

Numerical Optimization
Prof. Shirish K. Shevade
Department of Computer Science and Automation
Indian Institute of Science, Bangalore

Lecture - 6
Convex Sets

Hello and welcome to this series of lectures on numerical optimization. In the next few lectures, we will study about convex sets and convex functions. Remember that this convex sets and convex functions play a very important role in the optimization theory, and it is very important to understand this theory before we actually move on to the other problems. So, let us start looking at convex sets first and then the convex functions. Convex functions are very nice functions, which do not have problems of local minima as we will see later. Now, these convex functions are typically defined on convex sets. So, we first study convex sets before moving on to convex functions.

(Refer Slide Time: 01:12)

Line and line segment

Let $x_1, x_2 \in \mathbb{R}^n, x_1 \neq x_2$.

The diagram shows a line passing through points x_1 and x_2 . Points on the line are labeled with lambda values: $\lambda = 1.5$, $\lambda = 1$ (at x_1), $\lambda = 0$ (at x_2), and $\lambda = -0.2$.

Line passing through x_1 and x_2 :

$$\{y \in \mathbb{R}^n : y = \lambda x_1 + (1 - \lambda)x_2, \lambda \in \mathbb{R}\}$$

Line Segment, $LS[x_1, x_2]$:

$$\{y \in \mathbb{R}^n : y = \lambda x_1 + (1 - \lambda)x_2, \lambda \in [0, 1]\}$$

NPTEL

Now, before we get on to the actual theory of convex sets, so let me define some notations. So, on the display, you will see a line drawn passing through the points x_1 and x_2 , where x_1 is not equal to x_2 . Now, associated with each of the points on the line is a parameter lambda. So, let us assume that lambda is one at the point x_1 lambda is 0 at the point x_2 and beyond x_2 the value of lambda decreases. So, here you will see that for this point lambda is equal to minus 0.2 and it decreases further as you move away

from x_2 in this direction. Now, if you move away from x_1 in this direction, the value of λ increases.

So, any point on the line will have some real λ associated with that point. So, the line passing through x_1 and x_2 can be represented by a set y , so that y can be written as $\lambda x_1 + (1 - \lambda)x_2$, where λ belongs to the set of real numbers. So, if we plug in λ equal to 1, what we get is y is equal to x_1 as this quantity vanishes. If you plug in λ equal to 0, we will get this quantity vanishes and we get y equal to x_2 . So, a line is parameterized by a parameter called λ which is a real number and if we are given two points through which the line passes, one can appropriately set λ s at those points and find the equation of a line which passes through these points. Now, if you restrict our self only to the line segment x_1 and x_2 , then we will denote it less x_1, x_2 and that can be written as y is equal to $\lambda x_1 + (1 - \lambda)x_2$.

(Refer Slide Time: 03:34)

Affine sets


Definition
A set $X \subseteq \mathbb{R}^n$ is **affine** if for any $x_1, x_2 \in X$, and $\lambda \in \mathbb{R}$,

$$\lambda x_1 + (1 - \lambda)x_2 \in X.$$

If $X \subseteq \mathbb{R}^n$ is an affine set, $x_1, x_2, \dots, x_k \in X$ and $\sum_i \lambda_i = 1$, then the point $\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_k x_k \in X$.

Examples

- Solution set of linear equations: $\{x \in \mathbb{R}^n : Ax = b\}$ where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$.
- A subspace or a translated subspace

 MPTEL

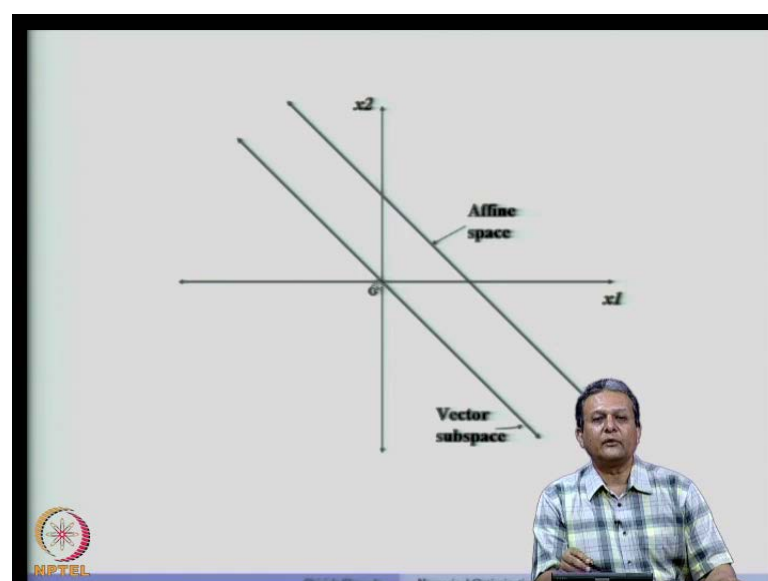
Now, remember that λ takes the value in the close interval 0 to 1, so unlike the line where λ can take any value from the set of real numbers here in when we talk about the line segment λ takes the value in the range 0 to 1. So, this is the n segment x_1, x_2 that we are talking about. Now, let us look at what are called affine sets, a set x which is subset of \mathbb{R}^n is said to be affine. If you take any two points in the set x_1

and x_2 and take any λ which is a real number, then $\lambda x_1 + (1 - \lambda)x_2$ should belong to X .

So, what this essentially means is that if you take any two points from the set X and take a line passing through those two points, then that line lies in the set X . So, such a set is called an affine set. Now, this definition can be generalized to this set of k points. So, earlier we took only two points, x_1 and x_2 . Suppose, we take k points x_1, x_2, \dots, x_k which are part of the set X , and if we choose λ_i such that $\sum_{i=1}^k \lambda_i = 1$, then if X is an affine set, then $\sum_{i=1}^k \lambda_i x_i$ belongs to the set X .

So, this is just an extension of this. So, remember that when we have an affine set, we take any points, any set of points from that set when $\sum \lambda_i = 1$ if we ensure that $\sum \lambda_i = 1$ for the affine set. Now, here are some examples of affine set. So, if we consider a solution set of linear equations, where A is $m \times n$ matrix and b is the m dimensional vector, then the set of all x such that $Ax = b$ form the solution set of this linear equations. Now, one can easily show that this set is an affine set. For example, if we take two solutions, x_1 and x_2 of this system of linear equations, then $Ax_1 = b$ and $Ax_2 = b$.

(Refer Slide Time: 06:24)



So, if we take λ , so λ into $x_1 + 1 - \lambda$ into x_2 , one can show that also satisfies. So, if we take axis that x equal to $\lambda x_1 + 1 - \lambda x_2$, that also satisfies the system $a x$ equal to b . So, any solution set of linear equations of this type is an affine set. Another example of an affine set is a subspace or a translated subspace. For example, this is the vector subspace because you will see that it is in to dimensional space. This is the vector subspace because it passes through the origin. Now, if I translate it that becomes an affine space. Now, both are affine sets because if you take any points, any two points in this set, then the line passing through those points belongs to the same set. So, these two are, both are affine sets.

(Refer Slide Time: 07:08)

Result
If X is an affine set and $x_0 \in X$, then $\{x - x_0 : x \in X\}$ forms a subspace.

Proof.
Let $Y = \{x - x_0 : x \in X\}$. To show that Y is a subspace, we need to show, $\alpha y_1 + \beta y_2 \in Y$ for any $y_1, y_2 \in Y$ and $\alpha, \beta \in \mathbb{R}$.
Let $y_1, y_2 \in Y$ and $\alpha, \beta \in \mathbb{R}$.
Therefore, $y_1 + x_0 \in X$ and $y_2 + x_0 \in X$.
 $\alpha y_1 + \beta y_2 = \alpha(y_1 + x_0) + \beta(y_2 + x_0) - (\alpha + \beta)x_0$
 $\alpha y_1 + \beta y_2 + x_0 = \alpha(y_1 + x_0) + \beta(y_2 + x_0) + (1 - \alpha - \beta)x_0$
Since X is an affine set, $\alpha y_1 + \beta y_2 + x_0 \in X$.
Thus, $\alpha y_1 + \beta y_2 \in Y \Rightarrow Y$ is a subspace. \square

In fact, one can show that if we have an affine space and if you translate it to the origin, then it becomes a vector space. So, here is a result which says that if (\cdot) affine at x and let us take a point x_0 from that set x , then the set of all points x minus x_0 , where x belongs to the set x , it forms a subspace and it is very easy to do that. So, let us look at the proof of this result. So, let us consider the set y . So, we are given x naught which will belong to the set x . So, let us consider the set y , such that which is the collection of x minus x naught where x belongs to x .

Now, we have to show that y is a subspace. That means what we have to show is that by the definition of subspace if you take any two vector, say y_1 and y_2 in the set y , then linear combination of those two vectors should also belong to the subspace y . So, that

means that if we choose α and β belong to \mathbb{R} and αy_1 and βy_2 should belong to Y , so this is what we want to show that Y is a subspace. Now, so let us consider y_1 and y_2 which belong to the set Y and let us consider two scalars which are real numbers, α and β . Now, since y_1 belongs to Y and x_0 belongs to X , so $y_1 + x_0$ clearly belongs to X and similarly, $y_2 + x_0$ belongs to X .

Now, let us see what happens to $\alpha y_1 + \beta y_2$. So, $\alpha y_1 + \beta y_2$ can be written as $\alpha(y_1 + x_0) + \beta(y_2 + x_0) - \alpha x_0 - \beta x_0$. Now, you will see that $y_1 + x_0$ belongs to X , $y_2 + x_0$ also belongs to X and x_0 obviously belongs to X that is given to us. So, all these three vectors belong to the set X . Now, this affine combination of these vectors. Now, let us sum the co-efficients which are here. So, $\alpha + \beta - \alpha - \beta$. So, that is 0, so that this is not an affine combination of the vectors in X . So, what we do is that to make it an affine combination, let us add x_0 .

So, once you add x_0 , so what do we get? So, these quantities remain the same and the one x_0 gets added to the third quantity. Now, you will see that $y_1 + x_0 + y_2 + x_0 + x_0$ and x_0 belong to X and this is affine combination of the three vectors in the set X because $\alpha + \beta + 1 - \alpha - \beta$ is 1. So, clearly $\alpha y_1 + \beta y_2 + x_0$ belongs to the set X because X is an affine set and we are considering an affine combination of three vectors in that set.

So, clearly this left hand side quantity belongs to the set X , which means that $\alpha y_1 + \beta y_2$ should belong to the set Y . So, since X is an affine set, the left side quantity belongs to X and therefore, $\alpha y_1 + \beta y_2$ belongs to Y which implies that Y is a subspace. So, if we have any affine set and if you translate it to the origin, then it becomes, so clearly you will see that since x_0 belongs to X origin, also belongs to the set $X - x_0$. So, clearly an origin is the part of the set Y and by this proof, it is a subspace.


(Refer Slide Time: 11:09)

Definition
Let $X = \{x_1, x_2, \dots, x_k\}$. A point x is said to be an **affine combination** of points in X if

$$x = \sum_{i=1}^k \lambda_i x_i, \quad \sum_{i=1}^k \lambda_i = 1.$$

Definition
Let $X \subseteq \mathbb{R}^n$. The set of all affine combinations of points in X is called the **affine hull** of X .

$$\text{aff}(X) = \left\{ \sum_{i=1}^k \lambda_i x_i : x_1, \dots, x_k \in X, \sum_{i=1}^k \lambda_i = 1 \right\}$$

 NPTEL

Now, let us look at some more definitions. Suppose, we are given a set of points x_1 to x_k . Now, a point x is said to be an affine combination of points in x . If you can write x as a linear combination of x_i 's and the linear combination is such that $\sum \lambda_i = 1$. So, the sum of the coefficients in the linear combination should be 1. Then we call x to be an affine combination of the points x_i 's, i going from 1 to k . Now, one can talk about this affine hull. So, affine hull is the set of all affine combinations of the points in x and it can be written as a set of all affine combinations of x_i 's in x .

(Refer Slide Time: 12:27)


Convex Sets


Definition
A set $C \subseteq \mathbb{R}^n$ is **convex** if for any $x_1, x_2 \in C$ and any scalar λ with $0 \leq \lambda \leq 1$, we have

$$\lambda x_1 + (1 - \lambda)x_2 \in C.$$

Examples

- Some simple convex sets



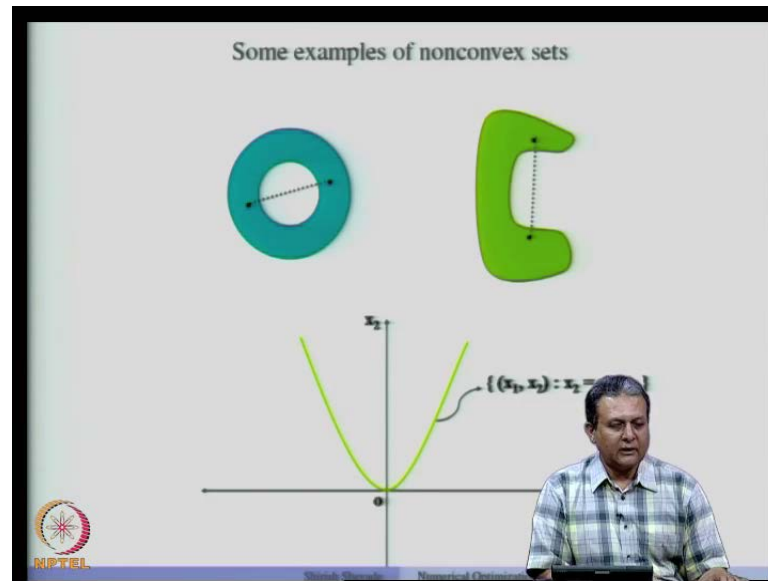
 NPTEL

So, remember that σ is going to be 1. So, you take any k points in the set x , take an affine combination of them and collect all such affine combinations. We will get affine hull of the set x . Now, we will move on to the definition of convex sets and as I said that this convex sets and the convex functions are very important topics in the theory of optimization, and some of these ideas are used to derive optimality conditions for general linear and non-linear programic problems. So, let us see the definition of a convex set. Now, a set which are subset of \mathbb{R}^n is said to be convex. If you take any two points x_1 and x_2 has belonging to that set and a scalar λ , such that $0 \leq \lambda \leq 1$ or equal to 1. So, λ lies in the close interval 0 to 1.

Then, we have $\lambda x_1 + (1 - \lambda) x_2$ which belongs to c . Now, what does this definition mean? So, this definition means that if we take any two points, x_1 and x_2 from the set c and take a line segment joining x_1 and x_2 , so $\lambda x_1 + (1 - \lambda) x_2$, where λ is in the close interval 0 to 1 denotes a line segment joining x_1 and x_2 . So, if you take any two points, x_1, x_2 belong to c , take a line segment joining those two points that the n segments should entirely lie within the set c . So, let us look at some examples.

So, here are a couple of examples of convex sets. Now, you can think of convex set has the following that along the boundary. Suppose, if you put a wall along the boundary of a given set, if you put a wall and then consider any two points in that set, then any two points, every point should see any other point of the set directly through a straight line path. So, if you take any two points here and if we take a line segment joining those two points that lies entirely within that set. Similarly, one can say that this is a convex set.

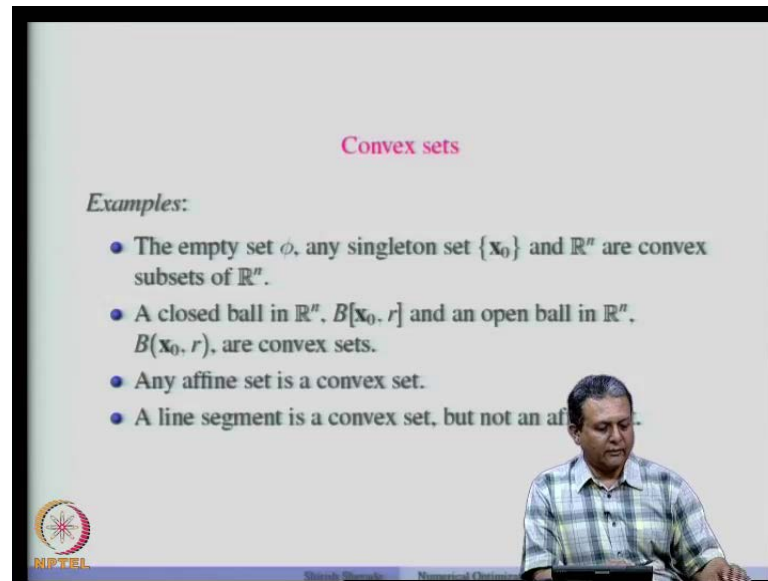
(Refer Slide Time: 14:46)



Now, let us see some examples of non-convex sets. So, on the left side, you will see that if I take this point and this point, so the line segment joining these two points does not entirely lie within the set. Although, some part of it lies, but not the entire line segments should lie within the set. So, if we again go by the earlier geometric interpretations, suppose if you put a wall around the boundary, then this point will not be able to see this point because of the wall.

Similarly, this is not a convex set because the line segment joining these two points is not entirely within the set. So, these are some examples of non-convex sets. So, here is one more example. So, let us consider the set of points in two-dimensional space, $x_1 \times x_2$, where x_2 is nothing but x_1 squared. So, this is the set of points which is shown by the green curve here. So, this set is not convex because if I take a point here and take a point here, the line segment joining them does not part of the set. So, these are some examples of non-convex sets.

(Refer Slide Time: 16:17)



The image shows a video frame of a lecture. The main content is a slide titled "Convex sets" in pink. Under the heading "Examples:", there is a bulleted list of four items. In the bottom right corner of the video frame, a man in a plaid shirt is visible, presumably the lecturer. In the bottom left corner of the slide, there is a logo for NPTEL (National Programme on Technology Enhanced Learning).

Convex sets

Examples:

- The empty set ϕ , any singleton set $\{x_0\}$ and \mathbb{R}^n are convex subsets of \mathbb{R}^n .
- A closed ball in \mathbb{R}^n , $B[x_0, r]$ and an open ball in \mathbb{R}^n , $B(x_0, r)$, are convex sets.
- Any affine set is a convex set.
- A line segment is a convex set, but not an affine set.

NPTEL

So, for convexity what we need is, for convexity of a set, what we need is that the line segment joining any two points of that set should lie entirely within that set. Now, let us look at some more examples of convex sets. Now, by convention, the empty set is a convex set. Then any singleton is also convex set because since we have only one point in the set, the set is obviously a convex set because there is no notion of a line segment in that set. Then the set of \mathbb{R}^n , the n dimensional set of a real numbers, these are all convex sets.

Now, if you take a close ball in \mathbb{R}^n . The ball with the center x_0 and radius r where r is greater than 0 or if you take an open ball in \mathbb{R}^n which are the center x_0 and the radius r which is again greater than 0, then they are all convex subsets of \mathbb{R}^n . Now, any affine set is a convex set because an affine set by definition is a set where if you take any two points, the line passing through the two points always lies in that set. So, if a line passes through the set, the line segment joining any two points also should lie in that set. So, affine set is also convex set. Then line segment if we just consider a line segment, line segment is a convex set, but the line segment is not affine because the entire line does not lie in that set. So, a line segment is a convex, but not affine.

(Refer Slide Time: 18:03)

• Possible to construct new convex sets using given convex sets

Definition
Let $P \subset \mathbb{R}^n, Q \subset \mathbb{R}^n, \alpha \in \mathbb{R}$.

• The *scalar multiple* αP of the set P is defined as

$$\alpha P = \{ \mathbf{x} : \mathbf{x} = \alpha \mathbf{p}, \mathbf{p} \in P \}$$

• The *sum* of two sets P and Q is the set,

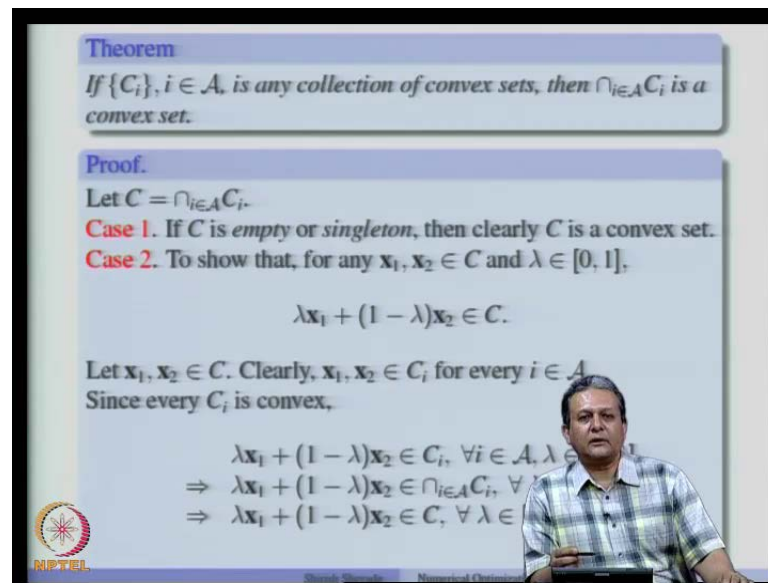
$$P + Q = \{ \mathbf{x} : \mathbf{x} = \mathbf{p} + \mathbf{q}, \mathbf{p} \in P, \mathbf{q} \in Q \}$$

The slide also features the NPTEL logo in the bottom left corner and a small inset video of the lecturer in the bottom right corner.

Now, it is possible to construct some new convex sets using given convex sets. So, then we will have some ways to derive new sets, new convex sets from the existing convex sets. Now, before we look in to those details, let us look at some definitions. So, let p and q be subsets of \mathbb{R}^n and let α be a real number. Now, the scalar multiple αp of the set is defined as the set of all points x , such that x is nothing but x equal to αp , where p belongs to p . So, we take given α which is the real number. We take every p from the set p and multiplied by α and get x .

Now, collect all such x 's from all possible p 's that will form the set αp . Now, similarly, one can also define the sum of two sets as p plus q is nothing but you take a vector p or point p from the set p and vector q from the set q and then sum them off. So, this is the vector sum of p and q . So, collect all such x 's from all possible pairs of p and q where p belongs to the set, p and q belongs to the set q and that will form the sum of two sets, p plus q .

(Refer Slide Time: 19:43)



Theorem
If $\{C_i\}, i \in \mathcal{A}$, is any collection of convex sets, then $\bigcap_{i \in \mathcal{A}} C_i$ is a convex set.

Proof.
Let $C = \bigcap_{i \in \mathcal{A}} C_i$.

Case 1. If C is empty or singleton, then clearly C is a convex set.

Case 2. To show that, for any $x_1, x_2 \in C$ and $\lambda \in [0, 1]$,

$$\lambda x_1 + (1 - \lambda)x_2 \in C.$$

Let $x_1, x_2 \in C$. Clearly, $x_1, x_2 \in C_i$ for every $i \in \mathcal{A}$.
Since every C_i is convex,

$$\lambda x_1 + (1 - \lambda)x_2 \in C_i, \forall i \in \mathcal{A}, \lambda \in [0, 1]$$
$$\Rightarrow \lambda x_1 + (1 - \lambda)x_2 \in \bigcap_{i \in \mathcal{A}} C_i, \forall \lambda \in [0, 1]$$
$$\Rightarrow \lambda x_1 + (1 - \lambda)x_2 \in C, \forall \lambda \in [0, 1]$$

MPTEL

Now, with this definition, we can generate new convex sets from the given sets of convex sets. So, let us look at some results to generate new convex sets. Now, if we have collection of convex sets denoted by C_i where i belongs to the set \mathcal{A} , then the theorem says that the intersection of this convex sets is a convex set. So, let us try to prove this figure.

So, let us consider the set C to be the intersection of all C_i 's where i is from the set, \mathcal{A} is the index set, i is the index which belongs to the set \mathcal{A} . Now, let us consider the first case where C is empty or singleton. Now, we have seen that by convention, it is empty. It is a convex set or even if C is singleton, it is a convex set. So, we have to worry about only those cases where C is non empty and non single term. So, let us look at that case. Now, whenever we want to show that any set is a convex set, what we need to show is that for any two points in that set and λ in the close interval 0 to 1 $\lambda x_1 + (1 - \lambda)x_2$ should belong to C . So, this is what we need to show.

So, let us consider any two points in the set C . Now, since they are in the intersection set C , they are in the individual sets as well for every i belong to \mathcal{A} . Now, remember that each of the C_i 's is a convex set, so by using the convexity of C_i , what we can do is that we can say that $\lambda x_1 + (1 - \lambda)x_2$ belongs to C_i or every i belong to \mathcal{A} , where λ is in the close interval 0 to 1. This is because of the fact that every C_i is a convex set. Now, that means that you take a line segment joining x_1 and x_2 that entirely

lies between within the set c_i for every i which means that if you take the line segment joining x_1 and x_2 , it should also lie in the intersection of this c_i 's because it belongs to every c_i , which means that $\lambda x_1 + (1 - \lambda)x_2$ belongs to the intersection of all c_i 's for all λ in the range 0 to 1.

(Refer Slide Time: 23:07)

Theorem
If C_1 and C_2 are convex sets, then $C_1 + C_2$ is a convex set.

Proof.
 Let $x_1, y_1 \in C_1$ and $x_2, y_2 \in C_2$.
 So, $x_1 + x_2, y_1 + y_2 \in C_1 + C_2$. We need to show,
 $\lambda(x_1 + x_2) + (1 - \lambda)(y_1 + y_2) \in C_1 + C_2 \forall \lambda \in [0, 1]$.
 Since C_1 and C_2 are convex,
 $z_1 = \lambda x_1 + (1 - \lambda)y_1 \in C_1 \forall \lambda \in [0, 1]$, and
 $z_2 = \lambda x_2 + (1 - \lambda)y_2 \in C_2 \forall \lambda \in [0, 1]$.
 Thus, $z_1 + z_2 \in C_1 + C_2 \dots \dots (1)$
 Now, $x_1 + x_2 \in C_1 + C_2$ and $y_1 + y_2 \in C_1 + C_2$.
 Therefore, from (1),
 $z_1 + z_2 = \lambda(x_1 + x_2) + (1 - \lambda)(y_1 + y_2) \in C_1 + C_2$.
 Thus, $C_1 + C_2$ is a convex set.

This intersection of c_i belong to a is nothing but the set c . So, $\lambda x_1 + (1 - \lambda)x_2$ belongs to the set c for all λ in the range 0 to 1. So, this result says that if we take a collection of convex sets, any collection of convex sets, then their intersection is the convex set. So, by using given convex sets, we can try to take their intersection and form new convex sets. So, this is the first result that we saw. Then the second result is about the sum of two sets. So, suppose if you are given c_1 and c_2 as convex sets, then $c_1 + c_2$ is a convex set.

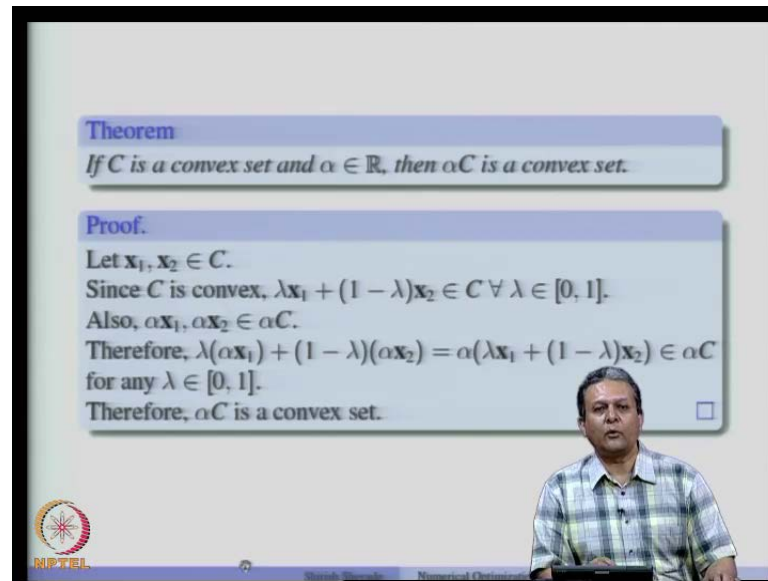
So, as I said earlier that to show that any set is the convex set, we need to take two points in that set, any two points and show that the line segment joining any two points lies entirely within the set. So, let us consider x_1, y_1 belongs to c_1 and x_2, y_2 belong to c_2 . So, these are the vectors in c_1 and c_2 respectively. So, we can say that $x_1 + x_2$ belongs to $c_1 + c_2$ and $y_1 + y_2$ belongs to $c_1 + c_2$. This is by the definition of $c_1 + c_2$. Now, we have got two points in the set $c_1 + c_2$. What we need to show is that $\lambda(x_1 + x_2) + (1 - \lambda)(y_1 + y_2)$ also belongs to $c_1 + c_2$ for all λ in the range 0 to 1. So, this is what we need to show.

Now, remember that what is given to us is that c_1 and c_2 are convex. So, if I take two vectors, x_1 and y_1 in c_1 , then the line segment joining x_1 and y_1 lies in the set c_1 . So, if I write z_1 to be $\lambda x_1 + (1 - \lambda) y_1$, so z_1 clearly belongs to c_1 for all λ in the range 0 to 1. Similarly, if I take two vectors, x_2 and y_2 in the set c_2 and take a line segment joining x_2 and y_2 , I can write z_2 to be $\lambda x_2 + (1 - \lambda) y_2$ which belongs to c_2 for all λ in the range 0 to 1.

So, now we have vector z_1 which belong to c_1 and a vector z_2 which belong to c_2 . So, what we can do is we can sum them up. So, $z_1 + z_2$ will belong to $c_1 + c_2$. So, this is the important point. The $z_1 + z_2$ belongs to $c_1 + c_2$. What is $z_1 + z_2$? Now, if you sum the right side, what do we get? So, remember that $x_1 + x_2$ belongs to $c_1 + c_2$ and $y_1 + y_2$ belongs to $c_1 + c_2$. So, if you sum these, the right side, what we get is $\lambda(x_1 + x_2) + (1 - \lambda)(y_1 + y_2)$ and that is nothing but $z_1 + z_2$ and since, $z_1 + z_2$ belong to $c_1 + c_2$ from 1, what we get is that the vector on the right side also belongs to $c_1 + c_2$ for all λ in the range 0 to 1.

So, we had two points, $x_1 + x_2$ in the set $c_1 + c_2$ and $y_1 + y_2$ in the set $c_1 + c_2$ and a line segment joining those two points which is written as $\lambda(x_1 + x_2) + (1 - \lambda)(y_1 + y_2)$ also belongs to the set $c_1 + c_2$ for all λ in the range 0 to 1, which means that $c_1 + c_2$ is a convex set. So, if we are given two convex sets, we can generate new convex sets. For example, if you take x axis in the two-dimensional space, if we take x axis as one convex set and y axis as one convex set, then the set $c_1 + c_2$ is entire \mathbb{R}^2 and that turns out to be a convex set. So, new convex sets could be generated by taking the vector sum of the convex sets easily.

(Refer Slide Time: 27:06)



The slide contains the following text:

Theorem
If C is a convex set and $\alpha \in \mathbb{R}$, then αC is a convex set.

Proof.
Let $x_1, x_2 \in C$.
Since C is convex, $\lambda x_1 + (1 - \lambda)x_2 \in C \forall \lambda \in [0, 1]$.
Also, $\alpha x_1, \alpha x_2 \in \alpha C$.
Therefore, $\lambda(\alpha x_1) + (1 - \lambda)(\alpha x_2) = \alpha(\lambda x_1 + (1 - \lambda)x_2) \in \alpha C$
for any $\lambda \in [0, 1]$.
Therefore, αC is a convex set. \square

The lecturer is a man with short grey hair, wearing a light-colored plaid shirt, standing in front of the slide.

Now, here is the third result which says that if c is a convex set and α is a real number, then αc is the convex set again. It is very easy to prove that. So, let us take two points in the set c and show that the line segment joining these two points also belong to the set c . So, let us take x_1 and x_2 in c and since, c is convex which is given to us, $\lambda x_1 + (1 - \lambda)x_2$ also belongs to the set c . Now, if you take αx_1 and αx_2 , where α is the real number, now clearly this αx_1 and αx_2 belongs to αc . What we have to show is that the line segment joining αx_1 and αx_2 should also belong to the set αc .

So, if we take the line segment joining αx_1 and αx_2 which is represented as $\lambda \alpha x_1 + (1 - \lambda)\alpha x_2$, we can rewrite it as $\alpha(\lambda x_1 + (1 - \lambda)x_2)$ and this is nothing but a point in the set c . So, this α into this quantity clearly lies in the set αc for any λ in the range 0 to 1, which means that αc is a convex set. So, this theorem tells that new convex sets can be generated by shrinking or expanding the given convex set appropriately.

(Refer Slide Time: 28:41)



Hyperplane

Definition
Let $b \in \mathbb{R}$ and $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a} \neq \mathbf{0}$. Then, the set

$$H = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} = b\}$$

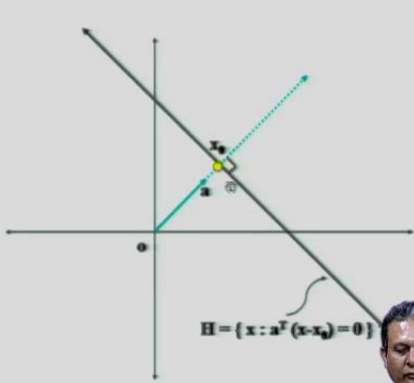
is said to be a *hyperplane* in \mathbb{R}^n .

- \mathbf{a} denotes the normal to the hyperplane H .
- If $\|\mathbf{a}\| = 1$, then $|b|$ is the distance of H from the origin.
- In \mathbb{R}^2 , hyperplane is a line
- In \mathbb{R}^3 , hyperplane is a plane





Now, let us look at the definition of hyperplane. So, suppose b is a real number and \mathbf{a} is a non-zero vector in \mathbb{R}^n . Then the set of all \mathbf{x} such that $\mathbf{a}^T \mathbf{x} = b$ is said to be a hyperplane in \mathbb{R}^n . So, let us see the significance of each of these terms in the set H . Now, the vector \mathbf{a} , denotes the normal to the hyperplane and if the norm of \mathbf{a} is 1, then the modulus of b , that absolute value of b is the distance of H from the origin. Now, in \mathbb{R}^2 in two-dimensional space, a hyperplane is a line and in three-dimensional space, a hyperplane is a plane and in high dimensional cases, we call this as a hyperplane.

(Refer Slide Time: 30:01)



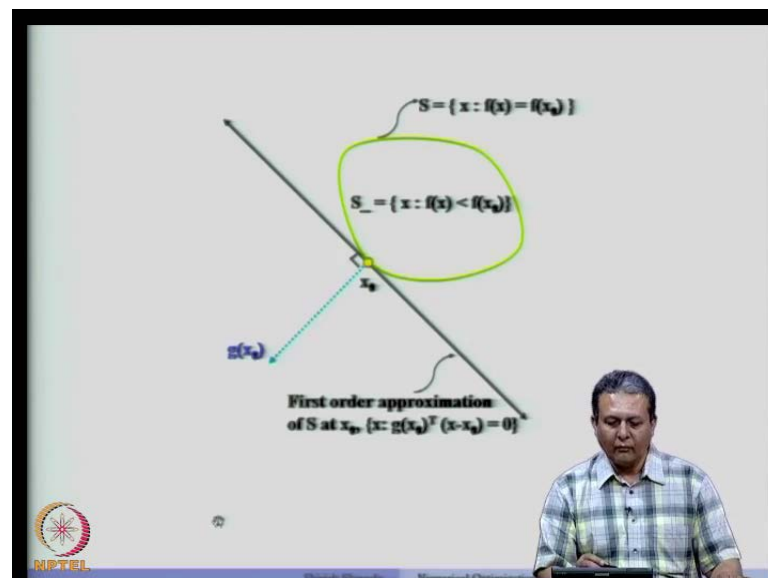
$H = \{\mathbf{x} : \mathbf{a}^T (\mathbf{x} - \mathbf{x}_0) = 0\}$



So, if you draw in two-dimensional space, it is a hyperplane. So, again let us consider a two-dimensional case, but the ideas are essentially the same problem, high dimensional cases. So, we have vector a , which is pointing in this direction and let us consider the set which is shown in by the black line as the set of all points x , such that $a^T x = b$. So, this is a transpose x equal to b is an equation of this hyperplane and if we collect all such x 's, that forms the set h and that is called the hyperplane.

So, remember that a is normal to this hyperplane. Now, if you are given a point which belongs to the hyperplane, then that this equation of the hyperplane can be written as, if x_0 is a point which lies on the hyperplane and we are still talking about the normal to the hyperplane as a , then the set of x , such that $a^T (x - x_0) = 0$. So, the $a^T (x - x_0) = 0$ forms the hyperplane. So, you will see that if we take any vector on this and form the vector $x - x_0$ and take a dot product of that with a , the dot product, they are perpendicular to any vector here and this vector a , they are perpendicular to each other or orthogonal to each other. So, their dot product is 0 and that is what is reflected in the definition of this hyperplane.

(Refer Slide Time: 31:48)



So, it is just another way to represent the hyperplane. If we know a point x_0 which belongs to the set to the hyperplane and if you know the normal a to the hyperplane, now if you are given a set s , if you are given some function f which is differentiable and the contour of that function, so the contour of that function we are denoting it by s , such that

set of all x , such that $f(x)$ is equal to $f(x_0)$ because the contour passes through the point x_0 . So, all the points on this contour will have the value of the function to be equal to $f(x_0)$ and as you move in the interior, the value of the function increases or decreases.

So, this is denoted by the set s minus. So, it is basically the set which is in the interior of this where $f(x)$ is strictly less than $f(x_0)$. So, as you move in the interior, the value of the function decreases. As you move away in the other side, away from this curve, then the value of the function increases. Now, if you take a point x_0 which belongs to this contour, then we can approximate this contour by a hyperplane and if we use the first derivative of the function, you remember that we have assumed the function f is differentiable. So, if we use the first derivative of the function and this vector g is a gradient vector of f at x_0 , so gradient of f at x_0 is nothing but $g(x_0)$.

So, then one can use the order approximation of s at x_0 as the set of all points, such that we have a normal to the hyperplane and the hyperplane passing through the points x_0 to set of all x , such that $g(x_0)^T(x - x_0) = 0$. So, this is going to be the first order approximation of the function of a differentiable function f at point x_0 . So, we will require this sometime later when we derive the optimality conditions. So, I just wanted to show that the first order approximation can be represented in the form like this.

(Refer Slide Time: 34:11)

Half-spaces

- The sets $H_+ = \{x : a^T x \geq b\}$ and $H_- = \{x : a^T x \leq b\}$ are called closed positive and negative *half-spaces* generated by H .

$\{x : a^T x = b\}$

H_+

H_-

a

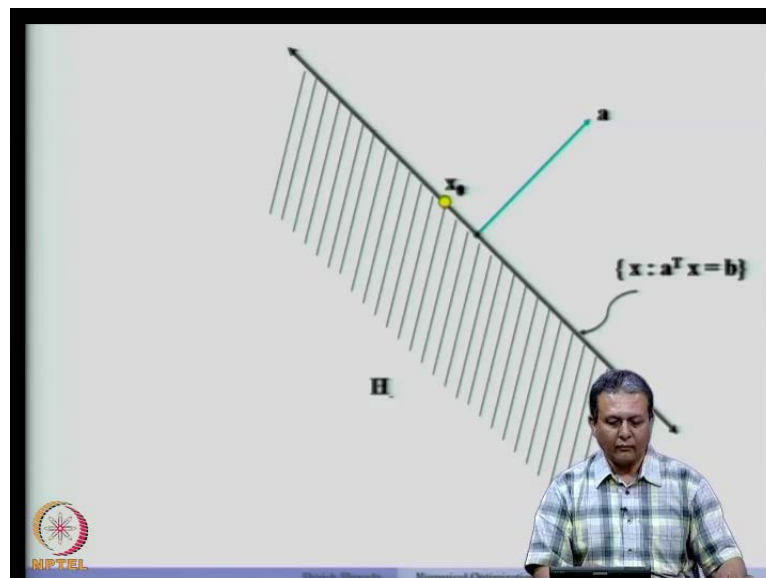
x_0

NPTEL

Now, let us look at half-spaces. So, the set h plus which is the collection of all x , such that $a^T x$ is greater than or equal to b , where a is again a non-zero vector in the space of n dimensions and b is a scalar. The set h plus is called a close positive half-space generated by the plane, hyperplane h . So, remember that the equation of the hyperplane h is set of all x , such that $a^T x$ equal to b . So, we are just considering the set of all points which lie on one side of this, which lies on or one side of this hyperplane.

Similarly, one can define the negative half-space set of all x , such that $a^T x$ less than or equal to b . Now, we will show this geometrically. So, we have hyperplane h and this a is normal to this hyperplane as this hyperplane is the set x , such that $a^T x$ equal to b and the points. This is the set of points which lie on this side of this hyperplane where a is pointing and that is called the positive half-space. So, the point on the hyperplane and on this side of the half of the hyperplane is called a closed half-space. Now, if we make the inequality strict that is the set of all x is that $a^T x$ greater than b , so these are the points which are not on the hyperplane, but are on this side of the hyperplane. So, that is called a open half-space, open positive half space.

(Refer Slide Time: 36:05)



(Refer Slide Time: 36:29)


Some more examples of convex sets

- $H = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} = b\}$ is a convex set
- Let $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m \in \mathbb{R}^n$ and $b_1, b_2, \dots, b_m \in \mathbb{R}$

Define, $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$

Then, $\{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}\}$ is a convex set.

- $H_+ = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} \geq b\}$ and $H_- = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} \leq b\}$ are convex sets.
- $\{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ and $\{\mathbf{x} : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ are convex sets.



© NPTEL

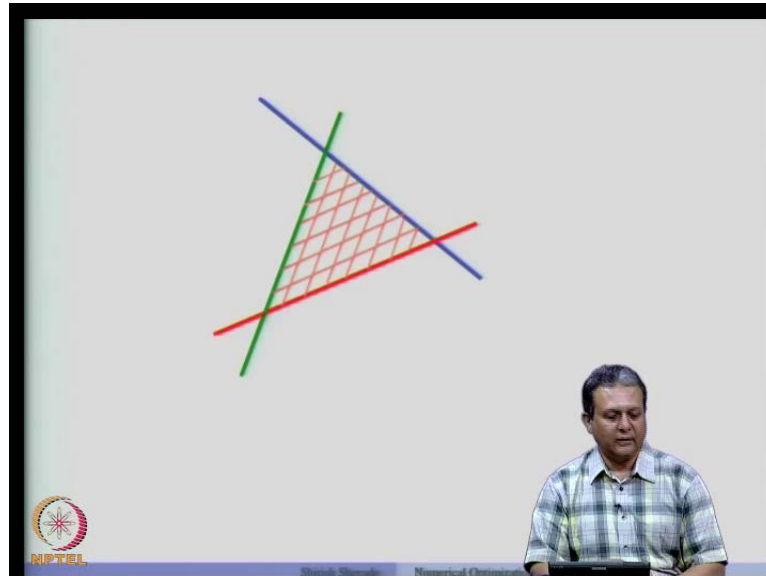
Similarly, one can talk about the closed negative half-space. So, again we are given a hyperplane whose normal is \mathbf{a} and the points on the hyperplane and points on the other side of the normal, they form a closed negative half-space. Now, with definitions of half-space, let us look at some more examples of convex sets. So, the hyperplane is a convex set. So, it is obvious because if you take any two points on the hyperplane h_1 and h_2 , the line segment joining those two points always lies in the hyperplane.

Now, let us consider m vectors in n dimensional space and m scalars b_1 to b_m . So, \mathbf{a}_1 to \mathbf{a}_m are the m vectors and suppose, we arrange them in the form of a matrix, where every row is \mathbf{a}_i^T denotes the transpose of the vector \mathbf{a}_i . So, you have m by n matrix and \mathbf{b} is a m dimensional vector. Now, the claim is that the set of all \mathbf{x} 's, that $\mathbf{A}\mathbf{x} = \mathbf{b}$. So, if you want to solve this system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$, then the solutions of that system of equations, it forms a convex set. Now, it is obvious because $\mathbf{a}_i^T \mathbf{x} = b_i$. This set can be written as $\mathbf{a}_1^T \mathbf{x} = b_1$, $\mathbf{a}_2^T \mathbf{x} = b_2$ and up to $\mathbf{a}_m^T \mathbf{x} = b_m$.

Now, we have already seen that set of all \mathbf{x} as that $\mathbf{a}_i^T \mathbf{x} = b_i$ is the convex set. So, each of $\mathbf{a}_i^T \mathbf{x} = b_i$, I am going from 1 to m is a convex set and what we are doing is that we are trying to take a intersection of this convex sets and we have seen earlier that the intersection of convex sets is a convex set. So, the solution set

of the system of linear equations x equal to b , where a is a matrix like this and b is a vector like this. It forms a convex set.

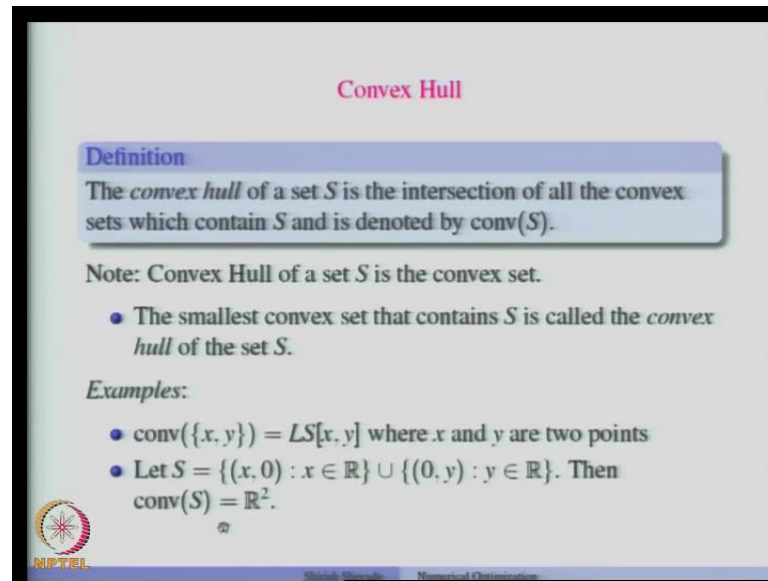
(Refer Slide Time: 39:05)



Now, if you take closed half-spaces, either positive or negative, both are convex sets. Now, again using the ideas similar to this, we can show that the set of all x 's, that $a x$ less than or equal to b or the set of all x 's at x greater than or equal to b , they form convex sets. Now, here is a simple example. So, let us consider a hyperplane. Although, I have shown arrows, but this is a hyperplane and then suppose we are interested in the closed half-space of this, formed by this hyperplane, the close half-space pointed by this arrow. So, clearly this is a convex set.

Now, let us consider another hyperplane. So, this another hyperplane and we are interested in the closed half-space of this hyperplane. The half-space which is formed which is denoted by this arrow. Now, if you take the intersection of the previous half-space and the current half-space, so it is basically the area between these two is the region between these two hyperplanes and that also is a convex set. If you add one more, so we get a triangle in this two-dimensional space. So, the interior of this triangle is a convex set. So, by using the existing convex sets is possible to construct new convex sets and that is illustrated here.

(Refer Slide Time: 40:22)



Convex Hull


Definition
The *convex hull* of a set S is the intersection of all the convex sets which contain S and is denoted by $\text{conv}(S)$.

Note: Convex Hull of a set S is the convex set.

- The smallest convex set that contains S is called the *convex hull* of the set S .

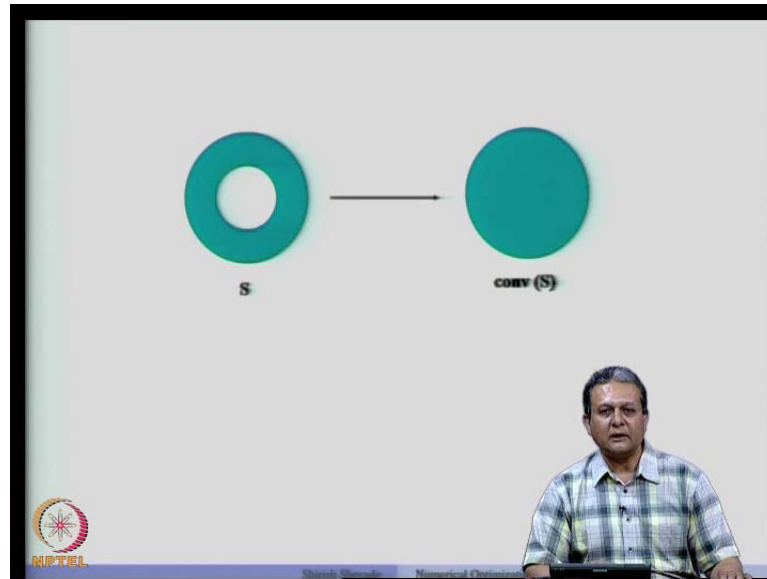
Examples:

- $\text{conv}(\{x, y\}) = LS[x, y]$ where x and y are two points
- Let $S = \{(x, 0) : x \in \mathbb{R}\} \cup \{(0, y) : y \in \mathbb{R}\}$. Then $\text{conv}(S) = \mathbb{R}^2$.

 NPTEL

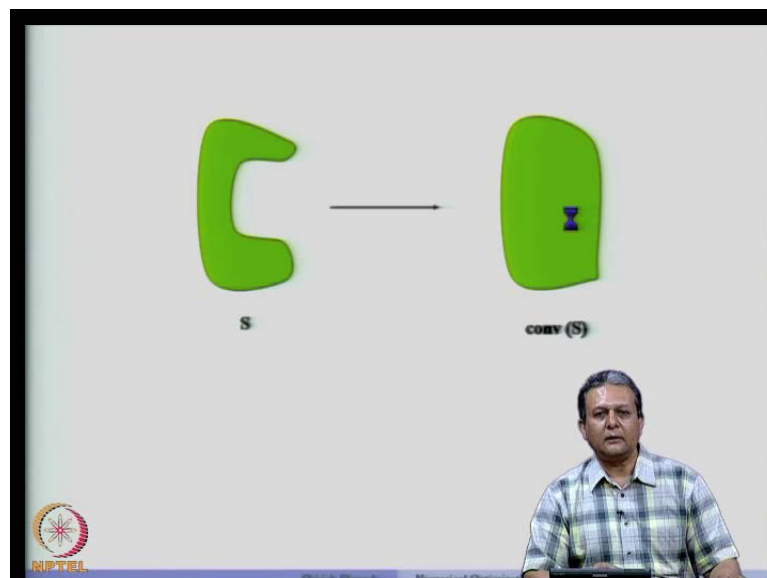
Now, let us look at the definition of a convex hull. The convex hull of a set s of a general set s is the intersection of all convex sets, which contain s and it is denoted by convex hull of x . So, it will be denoted like this. Now, by the very definition of convex hull, you will see that it is the intersection of all convex sets and we know that intersection of all convex sets is a convex set. So, every convex hull is a convex set. In fact, if you are given a set s , then sometimes we would like to form a convex set which contains s and such a convex set which is the smallest convex set is called the convex hull of the set s . So, many times this convex hulls are used for convex hull of a set s without bringing in any redundancy.

(Refer Slide Time: 41:59)

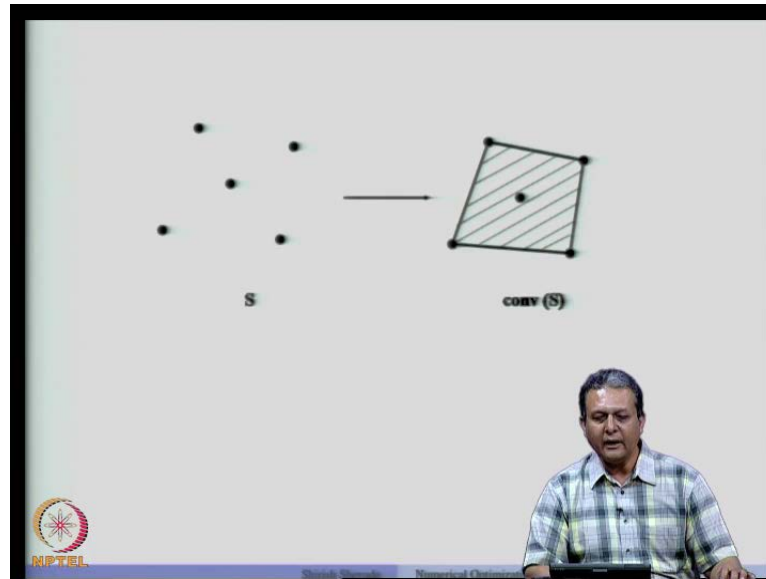


So, a convex hull can be thought of as a smallest convex set that contains the given set s , and we will see some examples. So, if you are given two points, then the line joining those two points, it is a convex hull of the set of those two points. Now, in two dimension, suppose if you are given x axis and y axis, then the convex hull of these two sets is entire two-dimensional space. Now, when we studied convex sets, we saw this example and we demonstrated that this s is not a convex set.

(Refer Slide Time: 42:26)

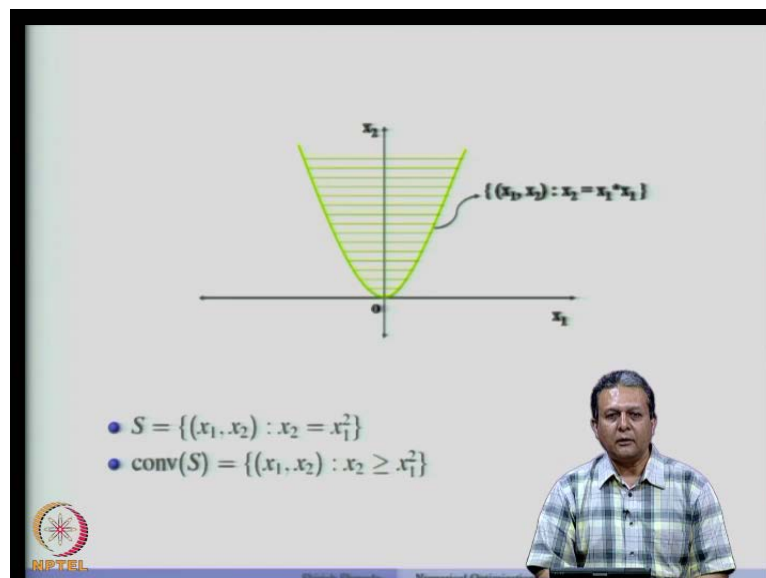


(Refer Slide Time: 42:39)



Now, one can form a convex set of this set. So, that becomes this set. Now, this set is a convex set as you will see that you take any two points and this set, it will always form a, that line segment always lies in the set. Now, here is another example which we saw that this is s is not a convex set, but then one can find the convex hull of this set and suppose, we are given a set of points and if you want to find the convex hull of this, so it is a smallest convex set that encompasses all the points. So, this is the smallest convex set which encompasses all the points.

(Refer Slide Time: 43:06)



So, remember that the points in the interior, they are not used in forming the convex set or representing the boundary of the convex hull. So, here is one more example that we consider the set x_1, x_2 , say that x_2 equal to a x_1 square. So, this was the given set and the convex hull of this set is the set of points on this curve, and the points above this curve. So, the shaded region here plus the points on the curve that together forms the convex hull of the set. So, this is basically the set x_1, x_1 such that x_2 greater than or equal to x_1 square. So, we will stop here and in the next lecture, we will continue with the convex sets and some more results related to the convex sets.

Thank you.