Advanced Mathematical Techniques in Chemical Engineering

(CH61015)

Course notes and practice problems on Linear Algebra

P. A. Deshpande

Department of Chemical Engineering

Indian Institute of Technology Kharagpur

parag@che.iitkgp.ac.in

Contents

Chapter 0

Background: A review of set theory

A set is a collection of well-defined objects. The objects present in the set are called its members or elements. We will generally represent a set by a capital letter and its members by small letters.

$$A = \{a, e, i, o, u\} \tag{1}$$

A is the set and a, e, i, o, u are its elements. To indicate that a is an element of A, the symbol $\in$ is used.

$$a \in A \tag{2}$$

The above is read as "a belongs to A" or "a is in A" or "a is an element of A".

The number of elements present in a set is called its cardinality or cardinal number. A set is said to be a finite set if its cardinality is finite. Else, the set is said to be an infinite set. A set is said to be an empty set or a null set or a void set if it does not contain any element. Such a set is generally denoted by ϕ. A few examples of empty sets are given below.

$$\phi = \{\} \tag{3}$$

$$\phi = \{x : x \in \mathbb{R} \wedge x^2 = -1\} \tag{4}$$

A set containing exactly one element is called a singleton set. Two sets A and B are said to be equal if they have same elements.

$$A = B \text{ iff } x \in A \implies x \in B \wedge x \in B \implies x \in A \tag{5}$$

Two sets are said to be equivalent if they have the same number of elements in them i.e. their cardinalities are equal.

$$A \sim B \text{ iff } |A| = |B| \tag{6}$$

B is said to be a subset of A if all elements of B are also elements of A.

$$B \subset A \text{ iff } x \in B \implies x \in A \; \forall x \in B \tag{7}$$

A set of cardinality $|B|$ has $2^{|B|}$ number of subsets. Two sets A and B are said to be comparable if either $A \subset B$ or $B \subset A$.
Power set of a set A is a set of all possible subsets of A.

$$\mathcal{P}(A) = \{S : S \subseteq A\} \tag{8}$$

The union of two sets A and B is a set which has elements which are either in A or in B or in both.

$$A \cup B = \{x : x \in A \vee x \in B \vee x \in A \wedge B\} \tag{9}$$

The intersection of two sets A and B is a set which has elements which are common to both A and B.

$$A \cap B = \{x : x \in A \wedge x \in B\} \tag{10}$$

Laws of sets
Idempotent laws:

$$A \cup A = A \tag{11}$$

$$A \cap A = A \tag{12}$$

Identity laws:

$$A \cup \phi = A \tag{13}$$

$$A \cap \phi = \phi \tag{14}$$

Commutative laws:

$$A \cup B = B \cup A \tag{15}$$

$$A \cap B = B \cap A \tag{16}$$

Associative laws:

$$A \cup (B \cup C) = (A \cup B) \cup C \tag{17}$$

$$A \cap (B \cap C) = (A \cap B) \cap C \tag{18}$$

Distributive laws:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \tag{19}$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{20}$$

Complement laws:

$$A \cap A' = \phi \tag{21}$$

$$A \cup A' = U \tag{22}$$

$$\phi' = U \tag{23}$$

$$U' = \phi \tag{24}$$

where U is the universal set.

Difference of two sets A and B is defined as

$$A - B = \{x : x \in A \wedge x \notin B\} \tag{25}$$

Symmetric difference of A and B is defined as

$$A \Delta B = (A - B) \cup (B - A) \tag{26}$$

Ordered pair: In any set, the order in which the elements are written is immaterial i.e.

$$A = \{a_1, a_2, a_3\} = \{a_1, a_3, a_2\} = \{a_3, a_2, a_1\} \ldots \tag{27}$$

An ordered pair is a collection of two elements described in a particular order i.e.

$$(a, b) = (b, a) \text{ iff } a = b \tag{28}$$

Two ordered pairs (a, b) and (c, d) are equal iff $a = c$ and $b = d$. As an extension to an ordered collection of n-elements, we can define ordered n-tuple.

Cartesian product of sets: Let A and B be two sets. The set of all ordered pairs (a, b) such that $a \in A$ and $b \in B$ is called the cartesian product of the two sets. It is denoted by $A \times B$.

$$A \times B = \{(a, b) : a \in A \land b \in B\} \tag{29}$$

$$A \times B \neq B \times A \text{ unless } A = B \tag{30}$$

As an extension, we can write

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2 \ldots a_n) : a_i \in A_i \forall i\} \tag{31}$$

Relations:

A relation R from a set A to B is a subset of $A \times B$.

$$R \subset A \times B \tag{32}$$

The domain of a relation R from A to B is the set of all first elements of the ordered pairs which belong to R.

$$dom(R) = \{a : a \in A \land (a, b) \in R\} \tag{33}$$

The range of a relation R from A to B is the set of all second elements of the ordered pairs which belong to R.

$$range(R) = \{b : b \in B \land (a, b) \in R\} \tag{34}$$

Consider $A = \{1, 2, 3\}$ and $B = \{1, 2, 4\}$.

$$A \times B = \{(1, 1), (1, 2), (1, 4), (2, 1), (2, 2), (2, 4), (3, 1), (3, 2), (3, 4)\} \tag{35}$$

We choose a subset of $A \times B$ such that the sum of the elements of the ordered pairs is greater than or equal to 4. Then

$$R = \{(1, 4), (2, 2), (2, 4), (3, 1), (3, 2), (3, 4)\} \tag{36}$$

It can be seen from Figure 1 that there are more than one arrows emerging from A. Hence, in relations, it is allowed to have more than one ordered pairs with the same first entry i.e. more than one elements in A can have a corresponding image in B. Now we define a special case of relations called a function.

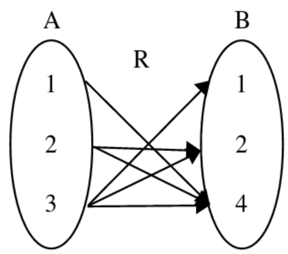


Figure 1: Representation of a relation given by Eq. (36)

Functions:

A function f from A to B, denoted as $f : A \to B$, is a subset of $A \times B$ such that

(i) every element of A is the first entry of some ordered pair, and

(ii) no two ordered pairs have the same first entry.

The set A is called the domain of f while the set B is called the co-domain of f. Different cases given in Figure 2 can be used to identify functions.

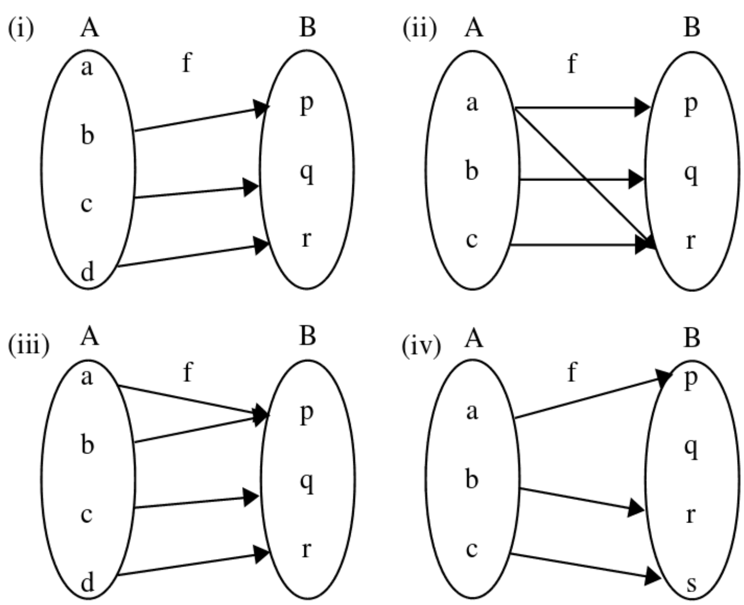


Figure 2: Different scenarios depicting a relation to be a function. (i) not a function, (ii) not a function, (iii) function, (iv) function

Into function: There is atleast one element in B such that there is no pre-image corresponding to it in A. q in Figure 3(i), for example, does not have any pre-image.

Onto or surjective function: Every element in B has a pre-image in A. In this case, $co - domain(f) = range(f)$ as can be seen from Figure 3(ii).

One-One or injective function: For every image, there is exactly one pre-image. But all images need not have a pre-image. See Figure 3(iii) for example.

One-one onto or bijective function: This function is both one-one, i.e. all the images have exactly one corresponding pre-image, and onto, i.e. every image has a pre-image. See Figure 3(iv) for example.

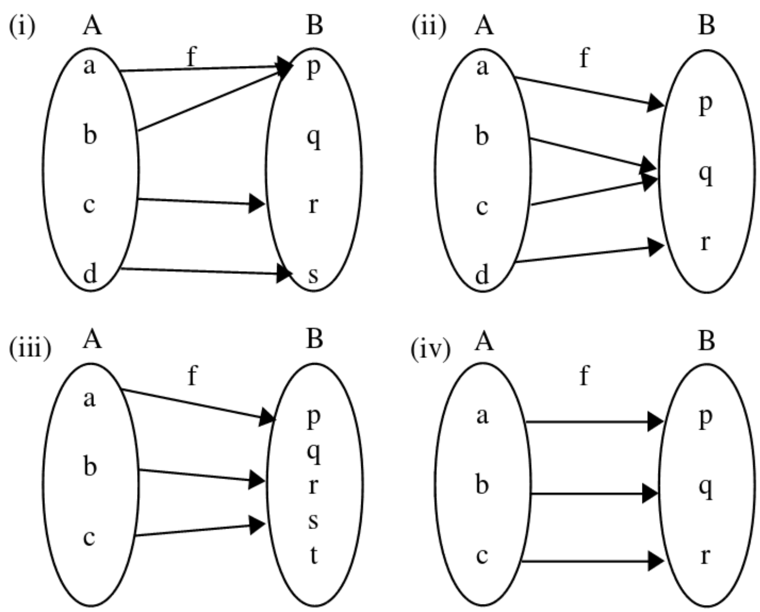


Figure 3: Different types of functions. (i) into function, (ii) onto or surjective function, (iii) one-one or injective function, (iv) one-one onto or bijective function

Problems

1. Write the following sets using set builder notation:

   (i) $A = \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4} \ldots\}$

   (ii) $A = \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}\}$

   (iii) $A = \{0, 0, 7, 26, 63 \ldots\}$

   (iv) $A = \{0, 2, 4, 6 \ldots\}$

   (v) $A =$ a set of all possible integers whose cube is an odd integer

   (vi) $A =$ a set of all real numbers which cannot be written as quotients of two integers

2. Show that a null set is unique.

3. Prove that

   (i) every set is a subset of itself

   (ii) ϕ is a subset of all sets

   (iii) number of subsets of a given set equals 2^N where N is the cardinality of the set

   (iv) number of proper subsets of a given set equals 2^N-2

4. If U is the universal set then identify the complement of A if

   $A = \{x : x \in \mathbb{N} \wedge x = 3n \; \forall \; n \in \mathbb{N}\}$.

5. Prove that the cardinality of the power set of A, $P(A)$, is $2^{|A|}$.

6. If $A = \{x : x = 4n + 1,\, n \in \mathbb{N},\, n \leq 5\}$ and

   $B = \{x : x = 3n,\, n \in \mathbb{N},\, n \leq 8\}$ then determine $A \Delta B$.

7. Prove the following set operations:

   (i) Idempotent laws:

   (a) $A \cup A = A$ (b) $A \cap A = A$

   (ii) Identity laws:

   (a) $A \cup \phi = A$ (b) $A \cap U = A$

   (iii) Commutative laws:

   (a) $A \cup B = B \cup A$ (b) $A \cap B = B \cap A$

   (iv) Associative laws:

(a) $A \cup (B \cup C) = (A \cup B) \cup C$ (b) $A \cap (B \cap C) = (A \cap B) \cap C$

(v) Distributive laws:

(a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

(b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(vi) De Morgan's laws:

(a) $(A \cup B)^c = A^c \cap B^c$ (b) $(A \cap B)^c = A^c \cup B^c$

(vii) Complement laws:

(a) $A \cap A^c = \phi$ (b) $A \cup A^c = U$

(c) $\phi^c = \mathrm{U}$ (d) $\mathrm{U}^c = \phi$

(viii) Involution law:

$(A^c)^c = A$

8. The Cartesian product $A \times A$ has 9 elements with two of the elements being (-1,0) and (0,1). Identify the set A and determine $A \times A$.

9. Prove that $|A \times B| = |A| \cdot |B|$.

10. Check whether $A \times B$ and $B \times A$ form equivalent sets.

11. For two sets $A = \{a_i\}, i = 1 \cdots n$ and $B = \{b_i\}, i = 1 \cdots n, n \in \mathbb{Z}$, verify that $< A, B > \subset \mathcal{PP}(A \cup B)$.

12. If a and b are elements of sets A and B, respectively, then verify the followings:

    (a) $a \cup b \subset (\cup A) \cup (\cup B)$

    (b) $< a, b > \subset \mathcal{PP}((\cup A) \cup (\cup B))$

    (c) $< x, y >, x \in A, y \in B \subset \mathcal{PPP}((\cup A) \cup (\cup B))$

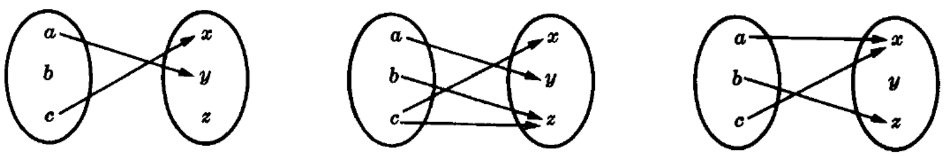
13. Prove the following set operations:

    (a) $A \times (B \cup C) = (A \times B) \cup (A \times C)$

    (b) $A \times (B \cap C) = (A \times B) \cap (A \times C)$

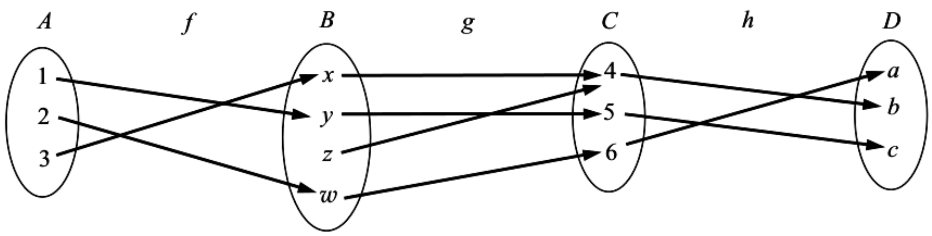
    (c) $A \times (B - C) = (A \times B) - (A \times C)$

14. State whether each diagram in the following figures defines a mapping from $A = \{a, b, c\}$ into $B = \{x, y, z\}$.

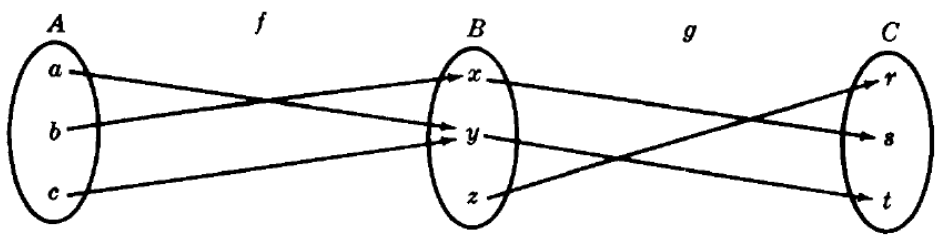


(a) (b) (c)

15. Let the mappings $f : A \to B$, $g : B \to C$, $h : C \to D$ be defined by the following diagram. Determine whether or not each function is (a) injective, (b) surjective, (c) bijective, (d) invertible.



16. Let $f : A \to B$ and $g : B \to C$ be defined by the diagram.



(a) Find the composition mapping $(g \circ f) : A \to C$.

(b) Find the images of the mappings f, g, $g \circ f$.

17. Consider the mapping $F : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $F(x, y, z) = (yz, x^2)$. Find (a) $F(2, 3, 4)$; (b) $F(5, -2, 7)$; (c) $F^{-1}(0, 0)$ that is, all $v \in \mathbb{R}^3$ such that $F(v) = 0$.

18. Consider the mapping $F : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $F(x, y) = (3y, 2x)$. Let S be the unit circle in $\mathbb{R}^2$, that is, the solution set of $x^2 + y^2 = 1$. (a) Describe $F(S)$. (b) Find $F^{-1}(S)$.

19. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 2x - 3$. If f is one-to-one and onto, find a formula for f^{-1}.

20. Suppose $f : A \to B$ and $g : B \to C$. Hence, $g \circ f : A \to C$ exists. Prove that

    (a) If f and g are one-to-one, then $g \circ f$ is one-to-one.

    (b) If f and g are onto mappings, then $g \circ f$ is an onto mapping.

    (c) If $g \circ f$ is one-to-one, then f is one-to-one mapping.

    (d) If $g \circ f$ is an onto mapping, then g is an onto mapping.

Chapter 1

Groups, rings and fields

Binary operation:

Let S be any non-empty set. A function $f : S \times S \to S$ is called a binary operation on the set S. The binary operation f on the set S associates every ordered pair $(a, b) \in S \times S$ to a unique element $f(a, b) \in S$.

From the above, it can be easily seen that addition is a binary operation on $\mathbb{N}$. Similarly, multiplication is a binary operation on $\mathbb{N}$. But subtration and division are not binary operations on $\mathbb{N}$.

Commutative binary operation:

A binary operation $\otimes$ on a set S is said to be commutative if

$$a \otimes b = b \otimes a \ \forall a, b \in S \tag{1.1}$$

Hence, multiplication of numbers is a commutative binary operation while multiplication of matrices is not.

Associative binary operation:

A binary operation $\otimes$ is said to be associative if

$$(a \otimes b) \otimes c = a \otimes (b \otimes c) \ \forall a, b, c \in S \tag{1.2}$$

Hence, multiplication and addition of numbers and matrices are associative binary operations.

Distributivity:

Let two operations $\otimes$ and $\odot$ be defined on S. The binary operation $\otimes$ is said to be

(a) left distributive over $\odot$ if

$$a \otimes (b \odot c) = (a \otimes b) \odot (a \otimes c) \ \forall a, b, c \in S \tag{1.3}$$

(b) right distributive over $\odot$ if

$$(b \odot c) \otimes a = (b \otimes a) \odot (c \otimes a) \ \forall a, b, c \in S \tag{1.4}$$

The binary operation $\otimes$ is said to be distributive over $\odot$ if it is both left as well as right distributive.

Closure property:

Let $\otimes$ be a binary operation on S. For $T \subseteq S$, T is said to be closed under $\otimes$ if $a \otimes b \in T \ \forall a, b \in T$.

Identity element:

Let $\otimes$ be a binary operation on S. An element $e_1 \in S$ is called a left identity if

$$e_1 \otimes a = a \ \forall a \in S \tag{1.5}$$

An element $e_2 \in S$ is called a right identity if

$$a \otimes e_2 = a \ \forall a \in S \tag{1.6}$$

An element $e \in S$ is called an identity element if it is both left and right identity i.e.

$$e \otimes a = a = a \otimes e \ \forall a \in S \tag{1.7}$$

Inverse of an element:

Let $\otimes$ be a binary operation on S. An element $b_1 \in S$ is called the left inverse of an element $a \in S$ if

$$b_1 \otimes a = e \tag{1.8}$$

where e is the identity element in S. An element $b_2 \in S$ is called the right inverse of an element $a \in S$ if

$$a \otimes b_2 = e \tag{1.9}$$

An element $b \in S$ is called an inverse element of $a \in S$ if it is both left as well as right inverse i.e.

$$b \otimes a = e = a \otimes b \tag{1.10}$$

Groups:

A group is a non-empty set G equipped with a binary operation $\otimes : G \times G \to G$ that associates an element $a \otimes b \in G\ \forall a, b \in G$ and having the following properties.

$$a \otimes (b \otimes c) = (a \otimes b) \otimes c \tag{1.11}$$

$$a \otimes e = e \otimes a = a \text{ i.e. } \exists e \in G \tag{1.12}$$

$$\forall a \in G \exists a^{-1} \in G \text{ such that } a \otimes a^{-1} = a^{-1} \otimes a = e \tag{1.13}$$

The group is called abelian if it is commutative i.e.

$$a \otimes b = b \otimes a\ \forall a, b \in G \tag{1.14}$$

A non-empty group M with a binary operation $\otimes : M \times M \to M$ is called a monoid if it satisfies only the first two conditions for a group.

The identity element of a group is unique. To prove this, we assume there be two identity elements of G.

$$e^I \otimes a = a\ \forall a \in G \tag{1.15}$$

$$a \otimes e^{II} = a\ \forall a \in G \tag{1.16}$$

$e^I, e^{II} \in G$. Therefore, for $a = e^{II}$,

$$e^I \otimes e^{II} = e^{II} \tag{1.17}$$

Similarly, for $a = e^I$,

$$e^I \otimes e^{II} = e^I \tag{1.18}$$

Hence, $e^I = e^{II}$. Therefore, a group has a unique identity element.

Every element in a group has a unique inverse element. To prove this, we assume two inverse elements of $a \in G$, $a^I, a^{II} \in G$.

$$a^I \otimes a = e \tag{1.19}$$

$$a \otimes a^{II} = e \tag{1.20}$$

$$(a^I \otimes a) \otimes a^{II} = e \otimes a^{II} = a^{II} \tag{1.21}$$

$$a^I \otimes (a \otimes a^{II}) = a^I \otimes e = a^I \tag{1.22}$$

Since $\otimes$ is associative,

$$(a^I \otimes a) \otimes a^{II} = a^I \otimes (a \otimes a^{II}) \tag{1.23}$$

$$\implies a^{II} = a^I \tag{1.24}$$

Subgroups:

Given a non-empty set G, a set $H \subseteq G$ is called a subgroup of G iff

$$\text{the identity element of } e \in G \text{ also belongs to H} \tag{1.25}$$

$$h_1 \otimes h_2 \in H \ \forall h_1, h_2 \in H \tag{1.26}$$

$$h^{-1} \in H \ \forall h \in H \tag{1.27}$$

Rings:

A ring is a non-empty set R equipped with two binary operations

$$+ : R \times R \to R \text{ (addition)} \tag{1.28}$$

$$* : R \times R \to R \text{ (multiplication)} \tag{1.29}$$

having the following properties.

(a) R is an abelian group with respect to $+$

(b) * is associative and has an identity element $1 \in R$

(c) * is distributive with respect to $+$

The identity element for addition is denoted by 0 (zero) and the additive inverse of $a \in R$ is denoted by $-a$. $\forall a, b, c \in R$, the following properties hold.

$$(a+b)+c = a+(b+c) \tag{1.30}$$

$$a+b = b+a \tag{1.31}$$

$$a+0 = 0+a = a \tag{1.32}$$

$$a+(-a) = (-a)+a = 0 \tag{1.33}$$

$$a*(b*c) = (a*b)*c \tag{1.34}$$

$$a*1 = 1*a = a \tag{1.35}$$

$$(a+b)*c = (a*c)+(b*c) \tag{1.36}$$

$$a*(b+c) = (a*b)+(a*c) \tag{1.37}$$

Fields:

A set K is a field if it is a ring and the following properties hold.

(a) 0≠1

(b) $K^* = K - \{0\}$ is a group i.e. every $a \neq 0$ has an inverse with respect to *

(c) * is commutative

Problems

1. Verify whether the set of integers $\mathbb{Z}$ is a group under addition? Identify whether $\mathbb{Z}^* = \mathbb{Z} - \{0\}$ is a group under multiplication?

2. Verify whether the set $\mathbb{Q}$ of rational numbers is a group under addition? Identify whether $\mathbb{Q}^* = \mathbb{Q} - \{0\}$ is a group under multiplication?

3. Verify whether a set of $n \times n$ invertible matrices is a group under (a) matrix addition and (b) matrix multiplication?

4. Identify whether the following operations make a binary operation on the given sets.

   (a) addition on $\mathbb{N}, \mathbb{Z}, \mathbb{I}^+, \mathbb{I}^-, \mathbb{Q}, \mathbb{R}, \mathbb{C}$

   (b) subtration on $\mathbb{N}, \mathbb{Z}, \mathbb{I}^+, \mathbb{I}^-, \mathbb{Q}, \mathbb{R}, \mathbb{C}$

   (c) multiplication on $\mathbb{N}, \mathbb{Z}, \mathbb{I}^+, \mathbb{I}^-, \mathbb{Q}, \mathbb{R}, \mathbb{C}$

   (d) division on $\mathbb{N}, \mathbb{Z}, \mathbb{I}^+, \mathbb{I}^-, \mathbb{Q}, \mathbb{R}, \mathbb{C}$

5. For each of the above operations or combination of operations, determine whether the operations are (a) commutative, (b) associative, (c) distributive.

6. For $\otimes : A \times A \to A$, is A closed under $\otimes$?

7. Verify whether a universal set U is closed under every possible binary operation $\otimes$ defined on U?

8. Verify whether $\mathbb{N}$ is a group under (a) addition, (b) multiplication? If yes, then check if the group is abelian?

9. Verify whether $\mathbb{N}$ is a monoid under (a) addition, (b) multiplication?

10. Verify whether a set of non-singular matrices $(n \times n)$ forms a group under (a) matrix addition, (b) matrix multiplication? Are the identity elements for the two operations the same? Is the group abelian under both the operations?

11. Verify whether the additive groups $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are commutative rings?

12. Verify whether the rings $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are fields?

13. Verify whether the group $\mathbb{R}[X]$ of polynomials in one variable with real coefficients is a ring under multiplication of polynomials? Check whether it is commutative?

14. Verify whether a group of $n \times n$ matrices is a ring under matrix multiplication? Check whether it is commutative?

15. Verify whether a set of fractions of polynomials $f(x)/g(x), f(x), g(x) \in \mathbb{R}(x), g(x) \neq 0$ is a field?

16. Let G be a group with elements of the form $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $a, b, c, d \in \{0, 1\}$ and $ad - bc \neq 0$. Prove that G is a group under matrix multiplication with order 6.

17. If G is a group such that $(ab)^2 = a^2b^2 \ \forall a, b \in G$ then prove that G is abelian.

Chapter 2

Linear vector spaces

A vector space is an algebraic structure $(V, F, \otimes, \odot)$ consisting of a non-empty set V, a field F, a binary operation $\otimes : V \times V \to V$ and external mapping $\odot : F \times V \to V$ associating each $a \in F, v \in V$ to a unique element $a \odot v \in V$ and satisfying the following.
(i) $(V, \otimes)$ is an abelian group
(ii) $\forall u, v \in V,$ and $a, b \in F$

$$a \odot (u \otimes v) = a \odot u \otimes a \odot v \tag{2.1}$$

$$(a \otimes b) \odot v = (a \odot v) \otimes (b \odot u) \tag{2.2}$$

$$(a \odot b) \odot u = a \odot (b \odot u) \tag{2.3}$$

$$1 \odot u = u \tag{2.4}$$

The elements of V are called vectors and they will be denoted as $\underline{\mathrm{u}}_i$ from now onwards. The elements of F are called scalars.

- The mapping $\odot$ can be identified as scalar multiplication and the binary operation $\otimes$ can be identified as vector addition.
- The vector space is called a real vector space if $F = \mathbb{R}$ and it is called a complex vector space if $F = \mathbb{C}$.
- Strictly speaking, the above properties define a linear vector space. But we will be using the term vector space to signify the same thing.

Subspaces:

Let V be a vector space over a field F. A non-empty $S \subset V$ is said to be a subspace of V if S itself is a vector space over F under the operations on V restricted to S.

Restriction:

If $S \subset V$ and a binary operation $\otimes$ is defined on S then $\otimes$ is said to be a restriction of S on V if $a \otimes b \in S \ \forall a, b \in S$.

Criteria for a subset to be a subspace:

Let V be a vector space over F. A non-empty set $S \subset V$ is a subspace iff $\forall \underline{\mathrm{u}}, \underline{\mathrm{v}} \in S$ and $\forall \alpha \in F$

$$\underline{\mathrm{u}} + \underline{\mathrm{v}} \in S \tag{2.5}$$

$$\alpha \underline{\mathrm{u}} \in S \tag{2.6}$$

Hence, S is an additive abelian group under vector addition and S is closed under scalar multiplication. The above criteria can be alternatively posed as follows.

A non-empty $S \subset V$ is a subspace of V iff $\alpha \underline{\mathrm{u}} + \beta \underline{\mathrm{v}} \in S, \ \underline{\mathrm{u}}, \underline{\mathrm{v}} \in S, \ \alpha, \beta \in F$.

Problems

1. Show that a set of all ordred n-tuples of elements of any field F is a vector space over F.

2. Show that a set of $F^{m \times n}$ matrices over a field F is a vector space over F with respect to the addition of matrices as the vector addition and multiplication of a matrix by a scalar as the scalar multiplication.

3. Show that the set $F[x]$ of all polynomials over a field F is a vector space over F.

4. Let V_1 and V_2 be two vector spaces over the same field F. Show that their Cartesian product $V_1 \times V_2 = \{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2\}$ is a vector space over F.

5. Let F be a field and V be a set of all ordered pairs $(a_1, a_2), a_1, a_2 \in F$. The following operations have been defined.

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$$
$$\lambda(a_1, a_2) = (\lambda a_1, a_2), \lambda \in F$$

   Verify whether the operations given above make V a vector space over F.

6. Show that a set V of all real valued continuous functions defined on a closed interval $[a, b]$ is a real vector space with vector addition and scalar multiplication defined as follows.

$$(f + g)(x) = f(x) + g(x)$$
$$(\lambda f)(x) = \lambda f(x)$$

   $\forall f, g \in V, \lambda \in \mathbb{R}$.

7. If X is a non-empty set and V is a vector space over F then show that $V^X = \{f : X \to V\}$ is a vector space over F under the vector addition and scalar multiplication defined below.

$$(f + g)(x) = f(x) + g(x)$$
$$(\lambda f)(x) = \lambda f(x)$$

$\forall f, g \in V^X, \lambda \in F$.

8. Let $V = \{(x, y) : x, y \in \mathbb{R}\}$. Show that V is not a vector space under the vector addition and scalar multiplication defined below.

$$(a_i, b_i) + (a_j, b_j) = (3b_i + 3b_j, -a_i - a_j)$$
$$k(a_i, b_i) = (3kb_i, -ka_i)$$

$\forall (a_i, b_i), (a_j, b_j) \in V, k \in \mathbb{R}$.

9. Let $V = \{(x, y) : x, y \in \mathbb{R}\}$. For any $\alpha = (x_i, y_i), \beta = (x_j, y_j) \in V, c \in \mathbb{R}$, the following operations are defined.

$$\alpha \otimes \beta = (x_i + x_j + 1, y_i + y_j + 1)$$
$$c \odot \alpha = (cx_i, cy_i)$$

(i) Prove that $(V, \otimes)$ is an abelian group.

(ii) Prove that V is not a vector space over $\mathbb{R}$ under the two operations.

10. For any $u = (x_i, x_j, x_k)$ and $v = (y_i, y_j, y_k)$, $u, v \in \mathbb{R}^3, a \in \mathbb{R}$,

$$u \otimes v = (x_i + y_i + 1, x_j + y_j + 1, x_k + y_k + 1)$$
$$a \odot u = (ax_i + a - 1, ax_j + a = 1, ax_k + a - 1)$$

Prove that $\mathbb{R}^3$ is a vector space over $\mathbb{R}$ under these two operations.

11. If $V = \mathbb{R}^+, u \otimes v = uv, a \odot u = u^a \; \forall u, v \in V, a \in \mathbb{R}$ then prove that V is a vector space over $\mathbb{R}$ under these two operations.

12. Identify whether $\mathbb{R}^3$ is a vector space over $\mathbb{R}$ under the following two operations.

$$(x_i, x_j, x_k) + (y_i, y_j, y_k) = (x_i + y_i, x_j + y_j, x_k + y_k)$$
$$a(x_i, x_j, x_k) = (ax_i, ax_j, ax_k)$$

$\forall (x_i, x_j, x_k), (y_i, y_j, y_k) \in \mathbb{R}^3, a \in \mathbb{R}$.

13. Let $V = \{(x, 1) : x \in \mathbb{R}\}$. For any $u = (x, 1), v = (y, 1) \in V, a \in \mathbb{R}$,

$$u \otimes v = (x + y, 1)$$
$$a \odot u = (ax, 1)$$

Verify whether V is a vector space over $\mathbb{R}$ under these two operations.

14. Let V be a set of ordered pairs $(a, b), a, b \in \mathbb{R}$. The following operations are defined.

$$(a, b) \otimes (c, d) = (a + c, b + d)$$
$$k(a, b) = (ka, 0)$$

Show that V is not a vector space over $\mathbb{R}$ under these two operations.

15. Let V be the set of ordered pairs $(a, b), a, b \in \mathbb{R}$ and $k \in \mathbb{R}$. Show that V is not a vector space over $\mathbb{R}$ with the vector addition and scalar multiplication defined by:

(i)

$$(a, b) + (c, d) = (a + d, b + c)$$
$$k(a, b) = (ka, kb)$$

(ii)

$$(a, b) + (c, d) = (a + d, b + c)$$
$$k(a, b) = (a, b)$$

(iii)

$$(a, b) + (c, d) = (0, 0)$$
$$k(a, b) = (ka, kb)$$

(iv)

$$(a, b) + (c, d) = (ac, bd)$$
$$k(a, b) = (ka, kb)$$

16. Show that $S = \{(0, b, c) : b, c \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$.

17. If $V = \{(x, y, z) : x, y, z \in \mathbb{R}\}$ and $S = \{(x, x, x) : x \in \mathbb{R}\}$ then show that S is a subspace of V. What is the geometrical interpretation of this?

18. Let $a_i, a_j, a_k \in F$. Show that the set S of all triads (x_i, x_j, x_k) of elements of $F : a_i x_i + a_j x_j + a_k x_k = 0$ is a subspace of F^3. What is the geometrical interpretation of this when $F = \mathbb{R}$.

19. Show that a set S of all $n \times n$ symmetric matrices over a field F is a subspace of the vector space $F^{n\times n}$ matrices over F.

20. Let V be the vector space of all real valued continuous functions over $\mathbb{R}$. Show that the set S of solutions of the differential equation

$$2\frac{d^2y}{dx^2} - 9\frac{dy}{dx} + 2y = 0$$

is a subspace of V.

21. Let $\mathbb{R}$ be a field of real numbers and S be the set of all solutions of the equation $x + y + 2z = 0$. Show that S is a subspace of $\mathbb{R}^3$.

22. Let S be the set of all elements of the form $(x + 2y, y, -x + 3y) \in \mathbb{R}^3, x, y \in \mathbb{R}$. Show that S is a subspace of $\mathbb{R}^3$.

23. Let V be the vector space of all 2×2 matrices over $\mathbb{R}$. Show that
(i) the set of all 2×2 singular matrices over $\mathbb{R}$ is not a subspace of V.
(ii) the set of all 2×2 matrices satisfying $A \times A = A, A \in V$ is not a subspace of A.

24. Let V be a vector space over $\mathbb{R}^3$. Which of the following subsets of V are subspaces of V?
(i) $S_1 = \{(a, b, c) : a + b = 0\}$
(ii) $S_2 = \{(a, b, c) : a = 2b + 1\}$
(iii) $S_3 = \{(a, b, c) : a \geq 0\}$
(iv) $S_4 = \{(a, b, c) : a^2 = b^2\}$

(v) $S_5 = \{(a, 2b, 3c) : a, b, c \in \mathbb{R}\}$

(vi) $S_6 = \{(a, a, a) : a \in \mathbb{R}\}$

(vii) $S_7 = \{(a, b, c) : a, b, c \in \mathbb{Q}\}$

25. Let $V = \{a_0 + a_1x + a_2x^2 + a_3x^3 : a_i \in \mathbb{R}\}$ be the set of all polynomials of degree less than or equal to 3 over $\mathbb{R}$. V is a vector space over $\mathbb{R}$. Prove that

(i) $S_1 = \{a_0 + a_2x^2 : a_0, a_2 \in \mathbb{R}\}$ is a subspace of V.

(ii) $S_2 = \{a_0 + a_1x + a_2x^2 + a_3x^3 : a_0 = a_2 + a_3, a_i \in \mathbb{R}\}$ is a subspace of V.

26. Let V be the set of all 2×3 matrices. V is a vector space over $\mathbb{R}$. Determine which of the following subsets of V are subspaces of V?

(i)

$$S_1 = \left\{ \begin{bmatrix} a & b & c \\ d & 0 & 0 \end{bmatrix} : a, b, c, d \in \mathbb{R} \right\}$$

(ii)

$$S_2 = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} : a + c = e + f, a, b, c, d, e, f \in \mathbb{R} \right\}$$

(iii)

$$S_3 = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} : a > 0, b = c, a, b, c, d, e, f \in \mathbb{R} \right\}$$

Chapter 3

Linear independence, basis, dimension and span

Linear combination:

If V is a vector space over a field F, $\underline{\mathrm{v}}_1, \underline{\mathrm{v}}_2, \ldots, \underline{\mathrm{v}}_n$ are n vectors in V and $\lambda_1, \lambda_2, \ldots \lambda_n$ are scalars in F then

$$\underline{\mathrm{v}}_{n+1} = \sum_{i=1}^{n} \lambda_i \underline{\mathrm{v}}_i \tag{3.1}$$

is called a linear combination of $\underline{\mathrm{v}}_i, \underline{\mathrm{v}}_{n+1} \in V$.

Linear span:

Let V be a vector space over a field F and $S \subset V$. A set of all possible finite linear combinations of vectors in S in called the linear span of S.

$$[S] = \left\{ \sum_{i=1}^{n} \lambda_i \underline{\mathrm{v}}_i : \lambda_i \in F, n \in \mathbb{N}, \underline{\mathrm{v}}_i \in S \right\} \tag{3.2}$$

Linear independence:

A set of vectors $\underline{\mathrm{v}}_1, \underline{\mathrm{v}}_2, \ldots, \underline{\mathrm{v}}_n$ in a vector space V over F are said to be linearly independent if

$$\sum_{i=1}^{n} \lambda_i \underline{\mathrm{v}}_i = 0 \implies \lambda_i = 0 \forall i, \lambda_i \in F \tag{3.3}$$

Basis and dimension: A non-empty subset $B \subset V$ is said to be a basis for a vector space V if

(i) B spans V

(ii) all vectors in B are linearly independent.

A vector space is said to be finite dimensional if there exists a finite subset of V that spans it. The followings are noteworthy.

(i) A null vector cannot be in the basis.

(ii) If the number of vectors in the basis is not finite then the vector space is called infinite dimensional.

If the number of vectors present in the basis are n then the vector space is said to be n-dimensional. If $V(F)$ is a vector space over F then for dimension n,

(i) any set of $n+1$ or more vectors in V are linearly dependent.

(ii) no set of $n-1$ or lesser vectors can span V.

Problems

1. Express the following vectors in $\mathbb{R}^3$ as a linear combination of $\underline{v}_1, \underline{v}_2, \underline{v}_3$.

   (i) $\underline{w}_1 = [1 \quad -2 \quad 5]^T$; (ii) $\underline{w}_2 = [1 \quad 1 \quad 1]^T$; (iii) $\underline{w}_3 = [0 \quad 0 \quad 0]^T$

$$\underline{v}_1 = [1 \quad 1 \quad 1]^T; \quad \underline{v}_2 = [1 \quad 2 \quad 3]^T; \quad \underline{v}_3 = [2 \quad -1 \quad 1]^T$$

2. Express the polynomial $f(x)$ as a linear combination of the polynomial functions $\phi_1(x), \phi_2(x), \phi_3(x)$.

$$f(x) = x^2 + 4x - 3$$
$$\phi_1(x) = x^2 - 2x + 5$$
$$\phi_2(x) = 2x^2 - 3x$$
$$\phi_3(x) = x + 3$$

3. Let $V = \mathbb{R}^{2\times 2}$ be a vector space of all 2×2 matrices over $\mathbb{R}$. Express $\underline{\underline{M}}$ as a linear combination of $\underline{\underline{M}}_1, \underline{\underline{M}}_2, \underline{\underline{M}}_3$.

$$\underline{\underline{M}} = \begin{bmatrix} 4 & 7 \\ 7 & 9 \end{bmatrix}; \quad \underline{\underline{M}}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}; \quad \underline{\underline{M}}_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; \quad \underline{\underline{M}}_3 = \begin{bmatrix} 1 & 1 \\ 4 & 5 \end{bmatrix}$$

4. Consider $\underline{v}_1 = [1 \quad 2 \quad 3]^T$ and $\underline{v}_2 = [2 \quad 3 \quad 1]^T$. Find conditions on a, b and c so that $\underline{v} = [a \quad b \quad c]^T$ is a linear combination of $\underline{v}_1$ and $\underline{v}_2$.

5. Let $V = P_2(t)$ be the vector space of all polynomials of degree lesser than or equal to 2. Express the polynomial $f(t) = at^2 + bt + c$ as a linear combination of the following polynomials.

$$f_1(t) = (t-1)^2; \qquad f_2(t) = t - 1; \qquad f_3(t) = 1$$

6. Express the vector $\underline{u} = [2 \quad 4 \quad 5]^T$ in $\mathbb{R}^3$ as a linear combination of $\underline{v}_1 = [1 \quad 2 \quad 3]^T$ and $\underline{v}_2 = [2 \quad 3 \quad 1]^T$.

7. Show that the vectors $\underline{v}_1 = [1 \quad 1]^T$ and $\underline{v}_2 = [1 \quad 2]^T$ span $\mathbb{R}^2$ over $\mathbb{R}$.

8. Show that the polynomials 1, $1+x$ and $(1+x)^2$ span the vector space $V = P_2(x)$ of all polynomials of degree at most 2 over $\mathbb{R}$.

9. Show that $\underline{v}_1$, $\underline{v}_2$ and $\underline{v}_3$ span $\mathbb{R}^3$ where $\underline{v}_1 = [1 \quad 1 \quad 1]^T$, $\underline{v}_2 = [1 \quad 2 \quad 3]^T$ and $\underline{v}_3 = [1 \quad 5 \quad 8]^T$.

10. Determine the condition(s) on a, b and c so that $\underline{v} = [a \quad b \quad c]^T$ in $\mathbb{R}^3$ over $\mathbb{R}$ belongs to the subspace spanned by $\underline{v}_1 = [1 \quad 2 \quad 0]^T$, $\underline{v}_2 = [-1 \quad 1 \quad 2]^T$ and $\underline{v}_3 = [3 \quad 0 \quad -4]^T$.

11. Let $\underline{v}_1 = [-1 \quad 2 \quad 0]^T$, $\underline{v}_2 = [3 \quad 2 \quad -1]^T$ and $\underline{v}_3 = [1 \quad 6 \quad -1]^T$ be three vectors in $\mathbb{R}^3$ over $\mathbb{R}$. Show that $[\underline{v}_1 \quad \underline{v}_2] = [\underline{v}_1 \quad \underline{v}_2 \quad \underline{v}_3]$.

12. Let $\underline{v}_1 = [1 \quad 2 \quad -1]^T$, $\underline{v}_2 = [2 \quad -3 \quad 2]^T$, $\underline{v}_3 = [4 \quad 1 \quad 3]^T$ and $\underline{v}_4 = [-3 \quad 1 \quad 2]^T$ be four vectors in $\mathbb{R}^3$ over $\mathbb{R}$. Show that $[\underline{v}_1 \quad \underline{v}_2] \neq [\underline{v}_3 \quad \underline{v}_4]$.

13. Let $\mathbb{R}^{2\times 2}$ be the vector space of all 2×2 matrices. Show that the matrices

$$\underline{\underline{M}}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \quad \underline{\underline{M}}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \quad \underline{\underline{M}}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}; \quad \underline{\underline{M}}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

span $\mathbb{R}^{2\times 2}$.

14. Verify whether the following vectors span $\mathbb{R}^4$.

$\underline{v}_1 = [3 \quad -1 \quad 0 \quad -1]^T$; $\underline{v}_2 = [2 \quad -1 \quad 3 \quad 2]^T$
$\underline{v}_3 = [-1 \quad 1 \quad 1 \quad 3]^T$; $\underline{v}_4 = [1 \quad 1 \quad 9 \quad -5]^T$

15. Find one vector in $\mathbb{R}^3$ over $\mathbb{R}$ that spans the intersection of subspaces of S and T where $S = \{(a, b, 0) : a, b \in \mathbb{R}\}$ and $T = [\underline{u}_1, \underline{u}_2]$, $\underline{u}_1 = [1 \quad 1 \quad 1]^T$, $\underline{u}_2 = [1 \quad 2 \quad 3]^T$.

16. Consider the vector space $V = P_n(t)$ consiting of all polynomials of degree $\leq n$. Show that the set of polynomials $1, t, t^2, t^3, \ldots, t^n$ span V.

17. Show that the following vectors are linearly independent in $\mathbb{R}^4$ over $\mathbb{R}$.

$\underline{v}_1 = [1 \quad 1 \quad 2 \quad 4]^T$; $\underline{v}_2 = [2 \quad -1 \quad -5 \quad 2]^T$
$\underline{v}_3 = [2 \quad 1 \quad 1 \quad 6]^T$; $\underline{v}_4 = [1 \quad -1 \quad -4 \quad 0]^T$

18. Determine whether the vectors

$$f(x) = 2x^3 + x^2 + x + 1$$
$$g(x) = x^3 + 3x^2 + x - 2$$
$$h(x) = x^3 + 2x^2 - x + 3$$

in the vector space $R[x]$ over $\mathbb{R}$ are linearly independent?

19. Let V be the vector space of functions from $\mathbb{R}$ into $\mathbb{R}$. Show that the functions $f(t) = sin(t), g(t) = e^t, h(t) = t^2$ are linearly independent in V.

20. Verify if the following matrices in $\mathbb{R}^{2\times 2}$ over $\mathbb{R}$ are linearly independent.

$$\underline{\underline{M}}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}; \quad \underline{\underline{M}}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \underline{\underline{M}}_3 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

21. Which of the following subsets S_i of $\mathbb{R}^3$ are linearly independent?
(i) $S_1 = \{[1 \quad 2 \quad 1]^T, [-1 \quad 1 \quad 0]^T, [5 \quad -1 \quad 1]^T\}$
(ii) $S_1 = \{[1 \quad 1 \quad 0]^T, [0 \quad 0 \quad 1]^T, [1 \quad 5 \quad 2]^T\}$
(iii) $S_1 = \{[1 \quad 3 \quad 2]^T, [1 \quad -7 \quad 8]^T, [2 \quad 1 \quad 1]^T\}$
(iv) $S_1 = \{[1 \quad 5 \quad 2]^T, [1 \quad 0 \quad 0]^T, [0 \quad 1 \quad 0]^T\}$

22. Which of the following subsets S_i of $R[x]$ are linearly independent?
(i) $S_1 = \{1, x - x^2, x + x^2, 3x\}$
(ii) $S_2 = \{x^2 - 1, x^2 + 1, x - 1\}$
(iii) $S_3 = \{x, x - x^3, x^2 + x^4, a + x^2 + x^4 + \frac{1}{2}\}$
(iv) $S_4 = \{1, 1 + x, 1 + x + x^2, x^4\}$

23. Which of the following subsets of all continuous function space are linearly independent?
(i) $S_1 = \{sin(x), cos(x), sin(x + 1)\}$
(ii) $S_2 = \{xe^x, x^2e^x, (x^2 + x - 1)e^x\}$
(iii) $S_3 = \{sin^2(x), cos(2x), 1\}$
(iv) $S_4 = \{x, sin(x), cos(x)\}$

24. If the set $\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$ is linearly independent in a vector space $V(F)$ then prove that the set $\{\underline{v}_1 + \underline{v}_2, \underline{v}_2 + \underline{v}_3, \underline{v}_3 + \underline{v}_1\}$ is also linearly independent.

25. If $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ is a linearly independent set of vectors in a vector space $V(F)$ and $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n, \underline{u}\}$ is a linearly dependent set then prove that $\underline{u}$ is a linear combination of $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$.

26. Find a maximal linearly independent subsystem of a system having the following vectors.
$\underline{v}_1 = [2 \quad -2 \quad -4]^T$; $\underline{v}_2 = [1 \quad 9 \quad 3]^T$
$\underline{v}_3 = [-2 \quad -4 \quad 1]^T$; $\underline{v}_4 = [3 \quad 7 \quad 1]^T$

27. Determine whether or not each of the following sets forms a basis of $\mathbb{R}^3(\mathbb{R})$.
(i) $B_1 = \{[1 \quad 1 \quad 1]^T, [1 \quad 0 \quad 1]^T\}$
(ii) $B_2 = \{[1 \quad 1 \quad 1]^T, [1 \quad 2 \quad 3]^T, [2 \quad -1 \quad 1]^T\}$
(iii) $B_3 = \{[1 \quad 1 \quad 2]^T, [1 \quad 2 \quad 5]^T, [5 \quad 3 \quad 4]^T\}$
(iv) $B_4 = \{[1 \quad 2 \quad 3]^T, [1 \quad 3 \quad 5]^T, [1 \quad 0 \quad 1]^T, [2 \quad 3 \quad 0]^T\}$

28. Let $\underline{v}_1 = [1 \quad i \quad 0]^T$, $\underline{v}_2 = [2i \quad 1 \quad 1]^T$ and $\underline{v}_3 = [0 \quad 1+i \quad 1-i]^T$ be three vectors in $\mathbb{C}^3(\mathbb{C})$. Show that the set $B = \{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$ is a basis of $\mathbb{C}^3(\mathbb{C})$.

29. Determine whether $\{[1 \quad 1 \quad 1 \quad 1]^T, [1 \quad 2 \quad 3 \quad 2]^T, [2 \quad 5 \quad 6 \quad 4]^T, [2 \quad 6 \quad 8 \quad 5]^T\}$ form a basis of $\mathbb{R}^4(\mathbb{R})$? If not then determine the dimension of the subspace they span.

30. Extend the set $\{[1 \quad 1 \quad 1 \quad 1]^T, [2 \quad 2 \quad 3 \quad 4]^T\}$ to a basis in $\mathbb{R}^4(\mathbb{R})$.

31. Let S be the set containing the following vectors in $\mathbb{R}^5(\mathbb{R})$.
$\{[1 \quad 2 \quad 1 \quad 3 \quad 2]^T, [1 \quad 3 \quad 3 \quad 5 \quad 3]^T, [3 \quad 8 \quad 7 \quad 13 \quad 8]^T,$
$[1 \quad 4 \quad 6 \quad 9 \quad 7]^T, [5 \quad 13 \quad 13 \quad 25 \quad 19]^T$
Find a basis of $[S]$ consisting of the originally given vectors. Also determine the dimension of $[S]$.

32. Let S be the set consisting of the following vectors in $\mathbb{R}^4$.
$[1 \quad -2 \quad 5 \quad 3]^T, \qquad [2 \quad 3 \quad 1 \quad -4]^T, \qquad [3 \quad 8 \quad -3 \quad -5]^T$

(i) Determine a basis and dimension of the subspace spanned by S.

(ii) Extend the basis of $[S]$ to a basis in $\mathbb{R}^4$.

(iii) Find a basis of $[S]$ consisting of the original vectors.

33. Let $S = \{\underline{u}_1, \underline{u}_2, \underline{u}_3\}$ and $T = \{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$ be subsets of $\mathbb{R}^3$ where

$\underline{u}_1 = [1 \quad 1 \quad -1]^T$; $\underline{u}_2 = [2 \quad 3 \quad -1]^T$; $\underline{u}_3 = [3 \quad 1 \quad -5]^T$

$\underline{v}_1 = [1 \quad -1 \quad -3]^T$; $\underline{v}_2 = [3 \quad -2 \quad -8]^T$; $\underline{v}_3 = [2 \quad 1 \quad -3]^T$

Show that $[S] = [T]$.

34. Which of the following subsets B_i form a basis for the given vector space V.

(i) $B_1 = \{[1 \quad 0]^T, \ [i \quad 0]^T, \ [0 \quad 1]^T\}$; $V = \mathbb{C}^2[\mathbb{R}]$

(ii) $B_2 = \{[1 \quad i \quad i+1]^T, \ [1 \quad i \quad i-1]^T, \ [i \quad -i \quad 1]^T\}$; $V = \mathbb{C}^3[\mathbb{C}]$

(iii) $B_3 = \{1, sin(x), sin^2(x), cos^2(x)\}$; $V = \mathbb{C}^3[-\pi, \pi]$

35. Show that the vectors $\underline{v}_1 = [1 \quad 0 \quad -1]^T$, $\underline{v}_2 = [1 \quad 2 \quad 1]^T$, $\underline{v}_3 = [0 \quad -3 \quad 2]^T$ form a basis in $\mathbb{R}^3$. Express each of these vectors as a linear combination of the standard basis in $\mathbb{R}^3$.

Chapter 4

Analysis of systems of simultaneous linear equations

Consider the following system of simultaneous linear equations.

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n$$

The above set of equations can be cast as a matrix equation given as follows.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{4.1}$$

4.1 can be written as $\underline{\underline{A}}\ \underline{\mathrm{x}} = \underline{\mathrm{b}}$ where $\underline{\underline{A}}$ is the coefficient matrix, $\underline{\mathrm{x}}$ is the vector of unknowns to be solved for, and $\underline{\mathrm{b}}$ is the vector consisting of the entries of the RHS.

One of the methods to solve such a system of equations is the Gauss elimination which involves converting $\underline{\underline{A}}$ into its echelon form followed by back substitution. However, here we would be interested in the analysis of the nature of the solutions without actually having to solve the system. We would be interested in answering the following questions.

(i) Does the solution exist?

(ii) If yes, then is the solution unique?

For this, we will make use of the concept of *null space* and *range space*. Let us first consider the homogeneous case.

$$\underline{\underline{A}}\ \underline{x} = \underline{0} \text{ (zero vector)} \tag{4.2}$$

The first task is to know whether the null space of $\underline{\underline{A}}$ is empty or it has elements in it. Null space is a space which consists of the vectors which satisfy $\underline{\underline{A}}\ \underline{x} = \underline{0}$ *i.e.* it is the space of solutions of the homogeneous case. If the null space is empty then the only solution is a trivial solution, $\underline{x} = [0\ \ 0\ \quad \ldots\ \ 0]^T$.

If the null space is non-empty then it is required to determine the dimension and a basis of the null space. All linear combinations of the basis in the null space satisfy $\underline{\underline{A}}\ \underline{x} = \underline{0}$.

How to check if the null space is empty? In such a case, the only solution to the homogeneous system is a trivial solution. For an $n \times n$ system, if $\mathbf{det}(\underline{\underline{A}})_{n\times n} = 0$ then the only solution is the trivial solution. In other words, the homogeneous system will have only a trivial solution if the rank $\rho(\underline{\underline{A}}) = n =$ number of unknowns. If this is not the case then we need to determine a basis of the null space of $\underline{\underline{A}}$. This can be determined following the procedure given below.

Consider a system of three simultaneous equations given below.

$$x_1 + x_2 + x_3 = 0$$

$$2x_1 + 3x_2 + x_3 = 0$$

$$5x_1 + 6x_2 + 4x_3 = 0$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 5 & 6 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_2 \to R_2 - 2R_1$

$R_3 \to R_3 - 5R_1$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_3 \to R_3 - R_2$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

No further operation will further reduce R_2 to a zero row. Therefore, we stop here and recast the matrix equation back to a system of simultaneous equations.

$$x_1 + x_2 + x_3 = 0$$
$$x_2 - x_3 = 0$$

We have two equations in three unknowns. Hence we have a degree of freedom of one. From the second equation, we assume $x_2 = \alpha$. Hence $x_2 = x_3 = \alpha$ and $x_1 = -2\alpha$. The solution to the homogeneous equations can, hence, be written as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \alpha \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

Hence, the dimension of the null space of $\underline{\underline{A}}$ is 1 since there is only one vector in its basis and a possible basis is $[-2 \;\; 1 \;\; 1]^T$. Every vector which is obtained by multiplying α to this vector will be a solution to the homogeneous equation. It can be seen that this vector or any multiple of it satisfy the original system of homogeneous equations. It can be noted that we had three unknowns (x_1, x_2, x_3) and two equations after row reduction. Hence the degree of freedom of the system was one. This resulted in the appearance of one parameter α. The dimension of the null space is also one. Hence, the dimension of the null space will be equal to the number of parameters or the degree of freedom of the row reduced system.

A system of homogeneous equations always has a solution since a trivial solution always exists for this case. Hence, for a homogeneous case, we have the following answers.

(i) Does the solution exist? $\rightarrow$ Yes, always

(ii) Is the solution unique? $\rightarrow$ The only possible unique solution in case of a homogeneous set of equations is a trivial solution. If the solution is not trivial then there are infinitely many solutions.

Now we focus on the non-homogeneous case.

$$\underline{\underline{A}}\ \underline{x} = \underline{b}; \quad \underline{b} \neq \underline{0}$$

If $\underline{\underline{A}}\ \underline{x} = \underline{0}$ has only a trivial solution then $\underline{\underline{A}}\ \underline{x} = \underline{b}$ has a unique solution. To prove this, let us assume that $\underline{\underline{A}}\ \underline{x} = \underline{b}$ has two distinct solutions $\underline{v}_1$ and $\underline{v}_2$.

$$\Longrightarrow \underline{\underline{A}}\ \underline{v}_1 = \underline{b}$$
$$\underline{\underline{A}}\ \underline{v}_2 = \underline{b}$$
$$\Longrightarrow \underline{\underline{A}}\ (\underline{v}_1 - \underline{v}_2) = \underline{0}$$

Hence, $\underline{v}_1 - \underline{v}_2$ must be the solution to the homogeneous case. But the only solution to the homogeneous case $\underline{\underline{A}}\ \underline{x} = \underline{0}$ has been given as the trivial solution. Hence $\underline{v}_1 - \underline{v}_2 = \underline{0}$ or $\underline{v}_1 = \underline{v}_2$ which means that the solution to $\underline{\underline{A}}\ \underline{x} = \underline{b}$ must be unique.

From the foregoing discussion, it can be concluded that to know the nature of solutions of $\underline{\underline{\mathrm{A}}}\ \underline{x}$=$\underline{b}$, we must first solve for $\underline{\underline{\mathrm{A}}}\ \underline{x}$=$\underline{0}$. If the only solution to $\underline{\underline{\mathrm{A}}}\ \underline{x}$=$\underline{0}$ is a trivial solution then $\underline{\underline{\mathrm{A}}}\ \underline{x}$=$\underline{b}$ must have a unique solution. If the solution to $\underline{\underline{\mathrm{A}}}\ \underline{x}$=$\underline{0}$ is not unique then $\underline{\underline{\mathrm{A}}}\ \underline{x}$=$\underline{b}$ may or may not be solvable. If it is solvable then it will have infinitely many solutions.

Now we focus on the question of determining whether $\underline{\underline{\mathrm{A}}}\ \underline{x}$=$\underline{b}$ has a solution or not. $\underline{\underline{\mathrm{A}}}\ \underline{x}$=$\underline{b}$ will have a solution if $\rho(\underline{\underline{A}}) = \rho(\underline{\underline{A}}|\underline{b})$. We make use of this fact to develop the range space of $\underline{\underline{A}}$. Range space of $\underline{\underline{A}}$ contains all the elements which satisfy $\underline{\underline{\mathrm{A}}}\ \underline{x}$=$\underline{b}$. Consider the previous case with $\underline{b} = [1 \;\; 2 \;\; 1]^T$. The corresponding matrix equation is given below.

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 5 & 6 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\underline{\underline{A}}|\underline{b} = \left[\begin{array}{ccc|c} 1 & 1 & 1 & b_1 \\ 2 & 3 & 1 & b_2 \\ 5 & 6 & 4 & b_3 \end{array}\right]$$

$R_2 \to R_2 - 2R_1$

$R_3 \to R_3 - 5R_1$

$$\underline{\underline{A}}|\underline{b} = \left[\begin{array}{ccc|c} 1 & 1 & 1 & b_1 \\ 0 & 1 & -1 & b_2 - 2b_1 \\ 0 & 1 & -1 & b_3 - 5b_1 \end{array}\right]$$

$R_3 \to R_3 - R_2$

$$\underline{\underline{A}}|\underline{b} = \left[\begin{array}{ccc|c} 1 & 1 & 1 & b_1 \\ 0 & 1 & -1 & b_2 - 2b_1 \\ 0 & 1 & -1 & b_3 - b_2 - 3b_1 \end{array}\right]$$

It can be seen from the above rearrangements that $\rho(\underline{\underline{A}}) = 2$. The condition of solvability dictates that $b_3 - b_2 - 3b_1 = 0$. Note that despite the entries known in $\underline{b}$, we continue with b_i's rather than the actual values given in $\underline{b}$. We now do the degree of freedom analysis.

Number of unknowns $= 3$ (b_1, b_2, b_3)

Number of equations $= 1$ $(b_3 - b_2 - 3b_1 = 0)$

Hence, we have two arbitarary parameters. Let $b_1 = \alpha$ and $b_2 = \beta \implies b_3 = 3\alpha + \beta$.

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Therefore, the dimension of the range space of $\underline{\underline{A}} = 2$ and a possible basis is $\underline{v}_1 = [1\ \ 0\ \ 3]^T$ and $\underline{v}_2 = [0\ \ 1\ \ 1]^T$. All possible linear combinations of $\underline{v}_1$ and $\underline{v}_2$ satisfy $\underline{\underline{A}}\ \underline{x}$=$\underline{b}$. Any $\underline{b}$ proposed must be expressed as a linear combination of $\underline{v}_1$ and $\underline{v}_2$ for the solution to exist. Therefore, as given originally in the problem,

$$\underline{b} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\implies \alpha = 1$$

$$\beta = 2$$

$$3\alpha + \beta = 1$$

No values of α and β satisfy the above set of equations. Hence, the solution does not exist for the system with $\underline{\mathrm{b}} = [1 \;\; 2 \;\; 1]^T$.

Problems

1. Determine whether the following set of simultaneous equations have a solution. If yes, then comment upon the nature of solutions.

   (i)

$$\begin{aligned} 2u + 3v &= 0 \\ 4u + 5v + w &= 0 \\ 2u - 5v - 3w &= 0 \end{aligned}$$

   (ii)

$$\begin{aligned} x + y &= 0 \\ x + 2y + z &= 0 \\ y + 2z + t &= 0 \\ z + 2t &= 5 \end{aligned}$$

   (iii)

$$\begin{aligned} 2u - v &= 0 \\ -u + 2v - w &= 0 \\ -v + 2w - z &= 0 \\ -w + 2z &= 5 \end{aligned}$$

   (iv)

$$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ x_1 + 2x_2 + 2x_3 &= 11 \\ 2x_1 + 3x_2 - 4x_3 &= 3 \end{aligned}$$

   (v)

$$\begin{aligned} x_1 + x_2 + x_3 &= -2 \\ 3x_1 + 3x_2 - x_3 &= 6 \\ x_1 - x_2 + x_3 &= -1 \end{aligned}$$

2. Determine the value of 'c' which makes it possible to solve $\underline{\underline{A}}\ \underline{\mathrm{x}} = \underline{\mathrm{b}}$ for

$$x_1 + x_2 + 2x_3 = 2$$
$$2x_1 + 3x_2 - x_3 = 5$$
$$3x_1 + 4x_2 + x_3 = c$$

3. Consider the system of equations given below.

$$x_1 + 2x_2 + 3x_3 + 5x_4 = b_1$$
$$2x_1 + 4x_2 + 8x_3 + 12x_4 = b_2$$
$$3x_1 + 6x_2 + 7x_3 + 13x_4 = b_3$$

   (i) Determine the dimension and basis for the null space of the coefficient matrix.

   (ii) Determine the dimension and basis for the range space.

   (c) Find the conditions on b_1, b_2, b_3 to have a solution.

4. For each of the following system of equations, find the conditions on b_i's so as to have a solution for the system.

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 2 & 5 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

5. Construct matrices subject to the following conditions.

   (i) The null space has a basis of $[2\ \ 2\ \ 1\ \ 0]^T$ and $[3\ \ 1\ \ 0\ \ 1]^T$.

   (ii) The column space has a basis of $[1\ \ 1\ \ 1]^T$ and the null space has a basis of $[2\ \ 2\ \ 1\ \ 0]^T$.

   (iii) The null space has a basis of $[4\ \ 3\ \ 2\ \ 1]^T$.

   (iv) The column space has a basis of $[1\ \ 1\ \ 0]^T$ and $[0\ \ 1\ \ 1]^T$ and the null space has a basis of $[1\ \ 0\ \ 1\ \ 0]^T$ and $[0\ \ 0\ \ 1\ \ 1]^T$.

   (v) The column space has a basis of $[1\ \ 1\ \ 5]^T$ and $[0\ \ 3\ \ 1]^T$ and the null space has a basis of $[1\ \ 1\ \ 2]^T$.

6. Determine the corresponding ranges space for the following systems of equations.

(i)

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 5 & 7 \\ 3 & 9 & 12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

(ii)

$$\begin{bmatrix} 1 & 3 & 1 & 2 \\ 2 & 6 & 4 & 8 \\ 0 & 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

(iii)

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

(iv)

$$\begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

(v)

$$\begin{aligned} x + 2y - 2z &= b_1 \\ 2x + 5y - 4z &= b_2 \\ 4x + 9y - 8z &= b_3 \end{aligned}$$

(vi)

$$\begin{aligned} x_1 + 2x_2 + 2x_3 + 4x_4 + 6x_5 &= b_1 \\ x_1 + 2x_2 + 3x_3 + 6x_4 + 9x_5 &= b_2 \\ x_3 + 2x_4 + 3x_5 &= b_3 \end{aligned}$$

7. Examine the solvability of the following system.

$$x_1 + 2x_2 = b_1$$
$$2x_1 + 4x_2 = b_2$$
$$x_1 = b_3$$

8. Consider the following matrix.

$$\underline{\underline{A}} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & -1 & 6 \\ 1 & 2 & 4 \end{bmatrix}$$

(i) Determine the dimension and a basis for the null space of $\underline{\underline{A}}$.

(ii) Determine the dimension and a basis for the range space.

(iii) Which of the following vectors $\underline{b}$ will yield a solution to $\underline{\underline{A}}\ \underline{x} = \underline{b}$?

$$\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 6 \\ 5 \\ 12 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 9 \\ 13 \end{bmatrix}$$

9. For the following set of equations, determine the dimension and a basis for the null space and the range space.

$$x_1 + 2x_2 + x_3 + 2x_4 - 3x_5 = 2$$
$$3x_1 + 6x_2 + 4x_3 - x_4 - 2x_5 = -1$$
$$4x_1 + 8x_2 + 5x_3 + x_4 - x_5 = 1$$
$$-2x_1 - 4x_2 - 3x_3 + 4x_4 - 5x_5 = 3$$

10. Consider the following matrix.

$$\underline{\underline{A}} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 2 \\ 4 & -1 & 9 \\ 2 & 1 & 1 \end{bmatrix}$$

Determine the dimension and a basis for the null and range spaces. Determine the solutions to $\underline{\underline{A}}\ \underline{x} = \underline{b}$ for $\underline{b} = [3\ \ 0\ \ 0\ \ 2]^T$.

11. For the following systems of equations, determine the dimension and bases for the null and range spaces. Determine the solutions to the systems on the basis of your analysis.

(i)

$$x_1 - x_2 + 3x_3 + 2x_4 = b_1$$
$$3x_1 + x_2 - x_3 + x_4 = b_2$$
$$-x_1 - 3x_2 + 7x_3 + 3x_4 = b_3$$

(ii)

$$x_1 + 2x_2 - x_3 + x_5 = b_1$$
$$3x_1 + 2x_2 + x_4 = b_2$$
$$x_1 - 2x_2 + 2x_3 + x_4 - 2x_5 = b_3$$

(iii)

$$x_1 + x_2 - x_3 = b_1$$
$$-2x_1 - x_2 + x_3 = b_2$$
$$x_1 + 2x_2 - 2x_3 = b_3$$

(vi)

$$5x_1 + 10x_2 + x_3 - 2x_4 = 6$$
$$-x_1 + x_2 - 2x_3 + x_4 = 0$$
$$2x_1 + 3x_2 + x_3 - x_4 = 2$$
$$6x_1 + 9x_2 + 3x_3 - 3x_4 = 6$$

Chapter 5

Linear transformations

Till now we considered how to analyze vectors in a given vector space. It is now desired to study the relationships among different vector spaces over the same field. Let V_1 and V_2 be two vector spaces over the same field F. A mapping $t : V_1 \to V_2$ is called a linear transformation or linear map or homomorphism if

(i) $t(\underline{u} + \underline{v}) = t(\underline{u}) + t(\underline{v})$

(ii) $t(\alpha \underline{u}) = \alpha t(\underline{u})$

$\forall \underline{u}, \underline{v} \in V_1, \alpha \in F$. A linear transformation from a vector space to itself is called a linear operator.

The above two conditions can be merged together to get a single condition as follows.

$t : V_1 \to V_2$ is called a linear transformation if $t(\alpha \underline{u} + \beta \underline{v}) = \alpha \underline{u} + \beta \underline{v}$ $\forall \underline{u}, \underline{v} \in V_1, \alpha, \beta \in F$.

Some other properties satisfied by linear transformations are given below.

(i) $t(\underline{0}_{V1}) = \underline{0}_{V2}$ where $\underline{0}_{V1}$ and $\underline{0}_{V2}$ are the zero vectors in V_1 and V_2, respectively.

(ii) $t(-\underline{u}) = -t(\underline{u})$

(iii) $t(\underline{u} - \underline{v}) = t(\underline{u}) - t(\underline{v})$

$\forall \underline{u}, \underline{v} \in V_1$

If V_1 and V_2 are two vector spaces over the same field F and $B = \{\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n\}$ is a basis in V_1 then there exists a unique transformation $t : V_1 \to V_2$ such that $t(\underline{b}_i) = \underline{b}'_i$ where $\underline{b}'_i$ are the vectors in V_2.

To prove the above, the following steps are adopted. Since B forms the basis in V_1,

$$\underline{v} = \sum_{i=1}^{n} \lambda_i \underline{b}_i, \underline{v} \in V_1 \tag{5.1}$$

$$\begin{aligned}
t(\underline{v}) &= t\left(\sum_{i=1}^{n} \lambda_i \underline{v}_i\right) \\
&= \sum_{i=1}^{n} (\lambda_i t(\underline{b}_i)) \\
&= \sum_{i=1}^{n} \lambda_i \underline{b}'_i
\end{aligned}$$

Since λ_i's need to be unique scalars, $\sum_{i=1}^{n} \lambda_i \underline{b}'_i$ shall also be unique. If $\underline{u}, \underline{v} \in V_1$ and $\alpha, \beta, \lambda_i, \mu_i \in F$,

$$\begin{aligned}
\underline{u} &= \sum_{i=1}^{n} \lambda_i \underline{b}_i \\
\underline{v} &= \sum_{i=1}^{n} \mu_i \underline{b}_i \\
t(\alpha\underline{u} + \beta\underline{v}) &= t\left(\alpha\left(\sum_{i=1}^{n} \lambda_i \underline{b}_i\right) + \beta\left(\sum_{i=1}^{n} \mu_i \underline{b}_i\right)\right) \\
&= t\left(\sum_{i=1}^{n} (\alpha\lambda_i + \beta\mu_i)\underline{b}_i\right) \\
&= \sum_{i=1}^{n} (\alpha\lambda_i + \beta\mu_i)\underline{b}'_i \\
&= \sum_{i=1}^{n} (\alpha(\lambda_i \underline{b}'_i) + \beta(\mu_i \underline{b}'_i)) \\
&= \alpha\left(\sum_{i=1}^{n} \lambda_i \underline{b}'_i\right) + \beta\left(\sum_{i=1}^{n} \mu_i \underline{b}'_i\right) \\
&= \alpha t(\underline{u}) + \beta t(\underline{v}) \\
\implies t(\alpha\underline{u} + \beta\underline{v}) &= \alpha t(\underline{u}) + \beta t(\underline{v}) \forall \underline{u}, \underline{v} \in V_1, \alpha, \beta \in F
\end{aligned}$$

Hence, $t : V_1 \to V_2$ is a linear transformation.

Till now we have shown that the transformation $t : V_1 \to V_2$ is linear and the vectors $t(\underline{v}) \in V_2$ are unique linear combinations of the basis in V_2. Now we show that the transformation itself is unique.

Let $t' : V_1 \to V_2$ be another linear transformation such that $t(\underline{\mathrm{b}}_i) = \underline{\mathrm{b}}'_i \ \forall i \in \mathbb{N}$.

$$\begin{aligned}
\underline{\mathrm{v}} &= \sum_{i=1}^{n} \lambda_i \underline{\mathrm{b}}_i \in V_1 \\
t'(\underline{\mathrm{v}}) &= t'\big(\sum_{i=1}^{n} \lambda_i \underline{\mathrm{b}}_i\big) \\
&= \sum_{i=1}^{n} \lambda_i t'(\underline{\mathrm{b}}_i) \\
&= \sum_{i=1}^{n} \lambda_i \underline{\mathrm{b}}'_i \\
&= t(\underline{\mathrm{v}}) \\
\implies t'(\underline{\mathrm{v}}) &= t(\underline{\mathrm{v}})
\end{aligned}$$

Hence, $t : V_1 \to V_2$ is unique.

If $\underline{\underline{A}}$ is an $m \times n$ matrix and a mapping t_A is defined such that $t_A : \underline{\underline{F}}^{n\times 1} \to \underline{\underline{F}}^{m\times 1}$

$$t_A(\underline{\underline{X}}) = \underline{\underline{A}}\ \underline{\underline{X}}\ \forall \underline{\underline{X}} = [x_1 \quad x_2 \quad \ldots x_n]^T \in \underline{\underline{F}}^{n\times 1}$$

then $\forall \underline{\underline{X}}, \underline{\underline{Y}} \in \underline{\underline{F}}^{n\times 1}$ and $\lambda \in F$, we have

$$\begin{aligned}
t_A(\underline{\underline{X}} + \underline{\underline{Y}}) &= \underline{\underline{A}}(\underline{\underline{X}} + \underline{\underline{Y}}) \\
&= \underline{\underline{A}}\ \underline{\underline{X}} + \underline{\underline{A}}\ \underline{\underline{Y}} \\
&= t_A(\underline{\underline{X}}) + t_A(\underline{\underline{Y}}) \\
\implies t_A(\underline{\underline{X}} + \underline{\underline{Y}}) &= t_A(\underline{\underline{X}}) + t_A(\underline{\underline{Y}})
\end{aligned}$$

Similarly, it can be easily shown that

$$t_A(\lambda \underline{\underline{X}}) = \underline{\underline{A}}(\lambda \underline{\underline{X}}) = \lambda(\underline{\underline{A}}\ \underline{\underline{X}}) = \lambda t(\underline{\underline{X}})$$

Thus, $t_A : \underline{\underline{F}}^{n\times 1} \to \underline{\underline{F}}^{m\times 1}$ is a linear transformation. Hence, every $m \times n$ matrix can be viewed as a linear transformation from $\underline{\underline{F}}^{n\times 1} \to \underline{\underline{F}}^{m\times 1}$.

Kernel and image of a linear transformation:

If V_1 and V_2 are the vector spaces over the same field F and $t : V_1 \to V_2$ then kernel of t, $ker(t)$, is defined as

$$ker(t) = \{\underline{\mathrm{v}} \in V_1 : t(\underline{\mathrm{v}}) = \underline{0}_{V_2}\} \tag{5.2}$$

It is easy to recognize that $ker(t)$ is nothing but the null space of t. Image of t, $Im(t)$, is defined as

$$Im(t) = \{t(\underline{v}) : \underline{v} \in V_1\} \tag{5.3}$$

It is easy to prove that

(i) $ker(t)$ is a subspace of V_1

(ii) $Im(t)$ is a subspace of V_2

The following conclusions follow the previous results.

- If $\{\underline{v}_1, \underline{v}_2 \ldots \underline{v}_n\}$ spans V_1 then $\{t(\underline{v}_1), t(\underline{v}_2) \ldots t(\underline{v}_n)\}$ spans V_2.
- If $\{\underline{v}_1, \underline{v}_2 \ldots \underline{v}_n\}$ is a set of linearly independent vectors in V_1 then $\{t(\underline{v}_1), t(\underline{v}_2) \ldots t(\underline{v}_n)\}$ is also linearly independent in V_2.
- As seen previously, $\underline{\underline{A}}(m \times n)$ is a linear transformation $t_A : \underline{\underline{F}}^{n\times 1} \rightarrow \underline{\underline{F}}^{m\times 1}$ and the kernel of t_A consists of all vectors $\underline{x}$ for which $t(\underline{x}) = \underline{0}$. Hence, $ker(t)$ is $\{\underline{x} : \underline{\underline{A}}\ \underline{x} = \underline{0}\}$.

Rank and nullity:

If V_1 and V_2 are the vector spaces over the same field F and $t : V_1 \rightarrow V_2$ is a linear transformation then

$$rank(t) = dim(Im(t)) \tag{5.4}$$

$$nullity(t) = dim(ker(t)) \tag{5.5}$$

Sylvester's law:

$$dim(Im(t)) + dim(ker(t)) = dim(V) \tag{5.6}$$

$$rank(t) + nullity(t) = dim(V) \tag{5.7}$$

Problems

1. Let F be a field. Show that the mapping $t : F^2 \to F^3$ given by
$$t(a,b) = (a,b,0) \ \forall a,b \in F$$
is a linear transformation.

2. Let $t : \mathbb{R}^3 \to \mathbb{R}^3$ be a mapping defined by
$$t(x,y,z) = (x,y,0) \ \forall x,y,z \in \mathbb{R}$$
Show that t is a linear transformation.

3. Let $V = \mathbb{R}[x]$ be the vector space of all polynomials over field $\mathbb{R}$ and let $D : V \to V$ be the mapping associating each polynomial $f(x)$ to its derivative $\frac{d}{dx}(f(x))$. Show that D is a linear transformation.

4. Let V be a real vector space of all continuous functions from $\mathbb{R}$ into itself. Show that the mapping $T : V \to V$ given by
$$T[f(x)] = \int_0^x f(t)dt \ \forall f(x) \in V, x \in \mathbb{R}$$
is a linear transformation from V to itself.

5. Let C be the vector space of all complex numbers over the field of complex numbers and let $t : C \to C$ be a mapping given by
$$t(x+iy) = x \ \forall x+iy \in \mathbb{C}$$
Show that t is not a linear transformation.

6. Show that the mapping $t : \mathbb{R}^2 \to \mathbb{R}^2$ given by
$$t(x,y) = (x+y,x) \ \forall x,y \in \mathbb{R}$$
is a linear transformation.

7. Show that each of the following mappings is a linear transformation.

   (i) $t : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

   $$t(x, y) = (ax + by, cx + dy)\ \forall a, b, c, d \in \mathbb{R}$$

   (ii) $t : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

   $$t(x, y, z) = (x + y + z, 2x - 3y + 4z)$$

   (iii) $t : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

   $$t(x, y, z) = (x + 2y - 3z, 4x - 5y + 6z)$$

8. Show that the following mappings are not linear transformations.

   (i) $t : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

   $$t(x, y) = (x + 1, y + 2)$$

   (ii) $t : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

   $$t(x, y) = (x^2, y^2)$$

   (iii) $t : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

   $$t(x, y, z) = (x + 1, y + z)$$

   (iv) $t : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

   $$t(x, y) = (xy, y)$$

9. Let $V = \mathbb{R}^{n\times n}$ be a vector space of $n \times n$ matrices and let M be a fixed non-null matrix in V. Which of the followings is a linear transformation?

   (i) $t(\underline{\underline{A}}) = \underline{\underline{M}}\ \underline{\underline{A}}$

   (ii) $t(\underline{\underline{A}}) = \underline{\underline{M}}\ \underline{\underline{A}} + \underline{\underline{A}}\ \underline{\underline{M}}$

   (ii) $t(\underline{\underline{A}}) = \underline{\underline{A}}\ \underline{\underline{M}} - \underline{\underline{M}}\ \underline{\underline{A}}$

   (iv) $t(\underline{\underline{A}}) = \underline{\underline{M}} + \underline{\underline{A}}$

   where $t : V \to V$.

10. Verify whether the operator L is linear where

    (i)

$$L = \frac{\partial}{\partial t} - \alpha \frac{\partial^2}{\partial z^2} + \beta$$

    (ii)

$$L = \frac{d^2}{\partial x^2} - m^2$$

11. If L is a linear operator then show that $L^n, n \in \mathbb{I}^+$ is also a linear operator.

12. Write the n^{th} order ordinary differential equation given below in operator form and identify whether the operator is linear.

$$a_n \frac{d^n u}{dx^x} + a_{n-1} \frac{d^{n-1} u}{dx^{n-1}} + \cdots + a_0 = f(x)$$

13. Let $t : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation for which $t(1,2) = (2,3)$ and $t(0,1) = (1,4)$. Find the formula for t and find $t(x,y)$.

14. Let $B = \{(-1,0,1), (0,1,-1), (1,-1,1)$ be a basis for $\mathbb{R}^3(\mathbb{R})$ and $t : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation such that

$$t(-1,0,1) = (1,0,0)$$
$$t(0,1,-1) = (0,1,0)$$
$$t(1,-1,1) = (0,0,1)$$

    Find the formula to compute $t(1,-2,3)$ from $t(x,y,z)$.

15. Let $t : \mathbb{R}^3 \to \mathbb{R}^3$ be such that $t(1,2,3) = (1,0,0)$, $t(1,2,0) = (0,1,0)$ and $t(1,-1,0) = (0,0,1)$. Find $t(a,b,c)\ \forall (a,b,c) \in \mathbb{R}^3$.

16. Let $t : \mathbb{R}^4 \to \mathbb{R}^3$ be a linear transformation defined by

$$t(x_1, x_2, x_3, x_4) = (x_1 - x_2 + x_3 + x_4, x_1 + 2x_3 - x_4, x_1 + x_2 + 3x_3 - 3x_4)$$

    Find a basis and dimension of (i) image of t, (ii) kernel of t.

17. Let $t : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation defined by

$$t(x, y, z) = (x + 2y - z, y + z, x + y - 2z)$$

Find a basis and dimension of (i) image of t, (ii) kernel of t.

18. Repeat the above for the followings.

(i) $t : \mathbb{R}^3 \to \mathbb{R}^3$

$$t(x, y, z) = (x + 2y - 3z, 2x + 5y - 4z, x + 4y + z)$$

(ii) $t : \mathbb{R}^4 \to \mathbb{R}^3$

$$t(x, y, z, u) = (x + 2y + 3z + 4u, 2x + 4y + 7z + 5u, x + 2y + 6z + 5u)$$

(iii) $t : \mathbb{R}^3 \to \mathbb{R}^2$

$$t(x, y, z) = (x + y + z, 2x + 2y + 2z)$$

19. Let $t_A : \mathbb{R}^{3\times 1} \to \mathbb{R}^{3\times 1}$ be a linear transformation given by $t_A(\underline{x}) = \underline{\underline{A}}\ \underline{x}$ where

$$\underline{\underline{A}} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

Determine

(i) basis for range of t_A

(ii) basis for kernel of t_A

(iii) rank and nullity of t_A

20. Let $t_A : \mathbb{R}^{4\times 1} \to \mathbb{R}^{3\times 1}$ be a linear transformation defined as

$$t_A \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 2 & 0 & -1 & 5 \\ 3 & 0 & -2 & 8 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix}$$

(i) Find the basis for the range of t_A.

(ii) Find the basis for the kernel of t_A.

(iii) What are $nullity(t_A)$ and $rank(t_A)$.

(iv) Does $[1 \quad 0 \quad 1 \quad 0]^T \in ker(t_A)$.

(v) Does $[3 \quad 5 \quad 8]^T \in range(t_A)$.

21. Show that the mapping $t : \mathbb{R}^2 \to \mathbb{R}^3$ given by $t(a, b) = (a + b, a - b, b)\ \forall (a, b) \in \mathbb{R}^2$ is a linear transformation. Find the range, rank, kernel and nullity of t.

22. Let $\mathbb{C}$ be the field of complex numbers and let $t : \mathbb{C}^3 \to \mathbb{C}^3$ be a mapping given by

$$t(a, b, c) = (a - b + 2c, 2a + b - c, -a - 2b)$$

Show that t is a linear transformation. Find its kernel.

23. Let $t : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation given by $t(a, b) = (2a - 3b, a + b)\ \forall (a, b) \in \mathbb{R}^2$. Find the matrix of t relative to the bases $B = \{\underline{\mathrm{b}}_1 = (1, 0), \underline{\mathrm{b}}_2 = (0, 1)\}$ and $B' = \{\underline{\mathrm{b}}'_1 = (2, 3), \underline{\mathrm{b}}'_2 = (1, 2)\}$.

24. Find the matrix of the linear transformation $t : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$t(a, b, c) = (2b + c, a - 4b, 3a)$$

relative to the ordered bases given below.
$B = \{\underline{\mathrm{b}}_1 = (1, 1, 1), \underline{\mathrm{b}}_2 = (1, 1, 0), \underline{\mathrm{b}}_3 = (1, 0, 0)\}$

25. B given above is an ordered basis for a linear transformation

$$t(x, y, z) = (3x + 2y - 4z, x - 5y + 3z)$$

Find the matrix of t relative to B.

26. The set $B = \{e^{3t}, te^{3t}, t^2e^{3t}\}$ is an ordered basis of the vector space V of all functions $f : \mathbb{R} \to \mathbb{R}$. Let D be the differential operator defined as d/dt. Find the matrix representation of D relative to the basis B.

27. Let $t : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation given by $t(1, 1) = (3, 7)$ and $t(1, 2) = (5, -4)$. Find the matrix of t relative to the standard basis of $\mathbb{R}^2$.

28. The matrix $\underline{\underline{A}}$ on $\mathbb{R}$ defines a linear transformation $t_A : \mathbb{R}^3 \to \mathbb{R}^3$ by the rule $t_A(\underline{x}) = \underline{\underline{A}}\ \underline{x}$. Find the matrix representing t_A relative to the basis $B = \{x_1, x_2, x_3\}$ where

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

29. Repeat question (26) with the following sets.

    (i) $B = \{e^t, e^{2t}, te^{2t}\}$

    (ii) $B = \{1, t, sin(3t), cos(3t)\}$

    (iii) $B = \{e^{5t}, \alpha te^{5t}, \beta t^2 e^{5t}\}$

30. Consider the following bases of $\mathbb{R}^2$.

$$B = \{\underline{b}_1 = (1,0), \underline{b}_2 = (0,1)\}$$
$$C = \{\underline{c}_1 = (1,3), \underline{c}_2 = (1,4)\}$$

    (i) Determine the change-of-basis matrix $\underline{\underline{P}}$ from B to C.

    (ii) Determine the change-of-basis matrix $\underline{\underline{Q}}$ from C to B.

31. Repeat the above problem with

$$B = \{\underline{b}_1 = (1,0,0), \underline{b}_2 = (0,1,0), \underline{b}_3 = (0,0,1)\}$$
$$C = \{\underline{c}_1 = (1,0,1), \underline{c}_2 = (2,1,2), \underline{c}_3 = (1,2,2)\}$$

32. Consider the linear transformation $t : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $t(x,y) = (5x - y, 2x + y)$ and the following bases in $\mathbb{R}^2$.

$$B = \{\underline{b}_1 = (1,0), \underline{b}_2 = (0,1)\}$$
$$C = \{\underline{c}_1 = (1,4), \underline{c}_2 = (2,7)\}$$

    (i) Find the matrix $\underline{\underline{P}}$ representing t relative to the basis B.

    (ii) Find the matrix $\underline{\underline{Q}}$ representing t relative to the basis C.

    (iii) Find the change-of-basis matrix $\underline{\underline{R}}$ from B to C.

    (iv) Find the change-of-basis matrix $\underline{\underline{S}}$ from C to B.

33. Repeat the above problem with the linear transformation $t : \mathbb{R}^3 \to \mathbb{R}^3$

$$t(x, y, z) = (3x + z, -2x + y, -x + 2y + 4z)$$

$$B = \{\underline{\mathrm{b}}_1 = (1, 0, 0), \underline{\mathrm{b}}_2 = (0, 1, 0), \underline{\mathrm{b}}_3 = (0, 0, 1)\}$$

$$C = \{\underline{\mathrm{c}}_1 = (1, 0, 1), \underline{\mathrm{c}}_2 = (-1, 2, 1), \underline{\mathrm{c}}_3 = (2, 1, 1)\}$$

Chapter 6

Inner product spaces

Let V be a real vector space. An inner product on V is a function $<,>: V \times V \to \mathbb{R}$ which assigns each ordered pair $(\underline{u}, \underline{v}) \in V \times V$ to a real number $< \underline{u}, \underline{v} >$ in such a way that the following axioms are satisfied.

(i) Linearity:

$$< a\underline{u}_1 + b\underline{u}_2, \underline{v} >= a <\underline{u}_1, \underline{v} > +b < \underline{u}_2, \underline{v} > \tag{6.1}$$

$$\forall \underline{u}_1, \underline{u}_2, \underline{v} \in V, a, b \in \mathbb{R} \tag{6.2}$$

(ii) Symmetry:

$$< \underline{u}, \underline{v} >=< \underline{v}, \underline{u} > \tag{6.3}$$

(iii) Positive definiteness:

$$< \underline{v}, \underline{v} > \geq 0 \ \forall \underline{v} \in V \tag{6.4}$$

$$< \underline{v}, \underline{v} > = 0 \text{ iff } \underline{v} = \underline{0} \tag{6.5}$$

A vector space equipped with an inner product is called an inner product space. The above conditions have been laid down for real vector spaces. As an extension, the most generalized conditions can be written as given below.

(i) Linearity:

$$< \underline{u}_1 + \underline{u}_2, \underline{v} > =< \underline{u}_1, \underline{v} > + < \underline{u}_2, \underline{v} > \tag{6.6}$$

$$< \alpha\underline{u}, \underline{v} > = \alpha < \underline{u}, \underline{v} > \tag{6.7}$$

(ii) Symmetry:

$$< \underline{u}, \underline{v} >= \overline{< \underline{v}, \underline{u} >} \tag{6.8}$$

(iii) Positive definiteness:

$$< \underline{v}, \underline{v} > \geq 0 \ \forall \underline{v} \in V \tag{6.9}$$

$$< \underline{v}, \underline{v} > = 0 \text{ iff } \underline{v} = \underline{0} \tag{6.10}$$

where overbar represents the complex conjugate. From linearity and symmetry, the following property follows.

$$< \underline{u}, \alpha\underline{v} >= \overline{\alpha} < \underline{u}, \underline{v} > \tag{6.11}$$

There can be several functions which can satisfy the conditions given above for an inner product. A *standard* inner product is defined as follows.

$$< \underline{u}, \underline{v} > = \sum_{i=1}^{n} u_i \overline{v}_i \tag{6.12}$$

$$< f, g > = \int_a^b f(x)\overline{g}(x)dx \tag{6.13}$$

Two vectors $\underline{u}$ and $\underline{v}$ are said to be orthogonal if their inner product is zero.

$$< \underline{u}, \underline{v} >= 0 \tag{6.14}$$

A collection of vectors is said to be an orthogonal set if

$$< \underline{v}_i, \underline{v}_j >= 0 \ \forall i \neq j \tag{6.15}$$

The above set is said to be orthonormal if

$$< \underline{v}_i, \underline{v}_j >= \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \tag{6.16}$$

Problems

1. Consider the vector space $\mathbb{R}^n$. Prove that $\mathbb{R}^n$ is an inner product space with the inner product defined by
$$<\underline{u}, \underline{v}>= u_1v_1 + u_2v_2 + \cdots + u_nv_n$$
where $\underline{u} = (u_1, u_2, \ldots, u_n)$ and $\underline{v} = (v_1, v_2, \ldots, v_n)$.

2. Consider the vector space $C[a, b]$ of all continuous functions defined on the interval $[a, b]$. Prove that the following operation makes it an inner product space.
$$< f, g >= \int_a^b f(t)\overline{g}(t)dt, \ \ f(t), g(t) \in C[a, b]$$

3. Identify whether $\mathbb{R}^2$ is an inner product space with an inner product defined by
$$<\underline{u}, \underline{v}>= a_1b_1 - a_2b_1 - a_1b_2 + 2a_2b_2$$
$$\underline{u} = (a_1, a_2), \underline{v} = (b_1, b_2) \in \mathbb{R}^2$$

4. Let V be a real vector space. Show that the sum of two inner products on V is an inner product on V. Is the same true for the difference of two inner product? What about positive multiples of an inner product?

5. Find the value of k so that the following is an inner product on $\mathbb{R}^2$.
$$<\underline{u}, \underline{v}>= a_1b_1 - 3a_1b_2 - 3a_2b_1 + ka_2b_2$$
$$\underline{u} = (a_1, a_2), \underline{v} = (b_1, b_2) \in \mathbb{R}^2$$

6. Let $<,>$ be the standard inner product on $\mathbb{R}^2$. If $\underline{u} = [1 \quad 2]^T$ and $\underline{v} = [-1 \quad 1]^T \in \mathbb{R}^2$ then find $\underline{w} \in \mathbb{R}^2$ satisfying
$$<\underline{v}, \underline{w}>= 3 \text{ and } <\underline{u}, \underline{w}>= -1$$

7. For each of the followings, determine whether the operation $<,>$ makes the space an inner product space.

   (i) $<\underline{u}, \underline{v}>= u_1v_1 - 2u_1v_2 - 2u_2v_1 + 5u_2v_2$

(ii) $<\underline{u},\underline{v}>= u_1v_1 - u_1v_2 - u_2v_1 + 3u_2v_2$

(iii) $<\underline{u},\underline{v}>= u_1^2 - 2u_1v_2 - 2u_2v_1 + v_1^2$

(iv) $<\underline{u},\underline{v}>= 2u_1v_1 + 5u_2v_2$

(v) $<\underline{u},\underline{v}>= u_1v_2u_3 + u_2v_1u_3$

8. Determine the value of λ so that the following is an inner product on $\mathbb{R}^2$.

$$<\underline{u},\underline{v}>= u_1v_1 + 2u_1v_2 + 2u_2v_1 + \lambda u_2v_2$$

9. Let $<,>$ be the standard inner product on $\mathbb{R}^2$. If $\underline{u} = [1 \quad 3]^T$ and $\underline{v} = [2 \quad 1]^T \in \mathbb{R}^2$ such that

$$<\underline{w},\underline{u}>= 3 <\underline{w},\underline{v}>= -1$$

then determine $\underline{w}$.

10. For any $\underline{u} = [u_1 \quad u_2]^T$ and $\underline{v} = [v_1 \quad v_2]^T \in \mathbb{R}^2$, the following operation is defined.

$$<\underline{u},\underline{v}>= \begin{bmatrix} u_1 & v_1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} u_2 \\ v_2 \end{bmatrix}$$

Determine whether $\mathbb{R}^2$ is an inner product space with the inner product defined by the above operation.

11. If $f(x) = x$ and $g(x) = e^{-ix}$ then show that $< f(x), g(x) >= \overline{< g(x), f(x) >}$.

12. Let $\underline{\underline{A}} = [a_{ij}]$ be a 2×2 matrix with real entries. For $\underline{x},\underline{y} \in \mathbb{R}^{2\times 1}$, let $f_A(\underline{x},\underline{y}) = \underline{y}^T \ \underline{\underline{A}} \ \underline{x}$. Show that f_A is an inner product space on $\mathbb{R}^{2\times 1}$ iff $\underline{\underline{A}}^T = \underline{\underline{A}}, a_{11} > 0, a_{22} > 0$ and $|\underline{\underline{A}}| > 0$.

13. Let $C[-\pi,\pi]$ be the inner product space of all continuous functions defined on $[-\pi,\pi]$ with the inner product as the standard inner product. Verify whether $sin(t)$ and $cos(t)$ are orthogonal in such a space.

14. Determine a non-zero vector in $\mathbb{R}^3$ that is orthogonal to the vectors $\underline{v}_1 = [1 \ 1 \ 2]^T$, $\underline{v}_2 = [2 \ 1 \ 3]^T$ and $\underline{v}_3 = [1 \ 2 \ 3]^T$.

15. Let S be the set consisting of vectors $\underline{v}_1 = [1\ \ 2\ \ 1]^T$, $\underline{v}_2 = [2\ \ 1\ \ \ -4]^T$ and $\underline{v}_3 = [3\ \ -2\ \ 1]^T$ with standard inner product defined on $\mathbb{R}^3$. Verify whether S is orthogonal and a basis of $\mathbb{R}^3$.

16. Let V be an inner product space and $S = \{\underline{v}_1, \underline{v}_2 \ldots, \underline{v}_n\}$ be an orthogonal set of vectors in V. Show that $\{\lambda\underline{v}_1, \lambda\underline{v}_2 \ldots, \lambda\underline{v}_n\}$ is also an orthogonal set for any scalar λ.

17. Let V be the vector space of all polynomials over $\mathbb{R}$ of degree ≤ 2 with the inner product defined as $< f, g >= \int_0^1 f(t)g(t)dt$. Find a basis of the subspace orthogonal to the polynomial $\phi(t) = 2t + 1$.

18. Consider the inner product space $\mathbb{R}^4$ with standard inner product defined. Let four vectors in $\mathbb{R}^4$ be as given below.

$$\underline{v}_1 = [1 \quad 1 \quad 0 \quad -1]^T, \underline{v}_2 = [1 \quad 2 \quad 1 \quad 3]^T$$
$$\underline{v}_2 = [1 \quad 1 \quad -9 \quad 2]^T, \underline{v}_4 = [16 \quad -13 \quad 1 \quad 3]^T$$

(i) Do the above vectors form a basis in $\mathbb{R}^4$?

(ii) Do they form an orthogonal basis?

(iii) Do they form an orthonormal basis?

(iv) Express an arbitarary vector $\underline{v} = [a \quad b \quad c \quad d]^T$ in terms of the above vectors.

19. Let $f(x)$ belong to a space of continuous functions with the standard inner product defined. If $f(x)$ is expressed as an infinite series as follows.

$$f(x) = \sum_{n=1}^{\infty} a_n sin(n\pi x)$$

then derive an expression for the coefficients of the expansion.

20. Determine whether the following functions

$$\phi_n(x) = exp(2\pi inx), n \in \mathbb{I}, 0 \leq x \leq 1$$

form an orthonormal set.

21. Consider a piecewise continuous function $f(x)$ defined on the interval $[-c, c]$ with a period $2c$. The function is to be represented as

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n cos \frac{n\pi x}{c} + b_n sin \frac{n\pi x}{c} \right)$$

Determine the expressions for a_n and b_n.

22. Consider the following function.

$$T(x,t) = \sum_{n=1}^{\infty} a_n exp\left(\frac{-\alpha n^2 \pi^2 t}{c^2} \right) sin\left(\frac{n\pi x}{c} \right)$$

If $T(x, 0) = f(x)$, obtain an expression for a_n.

23. Apply Gram-Schmidt orthogonalization to the basis $B = \{[1 \;\; 0 \;\; 1]^T, [1 \;\; 0 \;\; -1]^T, [0 \;\; 3 \;\; 4]^T\}$ of an inner product space $\mathbb{R}^3$ to obtain an orthogonal and an orthonormal basis.

24. Let B be a set of vectors in $\mathbb{R}^3$ with standard inner product defined. For the followings, identify whether B is a basis. Is the basis orthonormal? If not then obtain an orthonormal basis.

(i) $B = \{[2 \quad 0 \quad 1]^T, [3 \quad -1 \quad 5]^T, [0 \quad 4 \quad 2]^T\}$

(ii) $B = \{[1 \quad 0 \quad 0]^T, [1 \quad 1 \quad 0]^T, [1 \quad 1 \quad 1]^T\}$

25. A space of polynomial functions of the form $x^n \; \forall n \in \mathbb{I}^+ \cup 0$ of dimension 3 needs to be expanded using a suitable orthonormal basis. If the inner product in the space is defined as $< p, q >= \int_{-1}^{1} p(x)q(x)dx$ then determine a suitable orthonormal basis for the space.

Chapter 7

Norm and metric spaces

Consider a vector $\underline{v}$ in an inner product space. A scalar d, defined as

$$d = \sqrt{<\underline{v}, \underline{v}>} \tag{7.1}$$

gives the norm of the vector $\underline{v}$. It is also denoted as $||\underline{v}||$.

$$<\underline{v}, \underline{v}> = ||\underline{v}||^2 \tag{7.2}$$

Using the positive definiteness of inner products, we can write

$$<\underline{u} + \alpha\underline{v}><\underline{u} + \alpha\underline{v}> \geq 0 \tag{7.3}$$

for every scalar $\alpha \in \mathbb{R}$.

$$\Longrightarrow <\underline{u}, \underline{u}> + 2\alpha <\underline{u}, \underline{v}> + \alpha^2 <\underline{v}, \underline{v}> \geq 0 \tag{7.4}$$

The above inequality is quadratic in α. We want to know the value of α for which the above equation attains a minima. From the derivative test,

$$\alpha = \frac{-<\underline{u}, \underline{v}>}{<\underline{v}, \underline{v}>} \tag{7.5}$$

$$\Longrightarrow <\underline{u}, \underline{u}> - \frac{<\underline{u}, \underline{v}>^2}{<\underline{v}, \underline{v}>} \geq 0 \tag{7.6}$$

$$\Longrightarrow <\underline{u}, \underline{u}><\underline{v}, \underline{v}> \geq <\underline{u}, \underline{v}>^2 \tag{7.7}$$

The inequality of (7.7) is called Cauchy-Bunyakowski-Schwarz (CBS) inequality. In terms of norms, the inequality can be written as

$$||\underline{u}||||\underline{v}|| \geq | < \underline{u}, \underline{v} > | \tag{7.8}$$

Now consider

$$\begin{aligned} ||\underline{u}+\underline{v}||^2 &=< \underline{u}+\underline{v}, \underline{u}+\underline{v} > \\ &= ||\underline{u}||^2 + < \underline{u}, \underline{v} > + < \underline{v}, \underline{u} > + ||\underline{v}||^2 \\ &= ||\underline{u}||^2 + 2Re < \underline{u}, \underline{v} > + ||\underline{v}||^2 \end{aligned}$$

But following the CBS inequality,

$$Re < \underline{u}, \underline{v} > \leq ||\underline{u}||||\underline{v}||$$

$$\implies ||\underline{u}+\underline{v}||^2 \leq ||\underline{u}||^2 + 2||\underline{u}||||\underline{v}|| + ||\underline{v}||^2 \tag{7.9}$$

$$\implies ||\underline{u}+\underline{v}||^2 \leq (||\underline{u}|| + ||\underline{v}||)^2 \tag{7.10}$$

$$\implies ||\underline{u}+\underline{v}|| \leq ||\underline{u}|| + ||\underline{v}|| \tag{7.11}$$

The inequality given by (7.11) is called the triangle inequality. This inequality can easily be identified as a case of the definition of "distance" between vectors. Norm of a vector can be identified as the distance of that vector from the zero vector. We can now define a space, called the "metric space" if the followings are satisfied.

(i) Positivity

$$d(\underline{u}, \underline{v}) \geq 0 \tag{7.12}$$

(ii) Symmetry

$$d(\underline{u}, \underline{v}) = d(\underline{v}, \underline{u}) \tag{7.13}$$

(iii) Triangle inequality

$$d(\underline{u}, \underline{v}) \leq d(\underline{u}, \underline{w}) + d(\underline{w}, \underline{v}) \tag{7.14}$$

A more general definition of a metric is the p-metric defined as

$$d_p(\underline{u}, \underline{v}) = \left[\sum_{i=1}^{n} |v_i - u_i|^p \right]^{\frac{1}{p}} \tag{7.15}$$

For continuous functions,

$$d_p(f,g) = \left[\int_a^b |f(x) - g(x)|^p dx \right]^{\frac{1}{p}} \tag{7.16}$$

$\mathcal{L}^2$ space:

Consider a space $X[a,b]$ of functions defined in the interval $[a,b]$ such that $f : [a,b] \to \mathbb{C}$ and

$$\int_a^b |f(x)|^2 dx < \infty \tag{7.17}$$

In such a case, the space is called $\mathcal{L}^2$ space.

Consider two functions f and $g \in \mathcal{L}^2(a,b)$. From triangle inequality,

$$||\alpha f + \beta g|| \leq ||\alpha f|| + ||\beta g||$$

$$||\alpha f + \beta g|| \leq |\alpha|||f|| + |\beta|||g||$$

$$\alpha, \beta \in \mathbb{C}$$

Therefore, if $f, g \in \mathcal{L}^2(a,b)$ then all linear combinations in the space also belong to the space.

Problems

1. Prove the following (in)equalities.

   (i) $||\underline{u}+\underline{v}||^2 = ||\underline{u}||^2 + ||\underline{v}||^2$

   (ii) $|\ ||\underline{u}|| - ||\underline{v}||\ | \leq ||\underline{u}-\underline{v}||$

   (iii) $||\underline{u}+\underline{v}||^2 + ||\underline{u}-\underline{v}||^2 = 2(||\underline{u}||^2 + ||\underline{v}||^2)$

   (iv) $||\underline{w}-\underline{u}||^2 + ||\underline{w}-\underline{v}||^2 = \frac{1}{2}||\underline{u}-\underline{v}||^2 + 2||\underline{w}||^2 - \frac{1}{2}||\underline{u}+\underline{v}||^2$

   (v) $||\underline{u}-\underline{v}|| \leq ||\underline{u}-\underline{w}|| + ||\underline{w}-\underline{v}||$

2. For each of the followings, determine whether they belong to $\mathcal{L}^2$ and calculate the norm if it is defined.

   (i) $f(x) = \begin{cases} 1, & 0 \leq x \leq \dfrac{1}{2} \\ 0, & \dfrac{1}{2} \leq x \leq 1 \end{cases}$

   (ii) $f(x) = \frac{1}{\sqrt{x}}, 0 \leq x \leq 1$

   (iii) $f(x) = \frac{1}{\sqrt{x^3}}, 0 \leq x \leq 1$

   (iv) $f(x) = \frac{1}{x}, 1 \leq x \leq \infty$

3. Show that the infinite set of functions $\{1, cos(x), cos(2x) \ldots sin(x), sin(2x) \ldots\}$ is orthogonal in real inner product space $\mathcal{L}^2(-\pi, \pi)$.

4. Show that the infinite set of functions

$$\{e^{inx} : n \in \mathbb{Z}\}$$

   is orthogonal in complex space $\mathcal{L}^2(-\pi, \pi)$.

5. Verify the CBS inequality for the following functions on $[0, 1]$.

$$f(x) = 1$$

$$g(x) = x$$

6. Determine which of the following functions belong to $\mathcal{L}^2(0, \infty)$ and calculate the norm in cases it is defined.

(i) $f(x) = e^{-x}$

(ii) $f(x) = sin(x)$

(iii) $f(x) = cos(x)$

(iv) $f(x) = \frac{1}{1+x}$

7. Determine the real values of α for which x^{α} lies in (a) $\mathcal{L}^2(0,1)$ and (b) $\mathcal{L}^2(1,\infty)$.

8. Define a function $f(x) \in \mathcal{L}^2(-1,1)$ such that $< f(x), x^2 + 1 >= 0$ and $||f(x)|| = 2$.

Chapter 8

Adjoint operators

Consider an operator L defined on an inner product space V and two vectors $\underline{\mathrm{u}}, \underline{\mathrm{v}} \in V$ satisfying the following identity.

$$< L\underline{\mathrm{u}}, \underline{\mathrm{v}} >=< \underline{\mathrm{u}}, L^*\underline{\mathrm{v}} > \tag{8.1}$$

In such a case, L^* is said to be the adjoint of L. If $L^* = L$ then the operator is said to be self-adjoint.

Let us first consider a matrix as an operator.

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\underline{\mathrm{u}} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\underline{\mathrm{v}} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$\underline{\underline{A}}\ \underline{\mathrm{u}} = \begin{bmatrix} a_{11}u_1 + a_{12}u_2 \\ a_{21}u_1 + a_{22}u_2 \end{bmatrix}$$

$$\begin{aligned}
\Longrightarrow < \underline{\underline{A}}\ \underline{\mathrm{u}}, \underline{\mathrm{v}} > &= (a_{11}u_1 + a_{12}u_2)v_1 + (a_{21}u_1 + a_{22}u_2)v_2 \\
&= a_{11}u_1v_1 + a_{21}u_1v_2 + a_{12}u_2v_1 + a_{22}u_2v_2 \\
&= u_1(a_{11}v_1 + a_{21}v_2) + u_2(a_{12}v_1 + a_{22}v_2) \\
&=< \underline{\mathrm{u}}, \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \underline{\mathrm{v}} >
\end{aligned}$$

Clearly, from the previous equations,

$$\underline{\underline{A}}^* = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

It is clear that since in general, $\underline{\underline{A}} \neq \underline{\underline{A}}^*$, the operator given by matrix $\underline{\underline{A}}$ is not self-adjoint.

Now we extend the above discussion to the adjoint of a differential operator. Consider the following differential equation with the associated boundary conditions.

$$\begin{aligned}
&\frac{d^2f}{dx^2} + \alpha f = 0 \\
&\frac{df}{dx}(0) = 0 \\
&\frac{df}{dx}(1) = 0
\end{aligned}$$

Let us denote the differential operator by D.

$$\begin{aligned}
D &= \frac{d^2}{dx^2} + \alpha \\
< Df, g > &= \int_0^1 \left(\frac{d^2f}{dx^2} + \alpha f\right) g dx \\
&= \int_0^1 \left(\frac{d^2f}{dx^2} g\right) dx + \int_0^1 \alpha f g dx \\
&= g\frac{df}{dx}\bigg|_0^1 - \frac{dg}{dx}f\bigg|_0^1 + \int_0^1 f\frac{d^2g}{dx^2}dx + \int_0^1 \alpha f g dx \\
\Longrightarrow < Df, g > &= \int_0^1 f\left(\frac{d^2g}{dx^2} + \alpha g\right)dx + g\frac{df}{dx}\bigg|_0^1 - \frac{dg}{dx}f\bigg|_0^1
\end{aligned}$$

From the boundary conditions on f, we have $\frac{df}{dx}\Big|_0^1 = 0$. Now if we choose the boundary conditions on g such that

$$\begin{aligned}
&\frac{dg}{dx}(0) = 0 \\
&\frac{dg}{dx}(1) = 0
\end{aligned}$$

then

$$< Df, g > = \int_0^1 f\left(\frac{d^2 g}{dx^2} + \alpha g\right) dx \tag{8.2}$$

$$< Df, g > =< f, D^* g > \tag{8.3}$$

Hence, in this case, we have the differential operator given by D which is self-adjoint with the specified boundary conditions.

We showed the above protocol for obtaining the adjoint of an operator when the operator was a matrix or a differential operator. The adjoint of a matrix operator can be used to check the existance and uniqueness of solutions of $\underline{\underline{A}}\ \underline{x} = \underline{b}$.

Fredholm's alternative theorem:

Consider the system of linear equations $\underline{\underline{A}}\ \underline{x} = \underline{b}$. In order to determine the solvability of this system, we need to examine the homogeneous adjoint problem $\underline{\underline{A}}\ \underline{y} = \underline{0}$. According to Fredholm's alternative theorem, $\underline{\underline{A}}\ \underline{x} = \underline{b}$ will have a solution iff

$$< \underline{b}, \underline{y} >= 0\ \forall \underline{y} : \underline{\underline{A}}^*\ \underline{y} = \underline{0} \tag{8.4}$$

Problems

1. Determine the adjoint of the following matrix operators. Check whether they are self-adjoint.

   (i) $\underline{\underline{A}} = \begin{bmatrix} i & 0 \\ i & 1 \end{bmatrix}$

   (ii) $\underline{\underline{A}} = \begin{bmatrix} 1 & -2i \\ 3 & i \end{bmatrix}$

   (iii) $\underline{\underline{A}} = \begin{bmatrix} 3 & 2+i & 4 \\ 2-i & 2 & -i \\ 2 & i & 1 \end{bmatrix}$

   (iv) $\underline{\underline{A}} = \begin{bmatrix} -4+5i & 2+2i & 1-2i \\ 2+2i & -1+8i & -2-2i \\ 4+4i & -2-2i & -4+5i \end{bmatrix}$

2. The momentum operator in Quantum Mechanics is given as $\hat{p} = -i\frac{d}{dx}$. Check if this operator is self-adjoint for a region $[a, b]$.

3. The time-independent Schrödinger equation in 1-D is given as follows.

   $$\frac{-h^2}{8\pi^2 m}\frac{d^2\psi}{dx^2} + \hat{V}\psi = E\psi$$

   As a background study, find out the meaning and significance of all the quantities that appear in the above equation. Check whether the operator in the above equation is (a) linear, (b) self-adjoint.

4. Consider the following operator with homogeneous boundary conditions as shown below.

   $$\hat{L} = \frac{d^2}{dx^2} + 1, x \in [0, \pi]$$
   $$u(0) = u(\pi) = 0$$

   Verify whether $\hat{L}$ is self-adjoint.

5. Consider the following operator $\hat{L}$.

$$\hat{L} = p_2 \frac{d^2}{dx^2} + p_1 \frac{d}{dx} + p_0$$

where $p_i \in \mathbb{R}$. Determine the adjoint of $\hat{L}$ and suitable boundary conditions to make it a self-adjoint operator.

6. Identify the solvability condition and determine the range space using the alternative theorem for the following set of equations.

$$\begin{aligned} x_1 + x_2 + x_3 &= b_1 \\ 2x_1 - x_2 + x_3 &= b_2 \\ x_1 - 2x_2 &= b3 \end{aligned}$$

7. Using the alternative theorem, determine the solvability condition for the following set of equations.

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 3 \\ 2x_1 + x_2 + 6x_3 &= 2 \\ x_1 + 2x_2 + 4x_3 &= -1 \end{aligned}$$

8. Repeat the previous exercise with the following set.

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 - 3x_5 &= 2 \\ 3x_1 + 6x_2 + 4x_3 - x_4 + 2x_5 &= -1 \\ 4x_1 + 8x_2 + 5x_3 + x_4 - x_5 &= 1 \\ -2x_1 - 4x_2 - 3x_3 + 3x_4 - 5x_5 &= 3 \end{aligned}$$

9. Repeat the previous exercise with the following set.

$$\begin{aligned} x_1 - x_2 + 3x_3 + 2x_4 &= 2 \\ 3x_1 + x_2 - x_3 + x_4 &= -3 \\ -x_1 - 3x_2 + 7x_3 + 3x_4 &= 7 \end{aligned}$$

10. Solve all relevant problems of chapter 4 using Fredholm's alternative theorem.

11. Determine the adjoint of the linear transformation $t : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$t(x, y) = (x + 2y, x - y) \; \forall (x, y) \in \mathbb{R}^2$$

12. Determine the adjoint of the linear transformation $t : \mathbb{R}^3 \to \mathbb{R}^3$ given by

    (i) $t(x, y, z) = (x + 2y, 3x - 3z, y)$

    (ii) $t(x, y, z) = (3x + 4y - 5z, 2x - 6y + 7z, -5x + 7y + z)$

13. Show that the product of two self-adjoint operators is also a self-adjoint operator iff the two operators commute.

14. Show that for self-adjoint operators:

    (i) eigenvalues are always real

    (ii) eigenfunctions/eigenvectors are orthogonal

Chapter 9

Eigenvalue problems

Consider a family of two first order ordinary differential equations shown below.

$$\frac{dx}{dt} = ax + by \tag{9.1}$$

$$\frac{dy}{dt} = cx + dy \tag{9.2}$$

The family of equations shown above can be written as a single matrix equation as shown below.

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{9.3}$$

For the above equation, the zero vector $[0 \quad 0]^T$ is always a solution, which is a trivial solution. In order to know whether non-trivial solutions exist for this system, one needs to solve for the homogeneous system.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{9.4}$$

If the above equation is represented as a matrix equation as $\underline{\underline{A}}\ \underline{u} = \underline{0}$, and if $det(\underline{\underline{A}}) \neq 0$ then the trivial solution is the only solution. If $det(\underline{\underline{A}}) = 0$ then we get a number of solutions, all of which lie on a straight line. These are called straight line solutions.

It is desired to determine the general solution of $\underline{u}' = \underline{\underline{A}}\ \underline{u}$ where $\underline{u}'$ shows the derivative with respect to t. If $\underline{v}$ is an eigenvector of $\underline{\underline{A}}$ with the corresponding eigenvalue as λ then by definition, $\underline{\underline{A}}\ \underline{v} = \lambda\underline{v}$. We propose that $\underline{u}(t) = e^{\lambda t}\underline{v}$ is a solution to $\underline{u}' = \underline{\underline{A}}\ \underline{u}$. To test this, we compute the derivative of the test solution.

$$\begin{aligned}
\underline{u}'(t) &= \frac{d}{dt}(e^{\lambda t}\underline{v}) \\
&= \lambda e^{\lambda t}\underline{v} \\
&= e^{\lambda t}(\lambda \underline{v}) \\
&= e^{\lambda t}(\underline{\underline{A}}\ \underline{v}) \\
&= \underline{\underline{A}}(e^{\lambda t}\ \underline{v})
\end{aligned}$$

The above analysis shows that $e^{\lambda t}\ \underline{v}$ is a solution to $\underline{u}' = \underline{\underline{A}}\ \underline{u}$.

We can now generalize the above analysis to an n-dimensional system. If the coefficient matrix $\underline{\underline{A}}$ has n distinct eigenvalues then following the principle of linearity, the general solution of the system is given as follows.

$$\underline{u}(t) = \sum_{i=1}^{n} c_i e^{\lambda_i t}\underline{v}_i \tag{9.5}$$

The system of equations considered previously was a homogeneous system of the form $\underline{u}' = \underline{\underline{A}}\ \underline{u}$. Now we consider the solutions of non-homogeneous initial value problems of the form $\underline{u}' = \underline{\underline{A}}\ \underline{u} + \underline{b}(t)$ with an initial condition $\underline{u}(\underline{0}) = \underline{u}_0$. For solution of such systems, we make use of *similarity transformation.*

Similar matrices:

If $\underline{\underline{P}}$ is any non-singular matrix such that $\underline{\underline{P}}^{-1}\ \underline{\underline{A}}\ \underline{\underline{P}} = \underline{\underline{B}}$ then $\underline{\underline{A}}$ and $\underline{\underline{B}}$ are said to be similar matrices. For similar matrices $\underline{\underline{A}}$ and $\underline{\underline{B}}$ consider the following operations.

$$\begin{aligned}
\underline{\underline{B}} &= \underline{\underline{P}}^{-1}\ \underline{\underline{A}}\ \underline{\underline{P}} \\
\implies \underline{\underline{B}}\ \underline{\underline{P}}^{-1} &= \underline{\underline{P}}^{-1}\ \underline{\underline{A}}\ (\underline{\underline{P}}\ \underline{\underline{P}}^{-1}) \\
\implies \underline{\underline{B}}\ \underline{\underline{P}}^{-1} &= \underline{\underline{P}}^{-1}\ \underline{\underline{A}} \\
\implies \underline{\underline{B}}\ \underline{\underline{P}}^{-1}\ \underline{u} &= \underline{\underline{P}}^{-1}\ \underline{\underline{A}}\ \underline{u}
\end{aligned}$$

If $\underline{u}$ is an eigenvector of $\underline{\underline{A}}$ then

$$\underline{\underline{B}}\ \underline{\underline{P}}^{-1}\ \underline{u} = \underline{\underline{P}}^{-1}\ (\lambda\underline{u})$$
$$\Longrightarrow\ \underline{\underline{B}}\ (\underline{\underline{P}}^{-1}\ \underline{u}) = \lambda(\underline{\underline{P}}^{-1}\ \underline{u})$$
$$\Longrightarrow\ \underline{\underline{B}}\ \underline{v} = \lambda\underline{v}$$

where $\underline{v} = \underline{\underline{P}}^{-1}\ \underline{u}$. The above equation means that if λ and $\underline{u}$ are the eigenvalue and eigenvector of $\underline{\underline{A}}$, respectively, then λ will also be the eigenvalue of the similar matrix $\underline{\underline{B}}$ with the corresponding eigenvector as $\underline{\underline{P}}^{-1}\ \underline{u}$.

Diagonalization of matrices:
For similarity transformation, $\underline{\underline{B}} = \underline{\underline{P}}^{-1}\ \underline{\underline{A}}\ \underline{\underline{P}}$. Consider a matrix $\underline{\underline{P}}$ whose columns are made up of eigenvectors of $\underline{\underline{A}}$.

$$\begin{aligned}
\underline{\underline{A}}\ \underline{\underline{P}} &= \underline{\underline{A}}[\underline{v}_1 \mid \underline{v}_2 \mid \ \ldots \ \mid \underline{v}_n] \\
&= [\underline{\underline{A}}\ \underline{v}_1 \mid \underline{\underline{A}}\ \underline{v}_2 \mid \ldots \mid \underline{\underline{A}}\ \underline{v}_n] \\
&= [\lambda\ \underline{v}_1 \mid \lambda\ \underline{v}_2 \mid \ldots \mid \lambda\ \underline{v}_n] \\
&= \underline{\underline{P}}\ \underline{\underline{\Lambda}}
\end{aligned}$$

where

$$\underline{\underline{\Lambda}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ . & & & & \\ . & & & & \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix} \tag{9.6}$$

Hence, an $n \times n$ matrix can be diagonalized if it has n linearly independent eigenvectors. This property is of interest, as is shown below.

Power of matrices: n^{th} power of a matrix can be easily determined using similarity transformation. As seen before, when the matrix $\underline{\underline{P}}$ is made of the eigenvectors of $\underline{\underline{A}}$ then

$$
\begin{aligned}
& \underline{\underline{P}}^{-1}\,\underline{\underline{A}}\,\underline{\underline{P}} = \underline{\underline{\Lambda}} \\
\Longrightarrow\ & \underline{\underline{A}} = \underline{\underline{P}}\,\underline{\underline{\Lambda}}\,\underline{\underline{P}}^{-1} \\
\Longrightarrow\ & \underline{\underline{A}}^{n} = (\underline{\underline{P}}\,\underline{\underline{\Lambda}}\,\underline{\underline{P}}^{-1})^{n} \\
\Longrightarrow\ & \underline{\underline{A}}^{n} = (\underline{\underline{P}}\,\underline{\underline{\Lambda}}\,\underline{\underline{P}}^{-1})(\underline{\underline{P}}\,\underline{\underline{\Lambda}}\,\underline{\underline{P}}^{-1})\ldots(\underline{\underline{P}}\,\underline{\underline{\Lambda}}\,\underline{\underline{P}}^{-1}) \\
\Longrightarrow\ & A^{n} = \underline{\underline{P}}\,\underline{\underline{\Lambda}}^{n}\,\underline{\underline{P}}^{-1}
\end{aligned}
$$

Inverse of a matrix:

Following the above method, it is easy to determine the inverse of a matrix as shown below.

$$
\begin{aligned}
& \underline{\underline{A}} = \underline{\underline{P}}\,\underline{\underline{\Lambda}}\,\underline{\underline{P}}^{-1} \\
\Longrightarrow\ & \underline{\underline{A}}^{-1} = (\underline{\underline{P}}\,\underline{\underline{\Lambda}}\,\underline{\underline{P}}^{-1})^{-1} \\
\Longrightarrow\ & \underline{\underline{A}}^{-1} = \underline{\underline{P}}\,\underline{\underline{\Lambda}}^{-1}\,\underline{\underline{P}}^{-1}
\end{aligned}
$$

If an $n \times n$ matrix does not have n linearly independent eigenvectors then there exists a non-singular matrix $\underline{\underline{P}}$ such that

$$
\underline{\underline{P}}^{-1}\,\underline{\underline{A}}\,\underline{\underline{P}} = \underline{\underline{J}} \tag{9.7}
$$

where $\underline{\underline{J}}$ is called the *Jordan* matrix and has the following form.

$$
\underline{\underline{J}} = \begin{bmatrix} \underline{\underline{J}}_1 & 0 & 0 & \ldots & 0 \\ 0 & \underline{\underline{J}}_2 & 0 & \ldots & 0 \\ . & & & & \\ . & & & & \\ 0 & 0 & 0 & \ldots & \underline{\underline{J}}_n \end{bmatrix} \tag{9.8}
$$

where $\underline{\underline{J}}_i$'s are called Jordan blocks in which eigenvalues appear on the diagonal, 1's are on the first superdiagonal and the rest of the elements are all zero. As an example, a 3×3 Jordan block can be written as follows.

$$
\underline{\underline{J}} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}
$$

If an $n \times n$ matrix has k linearly independent eigenvectors then $\underline{\underline{P}}$ is constructed from k eigenvectors and $n-k$ generalized eigenvectors. For a 3×3 matrix, the following possibilities exist.

Only one eigenvector $\implies$ one Jordan block

$$\underline{\underline{J}} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

Two eigenvectors $\implies$ two Jordan blocks

$$\underline{\underline{J}} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

or

$$\underline{\underline{J}} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

Three eigenvectors $\implies$ three Jordan blocks

$$\underline{\underline{J}} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

Hence, the number of Jordan blocks equals the number of linearly independent eigenvectors of the matrix.

Now we look into the procedure of determining the generalized eigenvectors. We take the case of a 3×3 matrix. Let us first consider the case of a system with only one eigenvector available. The following holds true.

$$\underline{\underline{P}} = [\underline{\mathrm{v}} \mid \underline{\mathrm{q}}_1 \mid \underline{\mathrm{q}}_2]$$

where $\underline{\mathrm{v}}$ is the eigenvector and $\underline{\mathrm{q}}_i$'s are the generalized eigenvectors.

$$\underline{\underline{A}}\ \underline{\underline{P}} = [\underline{\underline{A}}\ \underline{\mathrm{v}} \mid \underline{\underline{A}}\ \underline{\mathrm{q}}_1 \mid \underline{\underline{A}}\ \underline{\mathrm{q}}_2]$$

$$\implies \underline{\underline{A}}\ \underline{\underline{P}} = [\lambda\ \underline{\mathrm{v}} \mid \underline{\underline{A}}\ \underline{\mathrm{q}}_1 \mid \underline{\underline{A}}\ \underline{\mathrm{q}}_2]$$

$$\underline{\underline{P}}\ \underline{\underline{J}} = [\underline{v} \mid \underline{q}_1 \mid \underline{q}_2] \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$\Longrightarrow \underline{\underline{P}}\ \underline{\underline{J}} = [\lambda\ \underline{v} \mid \underline{v} + \lambda\ \underline{q}_1 \mid \underline{q}_1 + \lambda\ \underline{q}_2]$$

For $\underline{\underline{J}} = \underline{\underline{P}}^{-1}\ \underline{\underline{A}}\ \underline{\underline{P}}$, we equate $\underline{\underline{A}}\ \underline{\underline{P}} = \underline{\underline{P}}\ \underline{\underline{J}}$.

$$[\lambda\ \underline{v} \mid \underline{\underline{A}}\ \underline{q}_1 \mid \underline{\underline{A}}\ \underline{q}_2] = [\lambda\ \underline{v} \mid \underline{v} + \lambda\ \underline{q}_1 \mid \underline{q}_1 + \lambda\ \underline{q}_2]$$

$$\Longrightarrow\ \underline{\underline{A}}\ \underline{q}_1 = \underline{v} + \lambda \underline{q}_1 \text{ and } \underline{\underline{A}}\ \underline{q}_2 = \underline{q}_1 + \lambda \underline{q}_2$$

$$\Longrightarrow\ (\underline{\underline{A}} - \lambda \underline{\underline{I}})\underline{q}_1 = \underline{v} \text{ and } (\underline{\underline{A}} - \lambda \underline{\underline{I}})\underline{q}_2 = \underline{q}_1$$

Hence, $\underline{q}_1$ and $\underline{q}_2$ can be obtained as the solutions of the above non-homogeneous problem. Let us now consider the case where two eigenvectors are available.

$$\underline{\underline{A}}\ \underline{\underline{P}} = [\underline{\underline{A}}\ \underline{v}_1 \mid \underline{\underline{A}}\ \underline{v}_2 \mid \underline{\underline{A}}\ \underline{q}]$$

$$\Longrightarrow \underline{\underline{A}}\ \underline{\underline{P}} = [\lambda\ \underline{v}_1 \mid \lambda\ \underline{v}_2 \mid \underline{\underline{A}}\ \underline{q}]$$

$$\underline{\underline{P}}\ \underline{\underline{J}} = [\underline{v}_1 \mid \underline{v}_2 \mid \underline{q}] \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$\Longrightarrow \underline{\underline{P}}\ \underline{\underline{J}} = [\lambda\ \underline{v}_1 \mid \lambda \underline{v}_2 \mid \underline{v}_2 + \lambda\ \underline{q}]$$

$$\Longrightarrow\ [\lambda\ \underline{v}_1 \mid \lambda\ \underline{v}_2 \mid \underline{\underline{A}}\ \underline{q}] = [\lambda\ \underline{v}_1 \mid \lambda \underline{v}_2 \mid \underline{v}_2 + \lambda\ \underline{q}]$$

$$\Longrightarrow\ (\underline{\underline{A}} - \lambda \underline{\underline{I}})\underline{q} = \underline{v}_2$$

The generalized vector can be obtained by solving the above equation. By exchanging $\underline{v}_1$ and $\underline{v}_2$, another generalized vector can be obtained by solving the following equation.

$$(\underline{\underline{A}} - \lambda \underline{\underline{I}})\underline{q} = \underline{v}_1$$

and further by linearity, the following can be written.

$$(\underline{\underline{A}} - \lambda \underline{\underline{I}})\underline{q} = \alpha \underline{v}_1 + \beta \underline{v}_2 \tag{9.9}$$

Now we use all the concepts developed till now for solving non-homogeneous initial value problems. Given

$$\begin{aligned}
&\frac{d}{dt}\underline{u}(t) = \underline{\underline{A}}\ \underline{u}(t) + \underline{b}(t) \\
\Longrightarrow\ &\frac{d}{dt}(\underline{\underline{P}}^{-1}\underline{u}) = \underline{\underline{P}}^{-1}\ \underline{\underline{A}}\ \underline{u} + \underline{\underline{P}}^{-1}\ \underline{b} \\
\Longrightarrow\ &\frac{d}{dt}(\underline{\underline{P}}^{-1}\underline{u}) = (\underline{\underline{P}}^{-1}\ \underline{\underline{A}}\ \underline{\underline{P}})\ (\underline{\underline{P}}^{-1}\ \underline{u}) + \underline{\underline{P}}^{-1}\ \underline{b}
\end{aligned}$$

We identify $\underline{\underline{P}}^{-1}\ \underline{\underline{A}}\ \underline{\underline{P}} = \underline{\underline{\Lambda}}$ in the above equation. Denoting $\underline{\underline{P}}^{-1}\ \underline{u} = \underline{v}$ and $\underline{\underline{P}}^{-1}\ \underline{b} = \underline{g}$ in the above equation, we can write

$$\frac{d}{dt}\underline{v}(t) = \underline{\underline{\Lambda}}\ \underline{v}(t) + \underline{g}(t) \tag{9.10}$$

The above equation can be solved using the method of integrating factor.

$$\frac{d}{dt}\left(e^{-\underline{\underline{\Lambda}}\,t}\underline{v}\right) = e^{-\underline{\underline{\Lambda}}\,t}\underline{\underline{\Lambda}}\ \underline{v}(t) + e^{-\underline{\underline{\Lambda}}\,t}\underline{g}(t) \tag{9.11}$$

The above equation can be solved using usual method except that we do not know the exponential of a matrix. This can be determined as follows.

$$e^{-\underline{\underline{\Lambda}}\,t} = \underline{\underline{I}} + \underline{\underline{\Lambda}}t + \frac{(\underline{\underline{\Lambda}}t)^2}{2!} + \frac{(\underline{\underline{\Lambda}}t)^3}{3!}\dots \tag{9.12}$$

$$\Longrightarrow\ e^{-\underline{\underline{\Lambda}}\,t} = \begin{bmatrix} e^{\lambda_1 t} & 0 & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & 0 & \dots & 0 \\ . & & & & \\ . & & & & \\ 0 & 0 & 0 & \dots & e^{\lambda_n t} \end{bmatrix} \tag{9.13}$$

Problems

1. Find the general solutions for each of the following problems.

   (i) $\underline{u}' = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \underline{u}$

   (ii) $\underline{u}' = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \underline{u}$

   (ii) $\underline{u}' = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \underline{u}$

   (iv) $\underline{u}' = \begin{bmatrix} 1 & 2 \\ 3 & -3 \end{bmatrix} \underline{u}$

2. Find the general solution of the system

$$\underline{u}' = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \underline{u}$$

   where $bc > 0$

3. For a harmonic oscillator governed by the following equation

$$\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx = 0$$

   determine all the values of b and k for which the system has real and distinct eigenvalues. Find the general solution and the solution which satisfies the initial condition $x(0) = 1$.

4. Consider $\underline{\underline{A}} = \begin{bmatrix} a & 1 \\ 0 & 1 \end{bmatrix}$. Determine the value of the parameter a for which $\underline{\underline{A}}$ has repeated real eigenvalues.

5. For each of the following systems, determine the general solutions.

   (i) $\underline{u}' = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \underline{u}$ (ii) $\underline{u}' = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \underline{u}$

   (iii) $\underline{u}' = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \underline{u}$ (iv) $\underline{u}' = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix} \underline{u}$

(v) $\underline{u}' = \begin{bmatrix} 1 & 1 \\ -1 & -3 \end{bmatrix} \underline{u}$ (vi) $\underline{u}' = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \underline{u}$

6. Determine the general solutions of the following harmonic oscillators.

$$\frac{d^2x}{dt^2} + \frac{dx}{dt} + x = 0$$
$$\frac{d^2x}{dt^2} + 2\frac{dx}{dt} + x = 0$$

7. Consider $\underline{u}' = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \underline{u}$, where $a + d \neq 0$ and $ad - bc \neq 0$. Determine the general solution of the system.

8. Consider the harmonic oscillator described by the following equation.

$$\underline{x}' = \begin{bmatrix} 0 & 1 \\ -k & -b \end{bmatrix} \underline{x}$$

where $b \geq 0$ and $k \geq 0$. For which values of k and b does the system have complex eigenvalues? real and distinct eigenvalues? repeated eigenvalues?

9. Solve all the above relevant problems using similarity tranformation.

10. Consider the following matrices.

$$\underline{\underline{A}} = \begin{bmatrix} 7 & -16 & -8 \\ -16 & 7 & 8 \\ -8 & 8 & -5 \end{bmatrix}$$
$$\underline{\underline{A}} = \begin{bmatrix} -2 & 0 & -2i \\ 0 & 1 & 0 \\ 2i & 0 & -2 \end{bmatrix}$$

(i) Diagonalize the above matrices.

(ii) Determine the square of the above matrice using similarity transformation. Verify using multiplication.

11. Solve the following problem using similarity transformation.

$$\frac{d\underline{u}}{dt} = \underline{\underline{A}}\ \underline{u}$$
$$\underline{u}(0) = [1 \quad 1 \quad 1]^T$$
$$\underline{\underline{A}} = \begin{bmatrix} 5 & -3 & -2 \\ 8 & -5 & 4 \\ -4 & 3 & 3 \end{bmatrix}$$

12. Consider the following matrix.

$$\underline{\underline{A}} = \begin{bmatrix} -2 & 1 \\ -1 & -2 \end{bmatrix}$$

Using similarity transformation, obtain the solution of

$$\frac{d}{dt}\underline{u}(t) = \underline{\underline{A}}\ \underline{u}(t) + \underline{b}(t)$$

where $\underline{b} = [1 \quad 1]^T$ and $\underline{u}(0) = [0 \quad 0]^T$.

13. Using similarity transform, solve

$$\frac{d}{dt}\underline{u}(t) = \underline{\underline{A}}\ \underline{u}(t) + \underline{b}(t)$$

where

$$\underline{\underline{A}} = \begin{bmatrix} -i & i & 0 \\ i & -i & 0 \\ 0 & 0 & -i \end{bmatrix}$$
$$\underline{b} = \begin{bmatrix} \sqrt{2}t \\ \sqrt{2}t \\ e^{-t} \end{bmatrix}$$
$$\underline{u}(0) = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 1 \end{bmatrix}$$

14. Solve the initial value problem

$$\frac{d^2u}{dt^2} + 5\frac{du}{dt} + 6u = e^{-t}$$

with

$$\left(\frac{du}{dt}\right)_0 = (u)_0 = 1$$

.

15. Convert the following initial value problem to a matrix equation and solve using similarity transformation and otherwise.

$$\frac{d^2u}{dt^2} + 3\frac{du}{dt} + 2u = 0; u(0) = 1; \frac{du}{dt}(0) = 3$$

16. Verify that $\{[e^{2t} \quad e^{2t} \quad e^{2t}]^T, [e^{-t} \quad 0 \quad e^{-t}]^T, [-e^{-t} \quad e^{-t} \quad 0]^T\}$ is a solution set to the system

$$\frac{d\underline{x}}{dt} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \underline{x}$$

17. Find the general solution of the following system.

$$\frac{d\underline{x}}{dt} = \begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} \underline{x} + \begin{bmatrix} 2e^t \\ 4e^t \\ -2e^t \end{bmatrix}$$

18. Find the general solution of the following system.

$$\frac{d\underline{x}}{dt} = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \underline{x} + \begin{bmatrix} -t-1 \\ -4t-2 \end{bmatrix}$$

19. Find the general solution of the following system.

$$\frac{d\underline{x}}{dt} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \underline{x} + \begin{bmatrix} -4cos(t) \\ -sin(t) \end{bmatrix}$$

Chapter 10

Sturm-Liouville theory

Consider an operator of the form given below.

$$\hat{L} = \frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + r(x) \tag{10.1}$$

The eigenvalue problem

$$\hat{L}u + \lambda\rho(x)u = 0, x \in (a, b) \tag{10.2}$$

subject to homogeneous boundary conditions

$$\alpha_1 u(a) + \alpha_2 \frac{du}{dx}(a) = 0; |\alpha_1| + |\alpha_2| > 0 \tag{10.3}$$

$$\beta_1 u(b) + \beta_2 \frac{du}{dx}(b) = 0; |\beta_1| + |\beta_2| > 0 \tag{10.4}$$

with α_i and β_i as real coefficients constitute a Sturm-Liouville problem. It can be shown with the method shown previously that the Sturm-Liouville operator is a self-adjoint operator. Hence, under the boundary conditions stated above, the eigenvalues are real and the eigenfunctions are orthogonal. When the interval (a, b) is bounded and $p(x)$ does not vanish on $[a, b]$ then the problem is called a regular Sturm-Liouville problem. Else, the problem is called singular Sturm-Liouville problem. The set of eigenvalues of a regular Sturm-Liouville problem are countably infinite and is a monotonically increasing sequence

$$\lambda_0 < \lambda_1 < \lambda_2 < \ldots \lambda_n < \lambda_{n+1} < \ldots$$

Numerous classes of Sturm-Liouville problems are identified and solutions derived in the following sections

Bessel equation

The Bessel equation is given as follows.

$$x^2\frac{d^2y}{dx^2}+x\frac{dy}{dx}+(x^2-n^2)y=0 \tag{10.5}$$

where n is a non-negative parameter. The solutions to the Bessel equation are called Bessel functions. It can be seen that the equation is an ordinary differential equation with variable coefficients. Such equations can be solved using a technique called series solution due to Frobenius. We first recast the equation to the following form.

$$\frac{d^2y}{dx^2}+\frac{1}{x}\frac{dy}{dx}+(1-\frac{n^2}{x^2})y=0 \tag{10.6}$$

The solution to the equation is of the following form.

$$y=\sum_{r=0}^{\infty}a_rx^{k+r} \tag{10.7}$$

$$\implies \frac{dy}{dx}=\sum_{r=0}^{\infty}(k+r)a_rx^{k+r-1} \tag{10.8}$$

$$\implies \frac{d^2y}{dx^2}=\sum_{r=0}^{\infty}(k+r-1)(k+r)a_rx^{k+r-2} \tag{10.9}$$

Substitution of the assumed solution and the derivatives gives

$$\sum_{r=0}^{\infty}(k+r-1)(k+r)a_rx^{k+r-2}+\sum_{r=0}^{\infty}(k+r)a_rx^{k+r-2}+\sum_{r=0}^{\infty}a_rx^{k+r}-\sum_{r=0}^{\infty}a_rn^2x^{k+r-2}=0$$

$$\implies \sum_{r=0}^{\infty}\left[(k+r+n)(k+r-n)x^{k+r-2}+x^{k+r}\right]=0 \tag{10.10}$$

The above equation is an important result and a major landmark in obtaining solutions of the type discussed here. We compare the coefficients of the above polynomial equation.

$$r=0,\ a_0(k+n)(k-n)=0 \implies k=\pm n$$

$$r=1,\ a_1[(k+1+n)(k+1-n)]=0 \implies a_1[(k+1)^2-n^2]=0 \implies a_1=0$$

We compute the coefficient of a random term, x^{k+r}, for $r=r$.

$$a_{r+2}[(k+r+2+n)(k+r+2-n)]+a_r=0$$

$$\implies a_{r+2} = \frac{-a_r}{(k+r+2+n)(k+r+2-n)} \tag{10.11}$$

The above equation is one of the most important results of this solution procedure and is called the *recurrence* relation. This relation relates different coefficients appearing in the solution. From $a_1 = 0$ and the above recurrence relation, we can write,

$$a_1 = a_3 = a_5 = \cdots = 0$$

For $k = +n$,

$$\begin{aligned}
a_{r+2} &= \frac{-a_r}{(2n+r+2)(r+2)} \\
\implies a_2 &= \frac{-a_0}{(2)(2n+2)} \\
\implies a_2 &= \frac{-a_0}{2^2(n+1)} \\
\implies a_4 &= \frac{-a_2}{(4)(2n+4)} = \frac{-a_2}{2 \cdot 2^2(n+2)} \\
\implies a_4 &= \frac{a_0}{2! \cdot 2^4(n+2)(n+1)} \\
\implies a_6 &= \frac{-a_4}{(6)(2n+6)} = \frac{-a_4}{3 \cdot 2^2(n+3)} \\
\implies a_6 &= \frac{-a_0}{3! \cdot 2^6(n+3)(n+2)(n+1)}
\end{aligned}$$

From the above expressions of the coefficients, we can write the solution to the Bessel equation as follows.

$$y = a_0 x^k \left[1 - \frac{x^2}{2^2(n+1)} + \frac{x^4}{2!2^4(n+2)(n+1)} - \frac{x^6}{3! \cdot 2^6(n+3)(n+2)(n+1)} - \dots \right]$$

Therefore, for $k = +n$, the solution to the Bessel equation can be written as follows.

$$y = a_0 x^n \sum_{r=0}^{\infty} \frac{(-1)^r x^{2r}}{r! 2^{2r}(n+1)(n+2)\dots(n+r)} \tag{10.12}$$

We have determined the solution of the Bessel equation. The only unknown is a_0. For a specific choice of a_0, we get a clean solution of the equation, as given below.

$$a_0 = \frac{1}{2^n \Gamma(n+1)} = \frac{1}{2^n n!} \tag{10.13}$$

$$\implies y = \frac{1}{2^n \Gamma(n+1)} x^n \sum_{r=0}^{\infty} \frac{(-1)^r x^{2r}}{r! 2^{2r}(n+1)(n+2)\dots(n+r)} \tag{10.14}$$

$$\implies y = \sum_{r=0}^{\infty} \frac{(-1)^r x^{2r+n}}{2^{2r+n} r!(r+n)!} \tag{10.15}$$

This series is called Bessel's function of the first kind of order n, and is conventionally denoted s $J_n(x)$. With $k = +n$ and $-n$, the Bessel's functions can be written as follows.

$$J_n(x) = \sum_{r=0}^{\infty} (-1)^r \left(\frac{x}{2}\right)^{2r+1} \frac{1}{r!(r+n)!} \tag{10.16}$$

$$J_{-n}(x) = \sum_{r=0}^{\infty} (-1)^r \left(\frac{x}{2}\right)^{2r+1} \frac{1}{r!(r-n)!} \tag{10.17}$$

The following plots show the first two Bessel functions $J_0(x)$ and $J_1(x)$.

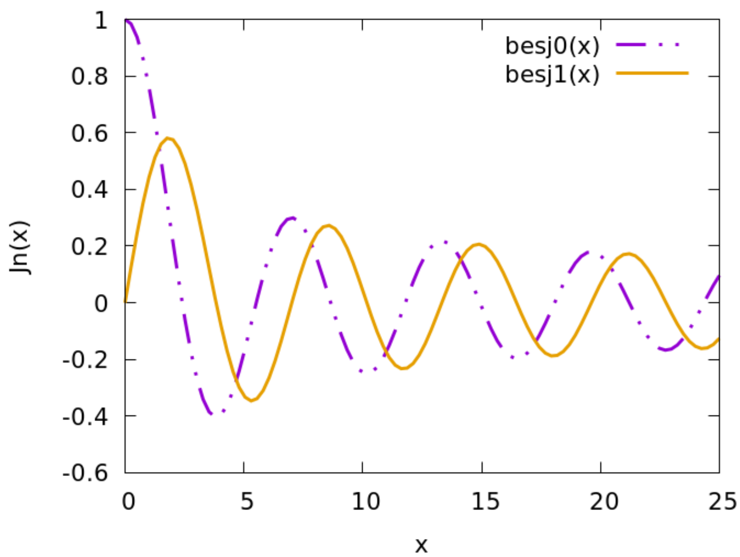


Figure 10.1: Bessel functions of the first kind $J_0(x)$ and $J_1(x)$

In the previous discussion, we used the concept of gamma function. Let us remind ourselves of the gamma function which is defined as follows.

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt \tag{10.18}$$

As seen before, gamma function is related to the factorial as follows.

$$\Gamma(x+1) = x! \tag{10.19}$$

We discussion only about the Bessel functions of the first kind. Bessel functions of the second kind can be derived from Bessel functions of the first kind. This is left as a background study and an exercise.

Legendre equation

The Legendre equation is given as

$$(1-x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + \lambda y = 0; \quad -1 < x < 1 \tag{10.20}$$

Following the Frobenius method, the series solution to the above equation can be represented as an infinite series given below.

$$y = \sum_{r=0}^{\infty} a_r x^{k+r} \tag{10.21}$$

$$\implies \frac{dy}{dx} = \sum_{r=0}^{\infty}(k+r)a_r x^{k+r-1} \tag{10.22}$$

$$\implies \frac{d^2y}{dx^2} = \sum_{r=0}^{\infty}(k+r-1)(k+r)a_r x^{k+r-2} \tag{10.23}$$

Hence, the original differential equation can be cast as follows.

$$\sum_{r=0}^{\infty} a_r(k+r-1)(k+r)x^{k+r-2} - \sum_{r=0}^{\infty} a_r(k+r-1)(k+r)x^{k+r} - 2\sum_{r=0}^{\infty} a_r(k+r)x^{k+r} + \lambda\sum_{r=0}^{\infty} a_r x^{k+r} = 0$$

$$\implies \sum_{r=0}^{\infty} a_r\left[(k+r-1)(k+r)x^{k+r-2} + \Big(\lambda - (k+r)(k+r+1)\Big)x^{k+r}\right] = 0 \tag{10.24}$$

The above equation puts us in a position to write the recurrence relation as given below.

$$a_{r+2} = \left[\frac{(k+r)(k+r+1)-\lambda}{(k+r+1)(k+r+2)}\right]a_r \tag{10.25}$$

Therefore, a knowledge of a_0 and a_1 will yield the complete series through the recurrence relation. When $k=0$ and $\lambda = n(n+1), n \in \mathbb{I}^+ \cup 0$, For such a case

$$a_{n+2} = a_{n+4} = a_{n+6} = \cdots = 0$$

Therefore, the infinite series solution reduces to a polynomial for the above choice of parameters. In such a case,

$$a_{r+2} = \left[\frac{r(r+1)-n(n+1)}{(r+1)(r+2)}\right]a_r$$

$$a_{r+2} = \left[\frac{(r-n)(r+n+1)}{(r+1)(r+2)}\right]a_r \tag{10.26}$$

With arbitrary a_0 and a_1, we get,

$$a_2 = \left[\frac{-n(n+1)}{2!}\right] a_0 \qquad a_4 = \left[\frac{(n-2)n(n+1)(n+3)}{4!}\right] a_0$$

$$.$$

$$a_3 = \left[\frac{-(n-1)(n+2)}{3!}\right] a_1 \qquad a_5 = \left[\frac{(n-3)(n-1)(n+2)(n+4)}{5!}\right] a_1$$

$$.$$

Hence, the solution to the Legendre equation takes the following form.

$$\begin{aligned} y = a_0 &\left[1 - \frac{n(n+1)}{2!} x^2 + \frac{(n-2)n(n+1)(n+3)}{4!} x^4 + \dots\right] \\ + a_1 &\left[x - \frac{(n-1)(n+2)}{3!} x^3 + \frac{(n-3)(n-1)(n+2)(n+4)}{5!} x^5 + \dots\right] \end{aligned}$$

If $u_0(x)$ and $u_1(x)$ are the power series with even only and odd only powers, respectively, then the above solution can be written as follows.

$$y = a_0 u_0(x) + a_1 u_1(x) \tag{10.27}$$

It can be seen from the above analysis that the solution depends upon n. For each n, a pair of linearly independent solutions is obtained.

$$\begin{aligned} n = 0,\ u_0(x) &= 1 \\ u_1(x) &= x + \frac{1}{3} x^3 + \frac{1}{5} x^5 + \dots \\ n = 1,\ u_0(x) &= 1 - x^2 - \frac{1}{3} x^4 + \dots \\ u_1(x) &= x \\ n = 2,\ u_0(x) &= 1 - 3x^2 \\ u_1(x) &= x - \frac{2}{3} x^3 - \frac{1}{5} x^5 + \dots \\ n = 3,\ u_0(x) &= 1 - 6x^2 + 3x^4 + \dots \\ u_1(x) &= x - \frac{5}{3} x^3 \end{aligned}$$

$$.$$

It can be observed that for every n, one of the solutions is a polynomial solution while the other is an infinite series. One is generally interested in the polynomial solutions. Such solutions can be written in the following form.

$$y = a_n x^n + a_{n-2}x^{n-2} + a_{n-4}x^{n-4} \dots \tag{10.28}$$

One specific case is when a_n is given as follows.

$$a_n = \frac{(2n)!}{2^n (n!)^2} = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{n!} \tag{10.29}$$

The resulting polynomial is called Legendre polynomial of degree n, denoted as $P_n(x)$. The general expression for Legendre polynomial of degree n can be written as follows.

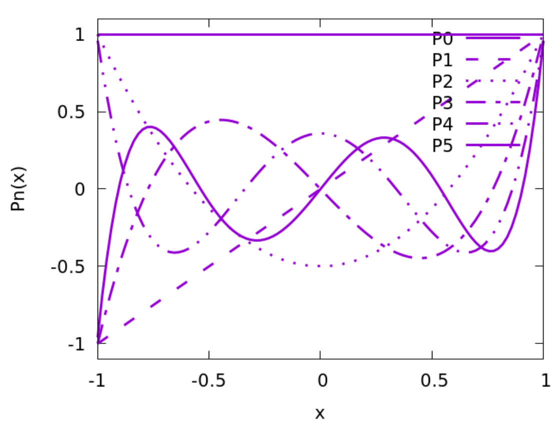
$$P_n(x) = \frac{1}{2^n} \sum_{k=0}^{[n/2]} (-1)^k \left[\frac{(2n-2k)!}{k!(n-k)!(n-2k)!} \right] x^{n-2k} \tag{10.30}$$

where

$$[n/2] = \begin{cases} n/2, & \text{if } n \text{ is even} \\ (n-1)/2, & \text{if } n \text{ is odd} \end{cases}$$

$$\begin{aligned} P_0(x) &= 1 & P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) & P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3) & P_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15x) \end{aligned}$$

The following figure shows the first five Legendre polynomials.



Laguerre equation

Laguerre equation is given as follows.

$$x\frac{d^2y}{dx^2} + (1-x)\frac{dy}{dx} + \lambda y = 0 \tag{10.31}$$

Using the series solution technique, we can write the followings.

$$y = \sum_{r=0}^{\infty} a_r x^{k+r} \tag{10.32}$$

$$\implies \frac{dy}{dx} = \sum_{r=0}^{\infty}(k+r)a_r x^{k+r-1} \tag{10.33}$$

$$\implies \frac{d^2y}{dx^2} = \sum_{r=0}^{\infty}(k+r-1)(k+r)a_r x^{k+r-2} \tag{10.34}$$

Substitution of the above quantities in the original equation gives

$$x\sum_{r=0}^{\infty}(k+r-1)(k+r)a_r x^{k+r-2} + (1-x)\sum_{r=0}^{\infty}(k+r)a_r x^{k+r-1} + \lambda\sum_{r=0}^{\infty} a_r x^{k+r} = 0$$

$$\implies \sum_{r=0}^{\infty}(k+r)^2 a_r x^{k+r-1} - \sum_{r=0}^{\infty}(k+r-\lambda)a_r x^{k+r} = 0 \tag{10.35}$$

The recurrence relation can be obtained from the above equation by collecting the terms of x^{k+r-1}.

$$(k+r)^2 a_r - [(r-1)+k-\lambda]a_{r-1} = 0$$

$$\implies a_r = \left[\frac{k+r-\lambda-1}{(k+r)^2}\right]a_{r-1} \tag{10.36}$$

Hence, with arbitrary $a_0 \neq 0$, we get the complete series. Further,

$$(k+r)^2 a_r = (k+r-\lambda-1)a_{r-1} = 0$$

with $a_{-1} = 0$. Therefore, $a_0 = 0$ or $k = 0$. Since $a_0 \neq 0$, $k = 0$.

$$\implies a_r = \left[\frac{r-\lambda-1}{r^2}\right]a_{r-1} \tag{10.37}$$

$$\implies a_{r+1} = -\left[\frac{r+\lambda}{(r+1)^2}\right]a_r \tag{10.38}$$

Hence, the solution of the Laguerre equation is of the following form.

$$y = \sum_{r=0}^{\infty} a_r x^r \tag{10.39}$$

with the coefficients obtained by the recurrence relation obtained above.

Now consider a case when $r = \lambda$. In such a case, the series terminates to a polynomial solution. This means that further coefficients all become zero ($a_{r+1}, a_{r+2} = \cdots = 0$). Hence, the solution looks like the one given below.

$$y = \sum_{r=0}^{\lambda} a_r x^r = a_r x^r + a_{r-1}x^{r-1} + \cdots + a_1 x + a_0 \tag{10.40}$$

This polynomial solution is called the Laguerre polynomial. From the previous recurrence relation, we can write

$$a_{r-1} = \left[\frac{r^2}{r - \lambda - 1}\right] a_r$$

We can substitute $r = \lambda, \lambda-1, \lambda-2 \ldots$ in the above recurrence relation to obtain the general coefficient of the polynomial solution. The expression for $a_{\lambda-n}$ can be written as follows.

$$a_{\lambda-n} = \left[(-1)^n \frac{(\lambda!)^2}{((\lambda-n)!)^2(n!)}\right] a_\lambda \tag{10.41}$$

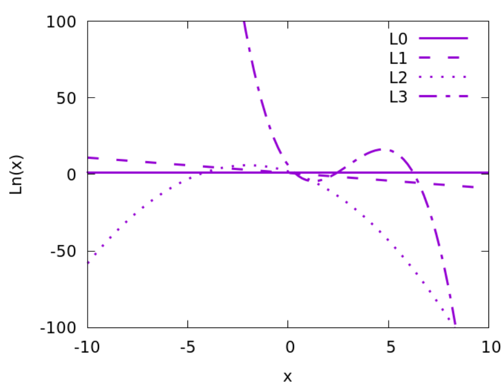
Hence, the solution of the equation as the Laguerre polynomial can be written as

$$L_\lambda(x) = \sum_{n=0}^{\lambda} \left[(-1)^n \frac{(\lambda!)^2}{((\lambda-n)!)^2(n!)} a_\lambda\right] x^{\lambda-n} \tag{10.42}$$

First four Laguerre polynomials are given below.

$$L_0(x) = 1 \qquad\qquad L_1(x) = 1 - x$$

$$L_2(x) = 2 - 4x - x^2 \qquad\qquad L_3(x) = 6 - 18x + 9x^2 - x^3$$



Hermite equation

The following differential equation is called the Hermite equation.

$$\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + 2\lambda y = 0 \tag{10.43}$$

Using the series solution with $k = 0$, we get the followings.

$$y = \sum_{r=0}^{\infty} a_r x^r \tag{10.44}$$

$$\implies \frac{dy}{dx} = \sum_{r=0}^{\infty} r a_r x^{r-1} \tag{10.45}$$

$$\implies \frac{d^2y}{dx^2} = \sum_{r=0}^{\infty} (r-1) r a_r x^{r-2} \tag{10.46}$$

$$\implies \sum_{r=0}^{\infty} (r-1) r a_r x^{r-2} - 2x \sum_{r=0}^{\infty} r a_r x^{r-1} + 2\lambda \sum_{r=0}^{\infty} a_r x^r = 0 \tag{10.47}$$

Proceeding like before, we get the recurrence relation as shown below.

$$a_r = \left[\frac{2(r-\lambda-2)}{r(r-1)}\right] a_{r-2} \tag{10.48}$$

Hence, in the polynomial solutions, only odd and only even powers are observed. With $r = 2n$ (even coefficients) and $r = 2n + 1$ (odd coefficients) we get the followings.

$$\begin{aligned}
a_{2n} &= \left[\frac{2(-\lambda+2n-2)}{2n(2n-1)}\right] a_{2n-2} \\
\implies a_{2n} &= \left[\frac{2(-\lambda+2n-2)}{2n(2n-1)}\right]\left[\frac{2(-\lambda+2n-4)}{(2n-2)(2n-3)}\right] a_{2n-4} \\
\implies a_{2n} &= \left[\frac{2(-\lambda+2n-2)}{2n(2n-1)}\right]\left[\frac{2(-\lambda+2n-4)}{(2n-2)(2n-3)}\right] \dots \left[\frac{2(-\lambda+2-2)}{2(2-1)}\right] a_0 \\
\implies a_{2n} &= \frac{2^n}{(2n)!}(2n-\lambda-2)(2n-\lambda-4)\dots(-\lambda) a_0
\end{aligned}$$

Similarly,

$$\begin{aligned}
a_{2n+1} &= \left[\frac{2(-\lambda+2n-1)}{(2n+1)2n}\right] a_{2n-1} \\
\implies a_{2n} &= -\left[\frac{2(-\lambda+2n+1)}{(2n+1)2n}\right]\left[-\frac{2(-\lambda+2n+3)}{(2n-1)(2n-2)}\right] a_{2n-3} \\
\implies a_{2n} &= \left[\frac{2(-\lambda+2n-1)}{(2n+1)2n}\right]\left[\frac{2(-\lambda+2n-3)}{(2n-1)(2n-2)}\right] \dots \left[\frac{2(-\lambda+2-1)}{(2+1)2}\right] a_1 \\
\implies a_{2n} &= \frac{2^n}{(2n+1)!}(2n-\lambda-1)(2n-\lambda-3)\dots(1-\lambda) a_1
\end{aligned}$$

Hence, the solution to the Hermite equation can be written as the summation of the even and odd series.

$$y = a_0 y_{even} + a_1 y_{odd} \tag{10.49}$$

Like before, we can see that the infinite series truncates to polynomials when λ is an integer (say N) in which case, the recurrence relation can be written as follows.

$$a_r = \left[\frac{2(r-N-2)}{r(r-1)}\right] a_{r-2}$$

It can be seen that the series will terminate when r reaches a values of $N+2$. Hence, we get polynomial solutions. Following the previous analysis, we list out first few Hermite polynomials below and leave it for you to plot and see their nature.

$$H_0(x) = 1$$

$$H_1(x) = 2x$$

$$H_2(x) = 4x^2 - 2$$

$$H_3(x) = 8x^3 - 12x$$

$$H_4(x) = 16x^4 - 48x^2 + 12$$

$$H_5(x) = 32x^5 - 160x^3 + 60x$$

Problems

1. Show that the Sturm-Liouville operator is:

   (i) a linear operator

   (ii) self-adjoint

2. Show that every second order ordinary differential equation

$$a_2(x)\frac{d^2y}{dx^2} + a_1(x)\frac{dy}{dx} + a_0(x)y = f(x)$$

   can be cast as a Sturm-Liouville problem

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + r(x)y = F(x)$$

   using the following transformations.

$$p(x) = exp\left(\int \frac{a_1(x)}{a_2(x)}dx\right)$$
$$r(x) = p(x)\frac{a_0(x)}{a_2(x)}$$
$$F(x) = p(x)\frac{f(x)}{a_2(x)}$$

3. Verify whether Legendre, Laguerre, Hermite and Bessel equations

   (i) can be cast as Sturm-Liouville problems?

   (ii) are linear?

   (iii) are self-adjoint?

4. Determine the eigenvalues and eigenfunctions of the boundary value problem

$$\frac{d^2y}{dx^2} + \lambda y = 0; \; a \leq x \leq b$$

$$y(a) = y(b) = 0$$

5. Use Gram-Schmidt method to construct an orthogonal set of polynomials out of the independent set

$$\{1, x, x^2, x^3, x^4, x^5 : -1 \leq x \leq 1\}$$

   Compare the result with the Legendre polynomials $P_i(x), i = 1 - 5$. Show that there is a linear relationship between the two set of polynomials.

6. Verify that $P_n(x)$ satisfies the Legendre's equation for $n = 3$ and $n = 4$.

7. Show that the substitutions

$$x = cos\theta$$

$$y(cos\theta) = u(\theta)$$

transform the Legendre equation to

$$sin\theta \frac{d^2u}{dx^2} + cos\theta \frac{du}{dx} + [n(n+1)sin\theta]u = 0, 0 \leq \theta \leq \pi$$

8. Verify *Rodrigues generating function* for Legendre polynomial as

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$$

9. Show that Legendre polynomials are orthogonal in $\mathcal{L}^2(-1, 1)$.

10. Verify whether the following is a correct generating function for Hermite polynomials.

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2})$$

11. Determine whether the set $\{H_n(x) : n \in \mathbb{N}_0\}$ is orthogonal in $\mathcal{L}^2(\mathbb{R})$.

12. Show that $||H_n(x)||^2 = 2^n n! \sqrt{\pi}$.

13. Verify whether the following is a correct generating function for Laguerre polynomials.

$$L_n(x) = \sum_{k=0}^{n} (-1)^k \frac{1}{k!} \binom{n}{k} x^k$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

14. Verify whether the following is a correct generating function for Laguerre polynomials.

$$L_n(x) = e^x \frac{d^n}{dx^n} (x^n e^{-x})$$

15. Show that Bessel functions J_n and J_{-n} are not linearly independent if $n \in \mathbb{N}_0$. Check the same for the case of $n \in \mathbb{I}$.