UNIT 7

CONVEX AND CONCAVE FUNCTION

Structure

7.1 INTRODUCTION

The need for ideas of convex set and convex function is required in optimisation problems discussed in courses MST-022, MST-026 and MSTE-011. In this unit, you will study some basic results related to the convex set and convex function. To define a convex set, we will use the idea of linear combination and convex combination of two points in the underlying set. The convex combination is a particular case of affine combination. So, linear, affine and convex combinations are discussed in Sec. 7.2. Conic combination is also defined in Sec. 7.2. Definition of affine and convex sets and their simple examples are discussed in Sec. 7.3. After understanding what is a convex set, some properties of convex set are discussed in Sec. 7.4. The domain of a convex function is a convex set and you have studied about convex set in Secs. 7.3 and 7.4. So, the convex function is defined in Sec. 7.5. Epigraph of a function and some properties of the convex function are discussed in Sec. 7.6. But the idea of the concave function is also related to the convex function so the concave function is also discussed in the same section.

What we have discussed in this unit is summarised in Sec. 7.7. 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, more questions based on the entire unit are given in Sec. 7.8 under the heading Terminal Questions.

Due to the reason mentioned in Sec. 1.1 of Unit 1 of this course, solutions of all the SAQs and Terminal Questions are given in Sec. 7.9.

In the next unit, you will study gamma and beta functions.

Expected Learning Outcomes

After completing this unit, you should be able to:

- define what are linear, affine, conic and convex combinations;
- explain what we mean by affine and convex sets and can provide a lot of examples of each;
- list and prove some properties of the convex set;
- define convex and concave functions; and
- list some properties of a convex function and derive some of them.

7.2 LINEAR, AFFINE, CONIC AND CONVEX COMBINATIONS

As mentioned in Sec. 7.1 in courses MST-022, MST-026 and MSTE-011 you will deal with the optimum solution to a given problem. Suppose you are interested in obtaining the optimum value of the function $f(x_1, x_2, x_3, \ldots, x_n)$ known as the objective function. The variables $x_1, x_2, x_3, \ldots, x_n$ are known as decision variables. If there is no restriction on decision variables then you have to obtain the optimum value of the objective function on the entire $\mathbb{R}^n$. But generally, there are restrictions on decision variables in terms of available man-hours, machine hours, money, storage capacity, etc. Restrictions on decision variables reduce $\mathbb{R}^n$ to some subset C of $\mathbb{R}^n$, where all restrictions agree. If this common reduced subset C of $\mathbb{R}^n$ is a convex set as well as objective function and all the constraints in an optimisation problem are linear functions of decision variables then the optimisation problem is known as a **linear programming problem**. You have learnt the graphical method to solve linear programming problems in earlier classes. You will study some more methods to solve LPP in the course MSTE-011. An optimisation problem where either objective function or constraints on decision variables or both are non-linear function(s) of decision variables, then it is known as a **non-linear programming problem**. In both linear and non-linear programming problems convex set and convex function play an important role. So, this unit is devoted to discuss both the convex set and convex function. But the idea of a convex set is based on the idea of a convex combination. So, in this section, we will discuss different types of combinations of points in a set given as follows.

- Linear Combination
- Affine Combination
- Conic Combination, and
- Convex Combination

Let us discuss these one at a time.

Linear Combination

Before defining linear combination first, we have to clarify what are coefficients (scalars) and points (vectors) in the world of convex analysis (linear algebra). Here we are discussing these concepts from a data science or machine learning point of view. So, in an expression of the form

$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \alpha_k x_k,$

we say that $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_k$ are **coefficients** or **scalars** and they may be any real numbers. While $x_1, x_2, x_3, \ldots, x_k$ are known as **points** or **vectors** in $\mathbb{R}^n$, where $n = 1, 2, 3, \ldots$

For example, in the expression: $3(4, 2) + 8(-2, 7)$, coefficients or scalars are $\alpha_1 = 3$, $\alpha_2 = 8$, while points or vectors are $x_1 = (4, 2)$, $x_2 = (-2, 7)$. In the expression $2(4, 1, 2) - 7(8, 9, 5)$ or $2(4, 1, 2) + (-7)(8, 9, 5)$, coefficients or scalars are $\alpha_1 = 2$, $\alpha_2 = -7$, while points or vectors are $x_1 = (4, 1, 2)$, $x_2 = (8, 9, 5)$. In the first expression points or vectors are elements of $\mathbb{R}^2$ and in the second expression points or vectors are elements of $\mathbb{R}^3$.

Remark 1: From here onwards we will call α_i's as scalars instead of coefficients and continue this terminology in the courses MST-022 and MST-026 while x_i's will be called points in this unit and we will call them vectors in the courses MST-022 and MST-026 due to standard practice that is followed in most of the books on these topics.

Now, we define linear combination. A point x is a **linear combination** of points $x_1, x_2, x_3, \ldots, x_k$ if and only if

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \alpha_k x_k, \quad \alpha_i \in \mathbb{R}, \ i = 1, 2, 3, \ldots, k \qquad \ldots (7.1)$$

For example, a point $x = (3, 7)$ is a linear combination of points $x_1 = (1, 1)$, $x_2 = (1, -1)$ as there exist real numbers $\alpha_1 = 5$, $\alpha_2 = -2$, such that

$$(3, 7) = 5(1, 1) + (-2)(1, -1) \quad \left[\begin{array}{l} \because \text{ if } (3, 7) = \alpha_1(1, 1) + \alpha_2(1, -1) \\ \Rightarrow (3, 7) = (\alpha_1 + \alpha_2, \alpha_1 - \alpha_2) \Rightarrow \alpha_1 + \alpha_2 = 3, \\ \alpha_1 - \alpha_2 = 7. \text{ After solving, we get } \alpha_1 = 5, \alpha_2 = -2 \end{array}\right]$$

Affine Combination

In a linear combination there was no restriction on values of α_i's. In linear combination α_i's. may be any real numbers. But in an affine combination, there is a restriction on their sum defined as follows.

A point x is an **affine combination** of points $x_1, x_2, x_3, \ldots, x_k$ if and only if

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \alpha_k x_k, \ \alpha_i \in \mathbb{R}, \ i = 1, 2, 3, \ldots, k \text{ and } \sum_{i=1}^{k} \alpha_i = 1 \ldots (7.2)$$

For example, point $x = (1, 5)$ is an affine combination of points $x_1 = (1, 1)$, $x_2 = (1, -1)$ as there exist real numbers $\alpha_1 = 3$, $\alpha_2 = -2$, such that

$$(1, 5) = 3(1, 1) + (-2)(1, -1) \quad \left[\begin{array}{r} \because 3(1, 1) + (-2)(1, -1) = (3, 3) + (-2, 2) \\ = (3 - 2, 3 + 2) = (1, 5) \end{array}\right]$$

Conic Combination

In affine combination there was a restriction on sum of all α_i's. But in conic combination there is no restriction on their sum but there are restrictions on individual α_i's defined as follows.

A point x is a **conic combination** of points $x_1, x_2, x_3, \ldots, x_k$ if and only if

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \alpha_k x_k, \quad \alpha_i \in \mathbb{R}, \text{ and } \alpha_i > 0, \quad i = 1, 2, 3, \ldots, k \quad \ldots (7.3)$$

For example, point $x = (7, -1)$ is a conic combination of points $x_1 = (1, 1)$, $x_2 = (1, -1)$ as there exist real numbers $\alpha_1 = 3, \ \alpha_2 = 4,$ such that $\alpha_1 > 0, \ \alpha_2 > 0,$ and

$$(7, -1) = 3(1, 1) + 4(1, -1) \quad \left[\begin{aligned} \because 3(1, 1) + 4(1, -1) &= (3, 3) + (4, -4) \\ &= (3+4, 3-4) = (7, -1) \end{aligned}\right]$$

Convex Combination

In affine combination there was a restriction only on sum of all α_i's. In conic combination there was a restriction only on individual α_i's. But in convex combination there are restrictions on both individual α_i's as well as on sum of all α_i's defined as follows.

A point x is a **convex combination** of points $x_1, x_2, x_3, \ldots, x_k$ if and only if

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k, \ \alpha_i \in \mathbb{R}, \ \alpha_i \ge 0, \ i = 1, 2, 3, \ldots, k \text{ and } \sum_{i=1}^{k} \alpha_i = 1 \quad \ldots (7.4)$$

For example, point $x = \left(1, -\frac{1}{2}\right)$ is a convex combination of points $x_1 = (1, 1)$, $x_2 = (1, -1)$ as there exist real numbers $\alpha_1 = \frac{1}{4}, \ \alpha_2 = \frac{3}{4},$ such that $\alpha_1 > 0, \ \alpha_2 > 0, \ \alpha_1 + \alpha_2 = 1$ and

$$\left(1, -\frac{1}{2}\right) = \frac{1}{4}(1, 1) + \frac{3}{4}(1, -1) \quad \left[\begin{aligned} \because \frac{1}{4}(1, 1) + \frac{3}{4}(1, -1) &= \left(\frac{1}{4}, \frac{1}{4}\right) + \left(\frac{3}{4}, \frac{-3}{4}\right) \\ &= \left(\frac{1}{4} + \frac{3}{4}, \frac{1}{4} + \frac{-3}{4}\right) = \left(1, -\frac{1}{2}\right) \end{aligned}\right]$$

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

SAQ 1

In each part identify whether given combination is affine or conic or convex.

(a) $-2(7, 5) + 3(0, -1)$ (b) $3(7, 5) + 2(0, -1)$ (c) $0.3(7, 5) + 0.7(0, -1)$

7.3 AFFINE AND CONVEX SETS

This section is divided into two subsections. In the first subsection we will discuss affine set and in the second subsection we will discuss convex set.

7.3.1 Affine Set

In Sec. 7.2 you have understood what we mean by an affine combination which will help you in understanding an affine set. So, let us define an affine set.

A set S is said to be an **affine set** if affine combination of any two members of it is in the set S. That is for all $x, y \in S$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ with $\alpha_1 + \alpha_2 = 1$ then $\alpha_1 x + \alpha_2 y \in S.$ Let us try to understand this definition geometrically. This definition says that:

$$\alpha_1 x + \alpha_2 y \in S, \ \ \forall \ x, y \in S \ \text{ and } \ \alpha_1, \alpha_2 \in \mathbb{R} \ \text{ with } \ \alpha_1 + \alpha_2 = 1 \qquad \ldots (7.5)$$

$$\Rightarrow (1 - \alpha_2)x + \alpha_2 y \in S, \ \ \forall \ x, y \in S, \ \alpha_2 \in \mathbb{R} \ \left[\because \alpha_1 + \alpha_2 = 1 \Rightarrow \alpha_1 = 1 - \alpha_2\right]$$

Or taking $\alpha_2 = \alpha,$ we get

$$(1 - \alpha)x + \alpha y \in S, \ \ \forall \ x, y \in S, \ \alpha \in \mathbb{R} \qquad \ldots (7.6)$$

On comparing equation (7.6) with equation (6.13) of the previous unit, we see that (7.6) represents all points of a straight line passing through the points x and y. But as per the requirements mentioned in (7.6) all these points belong to the set S. In other words, whole line lies inside the set S. So, **geometrically** a set S is said to be an **affine set** if we take any two points in the set S then the whole line joining the two points should lie in the set S.

After getting geometrical interpretation of an affine set, you can easily name some sets which are affine sets. Think geometrically what are the sets which contain whole line joining any two points on them. Recall that some such sets you have studied in the previous unit which contain whole line joining any two points in that set. Obviously, now some of the names which are coming in your mind includes: (i) a line in 2 or higher dimension (ii) plane in 3-dimension (iii) hyperplane in more than three dimension (iv) $\mathbb{R}^n$ itself for each $n = 1, 2, 3, \ldots,$ etc. Two more sets which always fall in this category are empty set and the singleton set. Why empty set is an affine set? It is affine set because there are no two points in it to give a counter example to fail the definition. So, requirements to become affine set automatically satisfied. Why singleton set is an affine set? Let us explain it. Suppose S = {x}. Let $u, v \in S,$ then u = x and v = x you know why we have taken both u and v as x because S has only one element x. So, there is no other option for u and v. Now, take any $\alpha_1, \alpha_2 \in \mathbb{R},$ such that $\alpha_1 + \alpha_2 = 1$ then

$$\begin{aligned}
\alpha_1 u + \alpha_2 v &= \alpha_1 u + (1 - \alpha_1)v \quad \left[\because \alpha_1 + \alpha_2 = 1 \Rightarrow \alpha_2 = 1 - \alpha_1\right] \\
&= \alpha_1 x + (1 - \alpha_1)x \quad \left[\because u = x \ \text{ and } \ v = x\right] \\
&= (\alpha_1 + 1 - \alpha_1)x = x \in S
\end{aligned}$$

Hence, $\alpha_1 u + \alpha_2 v \in S, \ \ \forall \ u, v \in S$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ with $\alpha_1 + \alpha_2 = 1$

Thus, S is an affine set and therefore,

every singleton set is an affine set. … (7.7)

Now, let us do some examples related to affine set.

Example 1: Give an example of a set which is not an affine set.

Solution: Consider the half plane

$$3x + 2y \le 6 \qquad \ldots (7.8)$$

discussed in Example 14 in the previous unit. You may refer Remark 5 mentioned in the previous unit for the notion of half plane. We shall prove that given half plane is not an affine set using two ways (i) graphically so that you can understand it intuitively (ii) using definition.

Graphically: Let us extend Fig. 6.11 (b) and added information is shown in Fig. 7.1 (a). Note that points $A(0, 1)$ and $B(1, 1)$ lie in the shaded region which represent the set S where $S = \{(x, y) \in \mathbb{R}^2 : 3x + 2y \le 6\}$. Now, the line passing through the points A and B contains the point C(3, 1) but the point C(3, 1) does not lie in the set S. So, set S is not an affine set because we know that a set is said to be an affine set if it contains whole line passing through any two points of it. This completes geometric proof. Remember there is nothing special in the points $A(0, 1)$ and $B(1, 1)$ you can take any other two points in S such that the line joining them does not wholly lie inside the set S.

Using Definition: Let S be as defined in graphical method. We have to show that S is not an affine set using definition. To do so we have to give a counter example where definition of affine set fails. So, let us consider two points $x = (-1, 0),\ y = (1, 0)$ obviously both belong to the half plane given by (7.8) $[\because 3(-1) + 2(0) \le 6 \text{ true and } 3(1) + 2(0) \le 6 \text{ true}]$. Now consider $\alpha_1 = -1,\ \alpha_2 = 2,$ so $\alpha_1, \alpha_2 \in \mathbb{R}$ and $\alpha_1 + \alpha_2 = 1$ but

$$\alpha_1 x + \alpha_2 y = -1(-1, 0) + 2(1, 0)$$
$$= (1, 0) + (2, 0) = (3, 0) \notin S \quad [\because 3(3) + 2(0) \le 6 \text{ which is False}]$$

$\Rightarrow$ S is not an affine set. To have a look on this argument graphically you may refer Fig. 7.1 (b).

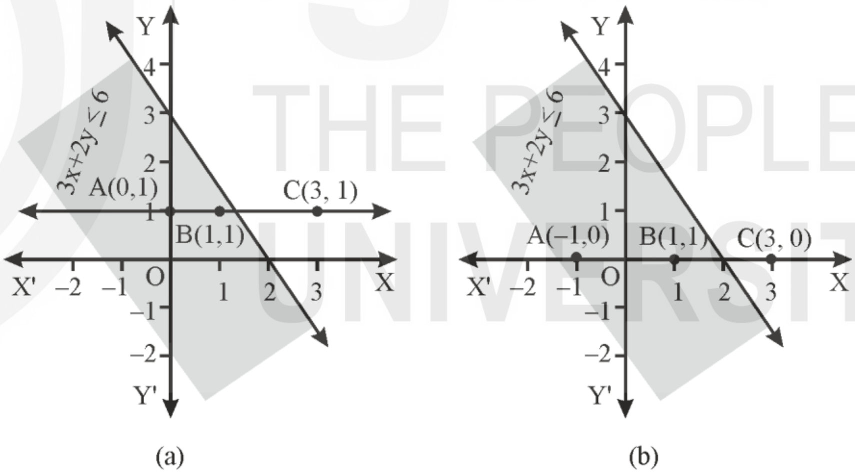

Fig. 7.1: Visualisation of (a) a half plane 3x + 2y ≤ 6 as an example of a set which is not an affine set (b) counter example given in the proof by definition

Example 2: Prove that hyperplane is an affine set.

Solution: We know that equation of a hyperplane is given by (you may refer equation (6.77) in Unit 6)

$$\omega^T x + b = 0 \qquad \dots (7.9)$$

Let u, v lie on the hyperplane given by equation (7.9). Required to prove for any $\alpha_1, \alpha_2 \in \mathbb{R}$ with $\alpha_1 + \alpha_2 = 1 \Rightarrow \alpha_1 u + \alpha_2 v$ lies on the hyperplane.

i.e., required to prove $\omega^T(\alpha_1 u + \alpha_2 v) + b = 0$

Since u and v lie on the hyperplane (7.9), therefore

$$\omega^T u + b = 0 \text{ or } \omega^T u = -b \qquad \dots (7.10)$$

$\omega^T v + b = 0$ or $\omega^T v = -b$ … (7.11)

Let $y = \alpha_1 u + \alpha_2 v$ then

$$\omega^T y = \omega^T (\alpha_1 u + \alpha_2 v) = \alpha_1 \omega^T u + \alpha_2 \omega^T v$$
$$= \alpha_1(-b) + \alpha_2(-b) \qquad [\text{Using } (7.10) \text{ and } (7.11)]$$
$$= (\alpha_1 + \alpha_2)(-b)$$
$$= -b \qquad [\because \alpha_1 + \alpha_2 = 1]$$

$\Rightarrow \omega^T (\alpha_1 u + \alpha_2 v) = -b$

$\Rightarrow \omega^T (\alpha_1 u + \alpha_2 v) + b = 0$

$\Rightarrow \alpha_1 u + \alpha_2 v$ lies on the hyperplane (7.9). Hence, hyperplane is an affine set.

7.3.2 Convex Set

In Sec. 7.2 you have understood what we mean by a convex combination which will help you in understanding a convex set. So, let us define a convex set.

A set S is said to be a **convex set** if convex combination of any two members of it is in the set S. That is if for all $x, y \in S$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ with $\alpha_1 \geq 0$, $\alpha_2 \geq 0$ and $\alpha_1 + \alpha_2 = 1$ then $\alpha_1 x + \alpha_2 y \in S$. Let us try to understand this definition geometrically. This definition says that:

$\alpha_1 x + \alpha_2 y \in S, \forall x, y \in S$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ with $\alpha_1 \geq 0, \alpha_2 \geq 0, \alpha_1 + \alpha_2 = 1$… (7.12)

$\Rightarrow (1-\alpha_2)x + \alpha_2 y \in S, \forall\ x, y \in S, \alpha_2 \in \mathbb{R}, \alpha_2 \in [0, 1]$ $[\because \alpha_1 + \alpha_2 = 1 \Rightarrow \alpha_1 = 1 - \alpha_2]$

Or taking $\alpha_2 = \alpha$, we get

$(1-\alpha)x + \alpha y \in S, \quad \forall\ x, y \in S, \ \alpha \in \mathbb{R}, \ \alpha \in [0, 1]$ … (7.13)

On comparing equation (7.13) with equation (6.13) of the previous unit, we see that (7.13) represents all points on the line segment joining points x and y. But as per the requirements mentioned in (7.13) all these points belong to the set S. In other words, whole line segment joining points x and y lies inside the set S. So, **geometrically** a set S is said to be a **convex set** if we take any two points in the set S then the whole line segment joining points x and y lies inside the set S.

After getting geometrical interpretation of a convex set, you can easily name some sets which are convex sets. Think geometrically what are the sets which contain whole line segment joining any two points on them. First note that all affine sets are convex but converse is not true. But there are many other sets which are convex but not affine the reason for it is **to become a convex set only line segment joining two points should be inside the set but to become affine set whole line joining two points should be inside the set. So, restriction of affine set is more hard compare to convex set**. The name of some sets which are not affine sets but are convex sets includes: circular disk, semicircular disk, square region, rectangle region, region of a parallelogram shape, triangular region, cubic region, half plane, etc. Recall that empty set and the singleton set are affine sets and so convex also…(7.14)

Now, let us do some examples related to convex set.

Example 3: Give an example of a set which is not a convex set.

Solution: We know that equation of a circle having centre at (h, k) and radius r is given by

$$(x-h)^2 + (y-k)^2 = r^2 \qquad \ldots (7.15)$$

Consider the circle having centre at origin (0, 0) and radius equal to 2. So, its equation is given by

$$(x-0)^2 + (y-0)^2 = 2^2 \text{ or } x^2 + y^2 = 4 \qquad \ldots (7.16)$$

We claim that it is not a convex set. To prove our claim, we have to show that there exist at least two points on (7.16) such that line segment joining those points does not completely lie on (7.16). Consider points A(2, 0) and B(0, 2) and join AB refer Fig. 7.2 (a). Now, midpoint of the line segment AB is $C\left(\frac{2+0}{2}, \frac{0+2}{2}\right) = C(1, 1)$. Obviously, point C does not lie on the circle given by (7.16). $\left[\because 1^2 + 1^2 = 4, \text{ i.e., } 2 = 4 \text{ which is not true}\right]$. Hence, circle given by (7.16) is not a convex set.

Remark 2: If instead of circle we have circular disk say $(x-0)^2 + (y-0)^2 \le 2^2$ or $x^2 + y^2 \le 4$ then it is a convex set. To see its proof in general case, refer Solution of Terminal Question 2.

Example 4: Prove that the square formed by four lines $x + y = 2,\ x - y = -2,\ x + y = -2,\ x - y = 2,$ is not a convex set.

Solution: Writing equations of 4 sides of the square in intercept form, we get

$$\frac{x}{2} + \frac{y}{2} = 1,\ \frac{x}{-2} + \frac{y}{2} = 1,\ \frac{x}{-2} + \frac{y}{-2} = 1,\ \frac{x}{2} + \frac{y}{-2} = 1 \qquad \ldots (7.17)$$

Square ABCD formed by these four lines is shown in Fig. 7.2 (b). Now, points A(2, 0) and $C(-2, 0)$ lie on the square but the midpoint of the line segment AC is the origin O(0, 0) which does not lie on the square ABCD. Hence, the square formed by four lines given by equations (7.17) is not a convex set.

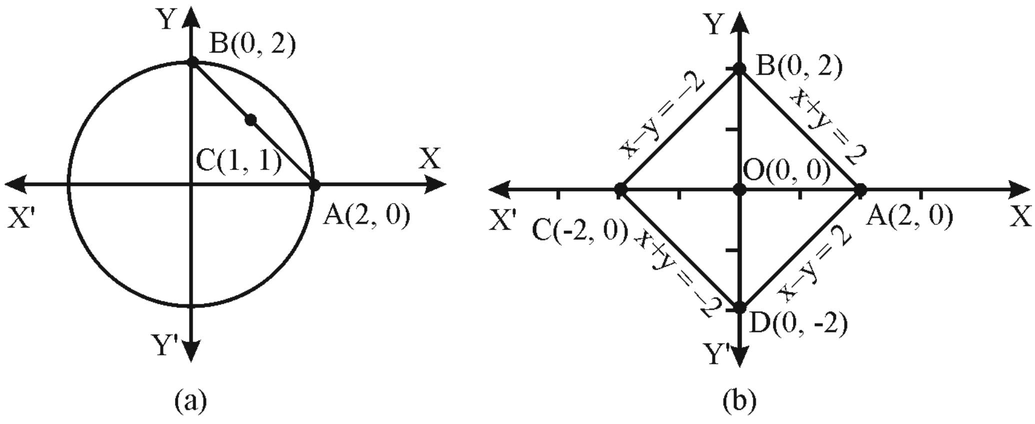

Fig. 7.2: Visualisation of a (a) a circle as an example of a set which is not convex (b) square as an example of a set which is not convex

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

SAQ 2

(a) Does the set of natural numbers is a convex set?

(b) Does union of two convex sets is always convex? Give proper justification in support of your answer.

7.4 PROPERTIES OF CONVEX SET

In the previous section you have seen many convex sets. You can generate more convex sets from given convex set(s) using some properties of convex set. In this section we will state and prove three important properties of convex set.

Property 1: Prove that intersection of convex sets is also a convex set.

Solution: Let $C_1, C_2, C_3, \ldots$ be a countable collection of convex sets.

Required to prove $\bigcap_{n=1}^{\infty} C_n$ is a convex set. That is required to prove for each $\alpha \in [0, 1]$

$$(1-\alpha)x + \alpha y \in \bigcap_{n=1}^{\infty} C_n \quad \forall\ x, y \in \bigcap_{n=1}^{\infty} C_n \text{ and } \alpha \in [0, 1]$$

Since $x, y \in \bigcap_{n=1}^{\infty} C_n$

$\Rightarrow x, y \in C_n \quad \forall\ n,\ n = 1, 2, 3, \ldots$

$\Rightarrow (1-\alpha)x + \alpha y \in C_n \ \forall\ n,\ n = 1, 2, 3, \ldots$ and $\forall \alpha \in [0, 1]$ $[\because$ Each C_n is convex$]$

$\Rightarrow (1-\alpha)x + \alpha y \in \bigcap_{n=1}^{\infty} C_n \quad \forall \alpha \in [0, 1]$ and $\forall\ x, y \in \bigcap_{n=1}^{\infty} C_n$

Hence, $\bigcap_{n=1}^{\infty} C_n$ is a convex set. … (7.18)

Property 2: If C is a convex set and $\alpha \in \mathbb{R},\ \alpha \geq 0,$ then prove that αC is also a convex set.

Solution: If $\alpha = 0,$ then $\alpha C = \{0\}$ and we know that every singleton set is a convex set and hence αC is also a convex set. Now, let $\alpha \in \mathbb{R}$ and $\alpha > 0,$ then required to prove $(1-\lambda)x + \lambda y \in \alpha C \quad \forall\ x, y \in \alpha C$ and $\lambda \in [0, 1]$

Since $x, y \in \alpha C$

$\Rightarrow \exists\ u, v \in C$ such that $x = \alpha u,\ y = \alpha v$ … (7.19)

Now, $(1-\lambda)x + \lambda y = (1-\lambda)\alpha u + \lambda \alpha v$ $\quad$ [Using (7.19)]

$$= \alpha\left((1-\lambda)u + \lambda v\right) \in \alpha C \quad \left[\begin{array}{l}\because \text{ C is convex so } (1-\lambda)u + \lambda v \in C \\ \Rightarrow \alpha\left((1-\lambda)u + \lambda v\right) \in \alpha C\end{array}\right]$$

Hence, $(1-\lambda)x + \lambda y \in \alpha C \quad \forall\ x, y \in \alpha C$ and $\lambda \in [0, 1]$

Hence, αC is a convex set. … (7.20)

Property 3: If C_1 and C_2 are two convex sets then prove that $C_1 + C_2$ is also a convex set.

Solution: By definition of sum of two sets, we know that

$$C_1 + C_2 = \{u + v : u \in C_1,\ v \in C_2\} \quad \ldots (7.21)$$

In order to prove that $C_1 + C_2$ is also a convex set we have to prove that

$(1-\alpha)x + \alpha y \in C_1 + C_2 \quad \forall\ x, y \in C_1 + C_2$ and $\alpha \in [0, 1]$

So, let $x, y \in C_1 + C_2$ then $\exists\ u_1, u_2 \in C_1$ and $v_1, v_2 \in C_2$ such that

$$x = u_1 + v_1 \text{ and } y = u_2 + v_2 \quad \ldots (7.22)$$

Now, $(1-\alpha)x + \alpha y = (1-\alpha)(u_1 + v_1) + \alpha(u_2 + v_2)$ [Using (7.22)]

$$= ((1-\alpha)u_1 + \alpha u_2) + ((1-\alpha)v_1 + \alpha v_2) \in C_1 + C_2$$

$$\left[\begin{array}{l}\because u_1, u_2 \in C_1 \text{ and } v_1, v_2 \in C_2 \text{ and } C_1, C_2 \text{ are convex sets} \\ \text{so } (1-\alpha)u_1 + \alpha u_2 \in C_1 \text{ and } (1-\alpha)v_1 + \alpha v_2 \in C_2\end{array}\right]$$

Hence, $(1-\alpha)x + \alpha y \in C_1 + C_2 \quad \forall\ x, y \in C_1 + C_2$ and $\alpha \in [0, 1]$

Hence, $C_1 + C_2$ is a convex set whenever C_1 and C_2 are convex sets… (7.23)

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

SAQ 3

If C_1 and C_2 are two convex sets then which of the following will be definitely convex sets.

(a) $5C_1$ (b) $7C_2 + 9C_1$ (c) $C_1 \cap C_2$ (d) $C_1 \cup C_2$

7.5 DEFINITION OF CONVEX AND CONCAVE FUNCTIONS

Before discussing convex and concave functions let us recall what is local minima, local maxima, global minimum and global maximum which you have studied in earlier classes. Let us explain these taken one at a time.

Local Minima (Maxima): A function y = f(x) is said to have local minima (maxima) at a point x = a in its domain if there exists a $\delta > 0$, such that

$$f(x) \geq f(a) \quad (f(x) \leq f(a)) \ \forall\ x \in (a - \delta, a + \delta) \quad \ldots (7.19)$$

For example, in Fig. 7.3 function y = f(x) has local minima at the points $x_2, x_4,$ and x_6 while this function has local maxima at the points $x_1, x_3,$ and x_5.

Global Minimum (Maximum): A function y = f(x) is said to have global minimum (maximum) at a point x = a in its domain D if

$$f(x) \geq f(a) \quad (f(x) \leq f(a)) \ \forall\ x \in D \quad \ldots (7.20)$$

For example, if $D = [x_1, x_6]$ then graphically (see Fig. 7.3) we see that global minimum of the function is at the point $x = x_2$, while global maximum of the function is at the point $x = x_3$.

So, global minimum means function takes minimum value at that point compare to values of the function at all other points of the domain of the function. Similarly, global maximum means function takes maximum value at that point compare to values of the function at all other points of the domain of the function.

Now, we discuss convex and concave functions. In optimisation problems we are generally interested in global minimum (maximum) instead of local minima (maxima). The main advantage of convex functions is that their local and global minimum (maximum) are the same. This is the reason we prefer to work with convex sets and functions in optimisation problems. We have already discussed about convex sets in the previous section. So, now let us discuss about convex and concave functions.

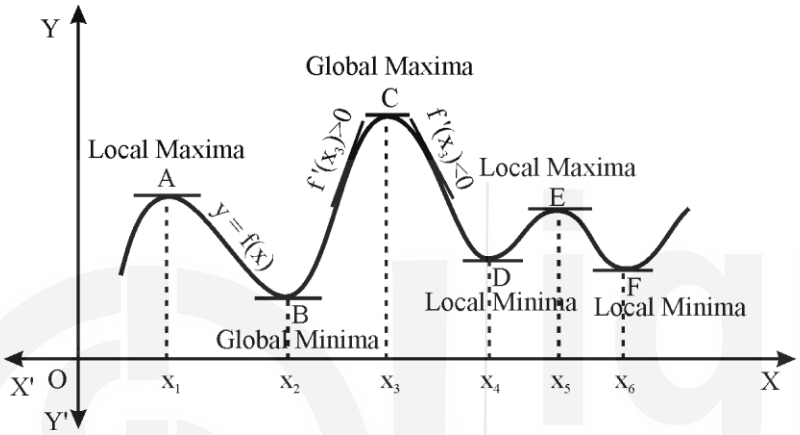

Fig. 7.3: Visualisation of points of local and global minimum and maximum

Convex Function: Let $D \subseteq \mathbb{R}^n,\ n = 1, 2, 3, \ldots, k$ be the convex set. Then a function $f : D \to \mathbb{R}$ is said to be a convex function if

$$f\left((1-\alpha)x + \alpha y\right) \le (1-\alpha)f(x) + \alpha f(y) \quad \forall\ x, y \in D \text{ and } \alpha \in [0, 1] \qquad \ldots (7.21)$$

and it is called concave function if

$$f\left((1-\alpha)x + \alpha y\right) \ge (1-\alpha)f(x) + \alpha f(y) \quad \forall\ x, y \in D \text{ and } \alpha \in [0, 1] \qquad \ldots (7.22)$$

From equations (7.21) and (7.22) note that if

$$f\left((1-\alpha)x + \alpha y\right) = (1-\alpha)f(x) + \alpha f(y) \quad \forall\ x, y \in D \text{ and } \alpha \in [0, 1] \qquad \ldots (7.23)$$

then function f will be both convex and concave.

Before discussing some examples, let us explain geometrically what the definitions of convex and concave functions say.

You know from equation (6.17) in 2-dimension and equation (6.67) in 3-dimension that:

- $(1-\alpha)x + \alpha y,\ \forall\ x, y \in D$ and $\alpha \in [0, 1]$ represents all points on the line segment in D joining points x and y, which are denoted by A and B refer Fig. 7.4 (a) where $D \subset \mathbb{R}$ … (7.24)
- $(1-\alpha)f(x) + \alpha f(y)\ \forall\ x, y \in D$ and $\alpha \in [0, 1]$ represents all points on the line segment joining points f(x) and f(y) in the range set of the function f, refer Fig. 7.4 (a) where points f(x) and f(y) are denoted by points C and D respectively. … (7.25)

Now, equation (7.21) says that the value of the function f obtained on any point of the line segment obtained by joining points x and y is less than or equal to the height of the corresponding point on the chord obtained by joining points (x, f(x)) and (y, f(y)). If you refer to Fig. 7.4 (a) it says that

$$EF \leq EG \qquad \ldots (7.26)$$

Similarly, equation (7.22) says that the value of the function f obtained on any point of the line segment obtained by joining points x and y is greater than or equal to the height of the corresponding point on the chord obtained by joining points (x, f(x)) and (y, f(y)). If you refer to Fig. 7.4 (b) it says that

$$EF \geq EG \qquad \ldots (7.27)$$

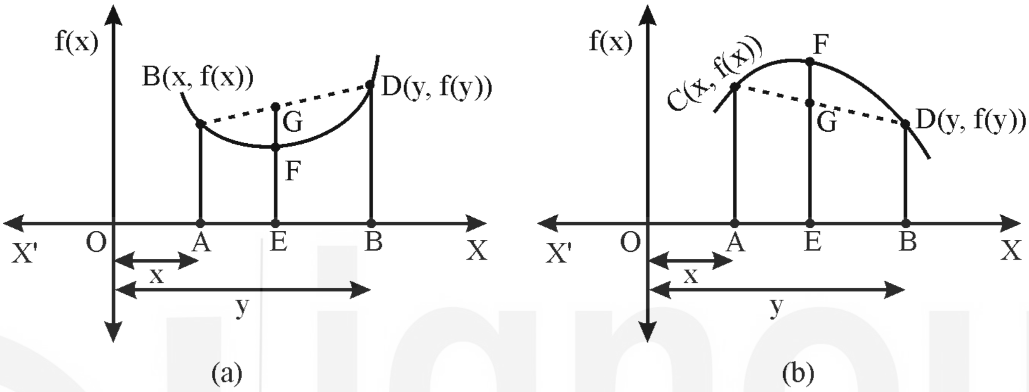

Fig. 7.4: Visualisation of (a) convex function (b) concave function

In other words, you can say that convex functions are open upward, while concave functions are open downward. To connect it with optimisation problems you can say that a convex function has global minimum value while a concave function has global maximum value.

Remark 3: Note that if a function f is convex then – f(x) will be concave function.

Now, let us do some examples.

Example 5: Prove that linear function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = ax + b, \quad a, b, x \in \mathbb{R} \qquad \ldots (7.28)$$

is both convex and concave.

Solution: Let $x,\ y \in \mathbb{R} =$ Domain of f and $\alpha \in [0, 1]$, then required to prove $f\left((1-\alpha)x + \alpha y\right) = (1-\alpha)f(x) + \alpha f(y) \quad \forall \ x, y \in D$ and $\alpha \in [0, 1]$

Since $x,\ y \in \mathbb{R} =$ Domain of f

$\therefore\ f(x) = ax + b$ ... (7.29) and $f(y) = ay + b$ ... (7.30)

Let $\alpha \in [0, 1]$, then

$$\begin{aligned} f\left((1-\alpha)x + \alpha y\right) &= a\left((1-\alpha)x + \alpha y\right) + b && \left[\text{Using } (7.28)\right] \\ &= (1-\alpha)ax + \alpha ay + b && \ldots (7.31) \end{aligned}$$

$$\begin{aligned} \text{Also, } (1-\alpha)f(x) + \alpha f(y) &= (1-\alpha)\left(ax + b\right) + \alpha\left(ay + b\right) && \left[\text{Using } (7.28)\right] \\ &= (1-\alpha)ax + \alpha ay + (1-\alpha+\alpha)b \\ &= (1-\alpha)ax + \alpha ay + b && \ldots (7.32) \end{aligned}$$

From equations (7.31) and (7.32), we get

$f((1-\alpha)x+\alpha y) = (1-\alpha)f(x)+\alpha f(y) \quad \forall\ x, y \in D$ and $\alpha \in [0, 1]$

Hence, linear function $f(x) = ax+b,\ \ a, b, x \in \mathbb{R}$ is both convex and concave.

Example 6: Explain graphically whether the function $f:[0, \pi] \to [-1, 1]$ defined by $f(x) = \sin x,\ \ x \in [0, \pi]$ … (7.33)

is convex or concave or both or none of them.

Solution: We claim that given function is not a convex function. As per the instruction of statement we have to explain it graphically. So, let us first draw its graph which is shown in Fig. 7.5 (a). Note that graph of this function in the domain $[0, \pi]$ is downward so if you will draw any chord by joining any two points on the graph then graph of the function will lie above the chord between those points. Hence, after observing graph of the function we can say that sine function is concave in the domain $[0, \pi]$.

Example 7: Explain graphically whether the function $f:[\pi, 2\pi] \to [-1, 1]$ defined by $f(x) = \sin x,\ \ x \in [\pi, 2\pi]$ … (7.34)

is convex or concave or both or none of them.

Solution: We claim that given function is a convex function. As per the instruction of statement we have to explain it graphically. So, let us first draw its graph which is shown in Fig. 7.5 (b). Note that graph of this function in the domain $[\pi, 2\pi]$ is upward so if you will draw any chord by joining any two points on the graph then graph of the function will lie below the chord between those points. Hence, after observing graph of the function we can say that sine function is convex in the domain $[\pi, 2\pi]$.

Example 8: Explain graphically whether the function $f:[0, 2\pi] \to [-1, 1]$ defined by $f(x) = \sin x,\ \ x \in [0, 2\pi]$ … (7.35)

is convex or concave or both or none of them.

Solution: We claim that given function is neither convex nor concave. As per the instruction of the statement we have to explain it graphically. Its graph is shown in Fig. 7.5 (c). By Example 6 given function is concave in the domain $[0, \pi]$, while by Example 7 given function is convex in the domain $[\pi, 2\pi]$.

Hence, given function is neither convex nor concave in the entire domain $[0, 2\pi]$.

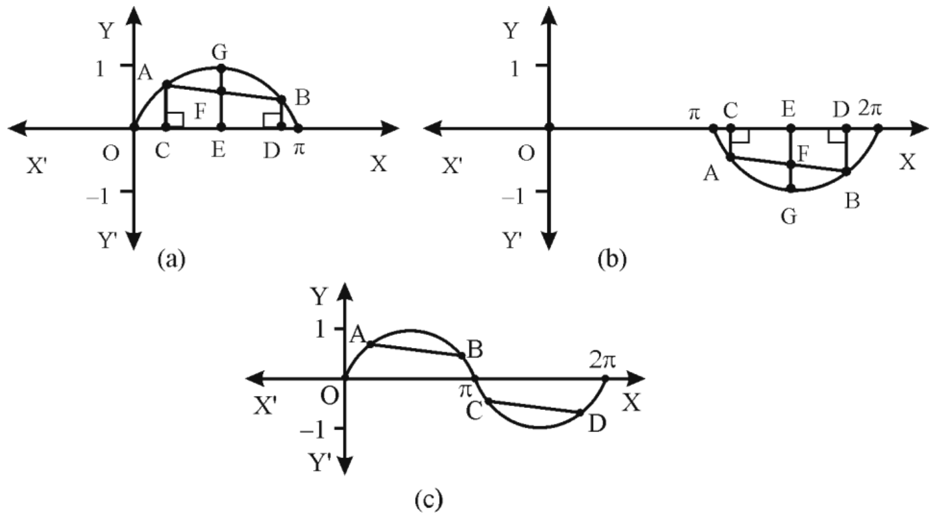

Fig. 7.5: Visualisation of graph of sine function in the domain (a) $[0,\pi]$ (b) $[\pi,2\pi]$ (c) $[0, 2\pi]$

7.6 EPIGRAPH AND PROPERTIES OF CONVEX FUNCTION

One thing that connect convex set with convex function is epigraph of a function via an important property refer Property 5. So, before discussing properties of a convex function let us first define epigraph of a function.

Epigraph: Let us first define epigraph of a function in layman language. The region above and on the graph of a function is known as epigraph of a function. In Fig. 7.6 (a) and (b) shaded regions in light blue colour (you may refer soft copy if this page is printed in black and white in eGyankosh) represent the epigraph of the functions $f(x) = 2 + x^2$ and $f(x) = x(x+2)(x-1)(x-2)$ respectively. Let us explain one more thing before giving definition of epigraph, which will help you to understand definition of epigraph of a function. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2 + 2,\ x \in \mathbb{R}$ the graph of this function is shown in Fig. 7.6 (a).

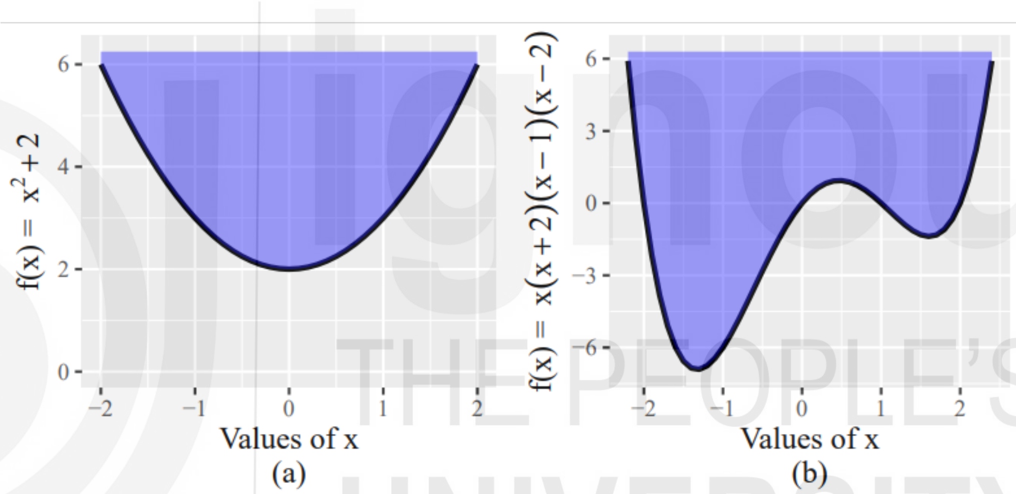

Fig. 7.6: Visualisation of epigraph of a function (a) epigraph of a convex function (b) epigraph of a function which is not convex

Note that: Domain of the function f is: $\mathbb{R}$ which is of 1-dimension but points on its graph are of 2-dimension. For example, $(0, 4) \in \mathbb{R}^2$ is a point in epigraph of this function and is of 2-dimension while domain of f was of 1-dimension. So, in general if $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ then points in epigraph of the function f will be of n + 1 dimension. That is dimension of the points in the epigraph of a function is one more than the dimension of the domain of the function. Further, if you want to present points on the graph of the function f as a set of points on it then it can be written as follows.

Set of points on the graph of the function f = $\{(x, f(x)) : x \in S\} \subseteq \mathbb{R}^{n+1}$ … (7.36)

Keep this important explanation in mind that **dimension of points on epigraph of a function is one more than the dimension of the domain of the function**. This explanation will help you to understand the definition of the epigraph of the function which is given as follows.

Let $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ then **epigraph** of the function f is defined as follows.

$$\text{efi}(f) = \{(x, y) \subseteq \mathbb{R}^{n+1} : x \in S,\ y \in \mathbb{R},\ y \geq f(x)\} \quad \ldots (7.37)$$

For example, for the function f whose epigraph is shown in Fig. 7.6 (a), if $0 \in \mathbb{R}$, then

$E = \{(0, y) \subseteq \mathbb{R}^2 : y \in \mathbb{R},\ y \geq f(0)\} = \{(0, y) \subseteq \mathbb{R}^2 : y \in \mathbb{R},\ y \geq 2\}$ which is a vertical line having lowest point at (0, 2).

Similarly, corresponding to each point x_0 of the domain of the function f we get a vertical line which has coordinate of the lowest point as $(x_0, f(x_0))$.

Now, we discuss some properties of convex function.

Properties of Convex Function

Properties of convex function mainly helps us in two ways.

(1) They are used to create more convex functions.

(2) They are used to prove convexity of functions.

Here, first we will list some properties of convex function then we will prove them.

Property 1: If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R},\ g : S \to \mathbb{R}$ are two convex functions then $f + g$ is also a convex function. … (7.38)

Property 2: If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}$ is a convex function then αf is also a convex function, where $\alpha \in \mathbb{R},\ \alpha \geq 0$. … (7.39)

Property 3: If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}$ is a convex function then every local minimum of the function f on S is global minimum. … (7.40)

Property 4: **Connection between Convex Set and Convex Function**: If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}$, then f is a convex function on S if and only if epigraph of the function f is a convex set. … (7.41)

Property 5: If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R},\ g : S \to \mathbb{R}$ are two convex functions then point wise maximum function, i.e., $\max_{x \in S}\{f(x), g(x)\}$ is also a convex function. … (7.42)

Proof of Property 1: Since functions f and g are two convex functions then by definition of convex function we have

$$f((1-\alpha)x + \alpha y) \leq (1-\alpha)f(x) + \alpha f(y) \quad \forall \ x, y \in D \text{ and } \alpha \in [0, 1] \qquad \ldots (7.43)$$

$$g((1-\alpha)x + \alpha y) \leq (1-\alpha)g(x) + \alpha g(y) \quad \forall \ x, y \in D \text{ and } \alpha \in [0, 1] \qquad \ldots (7.44)$$

Required to prove $f + g$ is convex.

$$(f+g)((1-\alpha)x + \alpha y) = f((1-\alpha)x + \alpha y) + g((1-\alpha)x + \alpha y) \quad \left[\begin{array}{l}\text{Using definition of}\\ \text{sum of two functions}\end{array}\right]$$

$$\leq (1-\alpha)f(x) + \alpha f(y) + (1-\alpha)g(x) + \alpha g(y) \quad [\text{Using (7.43) and (7.44)}]$$

$$\leq (1-\alpha)(f(x) + g(x)) + \alpha(f(y) + g(x))$$

$$= (1-\alpha)(f+g)(x) + \alpha(f+g)(y)$$

$$(f+g)((1-\alpha)x + \alpha y) \leq (1-\alpha)(f+g)(x) + \alpha(f+g)(y) \quad \forall \ x, y \in D \text{ and } \alpha \in [0, 1]$$

Hence, $f + g$ is also a convex function.

Proof of Property 2: We are given that f is convex function so we have

$$f\left((1-\alpha)x+\alpha y\right)\le(1-\alpha)f(x)+\alpha f(y) \quad \forall \ x, y\in D \ \text{and} \ \alpha\in[0, 1] \qquad \ldots (7.45)$$

Required to prove

$$(\alpha f)\left((1-\lambda)x+\lambda y\right)\le(1-\lambda)(\alpha f)(x)+\lambda(\alpha f)(y) \ \forall x, y\in D \ \text{and} \ \lambda\in[0, 1], \alpha\in\mathbb{R}, \alpha\ge 0$$

$$\begin{aligned}(\alpha f)\left((1-\lambda)x+\lambda y\right) &= \alpha f\left((1-\lambda)x+\lambda y\right)\\ &\le \alpha\left((1-\lambda)f(x)+\lambda f(y)\right) \quad [\text{Using } (7.19)]\\ &=(1-\lambda)\alpha f(x)+\lambda\alpha f(y)\\ &=(1-\lambda)(\alpha f)(x)+\lambda(\alpha f)(y)\end{aligned}$$

That is, we have proved that

$$(\alpha f)\left((1-\lambda)x+\lambda y\right)\le(1-\lambda)(\alpha f)(x)+\lambda(\alpha f)(y) \ \forall x, y\in D \ \text{and} \ \lambda\in[0, 1], \alpha\in\mathbb{R}, \alpha\ge 0$$

Hence, αf is a convex function.

Proof of Property 3: We are given that $f:S\to\mathbb{R}$ is a convex function so we have

$$f\left((1-\alpha)x+\alpha y\right)\le(1-\alpha)f(x)+\alpha f(y) \quad \forall \ x, y\in S \ \text{and} \ \alpha\in[0, 1] \qquad \ldots (7.46)$$

Required to prove every local minimum of the function f on S is global minimum.

Suppose x_0 be the local minimum of the function f having convex domain S. So, by definition of local minimum there will exists a neighbourhood $N_\delta(x_0)$ of x_0 for some $\delta>0$ such that

$$f(x)\ge f(x_0) \quad \forall \ x\in N_\delta(x_0)\cap S \qquad \ldots (7.47)$$

Required to prove that $x=x_0$ is global minimum. That is required to prove that

$$f(x)\ge f(x_0) \quad \forall \ x\in S \qquad \ldots (7.48)$$

Suppose, if possible $x=x_0$ is not global minimum of the function f. So, there will exist at least one point $x=x^*$ such that

$$f(x^*)<f(x_0) \qquad \ldots (7.49)$$

Consider the convex combination $(1-\alpha)x^*+\alpha x_0$ of points x^* and x_0. So, all the points of the line segment joining points x^* and x_0 are given by some value of $\alpha\in[0, 1]$. Also, $N_\delta(x_0)$ is neighbourhood of the point x_0 so there will exist some $\alpha=\alpha_0\in(0, 1)$ such that

$$(1-\alpha_0)x^*+\alpha_0 x_0\in N_\delta(x_0)\cap S$$

Let $x^{**}=(1-\alpha_0)x^*+\alpha_0 x_0$.

Now, $f(x^{**}) = f\left((1-\alpha_0)x^* + \alpha_0 x_0\right) \le (1-\alpha_0)f(x^*) + \alpha_0 f(x_0)$, $\left[\text{Using } (7.46)\right]$

$$< (1-\alpha_0)f(x_0) + \alpha_0 f(x_0) \quad \left[\text{Using } (7.49)\right]$$
$$= f(x_0)$$

i.e., $f(x^{**}) < f(x_0)$, where $x^{**} = (1-\alpha_0)x^* + \alpha_0 x_0 \in N_\delta(x_0) \cap S$

- a contradiction to equation (7.47). So, our supposition that $x = x_0$ is not global minimum of the function f is wrong. Hence, $x = x_0$ is global minimum of the function f. This completes the proof.

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

SAQ 4

Prove property 4 of convex function.

7.7 SUMMARY

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

- A point x is a **linear combination** of points $x_1, x_2, x_3, \ldots, x_k$ if and only if

  $x = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \alpha_k x_k, \;\; \alpha_i \in \mathbb{R}, \; i = 1, 2, 3, \ldots, k$

- A point x is an **affine combination** of points $x_1, x_2, x_3, \ldots, x_k$ if and only if

  $x = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \alpha_k x_k, \; \alpha_i \in \mathbb{R}, \; i = 1, 2, 3, \ldots, k$ and $\sum_{i=1}^{k} \alpha_i = 1\ldots$

- A point x is a **conic combination** of points $x_1, x_2, x_3, \ldots, x_k$ if and only if

  $x = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots + \alpha_k x_k, \;\; \alpha_i \in \mathbb{R}$, and $\alpha_i > 0, \; i = 1, 2, 3, \ldots, k$

- A point x is a **convex combination** of points $x_1, x_2, x_3, \ldots, x_k$ if and only if

  $x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k, \; \alpha_i \in \mathbb{R}, \; \alpha_i \ge 0, \; i = 1, 2, 3, \ldots, k$ and $\sum_{i=1}^{k} \alpha_i = 1$

- A set S is said to be an **affine set** if affine combination of any two members of it is in the set S.

- A set S is said to be a **convex set** if convex combination of any two members of it is in the set S.

- **Properties of Convex Set**:

  - Intersection of convex sets is also a convex set.
  - If C is a convex set and $\alpha \in \mathbb{R}, \; \alpha \ge 0$, then αC is also a convex set.
  - If C_1 and C_2 are two convex sets then $C_1 + C_2$ is also a convex set.

- Let $D \subseteq \mathbb{R}^n, \; n = 1, 2, 3, \ldots, k$ be the convex set. Then a function $f : D \to \mathbb{R}$ is said to be a **convex function** if

$$f\left((1-\alpha)x+\alpha y\right) \le (1-\alpha)f(x)+\alpha f(y) \quad \forall\ x, y \in D \text{ and } \alpha \in [0, 1]$$

and it is called **concave function** if

$$f\left((1-\alpha)x+\alpha y\right) \ge (1-\alpha)f(x)+\alpha f(y) \quad \forall\ x, y \in D \text{ and } \alpha \in [0, 1]$$

- Let $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ then **epigraph** of the function f is defined as follows $\text{efi}(f) = \left\{(x, y) \subseteq \mathbb{R}^{n+1} : x \in S, y \in \mathbb{R}, \ y \ge f(x)\right\}$

- **Properties of Convex Function**:
  - If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}, \ g : S \to \mathbb{R}$ are two convex functions then $f + g$ is also a convex function.
  - If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}$ is a convex function then αf is also a convex function, where $\alpha \in \mathbb{R}, \ \alpha \ge 0.$
  - If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}$ is a convex function then every local minimum of the function f on S is global minimum.
  - If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R},$ then f is a convex function on S if and only if epigraph of the function f is a convex set.
  - If $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}, \ g : S \to \mathbb{R}$ are two convex functions then point wise maximum function, i.e., $\max\limits_{x \in S}\left\{f(x), g(x)\right\}$ is also a convex function.

7.8 TERMINAL QUESTIONS

1. Does the set {4, 7} a convex set?
2. Prove that the circular disk in $\mathbb{R}^n$ is a convex set.
3. Prove that modulus function is a convex function.
4. Prove property 5 of convex function mentioned in Sec. 7.6.

7.9 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1. (a) Here $\alpha_1 = -2, \ \alpha_2 = 3.$ Since $\alpha_1 = -2 < 0,$ so it neither can be a conic nor a convex combination. But $\alpha_1 + \alpha_2 = -2 + 3 = 1,$ so it is an affine combination.

(b) Here $\alpha_1 = 3, \ \alpha_2 = 2.$ Since $\alpha_1 + \alpha_2 = 3 + 2 = 5 \ne 1,$ so it neither can be an affine nor a convex combination. But here $\alpha_1 = 3 > 0, \ \alpha_2 = 2 > 0,$ so it is a conic combination.

(c) Here $\alpha_1 = 0.3, \ \alpha_2 = 0.7.$ Since $\alpha_1 + \alpha_2 = 0.3 + 0.7 = 1,$ so it an affine combination. Also $\alpha_1 = 0.3 > 0, \ \alpha_2 = 0.7 > 0,$ so it is a conic combination. Further, $\alpha_1 = 0.3 > 0, \ \alpha_2 = 0.7 > 0$ as well as $\alpha_1 + \alpha_2 = 0.3 + 0.7 = 1,$ so it is also a convex combination. Hence, it qualifies requirements of all the three types of combinations affine, conic and convex.

2. (a) We know that the set of all natural numbers is $\mathbb{N} = \{1, 2, 3, \ldots\}$. We claim that the set $\mathbb{N}$ is not a convex set. We know that if $\mathbb{N}$ is a convex set, then $(1-\alpha)x + \alpha y \in \mathbb{N}, \forall\ x, y \in \mathbb{N}$ and for each $\alpha \in [0, 1]$.

Let $x = 1,\ y = 2 \in \mathbb{N},\ \alpha = 0.4,$ so $1-\alpha = 0.6.$ But
$\alpha 1 + (1-\alpha)2 = (0.4)1 + (0.6)2 = 0.4 + 1.2 = 1.6 \notin \mathbb{N}.$

Hence, the set $\mathbb{N}$ is not a convex set. The above argument is also shown in Fig. 7.7.

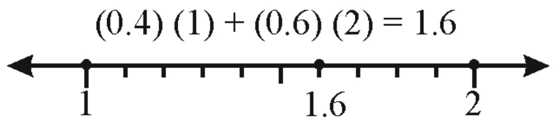

Fig. 7.7: Visualisation of the steps that the set of natural numbers is not a convex set

Remark 4: Similarly, we can show that (i) set of integers (ii) set of rational numbers, and (iii) set of irrational numbers are not convex sets.

2. (b) Union of two convex sets is not always convex. Here we will discuss one example where union of two sets will be convex and one example where union of two sets will not convex. Finally, we will discuss under what restriction union of two sets is always a convex set.

Example where union of two convex sets is a convex set:

Show that the two regions S_1, S_2 each bounded by four sides given as follows are convex sets. Further, show that their union is also a convex set.

$$S_1 = \left\{(x, y) \in \mathbb{R}^2 : x \le 2,\ x \ge 0,\ y \le 2,\ y \ge 0\right\}$$
$$S_2 = \left\{(x, y) \in \mathbb{R}^2 : x \le 3,\ x \ge 2,\ y \le 2,\ y \ge 0\right\}$$

Solution: Let $A(x_1, y_1)$ and $B(x_2, y_2)$ be two points in S_1, refer Fig. 7.8 (a). We shall prove that

$$(1-\alpha)(x_1, y_1) + \alpha(x_2, y_2) \in S_1 \quad \forall\ (x_1, y_1),\ (x_2, y_2) \in S_1 \text{ and } \forall\ \alpha \in [0, 1]$$

Now $(1-\alpha)(x_1, y_1) + \alpha(x_2, y_2) = \left((1-\alpha)x_1 + \alpha x_2,\ (1-\alpha)y_1 + \alpha y_2\right)$

Since $(x_1, y_1),\ (x_2, y_2) \in S_1 \Rightarrow 0 \le x_1, x_2 \le 2$ and $0 \le y_1, y_2 \le 2$

$$\Rightarrow 0 \le (1-\alpha)x_1 + \alpha x_2 \le 2 \text{ and } 0 \le (1-\alpha)y_1 + \alpha y_2 \le 2 \ \left[\because \alpha \in [0, 1]\right]$$

$$\therefore (1-\alpha)(x_1, y_1) + \alpha(x_2, y_2) \in S_1 \quad \forall\ (x_1, y_1),\ (x_2, y_2) \in S_1 \text{ and } \forall\ \alpha \in [0, 1]$$

Hence, S_1 is a convex set. Similarly, we can prove that S_2 is also a convex set. Now, we consider their union

$$S_1 \cup S_2 = \left\{(x, y) \in \mathbb{R}^2 : x \le 3,\ x \ge 0,\ y \le 2,\ y \ge 0\right\}$$

We claim that $S_1 \cup S_2$ is also a convex set.

Let $A(x_1, y_1)$ and $B(x_2, y_2)$ be two points in $S_1 \cup S_2$. We shall prove that
$$(1-\alpha)(x_1, y_1) + \alpha(x_2, y_2) \in S_1 \cup S_2$$
$$\forall (x_1, y_1), (x_2, y_2) \in S_1 \cup S_2 \text{ and } \forall \alpha \in [0, 1]$$

Now $(1-\alpha)(x_1, y_1)+\alpha(x_2, y_2)=((1-\alpha)x_1+\alpha x_2,\ (1-\alpha)y_1+\alpha y_2)$

Since $(x_1, y_1),\ (x_2, y_2)\in S_1\cup S_2 \Rightarrow 0\le x_1, x_2\le 3$ and $0\le y_1, y_2\le 2$

$\Rightarrow 0\le(1-\alpha)x_1+\alpha x_2\le 3$ and $0\le(1-\alpha)y_1+\alpha y_2\le 2$ $[\because \alpha\in[0, 1]]$

$\therefore\ (1-\alpha)(x_1, y_1)+\alpha(x_2, y_2)\in S_1\cup S_2$

$$\forall (x_1, y_1), (x_2, y_2)\in S_1\cup S_2 \text{ and } \forall \alpha\in[0, 1]$$

Hence, $S_1\cup S_2$ is a convex set.

Example where union of two convex sets is not a convex set:

Show that the two regions S_1, S_2 bounded by four sides given as follows are convex sets, but their union is not a convex set.

$S_1=\{(x, y)\in\mathbb{R}^2 : x\le 2,\ x\ge 0, y\le 4,\ y\ge 0\}$

$S_2=\{(x, y)\in\mathbb{R}^2 : x\le 6,\ x\ge 4, y\le 4,\ y\ge 0\}$

Solution: Following similar steps as in earlier part we can show that both the sets S_1 and S_2 are convex sets. We claim that their union is not a convex set. Consider the points $P(1, 2), Q(5, 2)\in S_1\cup S_2$, refer Fig. 7.8 (b). Now, the point $R(3, 2)$ lies on the line segment PQ but it does not lie in the region $S_1\cup S_2$. Implies this the region $S_1\cup S_2$ does not contain the whole line segment obtained by joining its two points P and Q. Hence, the region $S_1\cup S_2$ is not a convex set.

Finally, union of two convex sets is also a convex set if either $S_1\subseteq S_2$ or $S_2\subseteq S_1$.

Solution: We are given that both S_1 and S_2 are convex sets.

Case I: $S_1\subseteq S_2$

$\therefore\ S_1\cup S_2=S_2$ which is a convex set. Hence, $S_1\cup S_2$ is a convex set.

Case II: $S_2\subseteq S_1$

$\therefore\ S_1\cup S_2=S_1$ which is a convex set. Hence, $S_1\cup S_2$ is a convex set.

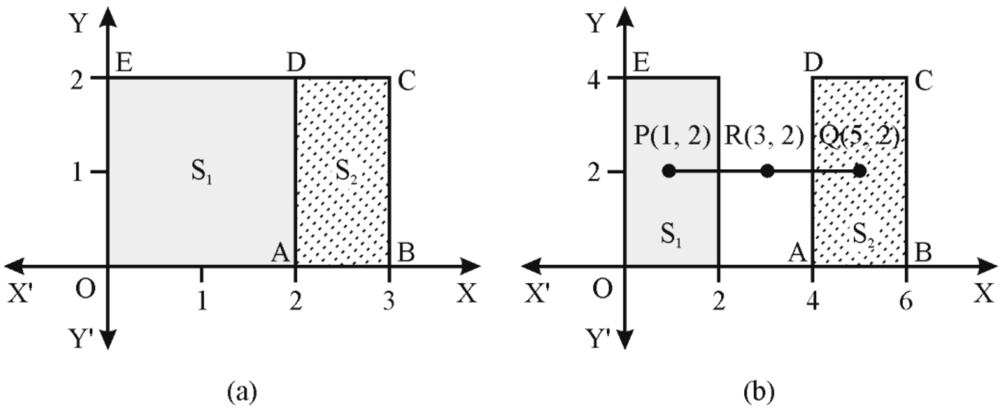

Fig. 7.8: Visualisation of (a) union of two convex sets is convex set (b) union of two convex sets is not a convex set

3. (a) We know that if C is a convex set then $\alpha C, \forall\ \alpha \geq 0$ is also a convex set. In our case $\alpha = 5 > 0.$ So, $5C_1$ is a convex set.

(b) We know that if C is a convex set then $\alpha C, \forall\ \alpha \geq 0$ is also a convex set. So, $7C_2$ and $9C_1$ are convex sets. We also know that if C_1 and C_2 are convex sets then their sum $C_1 + C_2$ is also a convex set. Hence, sum of $7C_2$ and $9C_1$, i.e., $7C_2 + 9C_1$ is also a convex set.

(c) We know that if C_1 and C_2 are convex sets then their intersection is also a convex set. Hence, $C_1 \cap C_2$ is a convex set.

(d) We know that in the case $C_1 \subseteq C_2$ or $C_2 \subseteq C_1$ union of two convex sets is convex. But in other cases, nothing can be said about convexity of union of two convex sets with surety. In other cases, union may be convex or may not be convex to see example you may refer solution of SAQ 2 (b).

4. **Proof of Property 4**: Required to prove f is convex function if and only if epigraph of the function f is a convex set. Let efi(f) denote the epigraph of the function f.

First suppose that f is convex function on its domain S. So, we have

$$f\left((1-\alpha)x_1 + \alpha x_2\right) \leq (1-\alpha)f(x_1) + \alpha f(x_2)\ \forall x_1, x_2 \in S \text{ and } \alpha \in [0, 1] \quad \ldots(7.50)$$

Let $(x_1, y_1), (x_2, y_2) \in \text{efi}(f)$, therefore by definition of epigraph, we have

$$y_1 \geq f(x_1),\ y_2 \geq f(x_2) \quad \ldots (7.51)$$

Consider the convex combination $(1-\alpha)(x_1, y_1) + \alpha(x_2, y_2)$ of points $(x_1, y_1), (x_2, y_2)$ and

$$(1-\alpha)(x_1, y_1) + \alpha(x_2, y_2) = (x, \alpha) \Rightarrow \left((1-\alpha)x_1 + \alpha x_2, (1-\alpha)y_1 + \alpha y_2\right) = (x, \alpha)$$
$$\Rightarrow (1-\alpha)x_1 + \alpha x_2 = x,\ (1-\alpha)y_1 + \alpha y_2 = \alpha \quad \ldots (7.52)$$

Now, $f(x) = f\left((1-\alpha)x_1 + \alpha x_2\right) \leq (1-\alpha)f(x_1) + \alpha f(x_2)$

$$\leq (1-\alpha)y_1 + \alpha y_2 \quad [\text{Using } (7.51)]$$
$$= \alpha \quad [\text{Using } (7.52)]$$

$$\Rightarrow f(x) \geq \alpha \quad \Rightarrow x \in \text{efi}(f)$$

$$\Rightarrow (1-\alpha)x_1 + \alpha x_2 \in \text{efi}(f)\ \forall\ (x_1, y_1), (x_2, y_2) \in \text{efi}(f),\ \alpha \in [0, 1]$$

Hence, efi(f) is a convex set.

Conversely: Suppose that efi(f) is a convex set. Let $(x_1, f(x_1)), (x_2, f(x_2)) \in \text{efi}(f)$. Since efi(f) is a convex set, so for all $\alpha \in [0, 1]$,

$$(1-\alpha)(x_1, f(x_1)) + \alpha(x_2, f(x_2)) \in \text{efi}(f)\ \forall\ (x_1, f(x_1)), (x_2, f(x_2)) \in \text{efi}(f)$$

$\Rightarrow \left((1-\alpha)x_1 + \alpha x_2,\ (1-\alpha)f(x_1) + \alpha f(x_2)\right) \in \text{efi}(f)$

$\Rightarrow f\left((1-\alpha)x_1 + \alpha x_2\right) \le (1-\alpha)f(x_1) + \alpha f(x_2)$ [By definition of epigraph]

i.e., $f\left((1-\alpha)x_1 + \alpha x_2\right) \le (1-\alpha)f(x_1) + \alpha f(x_2)\ \forall\, x_1, x_2 \in S$ and $\alpha \in [0, 1]$

Hence, f is a convex function.

Terminal Questions

1. Let C = {4, 7}. We claim that the set C is not a convex set. We know that if C is a convex set, then $(1-\alpha)x + \alpha y \in C,\ \forall\ x, y \in C$ and for each $\alpha \in [0, 1]$.

   Let $\alpha = 0.4$, then $1-\alpha = 0.6$. But $\alpha 4 + (1-\alpha)7 = (0.4)4 + (0.6)7 = 1.6 + 4.2 = 5.8 \notin C.$

   Hence, the set C is not a convex set.

2. Let us consider a circular disk with centre at origin and radius r, refer Fig. 7.9 to see shape of a circular disk in $\mathbb{R}^2$. Therefore, its equation is given by

   $x_1^2 + x_2^2 + \ldots + x_n^2 \le r^2$, where $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $r \in \mathbb{R}$, $r > 0$ … (7.53)

   Let $C = \left\{x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + x_2^2 + \ldots + x_n^2 \le r^2\right\}$

   Required to prove: The set C is convex. That is required to prove $(1-\alpha)x + \alpha y \in C,\ \forall\ x, y \in C$ and for each $\alpha \in [0, 1]$

   Let $x, y \in C$ then $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ and will satisfy (7.53)

   $\therefore\ x_1^2 + x_2^2 + \ldots + x_n^2 \le r^2$ … (7.54)

   and $y_1^2 + y_2^2 + \ldots + y_n^2 \le r^2$ … (7.55)

   Let $z = (1-\alpha)x + \alpha y = (1-\alpha)(x_1, x_2, \ldots, x_n) + \alpha(y_1, y_2, \ldots, y_n)$

   $= \left((1-\alpha)x_1 + \alpha y_1, (1-\alpha)x_2 + \alpha y_2, \ldots, (1-\alpha)x_n + \alpha y_n\right)$, $\alpha \in [0, 1]$

   Now, z is any point on the line segment joining points x and y including x and y. So, we will prove that point z lies on the circular disk given by (7.53) for all values of $\alpha \in [0, 1]$. Let us consider

$$\left((1-\alpha)x_1 + \alpha y_1\right)^2 + \left((1-\alpha)x_2 + \alpha y_2\right)^2 + \ldots + \left((1-\alpha)x_n + \alpha y_n\right)^2$$

$$= (1-\alpha)^2\left[x_1^2 + x_2^2 + \ldots + x_n^2\right] + \alpha^2\left[x_1^2 + x_2^2 + \ldots + x_n^2\right]$$

$$+ 2\alpha(1-\alpha)\left[x_1y_1 + x_2y_2 + \ldots + x_ny_n\right]$$

$$\le (1-\alpha)^2\left[r^2\right] + \alpha^2\left[r^2\right] \qquad \left[\text{Using } (7.54) \text{ and } (7.55)\right]$$

$$+ 2\alpha(1-\alpha)\left[x_1y_1 + x_2y_2 + \ldots + x_ny_n\right]$$

$$\le \left[(1-\alpha)r\right]^2 + \left[\alpha r\right]^2 + 2\alpha(1-\alpha)\left|x_1y_1 + x_2y_2 + \ldots + x_ny_n\right| \quad \left[\because x.y \le |x.y|\right]$$

$$\leq \left[(1-\alpha)r\right]^2 + \left[\alpha r\right]^2 + 2\alpha(1-\alpha)\|x\|\|y\| \quad \left[\text{Using } (6.45e)\right]$$

$$\leq \left[(1-\alpha)r\right]^2 + \left[\alpha r\right]^2 + 2\alpha(1-\alpha)(r)(r) \quad \begin{bmatrix} \because \|x\| = \|x-0\| \leq r \text{ and} \\ \|y\| = \|y-0\| \leq r \end{bmatrix}$$

$$= \left[(1-\alpha)^2 + \alpha^2 + 2\alpha(1-\alpha)\right]r^2 = (1-\alpha+\alpha)^2 r^2 = r^2$$

i.e., $\left((1-\alpha)x_1 + \alpha y_1\right)^2 + \left((1-\alpha)x_2 + \alpha y_2\right)^2 + \ldots + \left((1-\alpha)x_n + \alpha y_n\right)^2 \leq r^2$

$\Rightarrow z = (1-\alpha)x + \alpha y \in C, \ \forall \ x, y \in C$ and for each $\alpha \in [0, 1]$

Hence, circular disk in $\mathbb{R}^n$ is a convex set.

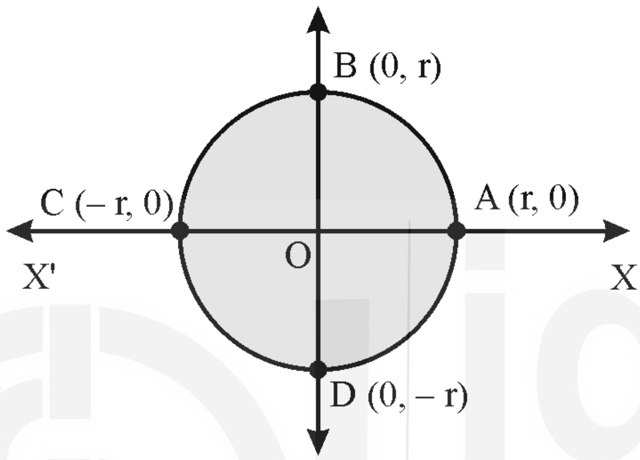

Fig. 7.9: Visualisation of a circular disk in two dimensions

3. We know that modulus function is defined as follows. $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = |x|, \ x \in \mathbb{R}$ … (7.56)

Let $x, \ y \in \mathbb{R}$ and $\alpha \in [0, 1]$, then required to prove

$f\left((1-\alpha)x + \alpha y\right) \leq (1-\alpha)f(x) + \alpha f(y) \quad \forall \ x, y \in \mathbb{R}$ and $\alpha \in [0, 1]$

Now,

$$f\left((1-\alpha)x + \alpha y\right) = |(1-\alpha)x + \alpha y| \quad \left[\text{Using } (7.56)\right]$$

$$\leq |(1-\alpha)x| + |\alpha y| \quad \left[\because |a+b| \leq |a| + |b|\right]$$

$$= (1-\alpha)|x| + \alpha|y| \left[\because 1-\alpha \geq 0, \alpha \geq 0 \text{ so } |1-\alpha| = 1-\alpha \text{ and } |\alpha| = \alpha\right]$$

$$= (1-\alpha)f(x) + \alpha f(y) \quad \left[\text{Again using } (7.56)\right]$$

i.e., $f\left((1-\alpha)x + \alpha y\right) \leq (1-\alpha)f(x) + \alpha f(y) \quad \forall \ x, y \in \mathbb{R}$ and $\alpha \in [0, 1]$

Hence, modulus function is a convex function. … (7.57)

4. **Proof of Property 5**: To prove this property we will use Property 4 already proved in SAQ 4. So, we know that

Function $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex on S iff efi(f) is a convex set. … (7.58)

Since functions f and g are convex functions, so efi(f) and efi(g) are convex sets. [Using (7.58) for functions f and g]

$\Rightarrow \text{efi}(f) \cap \text{efi}(g)$ is a convex set [Using (7.18)] … (7.59)

But using (7.37), we have

$$\Rightarrow \text{efi}(f) \cap \text{efi}(g) = \{(x, y) \subseteq \mathbb{R}^{n+1} : x \in S, y \in \mathbb{R}, \; y \geq f(x), \; y \geq g(x)\}$$

$$= \left\{(x, y) \subseteq \mathbb{R}^{n+1} : x \in S, y \in \mathbb{R}, \; y \geq \max_{x \in S}\{f(x), g(x)\}\right\} \quad \text{... (7.60)}$$

In view of (7.59) and (7.60), we have

$\left\{(x, y) \subseteq \mathbb{R}^{n+1} : x \in S, y \in \mathbb{R}, \; y \geq \max_{x \in S}\{f(x), g(x)\}\right\}$ is a convex set.

$\Rightarrow \max_{x \in S}\{f(x), g(x)\}$ is a convex function. [Using (7.58)]