in F$$

$$\Rightarrow \alpha x - \alpha x' \in W$$

$$\Rightarrow W + \alpha x = W + \alpha x'$$

$$\Rightarrow \alpha(W + x) = \alpha(W + x')$$

NOTES

Thus, scalar multiplication is also well defined. It should now be a routine exercise to check that all conditions in the definition of a vector space are satisfied.

$W + 0$ will be zero of $\dfrac{V}{W}$

$W - x$ will be inverse of $W + x$

Also

$$\alpha((W + x) + (W + y)) = \alpha(W + (x + y)) = W + \alpha(x + y) = W + (\alpha x + \alpha y)$$

$$= (W + \alpha x) + (W + \alpha y) = \alpha(W + x) + \alpha(W + y) \text{ etc.}$$

Hence, V/W forms a vector space over F, called the quotient space of V by W.

Check Your Progress

1. What is normed linear space?
2. How are 2- or 3-dimensional vectors defined through real valued entries?
3. Define the terms seminormed and normed vector spaces.
4. What is linear transformation?
5. Explain about the dim mn?
6. When is a linear transformation $T: V \to W$ non singular?

3.3 BANACH SPACES

The Hahn-Banach theorem is an essential tool in functional analysis. It permits the extension of bounded linear functionals defined on a subspace of some vector space to the complete space and also illustrates that there are 'Enough' continuous linear functionals defined on every normed vector space for studying the dual space. It is named for Hans Hahn and Stefan Banach who proved this theorem independently and a general extension theorem from which the Hahn-Banach theorem can be derived was proved in 1923 by Marcel Riesz.

The most general formulation of the theorem can be given for a vector space V over the field **R** of real numbers where a function $f : V \to \mathbf{R}$ is called sublinear if,

$f(\gamma x) = \gamma f(x)$ for any $\gamma \in$ R+and any $x \in V$ (Positive homogeneity).

$f(x + y) \leq (x) + f(y)$ for any $x, y \in V$ (Subadditivity).

Every seminorm on V (specifically, every norm on V) is sublinear. The Hahn-Banach theorem states that if $\mathcal{N}: V \to$ R is a sublinear function and φ : U

$\to \mathbb{R}$ is a linear functional on a linear subspace $U \subseteq V$ which is dominated by $\mathcal{N}$ on U,

$$\varphi(x) \leq \mathcal{N}(x) \qquad \forall x \in U$$

NOTES

Then there exists a linear extension $\psi : V \to$ R of φ to the whole space V, i.e., there exists a linear functional ψ such that,

$$\psi(x) = \varphi(x) \qquad \forall x \in U$$

and

$$\psi(x) \leq \mathcal{N}(x) \quad \forall x \in V.$$

Another description of Hahn-Banach theorem states that if V is a vector space over the scalar field **K** (either the real numbers **R** or the complex numbers **C**), if $\mathcal{N} : V \to \mathbb{R}$ is a seminorm and $\varphi : U \to \mathbb{K}$ is a **K**-linear functional on a **K**-linear subspace U of V which is dominated by on U in absolute value,

$$|\varphi(x)| \leq \mathcal{N}(x) \qquad \forall x \in U$$

Then there exists a linear extension $\psi : V \to \mathbb{K}$ of φ to the whole space V, i.e., there exists a **K**-linear functional ψ such that,

$$\psi(x) = \varphi(x) \qquad \forall x \in U$$

and

$$|\psi(x)| \leq \mathcal{N}(x) \qquad \forall x \in V.$$

Banach Spaces

A Banach space is a complete normed vector space or a Banach space is a vector space which is equipped with a norm and which is complete with respect to that norm. Two common types of Banach spaces are real Banach spaces and complex Banach spaces, which are Banach spaces whose underlying vector spaces are defined over the field of real numbers or complex numbers, respectively.

Various infinite dimensional function spaces evaluated in analysis are Banach spaces, including spaces of continuous functions (continuous functions on a compact Hausdorff space), spaces of Lebesgue integrable functions known as L^p spaces and spaces of holomorphic functions known as Hardy spaces. These are the most commonly used topological vector spaces and their topology is based on a norm.

A metric space X is considered as complete if every Cauchy sequence in X converges to a point in X. Normed spaces whose induced metric spaces are complete are specified with a special name. A Banach space is a normed space whose induced metric space is complete.

The following normed spaces are all Banach spaces:

l_p, L^p, $\mathbf{R}^n$, C^n and C [a, b]

A closed vector space of a Banach space is itself a Banach Space.

Let **K** stand for one of the fields **R** or **C**. The known Euclidean spaces $\mathbf{K}^n$, where the Euclidean norm of $x = (x_1, \ldots, x_n)$ is given by $\|x\| = (x_1^2 + \ldots + x_n^2)^{1/2}$ are termed as Banach spaces. Hence every finite dimensional **K** vector space becomes a Banach space being endowed with an arbitrary norm because all norms are equivalent on a finite dimensional **K** vector space.

Consider the space of all continuous functions $f : [a, b] \rightarrow \mathbf{K}$ defined on a closed interval $[a, b]$. This space becomes a Banach space if an appropriate norm $\|f\|$ is defined in it. Such a norm may be defined as $\|f\| = \sup \{|f(x)| : x. [a, b]\}$ known as the supremum norm. This is a well defined norm because continuous functions defined on a closed interval are bounded.

Since f is a continuous function on a closed interval then it is bounded and the supremum in the above definition is obtained using the Weierstrass extreme value theorem. Hence, we can replace the supremum by the maximum. In this case, the norm is also called the **maximum norm**.

The space is complete under this norm and the resulting Banach space is denoted by $C[a, b]$. This example can be generalized to the space $C(X)$ of all continuous functions $X \rightarrow \mathbf{K}$, where X is a compact space, or to the space of all bounded continuous functions $X \rightarrow \mathbf{K}$, where X is any topological space or indeed to the space B(X) of all bounded functions $X \rightarrow \mathbf{K}$, where X is any set.

For any open set Ω **C**, the set $A(\Omega)$ of all bounded, analytic functions $u : \Omega \rightarrow \mathbf{C}$ is a complex Banach space with respect to the supremum norm.

If $p \geq 0$ is a real number, we can consider the space of all infinite sequences $(x_1, x_2, x_3, \ldots)$ of elements in **K** such that the infinite series $\sum_i |x_i|^p$ is finite. The p-th root of this series' value is then defined to be the p-norm of the sequence. The space, together with this norm, is a Banach space; it is denoted by ℓ^p. The Banach space ℓ^∞ consists of all bounded sequences of elements in **K**; the norm of such a sequence is defined to be the supremum of the absolute values of the sequence's members.

Again, if $p \geq 1$ is a real number, we can consider all functions $f : [a, b] \rightarrow \mathbf{K}$ such that $|f|^p$ is Lebesgue integrable. The p-th root of this integral is then defined to be the norm of f. By itself, this space is not a Banach space because there are non-zero functions whose norm is zero. We define an equivalence relation as follows: f and g are equivalent if and only if the norm of $f-g$ is zero. The set of equivalence classes then forms a Banach space; it is denoted by $L^p([a, b])$. It is crucial to use the Lebesgue integral and not the Riemann integral here, because the Riemann integral would not yield a complete space. These examples can be generalized; see L^p spaces for details.

If X and Y are two Banach spaces, then we can form their direct sum $X . Y$, which has a natural topological vector space structure but no canonical norm. However, it is again a Banach space for several equivalent norms, for example

$$\|x \otimes y\| = \left(\|x\|^p + \|y\|^p\right)^{1/p}, \qquad 1 \leq p \leq \infty$$

This construction can be generalized to define ℓ^p-direct sums of arbitrarily many Banach spaces. When there is an infinite number of non-zero summands, the space obtained in this way depends upon p.

If M is a closed linear subspace of the Banach space X, then the quotient space X/M is again a Banach space.

NOTES

The **L^p spaces** are function spaces defined using a natural generalization of the p-norm for finite-dimensional vector spaces. They are sometimes called **Lebesgue spaces**, named after Henri Lebesgue. The length of a vector $x = (x_1, x_2, \ldots, x_n)$ in the n-dimensional real vector space $\mathbf{R}^n$ is usually given by the Euclidean norm:

$$\|x\|_2 = \left(x_1^2 + x_2^2 + \cdots + x_n^2\right)^{1/2}.$$

The Euclidean distance between two points x and y is the length $\|x-y\|_2$ of the straight line between the two points. In many situations, the Euclidean distance is insufficient for capturing the actual distances in a given space.

For a real number $p \geq 1$, the ***p*-norm** or ***L^p*-norm** of x is defined by,

$$\|x\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}.$$

The ***L$^\infty$*-norm** or maximum norm (or uniform norm) is the limit of the L^p-norms for $p \to \infty$. It turns out that this limit is equivalent to the following definition:

$$\|x\|_\infty = \max\left\{|x_1|, |x_2|, \ldots, |x_n|\right\}.$$

For all $p \geq 1$, the p-norms and maximum norm as defined above indeed satisfy the properties of a 'Length Function' or norm, which specify that:

- Only the zero vector has zero length.
- The length of the vector is positive homogeneous with respect to multiplication by a scalar.
- The length of the sum of two vectors is no larger than the sum of lengths of the vectors (By triangle inequality).

Abstractly speaking, this means that $\mathbf{R}^n$ together with the p-norm is a Banach space. This Banach space is the ***L^p*-space** over $\mathbf{R}^n$.

For example,

For $1 \leq p < \infty$, we define the p-norm on $\mathbb{R}^n$ (or $\mathbb{C}^n$) by

$$\|(x_1, x_2, \ldots, x_n)\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}.$$

For $p = \infty$, we define the ∞, or maximum, norm by

$$\|(x_1, x_2, \ldots, x_n)\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}.$$

Then $\mathbb{R}^n$ equipped with the p-norm is a finite-dimensional Banach space for $1 \leq p \leq \infty$.

For example,

The space $C([a,b])$ of continuous, real-valued (or complex-valued) functions on $[a,b]$ with the sup-norm is a Banach space. In general, the space $C(K)$ of continuous functions on a compact metric space K equipped with the sup-norm is a Banach space.

NOTES

For example,

The space $C^k([a,b])$ of k-times continuously differentiable functions on $[a,b]$ is not a Banach space with respect to the sup-norm $\|\cdot\|_\infty$ for $k \geq 1$, since the uniform limit of continuously differentiable functions need not be differentiable. We define the C^k-norm by

$$\| f \|_{C^k} = \| f \|_\infty + \| f' \|_\infty + \cdots + \| f^{(k)} \|_\infty .$$

Then $C^k([a,b])$ is a Banach space with respect to the C^k-norm. Convergence with respect to the -norm is uniform convergence of functions and their first derivatives.

For example,

For $1 \leq p < \infty$, the sequence space $\ell^p(\mathbb{N})$ consists of all infinite sequences $x = (x_n)_{n=1}^{\infty}$ such that

$$\sum_{n=1}^{\infty} |x_n|^p < \infty$$

with the p-norm,

$$\| x \|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}$$

For $p = \infty$, the sequence space $\ell^\infty(\mathbb{N})$ consists of all bounded sequences, with

$$\| x \|_\infty = \sup\{|x_n| \mid n = 1,2,\ldots\}.$$

Then $\ell^p(\mathbb{N})$ is an infinite-dimensional Banach space for $1 \leq p \leq \infty$. The sequence space $\ell^p(\mathbb{Z})$ of bi-infinite sequences $x = (x_n)_{n=-\infty}^{\infty}$ is defined in an analogous way.

3.3.1 Conjugate Spaces

The complex conjugate of a complex vector space V is the complex vector space $\overline{V}$ consisting of all formal complex conjugates of elements of V, i.e., $\overline{V}$ is a vector space whose elements are in one-to-one correspondence with the elements of V:

$$\overline{V} = \{\overline{v} \mid v \in V\},$$

It implies the following rules for addition and scalar multiplication:

$$\overline{v} + \overline{w} = \overline{v + w} \quad \text{and} \quad \alpha\overline{v} = \overline{\overline{\alpha}v}.$$

Here v and w are vectors in V, α is a complex number and $\overline{\alpha}$ denotes the complex conjugate of α. In the case where V is a linear subspace of $\mathbb{C}^n$, the formal complex conjugate $\overline{V}$ is obviously isomorphic to the real complex conjugate subspace of V in $\mathbb{C}^n$.

NOTES

Conjugate Linear Maps

Any linear map $f: V \to W$ induces a conjugate linear map $\overline{f}: \overline{V} \to \overline{W}$ defined by the formula,

$$\overline{f}(\overline{v}) = \overline{f(v)}.$$

The conjugate linear map $\overline{f}$ is linear. Furthermore, the identity map on V induces the identity map $\overline{V}$ and the following expression:

$$\overline{f} \circ \overline{g} = \overline{f \circ g}$$

This holds for any two linear maps f and g. Therefore, the rules $V \mapsto \overline{V}$ and $f \mapsto \overline{f}$ define a category of complex vector spaces to itself.

If V and W are finite dimensional and the map f is described by the complex matrix A with respect to the bases $\mathcal{B}$ of V and $\mathcal{C}$ of W then the map $\overline{f}$ is described by the complex conjugate of A with respect to the bases $\overline{\mathcal{B}}$ of $\overline{V}$ and $\overline{\mathcal{C}}$ of $\overline{W}$.

Structure of the Conjugate

The vector spaces V and $\overline{V}$ have the same dimension over the complex numbers and are therefore isomorphic as complex vector spaces. Though there is no standard isomorphism from V to $\overline{V}$. This implies that the map C is not an isomorphism, because it is antilinear.

The double conjugate $\overline{\overline{V}}$ is naturally isomorphic to V with the isomorphism $\overline{\overline{V}} \to V$ defined by,

$$\overline{\overline{v}} \mapsto v.$$

Typically the double conjugate of V is simply identified with V.

Conjugate of a Hilbert Space

Given a Hilbert space $\mathcal{H}$ (either finite or infinite dimensional), its complex conjugate $\overline{\mathcal{H}}$ is the same vector space as its continuous dual space $\mathcal{H}'$. There is one-to-one antilinear association between continuous linear functionals and vectors. Alternatively, any continuous linear functional on $\mathcal{H}$ is an inner multiplication to some fixed vector and vice versa.

3.3.2 Natural Embedding of a Normed Linear Space in its Second Dual

Determining if two given spaces are homeomorphic is one of the fundamental problems in topology.

Definition: A one-one and onto (bijection) continuous map $f: X \to Y$ is a homeomorphism if its inverse is continuous.

A bijection $f: X \to Y$ induces a bijection between subsets of X and subsets of Y and it is a homeomorphism iff this bijection restricts to a bijection,

$$\{\text{Open (or closed) subsets of } X\} \frac{U \to f(U)}{f^{-1}(V) \leftarrow V} \{\text{Open (or closed) subsets of } Y\}$$

between open (or closed) subsets of X and open (or closed) subsets of Y.

Definition: Suppose X is a set, Y a topological space and $f: X \to Y$ an injective map. The embedding topology on X (for the map f) is the collection,

$f^{-1}(\mathcal{T}_Y) = \{f^{-1}(V) | V \subset Y \text{ open}\}$ of subsets of X.

The subspace topology for $A \subset X$ is the embedding topology for the inclusion map $A \to X$.

Theorem 3.20 (Characterization of the Embedding Topology): Let X has the embedding topology for the map $f: X \to Y$. Then,

1. $X \to Y$ is continuous.
2. For any map $A \to X$ into X,

   $A \to X$ is continuous iff $A \to X \xrightarrow{f} Y$ is continuous.

The embedding topology is the only topology on X with these two properties. The embedding topology is the most common topology on X such that $f: X \to Y$ is continuous.

Proof: The reason is that $A \xrightarrow{g} X$ is continuous.

$$\Leftrightarrow g^{-1}(\mathcal{T}_X) \subset \mathcal{T}_A \Leftrightarrow g^{-1}(f^{-1}\mathcal{T}_Y) \subset \mathcal{T}_A \Leftrightarrow (fg)^{-1} (\mathcal{T}_Y) \subset \mathcal{T}_A \Leftrightarrow A \xrightarrow{g} X \xrightarrow{f} Y$$

is continuous by definition of the embedding topology. The identity map of X is a homeomorphism whenever X is equipped with a topology with these two properties.

Definition: An injective continuous map $f: X \to Y$ is an embedding if the topology on X is the embedding topology for f, i.e., $\mathcal{T}_X = f^{-1}\mathcal{T}_Y$.

Any injective map $f: X \to Y$ induces a bijection between subsets of X and subsets of $f(X)$ and it is an embedding iff this bijection restricts to a bijection,

$$\{\text{Open (or closed) subsets of } X\} \frac{U \to f(U)}{f^{-1}(V) \leftarrow V} \{\text{Open (or closed) subsets of } f(X)\}$$

between open (or closed) subsets of X and open (or closed) subsets of $f(X)$.

Alternatively, the injective map $f: X \to Y$ is an embedding iff the bijective corestriction $f(X)| f: X \to f(X)$ is a homeomorphism. An embedding is a homeomorphism followed by an inclusion. The inclusion $A \to X$ of a subspace is an embedding. Any open (or closed) continuous injective map is an embedding.

NOTES

For example, the map $f(x) = 3x + 1$ is a homeomorphism from $\mathbf{R} \to \mathbf{R}$.

Lemma: If $f: X \to Y$ is a homeomorphism (embedding) then the corestriction of the restriction $f(A)| f \setminus A: A \to f(A)$ $(B| f \setminus A: A \to B)$ is a homeomorphism (embedding) for any subset A of X (and any subset B of Y containing $f(A)$). If the maps $f_j: X_j \to Y_j$ are homeomorphisms (embeddings) then the product map $\prod f_j : \prod X_j \to Y_j$ is a homeomorphism (embedding).

Proof: In case of homeomorphisms employ that there is a continuous inverse in both cases. In case of embeddings, employ that an embedding is a homeomorphism followed by an inclusion map.

Lemma (Composition of Embeddings): Let $X \xrightarrow{f} Y \xrightarrow{g} Z$ be continuous maps. Then f and g are embeddings implies that $g \circ f$ is an embedding which in turn implies that f is an embedding.

Proof: For proving the second implication, first note that f is an injective continuous map. Let $U \subset X$ be open. Since $g \circ f$ is an embedding, $U = (g \circ f)^{-1} W$ for some open $W \subset Z$. But $(g \circ f)^{-1} = f^{-1} g^{-1} W$ where $g^{-1} W$ is open in Y since g is contimuous. This shows that f is an embedding.

Theorem 3.21 (Characterization of the Product Topology): Given the product topology $\prod Y_j$. Then,

1. The projections $\pi_j : \prod Y_j \to Y_j$ are continuous, and
2. For any map $f: X \to \prod_{j \in J} Y_j$ into the product space we have,

$$X \xrightarrow{f} \prod_{j \in J} Y_j \text{ is continuous} \Leftrightarrow \forall\, j \in J: X \xrightarrow{f} \prod_{j \in J} Y_j \xrightarrow{\pi_j} Y_j \text{ is continuous.}$$

The product topology is the only topology on the product set with these two properties.

Proof: Let $\mathbf{T}_X$ be the topology on X and $\mathbf{T}_j$ the topology on Y_j. Then $S_\Pi = \bigcup_{j \in J} \pi_j^{-1}(\mathcal{T}_j)$ is a subbasis for the product topology on $\prod_{j \in J} Y_j$. Therefore,

$$f: X \to \prod_{j \in J} Y_j \text{ is continuous} \Leftrightarrow f^{-1}\left(\bigcup_{j \in J} \pi_j^{-1}(\mathcal{T}_j)\right) \subset \mathcal{T}_X$$

$$\Leftrightarrow \left(\bigcup_{j \in J} f^{-1}(\pi_j^{-1}(\mathcal{T}_j))\right) \subset \mathrm{T}_X$$

$$\Leftrightarrow \forall\, j \in J:\ (\pi_j \circ f)^{-1}(\mathcal{T}_j) \subset \mathcal{T}_X$$

$$\Leftrightarrow \forall\, j \in J:\ \pi_j \circ f \text{ is continuous by definition of continuity}$$

Now, we have to show that the product topology is the unique topology with these properties. Take two copies of the product set $\prod_{j\in J} X_j$. Provide one copy with the product topology and the other copy with some topology that has the two properties of the above theorem. Then the identity map between these two copies is a homeomorphism.

Theorem 3.22: Let $(X_j)_{j\in J}$ be an indexed family of topological spaces with subspaces $A_j \subset X_j$. Then $\prod_{j\in J} A_j$ is a subspace of $\prod_{j\in J} X_j$.

1. $\overline{\prod A_j} = \prod \overline{A_j}$
2. $(\prod A_j)^\circ \subset \prod A_j^\circ$ and equality holds if $A_j = X_j$ for all but finitely many $j \in J$.

Proof: (1) Let (x_j) be a point of $\prod X_j$. Since $S_\Pi = \bigcup_{j\in J} \pi_j^{-1}(\mathcal{T}_j)$ is a subbasis for the product topology on $\prod X_j$, we have

$(x_j) \in \overline{\prod A_j} \Leftrightarrow \forall\, k \in J$: $\pi_k^{-1}(U_k) \cap \prod A_j \neq \varphi$ for all neighbourhoods U_k of x_k.

$\Leftrightarrow \forall\, k \in J$: $U_k \cap A_k \neq \phi$ for all neighbourhoods U_k of x_k

$\Leftrightarrow \forall\, k \in J$: $x_k \in \overline{A_k}$

$\Leftrightarrow (x_j) \prod \overline{A_j}$

(2) $(\prod A_j)^\circ \subset \prod A_j^\circ$ because π_j is an open map so that $\pi_j((\prod A_j)^\circ) \subset A_j^\circ$ for all $j \in J$. If $A_j = X_j$ for all but finitely many $j \in J$ then $\prod A_j^\circ \subset (\prod A_j)^\circ$ because $\prod A_j^\circ$ is open and contained in $\prod A_j$.

It follows that a product of closed sets is closed.

Note: A product of open sets need not be open in the product topology.

3.3.3 Embedding Lemma and Tychonoff Embedding

Theorem 3.23 (Embedding Lemma): Let $\mathcal{F}$ be a family of mappings where each member $f \in \mathcal{F}$ maps $X \to Y_f$. Then,

1. The evaluation mapping $e: X \to \prod_{f\in\mathcal{F}} Y_f$ defined by $\pi_f o e(x) = f(x)$, for all $x \in X$, is continuous.
2. The mapping e is an open mapping onto $e(X)$ if $\mathcal{F}$ distinguishes points and closed sets.

3. The mapping e is one-to-one if and only if $\mathcal{F}$ distinguishes points.
4. The mappping e is an embedding if $\mathcal{F}$ distinguished points F distinguishes points and closed sets.

NOTES

Proof: (1) Let $\pi_g: \prod_{f\in\mathcal{F}} Y_f \to Y_g$ be the projection map to the space Y_g. Then $\pi_g oe = g$ so that $\pi_g oe$ is continuous. Therefore e must be continuous as g is continuous.

(2) Suppose that U is open in X and $x \in U$. Choose $f \in\mathcal{F}$ such that $f(x) \notin \overline{f(X\setminus U)}$. The set $B = \{z \in e(X) | \pi_f(z) \notin \overline{f(X\setminus U)}\}$ is a neighbourhood of $e(x)$ as the set is open (it is defined for components not being in the closed set $\overline{f(X\setminus U)}$ and clearly $e(x) \in B$. Moreover $\pi_f(B) \subset f(U)$ by construction. It is now claimed that $f(U) \subset \pi_f(B)$. This follows trivially from the definition of a family of functions distinguishing points and closed sets. Therefore $f(U) = \pi_f(B)$ and subsequently $f(U)$ is an open subset of $\pi_g oe(X)$. Therefore the evaluation map is an open mapping.

(3) The definition of distinguishing points implies injectivity.

(4) Combining a, b and c, we see that $X \cong e(X)$ as e is a continuous, open, injective, surjective (as a continuous map is always surjective onto its image) map.

Definition: If X is a space and A, a set then by the power X^A we mean the product space $\Pi_\alpha X_\alpha$, where $X_\alpha = X$, for each $\alpha \in A$. Any power of $[0, 1]$ is called a cube. A map e: $X \to Y$ is an embedding iff the map e: $X \to e(X)$ is a homeomorphism. If there is an embedding e: $X \to Y$ then we say that X can be embedded in Y.

Theorem 3.24 (Tychonoff's Embedding Theorem): A space is Tychonoff iff it can be embedded in a cube.

Proof: $\Rightarrow$ Let X be a Tychonoff space and let $A = \{f: X \to [0, 1] / f \text{ is continuous}\}$. Define e: $X \to [0, 1]^A$ by $e(x)(f) = f(x)$.

(*i*) e is injective: If $x, y \in X$ with $x \neq y$, then there is $f \in A$ so that $f(x) = 0$ and $f(y) = 1$. Then $e(x)(f) \neq e(y)(f)$, so $e(x) \neq e(y)$.

(*ii*) e is continuous: This is immediate since $\pi_f e = f$.

(*iii*) e: $X \to e(X)$ carries open sets of X to open subsets of $e(X)$: For let U be open in X and let $x \in U$. Then there is $f \in A$ so that $f(x) = 0$ and $f(X - U) = 1$. Let $V = \pi_f^{-1}([0, 1))$, an open subset of $[0, 1]^A$. Then $e(x) \in V$ and if $y \in X$ is such that $e(y) \in V$, then $e(y)(f) \in [0, 1)$, so $f(y) < 1$ and $y \in U$. Thus $e(x) \in V \cap e(X) \subset e(U)$.

(*i*), (*ii*) and (*iii*) together imply that e is an embedding.

$\Leftarrow$: [0, 1] is clearly so $[0, 1]^A$ is Tychonoff for any A. Any subspace of a Tychonoff space is Tychonoff. Thus if X can be embedded in a cube, then X is homeomorphic to a Tychonoff space and so is itself Tychonoff.

NOTES

Theorem 3.25: Let $(\mathbf{T}, \mathcal{T})$ be the 3-point topological space defined by $\mathbf{T} = \{0, 1, 2\}$ and $\mathcal{T} = \{\varphi, \{0\}, \mathbf{T}\}$. Let $(X, \mathcal{U})$ be any topogical space and suppose that $\mathcal{U} \cap X = \varphi$. Then there is an embedding $e: X \to \mathcal{T}^{\mathcal{U} \cup X}$.

Proof: For each $U \in \mathcal{U}$, define $f_U: X \to \mathbf{T}$ by $f_U(y) = 0$ if $y \in U$ and $f_U(y) = 1$ if $y \notin U$. Then f_U is continuous. For each $x \notin X$, define $f_x: X \to \mathbf{T}$ by $f_x(y) = 2$ if $y = x$ and $f_x(y) = 1$ if $y \neq x$. Then f_x is also continuous.

Define e by $e_i(y) = f_i(y)$ for each $i \in \mathcal{U} \cup X$. Then

(*i*) e is injective, for if $x, y \in X$ with $x \neq y$ then $e_x(y) = 1$ but $e_x(x) = 2$, so $e_x(x) \neq e_x(y)$ and hence $e(x) \neq e(y)$.

(*ii*) e is continuous because each f_i is continuous.

(*iii*) e is open into $e(X)$, for if $U \in \mathcal{U}$ and $x \in U$ then $V = \pi_U^{-1}(0)$ is open in $\mathbf{T}^{\mathcal{U} \cup X}$. Furthermore, so $\pi_U e(x) = 0$, so $e(x) \in V$ while if $y \in X$ is such that $e(y) \in V$ then $\pi_U e(y) = 0$ and hence $y \in U$. Thus $V \cap e(X) \subset e(\mathcal{U})$.

3.3.4 Urysohn's Metrization Theorem

Theorem 3.26 (Urysohn's Metrization Theorem): Suppose $(X, \mathcal{T})$ is a regular topological space with a countable basis $\mathcal{B}$, then X is metrizable.

Proof: Let $(X, \mathcal{T})$ be a regular metrizable space with countable basis $\mathcal{B}$. For this proof, we will first create a countable collection of functions $\{f_n\}_{n \in \mathbf{N}}$, where $f_m: X \to \mathbf{R}$ for all $m \in \mathbf{N}$, such that given any $x \in X$ and any open neighbourhood U of x there is an index N such that $f_N(x) > 0$ and zero outside of U. We will then use these functions to imbed X in $\mathbf{R}^w$.

Let $x \in X$ and let U be any open neighbourhood of x. There exists $B_m \in \mathcal{B}$ such that $x \in B_m$. Now, since X has a countable basis and is regular, we know that X is normal. Next, as B_m is open, there exists some $B_n \in \mathcal{B}$ such that $\overline{B_n} \subset B_m$. Thus we now have two closed sets $\overline{B_n}$ and $X \setminus B_m$, and so we can apply Urysohn's lemma to give us a continuous function $g_{n,m}: X \to \mathbf{R}$ such that $g_{n,m}(\overline{B_n}) = \{1\}$ and $g_{n,m}(X \setminus B_m) = \{0\}$. Notice here that this function satisfies requisite: $g_{n,m}(y) = 0$ for $y \in X \setminus B_m$ and $g_{n,m}(x) > 0$. Here, g was indexed purposely, as it shows us that $\{g_{n,m}\}$ is indexed by $\mathbf{N} \times \mathbf{N}$, which is countable (since the cross product of two countable sets is countable). Considering this, relable the functions $\{g_{n,m}\}_{n,m \in \mathbf{N}}$ as $\{f_n\}_{n \in \mathbf{N}}$.

We now imbed X in the metrizable space $\mathbf{R}^w$. Let $F: X \to \mathbf{R}^w$ where $F(x) = (f_1(x), f_2(x), f_3(x), \ldots)$, where f_n are the functions constructed above. We claim that F is an imbedding of X into $\mathbf{R}^w$.

NOTES

For F to be an imbedding it is required of F to be homeomorphic onto its image. First, this needs that F should be a continuous bijection onto its image. We know that F is continuous as each of its component functions f_N are continuous by construction. Now we show that F is an injection.

Let $x, y \in X$ be distinct. From the Hausdorff condition there exist open sets U_x and U_y such that $x \in U_x, y \in U_y$ with $U_x \cap U_y = \phi$. By the construction of our maps f there exists an index $N \in \mathbf{N}$ such that $f_N(U_x) > 0$ and $F_N(X \setminus U_x) = 0$. It follows that $f_N(x) \neq f_N(y)$ and so $F(x) \neq F(y)$. Hence, F is injective.

Now, as it is clear that F is surjective onto its image $F(X)$, all that is left to show is that F is an embedding. We will show that for any open set $U \in X, F(U)$ is open in $\mathbf{R}^w$. Let $U \subset X$ be open and let $x \in U$. Pick an index N such that $f_N(x) > 0$ and $f_N(X \setminus U) = 0$. Let $F(x) = z \in F(U)$. Let $V = \pi_N^{-1}((0, \infty))$, i.e., all elements of $\mathbf{R}^w$ with a positive Nth coordinate. Now let $W = F(X) \cap V$. We claim that $z \in W \subset F(U)$ showing that $F(U)$ can be written as a union of open sets, hence making it open.

First we show that W is open in $F(X)$. We know that V is an open set in $\mathbf{R}^w$. $W = F(X) \cap V$, and W is open by the definition of the subspace topology.

Thereafter, we will first show that $f(x) = z \in W$ and then $W \subset F(U)$. To prove our first claim, $F(x) = z \Rightarrow (F(x)) = f_N(x) > 0 \Rightarrow - \pi_N(z) = \pi_N(F(x)) = f_N(x) > 0 \Rightarrow \pi_N(z) > 0$ which means that $z \in \pi_N^{-1}(V)$ and also $z \in F(X) \Rightarrow z \in F(X) \cap V = W$. Now we show that $W \subset F(U)$. Let $y \in W$. This means $y \in F(X) \cap V = W$.

Now we show that $W \subset F(U)$. Let $y \in W$. This means $y \in F(X) \cap V$. This means there exists some $w \in X$ such that $F(w) = y$. But, since $y \in V$ we have that:

$\pi_N(y) = \pi_N(F(w)) = f_N(w) > 0$ since $y \in V$, but $f_N(w) = 0$ for all $w \in X \setminus U$ and so $y \in F(U)$.

In conclusion, as we have shown that $F: X \to \mathbf{R}^w$ is a map that preserves open sets in both directions and bijective onto its image, we have shown that F is an embedding of the space X into the metrizable space $\mathbf{R}^w$ and X is therefore metrizable, the metric being given by the induced metric from $\mathbf{R}^w$.

Theorem 3.27: The topology generated by the dictionary ordering on $\mathbf{R}^2$ is metrizable.

Proof: From previous Theorem 3.26, all we have to do for showing that $\mathbf{R}^2$ is metrizable in the dictionary ordering is to prove that this space is regular with a countable basis.

Now, since the set $\{(a, b), (c, d) | a \leq c, b < d; a, b, c, d \in \mathbf{R}\}$ is a basis for the dictionary ordering on $\mathbf{R}^2$ and the set of intervals with rational end-points are a basis for the usual topology on $\mathbf{R}$, it follows that the set $\{(a, b), (c, d) | a \leq c, b < d; a, b, c, d \in \mathbf{Q}\}$ is a countable basis for the dictionary ordering.

Now we will show that the dictionary ordering is regular. Let $a \in \mathbf{R}^2$ and $B \subseteq \mathbf{R}^2$ such that B is closed in the dictionary ordering and $a \notin B$. Let $\varepsilon = \inf \{d(a, b) | \ b \in B\}$. We know that ε is greater than 0, for otherwise a would be an accumulation point of B, which is a contradiction. It follows that the open sets $((a, a-\varepsilon/2), (a, a+\varepsilon/2))$ and $\bigcup_{b\in B} ((b, b-\varepsilon/2), (b, b+\varepsilon/2))$ are disjoint open sets containing a and B, respectively. Hence, the dictionary ordering over $\mathbf{R}^2$ is metrizable, since it is regular and has a countable basis.

Note: In this proof we have shown that a sequence of functions $\{f_n\}_{n\in\mathbf{N}}$ with the property that for each $x \in X$ and each neighbourhood U of x there is some $n \in \mathbf{N}$ such that $f_n(x) > 0$ and $f_n(y) > 0$ for all $y \in X \backslash U$, gives us an imbedding $F: X \to \mathbf{R}^w$. Notice that we have the very similar result if we have a sequence of functions $\{f_j\}_{j\in J}$ with same properties as above: given any $x \in X$ and any neighbourhood U of x there exists $j \in J$ such that $f_j(x) > 0$ and $f_j(y) = 0$ for all $y \in X \backslash U$, then we have an imbedding from $X \to \mathbf{R}^j$ given by $F(x) = (f_j(x))_{j\in J}$. This is known as the **imbedding theorem** and is a generalization of Urysohn's metrization theorem.

3.4 UNIFORM BOUNDEDNESS PRINCIPLE AND ITS CONSEQUENCES

Theorem 3.28 (Baire): Suppose X is a complete metric space and $(X_n)_{n\geq 1}$ is a sequence of closed subsets in X. Also suppose that,

$$\text{Int } X_n = \emptyset \ \forall \ n \geq 1$$

Then

$$\text{Int}\left(\bigcup_{n=1}^{\infty} X_n\right) = \emptyset$$

Proof: Fix $O_n = X_n^c$ so that O_n is open and dense in $X \ \forall \ n \geq 1$. We will show that $G = \bigcap_{n=1}^{\infty} O_n$ is dense in X. Let ω be a nonempty open set in X. We shall prove that $\omega \bigcap G \neq \emptyset$. As usual set,

$$B(x, r) = \{y \in X; d(y, x) < r \}$$

Choose any $x_0 \in \omega$ and $r_0 > 0$ such that,

$$\overline{B(x_0, r_0)} \subset \omega$$

Now select $x_1 \in B(x_0, r_0) \bigcap O_1$ and $r_1 > 0$ such that,

$$\begin{cases} \overline{B(x_1, r_1)} \subset B(x_0, r_0) \bigcap O_1 \\ 0 < r_1 < \dfrac{r_0}{2} \end{cases}$$

This is true as O_1 is open and dense. Construct two sequences (x_n) and (r_n) by induction such that,

$$\begin{cases} \overline{B(x_{n+1}, r_{n+1})} \subset B(x_n, r_n) \bigcap O_{n+1}, \quad \forall n \geq 0 \\ 0 < r_{n+1} < \dfrac{r_n}{2}. \end{cases}$$

NOTES

Clearly, (x_n) is a Cauchy sequence. Suppose $x_n \to \ell$. Now as $x_{n+p} \in B(x_n, r_n)$ for every $n \geq 0$ and for every $p \geq 0$, as $p \to \infty$

$$\ell \in \overline{B(x_n, r_n)}, \quad \forall n \geq 0$$

or $\ell \in \omega \bigcap G$.

The Uniform Boundedness Principle

Consider two vector spaces E and F. Let $\mathcal{L}(E, F)$ denote the space of continuous linear operators from E into F set with the norm,

$$\| T \|_{\mathcal{L}(E,F)} = \sup_{\substack{x \in E \\ \|x\| \leq 1}} \| Tx \|$$

For $\mathcal{L}(E, E)$ we can write $\mathcal{L}(E)$.

Theorem 3.29 (Banach-Steinhaus Uniform Boundedness Principle): Consider two Banach spaces E and F be and let $(T_i)i \in I$ be a family of continuous linear operators from E into F. Let,

$$\sup_{i \in I} \| T_i x \| < \infty \ \forall x \in E \tag{3.1}$$

Then

$$\sup_{i \in I} \| T_i \|_{\mathcal{L}(E,F)} < \infty \tag{3.2}$$

Or

There exists a constant c such that

$$\| T_i x \| \leq c \| x \| \quad \forall x \in E, \quad \forall i \in I$$

Proof: Let,

$$X_n = \{x \in E; \quad \forall i \in I, \ \| T_i x \| \leq n\} \ \forall n \geq 1$$

so that X_n is closed. From Equation (3.1) we have,

$$\bigcup_{n=1}^{\infty} X_n = E$$

From the Baire category theorem, $\text{Int}(X_{n0}) \neq \emptyset$ for some $n_0 \geq 1$. Let us choose $x_0 \in$ E and $r > 0$ such that $B(x_0, r) \subset X_{n0}$. We have,

$$\| T_i(x_0 + rz) \| \leq n_0 \quad \forall i \in I, \qquad \forall z \in B(0,1)$$

$$\Rightarrow r \| T_i \|_{\mathcal{L}(E,F)} \leq n_0 + \| T_i x_0 \|$$

$$\Rightarrow \sup_{i \in I} \| T_i \|_{\mathcal{L}(E,F)} < \infty$$

NOTES

Corollary 1: Consider two Banach spaces E and F. Let (T_n) be a sequence of continuous linear operators from E into F such that $\forall x \in E$ as $n \to \infty$, $T_n x$ converges to a limit denoted by Tx. Then we have,

1. $\sup_n \| T_n \|_{\mathcal{L}(E,F)} < \infty$
2. $T \in \mathcal{L}(E, \mathrm{F})$
3. $\| T \|_{\mathcal{L}(E,F)} \leq \lim\ \inf_{n \to \infty} \| T_n \|_{\mathcal{L}(E,F)}$

Proof:

1. It follows straightforwardly from Theorem 3.20 and hence there exists a constant c such that

   $\| T_n x \| \leq c \| x \| \quad \forall n, \quad \forall x \in E$

   At the limit we find

   $\| Tx \| \leq c \| x \| \quad \forall x \in E$
2. It follows since T is linear.
3. Lastly we have $\| T_n x \| \leq \| T_n \|_{\mathcal{L}(E,F)} \| x \| \quad \forall x \in E$ which implies (3).

Corollary 2: Let G be a Banach space and let B be a subset of G. If for every $f \in G^*$ the set $f(B) = \{\langle f, x \rangle;\ x \in B\}$ is bounded in $\mathbb{R}$ (3.3)

then

B is bounded (3.4)

Proof: Apply Theorem 3.28 with $E = G^*$, $F = \mathbb{R}$ and $I = B$. For every $b \in B$, set

$T_b(f) = \langle f, b \rangle, \quad f \in E = G^*$

From Equation (3.3) this gives,

$\sup_{b \in B} | T_b(f) | < \infty \quad \forall f \in E$

From Theorem (3.28) there exists a constant c such that,

$| \langle f, b \rangle | \leq c \| f \| \ \forall f \in G^* \quad \forall b \in B$

Therefore,

$\| b \| \leq c \ \forall b \in B$

Corollary 3: Consider a Banach space G and a subset B^* of G^*. Suppose, for every $x \in G$ the set $\langle B^*, x \rangle = \{\langle f, x \rangle;\ f \in B^*\}$ is bounded in $\mathbb{R}$ (3.5)

NOTES

Then

B^* is bounded (3.6)

Proof: Apply Theorem 3.28 with $E = G, F = \mathbb{R}$ and $I = B^*$. For every $b \in B^*$ fix,

$$T_b(X) = \langle b, x\rangle (x \in G = E)$$

We get that there exists a constant c such that,

$$|\langle b, x\rangle| \leq c\, \| x \| \quad \forall b \in B^*, \quad \forall x \in G$$

From the definition of dual norm,

$$\| b \| \leq c \quad \forall b \in B^*$$

Theorem 3.30 (Open Mapping Theorem): Consider two Banach spaces E and F and let T be a continuous and onto linear operator from E into F. Then there exists a constant $c > 0$ such that,

$$T(B_E(0,1)) \supset B_F(0,c) \tag{3.7}$$

Equation (3.7) means that the image under T of any open set in E is an open set in F. Suppose U is open in E. We will show that $T(U)$ is open. Fix any point $y_0 \in T(U)$ such that $y_0 = Tx_0$ for some $x_0 \in U$. Let $r > 0$ be such that

B (x_0, r) $\subset$ U, i.e., $x_0 + B(0, r) \subset U$. Then $y_0 + T(B(0, r)) \subset T(U)$

From Equation (3.7) we get that,

$T(B(0, r)) \supset$ B $(0, rc)$

and hence,

$B(y_0, rc) \subset T(U)$

Corollary 4: Consider two Banach spaces E and F and let T be a continuous linear operator from E into F that is bijective. Then T^{-1} is also continuous.

Proof: From Equation (3.7) and the assumption that T is one-to-one we get that if $x \in E$ is chosen so that $\| Tx \| < c$, then $\| x \| < 1$. By homogeneity, we get

$$\| x \| \leq \frac{1}{c} \| Tx \| \quad \forall\, x \in E$$

Corollary 5: Consider a vector space E provided with two norms, $\| \|_1$ and $\| \|_2$. Let E be a Banach space for both norms and let there exists a constant $C \geq 0$ such that,

$$\| x \|_2 \leq C \| x \|_1 \quad \forall x \in E$$

Then, there is a constant $c > 0$ such that,

$$\| x \|_1 \leq C \| x \|_2 \quad \forall x \in E$$

i.e., the two norms are equivalent

Proof: We get the result by applying Corollary 4 with $E = (E, \| \|_1)$, $F = (E, \| \|_2)$ and $T = I$.

Proof of Theorem 3.29

Case 1: If T is a linear onto operator from E onto F, then there exists a constant $c > 0$ such that,

NOTES

$$\overline{T(B(0,1))} \supset B(0,2c) \tag{3.8}$$

Proof: Fix $X_n = n\overline{T(B(0,1))}$. As T is onto, we have $\bigcup_{n=1}^{\infty} X_n = F$. By the Baire category theorem there exists some n_0 such that $\text{Int}(X_{n_0}) \neq \emptyset$. This implies,

$$\text{Int}[\overline{T(B(0,1))}] \neq \emptyset.$$

Choose $c > 0$ and $y_0 \in F$ such that,

$$B(y_0, 4c) \subset \overline{T(B(0,1))} \tag{3.9}$$

Particularly, $y_0 \in \overline{T(B(0,1))}$ and by symmetry

$$-y_0 \in \overline{T(B(0,1))} \tag{3.10}$$

Summing Equations (3.9) and (3.10) we get,

$$B(0,4c) \subset \overline{T(B(0,1))} + \overline{T(B(0,1))}$$

Alternatively, since $\overline{T(B(0,1))}$ is convex we have

$$\overline{T(B(0,1))} + \overline{T(B(0,1))} = 2\overline{T(B(0,1))}$$

and Case 1 follows.

Case 2: Assume T is a continuous linear operator from E into F that satisfies Equation (3.8). Then we have,

$$T(B(0, 1)) \supset B(0, c) \tag{3.11}$$

Proof: Choose any $y \in F$ with $\|y\| < c$. We have to get some $x \in E$ such that,

$\|x\| < 1$ and $Tx = y$.

Equation (3.8) implies that,

$$\forall \varepsilon > 0 \quad \exists z \in E \text{ with } \| z \| < \frac{1}{2} \text{ and } \| y - Tz \| < \varepsilon \tag{3.12}$$

Picking $\varepsilon = c/2$, we find some $z_1 \in E$ such that,

$$\| z_1 \| < \frac{1}{2} \text{ and } \| y - Tz_1 \| < \frac{c}{2}$$

Applying the similar construction to $y - Tz_1$ with $\varepsilon = c/4$ we get some $z_2 \in E$ such that,

$$\| z_2 \| < \frac{1}{4} \text{ and } \| (y - Tz_1) - Tz_2 \| < \frac{c}{4}$$

Proceeding likewise, by induction we obtain a sequence (z_n) such that,

$$\| z_n \| < \frac{1}{2^n} \text{ and } \| (y - T(z_1 + z_2 + ...z_n \| < \frac{c}{2^n} \forall_n$$

NOTES

Clearly the sequence $x_n = z_1 + z_2 + ... + z_n$ is a Cauchy sequence. Let $x_n \to x$ with $\|x\| < 1$ and $y = Tx$ since T is continuous.

Theorem 3.31 (Closed Graph Theorem): Consider two Banach spaces E and F. Let T be a linear operator from E into F. Let the graph of T, $G(T)$ be closed in $E \times F$. Then T is continuous.

Proof: Let,

$$\| x \|_1 = \| x \|_E + \| Tx \|_F \text{ and } \| x \|_2 = \| x \|_E$$

be the two norms on E.

Note: The norm $\| \|_1$ is termed as the graph norm.

Since $G(T)$ is closed, E is a Banach space for the norm $\| \|_1$. Alternatively, E is also a Banach space for the norm $\| \|_2$ and $\| \|_2 \leq \| \|_1$. From Corollary 5, the two norms are equivalent and hence there exists a constant $c > 0$ such that $\|x\|_1 \leq c\|x\|_2$. Hence we can infer that $\| Tx \|_F \leq c \| x \|_E$.

3.5 QUOTIENT SPACE OF NORMED LINEAR SPACE AND ITS COMPLETENESS

Definition

Let M be a subspace of a linear space L and let the coset of an element x in L be defined by

$$x + M = \{x + m; m \in M\}$$

Then the distinct cosets form a partition of and if addition and scalar multiplication are respectively defined by

and

$$\alpha(x + M) \equiv \alpha x + M$$

then these cosets form a linear space denoted by L/M and called the quotient space of L with respect to M. The origin in L/M is the coset $0 + M = M$ and the negative of

$$x + M \text{ is } (-x) + M$$

Theorem 3.32

Let M be a closed linear subspace of a normed linear space$N$. If the norm of a coset $x + M$ in the quotient space N/M is defined by

$$\| x + M \| = \text{Inf}\{\| x + m \|; m \in M\} \quad ...(3.13)$$

then N/M is a normed linear space. Further if N is a Banach space, then so is N/M.

NOTES

Proof: We first check that (3.13) defines a norm in the required manner. It is clear that $\| x+M \| \geq 0$. since $\| x+m \|$ is a non-negative real number and every set of non-negative real numbers is bounded below, hence inf$\{\| x+m \|; m \in M\}$ is non negative, i.e.,

$$\| x+M \| \geq 0 \forall x+M \in N/M$$

Also $\| x+M \| = 0 \Leftrightarrow$ there exists a sequence $\{m_k\}$ in M such that $\|x+m_k\| \to 0$

$\Leftrightarrow$ x is in M

$\Leftrightarrow$ $x+M = M$= The zero element of N/M

Now

$$\begin{aligned}
\| (x+M) + (y+M) \| &= \| (x+y)+M \| \\
&= \text{Inf}\{\| x+y+m \|; m \in M\} \\
&= \text{Inf}\{\|x+y+m+m'\|; m \text{ and } m' \in M\} \\
&= \text{Inf}\{\|(x+m)+(y+m')\|; m, m' \in M\} \\
&\leq \text{Inf}\{\| x+m \|; m \in M\} + \text{Inf}\{\|y+m'\| \cdot m' \in M\} \\
&= \| x+M \| + \| y+M \| \\
\| \alpha(x+M) \| &= \text{Inf}\{\| \alpha(x+m) \|; m \in M\} \\
&= \text{Inf}\{|\alpha| \, \| x+m \|; m \in M\} \\
&= |\alpha| \text{Inf}\{\| x+m \|; m \in M\} \\
&= |\alpha| \, \| x+M \|
\end{aligned}$$

Let us assume that N is complete and we will show that N/M is also complete. If we start with a Cauchy sequence in N/M, then it is sufficient to show that this sequence has a convergent subsequence. It is clearly possible to find a subsequence {x_n+M} of the original Cauchy sequence such that

$$\|(x_1+M)-(x_2+M)\| < \frac{1}{2}$$
$$\|(x_2+M)-(x_3+M)\| < \frac{1}{4}$$

and in general

$$\|(x_n+M)-(x_{n+1}+M)\| < \frac{1}{2^n}$$

We will show that this sequence is convergent in N/M. Let us choose any vector y_1 in x_1+M and we select y_2 in x_2+M such that $\|y_1-y_2\| < \frac{1}{2}$. We now choose a vector y_3 in x_3+M such that $\|y_2-y_3\| < \frac{1}{4}$. Continuing in this way we obtain a sequence $\{y_n\}$ in N such that $\|y_n-y_{n+1}\| < \frac{1}{2^n}$. If $m < n$, then

$$\begin{aligned}
\|y_m - y_n\| &= \| y_m - y_{m+1}) + (y_{m+1} - y_{m+2}) + \cdots + (y_{n-1} - y_n) \| \\
&\leq \|y_m - y_{m+1}\| + \|y_{m+1} - y_{m+2}\| + \cdots + \|y_{n-1} - y_n\| \\
&< \frac{1}{2^m} + \frac{1}{2^{m+1}} + \cdots + \frac{1}{2^{n-1}} \\
&< \frac{1}{2^m} + \frac{1}{2^{m+1}} + \cdots + \\
&= \frac{(1/2)^m}{1-\frac{1}{2}} = \frac{1}{2^{m+1}}
\end{aligned}$$

So $\{y_n\}$ is a Cauchy sequence in N. Since N is complete, there exists a vector y in N such that $y_n \to y$. Finally

$$\| (x_n + M) - (y + M) \| = \| x_n - y + M \|$$
$$\leq \operatorname{Inf}\{\|x_n - y + m\|; m \in M\}$$
$$\leq \|x_n + m - y\| \text{ for all } m \in M$$

NOTES

But $y_n = x_n + m_n$ for some $m_n \in M$

$\leq \|y_n - y\| \to 0$ since $y_n \to y$.

Hence $x_n + M \to y + M \in N/M$

$\Rightarrow$ N/M is complete.

Definition

A series $\sum_{n=1}^{\infty} a_n$, $a_n \in X$ is said to be convergent to $x \in X$, where *X* is a normed linear space if the sequence of partial sums $< S_n >$ where $S_n = \sum_{i=1}^{n} a_i$ converges to *x* i.e. for every $\in > 0$, there exists $n_0 \in N$ such that $\|S_n - x\| < \in$ for $n \geq n_0$. A series $\sum_{n=1}^{\infty} a_n$ is said to be absolutely convergent if $\sum_{n=1}^{\infty} \|a_n\|$ is convergent.

Since every normed linear space is a metric space, hence every convergent sequence in it is Cauchy but not conversely.

Theorem 3.33

A normed linear space is complete if and only if every absolutely convergent series in X is convergent.

3.5.1 Bounded Linear Transformation

Definition

A linear transformation T is said to be bounded if ∃ a non negative real number K such that

$$\| T(x) \| \leq K \| x \| \; \forall x$$

where K is called bound for T.

Definition

Let T be a continuous linear transformation, then

$$\| T \| = \sup\{\| T(x) \|; \| x \| \leq 1\}$$

is called the norm of T.

Clearly norm of T is the smallest M for which $\| T(x) \| \leq M \| x \|$ holds for every

i. e. $\| T \| = \operatorname{Inf}\{M; \| T(x) \| \leq M \| x \|\}$

Theorem 3.34: Let N and N′ be normed linear spaces and let T be a linear transformation of into . Then the inverse T^{-1} exists and is continuous on its domain of definition iff ∃ exists a constant *m* > 0 *s*, that

$$m \| x \| \leq \| T(x) \| \; \forall x \in N. \qquad ...(3.14)$$

Proof: Let (3.13) hold. To show that T^{-1} exists and is continuous. Now exists iff T is one - one. Let $x_1, x_2 \in N$. Then

$$\|\leq\| T(x_1 - x_2) \| = 0 \quad T(x_1) = T(x_2) \quad \Rightarrow T(x_1) - T(x_2) = 0 \quad m \| x_1 - x_2$$
$$\Rightarrow \|x_1 - x_2\| = 0$$
$$\Rightarrow T(x_1 - x_2) = 0$$
$$\Rightarrow x_1 - x_2 = 0 \quad \text{by (3.13)}$$
$$\Rightarrow x_1 = x_2$$

NOTES

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

$$T(x) = y \Rightarrow x = T^{-1}(y)$$

Hence (3.13) is equivalent to

$$m\|T^{-1}y\| \leq \| y \| \Rightarrow \|T^{-1}(y)\| \leq \frac{1}{m} \| y \|$$
$$\Rightarrow T^{-1} \text{ is bounded}$$
$$\Rightarrow T^{-1} \text{ is continuous (by the above theorem)}$$

Conversely let T^{-1} exists and be continuous on its domain Y[N]. Let $x \in N$ Since T^{-1} exists, there is an $y \in T[N]$s. That

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

Again sinceT^{-1} is continuous, it is bounded so that $\exists a + ve$ constant Ks. That

$$\|T^{-1}y\| \leq K \| y \| \Rightarrow \| x \| \leq K \| T(x) \|$$
$$\Rightarrow m \| x \| \leq \| T(x) \| \text{ where } m = \frac{1}{K} > 0$$

Theorem 3.35

Let N and N′ be normed linear spaces and let T be a bounded linear transformation of N into N′: Put

$$a = \sup\{\| T(x) \|; x \in N, \| x \| = 1\}$$
$$b = \sup\{\| T(x) \| / \| x \|; x \in N; x \neq 0\}$$
$$c = \text{Inf}\{K; K \geq 0, \| T(x) \| \leq K \| x \| \; \forall x \in N\}$$

Then

$$\| T \| = a = b = c$$

and

$$\| T(x) \| \leq \| T \| \| x \| \quad \forall x \in N$$

Proof: By definition of norm

$$\| T \| = \sup\{\| T(x) \|; x \in N, \| x \| \leq 1\}$$

By definition of c, $\| T(x) \| \leq c \| x \| \; \forall x \in N$

and if $\| x \| \leq 1$, then $\| T(x) \| \leq c \forall x \in N$

and so $\sup\{\| T(x) \|; x \in N, \| x \| \leq 1\} \leq C$

i.e.

$$\| T \| \leq C.$$

Also by definition of b and c, it is clear that $c \leq b [\| T \| \leq c \leq b]$. A gain if $x \neq 0$,

Then

$$\| T(x) \| / \| x \| = \left\| T\left(\frac{x}{\| x \|}\right)\right\|$$

and $\frac{x}{\|x\|}$ has norm 1. Hence we conclude from the definitions of *b* and a that . But it is evident that

NOTES

$$a = \sup\{\| T(x) \|; x \in N; \| x \| = 1\} \leq \sup\{\| T(x) \|; x \in N; \| x \| \leq 1\}$$
$$\Rightarrow a \leq \| T \|.$$

Thus we have shown that

$$\| T \| \leq c \leq b \leq a \leq \| T \|$$
$$\Rightarrow \| T \| = a = b = c.$$

Finally definition of b shows that

$$\frac{\| T(x) \|}{\| x \|} \leq \sup\left\{\frac{\| T(x) \|}{\| x \|}; x \in N, x \neq 0\right\}$$
$$= b = \| T \|$$
$$\Rightarrow \| T(x) \| \leq \| T \| \| x \|$$

Remark : Now we shall denote the set of all continuous (or bounded) linear transformation of N into N′ by B(N, N′) [where letter B stands for bounded].

3.5.2 Normed Linear Space of Bounded Linear Transformations

Definition

Let *V*, *W* be normed vector spaces (both over $\mathbb{R}$ or over $\mathbb{C}$). A linear transformation or linear operator $T : V \to W$ is bounded if there is a constant *C* such that

$$\| Tx \|_W \leq C \| x \|_V \text{ for all } x \in V. \qquad ...(3.15)$$

Remark

We use the linearity of *T* and the homogeneity of the norm in *W* to see that

$$\left\| T\left(\frac{x}{\| x \|_V}\right)\right\|_W = \left\|\frac{T(x)}{\| x \|_V}\right\|_W = \frac{\| T(x) \|_W}{\| x \|_V}$$

Here *T* is bounded and satisfies (3.15), if and only if

$$\sup_{\|x\|_V = 1} \| T(x) \|_W \leq C.$$

Theorem 3.36

Let *V*,*W* be normed vector spaces and let $T : V \to W$ be a linear transformation. Then the following statements are equivalent:

1. *T* is a bounded linear transformation.
2. *T* is continuous everywhere in .
3. *T* is continuous at 0 in.

Proof: (1.) ⇒ (2.): Let *C* be the constant as defined in the definition of bounded linear transformation. By linearity of *T* we have

$$\| T(v) - T(\tilde{v}) \|_W = \| T(v - \tilde{v}) \|_W \leq C \| v - \tilde{v} \|_V$$

which implies (2.).

(2.) ⇒ (3.) is trivial.

(3.) ⇒ (1.): It is continuous at 0, then there exists $\delta > 0$ such that for all $v \in V$ with $\| v \| < \delta$ we have $\| Tv \| < 1$. Now let $x \in V$ and $x \neq 0$. Then

$$\left\| \delta \frac{x}{2 \| x \|_V} \right\|_V = \delta/2 \text{ and thus } \left\| T\left(\delta \frac{x}{\| x \|_V}\right) \right\|_W < 1.$$

But by the linearity of *T* and the homogeneity of the norm we get

$$1 \geq \left\| T\left(\delta \frac{x}{\| x \|_V}\right) \right\|_W = \left\| \delta \frac{T(x)}{2 \| x \|_V} \right\|_W = \frac{\delta}{2 \| x \|_V} \| Tx \|_W$$

and therefore $\| Tx \|_W \leq C \| x \|_V$ with $C = 2/\delta$.

Notation: If $T: V \to W$ is linear one often writes Tx for $T(x)$.

Definition

We denote the set of all bounded linear transformations $T: V \to W$ by $L(V, W)$. $L(V, W)$ forms a vector space. $S + T$ is the transformation with $(S + T)(x) = S(x) + T(x)$ and cT is the operator $x \mapsto cT(x)$. On $L(V, W)$ we define the operator norm (depending on the norms on *V* and *W*) by

$$\| T \|_{L(V,W)} \equiv \| T \|_{op} = \sup_{v \neq 0} \frac{\| Tv \|_W}{\| v \|_V}.$$

We can see $\| T \|_{L(V,W)}$ as the best constant for which (3.15) holds. Also note

$$\| Tx \|_W \leq \| T \|_{L(V,W)} \| x \|_V.$$

Using the homogeneity of the W-norm cab also be written as

$$\| T \|_{L(V,W)} = \sup_{\| x \|_V = 1} \| Tx \|_W.$$

We use the $\| \cdot \|_{op}$ notation if the choice of *V*, *W* and the norms are clear from the context.

Lemma

Let *V* and *W* be normed spaces. If *V* is finite dimensional then all linear transformations from *V* to *W* are bounded.

Proof: Let $v_1, \ldots, v_n$ be a basis of *V*. Then for $v = \sum_{j=1}^{n} \alpha_j v_j$ we have

$$\| Tv \|_W = \left\| \sum_{j=1}^{n} \alpha_j T v_j \right\|_W \leq \sum_{j=1}^{n} |\alpha_j| \, \|Tv_j\|_W \leq \sum_{j=1}^{n} \|Tv_j\|_W \max_{k=1,\ldots,n} |\alpha_k|$$

The expression $\max_{k=1,\ldots,n} |\alpha_k|$ defines a norm on *V*. Since all the norms on *V* are equivalent, there is a constant C_1 such that

$$\max_{j=1,\ldots,n} |\alpha_j| \leq C \left\| \sum_{j=1}^{n} \alpha_j v_j \right\|_V$$

for all choices of $\alpha_1, \ldots, \alpha_n$. Thus we get $\| Tv \|_W \leq C \| v \|_V$ for all $v \in V$, where the constant *C* is given by $C = C_1 \sum_{j=1}^{n} \|Tv_j\|$.

Lemma

On $\mathbb{R}^n, \mathbb{R}^m$ use the Euclidean norms $\| x \|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2}, \| y \|_2 = \left(\sum_{i=1}^{m} |y_i|^2 \right)^{1/2}$.

NOTES

Let A be an $m \times n$ matrix and consider the linear operator $T: \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(x)$ Ax. Let

$$\| A \|_{HS} := \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

Then

$$\| T \|_{op} \leq \| A \|_{HS}.$$

Proof:

$$\| Ax \|_2^2 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right|^2$$

By the Cauchy-Schwarz inequality,

$$\left| \sum_{j=1}^{n} a_{ij} x_j \right|^2 \leq \sum_{j=1}^{n} |a_{ij}|^2 \, \| x \|_2^2$$

and therefore

$$\| Ax \|_2 \leq \| A \|_{HS} \| x \|_2$$

for all $x \in \mathbb{R}^n$.

Thus $\| T \|_{op} \leq \| A \|_{HS}$.

Check Your Progress

7. Define vector spaces and their ordered basis with the help of an example.
8. Give the general formulation Hahn-Banach theorem.
9. What is Banach space?
10. When is metric space X considered as complete?
11. Define the term embedding.
12. What is embedding lemma?
13. Give the statement of Urysohn's metrization theorem.
14. State closed graph theorem.

3.6 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A normed linear space is a vector space X and a non-negative valued mapping $\|.\|$ on X termed as the norm, which satisfies the following properties:
   (i) $\|x\| = 0$ if and only if $x = 0$.
   (ii) $\|a\,x\| = |a|\,\|x\|$, for all scalars a.
   (iii) $\|x + y\| \leq \|x\| + \|y\|$

NOTES

Here ||x|| is considered as the length of x or the distance from x to 0. For a given vector x, if **y** is defined as (1/||x||) x, then y has unit length and is called the normalized vector for x.

2. The 2- or 3-dimensional vectors are defined through real valued entries and the 'Length' of a vector can easily be extended to any real vector space R^n.
3. A seminormed vector space is a pair (V,p) where V is a vector space and p a seminorm on V. A normed vector space is a pair $(V, \|\cdot\|)$ where V is a vector space and $\|\cdot\|$ a norm on V.
4. A linear transformation means a map T: $V \to W$, such that $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ where $x, y \in V$, $\alpha, \beta \in F$ and V, W are vector spaces over the field F.
5. If V and W be two vector spaces (over F) of dim m and n respectively. Then Hom (V, W) has dim mn.
6. A linear transformation T: $V \to W$ is non singular iff T carries each L.I. subset of V onto a L.I. subset of W.
7. If $U(F)$, $V(F)$ be vector spaces of dimension n and m respectively then $\beta = \{u_1, ..., u_n\}$, $\beta' = \{v_1, ..., v_m\}$ be their ordered basis respectively.
8. The most general formulation Hahn-Banach theorem can be given for a vector space V over the field R of real numbers where a function $f : V \to R$ is called sublinear.
9. A Banach space is a complete normed vector space or a Banach space is a vector space which is equipped with a norm and which is complete with respect to that norm.
10. A metric space X is considered as complete if every Cauchy sequence in X converges to a point in X. Normed spaces whose induced metric spaces are complete are specified with a special name. A Banach space is a normed space whose induced metric space is complete.
11. Suppose X is a set, Y a topological space and f : $X \to Y$ an injective map. The embedding topology on X (for the map f) is the collection,

    $f^{-1}(\mathrm{T}_Y) = \{f^{-1}(V) | V \subset Y \text{ open}\}$ of subsets of X.
12. Let $\mathcal{F}$ be a family of mappings where each member $f \in \mathcal{F}$ maps $X \to Y_f$. Then,
    - The evaluation mapping e: $X \to$ defined by $\pi_f \circ e(x) = f(x)$, for all $x \in X$, is continuous.
    - The mapping e is an open mapping onto $e(X)$ if $\mathcal{F}$ distinguishes points and closed sets.
    - The mapping e is one-to-one if and only if $\mathcal{F}$ distinguishes points.
    - The mappping e is an embedding if F distinguished points F distinguishes points and closed sets.
13. Suppose $(X, \mathcal{T})$ is a regular topological space with a countable basis $\mathcal{B}$, then X is metrizable.

14. Consider two Banach spaces E and F. Let T be a linear operator from E into F. Let the graph of T, $G(T)$ be closed in $E \times F$. Then T is continuous.

NOTES

3.7 SUMMARY

- A normed linear space is a vector space X and a non-negative valued mapping $||.||$ on termed as the norm.
- The 2- or 3-dimensional vectors are defined through real valued entries and the 'Length' of a vector can easily be extended to any real vector space Rn.
- The zero vector '0' has zero length whereas every other vector has a positive length.
- Multiplying a vector by a positive number changes its length without changing its direction.
- The triangle inequality holds, i.e., taking norms as distances, the distance from point A through B to C is never shorter than going directly from A to C or the shortest distance between any two points is a straight line.
- A vector space on which a norm is defined is then called a normed vector space.
- A seminormed vector space is a pair (V,p) where V is a vector space and p a seminorm on V.
- A normed vector space is a pair $(V, \cdot)$ where V is a vector space and $\cdot$ a norm on V.
- A Linear Transformation $(L.T.)$ means a map $T: V \to W$, such that $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ where $x, y \in V$, $\alpha, \beta \in F$ and V, W are vector spaces over the field F.
- The notation $L(V, W)$ is also used for denoting Hom (V, W).
- If V and W be two vector spaces (over F) of dim m and n respectively. Then Hom (V, W) has dim mn.
- A linear transformation $T: V \to W$ is non singular iff T carries each L.I. subset of V onto a L.I. subset of W.
- A function f: R $\to$ R is continuous on R iff for every open set G in R, $f^{-1}(G)$ is open in R.
- A function f: R $\to$ R is continuous on R iff for every closed set A in R $f^{-1}(A)$ is closed in R.
- If function g is continuous at a and f is continuous at $g(a)$ then the composite function $f \, o \, g$ is continuous at a.
- The function f defined by $f(x) = a^x$ and $a > 0$ is one-one strictly monotonic $(a \neq 1)$ and continuous on the domain R with range $(0, \infty)$. Therefore, the inverse function f^{-1} exists and is continuous strictly monotonic as ax on the domain $(0, \infty)$ with R as range.
- The most general formulation Hahn-Banach theorem can be given for a vector space V over the field R of real numbers where a function $f: V \to$ R is called sublinear.

- A Banach space is a complete normed vector space or a Banach space is a vector space which is equipped with a norm and which is complete with respect to that norm.
- A metric space X is considered as complete if every Cauchy sequence in X converges to a point in X. Normed spaces whose induced metric spaces are complete are specified with a special name. A Banach space is a normed space whose induced metric space is complete.
- The L^p spaces are function spaces defined using a natural generalization of the p-norm for finite-dimensional vector spaces. They are sometimes called Lebesgue spaces, named after Henri Lebesgue.
- Let X be a normed space and $X^{**} = (X^*)^*$denote the second dual space of X. The canonical mapdefined by gives an isometric linear isomorphism (embedding) from X into X^{**}.
- Suppose X is a set, Y a topological space and $f : X \rightarrow Y$ an injective map. The embedding topology on X (for the map f) is the collection, $f^{-1}(\mathcal{T}_Y) = \{f^{-1}(V) | V \subset Y \text{ open}\}$ of subsets of X.
- A space is Tychonoff iff it can be embedded in a cube.
- Suppose $(X, \mathcal{T})$ is a regular topological space with a countable basis $\mathcal{B}$, then X is metrizable.
- Consider two Banach spaces E and F be and let $(T_i)i{\in}I$ be a family of continuous linear operators from E into F.
- Consider two Banach spaces E and F and let T be a continuous and onto linear operator from E into F. Then such that, there exists a constant $c > 0$.
- Consider two Banach spaces E and F. Let T be a linear operator from E into F. Let the graph of T, $G(T)$ be closed in $E \times F$. Then T is continuous.

3.8 KEY TERMS

- **Norm:** A normed linear space is a vector space X and a non-negative valued mapping **||.||** on X termed as the norm.
- **Normed vector space:** A vector space on which a norm is defined is then called a normed vector space.
- **Seminormed vector space:** Seminormed vector space is a pair (V, p) where V is a vector space and p a seminorm on V.
- **Linear transformation:** A linear transformation means a map $T : V \rightarrow W$, such that $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ where $x, y \in V$, $\alpha, \beta \in F$ and V, $W$are vector spaces over the field F.
- **Banach space:** A Banach space is a complete normed vector space or a Banach space is a vector space which is equipped with a norm and which is complete with respect to that norm.
- **L^p spaces:** The L^p spaces are function spaces defined using a natural generalization of the p-norm for finite-dimensional vector spaces. They are sometimes called Lebesgue spaces, named after Henri Lebesgue.

3.9 SELF-ASSESSMENT QUESTIONS AND EXERCISES

NOTES

Short-Answer Questions

1. What is the importance of normed linear spaces?
2. Why is linear transformation used?
3. How will you define metric on normed linear spaces?
4. What are dual spaces in linear transformation?
5. What is bounded linear transformation?
6. Define Banach spaces.
7. Specify the term completeness.
8. What is the importance of conjugate spaces?
9. What is embedding?
10. State the embedding lemma.
11. Where is Urysohn's metrization theorem applied?
12. State Baire category theorem.

Long-Answer Questions

1. Show that image of a *L.I.* set by a *L.T.*, need not be *L.I.* (consider zero *L.T*).
2. Let dim $V = n$, T: $V \to V$ be a *L.T.* such that Range T = Ker T. Show that n is even. Prove that T: $R^2 \to R^2$, such that, $T(x_1, x_2) = (x_2, 0)$ is such a *L.T.*
3. Show that f: $R^4 \to R^4$, such that, $f(x, y, z, t) = (2x, 3y, 0, 0)$ is a *L.T.* Find its rank and nullity.
4. Find the *L.T.* from $R^3 \to R^3$ which has its range the subspace spanned by (1, 0, –1), (1, 2, 2).
5. Let G be the set of all invertible linear transformations from $V \to V$ then show that G forms a group under product of linear transformations.
6. Let T: $R^3 \to R^2$, S: $R^2 \to R^2$ be linear transformations. Show that ST is not invertible.
7. Show that it is possible to find two linear operators T, U on R^2 such that $TU = 0$ but $UT \neq 0$. (Consider $(x_1, x_2) \to (x_1, 0)$ and $(x_1, x_2) \to (0, x_1)$).
8. A linear transformation T: $V \to V$ is called *idempotent* or a *projection* if $T^2 = T$. Show that if S, T are idempotent and $ST = TS$ then ST and $S + T - ST$ are idempotent and if $ST + TS = 0$ then $S + T$ is idempotent.
9. If the *L.T.* T: $R^7 \to R^3$ has a four dimensional Kernel, show that range of T has dimension three.
10. Prove the characterization of the embedding topology.

11. Give the proof of Urysohn's metrization theorem.
12. State and prove Baire category theorem.
13. What do you mean by bounded linear transformation? Explain.

NOTES

3.10 FURTHER READING

Yang, Bicheng and Qiang Chen. 2016. *Integration Theory and Functional Analysis*. New York: Magnum Publishing LLC.

Kadets, Vladimir. 2018. *A Course in Functional Analysis and Measure Theory*. New York: Springer International Publishing.

Pundir, R. and S. K. Pundir. 2015. *Integration Theory and Functional Analysis*. Meerut (Uttar Pradesh): Pragati Prakashan.

Rudin, Walter. 1986. *Real and Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Ahlfors, Lars V. 1978. *Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Lang, Serge. 1998. *Complex Analysis*, 4th Edition. New York: Springer-Verlag New York Inc.

Shastri, Anant R. 2000. *An Introduction to Complex Analysis*. India: Macmillan Publishers India Ltd.

Gupta, S. L. and Nisha Rani. 2003. *Fundamental Real Analysis*, 4th Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, 1st Edition. United Kingdom: Cambridge University Press.

Basener, William F. *Topology and Its Applications*. New Jersey: John Wiley & Sons, Inc.

Bourbaki, Nicolas. 1966. *Elements of Mathematics: General Topology*, Part 1. United States: Addison-Wesley.

UNIT 4 FINITE DIMENSIONAL NORMED SPACES AND SUBSPACES

NOTES

Structure

4.0 INTRODUCTION

In mathematics, a norm is a function that assigns a strictly positive length or size to all vectors in a vector space other than the zero vector while a seminorm is allowed to assign zero length to some non-zero vectors. A normed vector space or normed space is a vector space over the real or complex numbers, on which a norm is defined. A norm is the formalization and the generalization to real vector spaces of the intuitive notion of 'Length' in the real world. Riesz's lemma (after Frigyes Riesz) is a lemma in functional analysis. It specifies (often easy to check) conditions that guarantee that a subspace in a normed vector space is dense. The lemma may also be called the Riesz lemma or Riesz inequality. It can be seen as a substitute for orthogonality when one is not in an inner product space.

In functional analysis, each bounded linear operator on a complex Hilbert space has a corresponding Hermitian adjoint (or adjoint operator). Adjoints of operators generalize conjugate transposes of square matrices to (possibly) infinite-dimensional situations. If one thinks of operators on a complex Hilbert space as generalized complex numbers, then the adjoint of an operator plays the role of the complex conjugate of a complex number. In a similar sense, one can define an adjoint operator for linear (and possibly unbounded) operators between Banach spaces. The adjoint of an operator A may also be called the Hermitian conjugate, Hermitian or Hermitian transpose (after Charles Hermite) of A and is denoted by A'' or $A\dagger$ (the latter especially when used in conjunction with the bracket notation). Confusingly, A'' may also be used to represent the conjugate of A.

A reflexive space is a locally convex Topological Vector Space (TVS) such that the canonical evaluation map from X into its bidual (which is the strong dual of the strong dual of X) is an isomorphism of TVSs. Since a normable TVS is reflexive if and only if it is semi-reflexive, every normed space (and so in particular, every Banach space) X is reflexive if and only if the canonical evaluation map from X into its bidual is surjective; in this case the normed space is necessarily also a Banach space. In 1951, R. C. James discovered a Banach space, now known as James' space that is not reflexive but is nevertheless isometrically isomorphic to its bidual (any such isomorphism is thus necessarily not the canonical evaluation map).

In this unit, you will learn about the equivalent norms, finite dimension normed linear spaces and compactness, Riesz lemma, adjoint operators and reflexive spaces.

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Explain the concept of equivalent norms
- Describe finite dimensional normed spaces and compactness
- Explain Riesz lemma
- Discuss adjoint operators and reflexive spaces
- State reflexive spaces

4.2 EQUIVALENT NORMS

A norm is a function that assigns a strictly positive length or size to all vectors in a vector space other than the zero vector while a seminorm is allowed to assign zero length to some non-zero vectors.

A simple example is the 2-dimensional Euclidean space $\mathbf{R}^2$ equipped with the Euclidean norm. Elements in this vector space, example are usually drawn as arrows in a 2-dimensional Cartesian coordinate system starting at the origin (0, 0). The Euclidean norm assigns to each vector the length of its arrow. Because of this, the Euclidean norm is often known as the magnitude. A vector space with a norm is called a normed vector space. Similarly, a vector space with a seminorm is called a seminormed vector space.

Given a vector space V over a subfield F of the complex numbers, a norm on V is a function $p: V \rightarrow F$ with the following properties:

For all a . F and all $\mathbf{u}$, $\mathbf{v}$. V,

1. $p(a\mathbf{v}) = |a|\, p(\mathbf{v})$, (positive homogeneity or positive scalability).
2. $p(\mathbf{u} + \mathbf{v}) \leq p(\mathbf{u}) + p(\mathbf{v})$ (triangle inequality or subadditivity).
3. If $p(\mathbf{v}) = 0$ then $\mathbf{v}$ is the zero vector (separates points).

A simple consequence of the first two axioms, positive homogeneity and the triangle inequality, is $p(\mathbf{0}) = 0$ and thus,

$$p(\mathbf{v}) \geq 0 \ (positivity).$$

A seminorm is a norm with the 3rd property (separating points) removed.

NOTES

Although every vector space is considered seminormed. Every vector space V with seminorm $p(\mathbf{v})$ induces a normed space V/W, called the quotient space, where W is the subspace of V consisting of all vectors $\mathbf{v}$ in V with $p(\mathbf{v}) = 0$. The induced norm on V/W is clearly well-defined and is given by,

$$p(W + \mathbf{v}) = p(\mathbf{v}).$$

A topological vector space is called normable (seminormable) if the topology of the space can be induced by a norm (seminorm). The norm of a vector $\mathbf{v}$ is usually denoted $\|\mathbf{v}\|$ and sometimes $|\mathbf{v}|$. The latter notation is generally not used because it is also used to denote the absolute value of scalars and the determinant of matrices. The following are some example of norms:

- All norms are seminorms.
- The trivial seminorm, with $p(\mathbf{x}) = 0$ for all $\mathbf{x}$ in V.
- The absolute value is a norm on the real numbers.
- Every linear form f on a vector space defines a seminorm by $\mathbf{x} \to |f(\mathbf{x})|$.

Euclidean Norm

On an n-dimensional Euclidean space $\mathbf{R}^n$, the perceptive notion of length of the vector $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ is illustrated by the formula,

$$\|x\| := \sqrt{x_1^2 + \cdots + x_n^2}.$$

The Euclidean norm is the most commonly used norm on $\mathbf{R}^n$.

On an n-dimensional complex space $\mathbf{C}^n$ the most common norm has the form,

$$\|z\| := \sqrt{|z_1|^2 + \cdots + |z_n|^2} = \sqrt{z_1\bar{z}_1 + \cdots + z_n\bar{z}_n}.$$

In both the cases we can express the norm as the square root of the inner product of the vector and itself as follows:

$$\|x\| : \sqrt{x^* x},$$

Here $\boldsymbol{x}$ is represented as a column vector $([x_1; x_2; ...; x_n])$, and $\boldsymbol{x}^*$ denotes its conjugate transpose.

This formula is applicable for any inner product space, including Euclidean and complex spaces. For Euclidean spaces, the inner product is equivalent to the dot product. Hence, in this specific case the formula can also be written with the following notation:

$$\|x\| : \sqrt{x \cdot x}.$$

The Euclidean norm is also called the Euclidean length, L^2 distance, ℓ^2 distance, L^2 norm or ℓ^2 norm. The set of vectors in $\mathbf{R}^{n+1}$ whose Euclidean norm is a given positive constant forms an n-sphere.

NOTES

The Euclidean norm of a complex number is the absolute value (also called the **modulus**) of it, if the complex plane is identified with the Euclidean plane $\mathbf{R}^2$. This identification of the complex number $x + iy$ as a vector in the Euclidean plane makes the quantity $\sqrt{x^2 + y^2}$ (first suggested by Euler) the Euclidean norm associated with the complex number.

Taxicab Norm or Manhattan Norm or L_1 Norm

$$\| x \|_1 := \sum_{i=1}^{n} | x_i | .$$

The name relates to the distance a taxi has to drive in a rectangular street grid to get from the origin to the point x.

The set of vectors whose 1-norm is a given constant forms the surface of a cross polytope of dimension equivalent to that of the norm minus 1. The Taxicab norm is also called the L_1 norm. The distance derived from this norm is called the Manhattan distance or L_1 distance.

p-Norm

Let $p \geq 1$ be a real number.

$$\| x \|_p := \left(\sum_{i=1}^{n} | x_i |^p \right)^{1/p} .$$

Note that for $p = 1$ we get the taxicab norm, for $p = 2$ we get the Euclidean norm and as p approaches ∞ the p-norm approaches the infinity norm or maximum norm.

The L^p class is a vector space and it is also true that the function,

$$\int_X | f(x) - g(x) |^p \, d\mu$$

(without p-th root) defines a distance that makes $L^p(X)$ into a complete metric topological vector space.

Other Norms

Other norms on $\mathbf{R}^n$ can be constructed as follows:

$$\| x \| := 2 \| x_1 \| + \sqrt{3 | x_2 |^2 + \max(| x_3 |, 2 | x_4 |)^2}$$

This is a norm on $\mathbf{R}^4$.

For any norm and any injective linear transformation A we can define a new norm of x equal to $\|Ax\|$.

Weak and Strong Convergence

Definition: A sequence (x_n) in a normed space X is said to be strongly convergent if there is an $x \in X$ such that,

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

This is written as,

NOTES

$$\lim_{n\to\infty} x_n = x$$

Or as,

$$x_n \to x$$

x is called the strong limit of (x_n) and we say that (x_n) converges strongly to x.

Definition: A sequence (x_n) in a normed space X is said to be weakly convergent if there is an $x \in X$ such that for every $f \in X'$,

$$\lim_{n\to\infty} f(x_n) = f(x)$$

This is written as,

$$x_n \xrightarrow{w} x$$

Or as,

$$x_n \rightharpoonup x$$

The element x is called the weak limit of (x_n) and we say that (x_n) converges weakly to x.

Check Your Progress

1. Define a normed vector space.
2. When is a topological space said to be completely normal?
3. What is Euclidean norm?
4. Define weak convergence.

4.3 FINITE DIMENSIONAL NORMED LINEAR SPACES AND COMPACTNESS

The motivating factor in rings was set of integers and in groups the set of all permutations of a set. A vector space originates from the notion of a vector that we are familiar with in mechanics or geometry. You would recall that a vector is defined as a directed line segment, which in algebraic terms is defined as an ordered pair (a, b) being coordinates of the terminal point relative to a fixed coordinate system. Addition of vectors is given by the rule:

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$$

You can easily verify that set of vectors under this forms an abelian group. Also, scalar multiplication is defined by the rule $\alpha (a, b) = (\alpha a, \alpha b)$ which satisfies certain properties. This concept is extended similarly to three dimensions. You can

NOTES

generalize the whole idea through the definition of a vector space and vary the scalars not only in the set of reals but in any field F. A vector space thus differs from groups and rings in as much as it also *involves* elements from outside itself.

Definition: Let $< V, + >$ be an abelian group and $< F, +, \cdot >$ be a field. Define a function $\times$ (called scalar multiplication) from $F \times V \rightarrow V$, such that, for all $\alpha \in F$, $v \in V$, $\alpha \cdot v \in V$. Then V is said to form a *vector space* over F if for all $x, y \in V$, $\alpha, \beta \in F$, the following hold

(*i*) $(\alpha + \beta)\ x = \alpha x + \beta x$

(*ii*) $\alpha\ (x + y) = \alpha x + \alpha y$

(*iii*) $(\alpha\beta)\ x = \alpha\ (\beta x)$

(*iv*) $1 \cdot x = x$, 1 being unity of F.

Also then, members of F are called *scalars* and those of V are called *vectors*.

Note: You can use the same symbol + for the two different binary compositions of V and F, for convenience. Similarly, the same symbol, is used for scalar multiplication and product of the field F.

Since $< V, + >$ is a group, its identity element is denoted by 0. Similarly, the field F would also have zero element which will also be represented by 0. In case of doubt, you can use different symbols like 0_v and 0_F, etc.

Since you generally work with a fixed field, you would only be writing V as a vector space (or sometimes $V(F)$ or V_F). It would always be understood that it is a vector space over F (unless stated otherwise).

You have defined the scalar multiplication from $F \times V \rightarrow V$. You can also define it from $V \times F \rightarrow V$ and have a similar definition. The first one is called a left vector space and the second a right vector space. It is easy to show that if V as a left vector space over F, then it is a right vector space over F and conversely. In view of this result, it becomes redundant to talk about left or right vector spaces. We will consider about only vector spaces over F.

You can also talk about the above system when the scalars are allowed to take values in a ring instead of a field, which leads to the definition of modules.

Theorem 4.1: In any vector space V(F), the following results hold:

(*i*) $0.x = 0$

(*ii*) $\alpha.0 = 0$

(*iii*) $(-\alpha)x = -(\alpha x) = \alpha(-x)$

(*iv*) $(\alpha - \beta)x = \alpha x - \beta x,\ \alpha, \beta \in F, x \in V$

Proof: (*i*) $0.x = (0 + 0).x = 0.x + 0.x$

$\Rightarrow\ 0 + 0.x = 0.x + 0.x$

$\Rightarrow\ 0 = 0.x$ (cancellation in V)

(*ii*) $\alpha.0 = \alpha.(0 + 0) = \alpha.0 + \alpha.0 \Rightarrow \alpha.0 = 0$

(*iii*) $(-\alpha)x + \alpha x = [(-\alpha) + \alpha]x = 0\ .\ x = 0$

$\Rightarrow\ (-\alpha x) = -\ \alpha x$

(*iv*) follows from above.

NOTES

The following examples illustrate Theorem 4.1

(i) If $< F, +, . >$ be a field, then F is a vector space over F as $< F, + > = < V, + >$ is an additive abelian group. Scalar multiplication can be taken as the product of F. All properties are seen to hold. Thus $F(F)$ is a vector space.

(ii) Let $< F, +, . >$ be a field

Let $\quad V = \{(\alpha_1, \alpha_2) \mid \alpha_1, \alpha_2 \in F\}$

Define + and . (scalar multiplication) by

$$(\alpha_1, \alpha_2) + (\beta_1, \beta_2) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2)$$
$$\alpha(\alpha_1, \alpha_2) = (\alpha\alpha_1, \alpha\alpha_2)$$

You can check that all conditions in the definition are satisfied. Here $V = F \times F = F^2$

One can extend this to F^3 and so on. In general we can take n–tuples $(\alpha_1, \alpha_2, ..., \alpha_n)$, $\alpha_i \in F$ and define F^n or $F^{(n)} = \{(\alpha_1, \alpha_2, ..., \alpha_n) \mid \alpha_i \in F\}$ as a vector space over F.

(iii) If $F \subseteq K$ be two fields then $K(F)$ will form a vector space, where addition of $K(F)$ is + of K and for any $\alpha \in F$, $x \in K$, $\alpha . x$ is taken as product of α and x in K.

Thus **C**(**R**), **C**(**C**), **R**(**Q**) would be some examples of vector spaces, where **C** = complex nos., **R** = reals and **Q** = rationals.

(iv) Let V = set of all real valued continuous functions defined on [0, 1]. Then V forms a vector space over the field **R** of reals under addition and scalar multiplication defined by:

$$(f + g)x = f(x) + g(x) \quad f, g \in V$$
$$(\alpha f)x = \alpha f(x) \quad \alpha \in \mathbf{R} \quad \text{for all } x \in [0, 1]$$

It may be recalled here that sum of two continuous functions is continuous and scalar multiple of a continuous function is continuous.

(v) The set $F[x]$ of all polynomials over a field F in an indeterminate x forms a vector space over F with respect to, the usual addition of polynomials and the scalar multiplication defined by:

For $\quad f(x) = a_0 + a_1x + ... + a_nx^n \in F[x], \quad \alpha \in F$

$$\alpha.(f(x)) = \alpha a_0 + \alpha a_1 x + ... + \alpha a_n x^n.$$

(vi) $M_{m \times n}(F)$, the set of all $m \times n$ matrices with entries from a field F forms a vector space under addition and scalar multiplication of matrices.

We use the notation $M_n(F)$ for $M_{n \times n}(F)$.

(vii) Let F be a field and X a non-empty set.

Let $F^X = \{f \mid f : X \to F\}$, the set of all mappings from X to F. Then F^X forms a vector space over F under addition and scalar multiplication defined as follows:

For $\quad f, g \in F^X, \alpha \in F$

Define $\quad f + g : X \to F, \alpha F : X \to F$ such that,

$$(f + g)(x) = f(x) + g(x)$$
$$(\alpha f)(x) = \alpha f(x) \quad \forall\, x \in X$$

NOTES

(viii) Let V be the set of all vectors in three dimensional space. Addition in V is taken as the usual addition of vectors in geometry and scalar multiplication is defined as:

$\alpha \in \mathbf{R}$, $\vec{v} \in V \Rightarrow \alpha\vec{v}$ is a vector in V with magnitude $|\alpha|$ times that of V. Then V forms a vector space over $\mathbf{R}$.

Subspaces

Definition: *A* non-empty subset W of a vector space $V(F)$ is said to form a *subspace* of V if W forms a vector space under the operations of V.

Theorem 4.2: A necessary and sufficient condition for a non-empty subset W of a vector space $V(F)$ to be a subspace is that W is closed under addition and scalar multiplication.

Proof: If W is a subspace, the result follows by definition.

Conversely, let W be closed under addition and scalar multiplication.

Let $x, y, \in W$ since $1 \in F, -1 \in F$

$\therefore\ -1.\, y \in W \Rightarrow -y \in W$

$x, -y \in W \Rightarrow x - y \in W$

$\Rightarrow\ < W, + >$ forms a subgroup of $< V, + >$.

Rest of the conditions in the definition follow trivially.

Theorem 4.3: A non-empty subset W of a vector space $V(F)$ is a subspace of V iff $\alpha x + \beta y \in W$ for $\alpha, \beta \in F, x, y \in W$.

Proof: If W is a subspace, result follows by definition.

Conversely, let given condition hold in W.

Let $x, y \in W$ be any elements. Since $1 \in F$

$$1\,.\,x + 1\,.\,y = x + y \in W$$

$\Rightarrow$ W is closed under addition.

Again, $x \in W$, $\alpha \in F$ then

$$\alpha x = \alpha x + 0.y \quad \text{for any } y \in W,\ 0 \in F$$

which is in W. (Note here 0 may not be in W)

Hence W is closed under scalar multiplication.

The result thus follows by previous theorem.

Remark: V and $\{0\}$ will be trivial subspaces of any vector space $V(F)$.

For example, consider the vector space $\mathbf{R}^2(\mathbf{R})$

then $W_1 = \{(a, 0) \mid a \in \mathbf{R}\}$

$W_2 = \{(0, b) \mid b \in \mathbf{R}\}$

are subspaces of $\mathbf{R}^2$

As for any $\alpha, \beta \in \mathbf{R}$, $(a_1, 0), (a_2, 0) \in W_1$, you find

$$\alpha(a_1, 0) + \beta(a_2, 0) = (\alpha a_1, 0) + (\beta a_2, 0)$$

$$= (\alpha a_1 + \beta a_2, 0) \in W_1$$

Hence W_1 is a subspace. Similarly, you can show W_2 is a subspace of $\mathbf{R}^2$.

Example 4.1: Show that union of two subspaces may not be a subspace.

NOTES

Solution: Consider the given under Theorem 4.3.

$W_1 \cup W_2$ will be the set containing all pairs of the type $(a, 0), (0, b)$

In particular $(1, 0), (0, 1) \in W_1 \cup W_2$

But $\quad (1, 0) + (0, 1) = (1, 1) \notin W_1 \cup W_2$.

Hence $W_1 \cup W_2$ is not a subspace.

You are referred to exercises for more results pertaining to intersection and union of subspaces.

A few more examples of subspaces are as follows:

(i) Let $V = \mathbf{R}[x]$ and suppose $W = \{f(x) \in V \mid f(x) = f(1-x)\}$

Then W is a subspace of V as

$W \neq \phi$ since $0 \in W$ as $f(x) = 0 = f(1-x)$

Again, if $f(x), g(x) \in W$, then $f(x) = f(1-x), g(x) = g(1-x)$

Let $\quad f(x) + g(x) = h(x)$

Then $\quad h(1-x) = f(1-x) + g(1-x)$

$$= f(x) + g(x) = h(x)$$

$\Rightarrow h(x) \in W$ or that $f(x) + g(x) \in W$

Again, for $\alpha \in \mathbf{R}$, let $\alpha f(x) = r(x)$

Then $\quad r(1-x) = \alpha f(1-x) = \alpha f(x) = r(x)$

$\Rightarrow \quad r(x) \in W \Rightarrow \alpha f(x) \in W$

Hence W is a subspace.

(ii) Let $V = F^X$ (Refer example (vii) of Theorem 4.1) and suppose $Y \subseteq X$

Then $\quad W = \{f \in V \mid f(y) = 0 \ \forall\ y \in Y\}$ is a subspace of V

Clearly $0 \in W$ and for $f, g \in W$,

$$f(y) = 0 = g(y) \quad \forall\ y \in Y$$

So $\quad (f+g)(y) = f(y) + g(y) = 0 \quad \forall\ y \in Y$

$\Rightarrow \ f + g \in W$

Again, if $\alpha \in F$, then

$$(\alpha f)y = \alpha(f(y)) = 0 \quad \forall\ y \in Y$$

$\Rightarrow \ \alpha f \in W$.

(iii) If $V = \mathbf{R}^n$, then

$W = \{(x_1, x_2, ..., x_n) \mid x_1 + x_2 + ... + x_n = 1\}$ will not be a subspace of V.

Notice, $(1, 0, 0, ..., 0) + (0, 1, 0, ..., 0) = (1, 1, 0, ..., 0) \notin W$.

(iv) Let $V = M_{2 \times 1}(F)$. Let A be a 2×2 matrix over F.

Then $\quad W = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in V \mid A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \right\}$ forms a subspace of V

NOTES

$$W \neq \phi \text{ as } \begin{bmatrix} 0 \\ 0 \end{bmatrix} \in W$$

For $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ in W, we have

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 = A \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\Rightarrow A \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right) = 0$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in W$$

Also $A \left(\alpha \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \alpha A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \Rightarrow \alpha \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in W$

Hence W is a subspace of V.

(v) Let $V = F_2^2$, where $F_2 = \{0, 1\}$ mod 2.

If
$$W_1 = \{(0, 0), (1, 0)\}$$
$$W_2 = \{(0, 0), (0, 1)\}$$
$$W_3 = \{(0, 0), (1, 1)\}$$

Then $W_1 \cup W_2 \cup W_3 = \{(0, 0), (1, 0), (0, 1), (1, 1)\} = V$

Thus we notice that here V is union of finite number of proper subspaces.

This result may, however, not hold if V happens to be a vector space over an infinite field.

Example 4.2: Let V be a vector space over a finite field F. Suppose $V = W_1 \cup W_2 \cup ... \cup W_k$, W_i being subspaces of $V \, \forall \, i$. If $o(F) \geq k$ then, show that $V = W_i$ for some i.

Solution: Suppose $V \neq W_i$ for any i

Now $W_k \not\subseteq W_1 \cup W_2 \cup ..... \cup W_{k-1}$

and $W_1 \cup W_2 \cup ..... \cup W_{k-1} \not\subseteq W_k$

$\Rightarrow$ $\exists \, x \in W_k$ such that, $x \notin W_1 \cup W_2 \cup ..... \cup W_{k-1}$

and $\exists \, y \in W_1 \cup ..... \cup W_{k-1}$ such that, $y \notin W_k$

Let $S = \{ax + y \mid a \in F\}$

Then no element of S can belong to W_k, as

$$ax + y \in W_k \Rightarrow ax + y - ax = y \in W_k, \text{ a contradiction}$$

So $ax + y \notin W_k \quad \forall \, a \in F$

$\Rightarrow$ $ax + y \in W_1 \cup W_2 \cup ...... \cup W_{k-1} \qquad \forall \, a \in F$

So $\exists \, \alpha, \beta \in F, \alpha \neq \beta$ such that

$$\alpha x + y \in W_j, \beta x + y \in W_j \text{ for some } j, 1 \leq j \leq k - 1$$

$\therefore$ $(\alpha x + y) - (\beta x + y) \in W_j$

$\Rightarrow$ $(\alpha - \beta)x \in W_j$

$\Rightarrow \quad x \in W_j \Rightarrow x \in W_1 \cup ..... \cup W_{k-1}$, a contradiction

$\therefore \quad V = W_i$ for some i

(You may notice here that example (v) of subspaces $o(F) = 2$ and you could write $V = W_1 \cup W_2 \cup W_3$, $V \neq W_i$ for any i)

NOTES

Sum and Direct Sum of Subspaces

If W_1 and W_2 be two subspaces of a vector space $V(F)$, then we define

$$W_1 + W_2 = \{w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2\}$$

$$W_1 + W_2 \neq \varphi \text{ as } 0 = 0 + 0 \in W_1 + W_2$$

Again, $x, y \in W_1 + W_2$, $\alpha, \beta \in F$ implies

$$x = w_1 + w_2$$

$$y = w'_1 + w'_2 \; w_1, w'_1 \in W_1, w_2, w'_2 \in W_2$$

$$\alpha x + \beta y = \alpha\,(w_1 + w_2) + \beta(w'_1 + w'_2)$$

$$= (\alpha w_1 + \beta w'_1) + (\alpha w_2 + \beta w'_2) \in W_1 + W_2$$

Showing thereby that sum of two subspaces is a subspace.

You can extend the definition, similarly, to the sum of n subspaces $W_1, W_2, ..., W_n$, which would also be a subspace and we write $W_1 + W_2 + ... + W_n = \sum_{i=1}^{n} W_i$

Definition: Let $W_1, W_2, ..., W_n$ be subspaces of V then $W_1 + W_2 + ... + W_n$ is called the direct sum if each $x \in W_1 + W_2 + ... + W_n$ can be expressed uniquely as $x = w_1 + w_2 + ... + w_n$, $w_i \in W_i$ and in that case we write

$$W_1 + W_2 + ... + W_n = W_1 \oplus W_2 \oplus ... \oplus W_n$$

We say, a vector space V is the direct sum of its subspaces $W_1, W_2, ..., W_n$ if $V = W_1 \oplus W_2 \oplus ... \oplus W_n$, i.e., if

$$V = W_1 + W_2 + ... + W_n$$

and each $v \in V$ can be expressed uniquely as $v = w_1 + w_2 + ... + w_n$, $w_i \in W_i$.

Theorem 4.4: $V = W_1 \oplus W_2 \Leftrightarrow V = W_1 + W_2$, $W_1 \cap W_2 = (0)$.

Proof: Let $V = W_1 \oplus W_2$

We need to prove $W_1 \cap W_2 = (0)$

Let $x \in W_1 \cap W_2$, then $x \in W_1$ and $x \in W_2$

$$\Rightarrow x = 0 + x \in W_1 + W_2 = V$$

$$\Rightarrow x = x + 0 \in W_1 + W_2 = V$$

Since x has been expressed as $x = x + 0$ and $0 + x$ and the representation has to be unique, we get $x = 0$

$$\Rightarrow W_1 \cap W_2 = (0).$$

Conversely, let $v \in V$ be any element and suppose

$$v = w_1 + w_2$$

$v = w'_1 + w'_2$ are two representations of v

then $\quad w_1 + w_2 = w'_1 + w'_2 \; (= v)$

$\Rightarrow \quad w_1 - w'_1 = w'_2 - w_2$

NOTES

Now L.H.S. is in W_1 and R.H.S. belongs to W_2

i.e., each belongs to $W_1 \cap W_2 = (0)$

$$\Rightarrow \quad w_1 - w'_1 = w'_2 - w_2 = 0$$

$$\Rightarrow \quad w_1 = w'_1, \; w_2 = w'_2.$$

Hence the result.

Note: The above theorem can also be stated as

$$W_1 + W_2 = W_1 \oplus W_2 \Leftrightarrow W_1 \cap W_2 = \{0\}.$$

Consider the Following Example: Consider the space $V(F) = F^2(F)$ where F is a field

Let
$$W_1 = \{(a, 0) \mid a \in F\}$$
$$W_2 = \{(0, b) \mid b \in F\}$$

then V is direct sum of W_1 and W_2

$$v \in V \Rightarrow \quad v = (a, b) = (a, 0) + (0, b) \in W_1 + W_2$$

thus $\quad V \subseteq W_1 + W_2$

or that $\quad V = W_1 + W_2$

Again if $(x, y) \in W_1 \cap W_2$ be any element then

$$(x, y) \in W_1 \text{ and } (x, y) \in W_2$$

$$\Rightarrow \quad y = 0 \text{ and } x = 0$$

$$\Rightarrow \quad (x, y) = (0, 0)$$

$$\Rightarrow W_1 \cap W_2 = (0)$$

Hence $\quad V = W_1 \oplus W_2.$

Example 4.3: Let V be the vector space of all functions from $\mathbf{R} \to \mathbf{R}$. Let $V_e = \{f \in V \mid f \text{ is even}\}$, $V_0 = \{f \in V \mid f \text{ is odd}\}$. Then V_e and V_0 are subspaces of V and $V = V_e \oplus V_0$.

Solution: Addition and scalar multiplication in V are given by the rule

$$(f + g)\,x = f(x) + g(x); \; (\alpha f)\,x = \alpha f(x)$$

Now $\quad V_e \neq \varphi$ as $0(x) = 0 \Rightarrow 0(x) = 0(-x)$

$$\Rightarrow 0 \in V_e$$

Again for $\alpha, \beta \in \mathbf{R}, f, g \in V_e$, we have

$$\begin{aligned}(\alpha f + \beta g)\,(-x) &= (\alpha f)\,(-x) + (\beta g)\,(-x)\\ &= \alpha(f(-x)) + \beta(g(-x))\\ &= \alpha f(x) + \beta g(x)\\ &= (\alpha f + \beta g)x\end{aligned}$$

$$\Rightarrow \alpha f + \beta g \in V_e$$

$$\Rightarrow V_e \text{ is a subspace of } V$$

Similarly, V_0 is a subspace of V.

Thus, $V_e + V_0$ is a subspace of V. We show $V \subseteq V_e + V_0$

Let $f \in V$ be any member

NOTES

Let $g : \mathbf{R} \to \mathbf{R}$ be such that $g(x) = f(-x)$, then $g \in V$

Also then $$f = \left(\frac{1}{2}f + \frac{1}{2}g\right) + \left(\frac{1}{2}f - \frac{1}{2}g\right)$$

Since $$\left(\frac{1}{2}f + \frac{1}{2}g\right)(-x) = \frac{1}{2}f(-x) + \frac{1}{2}g(-x) = \frac{1}{2}g(x) + \frac{1}{2}f(x)$$

$$= \left(\frac{1}{2}f + \frac{1}{2}g\right)x$$

we find $\frac{1}{2}f + \frac{1}{2}g \in V_e$

Similarly, $\frac{1}{2}f - \frac{1}{2}g \in V_0$

$\Rightarrow \quad f \in V_e + V_0 \quad \Rightarrow \quad V \subseteq V_e + V_0$

or that $\quad V = V_e + V_0$

Finally, $\quad f \in V_e \cap V_0 \quad \Rightarrow \quad f \in V_e, f \in V_0$

$\Rightarrow \quad f(-x) = f(x) \quad \text{and} \quad f(-x) = -f(x)$

$\Rightarrow \quad f(x) = -f(x)$

$\Rightarrow \quad f(x) + f(x) = 0 = 0(x)$

$\Rightarrow \quad 2f(x) = 0\,(x)$ for all x

$\Rightarrow \quad 2f = 0 \quad \Rightarrow \quad f = 0 \quad \Rightarrow \quad V_e \cap V_0 = (0).$

Hence the result.

Example 4.4: If L, M, N are three subspaces of a vector space V, such that $M \subseteq L$ then show that $L \cap (M + N) = (L \cap M) + (L \cap N) = M + (L \cap N)$.

Also give an example, where the result fails to hold when $M \not\subseteq L$.

Solution: We leave the first part for you to try. Recall a similar result was proved for ideals in rings. The equality is called modular equality.

Consider now the vector space $V = \mathbf{R}^2$

Let
$$L = \{(a, a) \mid a \in \mathbf{R}\}$$
$$M = \{(a, 0) \mid a \in \mathbf{R}\}$$
$$N = \{(0, b) \mid b \in \mathbf{R}\}$$

It is a routine matter to cheek that L, M, N are subspaces of V. Indeed

$$\alpha(a, a) + \beta(a', a') = (\alpha a, \alpha a) + (\beta a', \beta a')$$
$$= (\alpha a + \beta a', \alpha a + \beta a') \in L, \text{ etc.}$$

Now $\quad (x, y) \in L \cap M \quad \Rightarrow \quad (x, y) \in L$ and $(x, y) \in M$

$\Rightarrow \quad y = x$ and $y = 0$

$\Rightarrow \quad x = 0 = y \quad \Rightarrow \quad (x, y) = (0, 0)$

Similarly, $\quad L \cap N = \{(0, 0)\}$

$\Rightarrow \quad L \cap M + L \cap N = \{(0, 0)\}$

Again, $\quad M + N = \{(a, b) \mid a, b \in \mathbf{R}\}$ and as $(1, 1) \in M + N$

$(1, 1) \in L$

We find $(1, 1) \in L \cap (M + N)$, but $(1, 1) \notin L \cap M + L \cap N$

Hence $\quad L \cap (M + N) \neq (L \cap M) + (L \cap N)$, when $M \not\subseteq L$.

Example 4.5: Let $V = \mathbf{R}^X$ (Refer example of Theorem 4.1) and *fix* $x_0 \in X$. Define

$$W = \{f \in V \mid f(x_0) = 0\}$$
$$W' = \{g \in \mathrm{V} \mid g(x) = 0 \ \forall \ x \in X - \{x_0\}\}$$

then show that W, W' are subspaces of V and $V = W \oplus W'$.

Solution: We leave it for the reader to show that W, W' are subspaces.

Let $f \in W \cap W'$ then $f \in W$ and $f \in W'$

$$\Rightarrow \quad f(x_0) = 0, f(x) = 0 \quad \forall \ x \in X, x \neq x_0$$
$$\Rightarrow \quad f(x) = 0, \ \forall \ x \in X,$$
$$\Rightarrow \quad f = 0 \text{ and thus } W \cap W' = \{0\}.$$

Let $f \in V$ and let $f(x_0) = r$

Then $(f - r\,\delta x_0) \in W$, $\ r\,\delta x_0 \in W'$

and $\quad f = (f - r\,\delta x_0) + r\,\delta x_0 \in W + W'$

$\therefore \quad V = W + W'$

i.e., $\quad V = W \oplus W'$

Notice here δx_0 denotes the Kronecker delta, i.e., $\delta x_0\ (x_0) = 1$, $\delta x_0\ (x) = 0 \ \forall \ x \neq x_0$.

Quotient Spaces

If W be a subspace of a vector space $V(F)$ then since $< W, + >$ forms an abelian group of $< V, + >$, we can talk of cosets of W in V. Let $\frac{V}{W}$ be the set of all cosets $W + v$, $v \in V$, then we show that $\frac{V}{W}$ also forms a vector space over F, under the operations defined by:

$$(W + x) + (W + y) = W + (x + y) \quad x, y \in V$$
$$\alpha(W + x) = W + \alpha x \ \ \alpha \in F$$

Addition is well defined, since,

$$W + x = W + x'$$
$$W + y = W + y'$$
$$\Rightarrow \quad x - x' \in W, y - y' \in W$$
$$\Rightarrow \quad (x - x') + (y - y') \in W$$
$$\Rightarrow \quad (x + y) - (x' + y') \in W$$
$$\Rightarrow \quad W + (x + y) = W + (x' + y')$$

Again, $\quad W + x = W + x'$

$$\Rightarrow \quad x - x' \in W,$$
$$\Rightarrow \quad \alpha(x - x') \in W \qquad \alpha \in F$$
$$\Rightarrow \quad \alpha x - \alpha x' \in W$$
$$\Rightarrow \quad W + \alpha x = W + \alpha x'$$

$$\Rightarrow \qquad \alpha(W + x) = \alpha(W + x')$$

Thus, scalar multiplication is also well defined. It should now be a routine exercise to check that all conditions in the definition of a vector space are satisfied.

NOTES

$W + 0$ will be zero of $\dfrac{V}{W}$

$W - x$ will be inverse of $W + x$

Also $\alpha((W + x) + (W + y)) = \alpha(W + (x + y))$

$$= W + \alpha(x + y)$$
$$= W + (\alpha x + \alpha y)$$
$$= (W + \alpha x) + (W + \alpha y)$$
$$= \alpha(W + x) + \alpha(W + y) \text{ etc.}$$

Hence, V/W forms a vector space over F, called the quotient space of V by W.

4.3.1 Linear Transformation in Vector Spaces

In this section, you will learn about the concept of a homomorphism in case of vector spaces.

Definition: Let V and U be two vector spaces over the same field F, then a mapping $T : V \to U$ is called a homomorphism or a linear transformation if

$$T(x + y) = T(x) + T(y) \quad \text{for all } x, y \in V$$
$$T(\alpha x) = \alpha T(x) \quad \alpha \in F$$

You can combine the two conditions to get a single condition

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) \quad x, y \in V; \alpha, \beta \in F$$

It is easy to see that both are equivalent. If a homomorphism happens to be one-one onto also we call it an *isomorphism*, and say the two spaces are isomorphic. (Notation $V \cong U$).

This concept is illustrated with the help of the following examples:

(i) Identity map $\quad I : V \to V$, such that,

$$I(v) = v$$

and the zero map $\quad O : V \to V$, such that,

$$O(v) = 0$$

are clearly linear transformations.

(ii) For a field F, consider the vector spaces F^2 and F^3. Define a map $T : F^3 \to F^2$, by

$$T(\alpha, \beta, \gamma) = (\alpha, \beta)$$

then T is a linear transformation as

for any $x, y \in F^3$, if $x = (\alpha_1, \beta_1, \gamma_1)$

$$y = (\alpha_2, \beta_2, \gamma_2)$$

then $\quad T(x + y) = T(\alpha_1 + \alpha_2, \ \beta_1 + \beta_2, \ \gamma_1 + \gamma_2)$

$$= (\alpha_1 + \alpha_2, \ \beta_1 + \beta_2)$$
$$= (\alpha_1, \beta_1) + (\alpha_2, \beta_2) = T(x) + T(y)$$

and $\quad T(\alpha x) = T(\alpha\ (\alpha_1, \beta_1, \gamma_1)) = T((\alpha\alpha_1, \alpha\beta_1, \alpha\gamma_1)$

NOTES

$$= (\alpha\alpha_1, \alpha\beta_1) = \alpha(\alpha_1, \beta_1) = \alpha T(x)$$

(iii) Let V be the vector space of all polynomials in x over a field F. Define $T : V \to V$, such that,

$$T(f(x)) = \frac{d}{dx} f(x)$$

then
$$T(f + g) = \frac{d}{dx}(f + g) = \frac{d}{dx} f + \frac{d}{dx} g = T(f) + T(g)$$

$$T(\alpha f) = \frac{d}{dx}(\alpha f) = \alpha \frac{d}{dx} f = \alpha T(f)$$

shows that T is a linear transformation.

In fact, if $\theta : V \to V$ be defined such that

$$\theta(f) = \int_0^x f(t)\ dt$$

then θ will also be a linear transformation.

(iv) Consider the mapping

$$T : \mathbf{R}^3 \to \mathbf{R}, \text{ such that,}$$

$$T(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$$

then T is not a linear transformation.

Consider, for instance,

$$T((1, 0, 0) + (1, 0, 0)) = T(2, 0, 0) = 4$$

$$T(1, 0, 0) + T(1, 0, 0) = 1 + 1 = 2.$$

Theorem 4.5: Under a homomorphism $T : V \to U$,

(*i*) $T(0) = 0$ (*ii*) $T(-x) = -T(x)$.

Proof:
$$T(0) = T(0 + 0) = T(0) + T(0)$$

$$\Rightarrow \quad T(0) = 0$$

Again
$$T(-x) + T(x) = T(-x + x) = T(0) = 0$$

$$\Rightarrow \quad -T(x) = T(-x).$$

Definition: Let $T : V \to U$ be a homomorphism, then *kernel of T* is defined by

$$\text{Ker } T = \{x \in V \mid T(x) = 0\}$$

It is also called the *null space* of T.

Theorem 4.6: Let $T : V \to U$ be a homomorphism, then *Ker T* is a subspace of V.

Proof: Ker $T \neq \varphi$ as $0 \in$ Ker T

Let $\alpha, \beta \in F$, $x, y \in$ Ker T be any elements

then
$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$$

$$= \alpha \,.\, 0 + \beta \,.\, 0 = 0 + 0 = 0$$

$$\Rightarrow \quad \alpha x + \beta y \in \text{Ker } T.$$

Theorem 4.7: Let $T : V \to U$ be a homomorphism, then

Ker T $= \{0\}$ *iff T* is *one-one*.

Proof: Let Ker T $= \{0\}$. If $T(x) = T(y)$

NOTES

then $\quad T(x) - T(y) = 0$

$\Rightarrow \quad T(x - y) = 0$

$\Rightarrow \quad (x - y) \in \text{Ker } T = \{0\}$

$\Rightarrow \quad x - y = 0$

$\Rightarrow \quad x = y \quad \Rightarrow \; T$ is 1–1.

Conversely, let T be one-one

if $x \in \text{Ker } T$ be any element, then $T(x) = 0$

$\Rightarrow \quad T(x) = T(0)$

$\Rightarrow \quad x = 0$

$\Rightarrow \quad \text{Ker } T = \{0\}.$

Definition: Let $T : V \to U$ be a linear transformation then range of T is defined to be

$$T(V) = \{T(x) \mid x \in V\} = \text{Range } T = R_T$$
$$= \{u \in U \mid u = T(v),\ v \in V\}$$

Theorem 4.8: Let $T : V \to U$ be a linear transformation (linear transformation) then range of T is subspace of U.

Proof: Since $\quad T(0) = 0,\ 0 \in V$

$\therefore \quad T(0) \in \text{Range } T$

i.e., $\quad \text{Range } T \neq \varphi$

Let $\alpha, \beta \in F$, $T(x), T(y) \in T(V)$ be any elements

then $\quad x, y \in V$

Now $\quad \alpha T(x) + \beta T(y) = T(\alpha x + \beta y) \in T(V)$

as $\quad \alpha x + \beta y \in V$

Hence the result.

Note: $T(V) = U$ *iff* T is onto.

Theorem 4.9: Let $T : V \to U$ *be a* linear transformation then

$$\frac{V}{Ker\ T} \cong \text{Range } T = T(V).$$

Proof: Let $T : V \to U$ and put Ker $T = K$, then K being a subspace of V, we can talk of V/K.

Define a mapping $\theta : V/K \to T(V)$, such that,

$$\theta\,(K + x) = T(x), \quad x \in V$$

Then θ is well defined, one-one map as

$$K + x = K + y$$

$\Leftrightarrow \quad x - y \in K = \text{Ker } T$

$\Leftrightarrow \quad T(x - y) = 0$

$\Leftrightarrow \quad T(x) = T(y)$

$\Leftrightarrow \quad \theta(K + x) = \theta(K + y)$

If $T(x) \in T(V)$ be any element, then $x \in V$ and $\theta(K + x) = T(x)$, showing that θ is onto.

NOTES

$$\text{Finally, } \theta((K+x)+(K+y)) = \theta(K+(x+y))$$
$$= T(x+y)$$
$$= T(x)+T(y)$$
$$= \theta(K+x)+\theta(K+y)$$

and $\theta(\alpha(K+x)) = \theta(K+\alpha x) = T(\alpha x) = \alpha T(x) = \alpha\theta(K+x)$

shows θ is a linear transformation and hence an isomorphism.

Note: The above is called the fundamental theorem of homomorphism for vector spaces.

If the map T is also onto, then we have proved $\dfrac{V}{\text{Ker } T} \cong U$.

Theorem 4.10: If A and B be two subspaces of a vector space $V(F)$, then

$$\frac{A+B}{A} \cong \frac{B}{A\cap B}.$$

Proof: A being a subspace of $A+B$ and $A\cap B$ being a subspace of B, we can talk of $\dfrac{A+B}{A}$ and $\dfrac{B}{A\cap B}$.

Define a map $\theta : B \to \dfrac{A+B}{A}$ such that,

$$\theta(b) = A+b, \quad b\in B$$

Since $b_1 = b_2 \Rightarrow A+b_1 = A+b_2$, we find θ is well defined.

Again, as $\theta(\alpha b_1 + \beta b_2) = A + (\alpha b_1 + \beta b_2)$

$$= (A+\alpha b_1) + (A+\beta b_2)$$
$$= \alpha(A+b_1) + \beta(A+b_2) = \alpha\theta(b_1) + \beta\theta(b_2)$$

θ is a linear transformation

For any $A + x \in \dfrac{A+B}{A}$, we find $x \in A+B$

$$\Rightarrow \quad x = a+b, \quad a\in A, b\in B$$
$$A+x = A+(a+b)$$
$$= (A+a)+(A+b)$$
$$= A+(A+b) = A+b = \theta(b).$$

Showing that b is the required pre image of $A+x$ under θ and thus θ is onto.

Hence by Fundamental theorem

$$\frac{A+B}{A} \cong \frac{B}{\text{Ker }\theta}.$$

We claim $\quad \text{Ker }\theta = A\cap B$

Indeed $\quad x\in \text{Ker }\theta \Leftrightarrow \theta(x) = A$

$$\Leftrightarrow A+x = A$$
$$\Leftrightarrow x\in A, \text{ also } x\in \text{Ker }\theta \subseteq B$$
$$\Leftrightarrow x \in A\cap B$$

Hence $\quad \dfrac{A+B}{A} \cong \dfrac{B}{A\cap B}$

Note: By interchanging A and B, we get $\dfrac{B+A}{B} \cong \dfrac{A}{B \cap A}$

i.e., $$\frac{A+B}{A} \cong \frac{B}{A \cap B}.$$

NOTES

Corollary: If $A + B$ is the direct sum then as $A \cap B = \{0\}$

we get $$\frac{A}{(0)} \cong \frac{A \oplus B}{B}$$

But $\dfrac{A}{(0)} \cong A$ (Refer Note of Theorem 4.10) gives us $A \cong \dfrac{A \oplus B}{B}$.

Theorem 4.11: Let W be a subspace of V, then $\exists$ an onto linear transformation $\theta : V \to \dfrac{V}{W}$ such that, $Ker\ \theta = W$.

Proof: Define $\theta : V \to \dfrac{V}{W}$ such that,

$$\theta(x) = W + x$$

then θ is clearly well defined.

Also
$$\begin{aligned}\theta(\alpha x + \beta y) &= W + (\alpha x + \beta y)\\ &= (W + \alpha x) + (W + \beta y)\\ &= \alpha(W + x) + \beta(W + y) = \alpha\theta(x) + \beta\theta(y)\end{aligned}$$

Shows θ is a linear transformation

θ is clearly onto.

Again,
$$\begin{aligned} x \in \text{Ker}\ \theta &\Leftrightarrow \theta(x) = W\\ &\Leftrightarrow W + x = W\\ &\Leftrightarrow x \in W\end{aligned}$$

Hence $\text{Ker}\ \theta = W$.

θ is called the *natural homomorphism* or the *quotient map*.

Note: In case $W = (0)$ in the above, we find θ will be 1–1 also as

$$\begin{aligned} \theta(a) &= \theta(b)\\ \Rightarrow \quad W + a &= W + b\\ \Rightarrow \quad a - b \in W &= (0)\\ \Rightarrow \quad a - b &= 0\\ \Rightarrow \quad a &= b.\end{aligned}$$

Hence in that case $V \cong \dfrac{V}{W}$ or $V \cong \dfrac{V}{(0)}$.

Note $W = (0) \Rightarrow \text{Ker}\ \theta = (0) \Rightarrow \theta$ is one-one.

Example 4.6: Let W and U be subspaces of $V(F)$ such that $W \subset U \subset V$. Let $f : V \to V/W$ be the quotient map. Show that $f(U)$ is a proper subspace of V/W.

Solution: Since f is a linear transformation, $f(U)$ is a subspace of V/W.

If $f(U) = 0$ then $f(x) = 0$ for all $x \in U$

$$\Rightarrow \quad W + x = W \quad \text{for all } x \in U$$

$\Rightarrow \quad x \in W$ for all $x \in U$

$\Rightarrow \quad U \subseteq W$, a contradiction

Again since $U \neq V$, $\exists\, v_0 \in V$ such that, $v_0 \notin U$.

NOTES

If $f(v_0) \in f(U)$ then $f(v_0) = f(x)$ for some $x \in U$

$\Rightarrow f(v_0 - x) = 0$

$\Rightarrow W + (v_0 - x) = W$

$\Rightarrow v_0 - x \in W$

$\Rightarrow v_0 = x + w$ for some $w \in W$

$\Rightarrow v_0 \in U$, a contradiction

Hence $\quad f(v_0) \notin f(U) \Rightarrow f(U) \neq \dfrac{V}{W}$

or that $f(U)$ is a proper subspace of $\dfrac{V}{W}$.

Linear Span and Finite Dimensional Vector Space (FDVS)

Definition: Let $V(F)$ be a vector space, $v_i \in V$, $\alpha_i \in F$ be elements of V and F respectively. Then elements of the type $\sum_{i=1}^{n} \alpha_i v_i$ are called *linear combinations* of $v_1, v_2, ..., v_n$ over F.

Let S be a non-empty subset of V, then the set

$$L(S) = \left\{ \sum_{i=1}^{n} \alpha_i v_i \mid \alpha_i \in F, v_i \in S, n \text{ finite} \right\}$$

i.e., the set of all linear combinations of finite sets of elements of S is called *linear span* of S. It is also denoted by $< S >$. If $S = \varphi$, define $L(S) = \{0\}$.

Theorem 4.12: $L(S)$ is the smallest subspace of V, containing S.

Proof: $L(S) \neq \varphi$ as $v \in S \Rightarrow v = 1 \,.\, v,\ 1 \in F$

$\Rightarrow v \in L(S)$

thus, in fact, $S \subseteq L(S)$.

Let $\quad x, y \in L(S)$, $\alpha, \beta \in F$ be any elements

then $\quad x = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$

$y = \beta_1 v_1' + \beta_2 v_2' + ... + \beta_m v_m' \qquad v_i, v_j' \in S, \quad \alpha_i, \beta_j \in F$

Thus $\quad \alpha x + \beta y = \alpha\alpha_1 v_1 + \alpha\alpha_2 v_2 + ... + \alpha\alpha_n v_n + \beta\beta_1 v_1' + ... + \beta\beta_m v_m'$.

R.H.S. being a linear combination belongs to $L(S)$.

Hence $L(S)$ is a subspace of V, containing S.

Let now W be any subspace of V, containing S

We show $L(S) \subseteq W$

$x \in L(S) \Rightarrow x = \Sigma\alpha_i v_i \qquad v_i \in S, \alpha_i \in F$

$v_i \in S \subseteq W$ for all i and W is a subspace

$\Rightarrow \Sigma\alpha_i v_1 \in W \Rightarrow x \in W$

$\Rightarrow L(S) \subseteq W$

Hence the result follows.

NOTES

Theorem 4.13: If S_1 and S_2 are subsets of V, then

(*i*) $S_1 \subseteq S_2 \Rightarrow L(S_1) \subseteq L(S_2)$

(*ii*) $L(S_1 \cup S_2) = L(S_1) + L(S_2)$

(*iii*) $L(L(S_1)) = L(S_1)$.

Proof: (*i*) $x \in L(S_1) \Rightarrow x = \Sigma\alpha_i v_i \quad v_i \in S_1, \alpha_i \in F$

thus $\quad v_i \in S_1 \subseteq S_2$ for all i

$\Rightarrow \quad \Sigma\alpha_i v_i \in S_2 \Rightarrow x \in L(S_2)$

$\Rightarrow \quad L(S_1) \subseteq L(S_2)$.

(*ii*) $\quad S_1 \subseteq S_1 \cup S_2 \Rightarrow L(S_1) \subseteq L(S_1 \cup S_2)$

$\quad S_2 \subseteq S_1 \cup S_2 \Rightarrow L(S_2) \subseteq L(S_1 \cup S_2)$

$\Rightarrow \quad L(S_1) + L(S_2) \subseteq L(S_1 \cup S_2)$

Again, $\quad S_1 \subseteq L(S_1) \subseteq L(S_1) + L(S_2)$

$\quad S_2 \subseteq L(S_2) \subseteq L(S_1) + L(S_2)$

$\Rightarrow \quad S_1 \cup S_2 \subseteq L(S_1) + L(S_2)$.

Hence $L(S_1 \cup S_2) \subseteq L(S_1) + L(S_2)$

as $L(S_1 \cup S_2)$ is the smallest subspace containing $S_1 \cup S_2$ and $L(S_1) + L(S_2)$ is a subspace, being sum of two subspaces (and contains $S_1 \cup S_2$).

Thus $\quad L(S_1 \cup S_2) = L(S_1) + L(S_2)$.

(*iii*) Let $\quad L(S_1) = K$ then we show $L(K) = L(S_1)$

Now $\quad K \subseteq L(K) \;\therefore\; L(S_1) \subseteq L(L(S_1))$

Again $x \in L(L(S_1)) \Rightarrow x$ is linear combination of members of $L(S_1)$ which are linear combinations of members of S_1.

So x is a linear combination of members of S_1

$\Rightarrow \quad x \in L(S_1)$

Thus $\quad L(L(S_1)) \subseteq L(S_1)$

Hence $\quad L(L(S_1)) = L(S_1)$.

Theorem 4.14: If W is a subspace of V, then $L(W) = W$ and conversely.

Proof: $W \subseteq L(W)$ by definition and since $L(W)$ is the smallest subspace of V containing W and W is itself a subspace.

$\quad L(W) \subseteq W$

Hence $\quad L(W) = W$.

Conversely, let $L(W) = W$

Let $\quad x, y \in W, \alpha, \beta \in F$

Then $\quad x, y \in L(W)$

$\Rightarrow$ x, y are linear combinations of members of W.

$\Rightarrow$ $\alpha x + \beta y$ is a linear combination of members of W

$\Rightarrow \alpha x + \beta y \in L(W)$

$\Rightarrow \alpha x + \beta y \in W$

$\Rightarrow$ W is a subspace.

NOTES

Definition: If $V = L(S)$, we say S spans (or generates) V. The vector space V is said to be *finite-dimensional* (over F) if there exists a finite subset S of V such that are

$V = L(S)$. We use notation FDVS for a finite dimensional vector space.

From the results, it is proved that

If S_1 and S_2 are two subspaces of V, then $S_1 + S_2$ is the subspace spanned by $S_1 \cup S_2$

Indeed, $L(S_1 \cup S_2) = L(S_1) + L(S_2) = S_1 + S_2$.

Example 4.7: Let $S = \{(1, 4), (0, 3)\}$ be a subset of $\mathbf{R}^2(\mathbf{R})$. Show that (2, 3) belongs to $L(S)$.

Solution: $(2, 3) \in L(S)$ if it can be put as a linear combination of (1, 4) and (0, 3).

$$\begin{aligned}
\text{Now} \qquad (2, 3) &= \alpha(1, 4) + \beta(0, 3) \\
\Rightarrow \qquad (2, 3) &= (\alpha + 0, 4\alpha + 3\beta) \\
\Rightarrow \qquad 2 &= \alpha,\ 4\alpha + 3\beta = 3 \\
\Rightarrow \qquad \alpha &= 2,\ \beta = -\frac{5}{3} \\
\text{Hence} \qquad (2, 3) &= 2\,(1, 4) - \frac{5}{3}\,(0, 3) \\
\text{Showing that} \qquad (2, 3) &\in L(S).
\end{aligned}$$

Example 4.8: Let $V = \mathbf{R}^4(\mathbf{R})$ and let $S = \{(2, 0, 0, 1), (-1, 0, 1, 0)\}$. Find $L(S)$.

Solution: Any element $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \in L(S)$ is a linear combination of members of S.

$$\begin{aligned}
\text{Let} \quad (\alpha_1, \alpha_2, \alpha_3, \alpha_4) &= \alpha(2, 0, 0\ 1) + \beta(-1, 0, 1, 0), \quad \alpha, \beta \in \mathbf{R} \\
\text{then} \quad (\alpha_1, \alpha_2, \alpha_3, \alpha_4) &= (2\alpha - \beta, 0, \beta, \alpha) \\
\text{i.e.,} \qquad L(S) &= \{(2\alpha - \beta, 0, \beta, \alpha) \mid \alpha, \beta \in \mathbf{R}\}
\end{aligned}$$

Example 4.9: Show that the vector space $F[x]$ is not finite dimensional.

Solution: Let $V = F[x]$ and suppose it is finite dimensional.

Then $\exists\, S \subseteq V$, such that, $V = L(S)$ and S is finite.

Suppose $\quad S = \{p_1, p_2, ..., p_k\}$. We can assume $p_i \neq 0 \quad \forall i$

Let deg $p_i = r_i$ and let $t = \text{Max}\ \{r_1, r_2, ..., r_k\}$

Now $x^{t+1} \in V$ and since $V = L(S)$,

$$x^{t+1} = \alpha_1 p_1 + \alpha_2 p_2 + ... + \alpha_k p_k, \quad \alpha_i \in F$$

So $\qquad 0 = (-1)\, x^{t+1} + \alpha_1 p_1 + ... + \alpha_k p_k$

Since x^{t+1} does not appear in $p_1, p_2, ..., p_k$

We get $-1 = 0$, a contradiction. Hence V is not FDVS over F.

Note if $\quad S = \{1, x, ..., x^n, ...\}$ then $V = L(S)$.

NOTES

Linear Dependence and Independence: Basic Properties

Let $V(F)$ be a vector space. Elements $v_1, v_2, ..., v_n$ in V are said to be linearly dependent (over F) if $\exists$ scalars $\alpha_1, \alpha_2, ... \alpha_n \in F$, (not all zero) such that,

$$\alpha_1 v_1 + \alpha_2 v_2 + ... \alpha_n v_n = 0$$

($v_1, v_2, ..., v_n$ are finite in number, not essentially distinct).

Thus for linear dependence $\Sigma \alpha_i v_i = 0$ and at least one $\alpha_i \neq 0$.

If $v_1, v_2 ... v_n$ are not linearly dependent (*LD*), these are called linearly independent (*LI*)

In other words, $v_1, v_2, .., v_n$ are *LI* if

$$\Sigma \alpha_i v_i = 0 \Rightarrow \alpha_i = 0 \text{ for all } i$$

A finite set $X = \{x_1, x_2 ..., x_n\}$ is said to be *LD* or *LI* according as its n members are *LD* or *LI*

In general any subset Y of $V(F)$ is called *LI* if every finite non-empty subset of Y is *LI*, *otherwise* it is called *LD*

So, if some subsets are *LI* and some are *LD* then Y is called *LD*

Observations: (*i*) A non-zero vector is always *LI* as $v \neq 0$, $\alpha v = 0$ would mean $\alpha = 0$.

(*ii*) Zero vector is always *LD*

$$1 \,.\, 0 = 0 \qquad 1 \neq 0,\ 1 \in F$$

Thus, any collection of vectors to which zero belongs is always *LD*

In other words, if $v_1, v_2, ..., v_n$ are *LI* then none of these can be zero. (But not conversely, Refer example ahead).

(*iii*) v is *LI* iff $v \neq 0$.

(*iv*) Any subset of a *LI* set is *LI*

(*v*) Any super set of a *LD* set is *LD*

(*vi*) Empty set φ is *LI* since it has no non-empty finite subset and consequently it satisfies the condition for linear independence. In other words, whenever $\Sigma \alpha_i v_i = 0$ in φ then as there is no i for which $\alpha_i \neq 0$, set φ is *LI* We sometimes express it by saying that empty set is *LI* vacuously.

(*vii*) A set of vector is *LI* if and only if every finite subset of it is *LI*

Some examples of linear dependence and independence are given as follows:

(i) Consider $\mathbf{R}^2(\mathbf{R})$, $\mathbf{R}$ = reals.

$$v_1 = (1, 0),\ v_2 = (0, 1) \in \mathbf{R}^2 \text{ are } LI$$

as $\quad \alpha_1 v_1 + \alpha_2 v_2 = 0$ for $\alpha_1, \alpha_2 \in \mathbf{R}$

$$\Rightarrow \alpha_1(1, 0) + \alpha_2(0, 1) = (0, 0)$$

$$\Rightarrow (\alpha_1, \alpha_2) = (0, 0) \Rightarrow \alpha_1 = \alpha_2 = 0.$$

(ii) Consider the subset

$S = \{(1, 0, 0), (0, 1, 0), (0, 0, 1), (2, 3, 4)\}$ in the vector space $\mathbf{R}^3(\mathbf{R})$.

Since $\quad 2(1, 0, 0) + 3(0, 1, 0) + 4(0, 0, 1) - 1(2, 3, 4) = (0, 0, 0)$

we find S is *LD*

NOTES

(iii) In the vector space $F[x]$ of polynomials the vectors $f(x) = 1 - x$, $g(x) = x - x^2$, $h(x) = 1 - x^2$ are LD since $f(x) + g(x) - h(x) = 0$.

Example 4.10: Show that the vectors $v_1 = (0, 1, -2)$, $v_2 = (1, -1, 1)$, $v_3 = (1, 2, 1)$ are LI *in* $\mathbf{R}^3(\mathbf{R})$.

Solution: Let $\Sigma\alpha_i v_i = 0$ for $\alpha_i \in \mathbf{R}$

Then $\alpha_1(0, 1, -2) + \alpha_2 (1, -1, 1) + \alpha_3 (1, 2, 1) = (0, 0, 0)$

$\Rightarrow \quad (0, \alpha_1, -2\alpha_1) + (\alpha_2, -\alpha_2, \alpha_2) + (\alpha_3, 2\alpha_3, \alpha_3) = (0, 0, 0)$

$$\Rightarrow \quad \begin{aligned} 0 + \alpha_2 + \alpha_3 &= 0 \\ \alpha_1 - \alpha_2 + 2\alpha_3 &= 0 \\ -2\alpha_1 + \alpha_2 + \alpha_3 &= 0 \end{aligned}$$

Since the coefficient determinant $\begin{vmatrix} 0 & 1 & 1 \\ 1 & -1 & 2 \\ -2 & 1 & 1 \end{vmatrix}$ is $-6 \neq 0$ the above equations have only the zero common solution

$$\Rightarrow \alpha_1 = \alpha_2 = \alpha_3 = 0 \Rightarrow v_1, v_2, v_3 \text{ are } LI$$

Example 4.11: Show that $\{f(x), g(x), h(x)\}$ is LI in $F[x]$, whenever. $deg\, f(x)$, $deg\, g(x)$, $deg\, h(x)$ are distinct.

Solution: Let
$$\begin{aligned} f(x) &= a_0 + a_1x + ... + a_mx^m, \quad a_m \neq 0 \\ g(x) &= b_0 + b_1x + ... + b_nx^n, \quad b_n \neq 0 \\ h(x) &= c_0 + c_1x + ... + c_tx^t, \quad c_t \neq 0 \end{aligned}$$

Let $\alpha f(x) + \beta g(x) + \gamma h(x) = 0, \quad \alpha, \beta, \gamma \in F$

Let $m < n < t$ (without any loss of generality)

then $\gamma c_t = 0 \Rightarrow \gamma = 0$ as $c_t \neq 0$

$\therefore \quad \alpha f(x) + \beta g(x) = 0$

and so $\beta b_n = 0 \Rightarrow \beta = 0$ as $b_n \neq 0$

$$\Rightarrow \alpha f(x) = 0 \Rightarrow \alpha a_m = 0 \Rightarrow \alpha = 0 \text{ as } a_m \neq 0$$

Hence $\{f(x), g(x), h(x)\}$ is LI in $F[x]$ over F.

Example 4.12: Show that the vectors

$v_1 = (1, 1, 2, 4)$, $v_2 = (2, -1, -5, 2)$, $v_3 = (1, -1, -4, 0)$ and $v_4 = (2, 1, 1, 6)$ are LD in $\mathbf{R}^4(\mathbf{R})$.

Solution: Suppose $av_1 + bv_2 + cv_3 + dv_4 = 0, \ a, b, c, d \in \mathbf{R}$

then $a(1, 1, 2, 4) + b(2, -1, -5, 2) + c(1, -1, -4, 0)$
$$+ d(2, 1, 1, 6) = (0, 0, 0, 0)$$

or $(a, a, 2a, 4a) + (2b, -b, -5b, 2b) + (c, -c, -4c, 0)$
$$+ (2d, d, d, 6d) = (0, 0, 0, 0)$$

$$\Rightarrow \quad \begin{aligned} a + 2b + c + 2d &= 0 \\ a - b - c + d &= 0 \\ 2a - 5b - 4c + d &= 0 \\ 4a + 2b + 0c + 6d &= 0 \end{aligned}$$

NOTES

$$\Rightarrow \quad \begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & -1 & -1 & 1 \\ 2 & -5 & -4 & 1 \\ 4 & 2 & 0 & 6 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \to R_2 - R_1,\ R_3 \to R_3 - 2R_1,\ R_4 \to R_4 - 4R_1$

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & -3 & -2 & -1 \\ 0 & -3 & -2 & -1 \\ 0 & -3 & -2 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_4 \to \frac{1}{2} R_4,\ R_3 \to \frac{1}{3} R_3$

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & -3 & -2 & -1 \\ 0 & -1 & -2/3 & -1/3 \\ 0 & -3/4 & -1 & -1/2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_4 \to R_4 - R_2,\ R_3 \to R_3 - R_2$

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & -3 & -2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad \begin{aligned} a + 2b + c + 2d &= 0 \\ -3b - 2c + d &= 0 \\ 3b + 2c + d &= 0 \end{aligned}$$

$a = -1,\ b = -1,\ c = 1,\ d = 1$ satisfy the equations.

Since coefficients are non-zero, the given vectors are *LD*

Example 4.13: Show that

(*i*) $\{1, \sqrt{2}\}$ is *LI* in $\mathbf{R}$ over $\mathbf{Q}$.

(*ii*) $\{1, \sqrt{2}, \sqrt{3}\}$ is LI in $\mathbf{R}$ over $\mathbf{Q}$.

(*iii*) $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ is LI in $\mathbf{R}$ over $\mathbf{Q}$.

Solution: (*i*) Suppose $a + b\sqrt{2} = 0,\ a, b \in \mathbf{Q}$

Suppose $b \neq 0$, then $\sqrt{2} = -\frac{a}{b} \in \mathbf{Q}$, a contradiction

Hence $b = 0$ and so $a = 0$. Thus $\{1, \sqrt{2}\}$ is *LI* in $\mathbf{R}$ over $\mathbf{Q}$.

(*ii*) Let $a + b\sqrt{2} + c\sqrt{3} = 0, \quad a, b, c \in \mathbf{Q}$

Let $c \neq 0$, then

$$\sqrt{3} = -\frac{a}{c} - \frac{b}{c}\sqrt{2} = \alpha + \beta\sqrt{2}, \quad \alpha, \beta \in \mathbf{Q}$$

$$\Rightarrow \quad 3 = \alpha^2 + 2\beta^2 + 2\alpha\beta\sqrt{2}$$

$$\Rightarrow \quad \alpha\beta\sqrt{2} \in \mathbf{Q} \Rightarrow \alpha\beta = 0$$

Let $\alpha = 0$ then $\beta = \sqrt{\frac{3}{2}}$, a contradiction

So, $c = 0$ giving $a + b\sqrt{2} = 0 \Rightarrow a = b = 0$ by (*i*)

NOTES

Hence the result follows.

(*iii*) Let $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} = 0, \quad a, b, c, d \in \mathbf{Q}$

Then $(a + b\sqrt{2}) + \sqrt{3}(c + d\sqrt{2}) = 0$

Let $\quad c + d\sqrt{2} \neq 0$

Then $\quad \sqrt{3} = \dfrac{-(a + b\sqrt{2})}{(c + d\sqrt{2})} = \dfrac{-(a + b\sqrt{2})(c - d\sqrt{2})}{c^2 - 2d^2}$

$= \alpha + \beta\sqrt{2}, \quad \alpha, \beta \in \mathbf{Q}$

$\Rightarrow \quad \alpha.1 + \beta\sqrt{2} + (-1)\sqrt{3} = 0$

$\Rightarrow \quad -1 = 0$ by (*ii*), a contradiction

$\therefore \quad c + d\sqrt{2} = 0 \Rightarrow c = d = 0 \Rightarrow a + b\sqrt{2} = 0$

$\Rightarrow \quad a = b = 0$

Hence the result follows.

Theorem 4.15: If $S = \{v_1, v_2, ... v_n\}$ is a basis of *V,* then every element of V can be expressed uniquely as a linear combination of $v_1, v_2, ..., v_n$.

Proof: Since, by definition of basis, $V = L(S)$, each element $v \in V$ can be expressed as linear combination of $v_1, v_2, ..., v_n$.

Suppose $\quad v = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n, \quad \alpha_i \in F$

$v = \beta_1 v_1 + \beta_2 v_2 + ... + \beta_n v_n, \quad \beta_i \in F$

then $\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = \beta_1 v_2 + \beta_2 v_2 + ... + \beta_n v_n$

$\Rightarrow (\alpha_1 - \beta_1)\, v_1 + (\alpha_2 - \beta_2)\, v_2 + ... + (\alpha_n - \beta_n)\, v_n = 0$

$\Rightarrow \alpha_i - \beta_i = 0$ for all i ($v_1, v_2, ... v_n$ are *LI*)

$\Rightarrow \alpha_i = \beta_i$ for all i.

4.4 RIESZ'S LEMMA

Riesz's lemma, named after Frigyes Riesz, is a functional analysis lemma. It defines the circumstances under which a dense subspace in a normed vector space is guaranteed. The lemma may also be called the **Riesz lemma** or **Riesz inequality** the Riesz rearrangement inequality (also called Riesz-Sobolev inequality) states that for any three non-negative functions $h : \mathbb{R}^n \to \mathbb{R}^+, g : \mathbb{R}^n \to \mathbb{R}^+$ and $h : \mathbb{R}^n \to \mathbb{R}^+$ satisfies the inequality.

$$\iint_{\mathbb{R}^n + \mathbb{R}^n} f(x)g(x-y)h(y)dxdy \leq \iint_{\mathbb{R}^n + \mathbb{R}^n} f^*(x)g^*(x-y)h^*(y)dxdy,$$

It can be seen as a substitute for orthogonality when one is not in an inner product space.

Riesz's Lemma: Let X be a normed space, Y be a closed proper subspace of X and α be a real number with $0 < \alpha < 1$. Then there exists an x in X with $|x| = 1$ such that $|x - y| \geq \alpha$ for all y in Y.

1. For the finite-dimensional case, equality can be achieved, or there exists x of unit norm such that $d(x, Y) = 1$. When dimension of X is finite, the unit ball $B \subset X$ is compact. Also, the distance function $d(\cdot\,, Y)$ is continuous. Therefore its image on the unit ball B must be a compact subset of the real line, proving the claim.
2. The space ℓ_∞ of all bounded sequences shows that the lemma does not hold for $\alpha = 1$.

Consequences of Riesz's Lemma

Compact operators working on a Banach space have spectral features similar to matrices. Riesz's lemma is crucial in proving this point.

Riesz's lemma proved that any infinite-dimensional normed space contains a sequence of unit vectors $\{x_n\}$ with $|x_n - x_m| > \alpha$ for $0 < \alpha < 1$. This is useful in showing the non-existence of certain measures on infinite-dimensional Banach spaces. Riesz's lemma also shows that the identity operator on a Banach space X is compact if and only if X is finite-dimensional.

This lemma may also be used to define finite dimensional normed spaces: If X is a normed vector space, then X is finite dimensional if and only if X closed unit ball is compact.

Finite Dimensional Characterization

Riesz's lemma can be applied directly to show that the unit ball of an infinite-dimensional normed space X is never compact: Take an element x_1 from the unit sphere. Pick x_n from the unit sphere such that $d(x_n, Y_{n-1}) > \alpha$ for a constant $0 < \alpha < 1$, where Y_{n-1} is the linear span of $\{x_1 \dots x_{n-1}\}$ and $d(x_n, Y) = \inf_{y \in Y} |x_n - y|$.

Clearly $\{x_n\}$ contains no convergent subsequence and the noncompactness of the unit ball follows.

In general, a topological vector space X is finite dimensional if it is locally compact. This also holds true in reverse. A topological vector space is locally compact if it has a limited dimension. Therefore local compactness characterizes finite-dimensionality. This classical result is also attributed to Riesz. A short proof can be sketched as follows: Let C be a compact neighbourhood of $0 \in X$. By compactness, there are $c_1, \dots, c_n \in C$ such that,

$$C \subset \bigcup_{i=1}^{n} \left(c_i + \frac{1}{2} C \right).$$

We assert that the finite-dimensional subspace Y traversed by $\{c_i\}$ is dense in X, or that it is closed by X. Since X is the union of scalar multiples of C, it is sufficient to show that $C \subset Y$. Now, by induction,

$$C \subset Y + \frac{1}{2^m} C$$

for every m. But compact sets are bounded, so C lies in the closure of Y. This proves the result.

NOTES

4.5 ADJOINT OPERATORS

Definition: Consider two Hilbert spaces $\mathbf{H}_1$ and $\mathbf{H}_2$. Let $\mathbf{T} : \mathbf{H}_1 \to \mathbf{H}_2$ be a bounded linear operator. Then the Hilbert-adjoint operator $\mathbf{T}^*$ of T is operator $\mathbf{T}^* : \mathbf{H}_2 \to \mathbf{H}_1$ such that for all $x \in \mathbf{H}_1$ and $y \in \mathbf{H}_2$,

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \qquad \ldots(4.1)$$

Theorem 4.16: The Hilbert-adjoint operator **T*** of T exists, is unique and is a bounded linear oeprator with norm given by,

$$\| \mathbf{T}^* \| = \| \mathbf{T}^* \| \qquad \ldots(4.2)$$

Proof: Let $h(y, x) = \langle y, \mathbf{T}x \rangle$ $\qquad \ldots(4.3)$

We will show that h is sesquilinear. Now h is linear in the first argument and conjugate linear in the second argument since,

$$\begin{aligned} h(y, \alpha x_1 + \beta x_2) &= \langle y, \mathbf{T}(\alpha x_1 + \beta x_2) \rangle \\ &= \langle y, \alpha \mathbf{T}x_1 + \beta \mathbf{T}x_2 \rangle \\ &= \bar{\alpha} \langle y, \mathbf{T}x_1 \rangle + \bar{\beta} \langle y, \mathbf{T}x_2 \rangle \\ &= \bar{\alpha}\, h(y, x_1) + \bar{\beta}\, h(y, x_2). \end{aligned}$$

Hence sicne the inner product is sesquilinear, we infer that h is sesquilinear. By the Cauchy-Schwarz inequality,

$$|h(y, x)| = |\langle y, \mathbf{T}x \rangle| \le \| y \| \, \| \mathbf{T}x \| \le \| \mathbf{T} \| \, \|x\| \, \|y\|$$

Implying $\dfrac{|h(y, x)}{\|x\| \, \|y\|} \le \| \mathbf{T} \|$

We have,

$$\|h\| \le \| \mathbf{T} \| \qquad \ldots(4.4)$$

Also,

$$\| h \| = \sup_{\substack{x \neq 0 \\ y \neq 0}} \frac{|\langle y, \mathbf{T}x \rangle|}{\| y \| \, \| x \|} \ge \sup_{\substack{x \neq 0 \\ \mathbf{T}x \neq 0}} \frac{|\langle \mathbf{T}x, \mathbf{T}x \rangle|}{\| \mathbf{T}x \| \, \| x \|} = \| \mathbf{T} \| \qquad \ldots(4.5)$$

Combining Equations (4.4) and (4.5) gives $\|h\| = \|\mathbf{T}\|$. From Riesz representation theorem, substituting **T*** for S, we have

$$h(y, x) = \langle \mathbf{T}^* y, x \rangle \qquad \ldots(4.6)$$

where $\mathbf{T}^* : \mathbf{H}_2 \to \mathbf{H}_1$ is a uniquely determined bounded linear operator with norm,

$$\|\mathbf{T}^*\| = \|h\| = \|\mathbf{T}\| \qquad \ldots(4.7)$$

Combining Equations (4.3) and (4.6), we get

$$\langle y, \mathbf{T}x \rangle = \langle \mathbf{T}^* y, x \rangle$$

Taking the conjugate gives Equation (4.1).

NOTES

Lemma 1: Let X and Y be inner product spaces and $\mathbf{T}^* : \mathbf{X} \to \mathbf{Y}$ a bounded linear operator. Then,

1. $T = 0$ if and only if $\langle \mathbf{T}x, y \rangle = 0$ for all $x \in \mathbf{X}$ and $y \in \mathbf{Y}$.
2. For X, a complex vector space, if $\mathbf{T} : \mathbf{X} \to \mathbf{X}$ and $\langle \mathbf{T}x, x \rangle = 0$ for all $x \in \mathbf{X}$, then $T = 0$.

Proof: (1) If $T = 0$, then for all $x \in \mathbf{X}$, $\mathbf{T}x = 0$ and for any $u \in \mathbf{X}$ we have,

$$\langle \mathbf{T}x, y \rangle = 0 \ \langle 0, y \rangle = 0 \ \langle u, y \rangle = 0$$

Now let,

$$\langle \mathbf{T}x, y \rangle = 0 \text{ for all } x \in \mathbf{X},\ y \in \mathbf{Y}$$

Then $\mathbf{T}x = 0$ for all $x \in \mathbf{X}$ and $\mathbf{T} = 0$.

(2) If $\langle \mathbf{T}x, x \rangle = 0$ for all $x \in \mathbf{X}$, then for $w = \alpha x + y \in$ x we have,

$$\begin{aligned} \langle \mathbf{T}w, w \rangle &= \langle \mathbf{T}(\alpha x + y), \alpha x + y \rangle \\ &= |\alpha|^2 \langle \mathbf{T}x, x \rangle + \langle \mathbf{T}y, y \rangle + \alpha \langle \mathbf{T}x, y \rangle + \bar{\alpha} \langle \mathbf{T}y, x \rangle \end{aligned} \quad \ldots(4.8)$$

Now if we pick $\alpha = 1$, then Equation (4.8) becomes

$$\langle \mathbf{T}w, w \rangle = \langle \mathbf{T}x, x \rangle + \langle \mathbf{T}y, y \rangle + \langle \mathbf{T}x, y \rangle + \langle \mathbf{T}y, x \rangle \quad \ldots(4.9)$$

Now $\langle \mathbf{T}x, x \rangle$ and $\langle \mathbf{T}y, y \rangle$ are equal to 0 by our assumption. Hence Equation (4.9) becomes,

$$\langle \mathbf{T}x, y \rangle + \langle \mathbf{T}y, x \rangle = 0 \quad \ldots(4.10)$$

Adding Equations (4.9) and (4.10) we obtain $\langle \mathbf{T}x, y \rangle = 0$, and $\mathbf{T} = 0$ follows from (1).

Theorem 4.17: Consider two Hilbert spaces $\mathbf{H}_1$ and $\mathbf{H}_2$. Let $\mathbf{S} : \mathbf{H}_1 \to \mathbf{H}_2$ and $\mathbf{T} : \mathbf{H}_1 \to \mathbf{H}_2$ be bounded linear operators and β any scalar. Then,

1. $\langle \mathbf{T}^* y, x \rangle = \langle y, \mathbf{T}x \rangle$ for any $x \in \mathbf{H}_1, y \in \mathbf{H}_2$,
2. $(\mathbf{S} + \mathbf{T})^* = \mathbf{S}^* + \mathbf{T}^*$,
3. $(\beta\mathbf{T})^* = \bar{\beta}\mathbf{T}^*$,
4. $(\mathbf{T}^*)^* = \mathbf{T}$,
5. $\|\mathbf{T}^*\mathbf{T}\| = \|\mathbf{T}\mathbf{T}^*\| = \|\mathbf{T}\|^2$,
6. $\mathbf{T}^*\mathbf{T} = 0$ if and only if $\mathbf{T} = 0$,
7. $(\mathbf{ST})^* = \mathbf{T}^*\mathbf{S}^*$ (if $\mathbf{H}_2 = \mathbf{H}_1$).

Proof:

1. From the definition of Hilbert-adjoint operator, we have

$$\langle \mathbf{T}^* y, x \rangle = \overline{\langle x, \mathbf{T}^* y \rangle} = \overline{\langle \mathbf{T}x, y \rangle} = \langle y, \mathbf{T}x \rangle \quad \ldots(4.11)$$

NOTES

2. By the definition of the Hilbert-adjoint operator, for all x and y,

$$\begin{aligned}\langle x, (\mathbf{S}+\mathbf{T})^* y\rangle &= \langle (\mathbf{S}+\mathbf{T})x, y\rangle \\ &= \langle \mathbf{S}x, y\rangle + \langle \mathbf{T}x, y\rangle \\ &= \langle x, \mathbf{S}^* y\rangle + \langle x, \mathbf{T}^* y\rangle \\ &= \langle x, (\mathbf{S}^* + \mathbf{T}^*)y\rangle .\end{aligned}$$

Thus it follows from the equality of the inner product spaces that $(\mathbf{S}+\mathbf{T})^*$ $y = (\mathbf{S}^* + \mathbf{T}^*)\, y$ for all $y \in \mathbf{H}_2$ so that $(\mathbf{S}+\mathbf{T})^* = \mathbf{S}^* + \mathbf{T}^*$.

3. By the definition of the Hilbert-adjoint operator,

$$\begin{aligned}\langle (\beta\mathbf{T})^* y, x\rangle &= \langle y, (\beta\mathbf{T})\, x\rangle \\ &= \langle y, \beta(\mathbf{T}x)\rangle \\ &= \bar{\beta}\langle y, \mathbf{T}x\rangle \\ &= \bar{\beta}\langle \mathbf{T}^* y, x\rangle \\ &= \langle \bar{\beta}\mathbf{T}^* y, x\rangle .\end{aligned}$$

Hence by Lemma 1, $(\beta\mathbf{T})^* y = \bar{\beta}\mathbf{T}^* y$ for all $y \in \mathbf{H}_2$, which implies that $(\beta\mathbf{T})^* = \bar{\beta}\mathbf{T}^*$.

4. By the definition of the Hilbert-adjoint operator, and from (1) we have, $\langle (\mathbf{T}^*)^* x, y\rangle = \langle x, \mathbf{T}^* y\rangle = \langle \mathbf{T}x, y\rangle$ so that $\langle ((\mathbf{T}^*)^* - \mathbf{T})x, y\rangle$ and by Lemma 1, we have $(\mathbf{T}^*)^* = \mathrm{T}$.

5. We know that $\mathbf{T}^*\mathbf{T} : \mathbf{H}_1 \to \mathbf{H}_1$ and $\mathbf{T}\mathbf{T}^* : \mathbf{H}_2 \to \mathbf{H}_2$. By the Cauchy-Schwarz inequality and by the definition of the Hilbert-adjoint operator in the definition of the Hilbert-adjoint operator we have,

$\|\mathbf{T}x\|^2 = \langle \mathbf{T}x, \mathbf{T}x\rangle = \langle \mathbf{T}^*\mathbf{T}x, x\rangle \le \|\mathbf{T}^*\mathbf{T}x\|\, \|x\| \le \|\mathbf{T}^*\mathbf{T}\|\, \|x\|^2$

Taking the supremum over all x of norm 1 we obtain $\|\mathbf{T}\|^2 \le \|\mathbf{T}^*\mathbf{T}\|$. Now by Theorem 4.16, (Refer Unit 5) we have $\|\mathbf{T}^*\mathbf{T}\| \le \|\mathbf{T}^*\| \|\mathbf{T}\| = \|\mathbf{T}\|^2$. Hence $\|\mathbf{T}^*\mathbf{T}\| = \|\mathbf{T}\|^2$. Substituting $\mathbf{T}^*$ for T we get $\|\mathbf{T}^{**}\mathbf{T}^*\| = \|\mathbf{T}^*\|^2 = \|\mathbf{T}\|^2$. But by (4) we have $(\mathbf{T}^*)^* = \mathbf{T}$ sot that $\|\mathbf{T}\mathbf{T}^*\| = \|\mathbf{T}\|^2$.

6. From (5), if $\mathbf{T}^*\mathbf{T} = 0$, then $\mathbf{T} = 0$ and conversely if $\mathbf{T} = 0$ then $\mathbf{T}^*\mathbf{T} = 0$.

7. By the definition of the Hilbert-adjoint operator,

$\langle x, (\mathbf{ST})^* y\rangle = \langle (\mathbf{ST})x, y\rangle = \langle \mathbf{T}x, \mathbf{S}^* y\rangle = \langle x, \mathbf{T}^*\mathbf{S}^* y\rangle$. Hence by equality of inner product spaces we obtain $(\mathbf{ST})^* y = \mathbf{T}^*\mathbf{S}^* y$ for all $y \in \mathbf{H}_2$ giving $(\mathbf{ST})^* = \mathbf{T}^*\mathbf{S}^*$.

4.5.1 Reflexive Spaces

Let X be a normed space and $X^{**} = (X^*)^*$ denote the second dual space of X. The canonical map $x \to \hat{x}$ defined by $\hat{x}(f) = f(x), f \in X^*$ gives an isometric linear isomorphism (embedding) from X into X^{**}. The space is called reflexive if this map is surjective. This concept was introduced by Hahn in 1927.

For example, finite dimensional (normed) spaces and Hilbert spaces are reflexive. The space of absolutely summable complex sequences is not reflexive. James constructed a non-reflexive Banach space that is isometrically isomorphic to its second conjugate space.

NOTES

4.6 FINITE DIMENSIONAL NORMED SPACES AND SUBSPACES

Definition

A **norm** on a linear space X is a function $\| \quad \|: X \to R$ satisfying

1. $\| x \| \geq 0$ and $\| x \| = 0$ if and only if $x = 0$ for $x \in X$
2. $\| \alpha x \| = |\alpha| \cdot \| x \|$
3. $\| x + y \| \leq \| x \| + \| y \|$

We observe that a semi-norm becomes a norm if it satisfies one additional condition i.e.

$$\| x \| = 0 \text{ iff } x = 0$$

Further, $\| x \|$ is called norm of x. The non-negative real number $\| x \|$ is taken as the length of the vector x. A normed linear space is an ordered pair $(X, \| \cdot \|)$ where $\| \cdot \|$ is a norm on X.

Theorem 4.18

Every finite dimensional subspace Y of a normed space X is complete. Particularly, every finite dimensional normed space is complete. **To prove the theorem, we prove a Lemma.**

Lemma

Let $\{x_1, x_2, \dots \dots \dots, x_n\}$ be a linearly independent set of vectors in a normed space X (of any dimension). Then there is a number $C > 0$ such that for every choice of scalars $\alpha_1, \alpha_2, \dots \dots \dots, \alpha_n$, we have

$$\|\alpha_1 x_1 + \cdots \dots \dots + \alpha_n x_n\| \geq C(|\alpha_1| + \cdots \dots . + |\alpha_n|) \quad (C > 0)$$
…(4.12)

Proof: We write

$$S = |\alpha_1| + |\alpha_2| + \cdots \dots + |\alpha_n|.$$

If $S = 0$, all α_i are zero, so that (4.12) holds for any C. Let $S > 0$, then (4.12) is equivalent to the inequality which we obtain from (4.12) by dividing by S and writing $\beta_j = \alpha_j / S$ that is

$$\|\beta_1 x_1 + \cdots \dots \dots + \beta_n x_n\| \geq C \quad \left(\sum_{j=1}^{n} |\beta_j| = 1\right) \qquad \dots(4.13)$$

Hence it is sufficient to prove the existence of a $c > 0$ such that (4.13) holds for every n-tuple of scalars $\beta_1, \ldots \ldots \ldots, \beta_n$ with

$\Sigma|\beta_j| = 1.$

Suppose that this is false. Then there exists a sequence $< y_m >$ of vectors

$$y_m = \beta_1^{(m)} x_1 + \cdots \ldots \ldots + \beta_n{}^{(m)} x_n \quad \left(\sum_{j=1}^{n} |\beta_j{}^{(m)}| = 1\right)$$

such that

$$\|y_m\| \to 0 \quad \text{as } m \to \infty$$

Since $\Sigma|\beta_j{}^{(m)}| = 1$, we have $|\beta_j{}^{(m)}| \leq 1$. Hence for each fixed j, the sequence

$$\left(\beta_j^{(m)}\right) = \left(\beta_j^{(1)}, \beta_j^{(2)}, \ldots \ldots.\right)$$

is bounded. Consequently, by the Bolzano - Weierstrass theorem, $(\beta_1{}^{(m)})$ has a convergent subsequence. Let β_1 denote the limit of that subsequence and let $< y_{1,m} >$ denote the corresponding subsequence of $\langle y_m \rangle$. By the same argument, $< y_{1,m} >$ has a subsequence $< y_{2,m} >$ for which the corresponding subsequence of scalars $\beta_2{}^{(m)}$ converges, let β_2 denote the limit-continuing in this way, after n steps we obtain a subsequence

$$\langle y_{n,m} \rangle = (y_{n,1}, y_{n,2}, \ldots \ldots \ldots) \quad \text{of } \langle y_m \rangle$$

whose terms are of the form

$$y_{n,m} = \sum_{j=1}^{n} \gamma_j^{(m)} x_j \quad \left(\sum_{j=1}^{n} |\gamma_j^{(m)}| = 1\right)$$

with scalars $\gamma_j{}^{(m)}$ satisfying $\gamma_j{}^{(m)} \to \beta_j$ as $m \to \infty$.

Hence as $m \to \infty$,

$$y_{n,m} \to y = \sum_{j=1}^{n} \beta_j x_j$$

where $\Sigma|\beta_j| = 1$ so that not all β_j can be zero. Since $\{x_1, \ldots \ldots \ldots, x_n\}$ is a linearly independent set, we thus have $y \neq 0$. On the other hand, $y_{n,m} \to y$ implies $\|y_{n,m}\| \to \| y \|$ by the continuity of the norm. Since $\|y_m\| \to 0$ by assumption and $< y_{n,m} >$ is a subsequence of $\langle y_m \rangle$, we must have $\| y_{n,m} \| \to 0$. Hence $\| y \| = 0$, so that $y = 0$. But this contradicts that $y \neq 0$, and the lemma is proved.

NOTES

Proof of the theorem

We consider an arbitrary Cauchy sequence $\langle y_m \rangle$ in Y and show that it is convergent in Y, the limit will be denoted by y. Let $\dim Y = n$ and $\{e_1, e_2, \dots\dots, e_n\}$ any basis for Y. Then each y_m has a unique representation of the form

$$y_m = \alpha_1{}^{(m)} e_1 + \cdots \dots \dots + \alpha_n{}^{(m)} e_n$$

Since $< y_m >$ is a Cauchy sequence, for every $\in > 0$, there is an N such that $\| y_m - y_n \| < \in$ when $m, r > N$. From this and the above Lemma, we have for some $C > 0$,

$$\in > \|y_m - y_r\| = \left\| \sum_{j=1}^{r} \left(\alpha_j^{(m)} - \alpha_j^{(r)} \right) e_j \right\|$$
$$\geq C \sum_{j=1}^{r} \left| \alpha_j^{(m)} - \alpha_j^{(r)} \right|$$

where $m, r > N$. Division by $C > 0$ gives

$$\left| \alpha_j^{(m)} - \alpha_j^{(r)} \right| \leq \sum_{j=1}^{n} \left| \alpha_j^{(m)} - \alpha_j^{(r)} \right| < \frac{\in}{C} \quad (m, r > N)$$

This shows that each of the n sequences

$$\left(\alpha_j^{(m)} \right) = \left(\alpha_j^{(1)}, \alpha_j^{(2)}, \dots\dots\dots \right) \quad j = 1, 2, \dots, n.$$

is Cauchy in $\mathbb{R}$ or $\mathbb{C}$. Hence it converges let α_j denote the limit. Using these n limits, $\alpha_1, \alpha_2, \dots, \alpha_n$, we define

$$y = \alpha_1 e_1 + \alpha_2 e_2 + \cdots \dots \dots \dots + \alpha_n e_n$$

Clearly $y \in Y$. Further

$$\|y_m - y\| = \left\| \sum_{j=1}^{n} \left(\alpha_j^{(m)} - \alpha_j \right) e_j \right\| \leq \sum_{j=1}^{n} | \alpha_j^{(m)} - \alpha_j \| \cdot \| e_j \|$$

On the right $\alpha_j^{(m)} \to \alpha_j$. Hence $\|y_m - y\| \to 0$, that is $y_m \to y$. This shows that $< y_m >$ is convergent in Y. Since $< y_m >$ was an arbitrary Cauchy sequence in Y, This proves that Y is complete.

Remark

From the above theorem and the result "A subspace M of a complete metric space X is complete if and only if the set M is closed in X", we get the following :

Theorem 4.19

Every finite dimensional subspace Y of a normed space X is closed in X.

NOTES

Remark

Infinite dimensional subspaces need not be closed e.g. Let $X = C[0, 1]$ and $Y = \text{span}\{x_0, x_1, \ldots\ldots\}$ where $x_j(t) = t^j$ so that Y is the set of polynomials. Y is not closed in X.

4.7 HAHN-BANACH THEOREM FOR NORMED LINEAR SPACES

Let X be a real or complex normed linear space, let $M \subseteq X$ be a linear subspace, and let $\ell \in M^*$ be a bounded linear functional on M. Then there exists a linear functional $\tilde{\ell} \in X^*$ that extends ℓ (i.e. $\tilde{\ell} \upharpoonright M = \ell$) and satisfies $\| \tilde{\ell} \|_{X^*} = \| \ell \|_{M^*}$.

The proof of the Hahn-Banach theorem has two parts: First, we show that *l* can be extended (without increasing its norm) from M to a subspace one dimension larger: that is, to any subspace $M_1 = \text{span}\{M, x_1\} = M + \mathbb{R}x_1$ spanned by M and a vector $x_1 \in X \setminus M$. Secondly, we show that these one-dimensional extensions can be combined to provide an extension from M to all of X.

Section 4.7.1 is the first step.

4.7.1 Hahn-Banach Theorem for Real Linear Space

Let X be a real normed linear space, let $M \subseteq X$ be a linear subspace, and let $\ell \in M^*$ be a bounded linear functional on M. Then, for any vector $x_1 \in X \setminus M$, there exists a linear functional ℓ_1 on $M_1 = \text{span}\{M, x_1\}$ that extends ℓ (i.e. $\ell_1 \lceil M = \ell$) and satisfies $\|\ell_1\|_{M_1^*} = \| \ell \|_{M^*}$.

Proof: If $\ell = 0$ the result is trivial, so we can assume without loss of generality that $\| \ell \| = 1$ (why?) (this assumption is made only to simplify the formulae). Now every $x \in M_1$ can be uniquely represented in the form $x = \lambda x_1 + y$ with $\lambda \in \mathbb{R}$ and $y \in M$. To define ℓ_1 as an extension of ℓ, it suffices to choose the value of $\ell_1(x_1)$, call it c_1 : we then have

$$\ell_1(\lambda x_1 + y) = \lambda c_1 + \ell(y) \qquad \ldots(4.14)$$

We want to choose c_1 so that $|\ell_1(x)| \leq \| x \|$ for all $x \in M_1$, i.e.

$$-\|\lambda x_1 + y\| \leq \lambda c_1 + \ell(y) \leq \|\lambda x_1 + y\| \qquad \ldots(4.15)$$

for all $\lambda \in \mathbb{R}$ and $y \in M$. This holds for $\lambda = 0$ by hypothesis on ℓ, and for $\lambda \neq 0$ it can be rewritten as

$$-\left\|x_1 + \frac{y}{\lambda}\right\| - \ell(y/\lambda) \leq c_1 \leq \left\|x_1 + \frac{y}{\lambda}\right\| - \ell(y/\lambda) \qquad \ldots(4.16)$$

for all $\lambda \in \mathbb{R}$ and $y \in M$ (you should check that this is correct both for $\lambda > 0$ and for $\lambda < 0$), or equivalently

$$-\|x_1 + z\| - \ell(z) \le c_1 \le \|x_1 + z\| - \ell(z) \quad \ldots(4.17)$$

for all $z \in M$. But for $z_1, z_2 \in M$ we have

$$\ell(z_2) - \ell(z_1) = \ell(z_2 - z_1) \le \|z_2 - z_1\| \le \|x_1 + z_1\| + \|x_1 + z_2\| \quad \ldots(4.18)$$

NOTES

by $\| \ell \| = 1$ and the triangle inequality, so that

$$-\|x_1 + z_1\| - \ell(z_1) \le \|x_1 + z_2\| - \ell(z_2) \quad \ldots(4.19)$$

for all $z_1, z_2 \in M$. It follows that

$$\begin{aligned} c_- &\equiv \sup_{z_1 \in M} [-\|x_1 + z_1\| - \ell(z_1)] \\ c_+ &\equiv \inf_{z_2 \in M} [\|x_1 + z_2\| - \ell(z_2)] \end{aligned} \quad \ldots(4.20)$$

are finite and satisfy $c_- \le c_+$;so we can choose any $c_1 \in [c_-, c_+]$.

Definition

Let S be a set. Then a partial order on S is a binary relation $\preccurlyeq$ on S that satisfies

1. $a \preccurlyeq a$ (reflexivity);
2. $a \preccurlyeq b$ and $b \preccurlyeq a$ imply $a = b$ (antisymmetry); and
3. $a \preccurlyeq b$ and $b \preccurlyeq c$ imply $a \preccurlyeq c$ (transitivity)

for all $a, b, c \in S$. The pair $(S, \preccurlyeq)$ is called a partially ordered set (or poset). We sometimes also refer to S alone as a partially ordered set if the relation $\preccurlyeq$ is understood from the context.

Now let $(S, \preccurlyeq)$ be a partially ordered set. A subset $T \subseteq S$ is called totally ordered (with respect to $\preccurlyeq$) if for every pair $a, b \in T$ we have either $a \preccurlyeq b$ or $b \preccurlyeq a$. A totally ordered subset is also called a chain. An element $u \in S$ is said to be an upper bound for a subset $T \subseteq S$ if $a \preccurlyeq u$ for all $a \in T$. (Note that the upper bound *u* need not belong to T itself.) Finally, a maximal element of S is an element $m \in S$ such that $m \preccurlyeq x$ implies $m = x$. (A maximal element need not exist; and if one exists, it need not be unique.)

Examples

1. The usual order $\le$ on $\mathbb{R}$ is a total order. There is no maximal element.
2. The usual order $\le$ on $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ is also a total order. Now there is a unique maximal element $+\infty$.
3. The usual partial order $\le$ on $\mathbb{R}^n$ is defined by $x \le y$ if and only if $x_i \le y_i$ for $1 \le i \le n$. For $n \ge 2$ it is not a total order. There is no maximal element.
4. Consider the usual partial order $\le$ on $\mathbb{R}^2$ restricted to the three-element subset $S = \{(0,0), (0,1), (1,0)\}$. Then $(0,1)$ and $(1,0)$ are maximal elements.

NOTES

5. The lexicographic order on $\mathbb{R}^2$ is defined by $x \preccurlyeq y$ if and only if either $x_1 < y_1$ or else $x_1 = y_1$ and $x_2 \le y_2$. (Think of the ordering of words in a dictionary) This is a total order (why?). There is no maximal element.
6. Let A be an arbitrary set, and let $\mathcal{P}(A)$ be the set of all subsets of A. Then the relation $\subseteq$ of set inclusion is a partial order on $\mathcal{P}(A)$. (It is not a total order except in two degenerate cases - can you see what they are?) There is a unique maximal element A.

Let V be a vector space, and let $\mathcal{L}(V)$ be the set of all linear subspaces of V. Then the relation $\subseteq$ of set inclusion is a partial order on $\mathcal{L}(V)$. (It is not a total order except in two degenerate cases - can you see what they are?) There is a unique maximal element V.

We then have Zorn's lemma.

Zorn's Lemma

Let $(S, \preccurlyeq)$ be a partially ordered set in which every totally ordered subset has an upper bound. Then $(S, \preccurlyeq)$ contains at least one maximal element.

Zorn's lemma is a result of set theory that can be proven using the axiom of choice. More precisely, Zorn's lemma is equivalent to the axiom of choice in Zermelo-Fraenkel (ZF) set theory. Other important statements of set theory that are equivalent to the axiom of choice in ZF set theory are the well-ordering theorem and the Hausdorff maximal principle. We shall not enter into the details of these statements or the proof of their equivalence, which belong to a course in set theory or mathematical logic; rather, we shall simply take Zorn's lemma as a set-theoretic result that we can use without worry

We are now ready to prove the Hahn-Banach theorem.

Proof of the Hahn-Banach for Real Linear Space

Let $\mathcal{E}$ denotes the set of all extensions of ℓ to linear subspaces of X (not necessarily to all of X) that satisfy the properties claimed in the Hahn-Banach theorem. More precisely, $\mathcal{E}$ consists of all pairs (N, f) such that

1. N is a linear subspace of X that contains M;
2. f is a bounded linear functional on N;
3. $f \setminus M = \ell$; and
4. $\| f \|_{N^*} = \| \ell \|_{M^*}$.

Now equip $\mathcal{E}$ with a partial order $\preccurlyeq$ by declaring that

$$(N, f) \preccurlyeq (N', f') \Leftrightarrow N \subseteq N' \text{ and } f' \lceil N = f \qquad \dots(4.21)$$

In other words, $(N, f) \preccurlyeq (N', f')$ iff f is an extension of f. (It is easy to check that $\preccurlyeq$ is indeed a partial order; you should do this.)

Now suppose that $\mathcal{F}$ is a totally ordered subset of $\mathcal{E}$. I claim that $\mathcal{F}$ has an upper bound in $\mathcal{E}$ (in fact a least upper bound, though we do not need this fact), defined as follows: First let

$$Y = \cup_{(N,f)\in\mathcal{F}} N \qquad \ldots(4.22)$$

You should verify, using the fact that $\mathcal{F}$ is totally ordered, that Y is a linear subspace of X; it is, in fact, the smallest linear subspace containing all the subspaces N where $(N,f) \in \mathcal{F}$. Next define on Y a linear functional g as the union of all the linear functionals f with $(N,f) \in \mathcal{F}$, i.e.

$$g(y) = f(y) \quad \text{whenever} \quad (N,f) \in \mathcal{F} \text{ with } y \in N. \qquad \ldots(4.23)$$

You should verify, using again the total ordering of $\mathcal{F}$, that g is well-defined in the sense that $f(y) = f'(y)$ whenever $(N,f) \in \mathcal{F}$ and $(N',f') \in \mathcal{F}$ with $y \in N$ and $y \in N'$; and you should verify, using once again the total ordering of $\mathcal{F}$, that g is indeed linear. Finally, you should check that $\| g \|_{Y^*} = \| \ell \|_{M^*}$. It follows that $(Y,g) \in \mathcal{E}$ and that $(N,f) \preccurlyeq (Y,g)$ for all $(N,f) \in \mathcal{F}$. Hence (Y,g) is an upper bound for $\mathcal{F}$ (in fact the least upper bound, though we do not need this fact).

So all the hypotheses of Zorn's lemma are satisfied. We can therefore conclude that $\mathcal{E}$ has a maximal element (N_*, f_*).

4.7.2 Hahn-Banach Theorem for Complex Linear Space

Let V be a normed linear space over C Let W be a subspace of V and let g: W → C be a continuous linear functional on W. Then there exists a continuous linear extension $f: V \to \mathbb{C}$ of g such that $\| f \|_{V^*} = \| g \|_{W^*}$.

Hahn-Banach Theorem for Normed Linear Space)

Let Y be a subspace of a normed linear space X and *f* be a continuous linear functional on Y; i.e., $f \in L(Y, \mathbb{R})$. Then there exists a continuous linear functional g on X, i.e., an element of $L(X, \mathbb{R})$, such that

1. g is an extension of f.
2. $\| g \| = \| f \|$.

Proof: If ρ is defined on X by $\rho(x) = \| f \| \| x \|$, then ρ is a seminorm on X. It is obvious that,

$$f(y) \leq |f(y)| \leq \| f \| \| y \| = \rho(y)$$

for all $y \in Y$. By the seminorm version of the Hahn-Banach Theorem, there exists a linear functional g on X, which is an extension of f, such that $g(x) \leq \rho(x) = \| f \| \| x \|$, for all $x \in X$, and this implies that g is continuous, and $\| g \| \leq \| f \|$. Clearly $\| g \| \geq \| f \|$ since g is an extension of f.

NOTES

4.8 WEAK CONVERGENCE

Definition

Let X be a normed linear space, and let $x_n, x \in X$.

1. We say that x_n converges strongly, or converges in norm to x, and write $x_n \to x$, if

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

2. We say that x_n converges weakly to x, and write $x_n \xrightarrow{w} x$, if

$$\forall \mu \in X^*, \quad \lim_{n\to\infty} \langle x_n, \mu\rangle = \langle x, \mu\rangle.$$

Definition

Let X be a normed linear space, and suppose that $\mu_n, \mu \in X^*$. Then we say that μ_n converges weak* $*$ to μ, and write $\mu_n \xrightarrow{w^*} \mu$, if

$$\forall x \in X, \quad \lim_{n\to\infty} \langle x, \mu_n\rangle = \langle x, \mu\rangle.$$

Note that weak* convergence is just "pointwise convergence" of the operators μ_n !

Remark

Weak* convergence only makes sense for a sequence that lies in a dual space X^*. However, if we have a sequence $\{\mu_n\}_{n\in\mathbf{N}}$ in X^*, then we can consider three types of convergence of μ_n to μ: strong, weak, and weak*. By definition, these are:

$$\mu_n \to \mu \quad \Leftrightarrow \lim_{n\to\infty} \|\mu - \mu_n\| = 0,$$

$$\mu_n \xrightarrow{w} \mu \quad \Leftrightarrow \quad \forall T \in X^{**}, \quad \lim_{n\to\infty} \langle \mu_n, T\rangle = \langle \mu, T\rangle,$$

$$\mu_n \xrightarrow{w^*} \mu \quad \Leftrightarrow \quad \forall x \in X, \quad \lim_{n\to\infty} \langle x, \mu_n\rangle = \langle x, \mu\rangle.$$

Lemma

1. Weak* limits are unique.
2. Weak limits are unique.

Proof:

1. Suppose that X is a normed linear space, and that we had both $\mu_n \xrightarrow{w^*} \mu$ and $\mu_n \xrightarrow{w^*} \nu$ in X^*. Then, by definition,

$$\forall x \in X, \quad \langle x, \mu\rangle = \lim_{n\to\infty} \langle x, \mu_n\rangle = \langle x, \nu\rangle,$$

so $\mu = \nu$.

2. Suppose that we have both $x_n \xrightarrow{w} x$ and $x_n \xrightarrow{w} y$ in X. Then, by definition,

   $$\forall \mu \in X^*, \quad \langle x, \mu \rangle = \lim_{n \to \infty} \langle x_n, \mu \rangle = \langle y, \mu \rangle.$$

   Hence, by Hahn-Banach,

   $$\| x - y \| = \sup_{\|\mu\|=1} |\langle x - y, \mu \rangle| = 0$$

   so $x = y$.

Lemma

If X is a finite-dimensional vector space, then strong convergence is equivalent to weak convergence.

Check Your Progress

5. What are vector space and subspaces?
6. What is linear transformation in vector spaces?
7. State linear span and finite dimensional vector space.
8. Define reflexive space.

4.9 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A vector space with a norm is called a normed vector space.
2. A topological space is said to be topologically regular if every point has an open neighbourhood that is regular.
3. On an n-dimensional Euclidean space $\mathbf{R}^n$, the perceptive notion of length of the vector $x = (x_1, x_2, ..., x_n)$ is illustrated by the formula,

   $$\|x\| := \sqrt{x_1^2 + \cdots + x_n^2}.$$

   The Euclidean norm is the most commonly used norm on $\mathbf{R}^n$.
4. A sequence (x_n) in a normed space X is said to be weakly convergent if there is an $x \in X$ such that for every $f \in X$.
5. 5. Let $< V, + >$ be an abelian group and $< F, +, \cdot >$ be a field. Define a function $\times$ (called scalar multiplication) from $F \times V \to V$, such that, for all $\alpha \in F$, $v \in V$, $\alpha \cdot v \in V$.
6. Let V and U be two vector spaces over the same field F, then a mapping $T : V \to U$ is called a homomorphism or a linear transformation if

   $$T(x + y) = T(x) + T(y) \quad \text{for all } x, y \in V$$
   $$T(\alpha x) = \alpha T(x) \quad \alpha \in F$$
7. Let $V(F)$ be a vector space, $v_i \in V$, $\alpha_i \in F$ be elements of V and F respectively. Then elements of the type $\sum_{i=1}^{n} \alpha_i v_i$ are called *linear combinations* of $v_1, v_2, ..., v_n$ over F.

NOTES

Let S be a non-empty subset of V, then the set

$$L(S) = \left\{ \sum_{i=1}^{n} \alpha_i v_i \mid \alpha_i \in F, v_i \in S, n \text{ finite} \right\}$$

i.e., the set of all linear combinations of finite sets of elements of S is called *linear span* of S. It is also denoted by $< S >$. If $S = \varphi$, define $L(S) = \{0\}$.

8. Let X be a normed space and $X^{**} = (X^*)^*$ denote the second dual space of X. The canonical map defined by gives an isometric linear isomorphism (embedding) from X into X^{**}. This space is called reflexive if this map is surjective.

4.10 SUMMARY

- A norm is a function that assigns a strictly positive length or size to all vectors in a vector space other than the zero vector while a seminorm is allowed to assign zero length to some non-zero vectors.
- A topological vector space is called normable (seminormable) if the topology of the space can be induced by a norm (seminorm). The norm of a vector **v** is usually denoted $\|\mathbf{v}\|$ and sometimes $|\mathbf{v}|$.
- On an n-dimensional Euclidean space $\mathbf{R}^n$, the perceptive notion of length of the vector $x = (x_1, x_2, ..., x_n)$ is illustrated by the formula,

  $$\|x\| := \sqrt{x_1^2 + \cdots + x_n^2}.$$

  The Euclidean norm is the most commonly used norm on R^n.
- The Euclidean norm of a complex number is the absolute value (also called the modulus) of it, if the complex plane is identified with the Euclidean plane R^2.
- A sequence (x_n) in a normed space X is said to be strongly convergent if there is an $x \in X$ such that,

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

- A sequence (x_n) in a normed space X is said to be weakly convergent if there is an $x \in X$ such that for every $f \in X'$,

  $$\lim_{n\to\infty} f(x_n) = f(x)$$

- Let $< V, + >$ be an abelian group and $< F, +, \cdot >$ be a field. Define a function $\times$ (called scalar multiplication) from $F \times V \to V$, such that, for all $\alpha \in F$, $v \in V$, $\alpha \cdot v \in V$.
- *A* non-empty subset W of a vector space $V(F)$ is said to form a *subspace* of V if W forms a vector space under the operations of V.

- A necessary and sufficient condition for a non-empty subset W of a vector space $V(F)$ to be a subspace is that W is closed under addition and scalar multiplication.
- Let V and U be two vector spaces over the same field F, then a mapping $T : V \to U$ is called a homomorphism or a linear transformation if

$$T(x + y) = T(x) + T(y) \quad \text{for all } x, y \in V$$
$$T(\alpha x) = \alpha T(x) \quad \alpha \in F$$

- Let $V(F)$ be a vector space, $v_i \in V$, $\alpha_i \in F$ be elements of V and F respectively. Then elements of the type $\sum_{i=1}^{n} \alpha_i v_i$ are called *linear combinations* of $v_1, v_2, ..., v_n$ over F.

4.11 KEY TERMS

- **Euclidean norm:** On an n-dimensional Euclidean space R^n, the perceptive notion of length of the vector $x = (x_1, x_2, ..., x_n)$ is illustrated by the formula, $\|x\| := \sqrt{x_1^2 + \cdots + x_n^2}$.
  The Euclidean norm is the most commonly used norm on $\mathbf{R}^n$.
- **Taxicab norm:** $\| x \|_1 := \sum_{i=1}^{n} | x_i |$. The name relates to the distance a taxi has to drive in a rectangular street grid to get from the origin to the point x.
- **Subspaces:** *A* non-empty subset W of a vector space $V(F)$ is said to form a *subspace* of V if W forms a vector space under the operations of V.

4.12 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define strong convergence.
2. What is a perfectly normal space?
3. State the importance of Urysohn's lemma.
4. Define adjoint operators.
5. What is reflexive space?

Long-Answer Questions

1. Illustrate the concept of equivalent norms.
2. Describe weak and strong convergence with the help of examples.
3. Explain the concept of regular and normal spaces.
4. State and prove Urysohn's lemma.
5. Prove the characterization of the adjoint operators.
6. Explain Hahn Banach theorem for real linear space, complex linear space and normed linear space.

4.13 FURTHER READING

Yang, Bicheng and Qiang Chen. 2016. *Integration Theory and Functional Analysis*. New York: Magnum Publishing LLC.

Kadets, Vladimir. 2018. *A Course in Functional Analysis and Measure Theory*. New York: Springer International Publishing.

Pundir, R. and S. K. Pundir. 2015. *Integration Theory and Functional Analysis*. Meerut (Uttar Pradesh): Pragati Prakashan.

Rudin, Walter. 1986. *Real and Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Ahlfors, Lars V. 1978. *Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Lang, Serge. 1998. *Complex Analysis*, 4th Edition. New York: Springer-Verlag New York Inc.

Shastri, Anant R. 2000. *An Introduction to Complex Analysis*. India: Macmillan Publishers India Ltd.

Gupta, S. L. and Nisha Rani. 2003. *Fundamental Real Analysis*, 4th Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, 1st Edition. United Kingdom: Cambridge University Press.

Basener, William F. *Topology and Its Applications*. New Jersey: John Wiley & Sons, Inc.

Bourbaki, Nicolas. 1966. *Elements of Mathematics: General Topology*, Part 1. United States: Addison-Wesley.

UNIT 5 INNER PRODUCT SPACE AND HILBERT SPACE

NOTES

Structure

5.0 INTRODUCTION

In mathematics, an inner product space is a type of space in mathematics. A field of mathematics known as functional analysis, is a vector space with an additional structure called an inner product. This additional structure associates each pair of vectors in the space with a scalar quantity known as the inner product of the vectors. An inner product naturally induces an associated norm, thus an inner product space is also a normed vector space. A complete space with an inner product is called a Hilbert space. An incomplete space with an inner product is called a pre-Hilbert space, since its completion with respect to the norm, induced by the inner product, becomes a Hilbert space. Inner product spaces over the field of complex numbers are sometimes referred to as unitary spaces.

The mathematical concept of a Hilbert space, named after David Hilbert, generalizes the notion of Euclidean space. It extends the methods of vector algebra and calculus from the two-dimensional Euclidean plane and three-dimensional space to spaces with any finite or infinite number of dimensions. Hilbert spaces have inner products and so notions of orthogonality and orthogonal projection are available. Many of the applications of Hilbert spaces exploit the fact that Hilbert spaces support generalizations of simple geometric concepts like projection from their usual finite dimensional setting. In particular, the spectral theory of continuous self-adjoint linear operators on a Hilbert space generalizes the usual spectral decomposition of a matrix, and this often plays a major role in applications of the theory to other areas of mathematics and physics.

In this unit, you will learn about the inner product space, Hilbert spaces, orthogonal complements, conjugate space H*, self-adjoint operators on Hilbert spaces, projections on Hilbert spaces and positive, normal and unitary operators.

NOTES

5.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe inner product spaces and Schwarz's inequality
- Discuss Hilbert spaces and convex set in Hilbert spaces
- Explain the self-adjoint operators on Hilbert spaces
- Define orthogonal sets
- Describe the conjugate space
- Discuss reflexivity and projections on Hilbert spaces
- State normal and unitary operators

5.2 INNER PRODUCT SPACE

In general a vector space is defined over an arbitrary field F. In this section we restrict F to the field of real or complex numbers. In the first case, the vector space is called real vector space and in the second case it is called a complex vector space. We have dot or scalar product of two vectors which among other things satisfies the following:

(*i*) $\vec{v}\,.\,\vec{v} \geq 0$ and $(\vec{v}\,.\,\vec{v}) = 0 \Leftrightarrow \vec{v} = 0$

(*ii*) $\vec{v}\,.\,\vec{w} = \vec{w}\,.\,\vec{v}$

(*iii*) $\vec{u}\,.\,(\alpha\vec{v} + \beta\vec{w}) = \alpha(\vec{u}\,.\,\vec{v}) + \beta(\vec{u}\,.\,\vec{w})$

where $\vec{u}, \vec{v}, \vec{w}$ are vectors and α, β real numbers.

We wish to extend the concept of dot product to complex vector spaces also. We define a map on $V \times V$ to F (where V = vector space over F) with same property as dot product, called inner product and study the concept of length and orthogonality.

Definition: Let V be a vector space over field F (where F = field of real or complex numbers). Suppose for any two vectors $u, v \in V \; \exists$ an element $(u, v) \in F$ such that, [(u, v) here is just an element of F and should not be confused with the ordered pair.]

(*i*) $(u, v) = \overline{(v, u)}$ (*i.e.*, complex conjugate of (v, u))

(*ii*) $(u, u) \geq 0$ and $(u, u) = 0 \Leftrightarrow u = 0$

(*iii*) $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$

for any $u, v, w \in V$ and $\alpha, \beta \in F$.

Then V is called an *inner product space* and the function satisfying (*i*), (*ii*) and (*iii*) is called an *inner product*.

Thus inner product space is a vector space over the field of real or complex numbers with an inner product function.

Notes:

NOTES

1. Property (*ii*) in the definition of inner product space makes sense in as much as $(u, u) = \overline{(u, u)}$ by (*i*) $\Rightarrow (u, u) =$ real.
2. Property (*iii*) can also be described by saying that inner product is a linear map in 1st variable.
3. Can we say that inner product is linear in 2nd variable?

   Let's evaluate

   $$(u, \alpha v + \beta w) = \overline{(\alpha v + \beta w, u)} \text{ by } (i)$$
   $$= \bar{\alpha}\,\overline{(v, u)} + \bar{\beta}\,\overline{(w, u)}$$
   $$= \bar{\alpha}\,(u, v) + \bar{\beta}\,(u, w)$$

   So, it need not be linear in 2nd variable.
4. If F = field of real numbers, then the function inner product satisfies same properties as dot product seen earlier.
5. Inner product space over real field is called *Euclidean space* and over complex field is called *Unitary space*.
6. In the vector space of all vectors in 3-dimensional space over reals, the inner product will be the usual dot product of two vectors, *i.e.*,

   $$<\vec{u}, \vec{v}> = |\vec{u}||\vec{v}| \cos\theta.$$

Example 5.1: Let $V = F^{(n)}$, F = field of complex numbers.

Solution: $u = (\alpha_1, \alpha_2, ..., \alpha_n)$
$v = (\beta_1, \beta_2, ..., \beta_n)$ in $F^{(n)}$

Define $(u, v) = \alpha_1 \overline{\beta_1} + ... + \alpha_n \overline{\beta_n}$

It can be easily shown that (u, v) defines an inner product, called standard inner product.

Example 5.2: Let $V = \mathbf{R}^{(2)}$, $u = (\alpha_1, \alpha_2)$, $v = (\beta_1, \beta_2)$

Define $(u, v) = \alpha_1\beta_1 - \alpha_2\beta_1 - \alpha_1\beta_2 + 4\,\alpha_2\beta_2$

Then

Solution: (*i*) $(u, v) = (v, u) = \overline{(v, u)}$

(*ii*) $(u, u) = (\alpha_1 - \alpha_2)^2 + 3\alpha_2^2 \geq 0$

$(u, u) = 0 \Leftrightarrow \alpha_1 = \alpha_2, \alpha_2 = 0$
$\Leftrightarrow \alpha_1 = 0 = \alpha_2$
$\Leftrightarrow u = (\alpha_1, \alpha_2) = (0, 0) = 0$

(*iii*) $(\alpha u + \beta u, w) = \alpha(u, w) + \beta(v, w)$
can be easily verified.

Thus (u, v) defines an inner product.

Example 5.3: One may construct a new inner product from a given one. Let V, W be vector spaces over F and T, a one–one linear transformation from V into W.

Suppose (,) is an inner product on W. Then,

Solution: $< u, v > = (T(u), T(v))$

defines an inner product on V as

NOTES

$$(i) \quad \overline{< v, u >} = \overline{(T(v), T(u))}$$
$$= (T(u), T(v))$$
$$= < u, v >$$

$$(ii) \quad < u, v > = (T(u), T(u)) \geq 0$$

and $< u, u > = 0 \Leftrightarrow (T(u), T(u)) = 0$

$$\Leftrightarrow T(u) = 0 \Leftrightarrow u = 0 \text{ as } T \text{ is } 1\text{–}1$$

$$(iii) < \alpha u + \beta v, w > = (T(\alpha u + \beta v), T(w))$$
$$= (\alpha T(u) + \beta T(v), T(w))$$
$$= \alpha(T(u), T(w)) + \beta(T(v), T(w))$$
$$= \alpha < u, w > + \beta < v, w >$$

Example 5.4: Let $V = M_{m\times n}$ (**C**). Then $< A, B >$ = Trace (AB^*) where $B^* = \overline{B'}$, defines an inner product on V as

(i) $\quad \overline{< B, A >} = \overline{\text{Trace } BA^*}$

Let $\quad A = (a_{ij}), B = (b_{ij}), AB^* = C = (c_{ij})$

Solution (i): $\quad B^* = (d_{ij})$, where $d_{ij} = \bar{b}_{ji}$

$$\therefore \; c_{ik} = \Sigma a_{ij} d_{jk} = \Sigma a_{ij} \bar{b}_{kj}$$

$$\Rightarrow c_{ii} = \Sigma a_{ij} \bar{b}_{kj}$$

$$\Rightarrow \text{Trace } AB^* = \Sigma c_{ii} = \Sigma(\Sigma a_{ij} \bar{b}_{ij})$$

Let $\quad A^* = (e_{ij})$, where $e_{ij} = \bar{a}_{ji}$

Let $\quad BA^* = F = (f_{ij})$, then

$$f_{ik} = \Sigma b_{ij} \bar{a}_{kj}$$

$$\Rightarrow \text{Trace } BA^* = \Sigma f_{ii} = \Sigma(\Sigma b_{ij} \bar{a}_{ij})$$

$$\Rightarrow \overline{\text{Trace } BA^*} = \Sigma\Sigma a_{ij} \bar{b}_{ij} = \text{Trace } AB^*$$

$$\Rightarrow \quad \overline{< B, A >} = < A, B >$$

(ii) $\quad < A, B > = \text{Trace } AB^* = \Sigma(\Sigma a_{ij} \bar{b}_{ij})$

Solution (ii): $\quad \therefore \; < A, A > = \sum\sum a_{ij}\bar{a}_{ij} = \sum\sum | a_{ij} |^2 \geq 0$

and $\quad < A, A > = 0 \Leftrightarrow | a_{ij} | = 0 \quad \forall \; i, j$

$$\Leftrightarrow a_{ij} = 0 \quad \forall \; i, j$$
$$\Leftrightarrow A = 0$$

Similarly axiom (iii) can be verified.

NOTES

Example 5.5: Let V be an inner product space. Show that

(*i*) $(0, v) = 0$ for all $v \in V$

(*ii*) $(u, v) = 0$ for all $v \in V \Rightarrow u = 0$

Solution: (*i*) $(0, v) = (0, 0, v)$

$= 0\,(0, v) = 0$

(*ii*) $(u, v) = 0$ for all $v \in V$

$\Rightarrow (u, u) = 0 \Rightarrow u = 0.$

Example 5.6: Let W_1, W_2 be two subspaces of a vector space V. If W_1, W_2 are inner product spaces, show that $W_1 + W_2$ is also an inner product space.

Solution: Let $x, y \in W_1 + W_2$.

Then $x = u_1 + u_2$

$y = v_1 + v_2 \quad u_1, v_1 \in W_1;\ u_2, v_2 \in W_2$

Define $< x, y > = (u_1, v_1) + (u_2, v_2)$

Then

(*i*) $\overline{< y, x >} = \overline{(v_1, u_1) + (v_2, u_2)}$

$= \overline{(v_1, u_1)} + \overline{(v_2, u_2)}$

$= (u_1, v_1) + (u_2, v_2)$

$= < x, y >$

(*ii*) $< x, x > = (u_1, u_1) + (u_2, u_2) \geq 0$

and $< x, x > = 0 \Leftrightarrow (u_1, u_1) = 0 = (u_2, u_2)$

$\Leftrightarrow u_1 = 0 = u_2$

$\Leftrightarrow x = 0$

(*iii*) $< \alpha x + \beta y, z > = \alpha < x, z > + \beta < y, z >$

can be easily verified.

$\therefore < x, y >$ defines an inner product on $W_1 + W_2$

So, $W_1 + W_2$ is an inner product space.

Norm of a Vector

Let V be an inner product space. Let $v \in V$. Then *norm* of v (or length of v) is defined as $\sqrt{(v, v)}$ and is denoted by $\| v \|$.

In the vector space of all vectors in 3-dimensional space,

$$\| \vec{u} \| = \sqrt{< \vec{u}, \vec{u} >} = | \vec{u} | = \text{length of } \vec{u}\,.$$

For this reason, norm of vector in general is also called length of vector.

Example 5.7: $\| \alpha v \| = | \alpha |\ \| v \|$ for all $\alpha \in F$, $v \in V$

Solution: $\| \alpha v \|^2 = (\alpha v, \alpha v)$

$= \alpha \bar{\alpha}\ (v, v)$

$= | \alpha |^2\ \| v \|^2$

NOTES

$$\Rightarrow \| \alpha v \| = |\alpha| \, \| v \|$$

We now prove an important inequality known as *Cauchy-Schwarz inequality*.

Theorem 5.1: Let V be an inner product space.

Then $\quad |(u, v)| \le \| u \| \, \| v \|$ for all $u, v \in V$.

Proof: If $u = 0$, then $(u, v) = (0, v) = 0$

and $\quad \| u \| = \sqrt{(u,u)} = \sqrt{(0,0)} = 0$

$\therefore \quad$ L.H.S. = R.H.S.

Let $u \ne 0$. Then $\| u \| \ne 0$

(as $\| u \| = 0 \quad \Rightarrow \sqrt{(0,0)} = 0$

$\Rightarrow (u, u) = 0 \Rightarrow u = 0$)

Let $\quad w = v - \dfrac{(v, u)}{\| u \|^2} u$

Then
$$(w, w) = \left(v - \frac{(v, u)}{\| u \|^2} u, v - \frac{(v, u)}{\| u \|^2} u \right)$$
$$= (v, v) - \frac{(v, u)}{\| u \|^2} (u, v)$$
$$= \| v \|^2 - \frac{\overline{(u, v)}\,(u, v)}{\| u \|^2} = \| v \|^2 - \frac{|(u, v)|^2}{\| u \|^2}$$
$$= \frac{\| u \|^2 \| v \|^2 - |(u, v)|^2}{\| u \|^2}$$

Since $\quad (w, w) \ge 0$,

$$|(u, v)|^2 \le \| u \|^2 \, \| v \|^2$$
$$\Rightarrow |(u, v)| \le \| u \| \, \| v \|.$$

Notes:

(*i*) The above inequality will be an equality if and only if u, v are linearly dependent.

Proof: Suppose $|(u, v)| = \| u \| \, \| v \|$

If $u = 0$, then $u = 0.v \Rightarrow u, v$ are linearly dependent.

Let $u \ne 0$. Then from above

$$(w, w) = 0 \Rightarrow w = 0$$

$\therefore \quad v - \dfrac{(v, u)}{\| u \|^2} u = 0$

$\Rightarrow v = v - \dfrac{(v, u)}{\| u \|^2} u \Rightarrow u, v$ are linearly dependent.

Conversely, let $u = \alpha v$, $\alpha \in F$

Then $\quad |(u, v)| = |\alpha (v, v)| = |\alpha| \, \| v \|^2$

$$\| u \| \, \| v \| = |\alpha| \, \| v \| \, \| v \| = |\alpha| \, \| v \|^2$$
$$|(u, v)| = \| u \| \, \| v \|.$$

(*ii*) In the vector space of all vectors in 3-dimensional space, since

$$|<\vec{u}, \vec{v}>| = |\vec{u}|\ |\vec{v}|\ |\cos\theta|$$

$$\leq \|\vec{u}\|\|\vec{v}\| \quad \text{as } |\cos\theta| \leq 1$$

we find that Cauchy-Schwarz inequality holds.

Theorem 5.2: Let V be an inner product space.

Then (*i*) $\| x + y \| \leq \| x \| + \| y \|$ for all $x, y \in V$

(Triangle inequality)

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

(Parallelogram Law)

Proof: (*i*) $\| x + y \|^2 = (x + y, x + y)$

$$= (x, x) + (y, x) + (x, y) + (y, y)$$

$$= \| x \|^2 + \overline{(x, y)} + (x, y) + \| y \|^2$$

$$= \| x \|^2 + 2Re\ (x, y) + \| y \|^2$$

$$\leq \| x \|^2 + 2\ |(x, y)| + \| y \|^2$$

$$\leq \| x \|^2 + 2\ \| x \|\ \| y \| + \| y \|^2$$

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

Hence, $\| x + y \| \leq \| x \| + \| y \|$

This is called triangle inequality as

$\| x \| + \| y \|$ = sum of the lengths of two sides of a triangle

$\| x + y \|$ = length of the third side of the triangle showing that sum of two side of a triangle is less than its third side.

(*ii*) $\| x + y \|^2 + \| x - y \|^2$

$$= (x + y, x + y) + (x - y, x - y)$$

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

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

Note: $\| x + y \|^2 + \| x - y \|^2$ = sum of squares of lengths of diagonals of a parallelogram

$2 (\| x \|^2 + \| y \|^2)$ = sum of squares of sides of a parallelogram.

$\therefore$ sum of squares of lengths of diagonals of a parallelogram is equal to sum of squares of lengths of its sides. For this reason (*ii*) is called parallelogram law.

Example 5.8: Using Cauchy-Schwarz inequality, prove that cosine of an angle is of absolute value at most 1.

Solution: Let F = Field of real numbers and $V = F^{(3)}$

Consider standard inner product on V.

Let $u = (x_1, y_1, z_1)$, $v = (x_2, y_2, z_2) \in V$

Let $O = (0, 0, 0)$

Let θ be an angle between OU and OV.

Then $\cos\theta = \dfrac{x_1x_2 + y_1y_2 + z_1z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2}\,\sqrt{x_2^2 + y_2^2 + z_2^2}} = \dfrac{(u, v)}{\|u\|\,\|v\|}$

$\therefore \quad |\cos\theta| = \dfrac{|(u, v)|}{\|u\|\,\|v\|} \le \dfrac{\|u\|\,\|v\|}{\|u\|\,\|v\|} = 1$

5.3 HILBERT SPACE

Theorem 5.3: Suppose E is an inner product space and M a complete convex subset of E. Let $x \in E$, then the following will be equivalent:

1. $y \in M$ satisfies $\|x - y\| = (\min_{z\in M})\ \|x - z\|$.
2. $y \in M$ satisfies $\mathrm{Re}(y - x, y - z) \le 0 \ \forall z \in M$.

Additionally, there is unique $y \in M$ satisfying (1) and (2).

Proof: (1) $\Rightarrow$ (2): For $z \in M$ and $0 < \theta \le 1$, let

$$f(\theta) = \|x - \{(1-\theta)\,y + \theta z\}\|^2 = \|\,x - y + \theta\,(y - z)\|^2$$
$$= \|x - y\|^2 + \theta^2\,\|y - z\|^2 + 2\theta\,\mathrm{Re}(x - y, y - z).$$

Since $f(\theta) = f(0) = \|\,x - y\|^2$ for $0 < \theta \le 1$, we have

$$\lim_{\theta\downarrow 0}\frac{f(\theta) - f(0)}{\theta} = 2\,\mathrm{Re}(x - y, y - z) \ge 0$$

(2) $\Rightarrow$ (1): For $z \in M$ we have,

$$\mathrm{Re}\,(y - x, y - z) = -\mathrm{Re}(x - y, y - x + x - z)$$
$$= \|x - y\|^2 - \mathrm{Re}(x - y, x - z) \le 0$$

Hence

$$\|x - y\|^2 \le \mathrm{Re}(x - y, x - z) \le \|x - y\| \cdot \|x - z\|$$

and therefore

$$\|x - y\| \le \|x - z\|$$

for all $z \in M$.

It follows from (2) that there exists at most one such y, as if both y_1 and y_2 satisfy (2) for all $z \in M$ then,

$$0 \le \|y_1 - y_2\|^2 = (y_1 - y_2, y_1 - y_2) = (y_1 - x, y_1 - y_2) + (y_2 - x, y_2 - y_1)$$
$$= \mathrm{Re}(y_1 - x, y_1 - y_2) + \mathrm{Re}(y_2 - x, y_2 - y_1) \le 0,$$

so $y_1 = y_2$.

To prove that there exists y satisfying (1), consider

$$\alpha = \inf_{z\in M}\|\,x - z\,\|$$

Now consider a sequence $\{z_n\} \subset M$ that satisfies,

$$\alpha^2 \le \| x - z_n \|^2 \le \alpha^2 + \frac{1}{n}$$

Claiming that $\{z_n\}$ is a Cauchy sequence, we have

NOTES

$$\|z_n - z_m\|^2 = \|(z_n - x) - (z_m - x)\|^2$$
$$= \|(z_n - x)\|^2 + (z_m - x)\|^2 - 2 \text{ Re } (z_n - x, z_m - x)$$
$$4\left\|\frac{z_n + z_m}{2} - x\right\|^2 = \|(z_n - x)\|^2 + (z_m - x)\|^2 - 2 \text{ Re } (z_n - x, z_m - x)$$

Consequently

$$\|(z_n - z_m)\|^2 = 2\|z_n - x)\|^2 + 2\ \|(z_m - x)\|^2 - 4\ \left\|\frac{z_n + z_m}{2} - x\right\|^2$$

$$\le\ 2\left(\alpha^2 + \frac{1}{n}\right) + 2\left(\alpha^2 + \frac{1}{m}\right) - 4\alpha^2 = 2\left(\frac{1}{n} + \frac{1}{m}\right)$$

which shows that $\{z_n\}$ is a Cauchy sequence. Now as M is complete, there exists $y \in M$ with $y = \lim_{n\to\infty} z_n$. Noticeably, $\|x - y\| = \lim_{n\to\infty} \|x - z_n\| = \alpha$. This completes the proof.

The map $t : E \mapsto M$ defined by $tx = y$, where y is the unique element in M and satisfies (1) and (2) of Theorem 5.3 is called the projection from E onto M.

Corollary 1: Suppose M is a closed convex subset of a Hilbert space E then $t = t_M$ has the following properties:

1. $t^2 = t$, i.e., t is idempotent
2. $\|tx - ty\| \le \|x - y\|$, i.e., t is contractive
3. $\text{Re}(tx - ty, x - y) \ge 0$, i.e., t is monotone

Proof: (1) is evident.

(2) From $\text{Re}(tx - x, tx - ty) \le 0$ and $\text{Re }(ty - y, ty - tx) \le 0$ we get $\text{Re}(x - y - (tx - ty), tx - ty) \ge 0$.

Hence $\|tx - ty\|^2 \le \text{Re}(x - y, tx - ty) \le \|x - y\| \,.\, \|tx - ty\|$ from which $\|tx - ty\| \le \|x - y\|$ follows.

(3) Again from $\text{Re}(x - y - (tx - ty), tx - ty) \ge 0$ we get

$0 \le \|tx - ty\|^2 \le \text{Re}(x - y, tx - ty)$

Linear Transformation

Consider a linear transformation T from a normed vector space X into a normed vector space Y over the same field $\mathbb{R}$ or $\mathbb{C}$. T is continuous on X if and only if it is continuous at one point.

Theorem 5.4: T is continuous if and only if there is $C \geq 0$ such that,

$$\|Tx\| \leq C\|x\| \qquad ..(5.1)$$

for all $x \in X$.

NOTES

Proof: $\Rightarrow$ If there is $C \geq 0$ such that Equation (1) holds for all $x \in X$ then T is clearly continuous at $x = 0$ and hence it is continuous on X.

$\Leftarrow$ Conversely, let T is continuous on X and is hence continuous at $x = 0$. Then there exists $\delta > 0$, such that if $\|x\| \leq \delta$, then $\|Tx\| \leq 1$. Let now $x \in X$ and $x \neq 0$. This implies $\left\|\frac{\delta}{\|x\|}x\right\| = \delta$ and so $\left\|T\left(\frac{\delta}{\|x\|}x\right)\right\| \leq 1$. Thus $\|Tx\| \leq \frac{1}{\delta}\|x\|$.

If we choose $C = \frac{1}{\delta}$ then Equation (5.1) holds for $x \neq 0$. But when $x = 0$, Equation 5.1 holds always. This completes the proof.

From this theorem we get that if T is a continuous linear transformation from X into Y, then

$$\|T\| :=_{x \in X} \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} < +\infty \qquad ...(5.2)$$

and is the smallest C for which Equation (5.2) holds. $\|T\|$ is called the norm of T. $\|T\|$ can be defined for any linear transformation T from X into Y and T is continuous iff $\|T\| < +\infty$. Hence a continuous linear transformation is also known as a bounded linear transformation.

Theorem 5.5: $\|T\| = \sup_{x \in X, \|x\|=1} \|Tx\|$.

Theorem 5.6: Consider $L(X, Y)$ to be the space of all bounded linear transformations from X into Y. Then it is a normed vector space with norm given by Theorem 5.5.

Theorem 5.7: If Y is a Banach space then so is $L(X, Y)$.

Proof: Here we will show that $L(X, Y)$ is complete. Let $\{T_n\}$ be a Cauchy sequence in $L(X, Y)$. Now, as

$$\|T_n x - T_m x\| = \|T_n - T_m)x\| \leq \|T_n - T_m\| \, . \, \|x\|$$

$\{T_n x\}$ is a Cauchy sequence in Y for each $x \in X$. Set $Tx = \lim_{n\to\infty} T_n x$. T is evidently a linear transformation from X into Y. Claim now that $T \in L(X, Y)$. Since $\{T_n\}$ is Cauchy, $\|T_n\| \leq C$ for some $C > 0$ and for all n. Now,

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leq \liminf_{n\to\infty} \|T_n\| . \|x\|$$

$$\leq \left(\sup_n \|T_n\|\right) \|x\| \leq C\|x\|$$

$\Leftarrow$ $x \in X$. Therefore T is a bounded linear transformation. Now we will show, $\lim_{n\to\infty} \|T_n - T\| = 0$. Given $\varepsilon > 0$ there exists n_0 such that $\|T_n - T_m\| < \varepsilon$ if $n, m \geq n_0$. Let $n \geq n_0$. Then we have,

NOTES

$$\| T_n - T \| = \sup_{x \in X, \|x\|=1} \| T_n x - Tx \|$$

$$\sup_{x \in X, \|x\|=1} \lim_{m \to \infty} \| T_n x - T_m x \|$$

$$\leq \sup_{x \in X, \|x\|=1} \liminf_{m \to \infty} \| T_n - T_m \| . \| x \|$$

$$\leq \sup_{x \in X, \|x\|=1} \varepsilon \| x \| = \varepsilon,$$

This shows that $\lim_{n \to \infty} \| T_n - T \| = 0$ or $\lim_{n \to \infty} T_n = T$. Thus the sequence $\{T_n\}$ has a limit in $L(X, Y)$. This completes the proof.

$L(X, \mathbb{C})$ or $L(X, \mathbb{R})$ depending on whether X is a complex or a real vector space, is known as the topological dual of X and is denoted by X'. X' is a Banach space.

Theorem 5.8 (Riesz Representation Theorem): Let X be a Hilbert space and $\ell \in X'$, then there is $y_0 \in X$ such that,

$$\ell(x) = (x, y_0) \text{ for } x \in X$$

Furthermore, the mapping $\ell \mapsto y_0$ is conjugate linear and $\| \ell \| = \| y_0 \|$.

Proof: Let $\ell \neq 0$ and let $M = \ker \ell$. Then $M^{\perp}$ is one-dimensional. For $x \in X$, x can be uniquely expressed as $x = \upsilon + \lambda x_0$, where x_0 is a fixed nonzero element of $M^{\perp}$, $\upsilon \in M$, and λ is a scalar. We have then,

$$\ell(x) = \ell(\upsilon) + \lambda \ell(x_0) = \lambda \ell(x_0)$$

and

$$(x, x_0) = (\upsilon + \lambda x_0, x_0) = \lambda \| x_0 \|^2$$

Hence if we let $y_0 = \dfrac{\overline{\ell(x_0)}}{\| x_0 \|^2} x_0$, then $(x, y_0) = \lambda\, \ell(x_0) = \ell(x)$. Rest of the assertions are obvious. This completes the proof.

Let (Ω, Σ, μ) be a measure space and f be a Σ-measurable function on Ω.

If $\int_{\Omega} f d\mu$ has a meaning, then the set function ν defined by,

$$\nu(A) = \int_A f d\mu, \quad A \in \Sigma$$

is called the indefinite integral of f. Then $\nu(\phi) = 0$ and ν is σ-additive, i.e., if $\{A_n\} \subset \Sigma$ is a disjointed sequence, then

$$\nu\left(\bigcup_n A_n\right) = \sum_n \nu(A_n)$$

Also $\nu(A) = 0$ whenever $A \in \Sigma$ and $\mu(A) = 0$. This provides the definition of absolute continuity of a measure with respect to another measure. Let (Ω, Σ, μ)

NOTES

and (Ω, Σ, ν) be measure spaces. Then ν is said to be absolutely continuous with respect to μ if $\nu(A) = 0$ whenever $A \in \Sigma$ and $\mu(A) = 0$.

Theorem 5.9 (Lebesgue-Nikodym Theorem): Let (Ω, Σ, μ) and (Ω, Σ, ν) be measure spaces with $\mu(\Omega) < +\infty$ and $\nu(\Omega) < +\infty$ where ν is absolutely continuous with respect to μ. Then there exists a unique $h \in L^1(\Omega, \Sigma, \mu)$ such that,

$$\nu(A) = \int_A h d\mu, \quad A \in \Sigma$$

Moreover, $h \geq 0$ μ almost everywhere.

Proof: Let $\rho = \mu + \nu$. Then ρ is a finite measure on Σ. Consider the real Hilbert space $L^2(\Omega, \Sigma, \rho)$ and consider the linear functional ℓ on $L^2(\Omega, \Sigma, \rho)$ defined by,

$$\ell(f) = \int f d\nu$$

As

$$|\ell(f)| \leq \left(\int |f|^2\, d\nu\right)^{1/2} \left(\int 1 d\nu\right)^{1/2} \leq \nu(\Omega)^{1/2} \left[\int |f|^2\, d\rho\right]^{1/2}$$

$$= \nu(\Omega)^{1/2} \,\|f\|\, L^2(\rho)$$

ℓ is a bounded linear functional on $L^2(\Omega, \Sigma, \rho)$. By Riesz representation theorem there is a unique $g \in L^2(\Omega, \Sigma, \rho)$ such that,

$$\int f d\nu = \int fg d\rho = \int fg d\mu + \int fg d\nu$$

for all $f \in L^2(\Omega, \Sigma, \rho)$, or

$$\int f(1-g) d\nu = \int fg d\mu \qquad ...(5.3)$$

for all $f \in L^2(\Omega, \Sigma, \rho)$.

Claim 1: $0 \leq g(x) < 1$ for ρ almost everywhere x on Ω.

Let $A_1 = \{x \in \Omega : g(x) < 0\}$ and $A_2 = \{x \in \Omega : g(x) \geq 1\}$. If we let $f = \chi A_1$ in Equation (5.1), then $0 \leq \nu(A_1) \leq \int_{A_1} (1-g) dv = \int_{A_1} g d\mu$ which implies $\mu(A_1) = 0$ and hence $\nu(A_1) = 0$. Thus $\rho(A_1) = 0$. Now in Equation (5.1) choosing $f = \chi A_2$, we have $0 \geq \int_{A_2} (1-g) dv = \int_{A_2} g d\mu \geq \mu(A_2)$. This implies $\mu(A_2) = 0$ and hence $\nu(A_2) = 0$. Consequently, $\rho(A_2) = 0$. This proves Claim 1.

Claim 2: Equation (5.1) holds for all Σ-measurable and ρ almost everywhere nonnegative functions f. For each positive integer n, let $f_n = f \wedge n$. Since $1 - g > 0$ and $g \geq 0$ ρ almost everywhere, $0 \leq f_n(1-g) \nearrow f(1-g)$ and $0 \leq f_n g \nearrow fg$. Then from Monotone convergence theorem and Equation (5.1) we get,

$$\int f(1-g)d\nu = \lim_{n\to\infty}\int f_n(1-g)d\nu = \lim_{n\to\infty}\int f_n g d\mu = \int fg d\mu$$

which establishes the claim.

NOTES

For a Σ-measurable and ρ almost everywhere greater than or equal to 0 function z choose $f = \dfrac{z}{1-g}$ in Equation (5.1). Then,

$$\int z d\nu = \int z\frac{g}{1-g}d\mu = \int zh d\mu \qquad ...(5.4)$$

where $h = \dfrac{g}{1-g}$. If for $A \in \Sigma$ we take $z = \chi A$ in Equation (5.2), then

$$\nu(A) = \int I_A h d\nu = \int_A h d\mu$$

Since $\nu(\Omega) < +\infty$, we know that $\int h d\mu < +\infty$ and hence $h \in L^1(\Omega, \Sigma, \mu)$. The uniqueness of h is obvious. That $h \geq 0$ μ almost everywhere is also obvious. This completes the proof.

A measure space (Ω, Σ, μ) is said to be σ-finite if there are $A_1, A_2, \ldots$ in Σ such that $\bigcup A_n = \Omega$ and $\mu(A_n) < +\infty$, $n = 1, 2, \ldots$.

Theorem 5.10: Lebesgue-Nikodym theorem holds if both (Ω, Σ, μ) and (Ω, Σ, ν) are σ-finite. But in this case h may not be μ-integrable.

Let X be a Hilbert space. For definiteness, let X be a complex Hilbert space.

$B(.,.): X \times X \mapsto \mathbb{C}$ is called sesquilinear if for x, x_1, x_2 in X and λ_1, λ_2 $\in \mathbb{C}$ the following equalities hold:

$$B(\lambda_1 x_1 + \lambda_2 x_2, x) = \lambda_1 B(x_1, x) + \lambda_2 B(x_2, x)$$
$$B(x, \lambda_1 x_1 + \lambda_2 x_2) = \overline{\lambda}_1 B(x, x_1) + \overline{\lambda}_2 B(x, x_2)$$

B is said to be bounded if there is $r > 0$, such that $|B(x, y) \leq r\|x\| . \|y\|$ for all x and y in X and B is said to be positive definite if there exists $\rho > 0$ such that $|B(x, x) \geq \rho\ \|x\|^2$ for all x in X.

Theorem 5.11: Suppose that B is a bounded, positive definite and sesquilinear function on $X \times X$ and that $B(x, y) = \overline{B(y, x)}$ for all x and y in X. Let $((.,.))$. Then $(X, ((.,.)))$ is a Hilbert space which is equivalent to $(X, (.,.))$ as Banach space.

Theorem 5.12 (Lax-Milgram Theorem): Let X be a Hilbert space and B a bounded, positive definite and sesquilinear functional on $X \times X$. Then there is a unique bounded linear operator $S: X \mapsto X$ such that $(x, y) = B(Sx, y)$ for all x, y $- X$ and $\|S\| \leq \rho^{-1}$. Besides S^{-1} exists and is bounded with $\|S^{-1}\| \leq r$.

NOTES

Proof: Let $D = \{y \in X : \exists\, y^* \in X$. such that $(x, y) = B(x, y^*)\ \forall x \in X\}$. D ***
as $0 \in D$. Also y^* is uniquely determined by y as if

$B(x, y_1^*) = B(x, y_2^*) = (x, y) \forall x \in X$ then $B(x, y_1^* - y_2^*) = 0 \forall x \in X$, and hence 0 = $B(y_1^* - y_2^*, y_1^* - y_2^*) \geq \rho \| y_1^* - y_2^* \|^2$ implying $\| y_1^* - y_2^* \| = 0$ or $y_1^* = y_2^*$.

For $y \in D$, let $Sy = y^*$. As B is sesquilinear, D is a vector subspace of X and S is linear on D. Furthermore, from $\rho \| Sy \|^2 \leq B(Sy, Sy) \ = (Sy, y) \leq$ $\| y \| . \| Sy \|$, we get that $\| Sy \| \leq | \rho^{-1} \| y \|$ for $y \in D$. Thus S is bounded on D with $\| S \| \leq \rho^{-1}$. We will now show that $D = X$. For this we first show that D is closed. Let $\{y_n\}_{n=1}^{\infty} \subset D$ $\lim_{n\to\infty} y_n = y$ with for some $y \in X$.

Then, $(x, y) = \lim_{n\to\infty}(x, y_n) = \lim_{n\to\infty} B(x, Sy_n)$ for all $x \in X$. Since S is bounded on D, Sy_n is Cauchy in X and hence has a limit $z \in X$. This and the boundedness of B implies that, (x, y) = $\lim_{n\to\infty} B(x, Sy_n) = B(x, z)$ for all $x \in X$. Hence $y \in D$ and $z = Sy$. So D is closed. Now if $D \neq X$, there is $y_0 \in D^{\perp}, y_0 \neq 0$. Consider the linear functional ℓ defined on X by,

$\ell(x) = B(x, y_0), x \in X$

As B is bounded, ℓ is a bounded linear functional on X, and hence by Riesz representation theorem there is $x_0 \in X$ such that,

$B(x, y_0) = (x, x_0) x \in X$

Thus $x_0 \in D$ and $\rho \| y_0 \|^2 \leq B(y_0, y_0) = (x_0, y_0) = 0$. Hence $\|y_0\| = 0$. This contradicts the fact that $y_0 \neq 0$. Therefore $D = X$. Thus S is a bounded linear operator on X and $\|S\| \leq \rho^{-1}$.

As $Sy = 0$ implies $(x, y) = B(x, Sy) = 0 \forall x \in X$ and hence $y = 0$, S is a one-to-one map. Applying Riesz representation theorem again, for each y^* in X there exists $y \in X$ such that, $(x, y) = B(x, y^*) \forall x \in X$, i.e., $y^* = Sy$. Thus S is an onto map. Hence S^{-1} exists. But from $\| S^{-1} y \|^2 = | (S^{-1} y, S^{-1} y) | = \ | B | S^{-1} y, y |$ $\leq r \| (S^{-1} y \| . \| y \|$, it follows that $\| S^{-1} \| \leq r$.

5.3.1 Orthogonal Complements

Let V be an inner product space. Two vectors $u, v \in V$ are said to be *orthogonal* if $(u, v) = 0 \Leftrightarrow (v, u) = 0$. So, u is orthogonal to v iff v is orthogonal to u. Since $(0, v) = 0$ for all $v \in V$, 0 is orthogonal to every vector in V.

Conversely, if $u \in V$ is orthogonal to every vector in V, then $(u, u) = 0 \Rightarrow u = 0$.

Let W be a subspace of V.

NOTES

Define $W^{\perp} = \{v \in V \mid (v, w) = 0$ for all $w \in W\}$ ($W^{\perp}$ is read as W perpendicular). Then $W^{\perp}$ is a subspace of V as $0 \in W^{\perp} \Rightarrow W^{\perp} \neq \varphi$ and $v_1, v_2 \in W^{\perp}$, $\alpha, \beta \in F$

$\Rightarrow (\alpha v_1 + \beta v_2, w) = \alpha(v_1, w) + \beta(v_2, w) = 0$ for all $w \in W$

$\Rightarrow \alpha v_1 + \beta v_2 \in W^{\perp}$.

$W^{\perp}$ is called *orthogonal complement* of W.

Example 5.9: Let V be an inner product space. Let $x, y \in V$ such that, $x \perp y$

Then show that $\| x + y \|^2 = \| x \|^2 + \| y \|^2$. (This is Pythagoras Theorem when $F = \mathbf{R}$ as in triangle ABC with $AB \perp BC$, $AB^2 = \| x \|^2$, $BC^2 = \| y \|^2$, $AC^2 = \| x + y \|^2$)

Solution:
$$\begin{aligned} \| x + y \|^2 &= (x + y, x + y) \\ &= (x, x) + (y, y) + (x, y) + (y, x) \\ &= \| x \|^2 + \| y \|^2 \quad \text{as } (x, y) = 0 = (y, x). \end{aligned}$$

Orthonormal Set

A set $\{u_i\}_i$ of vectors in an inner product space V is said to be orthogonal if $(u_i, u_j) = 0$ for $i \neq j$. If further $(u_i, u_i) = 1$ for all i, then the set $\{u_i\}$ is called an *orthonormal set*.

Example 5.10: Let V be the real vector space of real polynomials of degree less than or equal to n. Define an inner product on V by

$$\left(\sum_{i=0}^{n} a_i x^i, \sum_{j=1}^{n} b_j x^j \right) = \sum_{1}^{n} a_i b_i$$

Then $\{1, x, ..., x^n\}$ is an orthonormal subset of V.

Theorem 5.13: Let S be an orthogonal set of non zero vectors in an inner product space V. Then S is a linearly independent set.

Proof: To show S is linearly independent, we have to show that every finite subset of S is linearly independent.

Let $\{v_1, ..., v_n\}$ be a finite subset of S.

Let, $\alpha_1 v_1 + ... + \alpha_n v_n = 0, \quad \alpha_i \in F$

$(\alpha_1 v_1 + ... + \alpha_n v_n, \alpha_1 v_1 + ... + \alpha_n v_n) = 0$

$\Rightarrow | \alpha_1 |^2 \| v_1 \|^2 + ... + | \alpha_n |^2 \| v_n \|^2 = 0$

$\Rightarrow | \alpha_i |^2 \| v_i \|^2 = 0$ for all $i = 1, ..., n$

$\Rightarrow | \alpha_i |^2 = 0$ for all i as $\| v_i \|^2 = 0 \Rightarrow \| v_i \| = 0 \Rightarrow v_i = 0$

which is not true

$\Rightarrow \alpha_i = 0$ for all $i = 1, ..., n$

$\Rightarrow S$ is linearly independent.

Corollary 2.: An orthonormal set in an inner product space is linearly independent.

Proof: Let S be an orthonormal set in an inner product space V. Let $v \in S$. Then $v \neq 0$ as $v = 0 \Rightarrow (v, v) = 0 \neq 1$, a contradiction. Therefore, S is an orthogonal set of non zero vectors and so linearly independent.

Theorem 5.14: (Gram-Schmidt Orthogonalization process)

Let V be a non zero inner product space of dimension n. Then V has an orthonormal basis.

NOTES

Proof: It is enough to construct an orthogonal basis of V. For let $S \subseteq V$ be an orthogonal set. Then $T = \left\{ \frac{x}{\|x\|} \mid x \in S \right\}$ is an orthonormal set.

Let $\{v_1,..., v_n\}$ be a basis of V.

Let $w_1 = v_1$. Define $w_2 = v_2 - \frac{(v_2, w_1)}{(w_1, w_1)} w_1$

$$= v_2 - \frac{(v_2, v_1)}{(v_1, v_1)} v_1$$

Then $(w_2, w_1) = (w_2, v_1)$

$$= (v_2, v_1) - \frac{(v_2, v_1)}{(v_1, v_1)} (v_1, v_1) = 0$$

Also $\quad v_2 = \alpha_1 v_1 + w_2 = \alpha_1 w_1 + w_2$

where $\alpha_1 = \frac{(v_2, v_1)}{(v_1, v_1)} \in F$.

(Note v_1 is linearly independent $\Rightarrow v_1 \neq 0 \Rightarrow (v_1, v_1) \neq 0$)

Define $\quad w_3 = v_3 - \frac{(v_3, w_2)}{(w_2, w_2)} w_2 - \frac{(v_3, w_1)}{(w_1, w_1)} w_1$

Then $\quad (w_3, w_2) = 0 = (w_3, w_1)$

Also $\quad v_3 = \alpha_1 w_1 + \alpha_2 w_2 + w_3$, where $\alpha_1, \alpha_2 \in F$.

In this way, we can construct an orthogonal set $\{w_1,..., w_n\}$ where each $v_i = \alpha_1 w_1 + ... + w_i, \quad \alpha_i \in F$

$\therefore \left\{ \frac{w_1}{\|w_1\|},..., \frac{w_n}{\|w_n\|} \right\}$ is an orthonormal set which is linearly independent by Corollary 1 to Theorem 5.13 and hence forms a basis of V as dim $V = n$.

Aliter: Let dim $V = n$. We use induction on n.

Let $n = 1$. Let $0 \neq x \in V$, then $v = \frac{x}{\|x\|} \in V$ such that, $\|v\| = 1$.

So, $\{v\}$ is an orthonormal basis of V.

Suppose now that the result holds for any inner product space of dimension less than or equal to $n - 1$.

Let V be an inner product space of dimension n

Let $0 \neq v \in V$ be such that, $\|v\| = 1$.

Define $\quad T_v : V \to \mathbf{C}$ such that,

$$T_v(v') = \langle v', v \rangle$$

Then T_v is a linear transformation.

Let $\alpha \in \mathbf{C}$, then $\alpha = \alpha \|v\|^2 = \alpha \langle v, v \rangle = \langle \alpha v, v \rangle = T_v(\alpha v)$

NOTES

and so T_v is onto, i.e., Range $T_v = \mathbf{C}$.

By Sylvester's law

$$\dim V = \dim \text{Ker } T_v + \dim \text{Range } T_v$$

$$\Rightarrow n = \dim \text{Ker } T_v + \dim \mathbf{C}$$

$$= \dim \text{Ker } T_v + 1$$

$$\Rightarrow \dim W = n - 1, \text{ where } W = \text{Ker } T_v$$

$$= \{x \in V \mid T_u(x) = 0\}$$

$$= \{x \in V \mid < v, x > = 0\}$$

By induction hypothesis, W has an orthonormal basis $\{w_1, w_2,..., w_{n-1}\}$

Now $\quad w_i \in W \Rightarrow < v, w_i > = 0 \quad \forall i = 1, 2,..., n-1$

Also $\quad < v, v > = \| v \|^2 = 1$

So $\{w_1, w_2,..., w_{n-1}, v\}$ is an orthonormal set.

i.e., $\{w_1, w_2,..., w_{n-1}, v\}$ is L.I. set by Corollary 1 to Theorem 5.13.

Since $\dim V = n$, $\{w_1, w_2,..., w_{n-1}, v\}$ is a basis of V and hence is an orthonormal basis of V. So, result follows by induction.

Example 5.11: Obtain an orthonormal basis, with respect to the standard inner product for the subspace of $\mathbf{R}^3$ generated by (1, 0, 3) and (2, 1, 1).

Solution: Let $v_1 = (1, 0, 3)$, $v_2 = (2, 1, 1)$

Then $\quad w_1 = v_1,\ w_2 = v_2 - \dfrac{(v_2, w_1)}{(w_1, w_1)}\ w_1$

Now $\quad (v_2, w_1) = (v_2, v_1) = 2 + 0 + 3 = 5$

$\quad (w_1, w_1) = (v_1, v_1) = 1 + 0 + 9 = 10$

$\therefore \quad \| w_1 \| = \sqrt{10}$

So, $\quad w_2 = (2, 1, 1) - \dfrac{5}{10}\ (1, 0, 3) = \left(\dfrac{3}{2}, 1, -\dfrac{1}{2}\right)$

$\therefore \quad \| w_2 \| = \sqrt{\dfrac{9}{4} + 1 + \dfrac{1}{4}} = \sqrt{\dfrac{7}{2}}$

$\therefore$ Required orthonormal basis is

$$\left\{\frac{w_1}{\| w_1 \|}, \frac{w_2}{\| w_2 \|}\right\} = \left\{\frac{1}{\sqrt{10}}(1, 0, 3),\ \frac{\sqrt{2}}{7}\left(\frac{\sqrt{3}}{2}, 1, -\frac{1}{2}\right)\right\}$$

Example 5.12: Let V be an inner product space over $\boldsymbol{R}$. *Let* $\{v_1, v_2,..., v_n\}$ be a basis of V such that, whenever $v = \Sigma\alpha_i v_i$ then $|| v ||^2 = \Sigma\alpha_i^2$. Show that $\{v_1, v_2,..., v_n\}$ is an orthonormal basis.

Solution: We have $v_i = 1.v_i \Rightarrow \| v_i \|^2 = 1 \quad \forall i$ by hypothesis

Consider $v_i + v_j$, $i \neq j$, then

$$\| v_i + v_j \|^2 = 2$$

$$\Rightarrow < v_i, v_i > + < v_j, v_j > + < v_i, v_j > + < v_j, v_i > = 2$$

$$\Rightarrow < v_i, v_j > + < v_j, v_i > = 0$$

$\Rightarrow$ $< v_i, v_j > + < v_i, v_j > = 0$ as V is an inner product space over **R**

$\Rightarrow$ $< v_i, v_j > = 0 \quad \forall i \neq j$

Hence $\{v_1, v_2,..., v_n\}$ is an orthonormal basis.

NOTES

Theorem 5.15: (Bessel's Inequality)

If $\{w_1,..., w_m\}$ is an orthonormal set in V, then

$$\sum_{i=1}^{m} |(w_i, v)|^2 \leq \| v \|^2 \textit{ for all } v \in V.$$

Proof: Let $x = v - \sum_{i=1}^{m} (v, w_i)\, w_i$

$\therefore \quad (x, w_j) = (v, w_j) - (v, w_j) = 0$ for all $j = 1,..., m$

Let $\quad w = \sum_{i=1}^{m} (v, w_i)\, w_i = \sum_{i=1}^{m} \alpha_i w_i, \ \alpha_i = (v, w_i)$

$\therefore \quad v = x + w$

Also $\quad (w, x) = (\alpha_1 w_1 + ... + \alpha_m w_m, x)$

$= \alpha_1 (w_1, x) + ... + \alpha_m (w_m, x) = 0$

Now $\quad \| v \|^2 = (v, v)$

$= (w + x, w + x)$

$= (w, w) + (x, x)$

$= \| w \|^2 + \| x \|^2 \geq \| w \|^2$

But $\quad \| w \|^2 = (w, w)$

$= (\alpha_1 w_1 + ... + \alpha_m w_m, \alpha_1 w_1 + ... + \alpha_m w_m)$

$= \alpha_1 \overline{\alpha_1}\ (w_1, w_1) + ... + \alpha_m \overline{\alpha_m}\ (w_m, w_m)$

$= | \alpha_1 |^2 + ... + | \alpha_m |^2$

as $\{w_1,....., w_m\}$ is an orthonormal set

$$= \sum_{i=1}^{m} | \alpha_i |^2 = \sum_{i=1}^{m} | (v_i, w_i) |^2 = \sum_{i=1}^{m} | \overline{(w_i, v)} |^2 = \sum_{i=1}^{m} | (w_i, v) |^2$$

$\therefore \quad \sum_{i=1}^{m} | (w_i, v) |^2 \leq \| v \|^2$ for all $v \in V$.

Corollary 3.: Equality holds if and only if $v = w$.

Proof: Suppose $v = w$

Then $\quad \| v \|^2 = \| w \|^2 = \sum_{i=1}^{m} | (w_i, v) |^2$

Conversely, suppose equality holds

Then $\quad \| v \|^2 = \| w \|^2$

$\Rightarrow \| x \|^2 = 0 \Rightarrow (x, x) = 0 \Rightarrow x = 0$

$\Rightarrow v = w + x = w.$

NOTES

Theorem 5.16: If V is a finite dimensional inner product space and W is a subspace of V, then $V = W \oplus W^{\perp}$.

Proof: Since V is an inner product space, so is W. By Theorem 5.14, W has an orthonormal basis $\{w_1,..., w_m\}$.

Let $v \in V$.

Let $w = \sum_{i=1}^{m}(v, w_i)w_i$, $w_i \in W$ and $x = v - w$

Then $(x, w_j) = 0$ as in Theorem 5.15, for all $j = 1,..., m$

$$\begin{aligned}
\therefore \quad (x, w) &= (x, \beta_1 w_1 + ... + \beta_m w_m) \\
&= \overline{\beta_1}\ (x, w_1) + ... + \overline{\beta_m}\ (x, w_m) \\
&= 0 \text{ for all } w \in W
\end{aligned}$$

$\therefore \quad x \in W^{\perp}$

So, $\quad v = w + x \in W + W^{\perp}$

$V \subseteq W + W^{\perp}$

$\Rightarrow \quad V = W + W^{\perp}$

Let $y \in W \cap W^{\perp} \Rightarrow (y, w) = 0$ for all $w \in W, y \in W$

$\Rightarrow (y, y) = 0$ as $y \in W$

$\Rightarrow y = 0$

$\therefore \quad W \cap W^{\perp} = \{0\}$

Hence $\quad V = W \oplus W^{\perp}$.

Corollary 4: If W is a subspace of a finite dimensional inner product space V, then $(W^{\perp})^{\perp} = W$.

By above theorem, $V = W \oplus W^{\perp}$

Let $\quad w \in W, x \in W^{\perp}$

Then $\quad x \in W^{\perp} \quad \Rightarrow < x, y > = 0 \quad \forall\, y \in W$

$\Rightarrow < x, w) = 0 \quad \forall\, x \in W^{\perp}$

$\Rightarrow w \in (W^{\perp})^{\perp}$

i.e., $\quad W^{\perp} \subseteq (W^{\perp})^{\perp}$

Let $v \in (W^{\perp})^{\perp}$ then $v = w + w'$, $w \in W$, $w' \in W^{\perp}$

$\Rightarrow 0 = < w', v > = < w', w + w') = < w', w > + < w', w' > = < w', w' >$

So $\quad w' = 0 \Rightarrow v = w \in W$

i.e., $\quad (W^{\perp})^{\perp} \subseteq W$ giving $W = (W^{\perp})^{\perp}$.

Corollary 5: If $S = \{x_1, x_2,..., x_r\}$ is a basis of W and $T = \{y_1, y_2,..., y_s\}$ is a basis of $W^{\perp}$ then

$\{x_1, x_2,..., x_r, y_1, y_2,..., y_s\}$ is an orthonormal basis of V.

By above theorem $V = W \oplus W^{\perp}$

Thus $S \cup T$ is a basis of V

Also $< x_i, y_j > = 0 \quad \forall\ i, j$ as $y_j \in W^{\perp}\ \forall j$

proving the result.

Note: Theorem 5.16 need not hold in case of infinite dimensional vector space. For instance, take

$$V = \left\{(a_n) \mid (a_n) \text{ is a sequence of complex numbers such that, } \sum_1^{\infty} |a_n|^2 < \infty\right\}.$$

Then V is a vector space with respect to componentwise addition and scalar multiplication

Take $\quad a = (a_n),\ b = (b_n) \in V$

Define $\quad < a, b > = \sum_1^{\infty} a_n \overline{b_n}$

Since $\quad (\,|a_n| - |b_n|\,)^2 \geq 0$

$|a_n|^2 + |b_n|^2 \geq 2\,|a_n|\,|b_n|$

Now $2\,|\Sigma a_n \overline{b}_n| \leq 2\Sigma\,|a_n|\,|\overline{b}_n|$

$\Rightarrow\ 2\,|\Sigma a_n \overline{b}_n| \leq 2\ \Sigma|a_n|\,|b_n|$ as $|b_n| = |\overline{b}_n|$

$\leq\ \Sigma|a_n|^2 + \Sigma|b_n|^2 < \infty$

Thus $< a, b >$ is well defined inner product on V.

Let $A_k \in V$ such that, k^{th} entry is 1 and zero elsewhere

Let $S = \{|\, A_k \mid k = 1, 2, \ldots, \} \subseteq V$

Then $< A_i, A_j > = \delta_{ij}$.

Let $W = L(S)$, then $W \neq V$ as $v = \left\{\frac{1}{n^2}\right\} \in V$ and $v \notin L(S)$.

[In fact $L(S)$, is the set of those sequences whose only finite number of entries are non zero].

Also $x \in W^{\perp} \Rightarrow\ < x, w > = 0 \qquad \forall\ w \in W$

$\Rightarrow\ < x, A_k > = 0\ \forall\ k = 1, 2, \ldots$

$\Rightarrow\ x_k = 0\ \forall k \quad$ where $x = (x_n)$

$\Rightarrow\ x = 0$ or that $W^{\perp} = \{0\}$

So $\quad V \neq W \oplus W^{\perp} = W$.

Notice V is not *F.D.V.S.* by Theorem 5.16.

Example 5.13: If W is a subspace of V and $v \in V$ satisfies

$$(v, w) + (w, v) \leq (w, w) \text{ for all } w \in W$$

prove that $(v, w) = 0$ for all $w \in W$, *where* V is an inner product space over F.

Solution: Let n be a +ve integer

Then $\quad w \in W \Rightarrow \frac{w}{n} \in W$

NOTES

$\therefore \qquad \left(v, \frac{w}{n}\right) + \left(\frac{w}{n}, v\right) \le \left(\frac{w}{n}, \frac{w}{n}\right)$

$\therefore \qquad (v, w) + (w, v) \le \frac{1}{n}\ (w, w)$

Let $\qquad n \to \infty$

Then $\qquad (v, w) + (w, v) \le 0 \quad$ for all $w \in W$

$(v, -w) + (-w, v) \le 0 \quad$ for all $w \in W$

$\Rightarrow -[(v, w) + (w, v)] \le 0$ for all $w \in W$

$\Rightarrow (v, w) + (w, v) \ge 0 \quad$ for all $w \in W$

$\Rightarrow (v, w) + (w, v) = 0 \quad$ for all $w \in W$

If $F \subseteq \mathbf{R}$, then $\quad (w, v) = (v, w)$

$\Rightarrow (v, w) + (v, w) = 0$

$\Rightarrow 2(v, w) = 0 \quad$ for all $w \in W$

$\Rightarrow (v, w) = 0 \quad$ for all $w \in W$

If $F \subseteq \mathbf{C}$, then $(v, iw) + (iw, v) = 0 \quad$ for all $w \in W$

$\Rightarrow -i\,(v, w) + i\,(w, \bar{z}) = 0 \quad$ for all $w \in W$

$\Rightarrow -i\,[z - \bar{z}] = 0,\ z = (v, w) = x + iy$

$\Rightarrow -i\,(2iy) = 0$

$\Rightarrow y = 0$

$\Rightarrow z = (v, w) =$ real $\quad$ for all $w \in W$

$\Rightarrow (v, w) + (v, w) = 0$

$\Rightarrow 2\,(v, w) = 0$

$\Rightarrow (v, w) = 0\ $ for all $w \in W$.

Example 5.14: If V is a finite dimensional inner product space and $f \in V$, prove that $\exists\, u_0 \in V$ such that $f(v) = (v, u_0)$ for all $v \in V$. Also show that u_0 is uniquely determined.

Solution: Let $\{v_1, ..., v_n\}$ be an orthonormal basis for V. Let $v \in V$.

Then $\qquad v = \alpha_1 v_1 + ... + \alpha_n v_n, \quad \alpha_i \in F$

Let $\qquad f(v_i) = \beta_i, \quad i = 1, 2, ..., n$

Define $\qquad u_0 = \overline{\beta_1}\, v_1 + ... + \overline{\beta_n}\, v_n \in V$

Then $\qquad (v, u_0) = (\alpha_1 v_1 + ... + \alpha_n v_n,\ \overline{\beta_1}\, v_1 + ... + \overline{\beta_n}\, v_n)$

$= \alpha_1 \beta_1 + ... + \alpha_n \beta_n \ $ as $(v_i, v_j) = \delta_{ij}$

$= f(v)\ $ for all $v \in V$

Suppose $\qquad \exists\, u_0' \in V$ such that $f(v) = (v, u_0')$

Then $\qquad (v, u_0) = (v, u_0')$ for all $v \in V$

$\Rightarrow (v, u_0 - u_0') = 0$ for all $v \in V$

$\Rightarrow (u_0 - u_0', u_0 - u_0') = 0$

$$\Rightarrow u_0 = u_0'$$

$\therefore u_0$ is uniquely determined.

NOTES

5.3.2 Conjugate Space H*

Consider a Hilbert space H and its conjugate space H*. Let y be a fixed vector in H. Define a function f_y on H by,

$f_y(x) = (x, y)$ for all x in H

Assert that f_y is linear for $f_y(x_1+x_2) = (x_1+x_2, y)$ for all x_1 and x_2 in H.

$$= (x_1, y) + (x_2, y) = f_y(x_1) + f_y(x_2)$$

Also

$f_y(\alpha x) = (\alpha x, y) = \alpha(x, y) = \alpha(f_y(x))$

which proves that $\|f_y\| \leq \|y\|$

This implies that f_y is continuous. Thus f_y is linear and continuous mapping and hence is a linear functional on H. On the other hand if y = 0, then $f_y(x) = (x, 0) = 0 \Rightarrow \|f_y\| = \|y\|$. If $y \neq 0$, then

$\|f_y\| = \sup\{|f_y(x)|; \|x\| = 1\}$

$$\geq \left| f_y\left(\frac{y}{\|y\|}\right) \right| \geq \left| \left(\frac{y}{\|y\|}, y\right) \right|$$

Hence $\|f_y\| = \|y\|$

Thus for each $y \in$ H, there is a linear functional $f_y \in$ H* such that $\|f_y\| = \|y\|$. Hence the mapping $y \to f_y$ is a norm preserving mapping of H into H*.

5.3.3 Reflexivity of Hilbert Space

Theorem 5.17: For every Hilbert space $\mathcal{H}$ there exists a surjective isometry $\Psi: \mathcal{H}^* \to \mathcal{H}$ of the dual $\mathcal{H}^*$ of $\mathcal{H}$ onto $\mathcal{H}$ which is additive and conjugate homogeneous (i.e., $\Psi(\alpha f) = \bar{\alpha}\Psi(f)$ for every $f \in \mathcal{H}^*$ and every $\alpha \in \mathbb{F}$).

Proof: Suppose $\mathcal{H}$ is a Hilbert space and let $\mathcal{H}^* = \beta[\mathcal{H}, \mathbb{F}]$ be the dual of $\mathcal{H}$. By the Riesz represenatation theorem, for each $f \in \mathcal{H}^*$ there exists a unique $y \in \mathcal{H}$ such that $f(x) = (x; y)$ for every $x \in \mathcal{H}$ and $\|f\| = \|y\|$. Conversely, for each $y \in \mathcal{H}$ the functional $f: \mathcal{H} \to \mathbb{F}$ given by $f(x) = \langle x; y \rangle$ for every $x \in \mathcal{H}$ is linear and bounded, i.e., $f \in \mathcal{H}^*$. This proves the surjective isometry $\Psi: \mathcal{H}^* \to \mathcal{H}$ of the dual $\mathcal{H}^*$ of $\mathcal{H}$ onto $\mathcal{H}$:

$\Psi(f) = y$ for every $f \in \mathcal{H}^*$

where $y \in \mathcal{H}$ is the unique Riesz representation of $f \in \mathcal{H}^*$. Therefore every f in $\mathcal{H}^*$ is such that, $f(x) = \langle x; \Psi(f) \rangle$ for every $x \in \mathcal{H}^*$

Notice that Ψ is additive. Clearly, if $f, g \in \mathcal{H}^*$, then

$$\langle x; \Psi(f+g) \rangle = (f+g)(x) = f(x) + g(x)$$

$$= \langle x;\Psi(f)\rangle + \langle x;\Psi(g)\rangle = \langle x;\Psi(f)+\Psi(g)\rangle$$

for every $x \in \mathcal{H}$, so that $\Psi(f+g) = \Psi(f)+\Psi(g)$. Moreover, if $f \in \mathcal{H}^*$ and $\alpha \in \mathbb{F}$ then, $\langle x;\Psi(\alpha f)\rangle = \alpha f(x) = f(\alpha x) = \langle \alpha x;\Psi(f)\rangle = \langle x;\bar{\alpha}\,\Psi(f)\rangle$

for every $x \in \mathcal{H}$ and hence $\Psi(\alpha f) = \alpha\Psi(f)$. This completes the proof.

From the above theorem we can conclude that every Hilbert space is isometrically equivalent to its dual. In particular, every real Hilbert space is isometrically isomorphic to its dual.

Theorem 5.18: Every Hilbert space is reflexive.

Proof: Let $\Psi : \mathcal{H}^* \to \mathcal{H}$ be the surjective isometry of Theorem 5.17 which is additive and conjugate homogeneous. Let the mapping $\langle ; \rangle_* : \mathcal{H}^* \times \mathcal{H}^* \to \mathbb{F}$ be given by, $\langle f;g\rangle_* = \langle \Psi(g);\Psi(f)\rangle$

for every $f, g \in \mathcal{H}^*$, where $\langle\,;\,\rangle$ is the inner product on $\mathcal{H}$. This defines an inner product on $\mathcal{H}^*$. $\langle\,;\,\rangle_*$ is additive Ψ since is additive. Now as Ψ is conjugate homogeneous,

$$\langle \alpha f;g\rangle_* = \langle \Psi(g);\Psi(\alpha f)\rangle = \langle \Psi(g);\bar{\alpha}\Psi(f)\rangle = \alpha\langle \Psi(g);\Psi(f)\rangle = \alpha\langle f;g\rangle_*$$

for every $f, g \in \mathcal{H}^*$ and every $\alpha \in \mathbb{F}$ and so $\langle\,;\,\rangle_*$ is homogeneous in the first argument. Evidently $\langle\,;\,\rangle_*$ is Hermitian symmetric and positive. Now, $\|f\|_* = \|\Psi(f)\| = \|f\|$

for every $f \in \mathcal{H}^*$, so that the norm $\|\;\|_*$ induced on $\mathcal{H}^*$ by the inner product $\langle\,;\,\rangle_*$ coincides with the usual induced norm on $\mathcal{H}^* = \beta[\mathcal{H},\mathbb{F}]$. Since the dual space of every normed space is a Banach space, $(\mathcal{H}^*, \|\;\|)$ is a Banach space and hence $(\mathcal{H}^*, \|\;\|_*)$ is a Hilbert space. We will now apply the Riesz representation theorem to the Hilbert space $\mathcal{H}^*$. Obtain an arbitrary $\varphi \in \mathcal{H}^{**}$. There exists a unique $g \in \mathcal{H}^*$ such that, $\varphi(f) = \langle f;g\rangle_* = \langle \Psi(g);\Psi(f)\rangle$

for every $f \in \mathcal{H}^*$. Every $f \in \mathcal{H}^*$ is given by $f(x) = \langle x;y\rangle$ for every $x \in \mathcal{H}$, where $y = \Psi(f) \in \mathcal{H}$. Fix $z = \Psi(g) \in \mathcal{H}$ such that,

$$f(z) = \langle z;y\rangle = \langle \Psi(g);\Psi(f)\rangle$$

Hence there exists $z \in \mathcal{H}$ such that, $\varphi(f) = f(z)$ for every $f \in \mathcal{H}^*$.

Therefore $\mathcal{H}$ is reflexive.

Check Your Progress

1. Define the term norm.
2. State Riesz representation theorem.
3. What can you say about the reflexivity of a Hilbert space?

5.4 SELF-ADJOINT OPERATORS ON HILBERT SPACE

NOTES

Definition: A bounded linear operator $\mathbf{T : H \to H}$ on a Hilbert space H is said to be self-adjoint or Hermitian if,

$$\mathbf{T}^* = \mathbf{T} \qquad \ldots(5.5)$$

Equivalently, a bounded linear operator T is said to be self-adjoint if,

$$\langle x, \mathbf{T}y \rangle = \langle \mathbf{T}x, y \rangle \text{ for all } x, y \in \mathbf{H} \qquad \ldots(5.6)$$

A linear map on $\mathbb{R}^n$ with materix A is self-adjoint iff A is symmetric ($\mathbf{A = A^T}$). A linear map on $\mathbb{C}^n$ with matrix A is self-adjoint iff A is Hermitian ($\mathbf{A = A^*}$).

Not: Self-adjoint operators on Hilbert spaces are used in quantum mechanics for representing physical observables like position, momentum, angular momentum and spin.

Definition: A bounded linear operator $\mathbf{T : H \to H}$ on a Hilbert space H is said to be unitary if T is bijective and

$$\mathbf{TT^* = T^*T} \qquad \ldots(5.7)$$

Hence

$$\mathbf{T^* = T^{-1}} \qquad \ldots(5.8)$$

Definition: A bounded linear operator $\mathbf{T : H \to H}$ on a Hilbert space H is said to be normal if,

$$\mathbf{TT^* = T^*T} \qquad \ldots(5.9)$$

Note: If T is self-adjoint or unitary, then T is normal; the converse is not generally true.

Theorem 5.19: Let $\mathbf{T : H \to H}$ be a bounded linear operator on a Hilbert space H. Then,

1. If T is self-adjoint then $\langle \mathbf{T}x, x \rangle \in \mathbb{R}$ for all $x \in \mathbf{H}$.
2. If H is complex and $\langle \mathbf{T}x, x \rangle \in \mathbb{R}$ for all $x \in \mathbf{H}$, then the operator T is self-adjoint.

Proof:

1. If T is self-adjoint, then for all x,

$$\overline{\langle \mathbf{T}x, x \rangle} = \langle x, \mathbf{T}x \rangle \qquad \ldots(5.10)$$

By definition $\langle \mathbf{T}x, y \rangle = \langle x, \mathbf{T}^* y \rangle$ and since T is self-adjoint, we have

$$\langle \mathbf{T}x, x \rangle = \langle x, \mathbf{T}x \rangle \qquad \ldots(5.11)$$

Combining Equations (5.10) and (5.11) gives,

$$\overline{\langle \mathbf{T}x, x \rangle} = \langle \mathbf{T}x, x \rangle \qquad \ldots(5.12)$$

Hence $\langle \mathbf{T}x, x \rangle$ is equal to its complex conjugate which implies that it is real.

2. If $\langle \mathbf{T}x, x\rangle \in \mathbb{R}$ for all $x \in \mathbf{H}$, then

$$\langle \mathbf{T}\,x, x\rangle = \overline{\langle \mathbf{T}\,x, x\rangle} = \overline{\langle x, \mathbf{T}^*\,x\rangle} = \langle \mathbf{T}^* x, x\rangle$$

Hence,

$$0 = \langle \mathbf{T}\,x, x\rangle - \langle \mathbf{T}^*\,x, x\rangle = \langle (\mathbf{T} - \mathbf{T}^*)x, x\rangle$$

and by Lemma 1, $\mathbf{T} - \mathbf{T}^* = 0$. Therefore, $\mathbf{T} = \mathbf{T}^*$.

Theorem 5.20: Consider $\mathbf{T}_n$ to be a sequence of bounded self-adjoint linear operators $\mathbf{T}_n : \mathbf{H} \to \mathbf{H}$ on a Hilbert space H. If $\mathbf{T}_n$ converges to T, then T is a bounded self-adjoint linear operator.

Proof: If $\mathbf{T}_n \to \mathbf{T}$ Then $\|\mathbf{T}_n - \mathbf{T}\| \to 0$.

Also,

$$\|\mathbf{T}_n^* - \mathbf{T}^*\| = \|(\mathbf{T}_n - \mathbf{T})^*\| = \|\mathbf{T}_n - \mathbf{T}\|$$

So that, $\|(\mathbf{T} - \mathbf{T}^*)\| \leq \|(\mathbf{T} - \mathbf{T}_n)\| + \|\mathbf{T}_n - \mathbf{T}_n^*\| + \|\mathbf{T}_n^* - \mathbf{T}^*\|$

$= \|(\mathbf{T} - \mathbf{T}_n)\| + \|\mathbf{T}_n - \mathbf{T}\| = 2\,\|\mathbf{T}_n - \mathbf{T}\|$

As $n \to \infty$, $\|\mathbf{T}_n - \mathbf{T}\| \to 0$. Hence $\|\mathbf{T} - \mathbf{T}^*\| = 0$ implying $\mathbf{T}^* = \mathbf{T}$, Hence T is self-adjoint.

Definition: Let H be a Hilbert space. Then S : H $\to$ H is positive operator denoted by $S \geq 0$ if $\langle S\,f, f\rangle$ is real and $\langle S\,f, f\rangle \geq 0$ for every $f \in H$.

A positive operator on a complex Hilbert space is self-adjoint.

5.4.1 Projections on Hilbert Spaces

Let M and N are subspaces of a linear space X such that every $x \in X$ can be written exclusively as $x = y + z$ with $y \in M$ and $z \in N$. Then we say that $X = M \oplus N$ is the direct sum of M and N and we call N a complementary subspace of M in X. The decomposition $x = y + z$ with $y \in M$ and $z \in N$ is unique iff $M \cap N = \{0\}$. A given subspace M has many complementary subspaces. For example, if $X = \mathbb{R}^3$ and M is a plane through the origin, then any line through the origin that does not lie in M is a complementary subspace. Every complementary subspace of M has the same dimension and the dimension of a complementary subspace is called the codimension of M in X. If $X = M \oplus N$, then the projection $P : X \to X$ of X onto M along N is defined by $Px = y$, where $x = y + z$ with $y \in M$ and $z \in N$. This projection is linear with rank $P = M$ and ker $P = N$, and satisfies $P^2 = P$.

Definition: A projection on a linear space X is a linear map $P : X \to X$ such that,

$$P^2 = P \qquad ...(5.13)$$

Any projection is associated with a direct sum decomposition.

Theorem 5.21: Suppose X is a linear space.

1. If $P : X \to X$ is a projection, then $X = \text{ran } P \oplus \ker P$.
2. If $X = M \oplus N$, where M and N are linear subspaces of X, then there is a projection $P : X \to X$ with ran $P = M$ and $\ker P = N$.

NOTES

Proof:

1. First prove that $x \in \text{ran } P$ iff $x = Px$. If $x = Px$, then noticeably $x \in \text{ran } P$. If $x \in \text{ran } P$ then $x = Py$ for some $y \in X$, and since $P^2 = P$, it follows that $Px = P^2y = Py = x$.

   If $x \in \text{ran } P \cap \ker P$ then $x = Px$ and $Px = 0$, so $\text{ran } P \cap \ker P = \{0\}$. If $x \in X$ then we have,

   $x = Px + (x - Px)$

   where $Px \in \text{ran } P$ and $(x - Px) \in \ker P$ since $P(x - Px) = Px - P^2x = Px - Px = 0$

   Thus $X = \text{ran } P \oplus \ker P$.

2. Observe that if $X = M \oplus N$, then $x \in X$ has the unique decomposition $x = y + z$ with $y \in M$ and $z \in N$, and $Px = y$ defines the required projection. This completes the proof.

When using Hilbert spaces, we are particularly interested in orthogonal subspaces. Suppose that $\mathcal{M}$ is a closed subspace of a Hilbert space $\mathcal{H}$. Then we have $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. We call the projection of $\mathcal{H}$ onto $\mathcal{M}$ along $\mathcal{M}^\perp$ the orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}$. If $x = y + z$ and $x' = y' + z'$, where $y, y' \in \mathcal{M}$ and $z, z' \in \mathcal{M}^\perp$, then the orthogonality of $\mathcal{M}$ and $\mathcal{M}^\perp$ implies that,

$$\langle Px, x' \rangle = \langle y, y' + z' \rangle = \langle y, y' \rangle = \langle y + z, y' \rangle = \langle x, Px' \rangle \qquad (5.14)$$

Equation (5.14) implies that an orthogonal projection is self-adjoint. The Equations (5.13) and (5.14) characterize orthogonal projections.

Definition: An orthogonal projection on a Hilbert space $\mathcal{H}$ is a linear map $P : \mathcal{H} \to \mathcal{H}$ that satisfies,

$$P^2 = P, \qquad \langle Px, y \rangle = \langle x, Py \rangle \text{ for all } x, y \in \mathcal{H}$$

An orthogonal projection is necessarily bounded.

Corollary 6: Let P be a nonzero orthogonal projection. Then $\|P\| = 1$.

Proof: If $x \in \mathcal{H}$ and $Px \neq 0$, then the use of the Cauchy-Schwarz inequality implies that,

$$\| Px \| = \frac{\langle Px, Px \rangle}{\| Px \|} = \frac{\langle x, P^2 x \rangle}{\| Px \|} = \frac{\langle x, Px \rangle}{\| Px \|} \leq \| x \|$$

Therefore $\| P \| \leq 1$. If $P \neq 0$ then there is an $x \in \mathcal{H}$ with $Px \neq 0$ and $\|P(Px)\| = \|Px\|$, so that $\|P\| \geq 1$.

There is a one-to-one correspondence between orthogonal projections P and closed subspaces $\mathcal{M}$ of $\mathcal{H}$ such that $\text{ran } P = \mathcal{M}$. The kernel of the orthogonal projection is the orthogonal complement of $\mathcal{M}$.

Theorem 5.22: Suppose that $\mathcal{H}$ is a Hilbert space.

1. If P is an orthogonal projection on $\mathcal{H}$, then $\text{ran } P$ is closed and $\mathcal{H} = \text{ran } P \oplus \ker P$ is the orthogonal direct sum of $\text{ran } P$ and $\ker P$.

2. If $\mathcal{M}$ is a closed subspace of $\mathcal{H}$ then there is an orthogonal projection P on $\mathcal{H}$ with ran $P = \mathcal{M}$ and ker $P = \mathcal{M}^{\perp}$.

Proof:

NOTES

1. Let P be an orthogonal projection on $\mathcal{H}$. Then by Theorem 21, we have $\mathcal{H}$ = ran $P \oplus$ ker P. If $x = Py \in$ ran P and $z \in$ ker P, then

   $$\langle x, z\rangle = \langle Py, z\rangle = \langle y, Pz\rangle = 0$$

   So ran $P \perp$ ker P. Hence $\mathcal{H}$ is the orthogonal direct sum of ran P and ker P. It follows that ran P = (ker P)$^{\perp}$, so ran P is closed.

2. Let $\mathcal{M}$ be a closed $\mathcal{H}$. Then $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$. Define a projection $P : \mathcal{H} \to \mathcal{H}$ by $Px = y$ where $x = y + z$ with $y \in \mathcal{M}$ and $z \in \mathcal{M}^{\perp}$.

Then ran $P = \mathcal{M}$ and ker $P = \mathcal{M}^{\perp}$. P is orthogonal. This completes the proof.

If P is an orthogonal projection on $\mathcal{H}$ with range $\mathcal{M}$ and associated orthogonal direct sum $\mathcal{H} = \mathcal{M} \oplus \mathcal{N}$, then $I - P$ is the orthogonal projection with range $\mathcal{N}$ and associated orthogonal direct sum $\mathcal{H} = \mathcal{N} \oplus \mathcal{M}$.

5.4.2 Positive, Normal, and Unitary Operators

Theorem 5.23 (Spectral): Suppose T is a self-adjoint operator on a finite-dimensional complex vector space V with a (Hermitian) inner product $\langle\, ,\rangle$. Then there is an orthonormal basis $\{e_i\}$ for V consisting of eigenvectors for T.

To prove this theorem we need to prove the following:

Lemma 1: Let W be a T-stable subspace of V, with $T = T^*$. Then the orthogonal complement $W^{\perp}$ is also T-stable.

Proof: Let $v \in W^{\perp}$ and $w \in W$. Then,

$$\langle Tv, w\rangle = \langle v, T^*w\rangle = \langle v, Tw\rangle = 0$$

since $Tw \in W$.

Proof of Theorem 5.23: For proving Theorem 5.23, we apply induction on the dimension of V. Let $v \neq 0$ be any vector of length 1 which is an eigenvector for T. We know that T has eigenvectors because $\mathbb{C}$ is algebrically closed and so the minimal polynomial of T factors into linear factors, and V is finite dimensional. Thus $\mathbb{C}\cdot v$ is T-stable and by Lemma 1, the orthogonal complement $(\mathbb{C}\cdot v)^{\perp}$ is also T-stable. With the restriction of the inner product to $(\mathbb{C}\cdot v)^{\perp}$ the restriction of T is still self-adjoint. So by induction on dimension, the theorem is proved.

Theorem 5.24: Suppose T is a normal operator on a finite-dimensional complex vector space V with a Hermitian inner product $\langle\, ,\rangle$. Then there is an orthonormal basis $\{e_i\}$ for V consisting of eigenvectors for T.

Lemma 2: Let T be an operator on V and W a T-stable subspace. Then the orthogonal complement $W^{\perp}$ of W is T^*-stable.

Proof: Let $v \in W^{\perp}$ and w $\in W$. Then,

$$\langle T^*v, w\rangle = \langle v, Tw\rangle = 0 \qquad \text{since } Tw \in W.$$

NOTES

Proof of Theorem 5.24: We will again apply induction on the dimension of V. Let λ be an eigenvalue of T and V_λ the λ-eigenspace of T on V. The assumption of normality is that T and T^* commute. So from the definition of commuting operators, T^* stabilizes V_λ. Then from Lemma 2, $T = T^{**}$ stabilizes $V_\lambda^\perp$. By induction on dimension, the proof is complete.

Corollary 7: Let T be a self-adjoint operator on a finite dimensional complex vector space V with inner product $\langle\ ,\ \rangle$. Let $\{e_i\}$ be an orthonormal basis for V. Then there is a unitary operator k on V, i.e., $\langle kv, kw\rangle = \langle v, w\rangle$ for all $v, w \in V$ such that, $\{ke_i\}$ is an orthonormal basis of T-eigenvectors.

Proof: Let $\{f_i\}$ be an orthonormal basis of T-eigenvectors, whose existence is assured by the spectral theorem. Let k be a linear endomorphism mapping $e_i \to f_i$ for all indices i. We claim that k is unitary. If $\upsilon = \sum_i a_i e_i$ and $w = \sum_j b_j e_i$, then

$$\langle kv, kw\rangle = \sum_{ij} a_i \bar{b}_j \langle ke_i, ke_j\rangle = \sum_{ij} a_i \bar{b}_j \langle f_i, f_j\rangle = \sum_{ij} a_i \bar{b}_j \langle e_i, e_j\rangle = \langle v, w\rangle$$

This is the unitariness and completes the proof.

A self-adjoint operator T on a finite dimensional complex vector space V with Hermitian inner product is positive definite if,

$\langle Tv, v\rangle \geq 0$ with equality only for $v = 0$.

The oeprator T is positive semidefinite if $\langle Tv, v\rangle \geq 0$ i.e., equality may occur for non-zero vectors v.

Lemma 3: The eigenvalues of a positive definite operator T are positive real numbers. When T is just positive semidefinite, the eigenvalue are nonnegative.

Proof: We have by now showed that the eigenvalues of a self-adjoint operator are real. Let v be a nonzero λ-eigenvector for T. Then,

$$\lambda\langle v, v\rangle = \langle Tv, v\rangle > 0$$

by the positive definiteness. Since $\langle v, v\rangle > 0$, essentially $\lambda > 0$. When T is just semidefinite, we get only $\lambda \leq 0$ by this argument.

Corrollary 8: Let $T = T^*$ be positive semidefinite. Then T has a positive semidefinite square root S, i.e., S is self-adjoint positive semidefinite and

$$S^2 = T$$

If T is positive definite, then S is positive definite.

Proof: From the spectral theorem, there is an orthonormal basis $\{e_i\}$ for V consisting of eigenvectors, with respective eigenvalues $\lambda_i \geq 0$. Define an operator S by,

$$Se_i = \sqrt{\lambda_i} \cdot e_i$$

Noticeably S has the same eigenvectors as T with eigenvalues the nonnegative real square roots of those of T and the square of this operator is T. Now, let $\upsilon = \sum_i a_i e_i$ and $w = \sum_i b_i e_i$ and compute

$$\langle S^* v, w\rangle = \langle v, Sw\rangle = \sum_{ij} a_i \bar{b}_j \langle e_i, e_j\rangle = \sum_{ij} a_i \bar{b}_j \sqrt{\lambda_j}\langle e_i, e_j\rangle = \sum_i a_i \bar{b}_i \sqrt{\lambda_i}\langle e_i, e_i\rangle$$

by orthonormality and the realness of $\sqrt{\lambda_i}$. We therefore get,

$$\sum_{ij} a_i \bar{b}_j \langle \sqrt{\lambda_i} e_i, e_j\rangle = \langle Sv, w\rangle$$

Since the adjoint is unique, $S = S^*$. This completes the proof.

The standard (Hermitian) inner product on $\mathbb{C}^n$ is,

$$\langle (v_1, \ldots, v_n), (w_1, \ldots, w_n)\rangle = \sum_{i=1}^{n} v_i \bar{w}_j$$

The n-by-n complex matrices give $\mathbb{C}$ linear endomorphisms by left multiplication of column vectors. With this inner product, the adjoint of an endomorphism T is, $T^* = T$, i.e., conjugate transpose. Certainly, we often write the superscript '*' to indicate conjugate transpose of a matrix and say that the matrix T is Hermitian. Similarly, an n-by-n matrix k is unitary if, $kk^* = 1_n$ where 1_n is the n-by-n identity matrix. This is equivalent to unitariness with respect to the standard Hermitian inner product.

Corollary 9: Suppose T is an Hermitian matrix. Then there is a unitary matrix k such that k^*Tk = Diagonal, with diagonal entries the eigenvalues of T.

Proof: Suppose $\{e_i\}$ is the standard basis for $\mathbb{C}^n$. It is orthonormal with respect to the standard inner product. Let $\{f_i\}$ be an orthonormal basis consisting of T-eigenvectors. From Corollary 7, let k be the unitary operator mapping e_i to f_i. Then k^*Tk is diagonal, with diagonal entries the eigenvalues.'

Corollary 10: Let T be a positive semidefinite Hermitian matrix. Then there is a positive semidefinite Hermitian matrix S such that,

$$S^2 = T$$

Proof: T is positive semidefinite self-adjoint with respect to the standard inner product. So $S^2 = T$.

5.5 COMPLETE ORTHOGONAL SETS

Definition: Orthogonal set

A set of functions

$$\{\phi_1(x), \phi_2(x), \cdots, \phi_n(x), \cdots\}$$

is called orthogonal if every function is orthogonal to every other function i.e. if

$$\int_a^b \phi_i(x)\phi_k(x)dx = 0 \quad \text{for i} \neq \text{k}$$

NOTES

Examples of orthogonal sets

1. The set of trigonometric functions

$$\{1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots\}$$

- on the interval $(-\pi, \pi)$. This set is one of the first and most important examples of orthogonal sets.

2. The set of Legendre polynomials

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n(x^2-1)^n}{dx^n}$$

- on the interval $(-1,1)$. The first few polynomials of the sequence are

$$P_0(x) = 1$$
$$P_1(x) = x$$
$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$
$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$

Definition: Orthonormal set

A collection of functions

$$\{\phi_1(x), \phi_2(x), \cdots, \phi_n(x), \cdots\}$$

is called orthonormal if it is orthogonal and if each of the functions is of unit length i.e.

$$\int_a^b \phi_i(x) = 1 \quad i = 1,2, \ldots$$

Example

The set of trigonometric functions

$$\{\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \cdots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \cdots\}$$

on the interval $(-\pi, \pi)$ is orthonormal. These functions are obtained by dividing the functions

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots$$

by their lengths.

Expansion by orthogonal systems of functions

Let e_1 and e_2 be any two mutually perpendicular vectors of unit length in a plane. Then any vector f in that plane can decomposed in the direction of these two vectors and written as

$$f = \mathrm{a}_1\mathbf{e}_1 + \mathrm{a}_2\mathbf{e}_2$$

NOTES

where $a_1 = f \cdot e_1$ and $a_2 = f \cdot e_2$. Similarly if e_1, e_2 and e_3 are any three mutually perpendicular unit vectors in a three-dimensional space, then any vector f in the space can be decomposed in the direction of these three vectors and written as

$$f = a_1 e_1 + a_2 e_2 + a_3 e_3$$

where $a_1 = f \cdot e_1, a_2 = f \cdot e_2$ and $a_3 = f \cdot e_3$. Likewise it is possible to represent any function f in Hilbert space as a linear combination of an orthonormal set of functions. For this, it is necessary for the orthonormal system to be complete.

Definition: Complete orthogonal set

An orthogonal set of functions is called complete if it is impossible to add to it even one function, not identically equal to zero, that is orthogonal to all the functions of the set.

We can easily give an example of an orthogonal set that is not complete. Suppose we are given any arbitrary orthogonal set and remove a single function from it, the remaining set will be incomplete. For example, if we remove $\cos x$ from orthogonal set

$$\{1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots\}$$

The remaining set

$$\{1, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots\}$$

is orthogonal as before, but it is not complete since the function $\cos x$ which we excluded is orthogonal to all functions of the set.

5.5.1 Parseval's Identity

The square of the length of a vector in Hilbert space is equal to the sum of the squares of its projections onto a complete set of mutually orthogonal directions. In other words, if

$$\{\phi_1(x), \phi_2(x), \cdots, \phi_n(x), \cdots\}$$

is a complete orthonormal set of functions in Hilbert space and if a function f is given by

$$f(x) = a_1\phi_1(x) + a_2\phi_2(x) + \cdots + a_n\phi_n(x) + \cdots$$

then

$$\sum_{i=1}^{\infty} a_i^2 = \int_a^b f^2(x)dx$$

NOTES

Proof: Denote by $r_n(x)$ the difference between $f(x)$ and the sum of the first n terms of its series representation i.e.

$$r_n(x) = f(x) - [a_1\phi_1(x) + a_2\phi_2(x) + \cdots + a_n\phi_n(x)]$$

Now the function $r_n(x)$ is orthogonal to each of the functions $\phi_1(x), \phi_2(x), \cdots, \phi_n(x), \cdots$. Let us show that it is orthogonal to the function $\phi_1(x)$, i.e.,

$$\int_a^b r_n(x)\phi_1(x)dx = 0.$$

We have

$$\int_a^b r_n(x)\phi_1(x)dx = \int_a^b [f(x) - a_1\phi_1(x) - a_2\phi_2(x) - \cdots - a_n\phi_n(x)]\phi_1(x)dx$$

or

$$\int_a^b r_n(x)\phi_1(x)dx = \int_a^b f(x)\phi_1(x)dx - a_1\int_a^b \phi_1^2(x)dx$$

where we employ the fact that, because the functions are orthogonal to each other,

$$\int_a^b \phi_i(x)\phi_j(x)dx = 0 \quad \text{for } i \neq j$$

Now in (1)

$$\int_a^b f(x)\phi_1(x)dx = a_1 \quad \text{and} \quad \int_a^b \phi_1^2(x)dx = 1$$

and thus (1) becomes

$$\int_a^b r_n(x)\phi_1(x)dx = \int_a^b f(x)\phi_1(x)dx - a_1\int_a^b \phi_1^2(x)dx = a_1 - a_1 = 0$$

Hence, in the equation

$$f(x) = a_1\phi_1(x) + a_2\phi_2(x) + \cdots + a_n\phi_n(x) + r_n(x)$$

the terms on the right side are all orthogonal to each other. Now, by the Pythagorean theorem, the square of the length of $f(x)$ is equal to the sum of the square of the summands on the right side of (2), i.e.

$$\int_a^b f^2(x)dx = \int_a^b [a_1\phi_1(x)]^2 dx + \cdots + \int_a^b [a_n\phi_n(x)]^2 dx + \int_a^b r_n^2(x)dx$$

Since the set of functions $\phi_1, \phi_2, \cdots, \phi_n$ is normalized, we have

$$\int_a^b f^2(x)dx = a_1^2 + a_2^2 + \cdots + a_n^2 + \int_a^b r_n^2(x)dx$$

NOTES

Now we are dealing with functions of a Hilbert space which consists of the functions f for which the Lebesgue integral of $|f^2|$ exists. This means that as n approaches ∞, the integral

$$\int_a^b f^2(x)dx$$

converges and the term

$$\int_a^b r_n^2(x)dx$$

on the right side of (3) approaches zero. Thus (3) becomes

$$\sum_{i=1}^{\infty} a_i^2 = \int_a^b f^2(x)dx$$

Check Your Progress

4. What is self-adjoint operator?
5. Define a projection.
6. Give the statement of spectral theorem.

5.6 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Let V be an inner product space. Let $v \in V$. Then norm of v (or length of v) is defined as and is denoted by $\| v \|$.
2. Let X be a Hilbert space and $\ell \in X'$, then there is $y_0 \in X$ such that,

   $\ell(x) = (x, y_0)$ for $x \in X$

   Furthermore, the mapping $\ell \mapsto y_0$ is conjugate linear and $\| \ell \| = \| y_0 \|$.
3. Every Hilbert space is reflexive.
4. A bounded linear operator $T : H \to H$ on a Hilbert space H is said to be self-adjoint or Hermitian if,

   $T^* = T$

   Equivalently, a bounded linear operator T is said to be self-adjoint if,

   $\langle x, Ty \rangle = \langle Tx, y \rangle$ for all $x, y \in H$
5. A projection on a linear space X is a linear map $P: X \to X$ such that, $P^2 = P$.
6. Suppose T is a self-adjoint operator on a finite-dimensional complex vector space V with a (Hermitian) inner product. Then there is an orthonormal basis $\{ei\}$ for V consisting of eigenvectors for T.

5.7 SUMMARY

NOTES

- Let V be a vector space over field F (where F = field of real or complex numbers). Suppose for any two vectors $u, v \in V \exists$ an element $(u, v) \in F$ such that, [(u, v) here is just an element of F and should not be confused with the ordered pair.]
- Inner product space over real field is called *Euclidean space* and over complex field is called *Unitary space*.
- Let V be an inner product space. Let v ∈ V. Then norm of v (or length of v) is defined as and is denoted by $\| v \|$.
- Let X be a Hilbert space and $\ell \in X'$, then there is $y_0 \in X$ such that,

  $\ell(x) = (x, y_0)$ for $x \in X$

  Furthermore, the mapping $\ell \mapsto y_0$ is conjugate linear and $\| \ell \| = \| y_0 \|$.
- Let (Ω, Σ, μ) and (Ω, Σ, ν) be measure spaces with $\mu(\Omega) < +\infty$ and $\nu(\Omega) < +\infty$ where ν is absolutely continuous with respect to μ.
- Lebesgue-Nikodym theorem holds if both (Ω, Σ, μ) and (Ω, Σ, ν) are σ-finite. But in this case h may not be μ-integrable.
- Let X be a Hilbert space and B a bounded, positive definite and sesquilinear functional on $X \times X$. Then there is a unique bounded linear operator $S : X \mapsto X$ such that $(x, y) = B(Sx, y)$ for all $x, y - X$ and $\|S\| \leq \rho^{-1}$. Besides S^{-1} exists and is bounded with $\|S^{-1}\| \leq r$.
- A set $\{u_i\}_i$ of vectors in an inner product space V is said to be orthogonal if $(u_i, u_j) = 0$ for $i \neq j$. If further $(u_i, u_i) = 1$ for all i, then the set $\{u_i\}$ is called an *orthonormal set*.
- An orthonormal set in an inner product space is linearly independent.
- Consider a Hilbert space H and its conjugate space H*. Let y be a fixed vector in H. Define a function fy on H by, fy(x) = (x, y) for all x in H. Then the mapping y → fy is a norm preserving mapping of H into H*.
- Every Hilbert space is isometrically equivalent to its dual.
- Every Hilbert space is reflexive.
- For every Hilbert space $\mathcal{H}$ there exists a surjective isometry $\Psi: \mathcal{H}^* \rightarrow \mathcal{H}$ of the dual $\mathcal{H}^*$ of $\mathcal{H}$ onto $\mathcal{H}$ which is additive and conjugate homogeneous.
- A bounded linear operator T : H → H on a Hilbert space H is said to be self-adjoint or Hermitian if,

  $T^* = T$

  Equivalently, a bounded linear operator T is said to be self-adjoint if,

  $\langle x, Ty \rangle = \langle Tx, y \rangle$ for all $x, y \in$ H
- A bounded linear operator T: H → H on a Hilbert space H is said to be self-adjoint or Hermitian if, T* = T.

NOTES

- A bounded linear operator T: H $\rightarrow$ H on a Hilbert space *H* is said to be unitary if *T* is bijective and TT* = T*T.
- A bounded linear operator T: H $\rightarrow$ H on a Hilbert space H is said to be normal if, TT* = T*T.
- A projection on a linear space *X* is a linear map $P: X \rightarrow X$ such that, $P^2 = P$.
- An orthogonal projection on a Hilbert space $\mathcal{H}$ is a linear map $P: \mathcal{H} \rightarrow \mathcal{H}$ that satisfies, $P^2 = P$, for all $x, y \in \mathcal{H}$
- Suppose *T* is a self-adjoint operator on a finite-dimensional complex vector space *V* with a (Hermitian) inner product. Then there is an orthonormal basis $\{e_i\}$ for *V* consisting of eigenvectors for *T*.

5.8 KEY TERMS

- **Unitary space:** Inner product space over real field is called *Euclidean space* and over complex field is called *Unitary space*.
- **Norm:** Let *V* be an inner product space. Let $v \in V$. Then *norm* of *v* (or length of *v*) is defined as $\sqrt{(v, v)}$ and is denoted by $\| v \|$.
- **Orthogonal vectors:** Let *V* be an inner product space. Two vectors $u, v \in V$ are said to be orthogonal if $(u, v) = 0 \Leftrightarrow (v, u) = 0$.
- **Orthogonal set:** A set $\{ui\}i$ of vectors in an inner product space *V* is said to be orthogonal if $(ui, uj) = 0$ for $i \neq j$. If further $(ui, ui) = 1$ for all *i*, then the set $\{ui\}$ is called an orthonormal set.
- **Self-adjoint operator:** A bounded linear operator T: H $\rightarrow$ H on a Hilbert space *H* is said to be self-adjoint or Hermitian if, T* = T.

5.9 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define inner product space.
2. Define the term unitary space.
3. What is a linear transformation?
4. Define an orthonormal basis.
5. State Gram-Schmidt orthogonalization process.
6. What do you understand by conjugate space?
7. When is a space said to be reflexive?
8. What do you mean by kernel of an orthogonal projection?
9. Define Hermitian inner product.
10. What is Parseval's identity?

Long-Answer Questions

1. Illustrate the properties of inner product spaces.
2. Prove Riesz representation theorem.
3. Explain and prove Lebesgue-Nikodym theorem.
4. Describe all the operators on the Hilbert space.
5. Discuss the concept of conjugate space.
6. Show that every Hilbert space is reflexive.
7. Let P be a nonzero orthogonal projection. Then prove that $\|P\| = 1$.
8. Describe and prove spectral theorem.
9. Explain orthonormal and complete orthonormal sets.

5.10 FURTHER READING

Yang, Bicheng and Qiang Chen. 2016. *Integration Theory and Functional Analysis*. New York: Magnum Publishing LLC.

Kadets, Vladimir. 2018. *A Course in Functional Analysis and Measure Theory*. New York: Springer International Publishing.

Pundir, R. and S. K. Pundir. 2015. *Integration Theory and Functional Analysis*. Meerut (Uttar Pradesh): Pragati Prakashan.

Rudin, Walter. 1986. *Real and Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Ahlfors, Lars V. 1978. *Complex Analysis*, 3rd Edition. London: McGraw-Hill Education – Europe.

Lang, Serge. 1998. *Complex Analysis*, 4th Edition. New York: Springer-Verlag New York Inc.

Shastri, Anant R. 2000. *An Introduction to Complex Analysis*. India: Macmillan Publishers India Ltd.

Gupta, S. L. and Nisha Rani. 2003. *Fundamental Real Analysis*, 4th Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, 1st Edition. United Kingdom: Cambridge University Press.

Basener, William F. *Topology and Its Applications*. New Jersey: John Wiley & Sons, Inc.

Bourbaki, Nicolas. 1966. *Elements of Mathematics: General Topology*, Part 1. United States: Addison-Wesley.

NOTES

NOTES