Indian Institute of Science

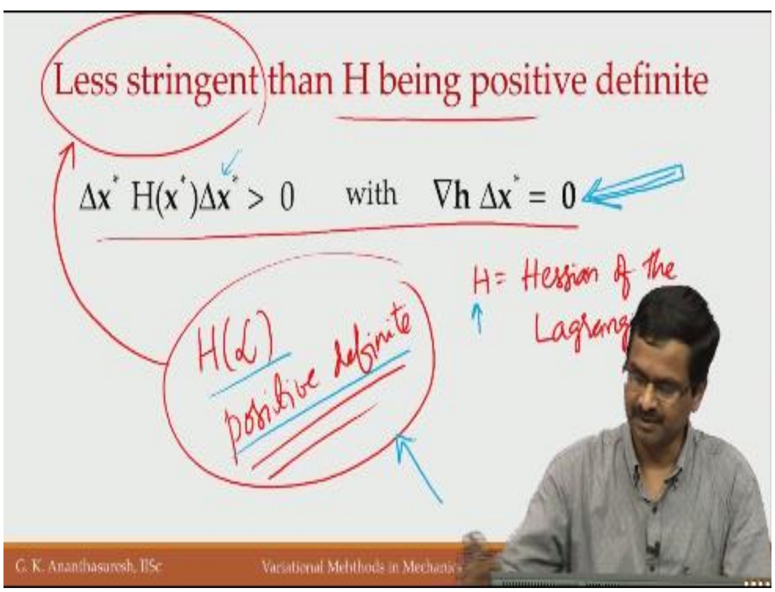
Variational Methods in Mechanics and Design

Prof. G.K Ananthasuresh
Department of Mechanical Engineering
Indian Institute of Science, Bangalore

NPTEL Online Certification Course

Yeah, hello again so we were talking about this sufficient conditions for constrained minima and after a long derivation we arrived at this condition.

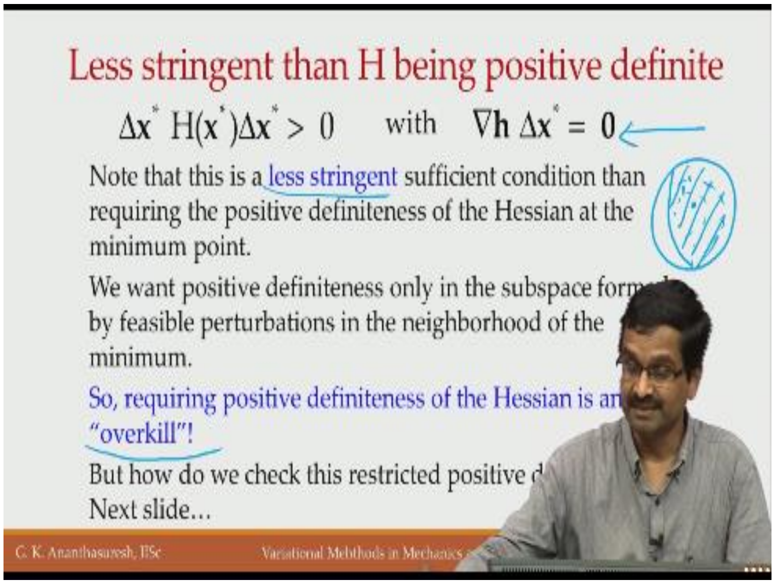
(Refer Slide Time: 00:30)



That is we would like to have Hessian of the Lagrangian note that H here is not the is, is not the Hessian objective function is the Hessian of Lagrangian that should be greater than 0 meaning that the perturbations that you have that is Δx* here U greater than 0 what Δx*, Δx* that satisfy this relationship okay, for such a Δx* you will greater than 0 we say that is a condition. But sometimes some books to make it probably simpler for beginners we will say that hessian of the Lagrangian if it is past definite they say that sufficiency condition.

But that is actually asking for more so what we have written here is less stringent requirement, but then how do you verify it that is a question, okay. If you do this by definite no problem but why check more than what is required okay.

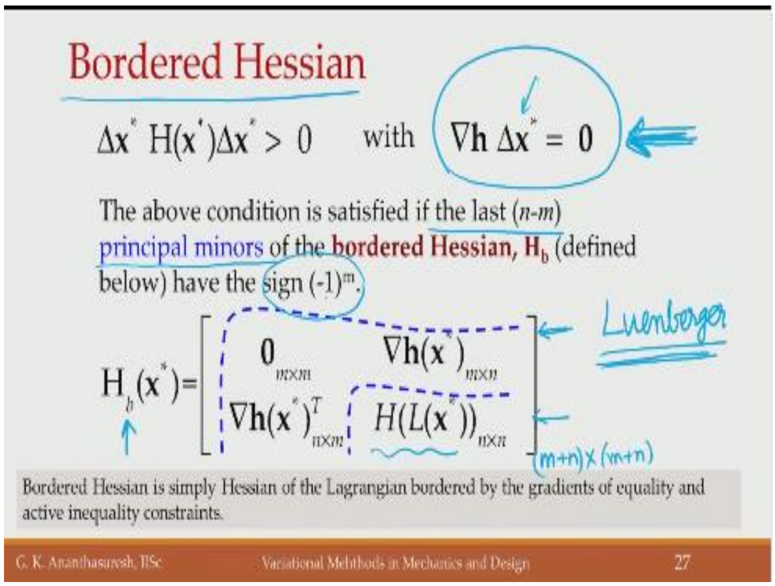
(Refer Slide Time: 01:30)



So in order to check this in the constraint where this is only those perturbation which are feasible right, that is what we said you have to consider the feasible perturbations for those perturbation how do you verify this you cannot be checking all visible perturbations, right if you have a point here in the vicinity let us say this part is the feasible perturbations, right. We cannot be checking everywhere that is unless know how much ever you check somebody may say oh you are not checked here ify ou check here, here, here they will take microscope and say we are not checked here and so forth, right.

Then we cannot be doing that we had to have a condition where we can easily verify this right, so this over care of the requiring the Hessian lagrangian new passed definite we solve that with this less stringent requirement in a way that is numerically easy, okay.

(Refer Slide Time: 02:26)



So checking this easily brings us to the concept of what is called bordered hessian, it is hessian but it is going to satisfy this perturbation restriction okay, that is equality constraints and active inequalities first order term should be equal to 0 when you bordered, okay. This bordered Hessian which is shown here okay, the b subscript there denotes it is the bordered Hessian as if it is a boundary I do not hessian is after a person has but bordered is not after a border person it is just that it is on the boundary, okay.

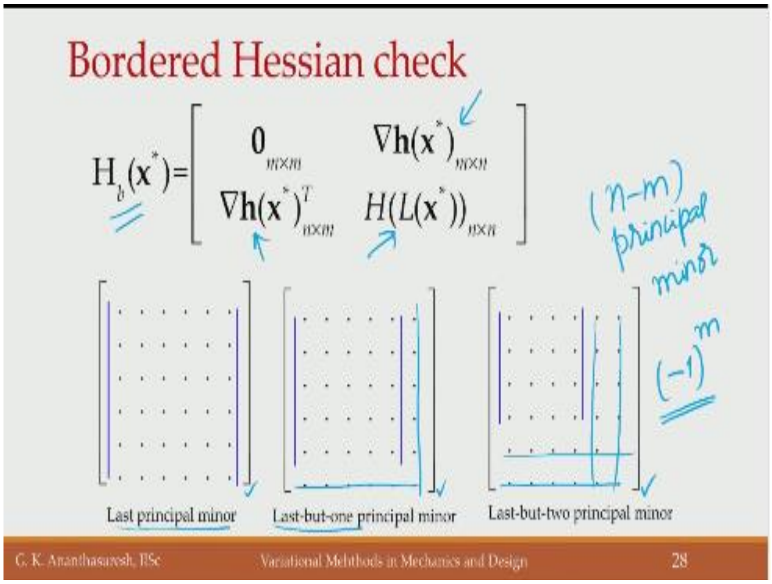
That is what it means that is shown here how do you construct the board Hessian what is size of it if you see this matrix we have m columns here n columns here we have m rows in the first one n rows so we have (m+n)x(m+n) okay, because the first part of it here you see there are m rows and m+n columns the second part of it has n rows and m+n also total (m+n)x(m+n) right, this is a bigger matrix.

See earlier checking this we said no exhaustively does not make sense we need a simple condition the condition is going ahead the board Hessian unfashionably is a much bigger matrix than history of the Lagrangian which is over here which was only n/n but now we are deal with (m+n)x(m+n) and what do we check we check that the last n-m principle minus or the board Hessian should have a sign which is -1 raise to m that is what you are saying.

So first understand how bordered hessian is written we are not going to go into the derivation of it, those of you are interested in knowing more about where it comes from you can look at a book by Leuenberger which is one of the best optimization book for this constrained minimization or unconditional minimization theory a book by Leuenberger you can look at that I think the title of this is I think there is one book called vector space methods, another book called nonlinear programming you can refer to vector space methods book where there are some hints given about why this bordered Hessian concept works to verify this condition,

Okay, this condition which is like an if statement and you know we will get up perturbations which are feasible and check their so this is a simple check, the check is if you construct the bordered Hessian the last (n-m) minus should have a sign -1 raise to m where m is the number of equalities plus the number of active inequalities again we emphasize that active inequalities or same as equalities, okay.

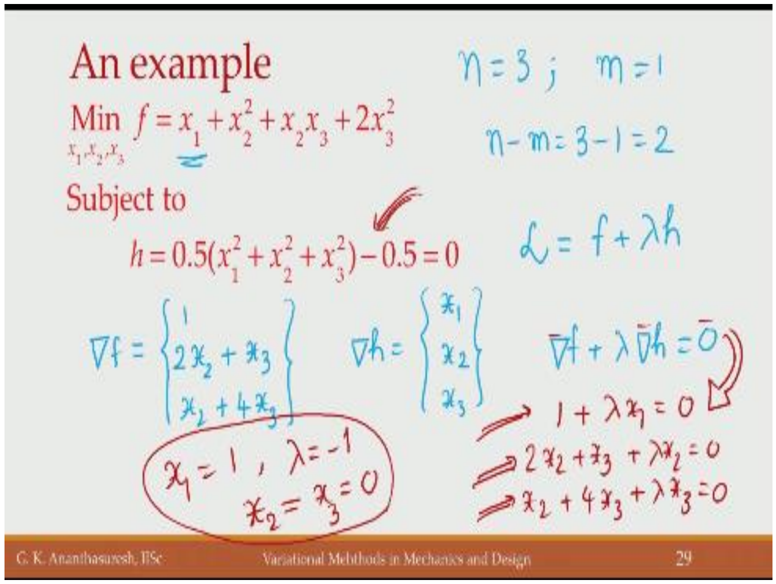
(Refer Slide Time: 05:40)



So what are the principal minus n-m last n-1 principal minor, so the last one is the determinant itself right, the last principle minor is determining the entire matrix. Last but one principle minor where you leave out I think last time I said other way around is not the first principle minor, second principle minor you have to leave out the last row and last column and take the rest and then you leave out last two rows and last two columns and take that, okay so in the last principal minor this is last but one is last but two and so forth you have to keep ongoing up to n-m principle minors.

These principle minors should have the sign -1 raise to m where m is the number of equalities and active inequalities okay, that is how we check very simple check you have to if you have let us say 10 variables and 6 constraints so n=10, m=6 therefore the four determinants you have to evaluate of the bordered hessian which is constructed in this manner with 0s here and the grant of the equalities and active inequalities and the transpose of that hessian Lagrangian all this particular in this form to be (m+n)x(m+n) matrix we have 10 for n, 6 for m will you have 16x16 matrix and that me teaser take this principle minors last n-m last four and then say that there they are all they all have sign -1 raise to m, okay.

(Refer Slide Time: 07:32)



Let us see this with an example, so that you understand how this works so here we have taken a three variable problem our f is a function of x_1 x_2 andx_3 in particular we have $x_1+x_2^2+x2x_3+2x3^2$ subject to one constraint so here in this problem n=3 then m=2 sorry m=1. Let me get the eraser yeah, m=1, 1 equality so n-m=3-1=2 so we have to check two principal minors of the bothered Hessian, okay. So first when you have f we can compute the gradient okay, so gradient because first we have to write the Hessian of the Lagrangian, right.

So you have to write L=f+λh we have only one constraint here so if I just do that gradient of f here is we have take derivative of f with respect to x_1 then give us one that is it x_1 is only the first term over here with respect to x_2 second term will give us x_2 and then third term will give us x_3 and then with respect to x_3 we have third term will give us x_2 and fourth term in the objective function f will give us $4x_3$ that is our gradient of the objective function.

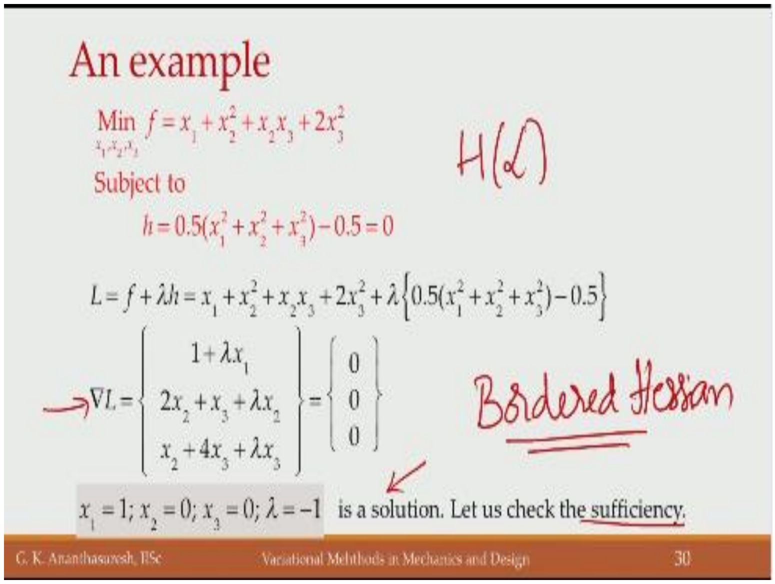
Now we also need to write gradient of the constraint that will give us 0.5 x_1^2X will just give us x_1 and then this will be x_2 so half and two will go and then x_3 okay, now what is the KKT condition, KKT condition the first condition so we have gradient of the objective function times λ times unit of this is equal to zero, so you have to write that and then we have equality constraint so if what I write this particular thing will give me that let me change the color of the ink so we are clear so I say $1+\lambda x_1=0$ okay, and all these are no gradient and we have three equations there that is so coefficient bar and then $2x_2+ x_3+\lambda x_2=0$ and then I have $x_2+ 4x_3+\lambda x_3=0$ so that is what we get that, right.

Three equations and we have three unknowns x_1, x_2, x_3 but then λ is also known for that we have an equation also okay, all those equations if you solve it together you will get some solutions here all linear is everything here this is nonlinear this h equality constraint that these are linear so we have to know some answer but one thing that we can see here if let us say $x_1=1$, in which case I say $\lambda=-1$ then the first equation is satisfied if I also said if we said other ones if I say x_2+x_2 not x2 and crystal in their ink color yeah, x_2 and x_3 are both equal to 0 then this is satisfied this is satisfied and this will also be satisfied.

Because x_1 we made it equal to one here that become 0.5-0.5=0 so this will be one solution that satisfies these three equations and the equality constraint, okay all of them are satisfied that is this is satisfied, this is satisfied this is satisfied and this is satisfied with this one. We will take that and then check it is also the KKT conditions been necessary conditions, right there is no complementarily here, because there is no equality constraint and λ can be positive or negative so we are in fact allowing it to be negative one here, okay.

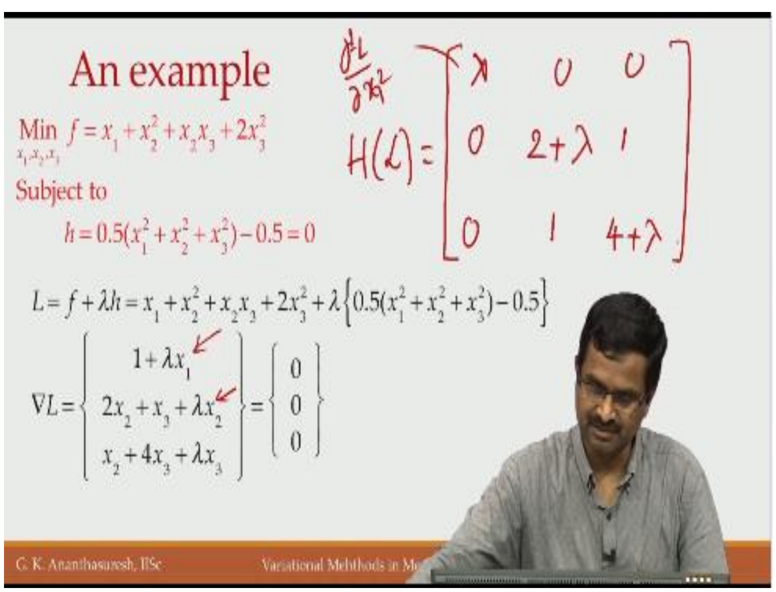
Now we have to check the sufficiency condition to see if this particular point that is x_1 is 1, x_2 is 0, x_3 is 0 okay, if I call this x^* it satisfies necessary conditions but is it sufficient is what we will check okay.

(Refer Slide Time: 12:59)



Whatever I have done there is you know typed up here okay, we said that this is a solution but it is not a solution until we check the sufficiency so we will check for sufficiency as by writing what we call bordered Hessian, okay that is what we will do for that what do we need to do whereas the Hessian of the Lagrangian we have already gradient of the Lagrangian okay, when you take Hessian meaning that this particular thing δl gradient L we take another derivative and write the Hessian of the Lagrangian, okay.

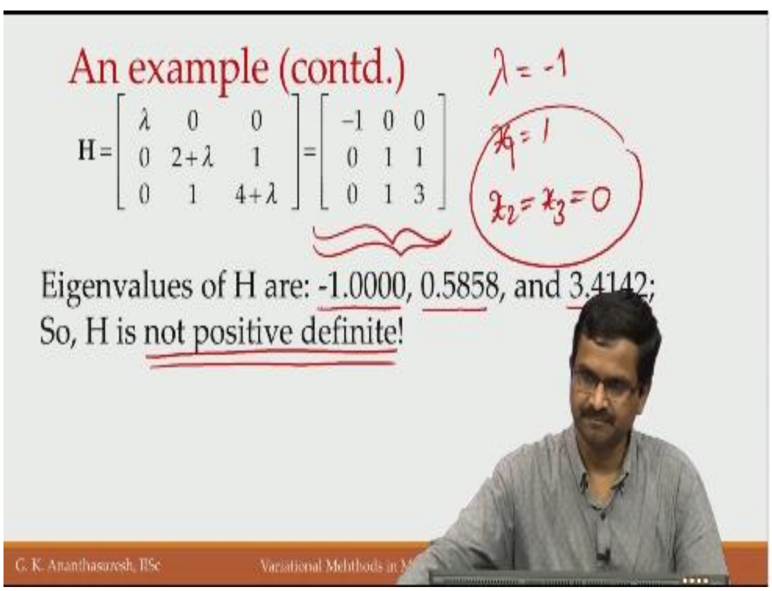
(Refer Slide Time: 13:48)



Well we can do it here the gate δL is here if I say Hessian of the Lagrangian okay, so we need to he said the gradient is a n by n matrix that is you have take the this gradient take another derivative respect to x1 if I do that here i will get 1 or not 1 it will be λ I am taking derivative of this thing with respect to x_1 okay, that is $\delta^2 l/\delta x_1^2$ this is $\delta/\delta x_1$ so if you that 0 is this there any way for necessary condition but what we need to have is the first one over there will be $\delta^2 l/\delta x_1^2$.

Next one will be $\delta^2 l/\delta x_1 \delta x_2$ okay, so we will take this particular one take their with respect to x_1 and write it over there that is actually 0 there and saying with this one and then take the next one okay with respect to x_2 if I do that is 0, 0 and then I come back here and look at this as x_2 I will have $2+\lambda$ $\delta^2 l/\delta x_2^2$ and then over here it will be x_3 that will be $4+\lambda$ whereas the x_2 x_3 if I look at here I have 1 and here it should be 1, okay this will be our Hessians hessian of the Lagrangian, okay.

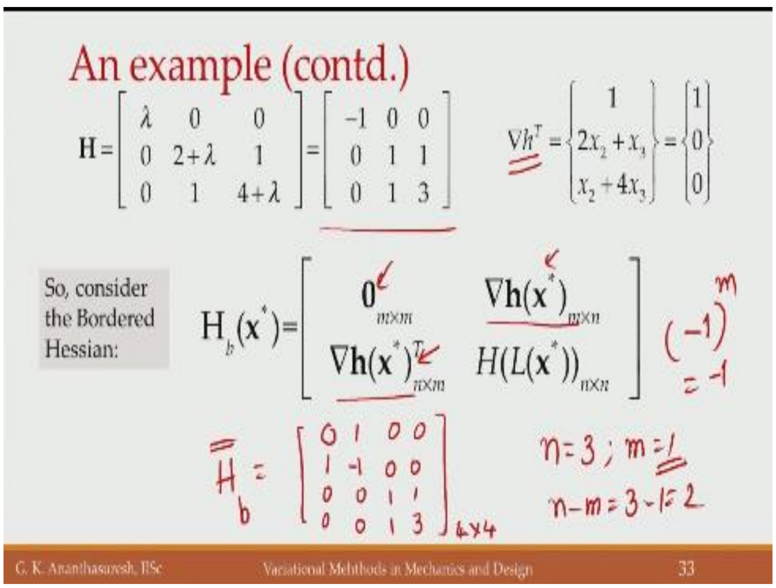
(Refer Slide Time: 15:39)



Which is what we get the previous slide okay, so which is typed up here and noting that λ in our case is -1 right, if you substitute i get -1, 2+-1 is 1, 4+-1 is 3 this is at this point which point that we have taken x1 equal to 1 and then x2, x3 or both 0 okay. Now if you look at this Hessian and find its eigen values which you cannot do just mentally but you can resort to a numerical analysis software like MATLAB and type this type this matrix and finds eigen values you will find that it has a negative eigen value and in the positive one another positive one okay, so this is not positive definite so if you look at some books that do not talk about bordered Hessian then would conclude that this point that is x1 equal to 1 x2 and x3 being equal to 0 you will conclude that it a minimum.

But for this problem indeed a minimum in fact that is why we are taking this example to make a point that the bordered Hessian has to be checked and not his sin of the Lagrangian being positive definite okay, that is an over you are asking for more one less stringent condition is good and that is what we need to check now, okay. Because we have constructed a problem where you do get non positive definite Hessian Lagrangian but yet the point is actually minimum because satisfies the constraint or feasible perturbations, okay.

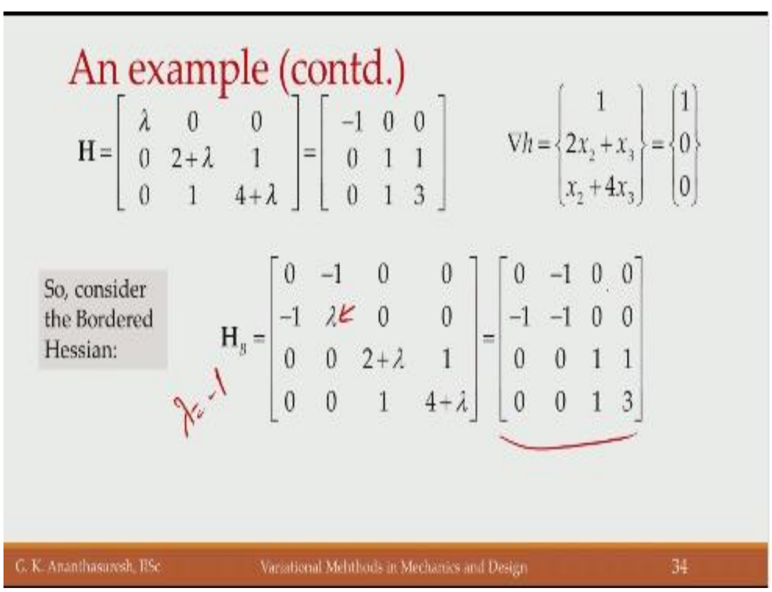
(Refer Slide Time: 17:17)



So we want to write this bordered Hessian right, we already wrote the Hessian which is in front of us yes know the Lagrangian we also know this thing okay, the gradient of the constraint so you can write this matrix so we have bordered Hessian when you write this H_b as a matrix we have that had 0 there mxm we have only one constraint m equal to 1 that will be just 1, 0 and then we have this thing which will be 100 there will be 100 that is we have written this and this and then 3x3 here we have to write this 1 minus 1 0 0 0 1 1 and 0 3,okay.

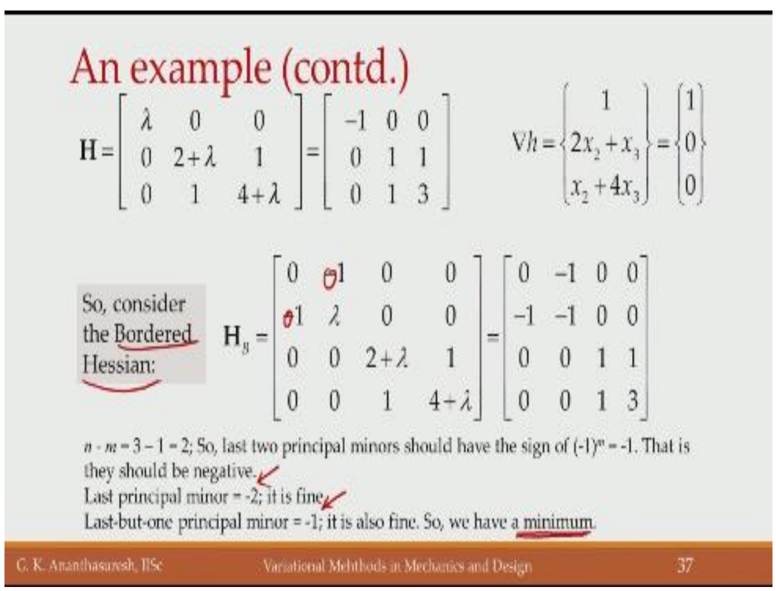
This is a four by four if you remember bordered hessian has a size of (m+n)/(m+n) here m is 1, n is 3 so we have this matrix now we have to look at the last n minus M principal minus again n equal to 3 here and m equal to 1 so n-m equal to 3 -1 equal to 2, last two principal minus and they should have a sign of 1 minus 1 raise to m, right in this case m equal to 1 so it will have minus 1 they should be negative the last two principle minor should be negative, right that is what we need to have.

(Refer Slide Time: 18:52)



So we have written again in terms of λ but then λ equal to in our case is minus 1 so if we do all that we get what we wrote in the previous slide right sorry this was -1 I think I made a mistake oh that was it should be minus 1 I plus 1let us see the gradient what we had taken to begin with that was that was one I think maybe this is the type 1 okay, so this is not -1 then okay you can check it that is not a problem so here if I look at it that is I think this is correct this should not be minus it should be plus that is what we have here okay and x1 was one for our solution okay, now if you take that last two principal minors that we had.

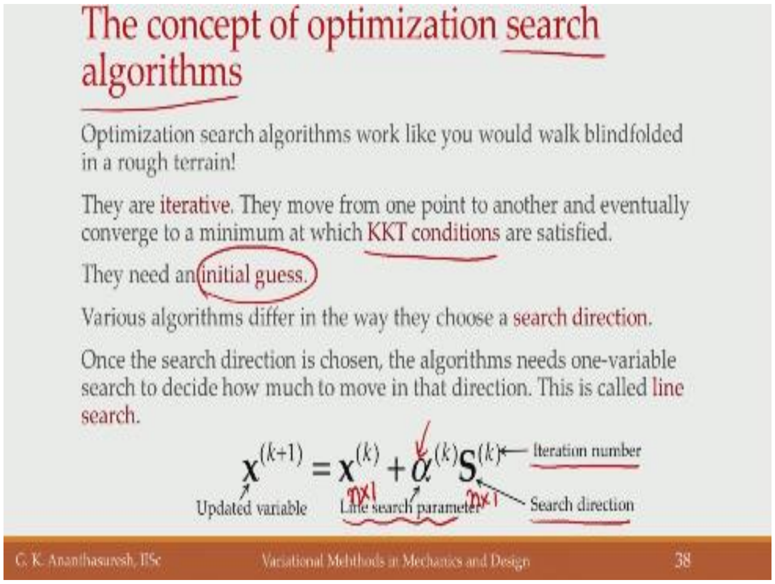
(Refer Slide Time: 19:45)



Example continued let us see yeah, all right we have to take this and last placement of determinant of the matrix itself four by four you have to take and then so that means that you are take determinant okay, so of 0then it is not minus 110100 minus 1000 11013 this determinant if you take that should have a sign which is negative minus 1 raise to m, m equal to 1 and then we leave out the last row and last column then we will let with 0 1 again that is not negative 0 and then 1 minus 1 0 0 0 1 so this determinant also should be negative, right.

If these two are negative then this will be a minimum sergeant sufficient see condition okay, so here if my calculation is right but for that -1 that is there okay, if last principle minor should be -2 you should be -1 so this actually indeed a minimum and not the conclusion that you would have we simply require that hessian Lagrangian is positive definite that is not the case here you will you concrete it's not a minimum but is indeed is a minimum so this is something that you should remember this concept about hessian to know more about it like I said you have to look at this book by Leuenberger, okay.

(Refer Slide Time: 21:33)

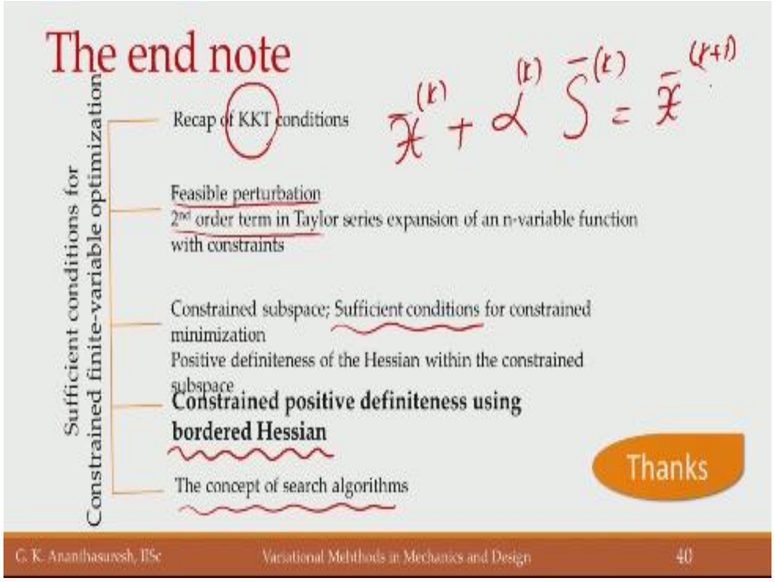


Before we leave this topic and move on to calculus of variations one more thing that I want you to note is how do this numerical search algorithms work right, all obtained algorithms you start with an initial guess so you will have an initial guess okay, you start there and then you move around the space to find the minimum local minimum that satisfies the KKT conditions how do you do that for that you need to have a search direction okay.

The search direction is this thing here s which will have a size equal to the size of x if x is n/1, s will also be n/1 then there will be this line search parameter what we call alpha which is a scalar that is once you choose which we have to go how far do you go right, that is that by this line such parameters so all the algorithms that are there that help you find the minimum with this gradient based methods differ from one another based on how they identify the such direction by the way this K here denotes the iteration number okay.

We start with k equal to zero initial guess we will have x_1 will put that in order to find next one x_2 you have to find the search direction at one and then find alpha 1and then try to find the x_{21} and so forth right, each time you take a search direction move as far as you should and then you will go to new point there evaluate such direction again and move from there and so forth one place to another place this is like taking a bus from here to there and then take a train there and then take a bus or fly or whatever finally end up at the local minimum okay, that is how all the algorithms work.

(Refer Slide Time: 23:36)



All right just to recapitulate this detour that we took in the last three lectures including this one we have to understand KKT conditions perturbations conditions and necessary conditions and then we also looked at this feasible perturbations consider the second order term to establish the sufficient conditions okay, and then particular we talked about this bordered hessian concept which is less stringent condition and more practical not too conservative compared to the hessian of the Lagrangian being positive definite.

And then finally we just talked about this consular algorithm that is there is a search direction okay, and then you move some distance add it to where you are okay, let us say you are the Kth iteration you find this alpha K you have this K and then you get your k+1 next equation keep on doing it until convergence that is how we find okay.

So now we close this detour on finite variable optimization in the next lecture we want to calculus of variations we will then have not this finite variables but will have functions as unknowns that is what we want to have and these three lectures are important because you will be harking back to these things many times to talk about or concert local minimum what is the gradient and what is necessary condition and so forth that is what we need to understand finitely optimization and then move on to calculus of variations, thank you.