UNIT 3

SET FUNCTION AND DISTANCE FUNCTION

Structure

3.1 INTRODUCTION

In Units 1 and 2 you have studied the functions of a single variable. You know that function is a rule which associates each element of a set with a unique element of another set. In this unit, you will study set function in Sec. 3.4 which is also a function from the set X to Y but here members of the set X are set itself. So, the domain of the set function is a collection of subsets of a set and the range may be any subset of the set of all extended real numbers. What we mean by extended real numbers and how we can perform operations in an extended real number system is explained in Sec. 3.2. You know that sum of two natural numbers is again a natural number. In such a case we say that the operation of addition is a binary operation on the set of all natural numbers. In Sec. 3.3 we have introduced the idea of binary operations on the family of numbers. Three special functions namely: characteristic, simple and step functions are defined in Sec. 3.5. In the course MST-026 you have to find out the distance between two points, so some distance functions are discussed in Sec. 3.6.

What we have discussed in this unit is summarised in Sec. 3.7. Self-Assessment Questions (SAQs) have been given in some sections which are generally based on the content discussed in that section. But to give you a good practice of what we have discussed in this unit, some more questions based on the entire unit are given in Sec. 3.8 under the heading Terminal Questions. Due to the reason mentioned in Sec. 1.1 of Unit 1, solutions of all the SAQs and Terminal Questions are given in Sec. 3.9.

In the next unit, you will study the convergence and divergence of sequences and series.

Expected Learning Outcomes

After completing this unit, you should be able to:

- ❖ explain the rules for addition and multiplication in the world of extended real numbers;
- ❖ define the set function and give some examples of the set function; and
- ❖ define some functions: characteristic function, simple function, step function and metric function.

3.2 EXTENDED REAL NUMBERS

You know when we say that x is a real number it means $-\infty < x < \infty.$ In other words, $-\infty$ and $+\infty$ are not included in the set of real numbers. If we include these two symbols in the set of all real numbers then this extended set is called the **extended set of real numbers** and is denoted by $\mathbb{R}^*$. So, $\mathbb{R}^* = (-\infty, +\infty) \cup \{-\infty, +\infty\} = [-\infty, +\infty].$ After including these two symbols we have to perform addition and multiplication with these two new symbols. So, we need to define what are the rules for addition and multiplication with these two new symbols. These rules are explained as follows:

Addition: If x is a real number, i.e., $-\infty < x < +\infty,$ then

- $x + (-\infty) = -\infty$ … (3.1)
- $x + (+\infty) = +\infty$ … (3.2)
- $(+\infty) + (+\infty) = +\infty$ ... (3.3)
- $(-\infty) + (-\infty) = -\infty$ ... (3.4)

Multiplication: If x is a real number, i.e., $-\infty < x < +\infty,$ then

- If x > 0 then $x(-\infty) = (-\infty)x = -\infty,$ $x(+\infty) = (+\infty)x = +\infty$ … (3.5)
- If x < 0 then $x(-\infty) = (-\infty)x = +\infty,$ $x(+\infty) = (+\infty)x = -\infty$ … (3.6)
- If x = 0 then $x(-\infty) = (-\infty)x = 0,$ $x(+\infty) = (+\infty)x = 0$ … (3.7)
- $(+\infty)(-\infty) = (-\infty)(+\infty) = -\infty,$ $(-\infty)(-\infty) = +\infty,$ $(+\infty)(+\infty) = +\infty$ … (3.8)

But remember, following are not defined even in the world of extended real numbers.

$(+\infty) + (-\infty)$ and $(-\infty) + (+\infty)$ … (3.9)

Now, you can try the following Self-Assessment Question.

SAQ 1

Suppose you are working in the world of extended real numbers then replace the ? sign by appropriate real numbers or the symbol $-\infty$ or $+\infty$ as per the rules of this world.

(i) $(-3)(-\infty) = ?$ (ii) $(5)(-\infty) = ?$ (iii) $(-10)(+\infty) = ?$

(iv) $-5 + (-\infty) = ?$ (v) $15 + (+\infty) = ?$ (vi) $-7 + (-\infty) = ?$

(vii) $-9+(+\infty)=?$ (viii) $(0)(-\infty)=?$ (ix) $(0)(+\infty)=?$

(x) $(+\infty)+(-\infty)=?$

3.3 BINARY OPERATIONS

Consider two numbers 2 and 5 then you know how to apply four fundamental operations addition, subtraction, multiplication and division on two numbers. So, we have $2+5=7,\ 2-5=-3,\ 2\times 5=10,\ 2\div 5=0.4.$ In mathematics we say that an operation $*$ (say) on a non-empty set X is said to be binary if $a*b\in X,$ for all $a,\ b\in X.$ If $*$ is a binary operation on a non-empty set X then we also say that X is **closed with respect to the operation** $*.$ If so then we also say that **closure property** holds in X with respect to the operation $*.$ For example, set $\mathbb{N}$ of all natural numbers is closed with respect to the operations addition (+) and multiplication $(\times)$ because sum and product of two natural numbers is always a natural number, i.e.,

$$a+b,\ a\times b\in\mathbb{N},\ \text{for all}\ a, b\in\mathbb{N}. \quad \ldots (3.10)$$

But the set of all natural numbers is not closed with respect to the subtraction $(-)$ and division $(\div)$ operations because $2, 5\in\mathbb{N}$ but $2-5=-3\notin\mathbb{N}$ and $2\div 5=0.4\notin\mathbb{N}.$

However, addition, subtraction, multiplication and division (except zero) all are binary operations in the set $\mathbb{R}$ of all real numbers.

You have gotten the idea of binary operation. Before defining it in general let us first define cartesian product of two sets.

Cartesian Product: Cartesian product of two sets A and B is denoted by $A\times B$ and define as follows. $A\times B=\{(x,\ y): x\in A,\ y\in B\}$ … (3.11)

For example, if A = {2, 5} and B = {4, 6, 7} then

$$A\times B=\{(2, 4), (2, 6), (2, 7), (5, 4), (5, 6), (5, 7)\}$$

Now we can define binary operation in general as follows.

A binary operation $*$ is a function $*: X\times X\to X$ defined by

$$*(a, b)=a*b,\ (a, b)\in X\times X \quad \ldots (3.12)$$

There are some other terms related to binary operation which are explained as follows.

Commutative Binary Operation: A binary operation $*: X\times X\to X$ is said to be commutative binary operation if it satisfies the following condition

$$a*b=b*a,\ \text{for all}\ a, b\in X \quad \ldots (3.13)$$

For example, binary operations addition (+) and multiplication $(\times)$ on the set $\mathbb{N}$ of all natural numbers are commutative since

$$a+b=b+a\ \text{and}\ a\times b=b\times a\ \text{for all}\ a, b\in\mathbb{N} \quad \ldots (3.14)$$

Operation subtraction $(-)$ is a binary operation on the set of all real numbers $\mathbb{R}.$ However, it is not commutative since there exists $2, 5\in\mathbb{R}$ such that $2-5=-3\in\mathbb{R}$ and $5-2=3\in\mathbb{R},$ but $2-5\neq 5-2,$ hence, subtraction $(-)$ is not commutative on the set of all real numbers $\mathbb{R}.$

Associative: A binary operation $*: X \times X \to X$ is said to be associative if it satisfies the following condition

$$(a * b) * c = a * (b * c), \text{ for all } a, b, c \in X \qquad \ldots (3.15)$$

For example, binary operations addition (+) and multiplication $(\times)$ on the set $\mathbb{N}$ of all natural numbers are associative since

$$(a + b) + c = a + (b + c) \text{ and } (a \times b) \times c = a \times (b \times c) \text{ for all } a, b, c \in \mathbb{N} \quad (3.16)$$

But operation subtraction $(-)$ is not associative on $\mathbb{R}$, the set of all real numbers because there exists $2, 5, 7 \in \mathbb{R}$ such that $(2-5)-7 \neq 2-(5-7)$. Hence, subtraction $(-)$ is not associative on the set of all real numbers $\mathbb{R}$.

Similarly, operation subtraction $(-)$ is not associative on any of the sets $\mathbb{Z}, \mathbb{Q}, \mathbb{R},$ and $\mathbb{C}$.

Existence of Identity: If $*: X \times X \to X$ is a binary operation then we say that an element $e \in X$ if it exists is called identity element with respect to the operation $*$ if

$$e * a = a = a * e, \ \forall \ a \in X \qquad \ldots (3.17)$$

For example, 0 is the identity element with respect to the binary operation addition (+) in the sets $\mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$. Also, 1 is the identity element with respect to the binary operation multiplication $(\times)$ in the sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$.

Subtraction $(-)$ is binary operations in $\mathbb{R}$ but there is no identity element in $\mathbb{R}$ with respect to subtraction operation.

Existence of Inverse: If $*: X \times X \to X$ is a binary operation and e be the identity element in X with respect to the operation $*$ then we say that an element $a \in X$ has its inverse in X if there exists an element $b \in X$ such that

$$b * a = e = a * b \qquad \ldots (3.18)$$

b is known as inverse of a and is generally denoted by $b = a^{-1}$. If each non-zero element of X has its inverse in X then we say that property 'existence of inverse' with respect to $*$ holds in X.

For example, property 'existence of inverse' with respect to addition (+) holds in $\mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$. But property 'existence of inverse' with respect to multiplication $(\times)$ holds in $\mathbb{Q}$ and $\mathbb{R}$ but not in $\mathbb{Z}$.

In the next section we will discuss set function and measure. So, before studying them it will be better if you have understanding of what we mean by a σ-field which is defined as follows.

Let Ω be a non-empty set and $\mathcal{F}$ be a collection of subsets of Ω then we say that the class $\mathcal{F}$ forms a **sigma algebra** (σ-algebra) or **sigma field** (**σ-field**) if it satisfies following three conditions:

(i) $\phi, \ \Omega \in \mathcal{F}$, i.e., empty set and the full space Ω should be members of $\mathcal{F}$

(ii) $A \in \mathcal{F} \Rightarrow A^c (= \Omega \setminus A) \in \mathcal{F}$, i.e., $\mathcal{F}$ is closed with respect to complement

(iii) $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ for any countable collection $\{A_n\}_{n=1}^{\infty}$ in $\mathcal{F}$, i.e., $\mathcal{F}$ is closed with respect to countable union $\qquad \ldots (3.19)$

Let us consider two examples.

Example 1: If Ω = {a, b, c, d} then what is the

(i) smallest σ-field on Ω

(ii) smallest σ-field containing the set A = {a} on Ω

(iii) largest σ-field on Ω.

Solution: We know that a collection $\mathcal{F}$ of subsets of Ω is called a σ-field if it satisfies the following three conditions.

(a) $\phi,\ \Omega \in \mathcal{F}$ (b) $A \in \mathcal{F} \Rightarrow A^c (= \Omega \setminus A) \in \mathcal{F}$ (c) $\{A_n\}_{n=1}^{\infty}$ is in $\mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

(i) So, the smallest σ-field on Ω is $\mathcal{F} = \{\phi,\ \Omega\}.$ It satisfies all the three requirements for a collection to be a σ-field. This is known as **trivial** σ-field on Ω.

(ii) Keeping in view the three requirements for a collection to be a σ-field the smallest σ-field containing a set A on Ω is given by $\mathcal{F} = \{\phi,\ A,\ A^c,\ \Omega\}.$

In our case A = {a} and Ω = {a, b, c, d}, therefore, required σ-field is given by

$$\mathcal{F} = \{\phi,\ \{a\},\ \{b, c, d\},\ \{a, b, c, d\} = \Omega\}$$

(iii) By definition of σ-field it is a collection of subsets of Ω which satisfies three conditions. Obviously, from school mathematics you know that the largest collection of subsets of any set is its power set. Further, power set is closed with respect to complement and union. So, the largest σ-field on Ω is the power set of Ω So, required σ-field on Ω is given by

$$\mathcal{F} = \left\{ \begin{matrix} \phi,\ \{a\},\ \{b\},\ \{c\},\ \{d\},\ \{a, b\},\ \{a, c\},\ \{a, d\},\ \{b, c\},\ \{b, d\}, \\ \{c, d\}, \{a, b, c\},\ \{a, b, d\}, \{a, c, d\},\ \{b, c, d\},\ \Omega \end{matrix} \right\} \quad \ldots (3.20)$$

Note 1: We have seen that both $\{\phi,\ \Omega\}$ and $P(\Omega) =$ power set, always form σ-field on Ω so you can say that both are trivial σ-fields on $\Omega.$ There exists lots of σ-fields on Ω other than these two. Example 2 given as follows and generalisation given after Example 2 explains a way of creating more σ-fields on Ω.

Example 2: If Ω = {a, b, c, d} then what is the smallest σ-field containing the sets A = {b} and B = {c, d} on Ω.

Solution: We know that a collection of subsets of a set Ω form a σ-field if it contains empty set, full set, closed under complement and closed under countable union. Therefore, a collection containing the sets A and B will form a σ-field on Ω if it contains following subsets of Ω

$$\mathcal{F} = \left\{ \begin{matrix} \phi,\ A,\ B,\ A^c,\ B^c,\ A \cap B,\ A \cap B^c,\ A^c \cap B,\ A^c \cap B^c,\ A \cup B, \\ A \cup B^c,\ A^c \cup B,\ A^c \cup B^c,\ A \Delta B,\ (A \Delta B)^c,\ \Omega \end{matrix} \right\}$$

In our case A = {b}, B = {c, d} and $\Omega = \{a, b, c, d\}$ so it becomes

$$\mathcal{F} = \{\phi,\ \{b\},\ \{c, d\},\ \{a, c, d\},\ \{a, b\},\ \{a\},\ \{b, c, d\},\ \{a, b, c, d\}\}$$

Generalisation: Every collection of subsets of a non-empty set Ω do not form a σ-field. But we can add more subsets of Ω in the given collection to form a

σ-field. How many subsets at the most we need in total to form a σ-field can be understood as follows.

From part (ii) of Example 1 note that if a collection has one set other than ϕ and Ω then we need at most $4\left(=2^{2^1}\right)$ subsets of Ω to form a σ-field. From Example 2 note that if a collection has two sets other than ϕ and Ω, then we need at most 16 $\left(=2^{2^2}\right)$ subsets of Ω to form a σ-field. In general, if a collection has n subsets of Ω other than ϕ and Ω then we need at the most 2^{2^n} subsets of Ω to form a σ-field with these given n subsets of Ω. But in practice we have less number of subsets compare to general case. For example, in Example 2 we have only 8 subsets instead of 16. … (3.21)

Now, you can try the following Self-Assessment Question.

SAQ 2

If Ω = {a, b, c, d} then what is the smallest σ-field containing the set $A = \{a\}$, B = {a, c] on Ω?

Now, we define set function and measure in the next section.

3.4 SET FUNCTION AND MEASURE

In Sec. 1.4 of Unit 1 of this course you have understood the definition of a function. You have seen that a function associates each element of a **set** to a unique element of another set or the same set. Set function is also a function but here domain is not a set like in the case of a function but it is a **collection of subsets of a non-empty set**. Because of the reason that elements of the domain of the set function are itself sets so it is given the name set function and is defined as follows.

Set Function: Let Ω be a non-empty set and $\mathcal{F}$ be any collection of subsets of Ω then a function $\mu : \mathcal{F} \to [0, +\infty]$ is known as a **set function**. … (3.22)

As mentioned earlier note that in a set function inputs are sets and outputs are real numbers or extended real numbers. Generally, output of a set function is some kind of measure of the input set. For example, it may be cardinality of the set, length or area or volume, etc. So, we can interpret it as a measure of the input set. As per the definition of a set function there is no requirement that members of its domain should form some algebraic structure. But many times, in particular, probability theory and measure theory we are interested to measure all members of a σ-field of subsets of Ω. So, instead of taking arbitrary collection of subsets of a set as a domain of a set function we generally take a σ-field of subsets of Ω as a domain of a set function. You will get more explanation regarding it in Unit 2 of the course MST-012. As mentioned earlier output of a set function is some kind of measure of the input set. So, let us define when a set function is called a measure.

Measure: Let Ω be a non-empty set and $\mathcal{F}$ be a collection of subsets of Ω which forms a σ-field then A set function $\mu : \mathcal{F} \to [0, +\infty]$ is called a measure if it satisfies the following two requirements … (3.23)

(i) $\mu(\phi) = 0$, i.e., measure of an empty set should be zero

(ii) μ is countably additive, i.e., whenever $A_1, A_2, A_3, \ldots \in \mathcal{F}$ with $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ and $A_m \cap A_n = \phi$ for all $m \neq n$ then $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$

Since μ is a function from $\mathcal{F}$ to $[0, +\infty]$ so it will associate a unique measure to each element of the σ-field $\mathcal{F}$ and so members of the σ-field $\mathcal{F}$ are known as **measurable sets**. … (3.24)

Thus, we can say that a measure is a set function which associates a measure to the members of the σ-field $\mathcal{F}$. In measure theory which is generally studied as a course in master degree programme of mathematics the triplet $(\Omega, \mathcal{A}, \mu)$ is known as **measure space**. Also Ω together with the σ-field $\mathcal{F}$ i.e., $(\Omega, \mathcal{F})$ is known as **measurable space**. Measurable space $(\Omega, \mathcal{F})$ means members of $\mathcal{F}$ are just one step away to get their measure, i.e., as soon as you define a measure μ on $\mathcal{F}$ then each member of $\mathcal{F}$ will get their corresponding measure. While measure space means members of $\mathcal{F}$ have been measured as per measure function μ. … (3.25)

Another important set function on extended real line is the length function which measures lengths of the intervals and is defined as follows.

Let $\mathcal{I}$ be the collection of all intervals on the real line. Let $I(a, b) \in \mathcal{I}$ be an interval with end points a and b where a and b are extended real numbers. The interval I(a, b) may be both sides open, both sides closed, one side open and one side closed. Then the set function $\lambda : \mathcal{I} \to [0, +\infty]$ defined by

$$\lambda(I(a, b)) = \begin{cases} |b - a|, & \text{if } a, b \in \mathbb{R} \\ +\infty, & \text{if either } a = -\infty \text{ or } b = +\infty \text{ or both} \end{cases} \quad I(a, b) \in \mathcal{I} \quad \ldots (3.26)$$

is called the **length function**.

Let us now consider one example of a set function.

Example 3: Let $\Omega = \{a, b, c, d\}$ and $\mathcal{F} = \{\phi, \{a\}, \{b, c, d\}, \Omega\}$ then define a set function $\mu : \mathcal{F} \to [0, +\infty]$ by

$$\mu(A) = \begin{cases} n(A), & \text{if } A \text{ is a finite set} \\ +\infty, & \text{if } A \text{ is not a finite set} \end{cases} \quad A \in \mathcal{F} \quad \ldots (3.27)$$

where n(A) represents cardinality of the set A. Obtain measure of all the members of $\mathcal{F}$ associated by the set function μ.

Solution: By definition of μ, measures of members of $\mathcal{F}$ are given as follows.

$$\mu(\phi) = n(\phi) = 0, \quad \mu(\{a\}) = n(\{a\}) = 1,$$
$$\mu(\{b, c, d\}) = n(\{b, c, d\}) = 3, \quad \mu(X) = n(X) = 4$$

Remark 1: The set function defined in Example 3 actually counts the number of elements in the input set. So, it applies counting measure on the members of $\mathcal{F}$ so it is known as **counting measure**. … (3.28)

Now, you can try the following Self-Assessment Question.

SAQ 3

Let $\mathcal{I}$ be the collection of all intervals on the real line and $\lambda : \mathcal{I} \to [0, +\infty]$ be the length function on $\mathcal{I}$ then write the image of the following members of $\mathcal{I}$ under the function λ.

(a) $[4, 9]$ (b) $(4, 9]$ (c) $[4, 9)$ (d) $(4, 9)$ (e) $[-\infty, 9]$
(f) $[-\infty, 9)$ (g) $[4, +\infty]$ (h) $[4, +\infty)$ (i) $[-\infty, +\infty]$ (j) $[-\infty, +\infty)$
(k) $(-\infty, +\infty]$ (l) $(-\infty, +\infty)$

3.5 CHARACTERISTIC, SIMPLE AND STEP FUNCTIONS

Suppose in a study we have 50 males and 52 female subjects. So, our universal set Ω which contains subjects of the study has 102 members/elements. If M and F denote respectively the sets of males and females subjects of the study then we have

- $M = \{m_1, m_2, m_3, \ldots, m_{50}\}$ is a subset of Ω, i.e., $M \subset \Omega$ and n(M) = 50, $n(\Omega) = 102$.
- $F = \{f_1, f_2, f_3, \ldots, f_{52}\}$ is a subset of Ω, i.e., $F \subset \Omega$ and n(F) = 52, $n(\Omega) = 102$.
- $\Omega = \{m_1, m_2, m_3, \ldots, m_{50}, f_1, f_2, f_3, \ldots, f_{52}\}, M^c = F, F^c = M, M \cup F = \Omega.$

If we define two functions $\chi_M : \Omega \to \{0, 1\}$ and $\chi_F : \Omega \to \{0, 1\}$ by

$$\chi_M(\omega) = \begin{cases} 1, & \omega \in M \\ 0, & \omega \notin M \end{cases} \quad \omega \in \Omega \quad \text{and} \quad \chi_F(\omega) = \begin{cases} 1, & \omega \in F \\ 0, & \omega \notin F \end{cases} \quad \omega \in \Omega$$

then what the functions χ_M and χ_F actually are indicating. The function χ_M is indicating whether the subject of the study is male or not. If outcome is 1 it means subject of the study is a male and if outcome is 0 it means subject is not a male. Similarly, the function χ_F is indicating whether the subject of the study is female or not. If outcome is 1 it means subject of the study is a female and if outcome is 0 it means subject is not a female. Such types of indicator functions have given the name characteristic function in mathematics because their outputs characterise the underlying set. So, in mathematics a characteristic function is defined as follows.

Characteristic Function: Let A be any subset of the universal set Ω then characteristic function of A is denoted by χ_A and defined as follows:

$\chi_A : \Omega \to \{0, 1\}$ define by

$$\chi_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases} \quad \omega \in \Omega \qquad \ldots (3.29)$$

Note 2: The characteristic function χ_A of the set A is also called the **indicator function** of A. … (3.30)

Let us do an example.

Example 4: Suppose a red die is thrown and E, F are the events defined as follows. E: getting an odd number, F: getting a multiple of 3. Write events E, F, $E \cup F$, and $E \cap F$. Visualise the characteristic functions of these four events in pictorial form.

Solution: When a red colour die (note here colour of the die is not any issue it is mentioned with the objective that you get exposure of such small points which are sometimes used in the statements of the question just to test the understanding level of the learner) is thrown then sample space is given by

$\Omega = \{1, 2, 3, 4, 5, 6\}$

Required four events are given by

$E = \{1, 3, 5\}, \quad F = \{3, 6\}, \quad E \cup F = \{1, 3, 5, 6\}, \quad E \cap F = \{3\}$

Now, visualisation of the characteristic functions of these four events in pictorial form are respectively shown in Fig. 3.1 (a) to (d).

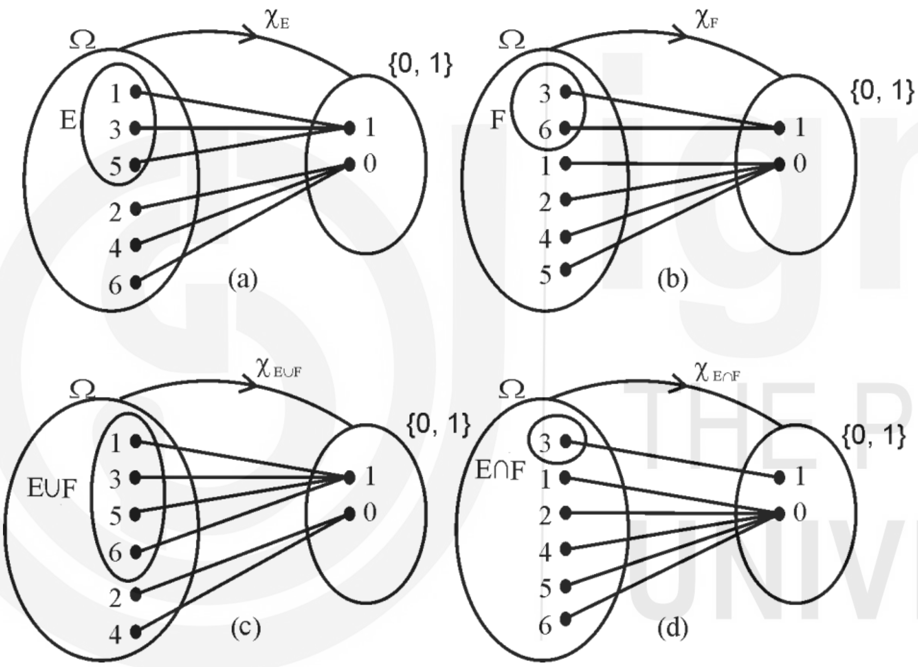

Fig. 3.1: Visualisation of the characteristic functions for the sets (events) (a) E (b) F (c) $E \cup F$ (d) $E \cap F$

Now, we discuss what is a simple function.

We just saw that characteristic function χ_A of a set A assumes only two values 0 and 1, refer (3.29). In Unit 1 of this course, you saw that signum function assumes only three values 0, – 1, 1, refer to (1.51). Similarly, a measurable function which assumes only a finite number of values is called a simple function. So, simple function is defined as follows.

Simple Function: Let $(\Omega, \mathcal{F}, \mu)$ be a measure space then a function $\phi : \Omega \to \mathbb{R}$ is said to be simple function if there is a finite disjoint class $\{E_1, E_2, E_3, \ldots, E_n\}$ of measurable sets (it means $E_i \in \mathcal{F}$) with $\bigcup_{i=1}^{n} E_i = \Omega$ and a finite set $\{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n\}$ of real numbers such that

$$\phi(x) = \alpha_i, \quad x \in E_i, \quad i = 1, 2, 3, 4, \ldots, n \qquad \ldots (3.31)$$

So, we can say that **simple function is a constant function in each** E_i. Characteristic function takes values 1 in E and 0 in E^c. So, it is a simple function. Finally, let us define a step function.

In Unit 1 of this course, we have defined Unit Step Function you may refer (1.49). Here we will define a step function in general. But before going through the definition of step function it is recommended that you should go through the definition of a partition of a closed and finite interval given by (5.20) and (5.21) in Unit 5 of this course.

Step Function: Let [a, b] be a finite and closed interval. A function $\phi : [a, b] \to \mathbb{R}$ is said to be a step function if:

(1) there is a partition P of [a, b], where by partition we mean a finite ordered set $\{x_0, x_1, x_2, x_3, \ldots, x_n\}$ of points of [a, b] such that
$a = x_0 < x_1 < x_2 < x_3 < \ldots < x_{n-1} < x_n = b.$

The n + 1 points $x_0, x_1, x_2, x_3, \ldots, x_n$ are called partition points of P and the n sub-intervals
$I_1 = [x_0, x_1], I_2 = [x_1, x_2], I_3 = [x_2, x_3], \ldots, I_i = [x_{i-1}, x_i], \ldots, I_n = [x_{n-1}, x_n]$
determined by n + 1 points of P are called the segments of the partition P.

(2) function ϕ assumes one and only one value in each **open** sub-intervals I_i, i = 1, 2, 3, …, n. So, there will exits n real numbers α_i such that

$\phi(x) = \alpha_i, \ \forall\ x \in (x_{i-1}, x_i), \ \ i = 1, 2, 3, \ldots, n$

(3) Also, values of the function ϕ at the end points of the sub-intervals I_i is irrelevant. … (3.32)

Remark 2: Like simple function step function is also constant in each sub-interval (x_{i-1}, x_i). Difference between them is in the definition of simple function $E_1, E_2, E_3, \ldots, E_n$ were sets while in the definition of step function we have intervals. We know that an interval is a set but a set may not be an interval. So, there may be simple functions which are not step functions but every step function is a simple function.

3.6 DISTANCE FUNCTION AND METRIC

So, far by distance between two points we mean straight line distance between them. For example:

(1) If you have two points on the real line, i.e., numbers x = 1, y = 4 $\in \mathbb{R}$ then distance between them is 3 units refer Fig. 3.2 (a).

(2) If you have two points $a = A(4, 0), \ b = B(0, \ 3) \in \mathbb{R}^2$ then using Pythagoras theorem in triangle OAB distance between them is 5 units refer to Fig. 3.2 (b).

But in machine learning we also use some other kinds of distances other than straight line distance between two points. The objective of this section is to make you familiar with some of those kinds of distances. First of all, let us explain what we mean by a distance function.

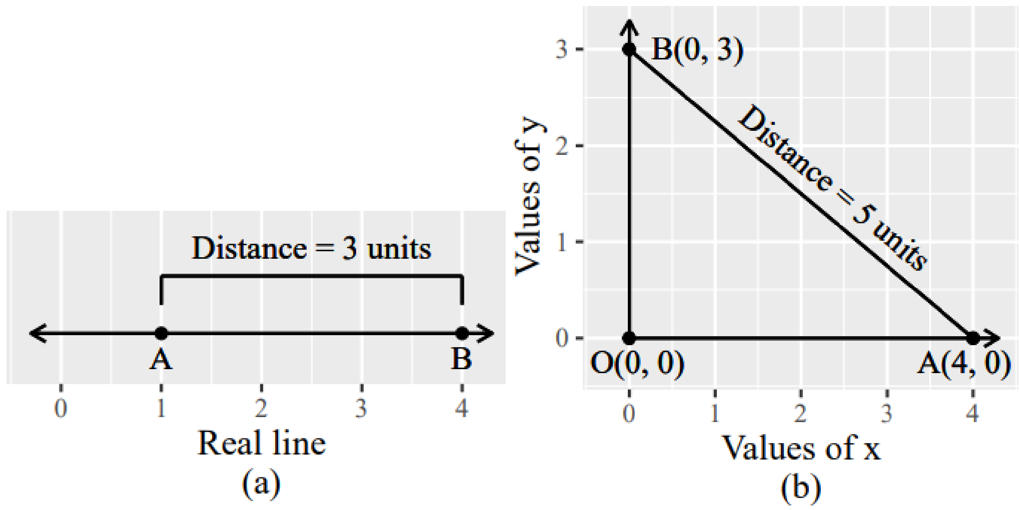

Fig. 3.2: Visualisation of the distance between two points A and B on (a) real line (b) the x-y plane

Distance Function: Let X be a non-empty set then distance function on X is a function which takes a pair of points from X as input and gives a non-negative real number as output. That is distance function looks like $d : X \times X \to [0, \infty]$, so domain of distance function d is $X \times X$ and range may be any subset of $[0, \infty]$. If $(a, b) \in X \times X$ then distance between two points a and b of X are denoted by d(a, b) and $d(a, b) \in [0, \infty]$. … (3.33)

We are interested in a distance function which satisfies following four properties (i) non-negative (ii) coincidence (iii) symmetry (iv) triangle inequality.

A distance function which satisfies these four properties is known as metric. A proper definition of a metric is given as follows.

Metric: Let X be a non-empty set then a distance function $d : X \times X \to [0, \infty]$ is called a metric on X if it satisfies following four properties.

(a) **Non-Negativity Property**: If x and y be any two points of X then we should have

$d(x, y) \geq 0 \ \forall \ x, y \in X.$ … (3.34)

i.e., distance between any two points of X should be non-negative real number or distance between two points should not be negative.

(b) **Coincidence Property**: distance function d should satisfy coincidence property

i.e., $d(x, y) = 0$ if and only if $x = y \ \ \forall \ x, y \in X$ … (3.35)

i.e., distance of each point of X should be zero from itself and vice versa.

(c) **Symmetric Property**: If $x, y \in X$ then distance function d should satisfy

$d(x, y) = d(y, x) \ \forall \ x, y \in X.$ … (3.36)

i.e., if x and y are any two points of X then distance of x from y and distance of y from x both should be equal.

(d) **Triangle Inequality**: If $x, y, z \in X$ then the distance function d should satisfy

$$d(x, y) \leq d(x, z) + d(z, y) \quad \forall\ x, y, z \in X \qquad \ldots (3.37)$$

i.e., if x and y are any two points of X then distance between x and y should be less than equal to the sum of distances of x to z and z to y for any point z of X.

In (3.25) we explained the meaning of measure space. Similarly, if d is a metric on X then (X, d) is known as **metric space**. When we say (X, d) is a metric space it means every two points x and y of X have a well-defined distance which is denoted by d(x, y) and given by as per definition of d.
… (3.38)

Some particular distance functions that we are going to discuss in this section are given as follows.

(a) Minkowski Distance

(b) Manhattan Distance

(c) Euclidean Distance

(d) Chebyshev Distance

(e) Hamming Distance

Let us discuss them one at a time.

(a) Minkowski Distance: A distance function $d: \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ defined by

$$d(x, y) = \sqrt[p]{|x_1 - y_1|^p + |x_2 - y_2|^p + |x_3 - y_3|^p + \ldots + |x_n - y_n|^p}$$

$$= \sqrt[p]{\sum_{i=1}^{n} |x_i - y_i|^p} \qquad \ldots (3.39)$$

where $x = A(x_1, x_2, x_3, \ldots, x_n)$, $y = B(y_1, y_2, y_3, \ldots, y_n) \in \mathbb{R}^n$, $p \geq 1$

is called Minkowski distance between two points A and B of $\mathbb{R}^n$. … (3.40)

(b) Manhattan Distance: If p = 1 in Minkowski distance function then corresponding distance function is known as Manhattan distance function. So, a distance function $d: \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ defined by

$$d(x, y) = |x_1 - y_1| + |x_2 - y_2| + |x_3 - y_3| + \ldots + |x_n - y_n| = \sum_{i=1}^{n} |x_i - y_i| \qquad \ldots (3.41)$$

where $x = A(x_1, x_2, x_3, \ldots, x_n)$, $y = B(y_1, y_2, y_3, \ldots, y_n) \in \mathbb{R}^n$

is called Manhattan distance between two points A and B of $\mathbb{R}^n$. If in a city, roads are like rectangular grids then distance travelled by a taxi to move from one point to other point is given by Manhattan distance formula. So, it is also known as **taxicab distance** or **city block distance**. In the area of Manhattan in New York roads are like rectangular grids.

For example, the distance shown in Fig. 3.2 (a), gives Manhattan distance between points A and B where $p = 1, n = 1, x_1 = 1, y_1 = 4,$ so

$$AB = d(x, y) = \left(\sum_{i=1}^{1} |x_i - y_i|^1\right)^{1/1} = |x_1 - y_1| = |1 - 4| = |-3| = -(-3) = 3$$

So, Manhattan distance between real numbers 1 and 4 is $d(1, 4) = 3.$ … (3.42)

Let us do an example to explain procedure of calculating Manhattan distance.

Example 5: Find the Manhattan distance between two points (1, 2) and (5, 7). Also, explain it geometrically.

Solution: We are given $(x_1, y_1) = (1, 2)$ and $(x_2, y_2) = (5, 7).$ By definition of Manhattan distance function distance between two points A(1, 2) and B(5. 7) is given by

$$\begin{aligned} AB = d(x, y) &= \sum_{i=1}^{2} |x_i - y_i| = |x_1 - y_1| + |x_2 - y_2| \\ &= |1 - 5| + |2 - 7| \\ &= |-4| + |-5| \\ &= 4 + 5 \\ &= 9 \end{aligned} \quad \text{... (3.43)}$$

So, Manhattan distance between points A(1, 2) and B(5, 7) is 9 units. … (3.44)

Let us explain it geometrically, refer Fig. 3.3 (a) and (b)

Geometrically Manhattan distance gives distance of two points A and B travel by a taxi under the condition that roads in the city are like rectangular grids. Here straight-line distance will not work because you cannot travel through buildings. So, due to buildings you cannot travel straight and we have assumed that roads are like rectangular grids so you have only two options either you can move horizontally or vertically. Thus, it gives sum of the units of the distance that you have to run horizontally (Run) plus the units of distance you have to go vertically (Rise) to reach at the point B(5, 7) by starting from the point A(1, 2). Note that x-coordinate of the point A is 1 and x-coordinate of the point B is 5 so you have to move 4 (= 5 – 1) units of distance horizontally (Run). Further, y-coordinate of the point A is 2 and y-coordinate of the point B is 7 so you have to move 5 (= 7 – 2) units of distance vertically (Rise). Hence, in total you have to move 9 units of distance including both horizontal and vertical to reach at the point B by starting from the point A. Thus, Manhattan distance between two points A and B is 9 units. Also, note that there are more than one options to reach at the point B by starting from point A and traveling exactly 9 units of distance in total. Four such options are shown in Fig. 3.3 (b) in red, magenta, black and green colours.

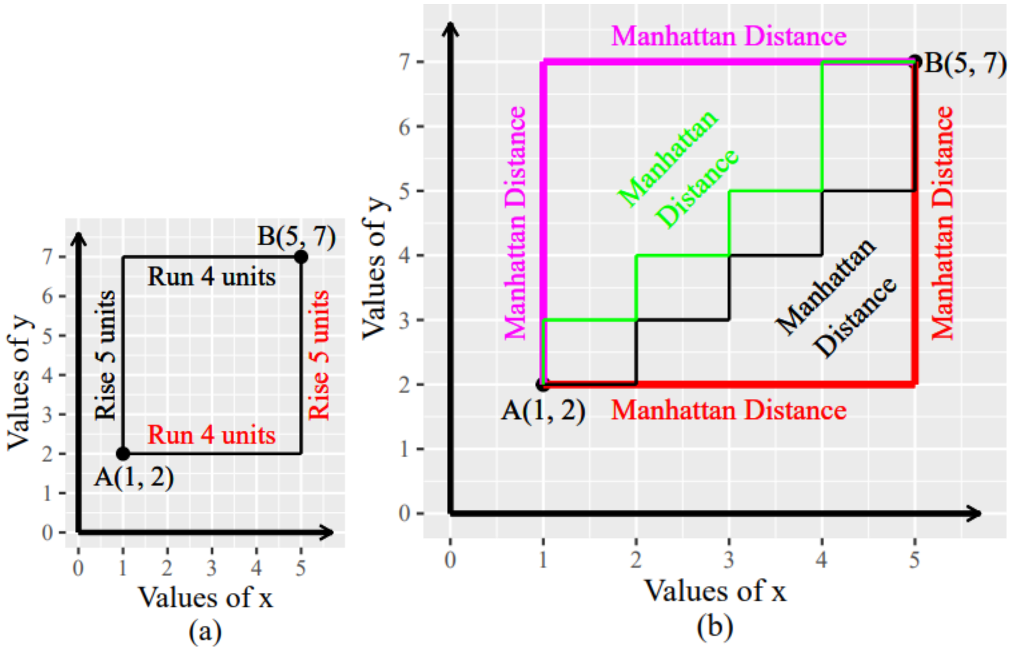

Fig. 3.3: Visualisation of (a) two-point A and B (b) Manhattan distance between two points A and B via four routes

(c) Euclidean Distance: If p = 2 in Minkowski distance function then corresponding distance function is known as Euclidean distance function. So, a distance function $d:\mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ defined by

$$d(x, y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2 + \ldots + |x_n - y_n|^2}$$

$$= \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} \qquad \ldots (3.45)$$

where $x = A(x_1, x_2, x_3, \ldots, x_n),\ y = B(y_1, y_2, y_3, \ldots, y_n) \in \mathbb{R}^n$

is called Euclidean distance between two points A and B of $\mathbb{R}^n$. This formula gives straight line distance between two points A and B. So, it is also known as **crow flying distance** because crow can fly over the buildings also. Here we are assuming that walking straight between two points of our interest is feasible.

For example, the distance shown in Fig. 3.2 (b), gives Euclidean distance between points A(4, 0) and B(0, 3) where
$p = 2, n = 2, (x_1, x_2) = (4, 0), (y_1, y_2) = (0, 3).$ So,

$$AB = d(x, y) = \left(\sum_{i=1}^{2} |x_i - y_i|^2\right)^{1/2} = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2} = \sqrt{|4-0|^2 + |0-3|^2}$$

$$= \sqrt{|4|^2 + |-3|^2} = \sqrt{16+9} = \sqrt{25} = 5$$

So, Euclidean distance between points A(4, 0) and B(0, 3) is 5 units. … (3.46)

Let us do an example to explain procedure of calculating Euclidean distance.

Example 6: Find the Euclidean distance between two points A(1, 2) and B(5, 7). Also, explain it geometrically.

Solution: We are given $(x_1, x_2) = (1, 2)$ and $(y_1, y_2) = (5, 7)$. By definition of

Euclidean distance function distance between two points A(1, 2) and B(5. 7) is given by

$$AB = d(x, y) = \sqrt{\sum_{i=1}^{2}|x_i - y_i|^2} = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2} = \sqrt{|1-5|^2 + |2-7|^2}$$

$$= \sqrt{|-4|^2 + |-5|^2} = \sqrt{16+25} = \sqrt{41} \simeq 6.403124 \text{units} \quad \ldots (3.47)$$

So, Euclidean distance between points A(1, 2) and B(5, 7) is $\sqrt{41}$ units ... (3.48)

Let us explain it geometrically, refer to Fig. 3.4 (a) and (b).

Geometrically Euclidean distance is the shortest distance between the two points A and B. So, it can be obtained using Pythagoras theorem in triangle ABC as follows.

$$AB = \sqrt{AC^2 + CB^2} = \sqrt{4^2 + 5^2} = \sqrt{16+25} = \sqrt{41} \simeq 6.403124 \text{units}$$

Thus, Euclidean distance between two points A and B is $\sqrt{41}$ units. If you are working in more than 2-dimension than you have to apply Pythagoras theorem more than once. Also, note that unlike Manhattan distance here you have only one option to reach at the point B by starting from point A and you have to travel a distance of $\sqrt{41}$ units. This unique path is shown in Fig. 3.4 (b).

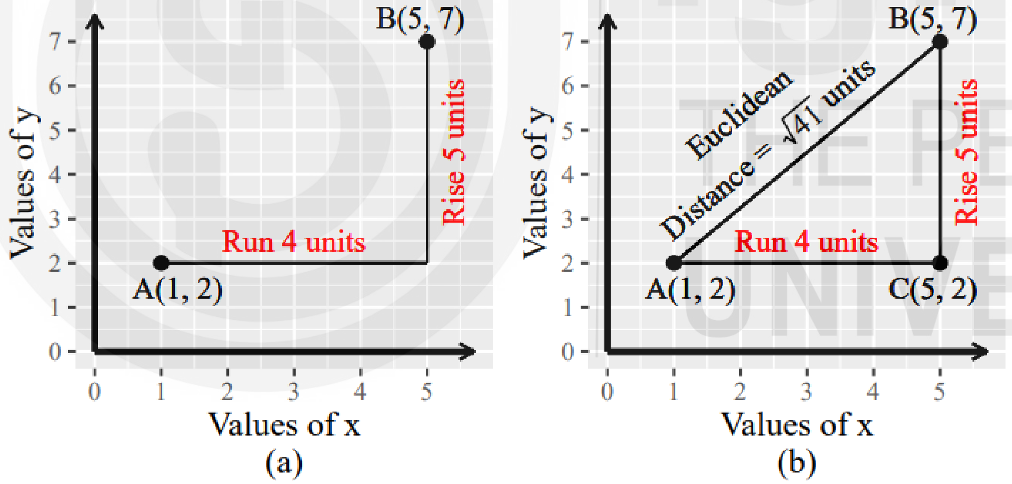

Fig. 3.4: Visualisation of (a) two points A and B (b) Euclidean distance between two points A and B via unique route

On comparing (3.44) and (3.48), we see that Euclidean distance between two points A and B is less than Manhattan distance between the same points A and B.

(d) Chebyshev Distance: If $p \to \infty$ in Minkowski distance function then corresponding distance function is known as Chebyshev distance function. So, a distance function $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty]$ defined by

$$d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|, |x_3 - y_3|, \ldots, |x_n - y_n|\} \quad \ldots (3.49)$$

where $x = A(x_1, x_2, x_3, \ldots, x_n)$, $y = B(y_1, y_2, y_3, \ldots, y_n) \in \mathbb{R}^n$

is called Chebyshev distance between two points A and B of $\mathbb{R}^n$. This formula looks at absolute distance between each corresponding coordinates of two points and select maximum among them.

For example, Chebyshev distance between points A(4, 0) and B(0, 3) is shown in Fig. 3.4 (a), and given as follows.

$$AB_{\text{Chebyshev}} = d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|, |x_3 - y_3|, \ldots, |x_n - y_n|\}$$

$$= \max\{|x_1 - y_1|, |x_2 - y_2|\}, \text{ where } x_1 = 4, x_2 = 0, y_1 = 0, y_1 = 3.$$

$$= \max\{|4 - 0|, |0 - 3|\} = \max\{|4|, |-3|\} = \max\{4, 3\} = 4.$$

So, Chebyshev distance between points A(4, 0) and B(0, 3) is 4 units. …(3.50)

Let us do an example to explain procedure of calculating Chebyshev distance.

Example 7: Find the Chebyshev distance between two points A(1, 2) and B(5, 7). Also, explain it geometrically.

Solution: We are given $(x_1, x_2) = (1, 2)$ and $(y_1, y_2) = (5, 7)$. By definition of Chebyshev distance function distance between two points A(1, 2) and B(5. 7) is given by

$$AB_{\text{Chebyshev}} = d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\},$$
$$\text{where } x_1 = 1, x_2 = 2, y_1 = 5, y_2 = 7.$$

$$= \max\{|1 - 5|, |2 - 7|\}$$
$$= \max\{|-4|, |-5|\} = \max\{4, 5\} = 5 \text{ units.} \quad \ldots (3.51)$$

So, Chebyshev distance between points A(1, 2) and B(5, 7) is 5 units… (3.52)

Let us explain it geometrically, refer to Fig. 3.4 (a).

Geometrically in 2-dimension Chebyshev distance is the maximum among Run and Rise refer Fig. 3.4 (a) where Run = 4 units and Rise = 5 units. So, Chebyshev distance between the two points A and B is max{Run, Rise}. So, $AB_{\text{Chebyshev}} = d(x, y) = \max\{\text{Run}, \text{Rise}\} = \max\{4, 5\} = 5.$

Remark 3: If points A and B both lie on a horizontal line or a vertical line then $AB_{\text{Chebyshev}} = AB_{\text{Euclidean}}$, otherwise $AB_{\text{Chebyshev}} \le AB_{\text{Euclidean}}$.

Let us do one example which will compare the shapes of a circle using different distance formulae with centre at origin and radius is of 1 unit.

Example 8: Draw a circle with centre at origin and radius of 1 unit when distance between two points is measured using (a) Manhattan distance formula (b) Euclidean distance formula (c) Minkowski distance formula for p > 2 (d) Chebyshev distance formula.

Solution: Centre of the circle is O(0, 0) and radius of the circle = r = 1 unit. We know that distance between two points $A(x_1, x_2)$ and $B(y_1, y_2)$ using different distance formulae are given by

$$AB_{Manhattan} = d(x, y) = |x_1 - y_1| + |x_2 - y_2| \quad \dots (3.53)$$

$$AB_{Euclidean} = d(x, y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2} \quad \dots (3.54)$$

$$AB_{Minkowski} = d(x, y) = \sqrt[p]{|x_1 - y_1|^p + |x_2 - y_2|^p} \quad \dots (3.55)$$

$$AB_{Chebyshev} = d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\} \quad \dots (3.56)$$

We know that locus of a point which moves in a plane in such a way that its distance from a fixed point always remains constant is called a circle. Using definitions of distance between two points given by (3.53) to (3.56) shapes of the circles are given in Fig. 3.5 (a) to (d) respectively.

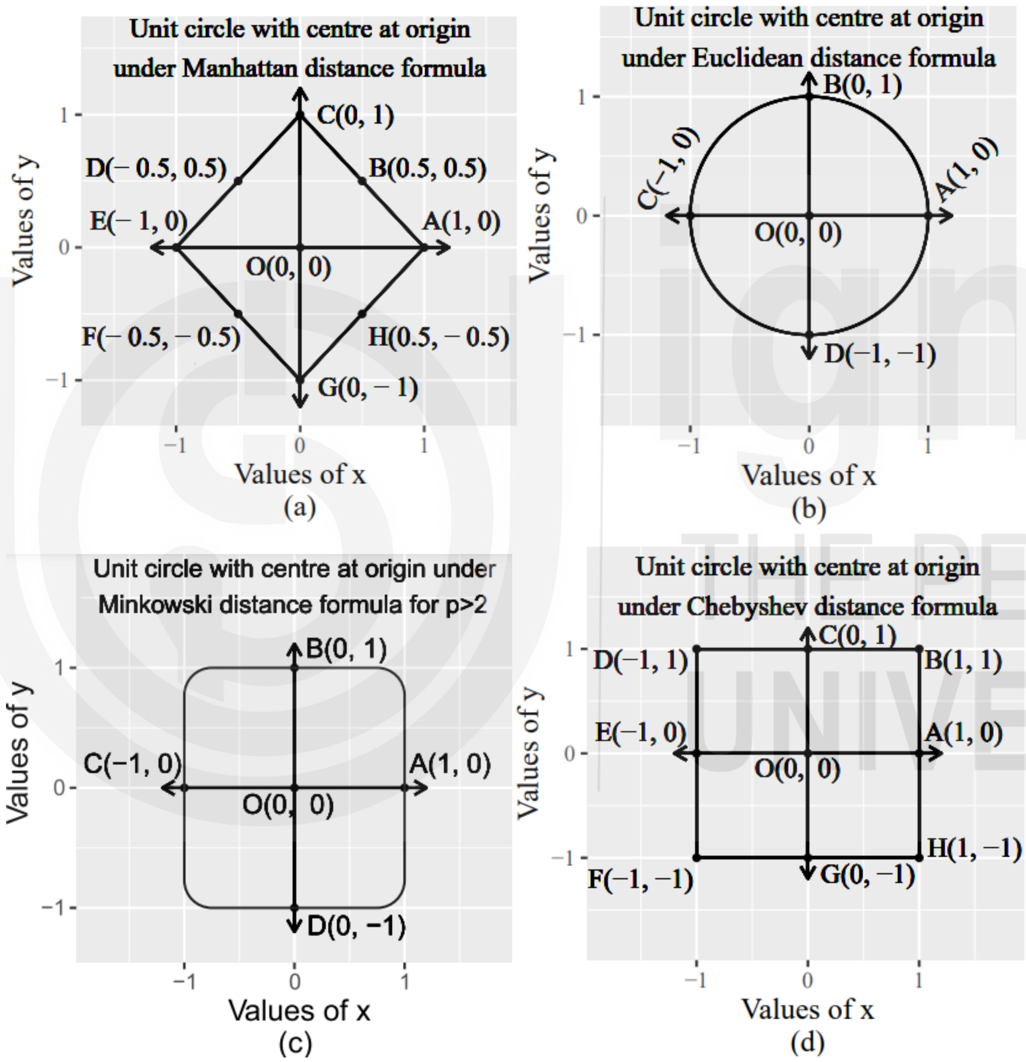

Fig. 3.5: Visualisation of unit circle with centre at the origin and formula for distance is (a) Manhattan (b) Euclidean (c) Minkowski for p > 2 (d) Chebyshev

(e) Hamming Distance: This method is used to obtain distance between two strings of the same length. Under this method distance between two strings is the number of counts of the positions at which the corresponding character in the two strings are different.

For example, Hamming distance between two words “PLANT” and “POINT” is given by counting the number of the positions at which the corresponding character in the two words “PLANT” and “POINT” are different.

Letter no in the two words	Letter in the word PLANT	Corresponding letter in the word POINT	Match or do not match	Count of distance
1	P	P	Match	0
2	L	O	Do not match	1
3	A	I	Do not match	1
4	N	N	Match	0
5	T	T	Match	0
Hamming Distance between two words “PLANT” and “POINT” =				2

Now, you can try the following Self-Assessment Question.

SAQ 4

In R we have a built-in data set “iris”. Screenshot of the first five rows of this data set is given as follows:

```
> head(iris,5)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1         3.5          1.4         0.2  setosa
2          4.9         3.0          1.4         0.2  setosa
3          4.7         3.2          1.3         0.2  setosa
4          4.6         3.1          1.5         0.2  setosa
5          5.0         3.6          1.4         0.2  setosa
```

Note that the first four variables of this data set are numeric. So, assuming each row of this data set as a point in 4-dimension. For example, point corresponding to the first row is (5.1, 3.5, 1.4, 0.2). Find the distance between the points corresponding to the first and third rows using the distance formula given by (a) Manhattan (b) Euclidean (c) Chebyshev (d) Minkowski for p = 4.

3.7 SUMMARY

A brief summary of what we have covered in this unit is given as follows:

- If we include two symbols $-\infty$ and $+\infty$ in the set of all real numbers then this extended set is called the **extended set of real numbers** and is denoted by $\mathbb{R}^*$. i.e., $\mathbb{R}^* = (-\infty, +\infty) \cup \{-\infty, +\infty\} = [-\infty, +\infty]$.
- An operation $*$ (say) on a non-empty set X is said to be binary if $a * b \in X$, for all $a, b \in X$. If $*$ is a binary operation on a non-empty set X then we also say that X is **closed with respect to the operation** $*$. If so then we also say that **closure property** holds in X with respect to the operation $*$.
- Let Ω be a non-empty set and $\mathcal{F}$ be a collection of subsets of Ω then we say that the class $\mathcal{F}$ forms a **sigma algebra** (σ-algebra) or **sigma field** (σ**-field**) if (i) $\phi, \Omega \in \mathcal{F}$, (ii) $A \in \mathcal{F} \Rightarrow A^c (= \Omega \setminus A) \in \mathcal{F}$, (iii) $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$
- **Set Function**: Let Ω be a non-empty set and $\mathcal{F}$ be any collection of subsets of Ω then a function $\mu : \mathcal{F} \to [0, +\infty]$ is known as a **set function**.
- **Measure**: Let Ω be a non-empty set and $\mathcal{F}$ be a collection of subsets of Ω which forms a σ-field then a set function $\mu : \mathcal{F} \to [0, +\infty]$ is called a

ignou THE PEOPLE'S UNIVERSITY

measure if it satisfies the following two requirements (i) $\mu(\phi) = 0$, (ii) μ is countably additive.

- The set function $\lambda : \mathcal{I} \to [0, +\infty]$ defined by

$$\lambda\left(I(a, b)\right) = \begin{cases} |b-a|, & \text{if } a, b \in \mathbb{R} \\ +\infty, & \text{if either } a = -\infty \text{ or } b = +\infty \text{ or both} \end{cases} \qquad I(a, b) \in \mathcal{I}$$

is called the **length function**.

- **Characteristic Function**: Let A be any subset of the universal set Ω then characteristic function of A is denoted by χ_A and defined as follows:

$\chi_A : \Omega \to \{0, 1\}$ define by

$$\chi_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases} \qquad \omega \in \Omega$$

- **Simple Function**: Let $(\Omega, \mathcal{F}, \mu)$ be a measure space then a function $\phi : \Omega \to \mathbb{R}$ is said to be a simple function if there is a finite disjoint class $\{E_1, E_2, E_3, \ldots, E_n\}$ of measurable sets (it means $E_i \in \mathcal{F}$) with $\bigcup_{i=1}^{n} E_i = \Omega$ and a finite set $\{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n\}$ of real numbers such that

$\phi(x) = \alpha_i, \quad x \in E_i, \; i = 1, 2, 3, 4, \ldots, n$

- **Step Function**: Let [a, b] be a finite and closed interval. A function $\phi : [a, b] \to \mathbb{R}$ is said to be a step function if:

$\phi(x) = \alpha_i, \; \forall \; x \in (x_{i-1}, x_i), \; i = 1, 2, 3, \ldots, n$ where the ordered set $\{x_0, x_1, x_2, x_3, \ldots, x_n\}$ forms a partition of [a, b].

- **Distance Function**: Let X be a non-empty set then distance function on X is a function which takes a pair of points from X as input and gives a non-negative real number as output. That is distance function looks like $d : X \times X \to [0, \infty]$.

- A distance function which satisfies four properties (i) non-negative (ii) coincidence (iii) symmetry, and (iv) triangle inequality is called a metric.

- Here we have discussed five types of distance functions namely:

(a) Minkowski Distance (b) Manhattan Distance (c) Euclidean Distance

(d) Chebyshev Distance, and (e) Hamming Distance

3.8 TERMINAL QUESTIONS

1. If Ω = {a, b, c, d} then what is the smallest σ-algebra containing the sets A = {a, b} and B = {b, c} on Ω.

2. Consider the experiment of throwing a fair die. If Ω be the sample space of this experiment and suppose $\mathcal{F}$ is the smallest σ-algebra on Ω containing the event E = {2, 5} then answer the following questions.

(a) Write Ω as a set (b) Write $\mathcal{F}$ (c) Define appropriate probability measure (d) Specify probability space (e) What is measure of the event {2, 5}.

3. Prove that every non-empty set can be treated as a metric space.

3.9 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1. As per the rules of the world of extended real numbers appropriate number or symbol $-\infty$ or $+\infty$ at the place of ? sign are given as follows.

(i) $(-3)(-\infty) = +\infty$ (ii) $(5)(-\infty) = -\infty$ (iii) $(-10)(+\infty) = -\infty$

(iv) $-5 + (-\infty) = -\infty$ (v) $15 + (+\infty) = +\infty$ (vi) $-7 + (-\infty) = -\infty$

(vii) $-9 + (+\infty) = +\infty$ (viii) $(0)(-\infty) = 0$ (ix) $(0)(+\infty) = 0$

(x) $(+\infty) + (-\infty) = \text{Not defined}$

2. We know that the smallest σ-field containing the sets A and B on Ω is given by

$$\mathcal{F} = \begin{Bmatrix} \phi,\ A,\ B,\ A^c,\ B^c,\ A \cap B,\ A \cap B^c,\ A^c \cap B,\ A^c \cap B^c,\ A \cup B, \\ A \cup B^c,\ A^c \cup B,\ A^c \cup B^c,\ A \Delta B,\ (A \Delta B)^c,\ \Omega \end{Bmatrix}$$

In our case A = {a}, B = {a, c} and $\Omega = \{a, b, c, d\}$ so it becomes

$$\mathcal{F} = \{\phi, \{a\}, \{a, c\}, \{b, c, d\}, \{b, d\}, \{c\}, \{a, b, d\}, \{a, b, c, d\}\}$$

3. We know that length function $\lambda : \mathcal{I} \to [0, +\infty]$ is defined by

$$\lambda(I(a, b)) = \begin{cases} |b - a|, & \text{if } a, b \in \mathbb{R} \\ +\infty, & \text{if either } a = -\infty \text{ or } b = +\infty \text{ or both} \end{cases} \qquad I(a, b) \in \mathcal{I}$$

Therefore, images of the given members of $\mathcal{I}$ under the length function λ are given by

(a) $\lambda([4, 9]) = |9 - 4| = |5| = 5$

(b) $\lambda((4, 9]) = |9 - 4| = |5| = 5$

(c) $\lambda([4, 9)) = |9 - 4| = |5| = 5$

(d) $\lambda((4, 9)) = |9 - 4| = |5| = 5$

(e) $\lambda([-\infty, 9]) = +\infty$

(f) $\lambda([-\infty, 9)) = +\infty$

(g) $\lambda([4, +\infty]) = +\infty$

(h) $\lambda([4, +\infty)) = +\infty$

(i) $\lambda([-\infty, +\infty]) = +\infty$

(j) $\lambda([-\infty, +\infty)) = +\infty$

(k) $\lambda\left((-\infty, +\infty]\right) = +\infty$

(l) $\lambda\left((-\infty, +\infty)\right) = +\infty$

4. Let us denote points corresponding to the first and the third rows by A and B respectively. So, we have

$A(5.1, 3.5, 1.4, 0.2) = A(x_1, x_2, x_3, x_4)$ and

$B(4.7, 3.2, 1.3, 0.2) = B(y_1, y_2, y_3, y_4)$

Now, required distances are given by

(a) $$AB_{\text{Manhattan}} = d(x, y) = |x_1 - y_1| + |x_2 - y_2| + |x_3 - y_3| + |x_4 - y_4|$$
$$= |5.1 - 4.7| + |3.5 - 3.2| + |1.4 - 1.3| + |0.2 - 0.2|$$
$$= |0.4| + |0.3| + |0.1| + |0| = 0.4 + 0.3 + 0.1 + 0 = 0.8$$

(b) $$AB_{\text{Euclidean}} = d(x, y) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2 + |x_4 - y_4|^2}$$
$$= \sqrt{|5.1 - 4.7|^2 + |3.5 - 3.2|^2 + |1.4 - 1.3|^2 + |0.2 - 0.2|^2}$$
$$= \sqrt{|0.4|^2 + |0.3|^2 + |0.1|^2 + |0|^2} = \sqrt{0.16 + 0.09 + 0.01 + 0}$$
$$= \sqrt{0.26} \simeq 0.509902$$

(c) $$AB_{\text{Chebyshev}} = d(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|, |x_3 - y_3|, |x_4 - y_4|\}$$
$$= \max\{|5.1 - 4.7|, |3.5 - 3.2|, |1.4 - 1.3|, |0.2 - 0.2|\}$$
$$= \max\{|0.4|, |0.3|, |0.1|, |0|\} = \max\{0.4, 0.3, 0.1, 0\} = 0.4$$

(d) $$AB_{\substack{\text{Minkowski}\\(p=4)}} = d(x, y) = \sqrt[4]{|x_1 - y_1|^4 + |x_2 - y_2|^4 + |x_3 - y_3|^4 + |x_4 - y_4|^4}$$
$$= \sqrt[4]{|5.1 - 4.7|^4 + |3.5 - 3.2|^4 + |1.4 - 1.3|^4 + |0.2 - 0.2|^4}$$
$$= \sqrt[4]{|0.4|^4 + |0.3|^4 + |0.1|^4 + |0|^4} = \sqrt[4]{0.0256 + 0.0081 + 0.0001 + 0}$$
$$= \sqrt[4]{0.0338} \simeq 0.428775$$

Terminal Questions

1. We know that the smallest σ-algebra containing the sets A and B on Ω is given by

$$\mathcal{F} = \left\{\begin{array}{l} \phi, A, B, A^c, B^c, A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c, A \cup B, \\ A \cup B^c, A^c \cup B, A^c \cup B^c, A \Delta B, (A \Delta B)^c, \Omega \end{array}\right\}$$

In our case A = {a, b}, B = {b, c} and $\Omega = \{a, b, c, d\}$ so it becomes

$$\mathcal{F} = \left\{\begin{array}{l} \phi, \{a, b\}, \{b, c\}, \{c, d\}, \{a, d\}, \{b\}, \{a\}, \{c\}, \{d\}, \{a, b, c\}, \\ \{a, b, d\}, \{b, c, d\}, \{a, c, d\}, \{a, c\}, \{b, d\}, \{a, b, c, d\} \end{array}\right\}$$

i.e., $\mathcal{F} = P(\Omega)$ = the largest σ-algebra on Ω

2 (a) $\Omega = \{1, 2, 3, 4, 5, 6\}$

(b) The smallest σ-algebra on Ω containing the event E = {2, 5} is given by $\mathcal{F} = \{\phi, E, E^c, \Omega\} = \{\phi, \{2, 5\}, \{1, 3, 4, 6\}, \Omega\}$

(c) Appropriate probability measure is defined by $\mu(E) = \dfrac{n(E)}{n(\Omega)}$, $E \in \mathcal{F}$

(d) Probability space in this case is

$$(\Omega, \mathcal{F}, \mu) = \left(\{1, 2, 3, 4, 5, 6\}, \{\phi, \{2, 5\}, \{1, 3, 4, 6\}, \Omega\}, \mu(E) = \frac{n(E)}{n(\Omega)}\right)$$

(e) By definition of probability measure as define in part (c), probability measure of the event E = {2, 5} is given by $\mu(\{2, 5\}) = \dfrac{n(\{2, 5\})}{n(\Omega)} = \dfrac{2}{6} = \dfrac{1}{3}$

3. Let X be a non-empty set. Define a function $f : X \times X \to \mathbb{R}$ by

$$d(x, y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases} \quad x, y \in X \qquad \dots (3.57)$$

We claim that d is a metric on X.

(i) **Non-negativity**: By definition of d it takes only two values 0 and 1. So, d(x, y) ≥ 0

(ii) **Coincidence**: d(x, y) = 0 iff x = y [Using (3.57)]

(iii) **Symmetry**: **Case I**: If x = y, then y = x. So, d(x, y) = 0 = d(y, x)

Case II: If x ≠ y, then y ≠ x. So, d(x, y) = 1 = d(y, x)

(iv) **Triangle Inequality**: Required to prove

$$d(x, y) \leq d(x, z) + d(z, y) \qquad \dots (3.58)$$

Case I: x = y, but y ≠ z, then x ≠ z. So, (3.58) reduces to 0 ≤ 1 + 1 which is true

Case II: x = y, and y = z, then x = z. So, (3.58) reduces to 0 ≤ 0 + 0 which is true

Case III: x ≠ y, but y = z, then x ≠ z. So, (3.58) reduces to 1 ≤ 1 + 0 which is true

Case IV: x ≠ y, and y ≠ z, then

Sub case I: x = z. So, (3.58) reduces to 1 ≤ 0 + 1 which is true

Sub case II: x ≠ z. So, (3.58) reduces to 1 ≤ 1 + 1 which is true

Thus, (3.58) holds for all $x, y, z \in X$.

Hence, d is a metric on X and so (X, d) is a metric space.

Remark 4: This metric d is called discrete metric on X. … (3.59)