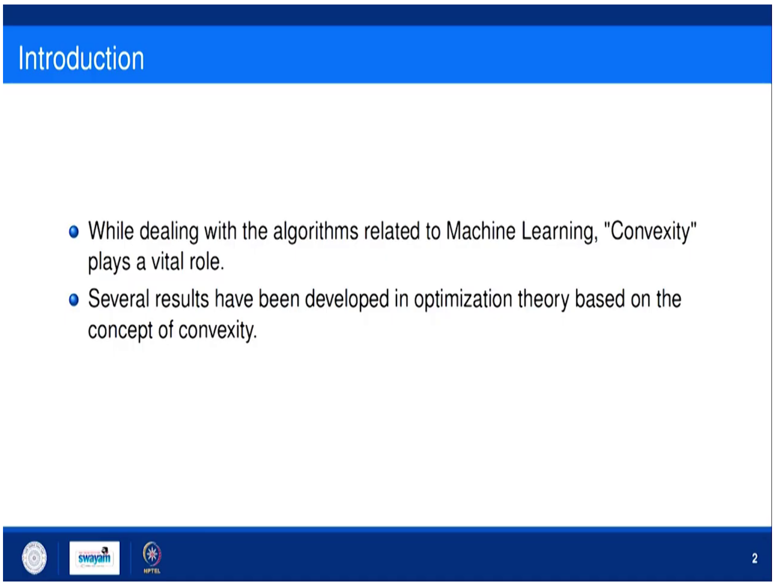
Essential Mathematics for Machine Learning
Prof. S. K. Gupta
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture – 23
Convex Sets and Functions

Hello friends. So, welcome to lecture series on Essential Mathematics for Machine Learning. In the last lectures, we have seen the basic concepts of calculus. We have seen that how we can find out gradient, Jacobian, hessian matrix, partial differentiation, etcetera, if a function of several variable is known to us.

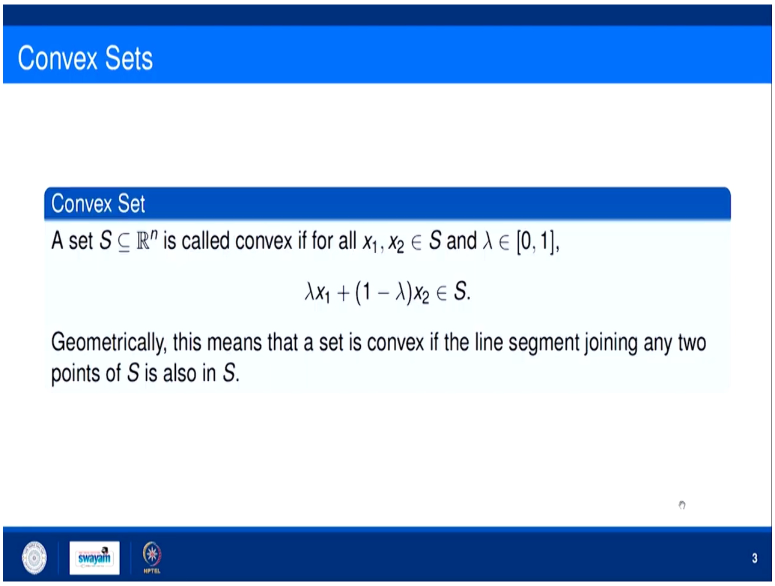
Now, in these lectures we will see that, what are the basic concepts of optimization which is required in machine learning. So, let us start with Convex Sets. What convex sets are?

(Refer Slide Time: 00:54)



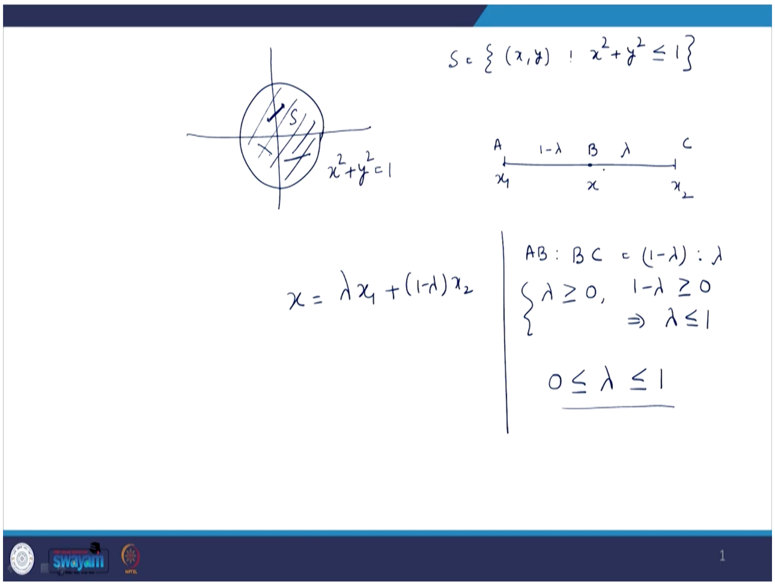
So, while dealing with the algorithm related to machine learning convexity plays a vital role. Several results have been developed in optimization theory based on the concepts of convexity, what are they we will discuss later on. First we will go the basic definitions of convex set and convex functions.

(Refer Slide Time: 01:14)



So, when a set is said to be a convex set. So, a set S subset of R n is called the convex if for all x 1 x 2 in S and lambda between 0 and 1 lambda x 1 plus 1 minus lambda x 2 is also in S. So, what do you mean by this let us see, let us understand try to understand.

(Refer Slide Time: 01:38)



Say you have a circle, say the circle is x square plus y square equal to 1. And suppose we are dealing with the region inside the circle, ok, suppose we are dealing with the region inside the circle that means, that means this region, ok. So, this is S. So, what is S? Basically, x y is all those x y such that x square plus y square is less than equal to 1, ok.

So, we are dealing with all those x y in R 2 such that x square plus y square is less than equal to 1. So, what this S is? This is including boundary and including all the point inside this circle. Now, if you take any two point in this region, you take any two point, anywhere in this region join this line segment, this line segment is always inside this region. You take two points here and here join; the line segment the line segment is inside this region. You take two points here and here join the line segment; line segment is inside the region. So, this set is a convex set.

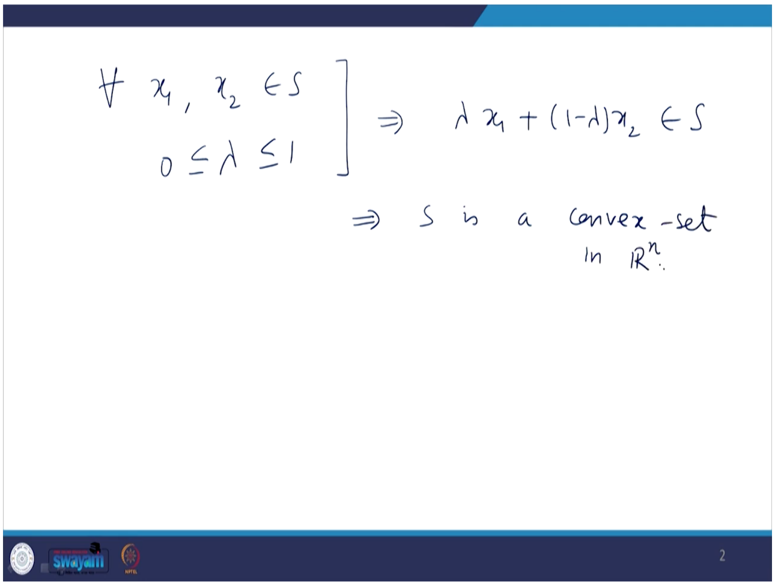
So, convex set means that if you take any two arbitrary point in the set, the line segment joining those two point lies completely inside that set. So, that set is called a convex set. So, geometrically, if we say that a set is convex or set may not be convex, if a set is convex that means, you take any two arbitrary point in the set, join the line segment you take two points, join the line segment, the line segment joining these two point lies totally inside the set so that means, that is a convex set.

So, suppose you have two points say x 1 and you have another point say x 2. And say you have a any point x in between these two points, supposes x divide this line segment, this line segment in the ratio say 1 minus lambda is to lambda, ok. This is A, this is B, suppose this is C, then I am taking AB is to BC as 1 minus lambda is to lambda, ok. So, so if you want to find out. So, first of all this B must be in between x 1 and x 2.

So, to ensure that this B is in between x 1 and x 2 that means, these two ratios must be non-negative because if any one of them is negative that means, this point B will go outside C or outside A. We want B to B in between x 1 and x 2, ok. So, that means, ratios must be non-negative, lambda should be non-negative and 1 minus lambda should be non-negative. So, this implies lambda is less than equal to 1. So, from these two we can say that lambda is less than equal to 1 and greater than equal to 0.

And if you take x; so, what are the coordinates of x? What x will be? So, x will be nothing, but for from the section formula one can easily see that this x is nothing, but lambda x 1 plus 1 minus lambda x 2, ok. Lambda varies between 0 and 1; you take any lambda between 0 and 1 this x will move this x will move in between x 1, x 2. You take lambda equal to half it will give a midpoint, you take lambda equal to 2 by 3 then this will be a some point over here. So, in this way this lambda will move in this region in this line segment x 1 and x 2.

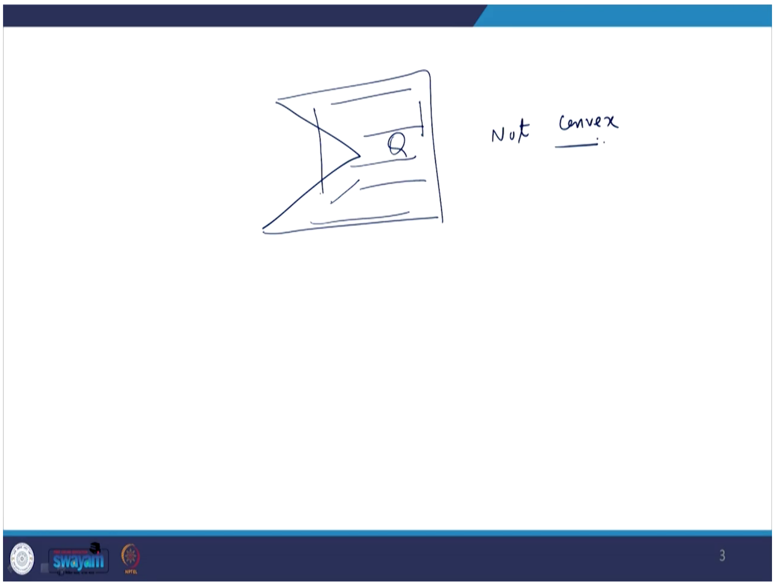
(Refer Slide Time: 05:35)



We want we want that any x which is in between x 1 and x 2 should they must belongs to S so that means, for all x 1, x 2 in S and lambda between 0 and 1 lambda x 1 plus or minus lambda x 2 must be in S.

So, what we have; what we have shown that if you take any two arbitrary point in S you take lambda between 0 and 1, then this represent the line segment joining x 1 and x 2, this. Because as lambda vary, this will vary, and this will give all the points in the line segment joining x 1, x 2 and this must belongs to S. That means, the line segment joining x 1, x 2 lies totally inside S, this means S is convex, S is a convex set in R n, ok. Suppose, you have this type of figure, ok. This is a region. Say this region is Q.

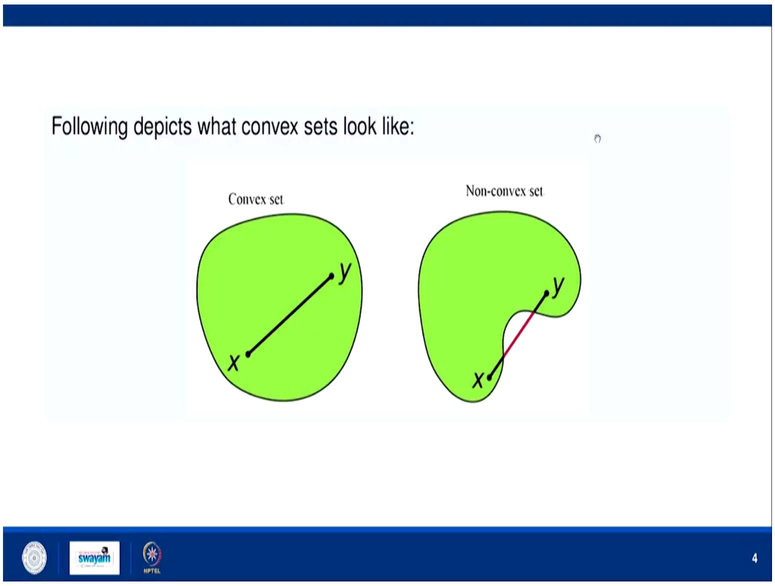
(Refer Slide Time: 06:30)



And you want to see whether this region Q is convex or not. If you take two points here and here join the line segment, the line segment joining these two points is totally inside the region, ok. If you take two points here and here joining the line segment the line segment is totally inside the region. But if you take two points here and here joining the line segment the line segment is not totally inside the region. So, this set is not convex.

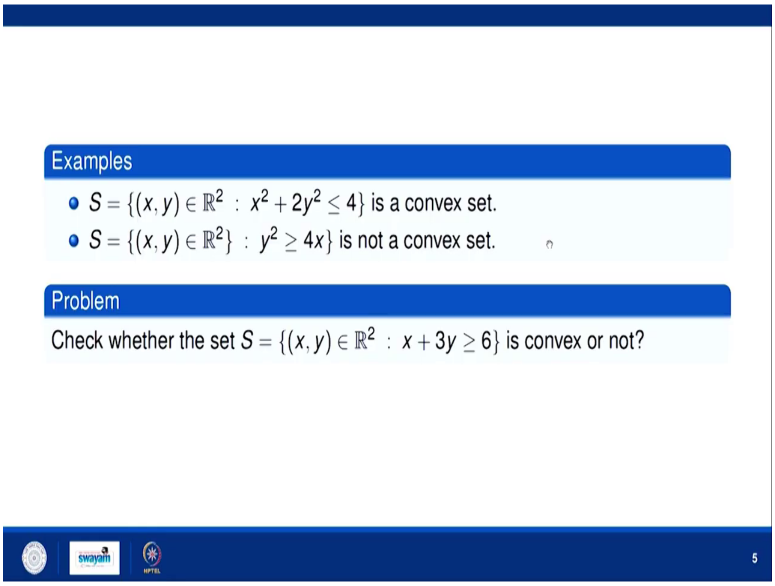
So, we can see that if a set is convex that means, that means you take any two arbitrary point in the set, join the line segment, the line segment must be totally inside the inside the set S. But if there exists at least one pair such that the line segment joining those two point does not lies completely inside the region S, then set is not a convex set, ok.

(Refer Slide Time: 07:34)



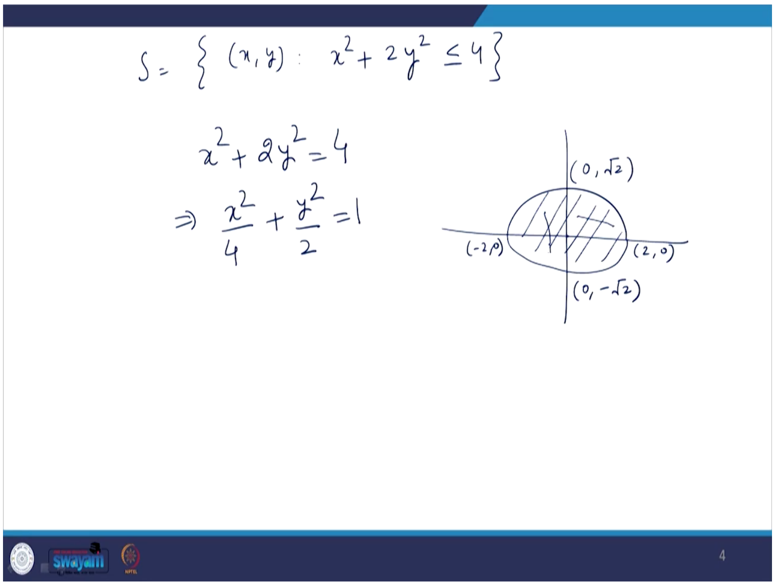
So, these are the figures where set is convex or set is not convex because this line segment is outside the region outer this green portion. So, this not is, this set is not convex however, this is a convex set. So, let us discuss few examples based on this.

(Refer Slide Time: 07:46)



So, the first example is all those x, y such that x square plus 2 y square less than equal to 4.

(Refer Slide Time: 07:58)

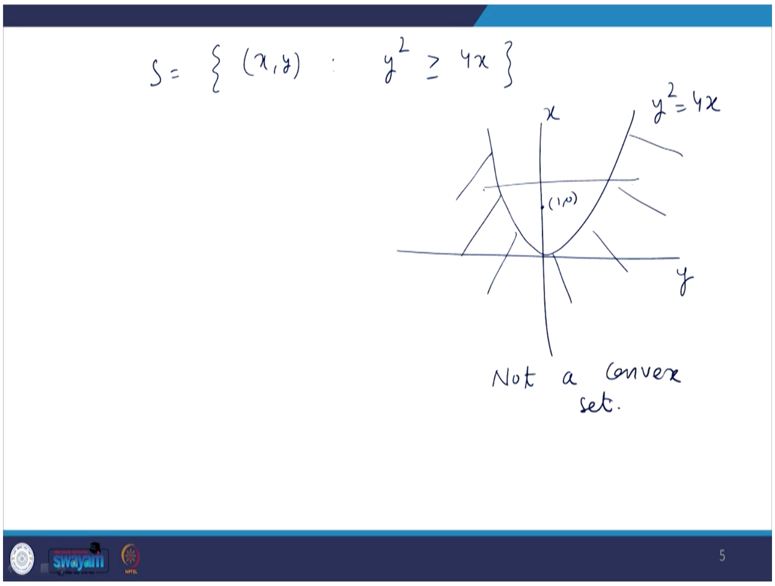


So, if you take S which is all those x, y such that x square plus 2 y square less than equal to 4, then this represent what? Let us see. This represent, see this is this is x square plus 2 y square, first you take equality sign. So, this is basically x square by 4 plus y square upon 2 equal to 1 and this is nothing but ellipse, ok. So, ellipse with a square 2, a square 4 and b square 2. So, that is something this kind of ellipse, ok. So, this point is 2 comma 0, this point is minus 2 comma 0, this point is 0 comma under root 2 and this point is 0 comma under root minus 2.

Now, if you take a point say 0 0. So, 0 plus 0 is less than equal to 4 which is true, that means, shade inside the region otherwise we shade outside the region. Now, if you take any two point any two arbitrary point anywhere in this region and join the line segment the line segment is totally inside the region so that means, the set is a convex set, you take any two arbitrary point anywhere. So, this set is a convex set, ok.

Now, if you see the second example the second example where x, y, all those x, y in R 2 such that y square is greater than equal to 4 x.

(Refer Slide Time: 09:30)

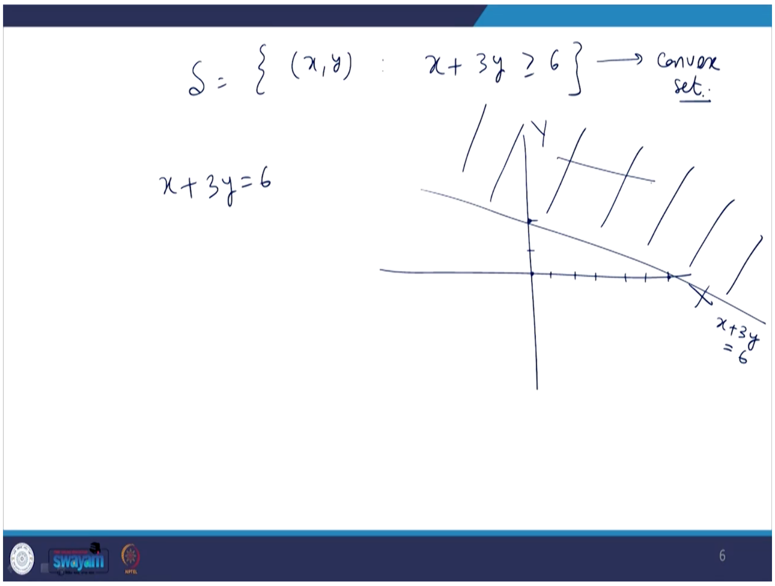


If you see this it is all those x, y such that y square is greater than equal to 4 x. So, first you first you take the equality sign which is nothing, but a parabola. So, you draw a parabola. So, this is y square equal to 4 x, ok. So, this is basically x and this is y, ok. So, if you take in this way then it is ok, because symmetrical about it is a symmetrical about x axis, so it is y square equal to 4 x.

Now, if you take a point say you take a point here say it is 1 comma 0, if you take a point on x axis which is 1 comma 0, if you take a point 1 comma 0 then it is y 0 and it is 1. So, that means, this is not satisfied so that means, region outside the parabola, ok.

Now, if you take two points, one point here one point here and join the line segment, the line segment goes outside the region. So, this set is not a convex set, ok. Now, we have a problem here check whether the given set this is convex or not. So, let us let us draw this graph of this S and let us see whether given region is convex or not. So, what is the region? What is the region S?

(Refer Slide Time: 11:02)

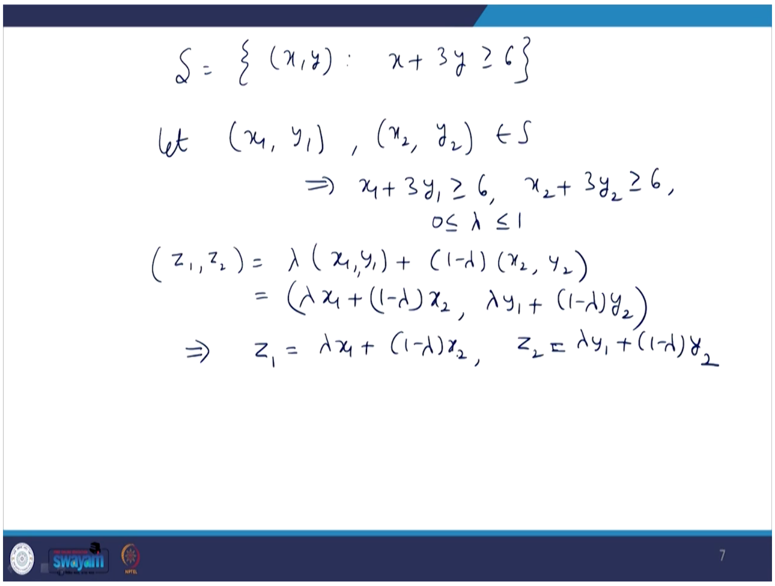


The region S is all those x, y in R 2 such that x plus 3 y is greater than equal to 6, ok. So, first you draw the line. Line is x plus 3, y equal to 6. So, let us say x axis here, y axis here. So, when you put y equal to 0 x is 6, so 1, 2, 3, 4, 5, 6, somewhere here and when x is 0, y is 2. So, y is 2 that is somewhere here. So, you draw this line, ok. So, this is x plus 3 y equal to 6.

Now, you take a point on either side of the line, either this side or this side, you suppose take a point 0 0. Now, 0 plus 0 is greater than equal to 6 it is not true, it is false that been shade opposite to this region, shade opposite to this point that is here. So, one can clearly see that this region is convex region because if you if you take any two arbitrary point in the region, join the line segment, the line segment lies totally inside the region. So, this is a convex set.

So, if we want to prove this is the geometric interpretation that you want to show that whether given set is convex or not, so you have shown by drawing the region that whether given set is convex or not. But if you want to prove analytically, mathematically that also you can show. Suppose, you have to show that this given set as is a convex set very simple example. So, let us see whether this given set is convex or not.

(Refer Slide Time: 12:42)



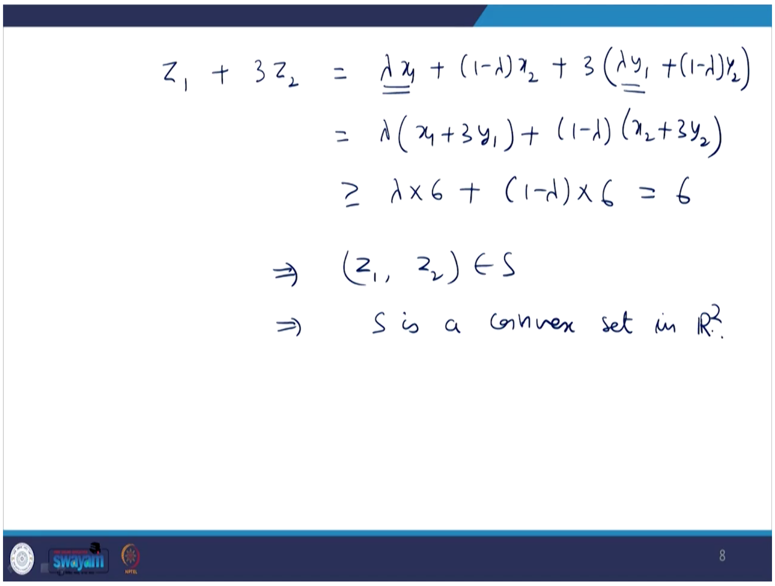
So, this is this S is what? All x, y such that x plus 3 y greater than equal to 6. Now, we have seen we have seen geometrically that this set is a convex set and you want to prove it mathematically. So, how we can show?

In order to show that a given set is a convex set or not, you take any two arbitrary point in S, and try to show that that lambda x 1 plus 1 minus lambda x 2 is also in S. So, lambda x 1 plus 1 minus lambda x 2 is also called convex linear combination of x 1 and x 2, ok. So, so we take two point, let x 1, y 1 and x 2, y 2 belongs to S. These are two arbitrary point we have taken in S. So, this implies x 1 plus 3 y 1 less than equal to greater than equal to 6 and x 2 plus 3 y 2 greater than equal to 6. Now, take lambda between 0 and 1.

Now, take another point say z 1, z 2, which is lambda time the first point and 1 minus lambda time the second point. And we have to show that this point lambda x plus 1 minus lambda y or lambda x 1 plus 1 minus lambda y lambda y 1 must belongs to S. Then, we then only it should be a convex set because lambda is varying and we are taken two arbitrary point in S.

So, what this is? This is equal to lambda x 1 plus 1 minus lambda x 2, this is lambda y 1 plus 1 minus lambda y 2. So, this implies z 1 is because first component will be equal to first component, second will be equal to a second, so this will be equal to lambda x 1 plus 1 minus lambda x 2 and z 2 will be equal to lambda y 1 plus 1 minus lambda y 2. So, we have to show that z 1, z 2 belongs to S. So, that means, we have to show that z 1 plus 3 z 1 plus 3 z 2 must be greater than or equal to 6, then only z 1 z 2 belongs to S, ok.

(Refer Slide Time: 15:02)

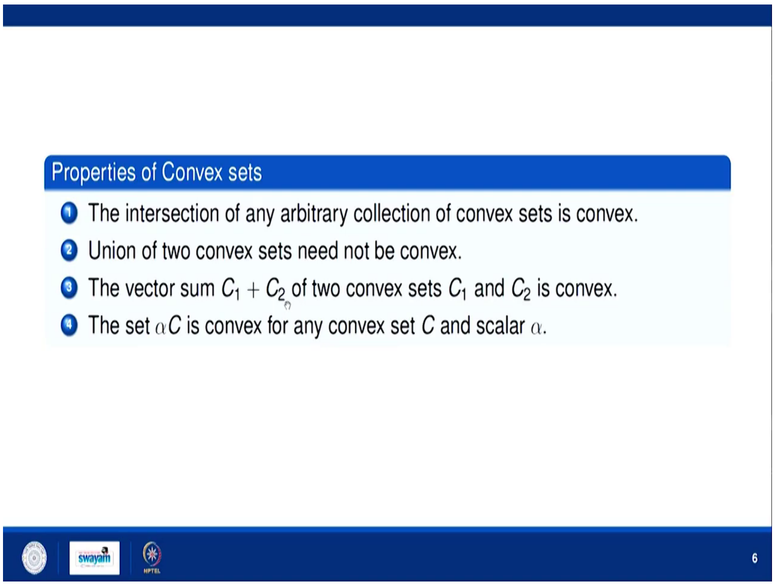


So, let us take z 1 plus 3 z 2. So, what is z 1? See z 1 is lambda x 1 plus 1 minus lambda x 2. So, substitute lambda x 1 plus 1 minus lambda x 2 plus 3 times. What is z 2? z 2 is lambda y 1 plus 1 minus lambda y 2, you substitute it is lambda y 1 plus 1 minus lambda y 2.

So, this is further equals to lambda times you take lambda common from this and this. So, lambda times x 1 plus 3 y 1 and 1 minus lambda times x 2 plus 3 y 2. Now, but x 1 plus 3 y 1 greater than or equal to 6 and x 2 plus 3 y 2 is also greater than or equal to 6 and lambda between 0 and 1, so that means, this is greater than equal to lambda into 6 plus 1 minus lambda into 6, this is equal to 6 and this implies z 1 z 2 belongs to S and this implies S is a convex set in R 2.

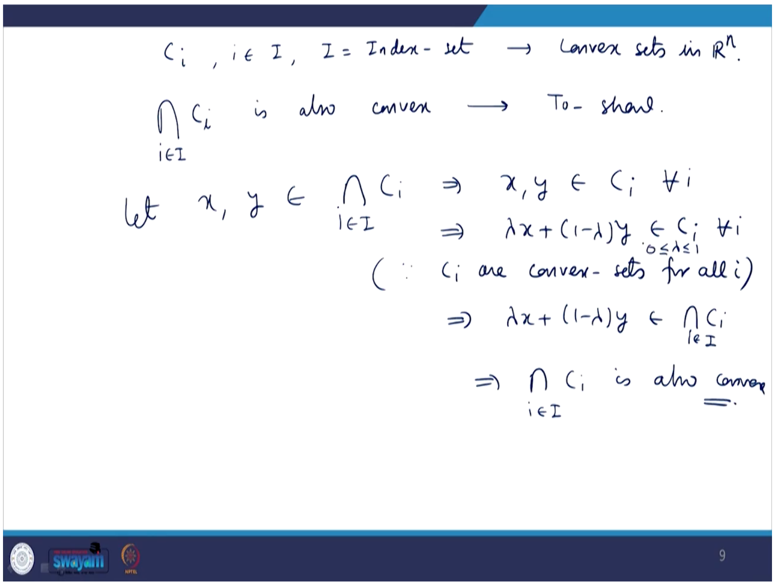
So, in this way we can show that we can show analytically also that a given set is convex. And he want to show that a given set is not convex, it is very simple you try to give a counter example such that two point is in S, but their CLC may not be in S for some lambda between 0 and 1, ok.

(Refer Slide Time: 16:40)



Now, these are the few properties of convex sets. The first property is the intersection of any arbitrary collection of convex set is convex. So, we will see one these properties one by one. The first property is very easy to show.

(Refer Slide Time: 16:56)



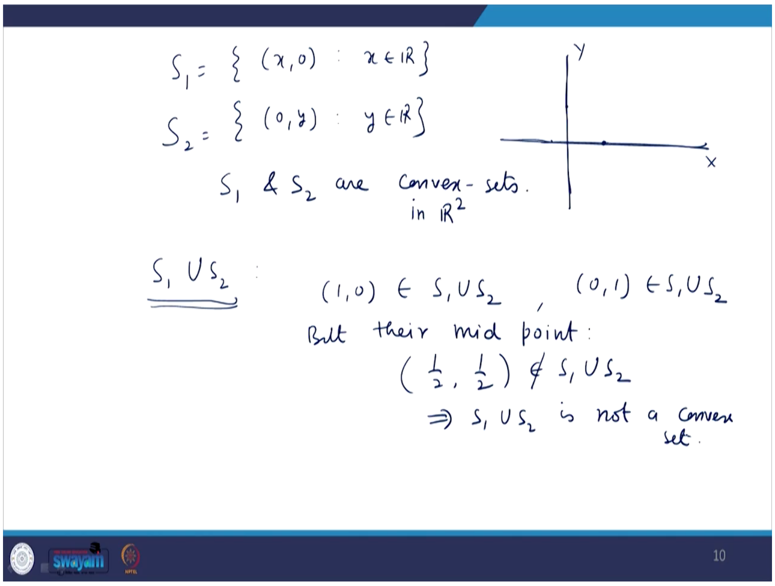
Suppose, C i and where i belongs to I, where I is an index set, ok. Suppose, these are convex sets in R n. And you have to show that i, i belongs to I C i is also convex is also a convex set, this to show. Then only we can show that arbitrary intersection of convex set is convex.

So, in order to show that this set is convex you take any two arbitrary point in a in this set and try to show that the CLC of those two point is also in the same set. So, let x and y belongs to this. So, this implies x, y belongs to C i for all I, because if x y is an intersection that means, x, y is in C i for all I and this implies lambda x plus 1 minus lambda y is also in C i for all I, because all C i's are convex sets, because since C i's are convex sets for all I.

So, this implies now since this is belongs to C i for all I this means lambda x plus 1 minus lambda y belongs to intersection of C i also, i belongs to I and this implies intersection i belongs to I of C i is also convex. So, we have shown that arbitrary intersection of convex set is also convex. We have taken two arbitrary point and we have shown that the CLC of those two points is also in the same set that means, this set is a convex set.

Of course, lambda is in between 0 and 1 for this lambda must be in 0 and 1. So, we have to write somewhere that because lambda is in between 0 and 1, ok. The second one is union of two convex set need not be a convex set. So, it is very easy. Union, if you if you see the union of two convex sets.

(Refer Slide Time: 19:31)



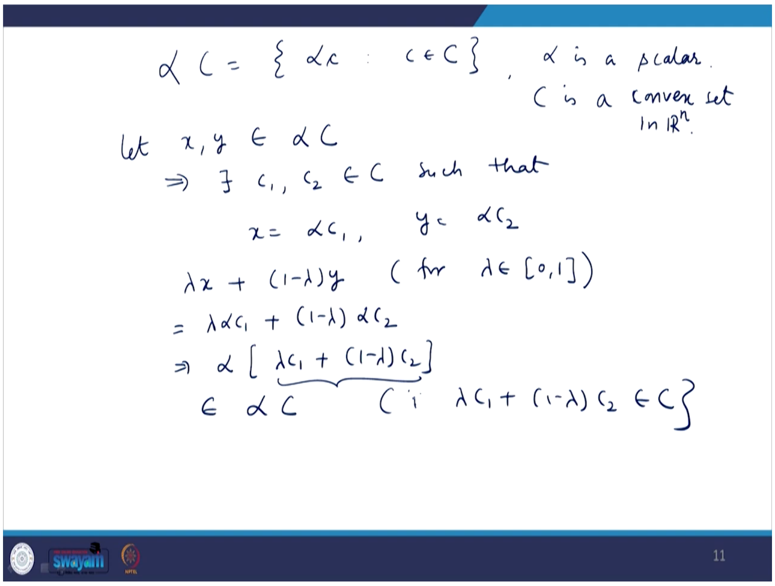
Suppose one set is, suppose x comma 0 where x belongs to R that means, x axis. So, x axis is clearly a convex set, you can easily see. You take any two arbitrary point on the set join the line segment, line segment is totally inside the see. You take S 2, suppose it is 0 comma y such that y belongs to R. This is y axis basically, is also convex. So, S 1, S 2 are convex sets in R 2, ok.

Now, if you take the union of S 1 and S 2, union means all the points on x axis and all the points on y axis. Now, if you want to show that this is not convex, so you take say one comma 0 belongs to S 1 union S 2. Of course, it is an x axis. And 0 comma one also belongs to S 1 union S 2 because it is in S 2. But if you take the midpoint of these two, but their midpoint. What is their midpoint? 1 by 2 comma 1 by 2. Now, this point is neither on x axis nor on y axis. So, this point does not belongs to S 1 union S 2, and this implies S union S 2 is not convex, is not a convex set.

So, to disprove that a given set is not a convex set, we have to give a counter example. So, here is a counter example. We can construct infinite number of counter example of such type, where S 1, S 2 are convex, but union need not be a convex set, ok. Now, in the same way we can easily show that the vector sum that is C 1 plus C 2 of two convex sets C 1 and C 2 is also convex. So, how to show this? You have to take two arbitrary point in C C 1 plus C 2 and try to show that the CLC of those two point is also in the same set.

Now, the set alpha C, alpha times C is a convex for is a convex set for any convex set C and the scalar alpha. So, this is also easy to show. Let us see how we can prove this.

(Refer Slide Time: 22:10)

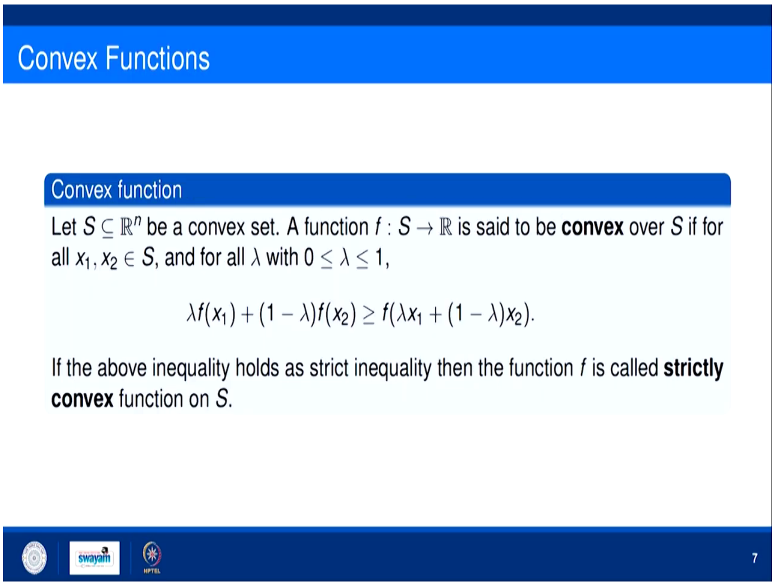


So, how alpha times C is defined? It is all alpha times C such that alpha such that C belongs to C. Here alpha is a fixed scalar, alpha is a scalar, ok. So, here it is given to us that C is a convex set in R n, C is a convex set in R n, ok. And you have to show that alpha C is also convex. So, again you have to take two arbitrary point in this set, let x and y belongs to alpha C and we have to show that the CLC of these two point is also in the same set.

So, if x and y belongs to alpha C, this implies that there will exist some C 1, C 2 belongs to C such that x will be alpha time C 1 and y will be alpha time C 2, ok. Now, you take up you take the CLC of these two point lambda x plus 1 minus lambda y, for lambda between 0 and 1. You take this. So, this will be nothing but lambda times alpha C 1 plus 1 minus lambda times alpha C 2. So, this imply you take alpha common. So, it is lambda C 1 plus 1 minus lambda C 2.

Now, this point will be in C because C is a convex set, because C 1, C 2 is in C and lambda between 0 and 1. So, this point would be some point in C. And this is alpha times that point. So, alpha times a point in C is in alpha C. So, this means this belongs to alpha C because lambda C 1 plus 1 minus lambda C 2 is in C. So, in this way we have we have shown that alpha C is also a convex set for a scalar alpha and C is a convex set, ok. So, these are the few properties of convex set which we will used in optimization basically. So, one should know that these properties are important, ok.

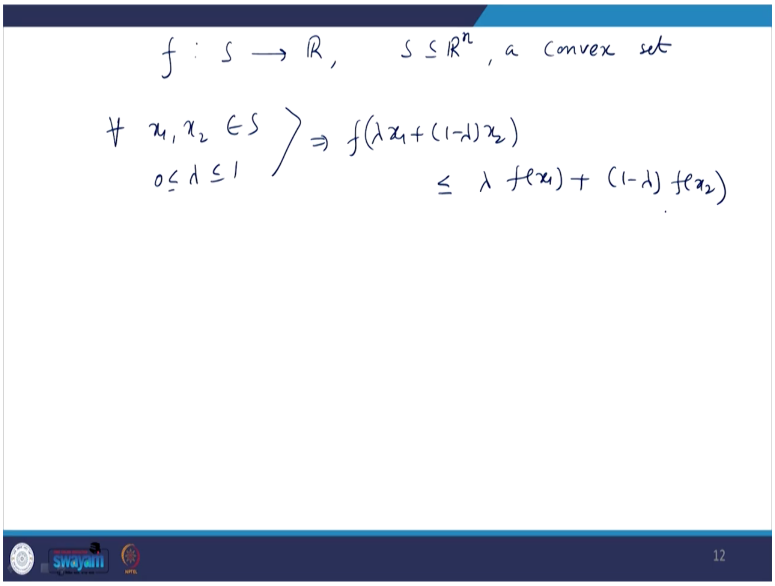
(Refer Slide Time: 24:31)



Now, let us come to convex functions. What convex functions are? So, let us try to discuss that let S be a subset of R n be a convex set. A function f from S to R, S is a convex set, from a convex set to R a set to be convex over S if for every x 1, x 2 in S and for any lambda between 0 and 1 this inequality hold. And if this inequality hold in a stick sense that means, instead of greater than equal to it is strictly greater than this thing then the function is called strictly convex function.

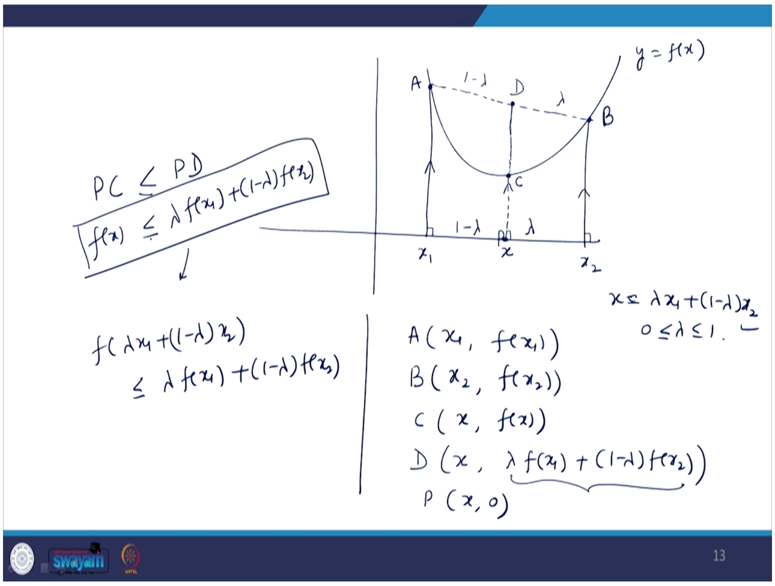
So, let us try to analyze the definition of convex functions geometrically, so that one can understand that how these functions are important.

(Refer Slide Time: 25:27)



So, what is the definition of convex function? That you have a function f from S to R where S is a subset of R n a convex set, and a convex set sorry it is a convex set, and for all x 1, x 2 in S and lambda between 0 and 1 if this implies that f of lambda x 1 plus 1 minus lambda x 2 is less than equals to lambda f x 1 plus 1 minus lambda f x 2. So, what does what is a geometrical interpretation of this definition, so let us try to understand that also.

(Refer Slide Time: 26:32)



So, suppose we have this type of this type of curve. This is y equal to f x may we take any two arbitrary point on the on the curve say x 1 and x 2, ok. Now, what are coordinates of this point A? The coordinate of point A will be x 1 comma f x 1, because x is x 1 and y is given by this expression y equal to f x 1. Now, what is the coordinate of point B? Coordinate of B point will be x 2 comma f x 2, ok.

Now, you take any point x in between x 1 and x 2 and suppose this point divide this line segment to the ratio 1 minus lambda is to lambda. So, of course, this x will be nothing but lambda x 1 plus 1 minus lambda x 2 for lambda between 0 and 1. So, you have taken an arbitrary point x in between x 1 and x 2. And if you if you draw this line vertically, then this point C this point C will be nothing but x comma f x.

And what is x? x is CLC of x 1 and x 2 that is lambda x 1 plus 1 minus lambda x 2. Now, if you draw this chord AB, if you draw this chord AB and you extend this line up to this chord, now suppose this point is D, ok. Now, these 3 lines are parallel, these 3 lines are parallel, all are making 90 degree angle with x axis. So, the ratio which is here the same ratio will be here.

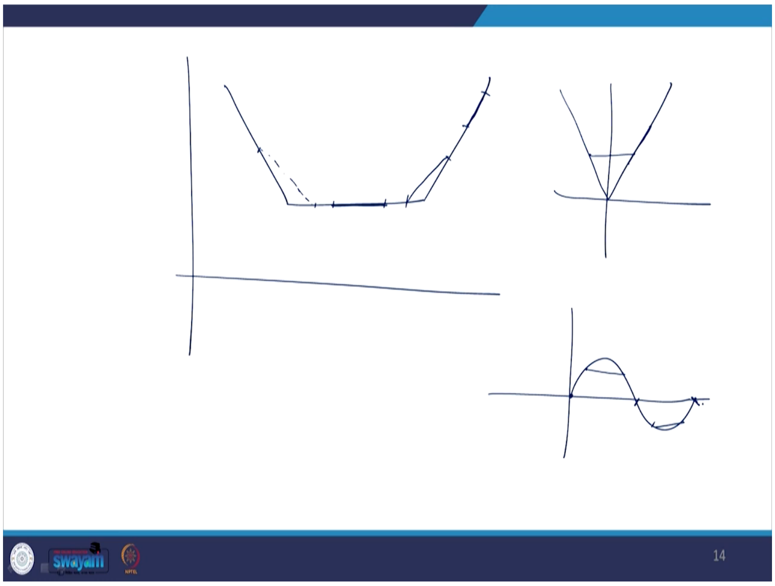
So, what will be the coordinates of D point then? The coordinate of D point will be x will remain x and the f of this point D will be given by because you know A point you know B point, by the section formula you can easily find out point D. So, that will be lambda f x 1 plus 1 minus lambda f x 2 for that will be D point. Now, if this point is suppose P, this x point, this x comma 0 is suppose P, P is x comma 0. So, one can easily see that P C if it is the straight line, here, here if it is the straight line it may be equality also, ok. And what is P C? P C is nothing but P C is nothing but f x is less than equals to.

And what is P D? P D is this entire line and this entire, this entire line is nothing but this thing this thing is the P D, which is lambda f x 1 plus 1 minus lambda f x 2. So, which is nothing but, this is nothing but the definition of convex function; x is x is this thing, ok. So, we can write it like f of lambda x 1 plus 1 minus lambda x 2 is less than equals to lambda f x 1 plus 1 minus lambda f x 2.

So, what we have analyzed basically? See, that means, if you take any two arbitrary point on the function, on the curve and join the chord then the chord is always above or on the curve. See, this chord is above the curve that is why P C is less than equal to P D, ok. That is why this inequality is holding.

So that means, that if you take any two arbitrary point on the curve, join the chord, the chord is always, the chord joining those two points is always lie above or on the curve, suppose this type of function, suppose you have this type of function.

(Refer Slide Time: 30:53)

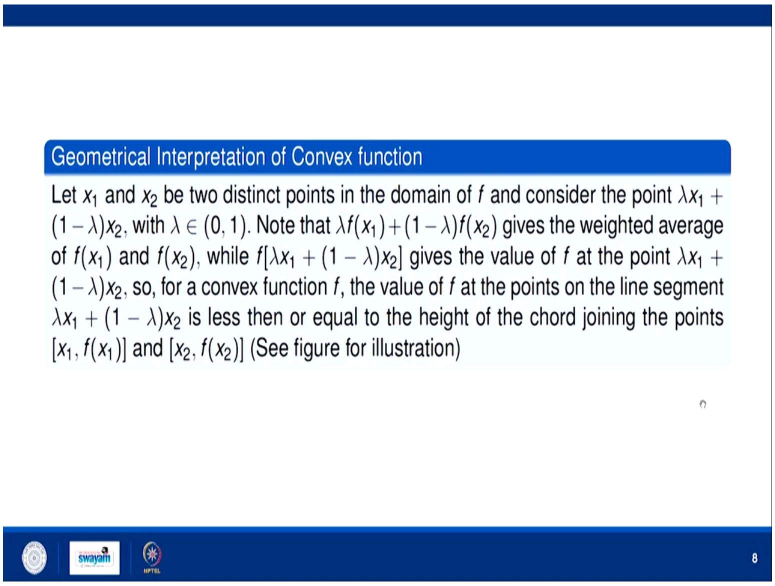


So, now, you have to see that this is convex or not. See, if you take any two arbitrary point on this on the on the function, on the curve, then join the chord, the chord is always above the above the curve. You take two point here, chord is on the curve. You take two point here, join the chord, chord is above the curve; two point here, chord is on the curve. So, chord is all either above the curve or on the curve. So, this function is convex.

You take a function like this say this function. You take any two arbitrary point on, this join the chord, chord is always above the curve. You take two, one point here one point here join the chord, chord is always here on the curve. So, in this way we can see that this function is also a convex function.

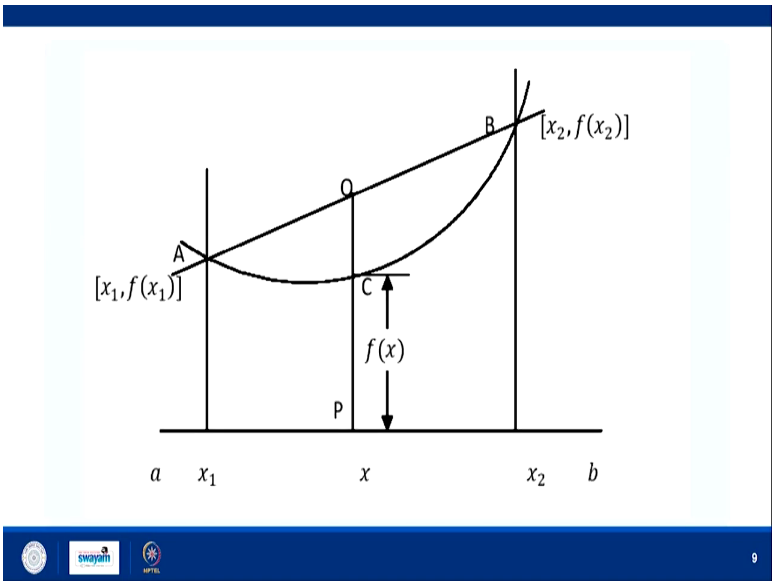
If you say this have a function, in this function. Now, if you take one point here one point here join the chord, chord is below the curve, so this function is not convex function. However, in this portion from here to here it is convex. If you take one point here one point here join the chord, chord is above the curve, in this portion it is convex. But if you see as a whole from here to here, so from here to here, it is not a convex function, ok.

(Refer Slide Time: 32:16)



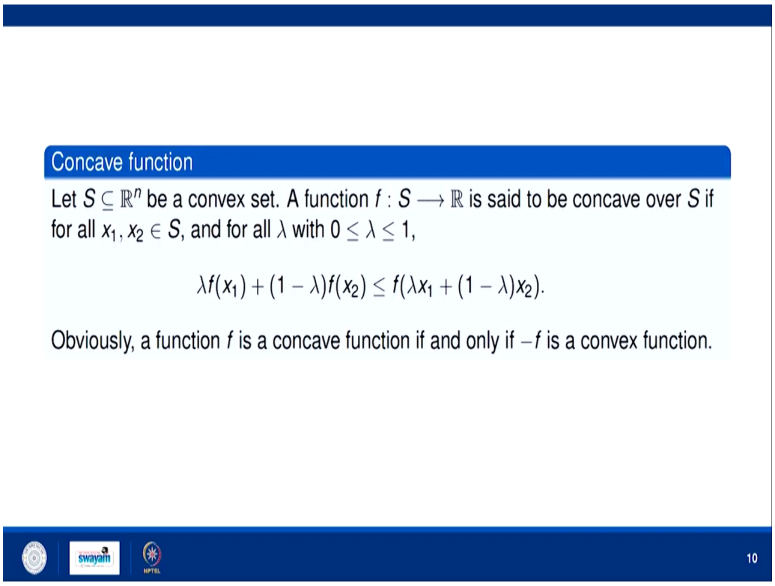
So, here is a geometric interpretation of convex functions which we have just discussed.

(Refer Slide Time: 32:18)



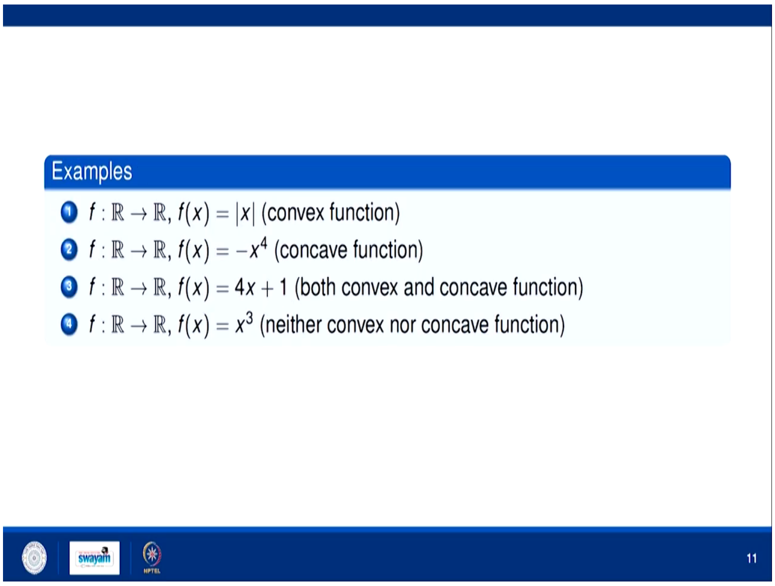
Now, concave function. So, a function f is called a concave function if this inequality is reversed, ok. That means, a concave function if and only if minus f is a convex function.

(Refer Slide Time: 32:22)



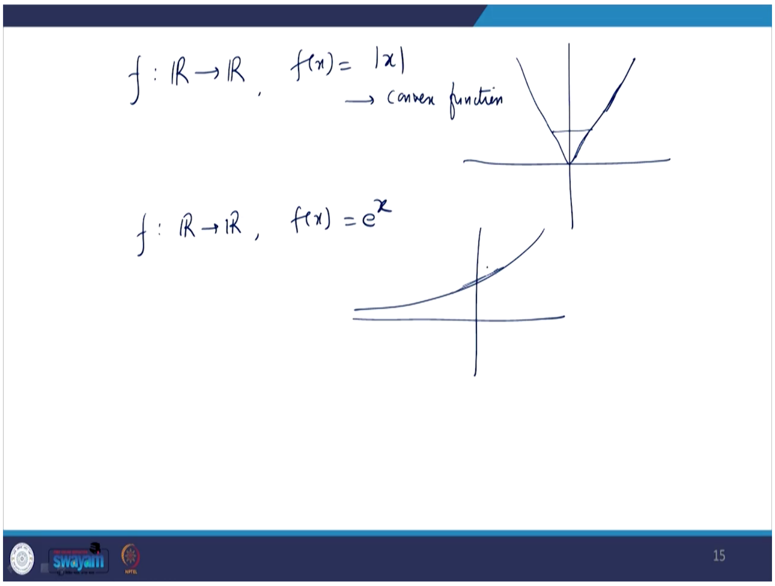
So, geometrically, how can we explain for concave function? Geometrically, we can we can say that a function is a concave function if you take any two arbitrary point on the curve join the chord, chord is always lie below or on the curve, ok.

(Refer Slide Time: 32:57)



Like we have various examples; see, we can simply draw and we can simply see that whether these functions are convex or not.

(Refer Slide Time: 33:06)

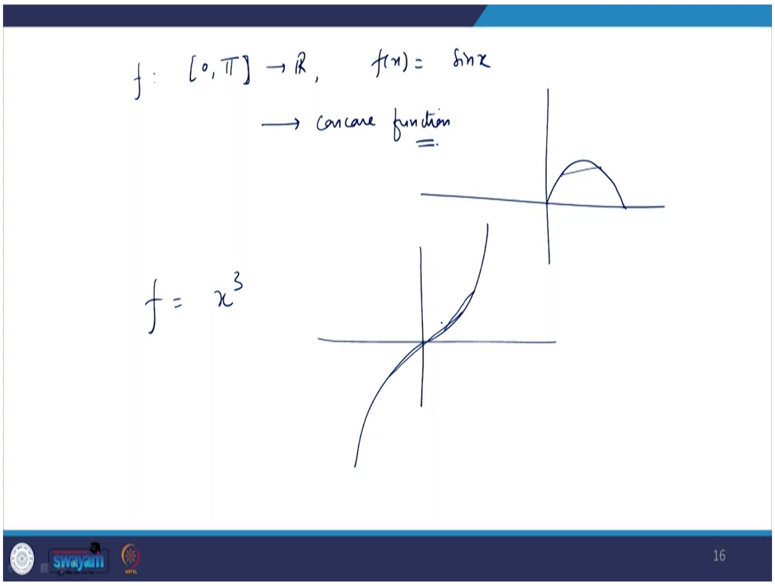


So, say we have functions from R to R and function is given by f x equals mod x, suppose mod x is given to you. So, what is mod x? How we can draw mod x? Mod x is like this. Now, this function is clearly a convex function because if you take any two arbitrary point join the chord, chord is always above the curve.

You take two points one point here one point here join the chord, chord is always above the on the curve. Here either it is above the curve or on the curve, on the curve or above the curve chord that means, the function is a convex function. So, it is a convex function.

So, you take function from R to R as e raised to power x suppose. So, if you take e raised to power x it is something like this function. So, if you take two any two point you join the chord, chord is always above the curve only it is never on the curve. So, this function is in fact, strictly convex; not only convex, so it is basically strictly convex function.

(Refer Slide Time: 34:14)

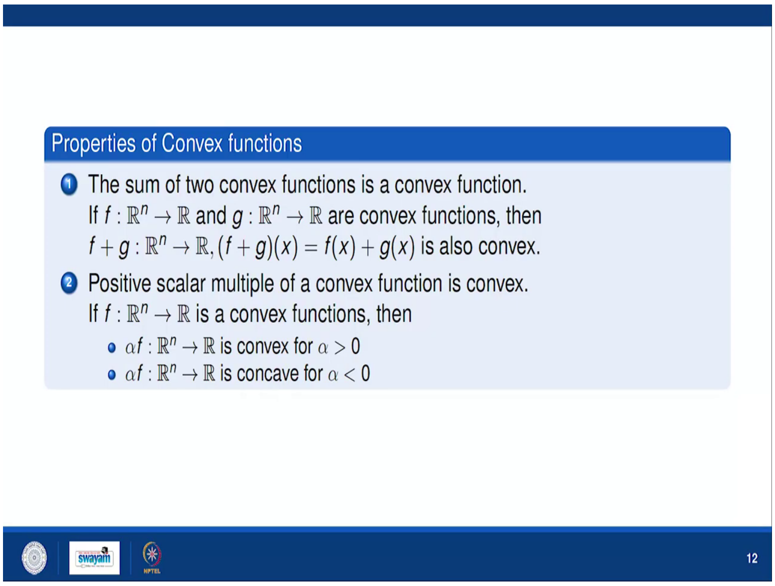


Suppose, you take a function from 0 to pi by 2, 0 to pi by 2 to R and the function is sin x. So, if you take 0 to 0 to pi suppose you take from 0 to pi, instead of pi by 2 you take it from 0 to pi. So, if you take a function 0 to pi like this sin x, so this function is what? It is a concave function because if you take any two point join the chord, chord is always below the curve. So, this function is a concave function.

So, in this way, so we if you x cube, so it is neither convex nor concave function because you can you can draw the graph of x cube; if you take x cube, so it is something like this. So, if you take two points here join the chord, chord is above the curve, but if you take one point here, one point here join the chord. Chord is sometimes below and sometimes above the curve.

So, this function is neither convex nor concave. So, in this way we can see that geometrically we can see that whether a given function is convex function, concave function, neither convex nor concave, or both convex or concave function, ok.

(Refer Slide Time: 35:35)



So, here is a very simple property of convex function which we can derive very easily using the basic definition of convex function. There are sum of two convex function is also convex, if you take f and g as a convex function, and if you sum these two function then the resultant is also a convex function. And the second property is the alpha times a convex function is convex if alpha is greater than 0. And alpha times f is a concave function for alpha less than 0.

So, one can easily prove these simple properties of convex function using that basic definition of convex functions.

So, in this lecture, we have seen that what convex sets are, how we can geometrically analyze that a given set is a convex set or not, and then we have seen convex functions, concave functions, their geometrical interpretations, various examples. In the next lecture, we will see that what are the properties of convex functions.

Thank you.