INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NPTEL

NPTEL ONLINE CERTIFICATION COURES

Nonlinear Programming

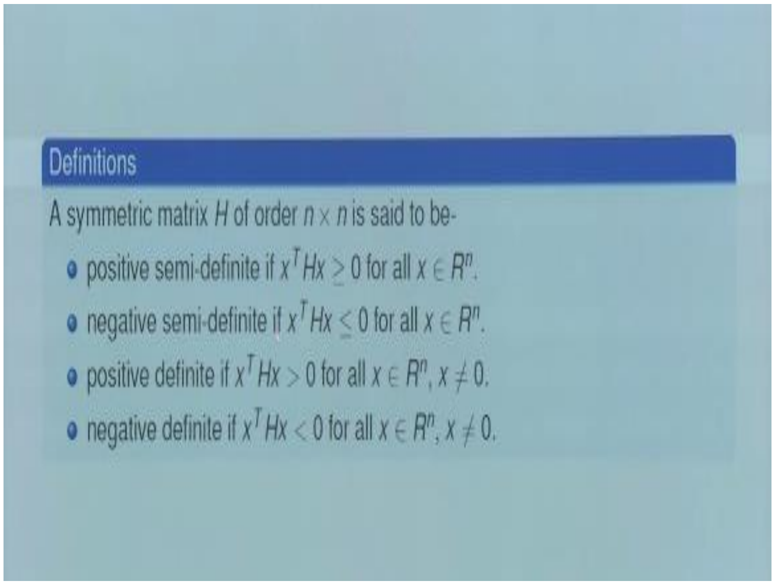
Lec – 04
Properties of Convex Functions-III

Dr. S. K. Gupta
Department of Mathematics
Indian Institute of Technology Roorkee

So welcome to lecture theories on Non linear programming, we have seen that what convex functions are and also some of their properties, now we will see some more properties of convex functions. So in our last classes we have seen that we can prove our function to be convex functions simple by definition of the convex function, or we can use the definition of epigraph because a function is a convex function if and only if its epigraph is a convex set or if function is once differentiable on a open convex subset S(r) then f is a convex if and only if $f(x1) – f(x2) \geq x1 – x2^{T}$ gradient of fx2 okay.

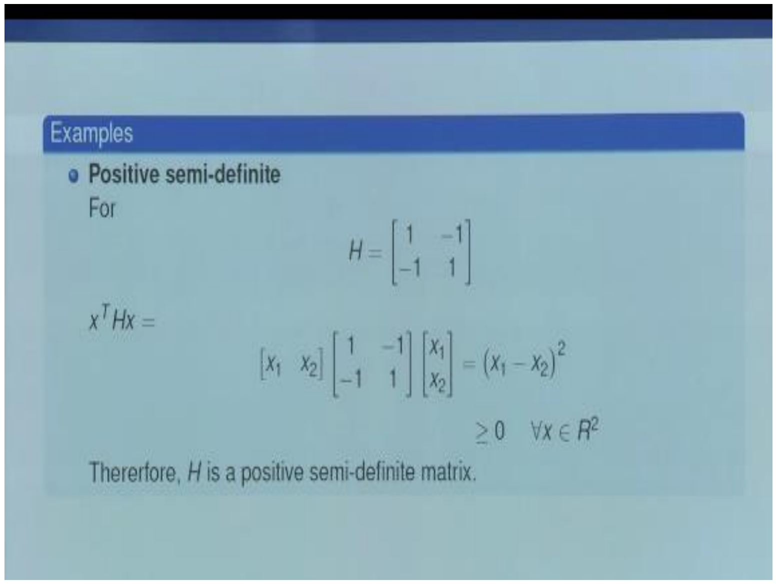
So these three properties we have defined for the convex function, now we will see that if F function is twice differentiable then what is the property for convex function how we will see F function is convex or not okay. So before studying twice differentiability of a convex function we will see that definiteness of a symmetric matrix H so a symmetry matrix H of order n x n is set to be positive have a definite if $x^{T} H(x) \geq 0$ for all x in R^{n} negative semi definite.

(Refer Slide Time: 01:43)



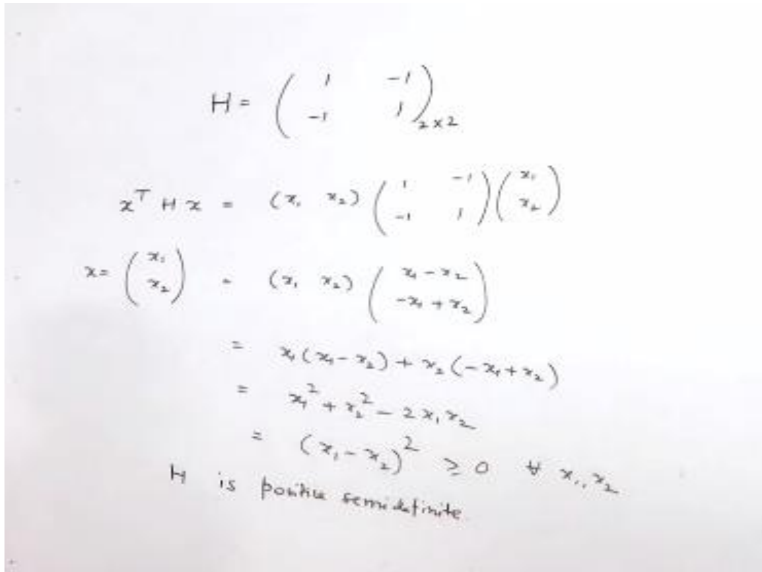
If $x^T Hx \leq 0$ for all x in R^n, positive definite if this inequality hold as strict inequality for x in R^n, x should not equal to 0 and it is negative definite if this inequality for negative semi definite hold as a strict inequality x should not be for all X in R^n and x should not equal to 0. So what these are let us discuss this by few examples.

(Refer Slide Time: 02:19)



Now suppose we have this matrix H.

(Refer Slide Time: 02:24)

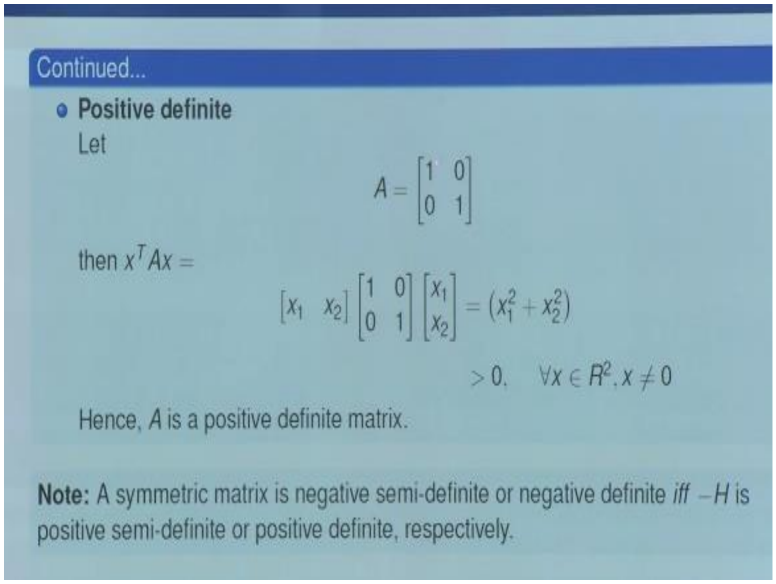


What is matrix H is, this is 1, -1, -1, 1 they are symmetric matrix and this I have in order to cross 2, now let us find x^T H(x) since this is of order 2 x 2 so x^T will be x1, x_2 it is 1, -1, -1, 1 and x is x_1, x_2, okay we are taking x as x1, x_2. So x^T will be x1, x2 as a row vector H is this matrix and X is a column vector now you multiply these3 so what we obtain, when you multiply these 3 so it is $x_1 - x_2$ it is $-x_1 + x_2$ and this is equals to x1($x_1 - x_2$) + x_2($-x_1$ + x2) = $x_1^2 + x_2^2 - 2x_1x_2$ and this is $(x_1 - x_2)^2$.

And of course this quantity is ≥ 0 for all x_1, x_2 okay so we say that this matrix H is positive semi definite, okay. This matrix excess positive semi definite since there exits some non zero X also where it is equal to 0 like 1, 1 = 0. So you cannot call it positive definite because a positive definite as you see in the definition positive definite this must be greater than 0 for all x in R^n and x should not equal to 0, okay.

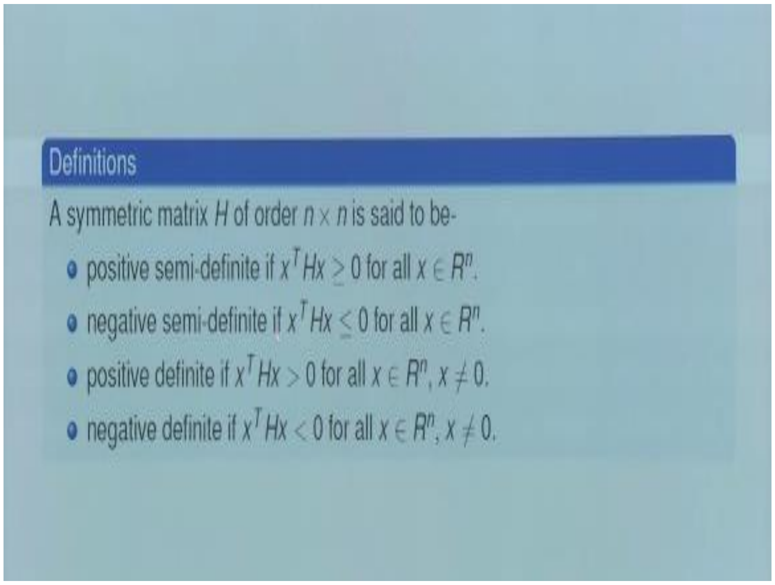
This must be positive strictly greater than 0, however this is ≥ 0 and there are non zero x where it holds and equal to 0, so therefore it is not positive definite however it is positive semi definite okay.

(Refer Slide Time: 04:55)



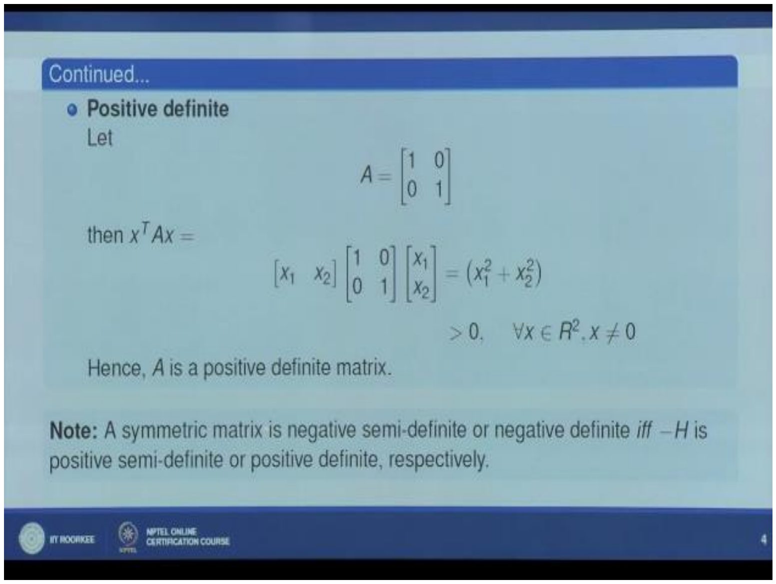
Now we take another example you take identity matrix of order 2 x 2, I(1, 0, 0, 1) and when you take x^T Ax so it is nothing but $x_1^2 + x_2^2$ you can simplify this it is strictly return 0 for every x^n or 2 because it is 0 only when x_1, x_2 both are 0 and x should not equal to 0 so that is not here so that means it is strictly return 0 so therefore this matrix are positive definite matrix. Now from these definitions.

(Refer Slide Time: 05:32)



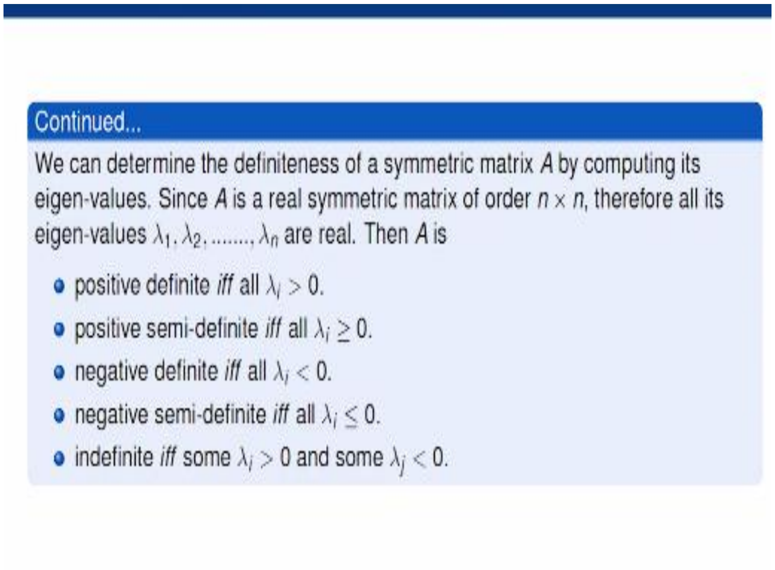
You can easily see that if a matrix H is positive semi definite then –(H) will be negative semi definite because when you multiply with -1 with both the sides then –(H) will be definitely negative semi definite and similarly if a matrix H is positive definite so –(H) will be negative definite because this inequality hold.

(Refer Slide Time: 06:01)



So that is there in this note symmetry matrix is negative semi definite or negative definite if and only if – (H) is positive semi definite or positive definite respectively, okay.

(Refer Slide Time: 06:13)



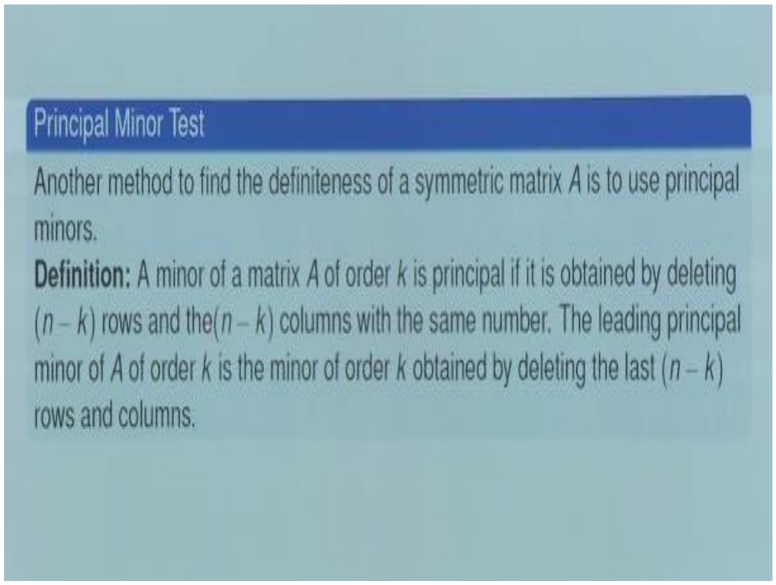
Now how can we check whether the given matrix is positive definite positive semi definite negative definite or negative semi definite so we have two tests first is Eigen value test okay now since the given matrix A is real and symmetric matrix so we know that if a matrix are real and symmetric matrix then all its Eigen values are real so it is of order n x n this means it will be having n number of Eigen values λ_1, λ_2 and λ_n and all will be real.

Then A is positive definite if and only if all it Eigen values are strictly kept in 0 okay, you find all the Eigen values of the real symmetric matrix A if the Eigen value all the Eigen values are strictly kept in 0 the we say that it is positive definite, positive semi definite if and only if all Eigen values are ≥ 0 negative definite if and only if all Eigen values are strictly less than 0, negative semi definite if and only if all λ I are ≤ 0.

And indefinite okay if and only if some λ_I >0 and some λ_g <0, now finding an Eigen value always is not a simple task if we have a matrix of order of 4 x 4 or 5 x 5 then always finding Eigen values of a matrix to check the definiteness of the symmetric matrix, matrix A is not an easy task, so we have another test also which we call as principle minor test, now what it is? First we understand what principle minor is then we go to a test.

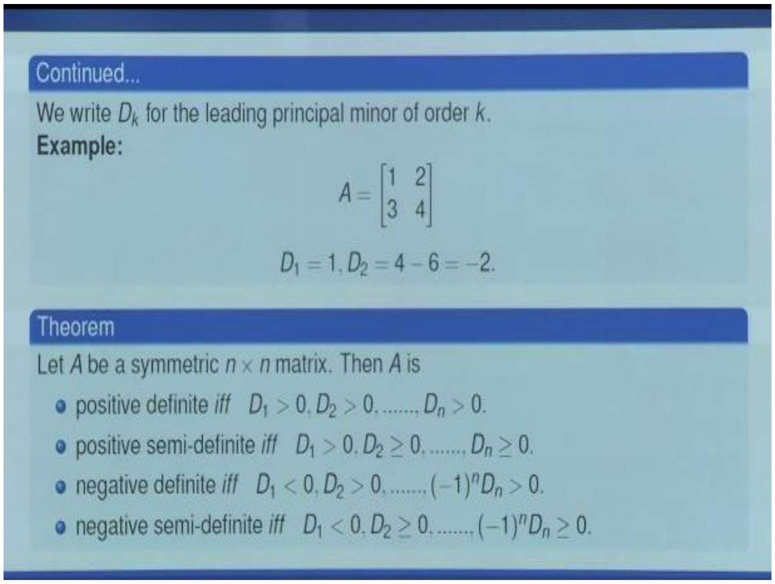
That how using this test we can check we can find whether a matrix A is positive or negative definite okay. So a minor of matrix A of order K is principle we which you call as principle minor if it is obtained by deleting (n-k) rows.

(Refer Slide Time: 08:28)



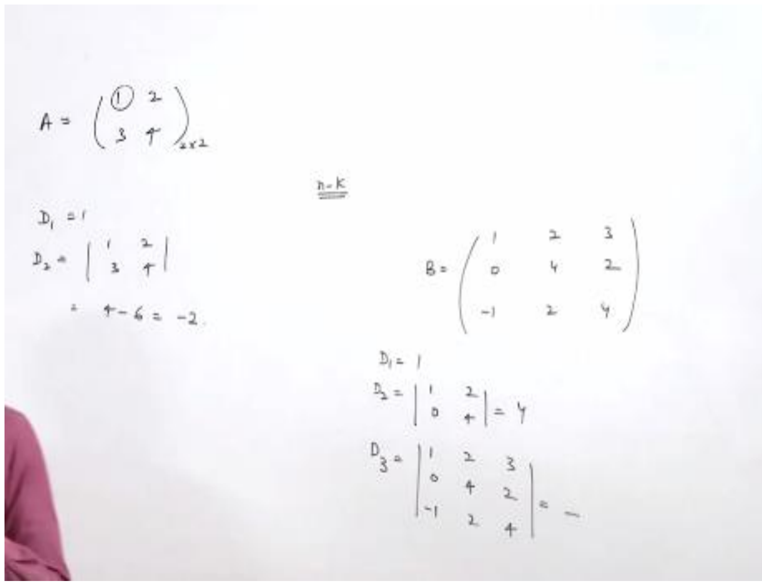
And (n-k) columns you delete n-k rows and n-k columns from n x n matrix you got a matrix of order k x k that determinant of that matrix called principle minor matrix, okay. The leading principle minor (A) of order K is the minor of order K of n by deleting the last and –k rows and columns we will understand this by an example okay what this means.

(Refer Slide Time: 09:02)



Let us understand this by an example.

(Refer Slide Time: 09:06)



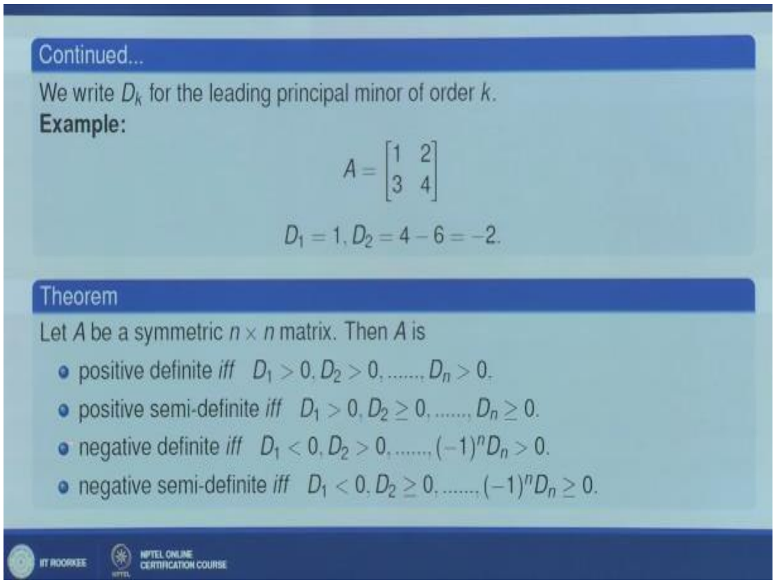
Suppose you have matrix given here A = 1, 2, 3, 4 if you delete it is 2 x 2 okay and you have to delete n-k rows and columns from the last okay now n is 2, suppose k is in this case is either 1 or 2 okay suppose k is 1, 2-1 is 1 so you delete last one row and one column you delete last one row and one column so you will left with D_1, D_1 means precise minor of order 1 which is 1 which is this element.

Now when you take D_2, D_2 is simply determinant of n – you delete you do not delete any rows and columns 2-2 is 0 so that means the full matrix 2, 3, 3 ,4 that is nothing but 4 – 6 which is -2 so this is how we can find the determinants of leading principle minors okay. So here decay represents leading principle minor of order k okay. Suppose you have a matrix B of order 3 x 3 you can take any matrix 1, 2, 3, 0, 4 , 2, -1, 2, 3, 2, 4.

Suppose you have this matrix, now suppose you have this matrix B so what D_1 means, D_1 means you start from the first element and take some determinant of order 1 x 1 okay so that is one only, now D_2 is simply you start with this element take the matrix of order 2 x 2 and take the determinate so that is 1, 2, 0, 4 which is nothing but 4 and D3 is nothing but determinant of 3 x 3 matrix which is itself that is 1, 2, 3, 0, 4, 2, -1, 2, 4.

So this we can find this will be D_3 so that is how we can find out the principle the leading principle minors okay. Now finding how then finding leading principle minors how can we check whether a matrix is positive definite negative definite or semi definite, so let us see the result here let A be a symmetric n x n matrix.

(Refer Slide Time: 11:50)



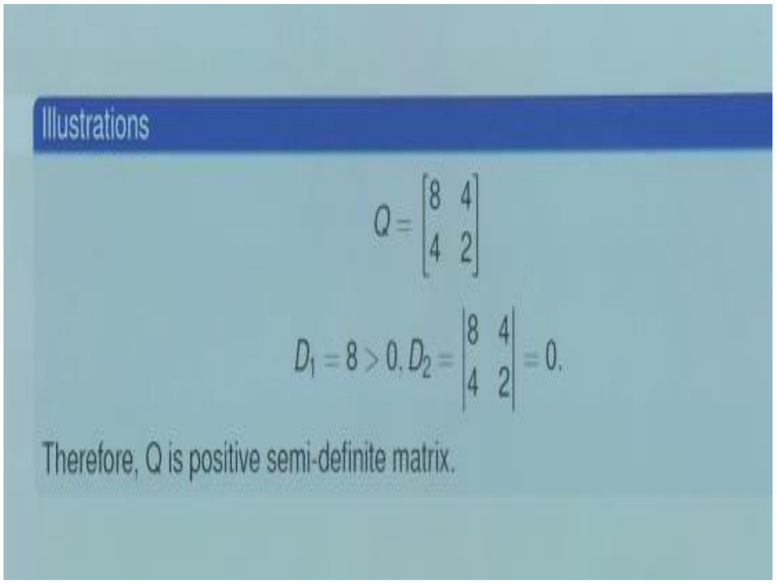
Then A is positive definite if all the leading principle minors all the matrix A are strictly greater than 0 okay. It is positive semi definite if and only if D_1 is strictly greater than 0 and all other principle minors are ≥ 0 it is negative definite if and only if $D_1 < 0$, $D_2 > 0$ that means $-1^n D_n$ is strictly greater than 0 and negative semi definite if and only if $D_1 < 0$, $D_2 \geq 0$ and so on such that this quantity is greater than equal to 0, okay.

We have the alternate signs in negative definite because of matrix if a matrix A is negative definite that means negative of A is positive definite if a matrix A is negative definite this means negative A is positive definite and what is the determinant of negative of A that is nothing but.

-1^n determinant of A that we already know by the properties of determinant okay, now D_1 is of order 1 x 1, so it will be negative because we substitute n = 1 here and D_2 is of order 2 x 2.

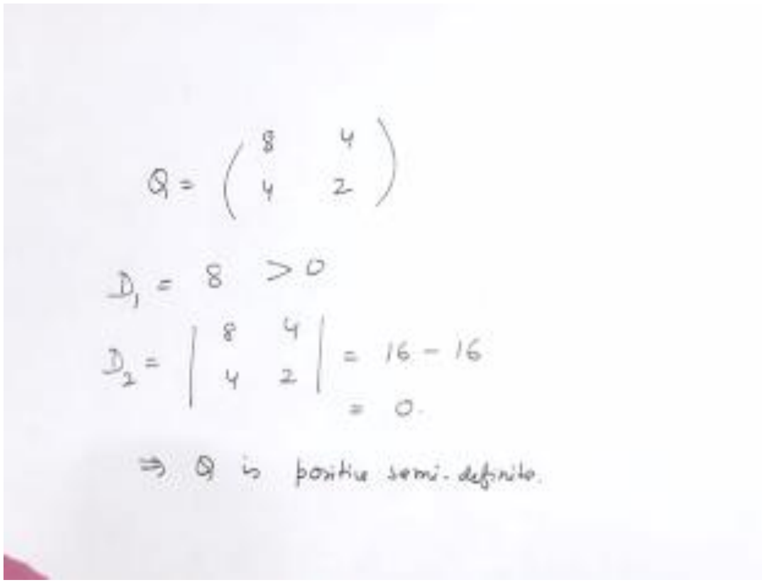
So it will remain positive, okay. And D_3 is of order 3 x 3 and when you substitute n = 3 so it will become negative, so from here we can conclude that if we have positive definite that is all plus leading plus or minus are strictly greater than 0. Than for negative definite we have an alternate signs, okay.

(Refer Slide Time: 13:52)



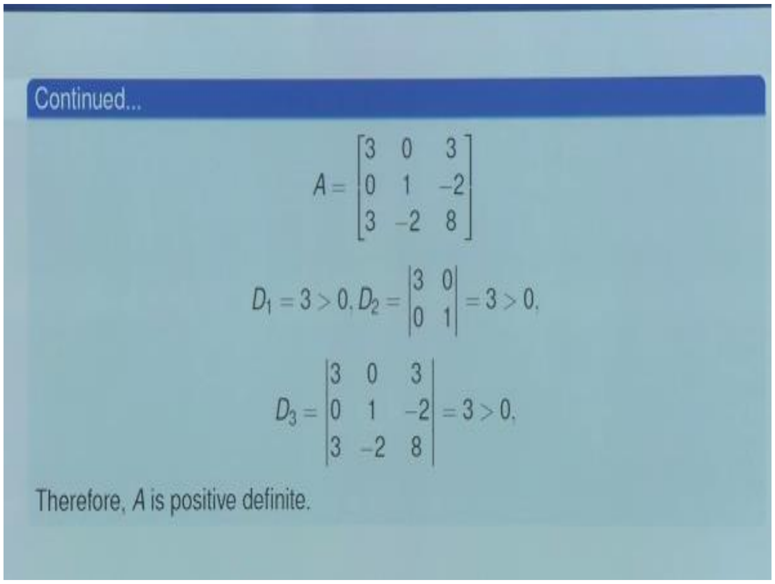
Now let us try let us see some examples based on this, we take the definiteness of a matrix A now suppose Q is a matrix and Q is given by.

(Refer Slide Time: 14:02)



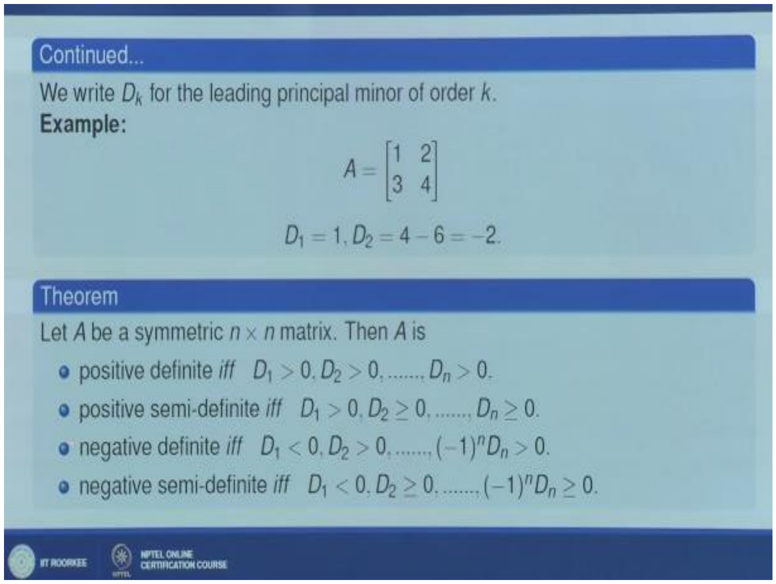
Is 8, 4, 4, 2 so what will be the D_1 here, D_1 is 8, 1 x 1 where is strictly greater than 0, what s D_2? D_2 is determinant of 8, 4, 4, 2 which is 16 – 16 which is 0 so D_1 is strictly greater than 0 and D_2 = 0 that means by this test what we can say by this test we can say that it is positive semi definite because D_1 is strictly greater than 0 and $D_2 \geq 0$ so we can say is a matrix q is this implies matrix Q is positive semi definite, okay. Now we take another example suppose we have matrix A here okay.

(Refer Slide Time: 14:53)



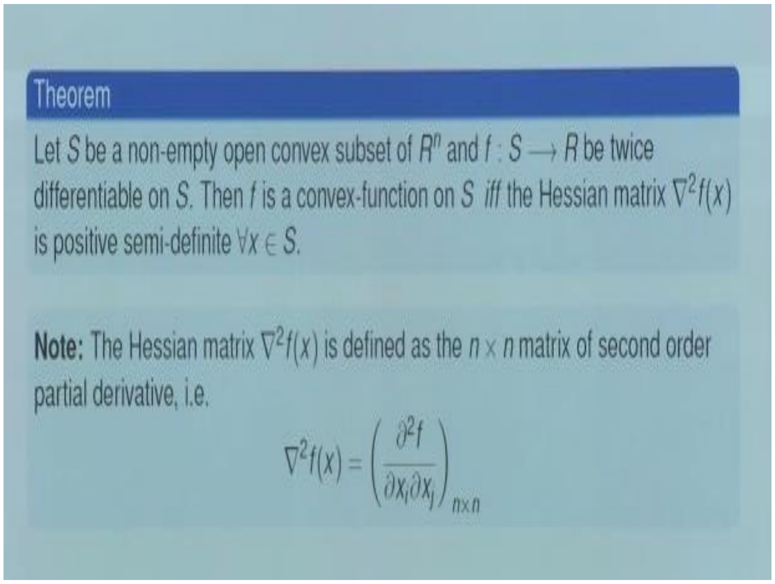
Now what is matrix A, you can easily see the matrix A is 3, 0, 3, 0, 1, -2, 3, -2, 8 now when you take D_1, D_1 is the leading principle minor of order 1 x 1 which is 3 and it is strictly greater than 0 D_2 is leading principle minor of order 2 x 2 so that is determinant of 3 is 0, 0, 1 which is 3 only and greater than 0, now D_3 is determinant of full matrix A which is 3 again and it is strictly greater than 0. All the leading principle – R is strictly greater than 0 this means matrix A is positive definite, okay because.

(Refer Slide Time: 15:43)



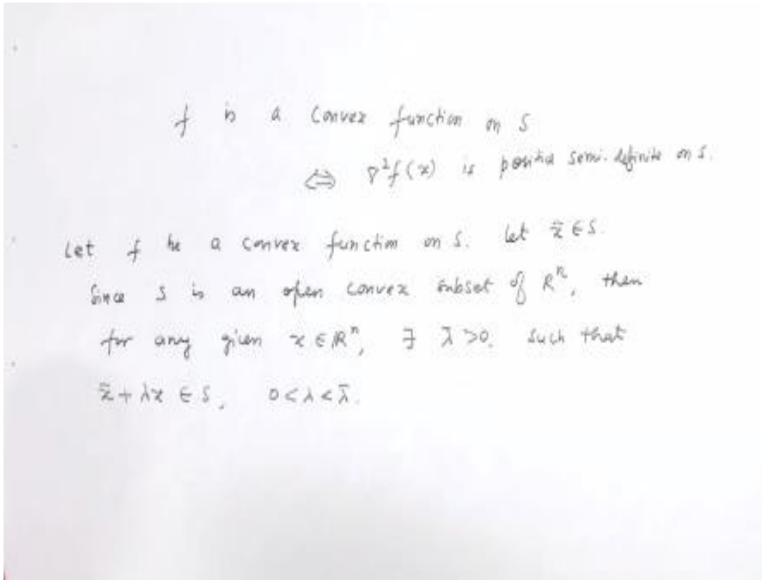
Of this listen it positive definite if and only if all principle reading ± S strictly greater than 0.

(Refer Slide Time: 15:52)



Now we have the result for the convex function okay we understand what causes semi definite or positive definite means, now we can state a result for a convex function. Let S be a non empty open convex subset of R^n and S is a function from S → R be twice differentiable on S, okay. Then F is a convex function on S if and only if the Hessian matrix which is gradient square f(x) is positive semi definite for all F belongs to S, so this is Hessian matrix we have already defined so F is a convex function.

(Refer Slide Time: 16:34)



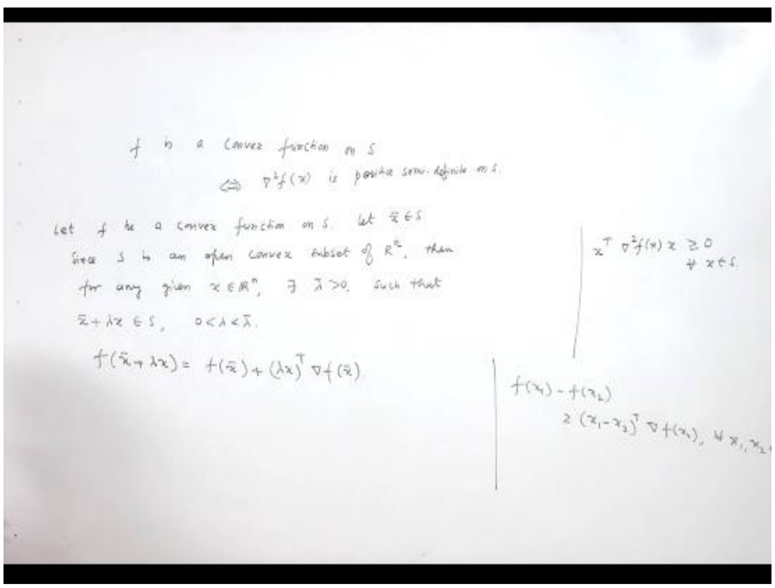
On S if and only if hessian matrix of F on S is positive semi definite on S so now they have to show that if F is a convex function then this is an hold and if this is set old then it is a convex function, so let us start to prove this result. So let F be a convex function on S so we have to prove that wave function is twice differentiable, now the hessian matrix is positive semi definite that means we have to show that for any x, x^T gradient square of f(x) x > 0 for all x in s, this we have to prove, okay.

Because we have to show that this is a positive semi definite okay. So let x1 belongs to s now since s is an open convex subset of R^n if then for any given x in R^n their exist some λ bar return 0 such that x bar + λ x will definitely belongs to s for λ between λ bar and 0 you see that s is an open convex subset of R^n okay. And x bar is a some fix point in s okay, now for if this is an open convex subset of R^n so for any given x you take any arbitrary point x in R^n there will always exist some λ bar.

No matter how it is small it may be but there will always with some λ bar such that this point which is x bar + λ x will definitely belongs to s for λ between 0 and λ bar because S is a open set, okay.

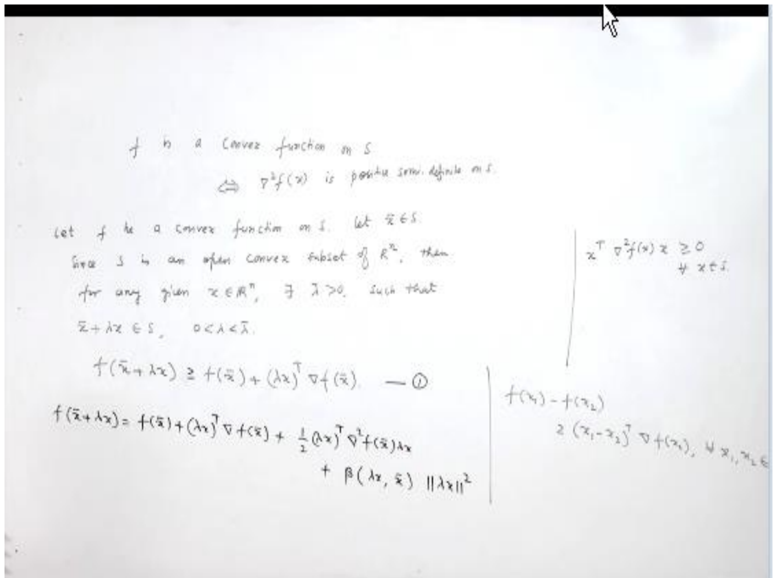
Then for the small value of the λ this x bar + λ x will belongs to S. Now sense the f is a convex function so below that $f(\bar{x}+ \lambda f(\bar{x})= f(\bar{x}+(\lambda x)T \Delta f(\bar{x})$. because if function is convex this means $f(x_1)-f(x_2)>=(x_1-X_2)^T \nabla +(x2)$ for all x_1,x_2 in S. this way already seen that is f is a convex function then $(x_1)-f(x2)>=(x_1-X_2)^T \nabla +(x_2)$ will satisfy the property all x_1, x_2. Now suppose this x1 and this is x_2 this belongs to S and this is belongs to S.

(Refer Slide Time: 20:04)



So these will volt this property will volt this also because f is a convex function. So f of this equal to $f(\bar{x})$ so this x_1 and this is x_2 plus x_1-x_2 which is λ x $\nabla f(x_2)$ okay. Then this is not equal this is a grater equal to okay f is a convex function. Now since f is a twice differentiable for all points in S. So we can apply a definition of twice differentiability. So by twice differentiability what we obtain f $(\bar{x}+\lambda x)$ will be equal to $f(\bar{x}+(\lambda x)^T+ ½(\lambda x)T\nabla^2 f(\bar{x}) \lambda x + \beta (\lambda x, \bar{x}) || \lambda x ||^2$.

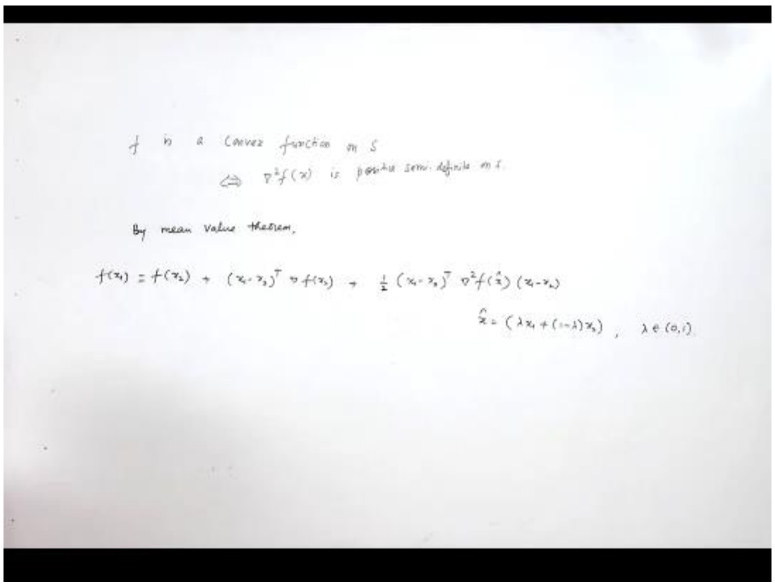
(Refer Slide Time: 21:50



This is the definition of twice differentiability of f okay now when you take this quantity when you apply this condition in this expression, expression 2. So what we obtain now we are this telling to 0 as λ =10 to 0. Okay this is by the definition of twice differentiability of f and x part okay. So now you applied the definition- 1 condition-1 on this expression. So what we obtain so this implies from 1 and 2; that 1/2 λ^2 x_T ∇^2 $f(\bar{x})x$+ λ^2 ||x2>= 0.because this quantity when you take on the left hand side is grater or equal to 0 okay.

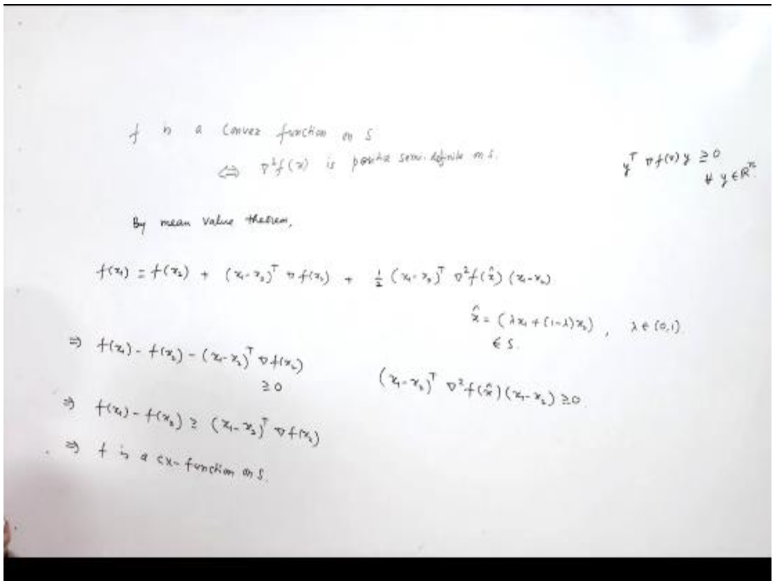
Now you can divide by λ2 both the sides what will obtain we obtain 1/2x $_T$ ∇^2 $f(\bar{x})x$ +$\beta(\lambda$ x, $\bar{x})$ λ^2 ||x2|| >= 0 okay. Now take $\lambda \rightarrow 0$, so this quantity $\beta(\lambda$ x, $\bar{x}) \rightarrow 0$? So this implies $\bar{x}^T \nabla^2 f(x)x$ ½ will return equal to 0. Because when you take $\lambda \rightarrow$10 to 0 both the side this will taking to 0 okay. And this is 0, so this quantity will return equal to 0. So this implies x^T ∇^2 $f(\bar{x})$ x >= 0 and this ➔ $\nabla^2 f(x)$ is positive semi definite on S. if it is positive at $\bar{x}$€S you can vary $\bar{x}$ so we can get that this should met the positive semi definite on S. that is for every x in S. so this is how we can obtain that if f is a convex function. Then it we should matrixes is positive semi definite okay. Now we will try to prove the convex part of this theorem, convex part means assume that the hessian matrixes is positive semi definite on S. and we will try to obtain that the function is a convex function. So by the mean value theorem, we have to plan mean value theorem function is twice differentiable. So it is $f(x_1)-f(x_2) = f(x_2)+(x_1-x_2)^T \nabla f(x_2)+1/2(x_1-x_2)T \nabla 2f(x2)(x_1-x_{2})$ you can take x^ here x^ is nothing but convex the combination of $(\lambda x1 + (1-\lambda)x2)$, $\lambda€(0,1)$.

(Refer Slide Time: 25:52)



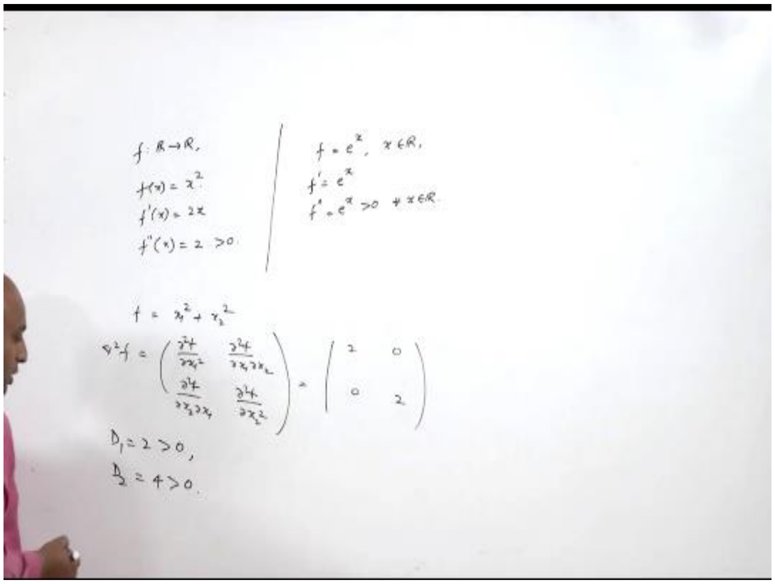
Is where the mean value theorem function f, the function twice differentiable. Now it is given to us that hessian matrixes positive semi definite. This means $y^T \nabla f(x)y>=0$ ¥ every y€Rn. This means this x cap is in R because x is a convex set. And this is an any point in Rn. So since this is an hessian matrixes positive semi definite so this quantity $(x_1-x_2)^T \nabla e\ f(x^\wedge)(x_1-x_2)>=0$. Because this is some point in R^n and this is in S. And hessian matrixes positive semi definite on S. so this quantity will be getting equal to 0. Now this quantity return equal to 0 means ➔ $f(x_1)-f(x_2)-(x_1$ $x_2^T \nabla f(x_2) >=0$. So this ➔ $f(x_1)-f(x_2)>= (x_1-x_2)^T \nabla f(x_2)$ so this ➔ f is a convex x function on S.

(Refer Slide Time: 27:59)



So that is how we can show. That f is convex if and only if is a hessian method -positive semi definite on S okay. We have various examples, now to illustrate to see. Suppose function is from $R \rightarrow R$ and $f(x) = x^2$. Now function is from $R \rightarrow R$ what the hessian matrix of this function is. Hessian matrix a simply second irrigative S. so the first irrigative is $f(x) = 2x$, and second irrigative is $f(x)=2>0$.that means positive semi definite. And that means function is convex okay. Now similarly you take f is equal to suppose e^x , x€ R, you take the first irrigative $f^1 = e^x$ second irrigative again to $e^x > 0$ ¥ x € R. for that means the function is convex function okay.

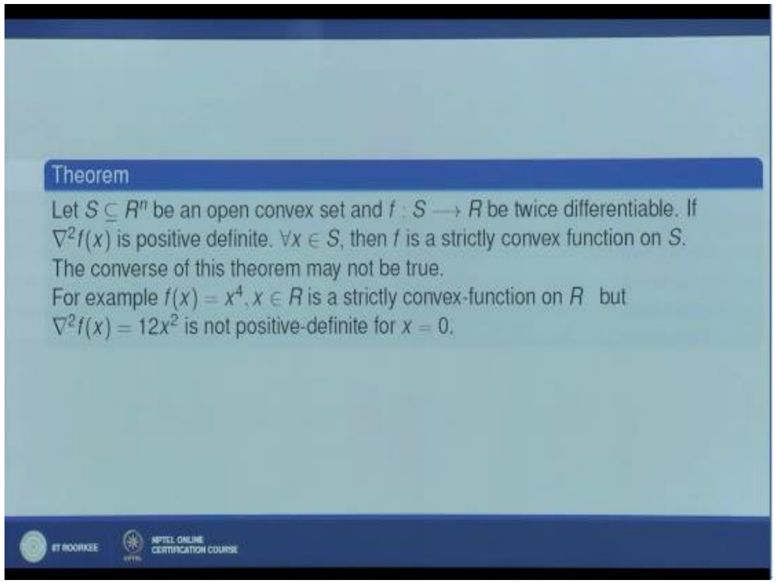
(Refer Slide Time: 30:39)



In fact these are strictly convex, because if hessian matrixes is strictly I mean hessian matrixes is positive definite. Then function is strictly convex okay, because if a function is hessian matrixes is strictly positive definite. Then the function is strictly convex. We have a result it is okay. Similarly suppose you have $f=x_1^2+x_2^2$. so you should find hessian matrixes function.

Hessian matrixes nothing but del^2 f/del x_1^2 del ^{2} f upon del x_1 x_2 del^2 f upon del x_2 delx_1 del^2 f upon del x_2^2 and that is nothing but when you take the irrigative the f(x_1) it is 2 0 0 2 so you should check whether hessian matrixes is positive semi definite or not we will find leading will to minus D1 is 2 which is returns 0, and D2 is 4which is again returns 0. So that means if is a positive definite that means function is technique convex okay. So we have the same results for strictly convex function also.

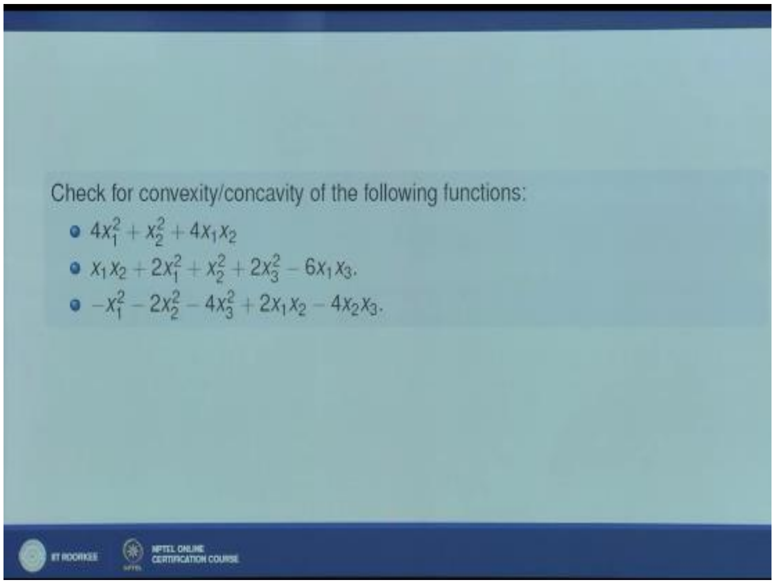
(Refer Slide Time: 30:51)



That if a hessian matrix is definite then f is strictly convex on S. but the convex may not be true. Like in this theorem we have if and only if condition. That is a function in a convex function if an only if. The hessian matrixes is positive semi definite for all x in S. but for strictly convex we don't have a if and only if condition. We are only the condition that if hessian matrixes is a positive definite. Then f is strictly convex.

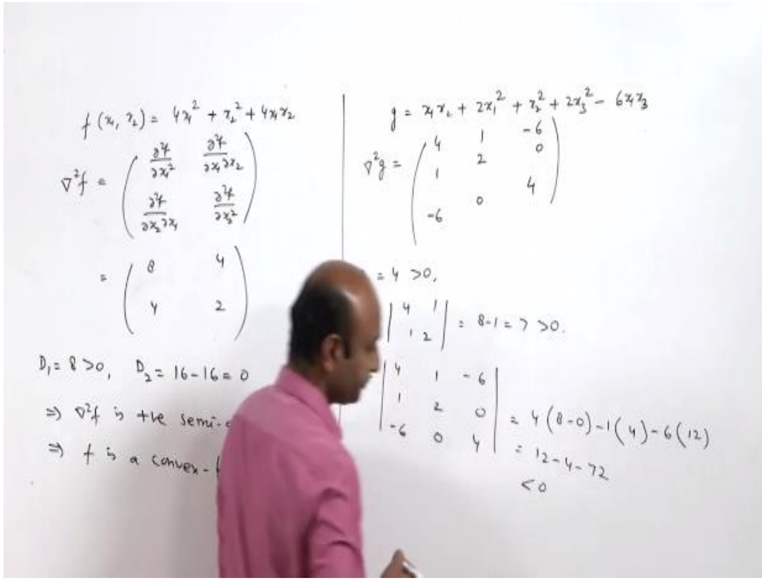
So why the convex is not true we have a convex example like ax has to power 4. When you take this is graphically you can see that is the function strictly convex okay. But when you take the second irrigative of this function. This is 12 x^2 we can easily check. And which is not positive definite for x = 0. So convex may not be true okay. Now let us check solve few examples, using twice differentiability of a function f to see whether the function is convex or not okay.

(Refer Slide Time: 32:08)



Now for concavity if you want to see the concavity the function is a concave function on a if and only if hessian matrixes negative semi definite on S okay. The same line e can obtain the proof.

(Refer Slide Time: 32:24)



Now you suppose the first example f(x1 x2) which is given as 4x12 +x22 + 4x1x2 . now we will obtain hessian matrixes of this function what are the hessian matrixes of the function it is del2 upon del x12 del2 upon del x22 del2 upon del x1del x2 del2 upon del x22.

What it is it is 8 4 4 2. And what is D1=8<.0, D2=16-16=0, so that implies this del2 f is positive semi definite and that implies f is a convex function. Of similarly you can check for second problem is nothing but which is g is equal to x1 x2 + 2x12 +x22 +2x32-6x1x3 okay. If you calculate hessian matrixes of g, what are the hessian matrixes of g, first is ∇^2 g (∇ x^2) which is 4 then ∇ x^2 g(∇ x1 ∇ x2) which is 1 and then ∇^2 f(∇ x1, ∇ x3) which is -6 now since it is a symmetric so these terms will come here now ∇^2 f g $\rightarrow$ ∇ $x2^2$ which is 2 and x2, x3 which is 0 again symmetric this will come and what it is ∇^2 g upon ∇ $x3^2$ from here it is 4.

So this will be this matrix, now what is D1?, D1 here is 4 strictly return 0 what is D2, represent leading this is minor of order 2 x 2 which is determinant of 4, 1, 1, 2 which is 8 – 1 is 7 strictly written 0 again, D3 determinant of 3 x 3 which is 4, 1, -6, 1, 2, 0, -6, 0, 4 when you take this determinant so this is nothing but 4(8 – 0) – 1, 4 (-6) and it is 12, it is 12 – 4 – 72 which is negative it is started to so it is negative, okay.

So one is positive, positive and negative so that means matrix is indefinite hence G is neither convex nor concave so in next class we will see that how can we solve some problems if it is a convex programming problem, so thank you.

For Further Details Contact

Coordinator, Educational Technology Cell
Indian Institute of Technology Roorkee
Roorkee – 247667
E Mail: etcell.iitrke@gmail.com. etcell@iitr.ernet.in
Website: www.iitr.ac.in/centers/ETC, www.nptel.ac.in

Camera
Jithin. K
Graphics
Binoy. V. P
Online & Video Editing
Mohan Raj. S
Production Team
Sarath Koovery
Arun. S
Pankaj Saini
Neetesh Kumar
Jitendar Kumar
Vivek Kumar
Nibedita Bisoyi

An Educational Technology cell
IIT Roorkee Production