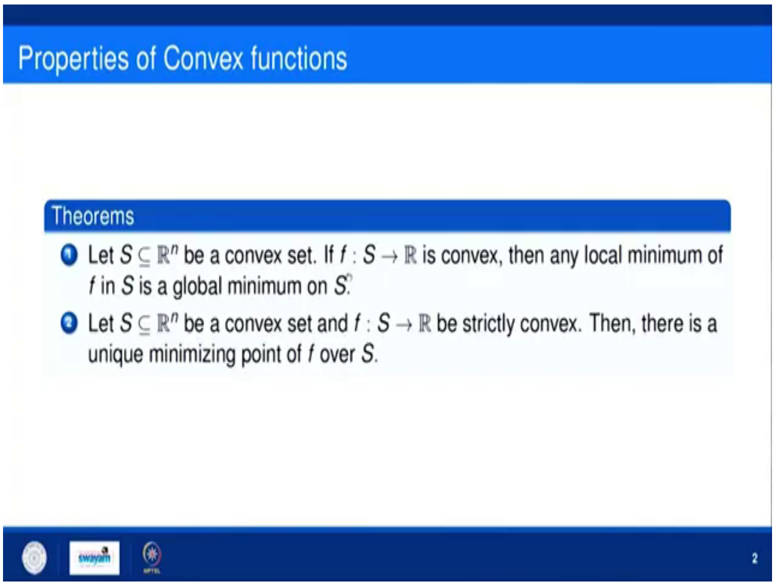
Essential Mathematics for Machine Learning
Prof. S. K. Gupta
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture – 24
Properties of Convex Functions – I

Hello friends. Welcome to lecture series on Essential Mathematics for Machine Learning. In the last lecture we have seen that what convex sets and convex functions are, their geometric interpretation also we have seen. Now, we will see some of the basic Properties of Convex Functions in this lecture.

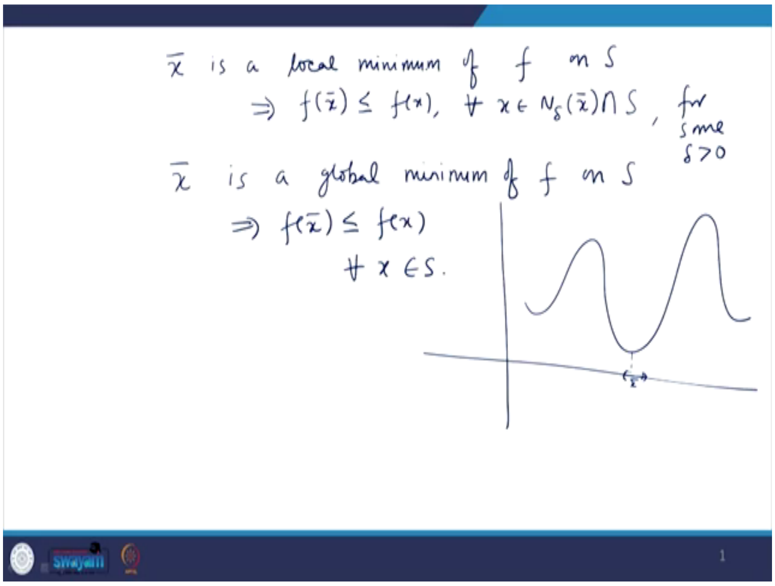
(Refer Slide Time: 00:42)



So, the first and most property of convex function is that if you have a convex function f from a convex set to R, then to take any local minimum of f in S that is a global minimum ok. So that means, if in the if in a small neighbourhood you have find a local minimum and you know that the function is a convex function, then that local minimum is nothing, but a global minimum.

This is a property of; this is a peculiar property of a convex function ok. So, let us try to see that what is the proof of this theorem. So, we have seen that what convex functions are that we have already seen.

(Refer Slide Time: 01:28)

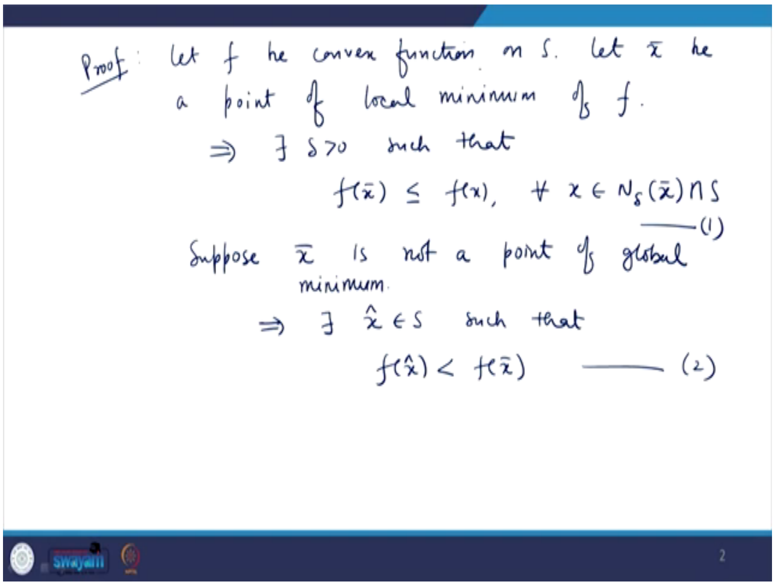


So, what do you mean by local minimum? Suppose, x bar is a point of local minimum local minimum of f on S. So, what does it mean? So, it is a local minimum, local minimum means that if you have such type of function; local minimum means if you take a small neighbourhood. In their small neighbourhood if this is x bar then the small neighbourhood, the value of this function if the function is of minimizing type; that means, the value of the function that of x bar is always less than equal to f x, for every x in this small neighbourhood.

So, this implies f x bar is always less than equal to f x for every x belongs to delta neighbourhood of x bar; that means, the small neighbourhood of x bar ok. And of course, this x is in S so, intersection with S. So, if I am saying that S y is a local minimum of f on S; that means, f x bar is always less than equal to f x for every x in delta neighbourhood of x bar intersection with S. And, if x bar we are saying it is the global minimum; what does it mean?

Global minimum of f on S; so, this means that f x bar is less than equal to f x for every x in S. So, if you are taking in entire S that this inequality is holding; that means, global minimum. And, if you are taking in a small neighbourhood for some delta greater than 0, if you are taking a small neighbourhood a intersection with S that this inequality is holding; that means, local minimum.

(Refer Slide Time: 03:33)

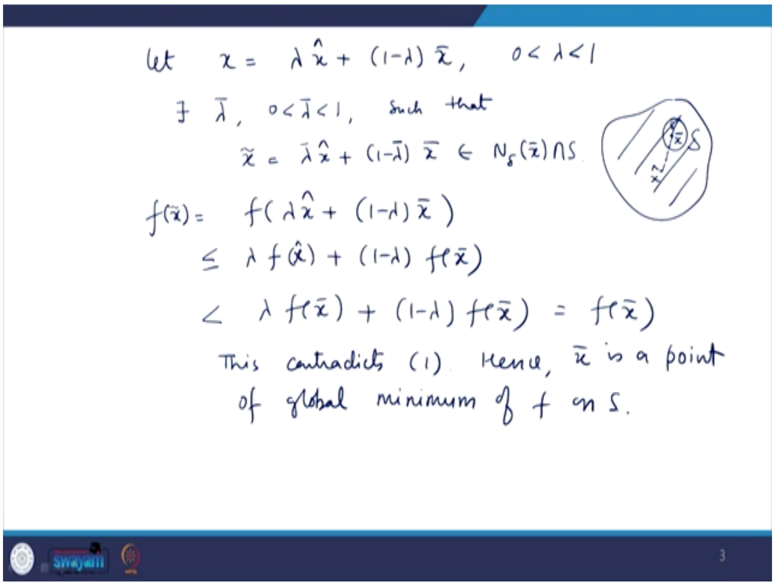


So, in that in that statement suppose x bar is a local minimum. So, let us try to prove that theorem; let f be a convex function on S. Now, now let x bar be a point of local minimum of f; if x bar is a point of local minimum; that means, this implies that there exists some delta greater than 0 such that f x bar is less than equals to f x for all x belongs to delta neighbourhood of x bar intersection with S.

So, suppose this expression is 1. Now, we have to show that x bar is nothing, but a point of global minimum. So, we will try to prove the result by a method of contradiction. So, let us suppose x bar is not a point of global minimum. So, suppose x bar is not a point of global minimum.

So, if it is not a point of global minimum; that means, there exists some point x in S where value of f x is still less than f x bar, because f x bar is not a point of global minimum. So, this implies that there exist some say x cap belongs to S such that f x cap is still less than f x bar, suppose it is 2. Now, still we have not used convex function, the definition of convex function. So, let us try to include that also in the proof, then only we can obtain the contradiction.

(Refer Slide Time: 05:47)



Let x equal to lambda x cap plus 1 minus lambda x bar where, lambda between 0 and 1. So, suppose this is region S, suppose this is S ok. Suppose this is some x cap, this is some x bar and this is a small neighbourhood of x bar for some delta this radius is delta and this is some x cap ok. So, in this in this neighbourhood, in this neighbourhood f x bar is always less than or equal to f x for every x belongs to delta neighbourhood x bar and for this x cap f x cap is less than f x bar.

Now, now this as clc this is x, clc of these two point. Now, there will always exist some lambda no matter how small lambda maybe, but there will always exist some lambda between 0 and 1 such that this clc will get a point at least 1 which such that this point belongs to delta neighbourhood of f x bar delta neighbourhood of x bar. So, we can say that there will always exist sum lambda bar, lambda bar between 0 and 1 such that such that x which is clc of these two points belongs to delta neighbourhood of x bar intersection with S ok.

Because it is x bar and it is sum x cap; so, there will always exist some lambda such that that x is x belongs to this neighbourhood. Now, let us take f x. What will be f x? f x will be f of lambda x cap plus 1 minus lambda f x bar and since function is convex. So, this is less than equals to lambda f x f x cap plus 1 minus lambda f x bar. What we are having from 2? f x cap is less than f x bar.

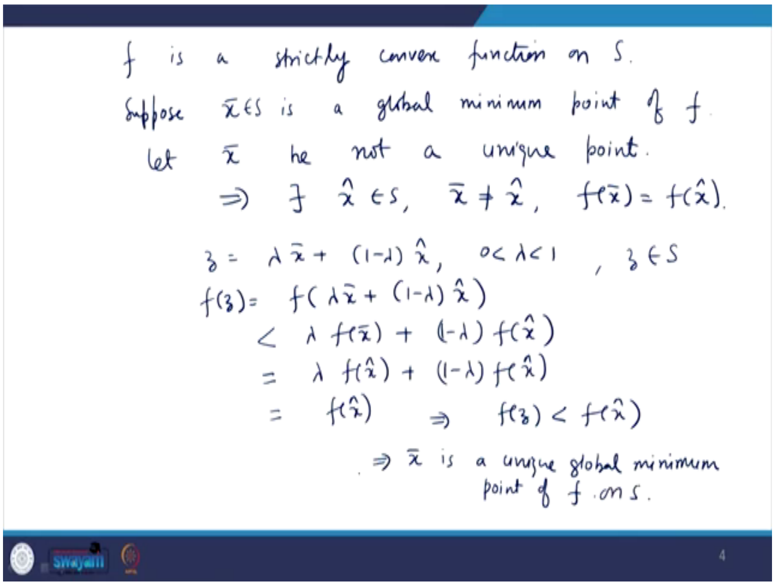
So, this is less than lambda f x bar plus 1 minus lambda f x bar which is equal to f x bar ok. So that means, now where this x is? This x is in delta neighbourhood of x bar ok, this x is in you can take this x tilde; this x tilde is in for some lambda between 0 and 1. So, some lambda bar sorry this is lambda bar, you have taken lambda bar. For some lambda bar between 0 and 1, this x tilde is in delta neighbourhood of f x bar x bar.

So, we have shown that f x tilde is less than f x bar, but from 1 f x bar is always less than equal to f x for every x in delta neighbourhood. But, we have find the x in delta neighbourhood where this inequality is reversed; that means, x bar is not a point of local minimum. So, this is this contradicts our assumption that this is not a global minimum. Hence, we can see hence we can say that x bar is a point of global minimum.

So, this implies this contradicts 1 and hence x bar is a point of global minimum of f on S. So, so we can say that if we have a convex function and we have find the local minimum then that local minimum is nothing, but a global minimum of f. The next result is that, if this function is a strictly convex then that local minimum is local minimum is global of course, but that is unique.

In this case for a normal convex function it may not be unique, but for if a function is strictly convex it is unique. So, let us try to prove this also.

(Refer Slide Time: 10:31)



So, now it is given to us that f is a strictly convex function on S and we have to show that global minimum is unique ok. Suppose, x bar is a point of global minimum, global minimum point of f x bar belongs to S ok. So, let x bar be not be not a unique point; be not a unique point ok. So; that means, this implies that there exists some x cap belongs to S, x bar not equal to x cap such that f x bar is equals to f x cap.

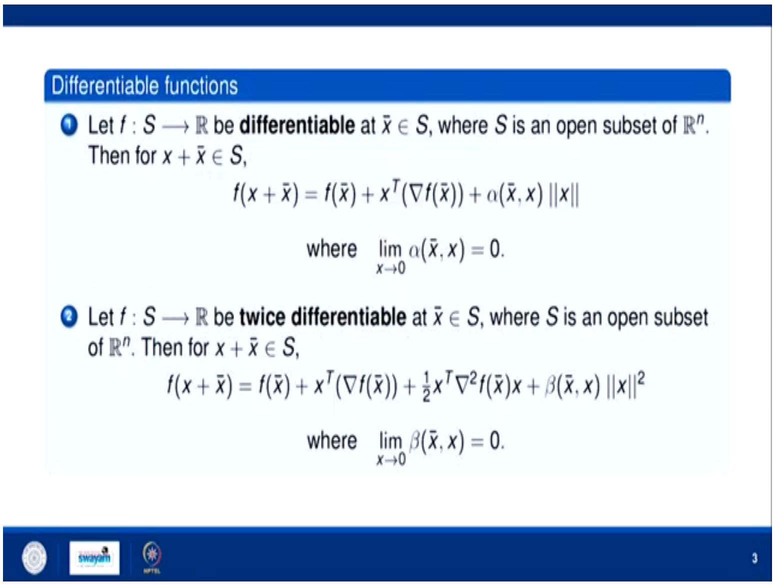
Because if it is not unique; that means, there will exist some other point also which is global minimum distinct from x bar such that f x bar is equals to f x cap. Now, you take you take say z which is lambda x bar plus 1 minus lambda x cap, lambda between 0 and 1 ok. Now, if you take f z; so, f z is nothing, but f of lambda x bar plus 1 minus lambda x cap.

So, it is equals to lambda times f x cap f x bar plus 1 minus lambda times f x cap so, but these two are equal; these two are equal; that means so, here it is function is strictly convex. So, this will be this will be strictly less than this. So, this is further equal to because f x bar is equal to f x cap. So, it is lambda times f x cap plus 1 minus lambda times f x cap. So, this after simplification we get f x cap.

So, this implies f x f z is less than f x cap. So, this contradicts that x cap is global minimum point ok; because there is a z of course, this z belongs to S because z is a convex set. So, this we have a point z on S such that f z is less than f x cap; that means, this x cap is not a point of global minimum.

So, this is a this is a contradiction and hence x bar is a unique global minimum point of f on S ok. So, we have seen that if we have a if we have a convex function then if you take any local minimum that is global and if function is strictly convex then this global minimum is unique.

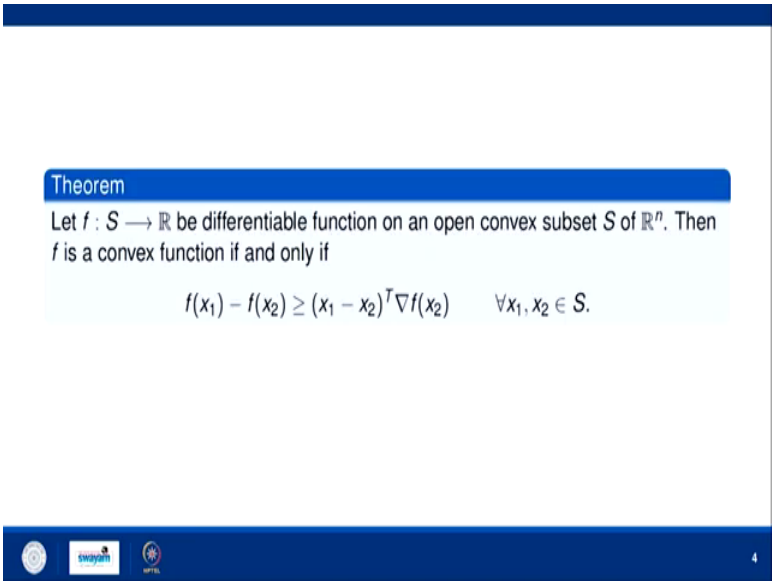
(Refer Slide Time: 14:02)



Now, let us see some other properties of convex functions. So, this we will use in this other theorems. So, you just have to recall this definition. So, if a function is differentiable at x bar belongs to S, where S is an open subset of R n. Then for any x plus x bar belongs to S we have this by Taylor's theorem that f of x plus x bar is equal to f of x bar plus x transpose gradient of f x bar plus alpha is a function of x bar and x norm of x, where this will tends to 0 as x tends to 0.

If function is once differentiable and if it is given to us twice differentiable then this definition will extend up to second derivative; I means Hessian matrix. And, then it is plus beta function of x bar and x norm of x square where this will tends to 0 as x tends to 0. So, this we will use if it is given to us that the function is once differentiable or twice differentiable accordingly in the proof.

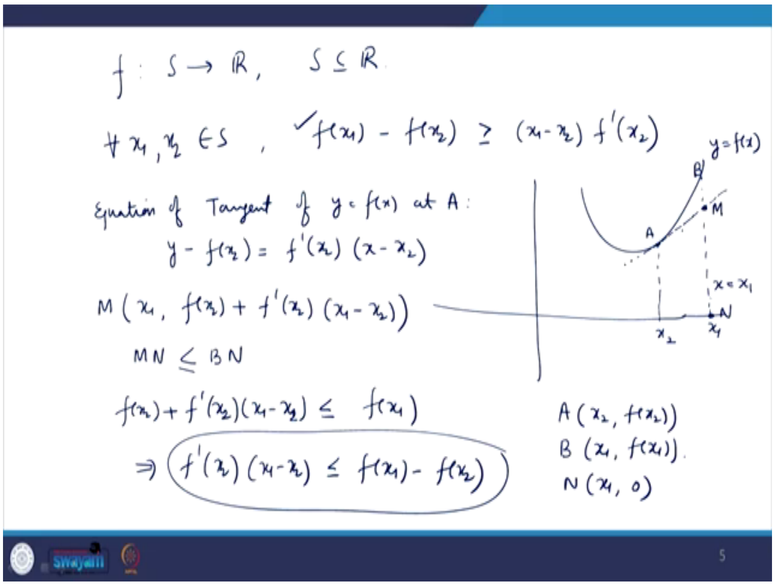
(Refer Slide Time: 15:18)



Now, we have the next result on convex function which is an important result and we will use to we will try to we will use this result in proving other properties of convex functions. But, in point of in the view of machine learning the proof of this theorem is not so important. We will try to simply see the geometric interpretation of this result ok.

So, first we will see what this theorem is basically. So, if function is once differentiable on an open convex subset S of R n. Then f is convex if and only if; that means, both side if f is convex then this holds and if this holds then f is convex, then this result hold. So, what do you mean by this? We will just understand it by geometrically.

(Refer Slide Time: 16:05)



So, suppose function is from S to R, where S is a sub set of R ok. So, this property will be nothing but for every x 1 x 2 in S. So, this is basically f of x 1 minus f of x 2 is greater than equals to x 1 minus x 2 f dash of x 2 ok. So, this is the same property if I am talking if here S is in R n subset of R n, I am taking S is subset of R to understand the geometric interpretation of this property ok.

So, now, let us let us take a function this, let us take a point say this point is say x 2, say this point is x 1. So, this is A point, suppose this is B point. So, what are coordinates of A point? A point is say this function is y equal to f x. So, coordinate of A point will be x 2 comma a of f x 2 and coordinate of B point will be x 1 comma f x 1 ok. Now, you draw a tangent at this point ok.

So, what is a equation of tangent at A; equation of tangent of y equal to f x at A? That will be y minus y 1, y 1 is f x 2 is equals to f dash x 2 x minus x 2 ok. This is a equation of this tangent. Now, what this point will be? Suppose, this point is M; now this tangent intersection of the tangent this and the line x equal to x 1, this line is x equal to x 1.

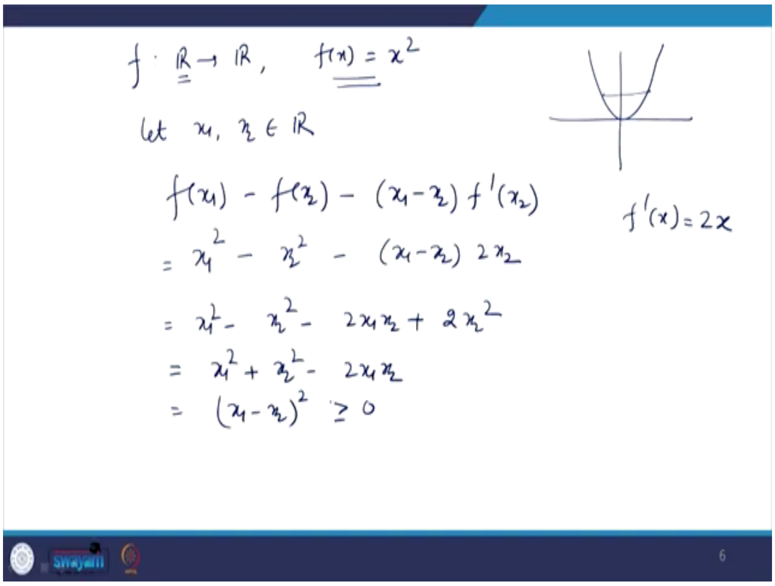
The intersection of these two line give this point M. So, the point M will be x is x 1 and what will be y for point M? For point M, you simply substitute x equal to x 1 find y. So, this is f x 2 plus f dash x 2 x 1 minus x 2 ok. So, we got this point. Now, you can clearly see that this if this point is suppose N, this point if suppose this point is N. So, where N point is x 1 comma 0 so, you can see that MN, MN is less than BN.

And, if it is a straight line, it may hold as an equation also ok. So, what is MN? MN is basically MN means y coordinate of M point this height; that means, y coordinate of this point this height. So, y coordinate this point is what? f x 2 plus f x f dash x 1 f dash x 2 x 1 minus x 2 is less than equal to BN, BN means y coordinate of B point; y coordinate of B point is what? f x 1.

So, this implies f dash x 2 x 1 minus x 2 is less than equals to f x 1 minus f x 2; that means, the same; that means, the same point which we are having here. So, so what we can say geometrically? So, geometrically we can say that that if you draw a tangent at any point on a convex function, then tangent is always below the curve or on the curve if it is a straight line. Here it is, this function is strictly convex.

So, that is why this is always lie below the curve. But, if it is a function like boat shape function then it may lie on the curve also. So; that means, if you are having a convex function, if you draw a tangent at any point; the tangent will lie either below the curve or on the curve always. So, that is another important interesting property of convex functions.

(Refer Slide Time: 20:50)



So, suppose function is given to say function is from R to R and function is defined as say f x equal to x square. You want to show that this function is convex ok. I mathematically, geometrically you know; geometrically you know that this function is a convex function. Because if you draw any two point join the chord, chord is always above the curve.

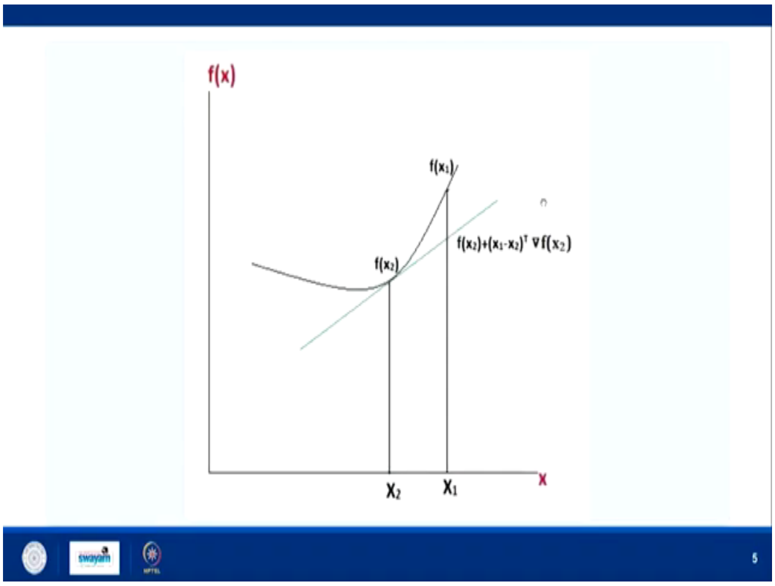
So, this function is always convex; I mean is a convex function. But if you want to show mathematically; so, you can use this property. What this property is? This property means for if you take let x 1 x 2 are any point in S in R, here S is R ok; then by then by this property you have to show this inequality holds. So, you can bring this expression on the left hand side and try to show that the entire expression is greater than equal to 0, for any x 1 x 2.

So, let us let us try to see, you take f x 1 minus f x 2 minus x 1 minus x 2 f dash x 2. And what we have to show? We have to show that this expression is this expression is greater than equal to 0; if you have shown this then by this theorem we can say that the given function is a convex function. So, so let us try to see what is x 1, what is f x 1?

f x 1 is x 1 square by this definition, x 2 is x 2 square minus x 1 minus x 2. What is f dash x 1, f dash x? f dash x is 2 x; so, it is 2 x 2. So, this is further equal to x 1 square minus x 2 square minus 2 x 1 x 2 minus minus plus 2 x 2 square. So, this is equals to x 1 square plus x 2 square minus 2 x 1 x 2. So, this is equal to x 1 minus x 2 whole square and this is always greater than equal to 0 for all for any x 1 x 2 in R.

So, in this way we have shown analytically also that the given function is a convex function. Similarly, if you take other examples say e raise to power x; so, using the same concept, using the same definition you can show easily show that the given function is a convex function ok.

(Refer Slide Time: 23:15)



So, geometrically we have already discussed this thing. So, in this lecture we have seen that the two important properties of convex function. Number 1, that every local minimum is global minimum and the second property is the if a function is once differentiable and a function is convex it is given to you, then then this inequality then then this inequality hold ok.

In the next lecture, we will see few more properties of convex functions.

Thank you.