Indian Institute of Science

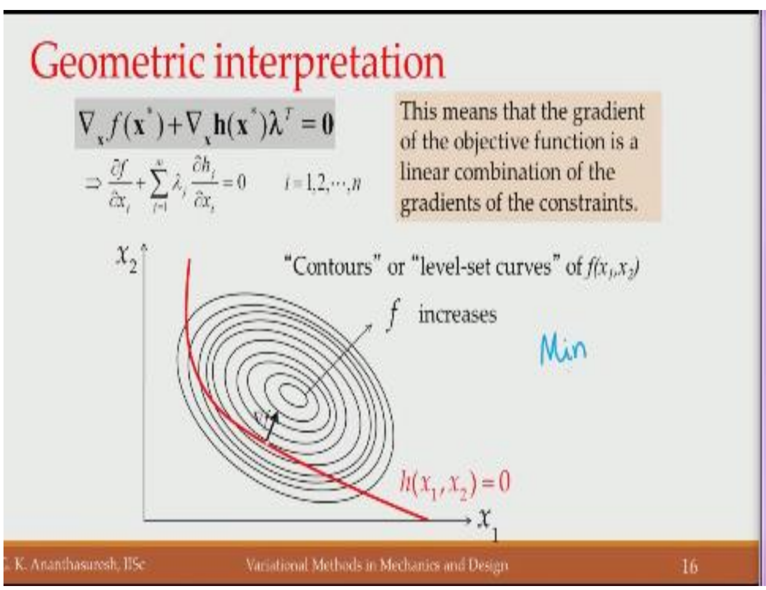
Variational Methods in Mechanics and Design

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

NPTEL Online Certification Course

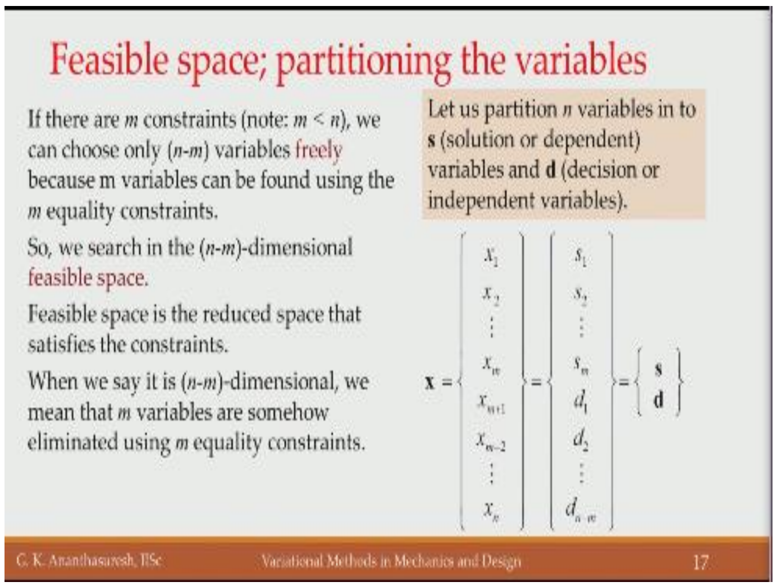
Okay welcome back so we were talking about necessary conditions for constrained minimization with two variables we discussed now we want to discuss the case of several variables so that they say the n variables so we have let us say a problem that we want to say.

(Refer Slide Time: 00:40)



Minimize with respect to when I put x power bar that I mean that there is a vector x1 x2 xn so f is a function of all of those and then we have a vector of constraints so when I put H I have h1 h2 h3 up to HM and it is a function of this X and that is going to be equal to 0 vector okay when you have a problem like this how do we do it when we had only x1 x2 and we could not eliminate one in terms of the other using the Equality constraint we did that in the perturbation variables we expressed Δ x2 in terminal 2 x1 and get that we will do the similar thing to derive this general necessary condition that is rendered objective function being a linear combination of the gradient of the gradients of the constraint okay.

(Refer Slide Time: 01:44)

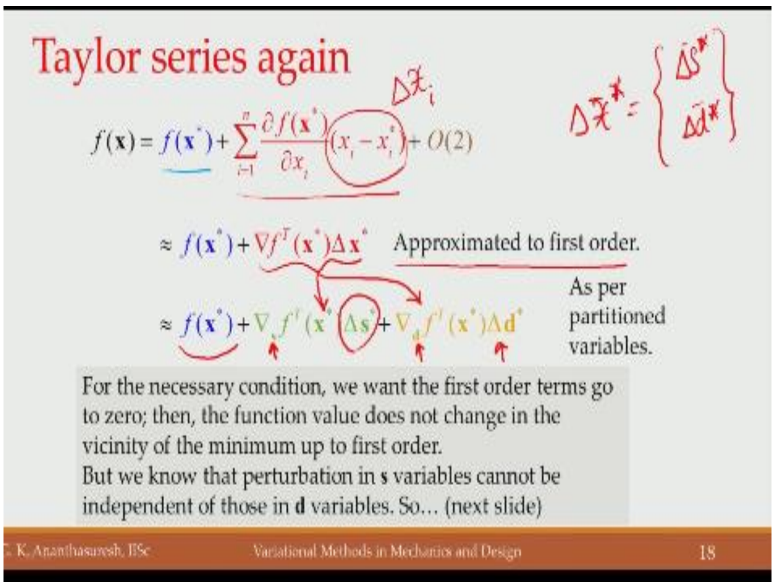


For that first let us understand that our X now is a n/ 1 vector so this is n/1because we have X 1 X 2 xn this one we are going to partition into two we are going to partition somewhere into 2 because we have m equality constraints again we want m to be less than n right so when we have that we will be left with n-M variables so we have a reduced search space which you call feasible space only in that we have to search.

Because when you have two variables and equality constraint limited either x1 or x2 in terms of the other then we had only one variable similarly here when we have n variables and we have M constraints we can eliminate s 1 s 2 s m for example and left with D 1 D 2 D n – M will be having a reduced space okay that is idea so this X now is partitioned into s and D we use s here because that is like solution variables we call them ok solution variables this D our decision variables.

They are independent because we can freely specify n minus M variables once we know them using equality constraint we can find the other ones so we have partitioned this X into s and D solution variables as decision variables d okay.

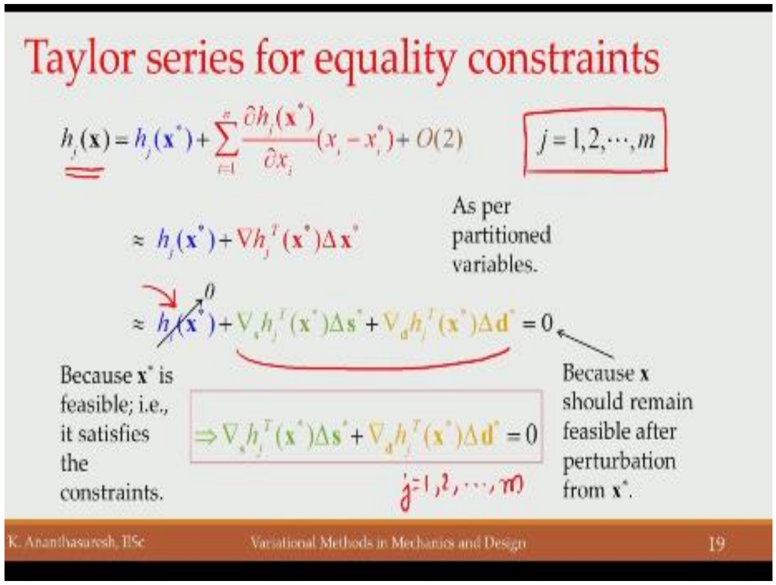
(Refer Slide Time: 03:29)



Now we consider Taylor series approximation again we have f of X f of X star and then let me change the ink color yeah and then we have this first order term ∂ f / ∂ X I, I going from 1 to N and then this is what we had earlier called Δ X I can call it I now Δ X 1 Δ X 2 and so forth this is the approximated to first order okay, now this thing okay this one thing is being split in to two that term and then this term how do you beat split because this Δ X star that we had we are partitioning into Δ s star and then Δ D star okay.

All these are vectors so I should put a bar they really ok so this derivative with respect to gradient respect to s these respect to D right this is how we are doing the first order thing now we say this first order thing should be equal to04 f of X star to a local minimum that is what we say but then we want to express Δ s star in terms of Δ D star because we have freedom only in n -m mm variables we want to do that in order to do that we do the same thing to.

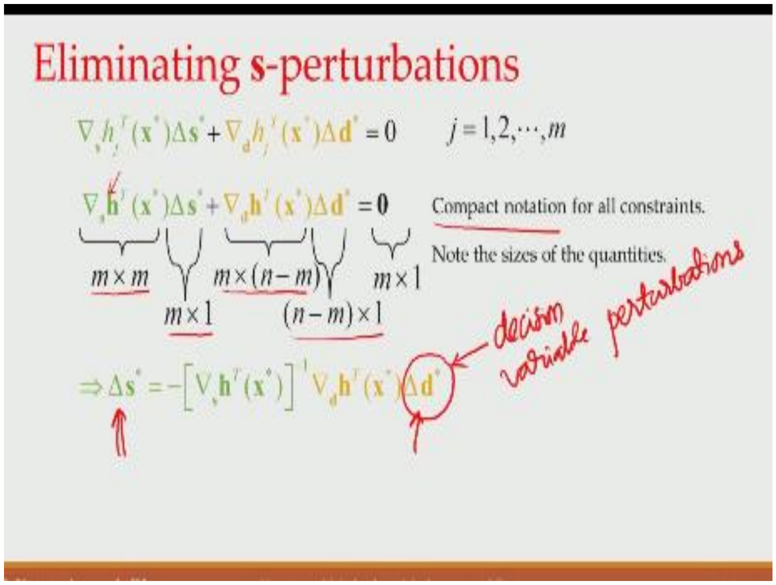
(Refer Slide Time: 05:02)



The Equality constraints we have many so we have ever taken hj where j goes from 1 2 3 up to m right as it is shown here j goes from 1 to m each of them expand and anyway the zeroth order term must be equal to 0 if X star whatever optimum because equal to constraints must be satisfied and then we have this first order term that should be equal to 0 to first order because H each equality constraint h1 h2 up to HM should be equal to 0 you not a perturbation that should equal to 0.

So we get an equation such as this for all j, j equal to one ton so here j equal to 1 2 up to m okay when we have this.

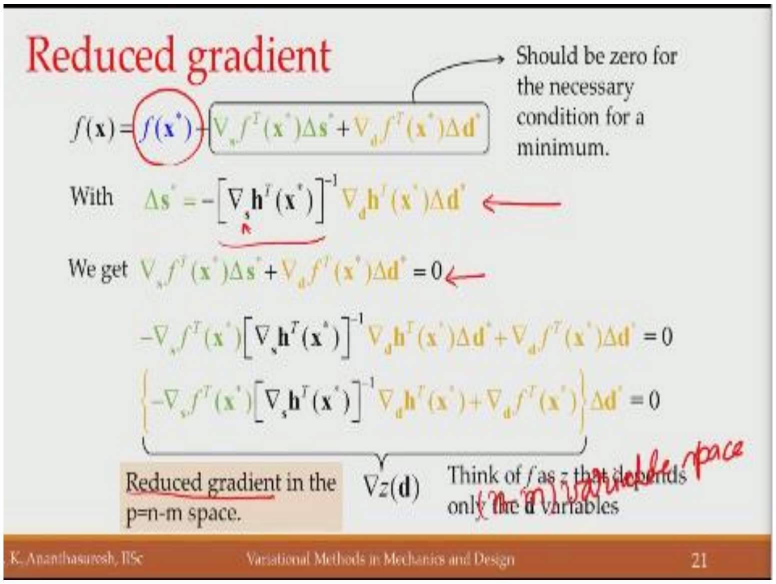
(Refer Slide Time: 05:57)



We remove that j equal to 1 to m and express it as a matrix form a compact notation right we have gradient of s of the entire H naught H J but entire edge is what we are showing here when you do this if you look at the sizes this will be M by M matrix and this is M by one and this is M by n minus m and this is n minus M by 1overall we will get m by one this is also m by one the 0 vector will be my one because we have m equality constraints.

When we have this then we can express this Δ s star in terms of Δ G Star because these are the ones that we can independently very this is what we called decision variables right DS now we have decision variable perturbation in order to see if it is minimum or not we have decision variable perturbations okay in terms of that we can express the perturbation of this solution variables that eliminated m perturbations right. And this Δ s star that we have here we can go back and substitute into the.

(Refer Slide Time: 07:37)

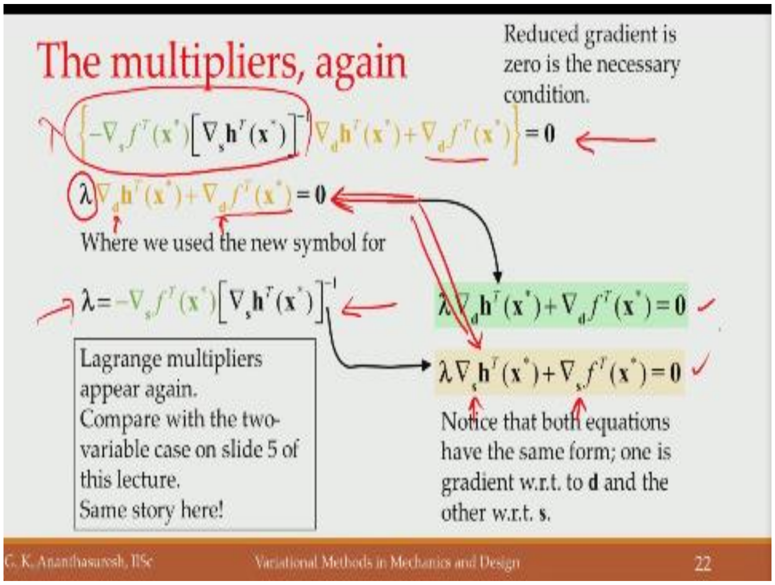


First order f approximation over here right then we get some equation which will be the necessary condition so now we have done just that this is what we have Δ s star and Δ D star will substitute for Δ s star from the previous slide what we had that we have this inverse here of the quadratic and with respect to s right we will have this right so now the first order term is going to look something like this.

So we have this anyway has to be the least value for if X star is a minimize so the first order term should be equal to 0 we return that first order term and say that equal to 0 right now we substitute to be Δ s star we get something like this so what we have here is minus gradient of F with respect to s times gradient of H with respect to s inverse of that times Δ D H evaluate at that point and then Δ F all of that we have right.

So now just like what we had done earlier first of all this is called the reduced gradient reduced because it is reduced in the reduced ton minus M variables face instead of n variable space we eliminated M perturbations which we call solution variable perturbations so we are left with n minus seven is called reduce gradient it is in a smaller space and here the first order term has to be equal to 0 okay.

(Refer Slide Time: 09:08)

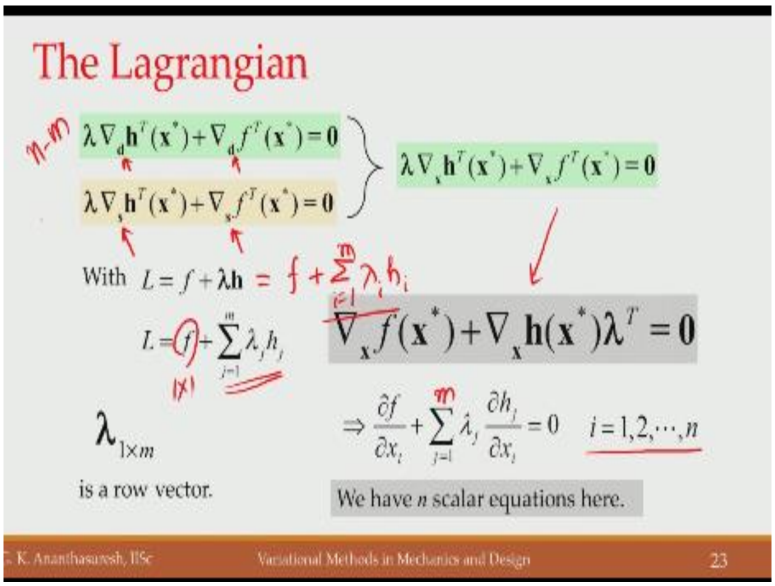


That means that think of let me erases something that we have think of Z okay let me see erase think of this Z okay as a function of so this is adhered think of this as a function of D decision variables okay then becomes a reduced gradient of that new function rather than F. Now we have the first order term being equal to 0 okay then we did something you introduce a λ what do you define λ this whole thing so this is our definition of λ okay.

We have λ times this times this where decision variables that is any arbitrary so first order should be equal to 0 we get this term as it is here now this one the whole thing is λ okay just like we are done in two variable case that gives us λ times this gradient of H which we r inverse we take it by multiplying again we get this equal to 0 okay and this is what we get over here also.

So if we compare this equation and this equation this is respect s here this is respect to D here just like we had earlier the way we define λ gives rise to this one and the first order perturbation of the reduced gradient gives rise to this overall what we have is that if I put together this first one and second one together what I get will be this okay.

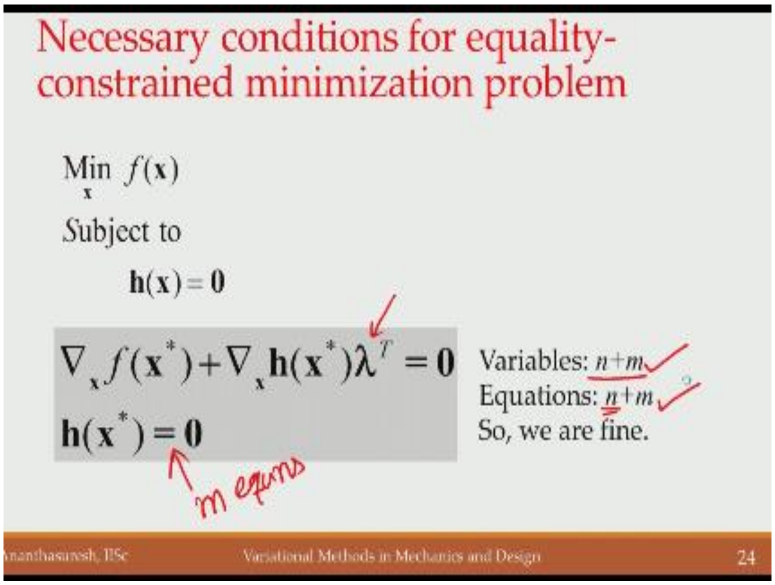
(Refer Slide Time: 11:03)



So the here with respect to D and D that is a first order perturbation in the of the radial gradient vanishing so now the way we defined λ we got with respect to s both are identical is the decision variable the solution variables overall since x is nothing but decision variable solution to put together we get this thing so this is how we get that when we can define a Lagrangian such that it is λ times H basically what we say here is f plus summation I equal to 1 to m λ I times H I.

Each constraint to add by multiplying with the Lagrange multiplier okay linear combination of all the constraints you get lagrangian when you have that lagrangian note that it is a λ 1 by m ρ vector when we write it like this already put λ transpose now I make it like a row vector x let us say λ H this is 1by M this is 1 by M this is M by one so m in bare with one by one that is what is this is one by one right is a scalar.

(Refer Slide Time: 13:54)

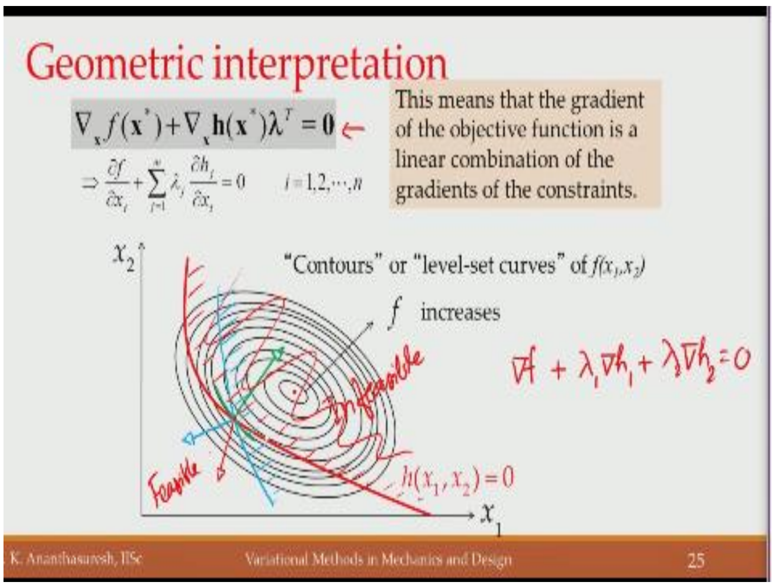


So what we get is a scalar other interpretation what I have written or what is here is that multiplying this okay we get the lagrangian that becomes unconstrained minimization right which is also is going to lead to the same thing that we have here right gradient of the objective function plus λ times gradient of the gradients of the constraints that are the geometric interpretation so great the objective function can be expressed in a combination of the gradients of the constraints okay.

That is what is shown here let me raise a few things so you can see this clearly okay so what we have here is that okay for I equal to one to n now okay we have ∂ f by ∂ X I plus summation j equal to 1 to M here λ J do HJ by ∂ X I equal to that becomes the necessary condition okay here again we have to count how many questions do you have here we have I equal to one to n so these were n minus M these were m put together we get total n so we have n equations in n unknowns are x1 x2 x3 up to xn we have so many equations.

So we are fine we can solve them okay but then we also have the Lagrange multipliers write λ so we have n plus M variables but then where are the equations n equations we already saw necessary conditions other m equations are right here which are the Equality constraint.

(Refer Slide Time: 14:24)



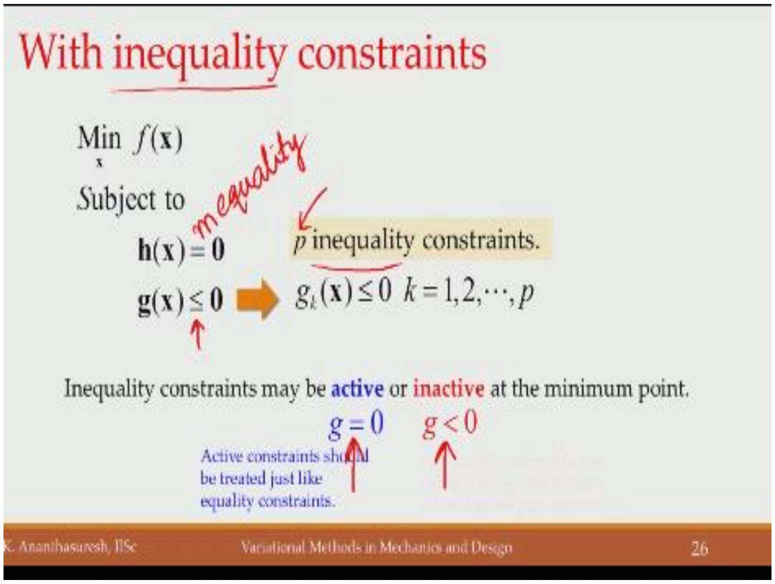
So we have n plus M variables and plus M equations we are fine we can solve okay that is how we deal with equality constraints right and there is geometric interpretation as we already discussed that when you have something like this if there are multiple constraint let us say I choose a different pen color let us say that there is this constraint let us say there is another constraint so we are saying that that is not feasible we can go there similarly for the red one we cannot be on that side we have to be below this curve.

So for this F which has a clear unconcerned minimization over there and catch a minimum and fathers in the infeasible space this side is a feasible this is feasible this side it is infeasible okay so on the feasible site we have to see where the optimum lies right graphically if you look at it say take the green color so at this point the grained objective function will be something like this okay because the contour is like that there that is how the gradient item increases that way grant always points to the most increasing direction.

Now let us at that point put the gradient of this sign one which will be like that okay and similarly let us put the gradient of this one here like that now these three arrows the green one blue one and redone should sum to zero that is what is we are saying here so the gradient of the objective function okay plus λ 1 times the gradient of the first constraint and λ 2 times gradient of the second constraint all these two sum to zero and that is what this is say okay the geometric interpretation is what we have on this slide okay that is optimum right.

So there you cannot make the objective function smaller than what it is here without violating one or the other constraints that is what means here's when I am here this is all my infeasible space right all of this main feasible space I cannot go there how to be on this side when I move their update function value increases or I go into the infeasible space that is a German interpretation of the constraint minimum.

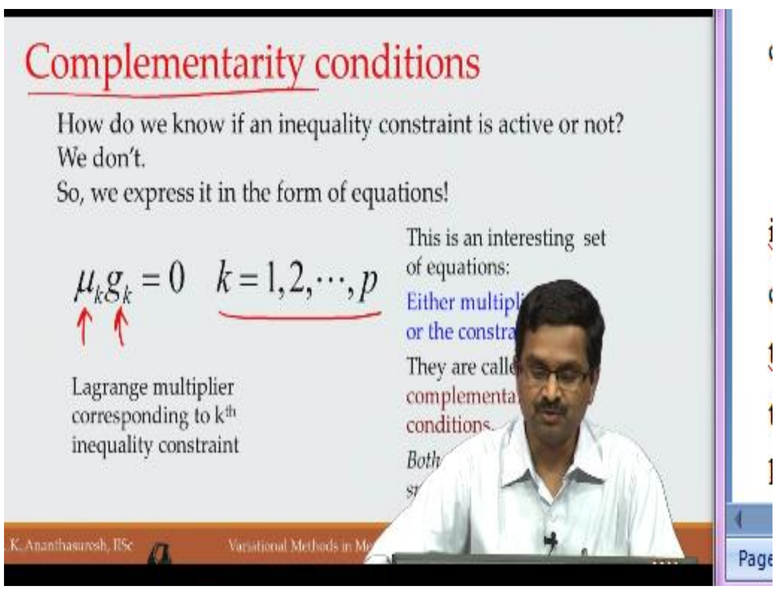
(Refer Slide Time: 17:11)



So two variables and n variables okay now let us move on to inequality constraints now inequality constraints will be denoted like that less than or equal to if it also do greater than equal to but we shoot to the convention of less than or equal to we have let us say P in equal constraints we have m equality here let m equality constraints now p inequality constraints like less than or equal to when you have inequality constraint.

There are two possibilities one is called the active when G inequality constraint expression is zero are strictly less than 0 okay in which case is in why is it or inactive we will see that little later but know that there are two ways that are possible for an equality constraint one is it is active meaning it is equal to zero strictly other is strictly in unequal g is less than zero at the minimum point okay.

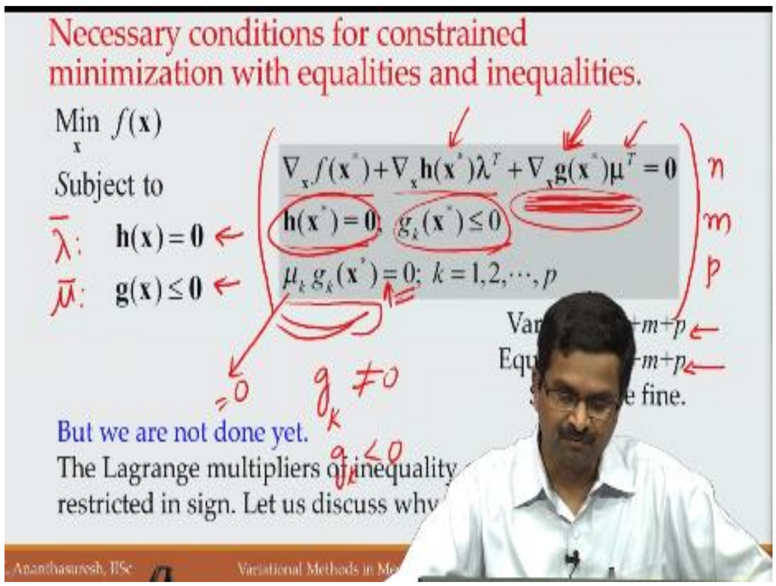
(Refer Slide Time: 18:19)



Now in this case we can write what are called complementarity conditions by writing some variable μ like λ s we had for the Equality constraints we can have Lagrange multiplier inequality and denoted by μ corresponding this k equal to 1 to P we have P inequality constraints we write them like this the nice thing about this way of writing what is called complementarity is that either that or this equal to 0 when you say μK times G k equal to 0 it can be satisfied by making μK 0 RG k 0.

Both can be 0 that is a special case but atleast one of them has to be equal to 0 there are two cases these are called complemented conditions and this will become useful for us okay.

(Refer Slide Time: 19:12)



So now when we have now let us put both back together when we have equality's and inequalities as shown here then we write our necessary conditions in this manner let us see this carefully the first one okay the first one here objective function gradient is a linear combination of the Equality can that we already proved we discussed that now we took the liberty of writing it for the inequalities also by defining this muse here the corresponding supplier was λ now we have defined μ for inequality constraint we have taken liberty to write this.

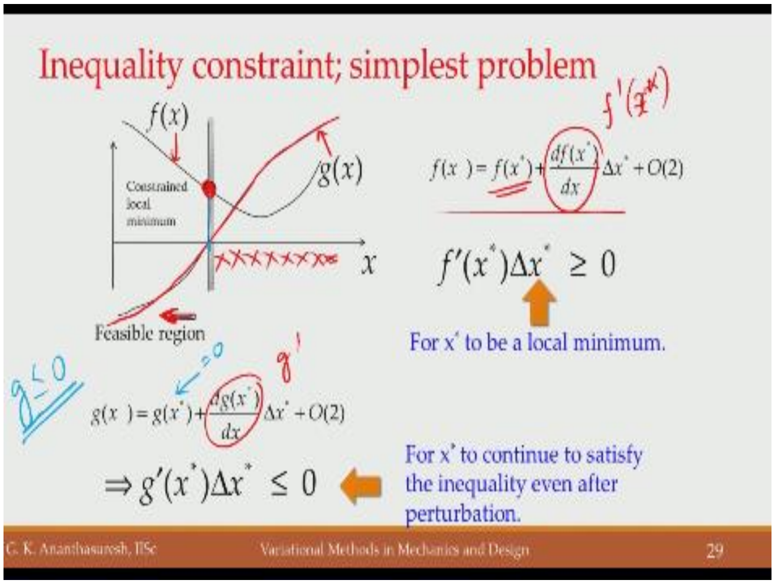
The reason is that whenever G is active meaning a particular inequality constraint is equal to zero strictly that becomes like an equality constraint whatever arguments we put him here will also apply here okay now one might think this is applicable only for those where equality cans any contention attractive but then when it is inactive we have this complementarity condition right whenever G a certain GK is let us say is not equal to 0 it must be definitely less than zero all right in which case to satisfy this complementarity condition.

We must have this μ should be equal to 0 right when that μ is put there is no problem your only adding something that has no meaning because it is basically 0 we can write it in general irrespective of a particular in equal to constraint is active or inactive we can add this term okay we can add this term without any problem okay and then we put the feasibility equal to consume must be satisfied it is must be satisfied with this less than or equal to and then we have this complementarity conditions.

All of these put together there are n here okay and these are not equations right there in equality we do not count them just this part will give me m but these are equation C equality sign is there are P so total variables and equations if you see how many equations n plus M plus P variables are also n plus M plus P because we haven in X x1 x2 xn and then M in λ 1 λ 2 λ m and then p in μ1 mu2 up 2 μP.

So we have enough equations to solve for the variables and total here is n plus M plus p1again it is a number of variables m is the number of equality constraints and pi s the number of inequality constraints okay but we are not done yet so this is not an the end we had say something about we have used different symbol for Lagrange multipliers also to the inequality constraints called μ what is the difference between λ and μ is there anything special that forced us to use a different symbol.

(Refer Slide Time: 22:38)

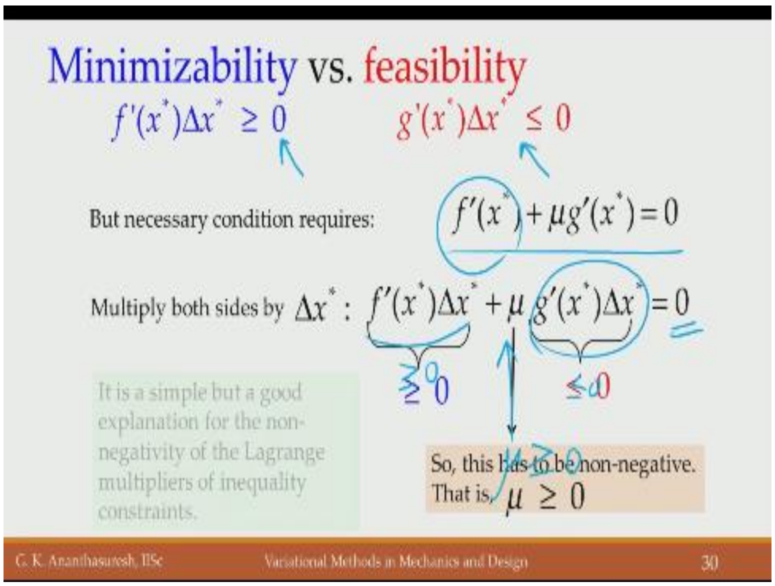


And there is okay this moose have the special thing that they have to be non negative meaning that they cannot be less than zero you will see why let us take a one variable problem for any four inequalities we can take a one variable function there is no problem let us say f of X is like that and g of x is like this okay so this is f of X actually G of X is this ok that is G of X this is f of X now we want to minimize f of X such that g of x is less than or equal to 0 right.

That means that this is our feasible space ok all this is infeasible I cannot go to that right side have to be on this side right then if you look at F clearly this is my minimum constrained minimum which lies always on the boundary right so it is called boundary minimum sometimes or boundary optimum now you consider the first order perturbation for this which is we have denoted this as f prime here okay when I say f prime evaluated at X star times Δ X star should be greater than equal to 0.

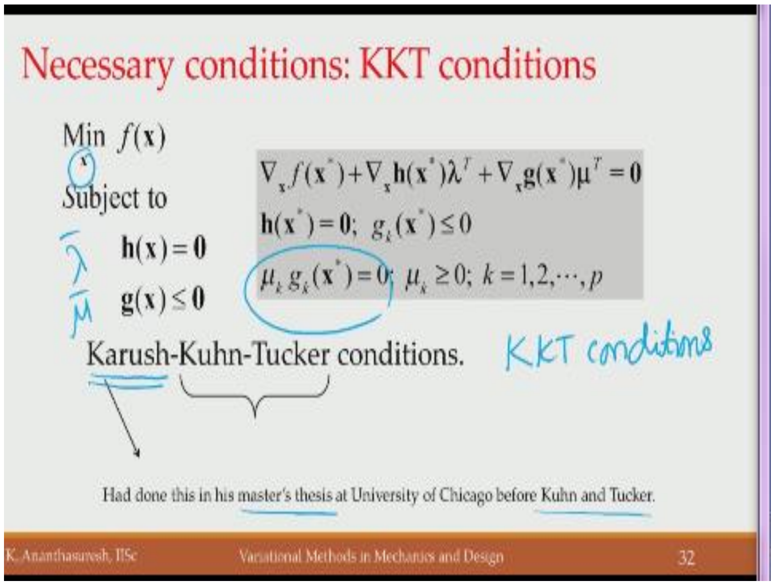
Only then we say f of X R is a minimum right any other perturbation on this side okay should give us a quantity for the up to four star should be greater than equal to 0 right so if you look at the G ok we have G of X is G X Star Plus again you put this thing is called G prime let me erase what going to distracters yeah so we calling it G prime so here we know that GE has to be less than or equal to 0. So g of x star because that is optimum that is equal to 0 because that is at that point z is 0 right so the first order thing Δ x star that should be less than or equal to 0 in order to make this satisfied so what we have now is.

(Refer Slide Time: 24:46)



F prime x Δ x star greater than equal to 0 G prime time Δ X star is less than equal to 0 right so now when we consider the necessary condition that we put their f prime plus μ times G prime equal to 0 we know that this one okay let us multiply by Δ X star this equation so we know that this has to be greater than equal to 0 that and this has to be less than equal to 0 because of that and this should be total should be equal to 0 that makes our μ to be non negative meaning that you should be greater than equal to 0 positive or 0 that is a special condition that comes on this okay.

(Refer Slide Time: 25:37)



So now we also add that necessary conditions now there are n equations here there are M equations here there are P equations there additionally we have this inequality is there and then further inequalities here that Lagrange multiplier corresponding equality constraints have to be nonnegative to positive or zero okay and this actually gives us what we can call what we can condition we have are called KKT conditions these are something that one should be able to write and explain what they mean even if somebody wakes you up in the middle of the night okay.

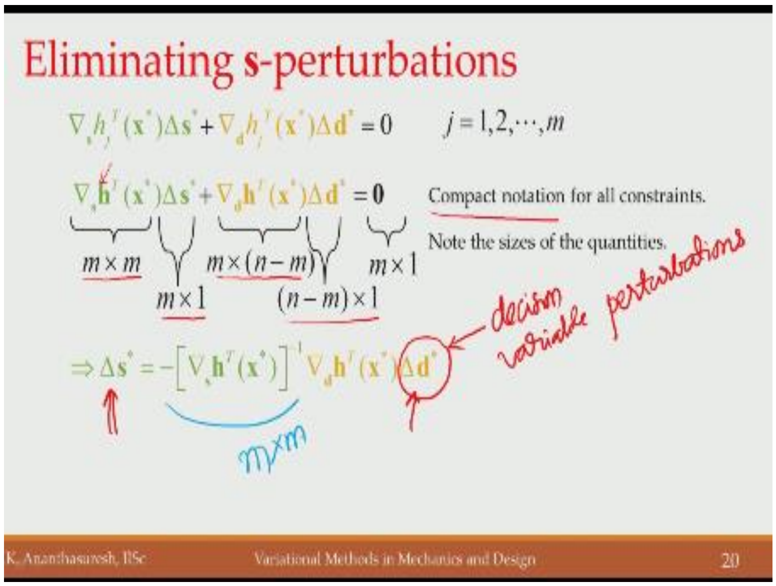
KKT stands for Karush-Kuhn-Tucker conditions three people sometimes people use only KTT conditions but that is not right because Karush she had proved at least a couple of decades before Kuhn-tucker derived these conditions okay they are called KKT conditions and we have enough equations and enough variable to solve for the variables that is X and this λ and then μ okay do not forget this complementarity condition k gk equal to zero that k goes from 1 to up to p the number of inequality okay.

By the way Karush had done that in a master's thesis at University of Chicago couple of decades before Kuhn Tucker did it at Princeton University okay.

(Refer Slide Time: 27:20)

But there is one little caveat that we need to say this KKT conditions are applicable only if this constraint qualification is satisfied okay because it is a necessary condition they must be necessary right but then sometimes you can construct a special problem where there is clearly a constrained minimum but that minimum point may not satisfy the KKT conditions as written here right in such cases the constraint qualification test might be not applicable. So for that reason actually we have to go back to slide number 17.

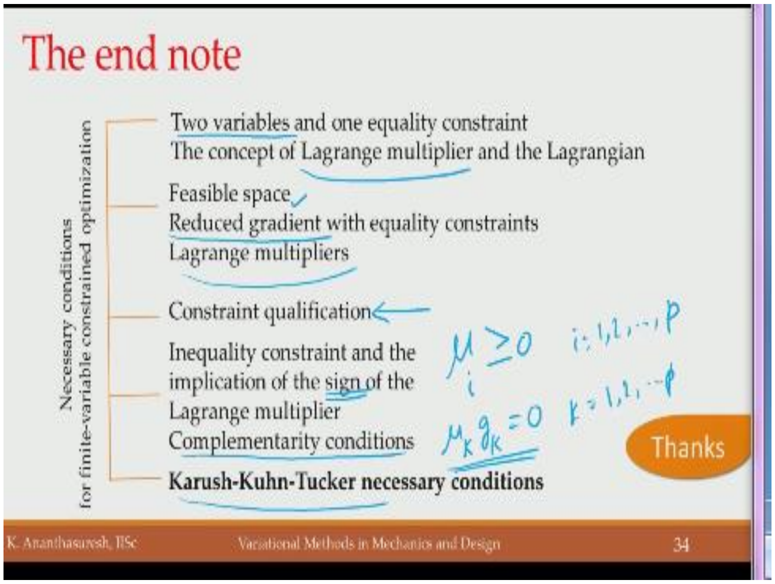
(Refer Slide Time: 28:02)



Of the previous lecture which will quickly go by number 20 in at 17 so Δ s star we are expressing into Δ D star where we had the inverse of a matrix inverse of a matrix cannot be always done right what if this matrix which is M by M matrix here right what if it is singular we cannot do that right so that is constraint qualification we say that gradient of the Equality constraints must be linearly independent then M by M will not be singular we can take inverse okay.

So that is basically constraint qualification so here since we put that caveat here that KKT conditions are applicable when we have the so-called considered qualification this is slide number 20 okay so that is important when you have a problem please check and ensure that the gradients of the Equality constraints are linearly independent so that when you try to eliminate some very perturbations interms of the other you take inverse you need that non singular matrix there that is satisfied only if they are independent okay.

(Refer Slide Time: 29:21)



That is concerned qualification that we need to remember so to summarize we first consider the two variable problem and eliminated one equality constraint when we cannot do it we went for the first order approximation eliminated one part of interms of the other that gave rise to this concept of lagrange multiplier okay and then we have the feasible space where the constraints are satisfied in the case of n variables we talked about reduced gradients where we eliminated M variables perturbation in terms of the remaining n minus M variables then we had m Lagrange multipliers.

We just talked about this constraint qualification concept which simply means that the gradients of the Equality constraints must be linearly independent when you also add in any quality constraints it is the active inequalities which are g equal to 0 when you have that those gradients all should be linearly independent along with the rest of the Equality constraint gradients and the sign of the leg Rommel tip liar for the inequality constraint that μ we search should be greater than or equal to 0 μ I where I goes from 1to P.

Without the complementary conditions which basically will have a very nice well disk is there are the same I equal to 1 to P and what we have written k is also equal to Ø that is they will tell you where the constraint is active or inactive okay finally we had karush Kuhn tucker conditions that capture all of this so this finishes the necessary conditions sufficient conditions we can discuss but then this is a detour so we get back to calculations in the next lecture.

So basically Sofia condition also exists in the case of n variables m equality constraint spin equality constraints when you have you read necessary conditions you can talk about the Hessian like we discussed in the unconcerned minimization you can talk about the same thing here but instead of going too much into that since our course is about this variation methods we will get back to that in the next lecture okay, thank you.