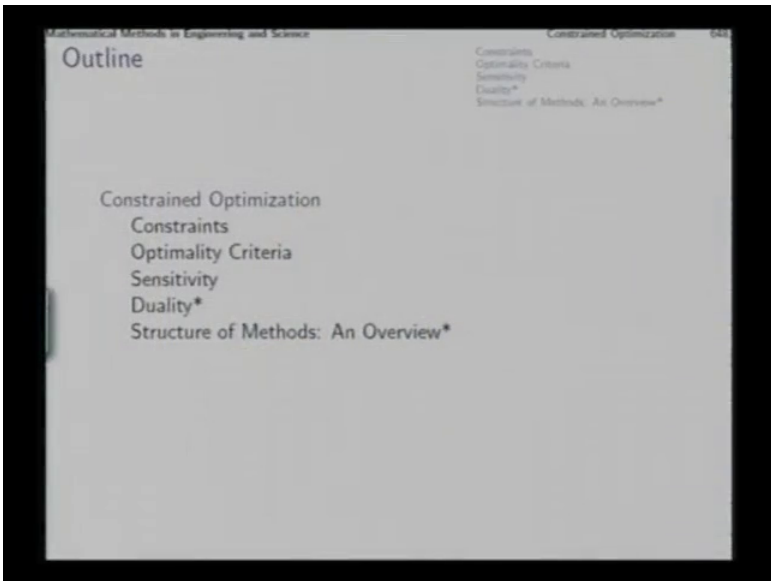
Mathematical Methods in Engineering and Science
Prof. Bhaskar Dasgupta
Department of Mechanical Engineering
Indian Institute of Technology, Kanpur

Module - IV
An Introductory Outline of Outline of Optimization Techniques
Lecture – 19
Constrained Optimization: Optimality Criteria

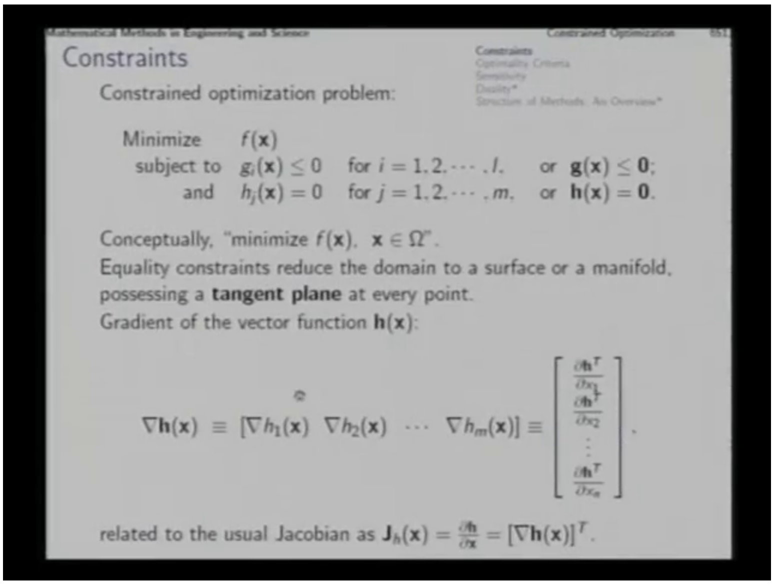
In the last 2 lectures, we discussed unconstrained optimization.

(Refer Slide Time: 00:22)



Now, in this lecture, we will discuss the basic frame work of constrained optimization first we discuss constraints.

(Refer Slide Time: 00:32)



As I told you earlier the typical form of an optimization problem is like this in which you want to minimized a function of several variables in the vector x subject to certain inequality conditions and certain equality conditions.

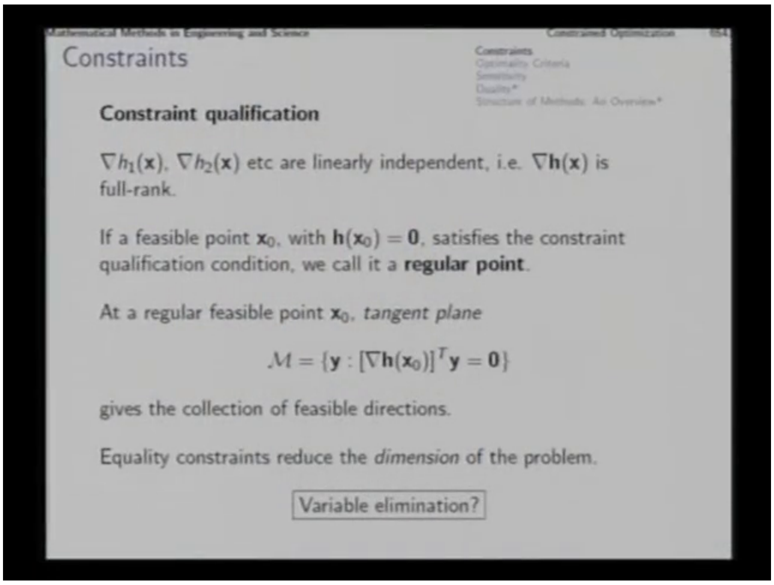
Conceptually you can say that the statement of the problem is minimized function of x in which x must belong to a given domain omega now the description of the domain can be made in terms of these constraints equality constraints constraint the feasible space to a lower dimensional set of the original solution space as if you can talk of it as a surface 3 dimensional space. So, in the; if the space of x is 3 dimensional the one equality constraints like this we will define a surface and any point outside that surface will be deemed as in infeasible inequality constraints do not reduce the dimension of the solution space, but they restrict certain regions of the solution space as invisible.

So, out of the 3 dimensional space expressing inequality constraint as on this side of that wall of the room on this side of this wall of the room like that we can restrain the feasible domain to this room rather than the entire infinite space. So, when we have such constraints you can talk of that a tangent plane at every point on the surface describing the equality constraint. So, for equality constraint if you have a surface then at that point at that surface every point is feasible and at that point you can talk of a tangent plane.

So, how would you describe the tangent plane? So, for the vector function h in which the components are in h 1, h 2, h 3, etcetera, if you put their gradients like this grad h 1 grad h 2 grad h 3 and so on and construct this entire matrix which turns out to be basically the transpose of the Jacobean of h which respect to x then you can say you can see that grad h 1 will be the normal to the surface describing hx equal to h 1 x equal to 0. Similarly grad h 2 will give you a normal to the surface h 2 x equal to 0 and so on.

What remain as the tangent plane will have directions which are perpendicular to all these gradient vectors, right. So, then before proceeding towards the theory of constrained optimization, we need to keep in mind one important point is that.

(Refer Slide Time: 03:37)



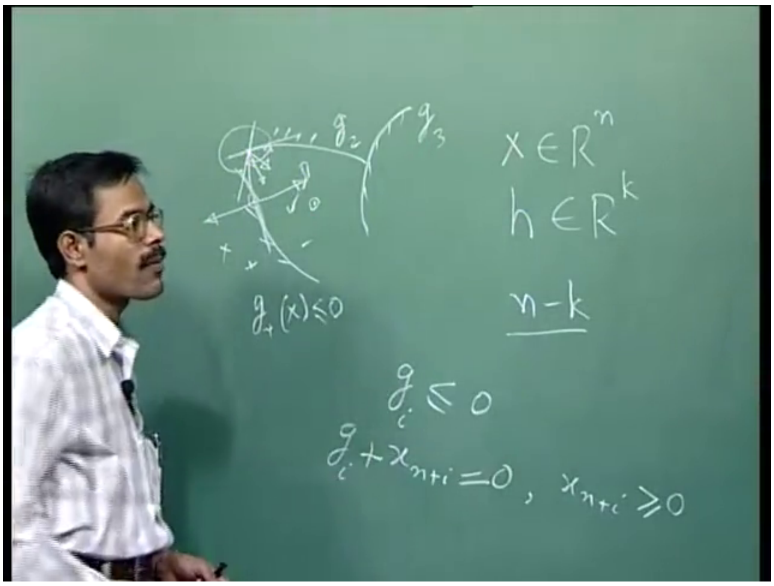
In entire discussion that we make subsequently it will be understood that grad h 1 grad h 2 all these are linearly independent. So, these columns are linearly independent a point x which satisfies that condition is called a regular point that is and this condition the linear.

independence of the gradients of the functions h 1, h 2, etcetera that the property the quality of linear independence that is called the that condition is called constraint qualification the idea of naming it as constraint qualification condition is that if at a particular point there are 2 constraints, but both of them have the same gradient or the gradients in the same direction and that will mean that the tangent plane that is allowed by one of them is exactly the same as what is allowed by the other similarly for the case of more than 2 such constraints and more than 2 such gradients if it happens that they are linearly dependent then that will mean that all the normals are not independence all the normal to the surface is not independent.

So, for that local neighborhood the immediate first order region around that point for that region for that small region one of the constraints would be dropped out; so, because they are not locally all independent. So, that is why when we say that all of these are linearly independent then we mean that in the neighborhood in the vicinity of that point all these constraints really qualify as independent constraints. So, it make sense to retrain all of them otherwise we could have dropped one of them for the local neighborhood. So, in all our theoretical discussion that will make subsequently it will be understood that we are talking about regular points where the constraint qualification condition is satisfied now at such a point at such a feasible point which is regular we have already discussed that in the tangent plane any direction is a direction which is perpendicular to orthogonal to all the grad h 1, grad h 2, grad h 3, etcetera, right.

So, these are the directions which are feasible directions that is along the direction you can make an inf- decimal moment without violating any of the constraints if your room ent has a component along the gradient of h; that means, you are moving out of that surface out of that tangent plane. So, that will be infeasible. So, this tangent plane this sub space m consisting of those ys; those vector those directions which are orthogonal to all the grad h 1, grad h 2 vectors; they are the correction of feasible directions so far as the equality constraints are concerned.

(Refer Slide Time: 06:54)



So, we find that if in a n variable problem in which at any point in the solution space there are n linearly independent directions possible if there are k constraints say if there are k constraints then and all of them are locally independent that is they satisfy the constraint qualification condition then on the vector y on the feasible direction y there will be k such conditions that is that y must be orthogonal to k different linearly independent vectors. So, if there are k such conditions then the number of linearly independent vectors that can be taken as y will be n minus y n minus k.

So, n minus k linearly independent vectors can be there in the tangent plane that way you will find that from that point if we want to conduct a search we do not really need to conduct a search in the n dimensional space because the tangent plane is n minus k dimensional and in fact, all the k equally constraint together will define a manifold of dimension n minus k. So, that way you need to conduct the search on n minus k dimensional manifold and that tangent plane is a planar entity of that many dimension. So, that way you find that equality constraints reduce the dimension of the problem.

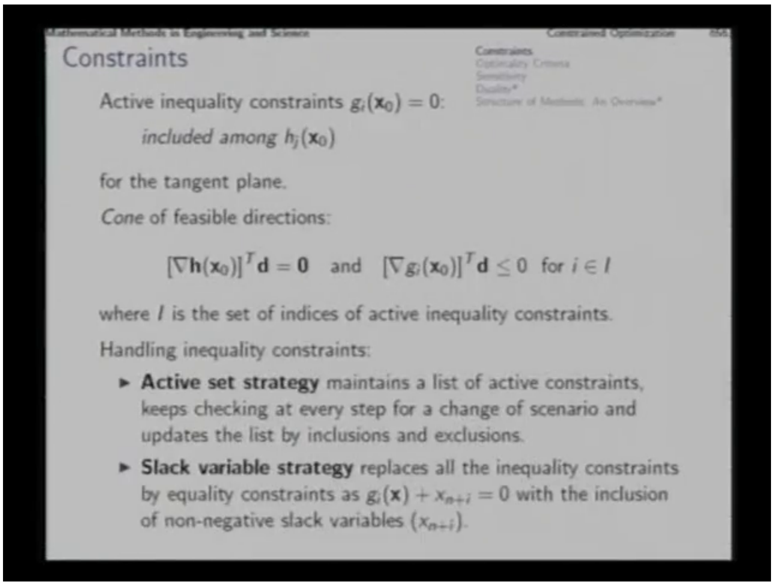
So, that maybe positive thing that may be something which is actually which in a way reducing our job but since in general the surfaces h x equal to 0; so, those surfaces are non-linear they are not all planar surfaces. So, that is how that is why we will not be a good passion to take advantage of this reduction of dimension because the satisfaction of the constraints itself will be to remain on the feasible domain; however, there is a method called variable elimination method which tries to use this fact and solve k of the variables in terms of the other n minus k variables and then conduct the optimization process in that n minus k variables that is possible only in case of very simple constraint surfaces very simple constraint functions.

However; so for as the tangent plane is concerned and that is analysis of feasible direction in the tangent plane are concerned effectively we keep in mind the concept of this elimination of certain direction and the restriction of the feasible direction set to a sub-face or to a tangent plane now what about inequality condition a constraints for inequality constraints all of the inequality constraints need not be active where a the point for example, if this is an inequality constraint say g 1 x is less than equal to 0. So, this is the feasible side this is the infeasible side that is g 1 x is positive here negative here and g 1 the boundary similarly g 2 x, similarly g 3 x.

So, you will find that at this point at this point none of the constraints is active in what sense that is in the immediate neighborhood of this point whichever way we move in whichever way we try to displace the points these constraints will not play a role on the other hand at this point g 1 is called an active constraint because from here there are some directions in which g 1 is not highlighted on the other hand there are some directions on along which g 1 will be highlighted. So, g 1 is an active constraint g 2, g 3 are same in active at this point g 1 g 2 both are active constraints g 3 in active and so on.

So, when a particular constraint is active; that means, the value of the corresponding constraint function is 0; that means. So, as for that constraint is concerned the point is on the boundary of the feasible domain; so, for the description of the tangent plane.

(Refer Slide Time: 11:28)



We include the active inequality constraints in the list of the equality constraints of this because here if you try to work out the tangent plane you will need to be consider this that is excluding this normal direction because this will be the gradient direction this will be the actually gradient direction gradient of g 1.

So, whatever we do for grad h 1 grad h 2 grad h 3 same thing, we will do for grad grad g 1 here, but we will not do that for grad g 2 grad g 3 because they are inactive constraints at this point this constraint are inactive. So, active equal inequality constraints get added get included among the equality constraints for the description of the tangent plane; that means, if an inequality constraint is active; that means, the tangent plane gets its dimension reduce by one for the value one for the dimension apart from that for inequality constraints there is a concept of cone of feasible direction.

In the case of an equality constraint a point which is feasible from that you cannot move along grad h you cannot move along minus grad h both ways, you will be leaving that constraint manifold constraint surface and going out of the feasible domain on the other hand for an inequality constraint you cannot move along grad g, but you can move along negative grad g because that you will be coming to the interior of the domain you will not leave the domain for example, at this point it is possible to move in all this directions. So, at in this point if you try to do draw 2 tangents like this and like this; this is tangent 2 g 1; g 1 equal to 0 surface this is tangent to their boundary of g 2; now in this entire zone all the directions are feasible right.

So, you find here that a cone like structure appears in to pictures in which the directions are feasible the opposite directions will not be feasible. So, these direction will not be feasible this will be feasible those directions which goes towards the interior of the domain leaving constraint boundary, they will be feasible and opposite once will not be feasible. So, you define the cone of feasible direction in this manner at that boundary point the value of g is 0 and any direction that takes the x point towards a direction in which g transport d is negative will be all right because g will become negative.

So, now when we have to algorithmically handle the constraints inequality constraints which are active and which are inactive there are 2 possibilities of handling 2 ways of 2 strategy to handle inequality constraint the they are being active or inactive one is active state strategy in which at every retraction as we move from one point to another a list of active constraints is maintained and after every iteration we make a check regarding which of the active constraints as become inactive due to this particular step and which other constraints which were earlier inactive we have become active.

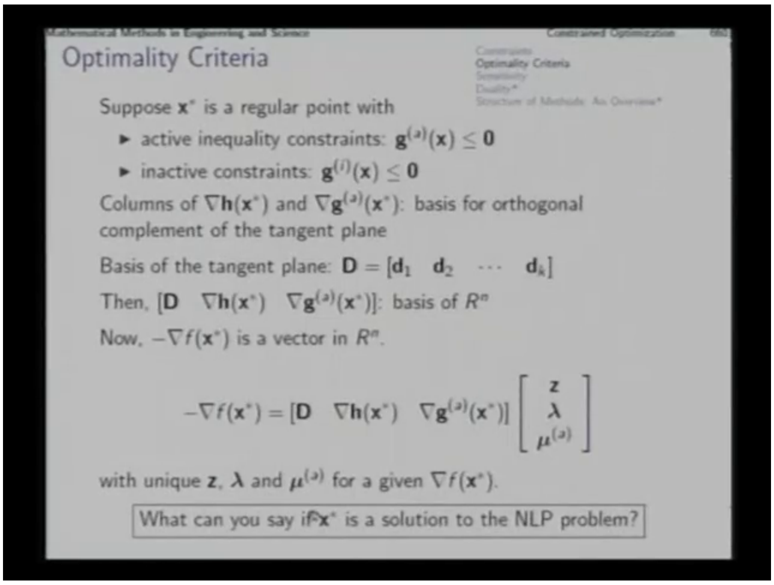
So, this is called the active set strategy at every retraction we make this update of active set active set update is basically in the form of a list a set in which we enter the indices of the active constraints there is another strategy called stack variable strategy in which every inequality constraint like this is replaced with a corresponding equality constraint by addition of another variable which we call as the stack variable with the condition that the stack variable sorry this is equal equality with the condition that the stack variable must be non negative. So, we put a non negative number here. So, if it is its value is positive; that means, this constraint is inactive if its value is 0 then this constraint is active in negative value for the stack variable is not allowed because that will make this constraint violated that will violate this constraint.

So, this is another strategy to inequality constraints that is which of them have to be taken in the active set and which are not. So, for that you do not need to maintain a list the value of this will signify whether it is inactive or active now with this understanding we go to find out we proceed to find out the optimize criteria for a non-linear optimization problem in which there are constraint. So, for example, as we have been discussing that for this corner point these directions are feasible and any direction in that tangent plane is feasible in a tangent plane is feasible.

Now if these are feasible directions, then what is the necessary condition for this point to be a local minimum point the condition should be that along the feasible directions if we make a move, then the function value should not decrease because if along a feasible direction the function value can decrease that will mean that the current point cannot be a local minimum. So, a direction along which the function decreases is a is in direction and the direction along which we can move without violating any constraint is a feasible direction. So, for the point to be a local minimum point it is necessary that there is no direction which is at a same time a feasible direction as well as a descent direction.

So, if we can check that all the feasible directions are such that along that the function increases function do not decrease then we will be satisfying the necessary condition for the point to be a local minimum.

(Refer Slide Time: 18:11)

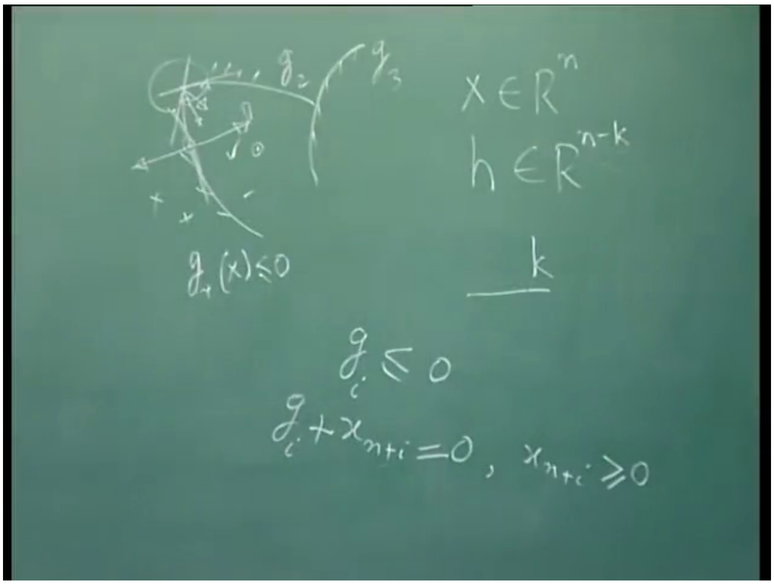


So, suppose x star is a regular point in which there are a few active inequality constraints and some inactive inequality constraints and of course, equality constraints are always active. So, we collect the all the active constraints grad h full and grad g only the active part and then we say that grad h 1, grad h 2, grad h 3 and out of these also gradients of the active inequality constraint function we collect.

Now, these gradients together which are all linearly independent because they satisfy their constraint qualification condition; so, all these together will be those vectors to which a tangent plane vector must be orthogonal right. So, these collection this collection of gradient vector is the full set up linearly independent vectors which must be orthogonal to any vector in a tangent plane.

So, therefore, these together these columns together give us a bases for the orthogonal compliment of the tangent plane if tangent plane is this then whichever subspace which is orthogonal to it and completes the R n full vector space then that is the also complement of the tangent plane for which these gradient vectors together offer a basis for the tangent plane also we can work out the basis suppose that basis is d having these many vectors i which I; this case we described n minus k as dimension a tangent plane, but in this notation here actually the number of constraints n minus k.

(Refer Slide Time: 20:04)

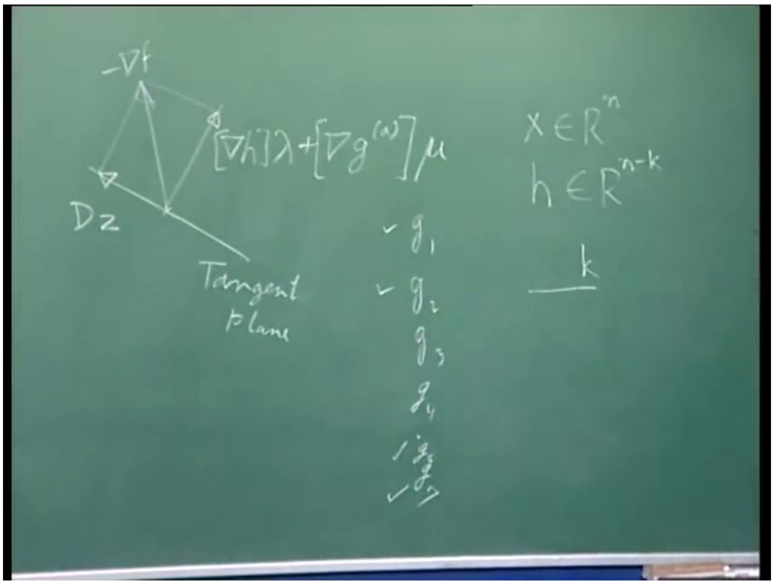


And dimension of that tangent manifold tangent plane is k. So, if we think of a basis for the tangent plane itself, then this basis for the tangent plane and this basis for the orthogonal compliment of it that is what is perpendicular to the tangent plane together gives us a basis for r n that is full n dimensional space.

Now, x is an n dimension vectors. So, when the space of x any vector is in n dimension vector in particular gradient of f is also an n dimension vector. So, it can be expressed in this basis. So, if you try to do that the negative gradient is a vector in r n. So, we try to express that in this basis one part of it describes the tangent plane and the other part describes all orthogonal vectors.

Now, we say that d z is the component of negative gradient in the tangent plane and these 2 together is component of the negative gradient orthogonal to the tangent plane.

(Refer Slide Time: 21:22)

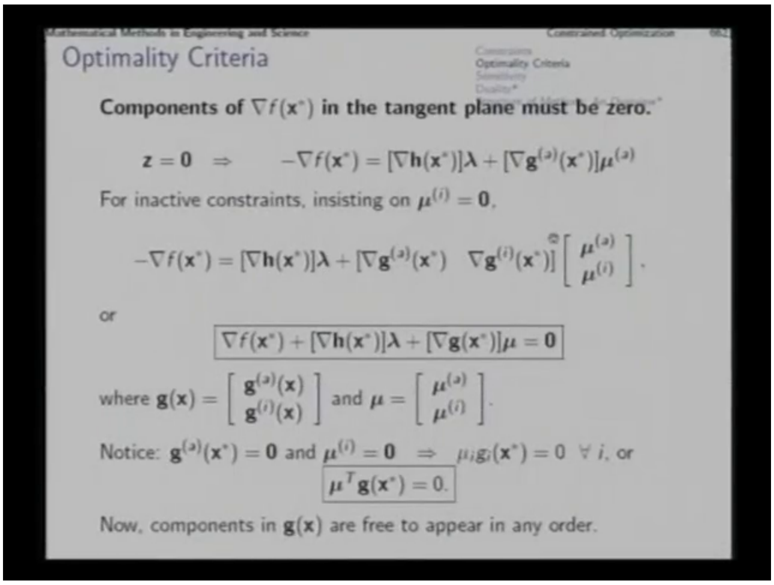


So, if this is the tangent plane then at this point if this is negative gradients then you would say that this vector is d z that is component of the negative gradient in the tangent plane and this vector is the rest of it, right; now if this is non zero say positive or negative this way. So, along all the directions d 1 d 2 d 3 d 4 positive or negative whatever if it is non zero non-zero component then you see that this is tangent plane in which every direction is feasible in the tangent plane every direction is feasible now if negative gradient has a non zero component in this tangent plane then moving in that direction we should able to reduce the value of the function right.

So, then this point cannot be a local minimum point because along this direction which is feasible we can have a positive component of negative gradients and we can reduce the function value along this direction. So, that is why for this point to be a local minimum point it is necessary that the negative gradient vector has no non zero component in the tangent plane; that means z work with 0.

So, this is what we can say if x star this current point is a solution to the non-linear programming problem the non-linear optimization problem minimize their subject to those constraints. So, this is what we can say about the components z lambda mu out of that we can say that z cannot be non zero if x star is a solution to the problem solution to the minimization problem.

(Refer Slide Time: 23:38)



So, components of gradient of f in the tangent plane must be 0 z is 0 so; that means, this negative gradient cannot have a component in the tangent plane and it must be orthogonal to the tangent plane; that means, the negative gradients should be completely describable in terms of the normal vectors gradient vectors of h 1, h 2, h 3, etcetera and active fellows from here active members from here. Now what else we can say; now this is going to be a little complicated keeping track of active constraint because suppose in a particular problem there are 7 constraints.

So, at a particular point suppose this is active this is active and this is active. So, at that particular point you will be taking g 1, g 2, g 4 and assembling them in this vectors g active and g 3, g 5, g 6, g 7; you will be assembling in g active at another point when you analyze; suppose there g 4 is not active, but g 7 and g 6 are active. So, then again you will be re assembling this. So, this problem is typically handled by saying that we will keep all the gradients here and all the gs we will consider together except that and that way if we keep rather than 3 if we keep 7 columns here than 7 values, we will say that let it has 7 values, but we will insist that if a particular constraint is inactive the corresponding mu should be 0.

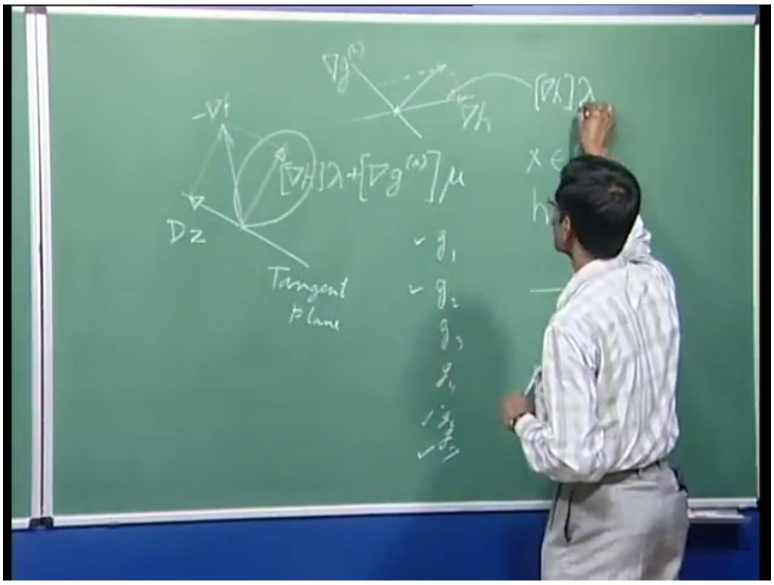
So, if g 1, g 2, g 6, g 7 are active then mu 1 mu 2 mu 6 mu 7 can be non zero g 3, g 4, g 5 must be 0 mu, 1 mu, 2 mu, mu 3, mu 4, mu 5 must be 0 that is corresponding to those constraints which are inactive let the columns stay in their place what mu's will be insist on being 0, if we follow that kind of an understanding that is called inactive constraints the corresponding mu's should be 0 then rather than having only active constraints here we can put active in active constraints in any order for that matter and say that inactive constraints in a we will have with corresponding mu's which are 0 anyway.

So, that they do affect this expression. So, extra columns will be there, but there corresponding multipliers will be 0. So, that way we can keep the entire g together and this long thing we can concisely write as grad h lambda plus grad g mu with understanding those mu's which correspond to inactive constraints will be 0 anyway and that gives us this requirement that is if a particular constraint is inactive then the corresponding g i is not 0. So, the mu i has to be 0 on the other hand for an active constraints g i itself is 0. So, mu i can be allowed to be non 0.

So, mu i g i the product is always 0. So, that is called the complementarity conditions that is between mu and g each in for each i the pair mu i and g i are complementary with each other if one of them is non zero the other one must be 0. So, together you can write like this also that is the sum of mu i g i is also 0 now this condition we arrived at from here in which we say that the negative gradient is a combination of grad h lambda and grad g mu only d z part is 0.

So, based on that we are arrived at this condition just take this negative gradient on the other side it goes as positive and this is the first order necessary condition arrived from the requirement that along the feasible directions in the tangent plane there should be no scope of improvement of the function value what about the directions in the cone of feasible directions when you explore that we say that now take this itself now take this itself in schematic diagram now that you have already decided that grad f will have no component in the tangent plane then that as expand.

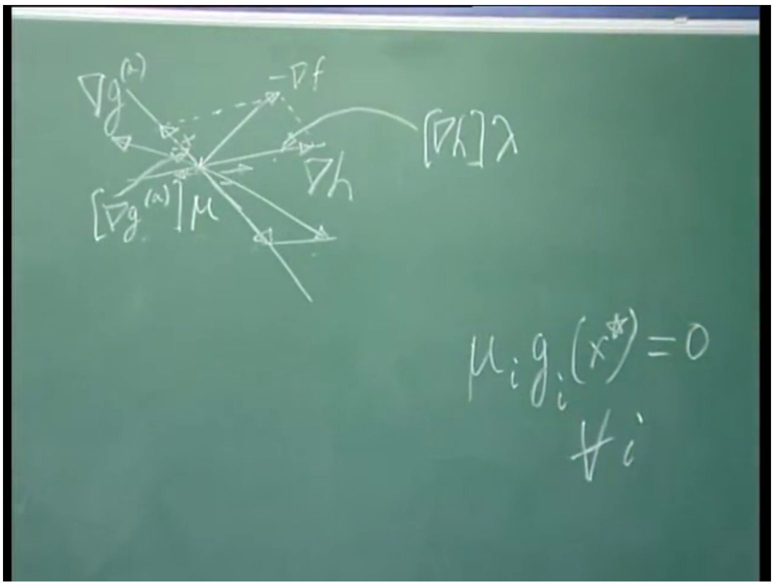
(Refer Slide Time: 28:29)



This part itself which is the direction which is the component in the direction normal to the tangent plane and resolve these 2 parts for this subspace only; now if we draw directions and this is the direction of grad h and this is the direction of grad g h and we are currently at this point; now in this along this and this directions the negative gradient can have a component.

So, if the negative gradient is this way now; that means that with the negative gradient will have one component along this.

(Refer Slide Time: 29:28)



Which is grad h into lambda and another component along this which is grad g a into mu now let us remove everything else to make this thing clearly feasible.

Now, note that grad h is a direction along which if you move then you will be leaving the constraint manifold and constraint surface you will be leaving and going out similarly if you try to move in this directions rather than this way this way and this way either way you move out of the constraint manifold in the direction of grad h or in the direction minus grad h; that means, that constraint manifold you are going to leave you are leaving tangent plane and going out of the surface this way or going out surface this way. So, therefore, minus grad f having a component along gradient of h or its negative will not be a problem because this direction is anyway infeasible.

So, the condition for this point for to be a local minimum is that along a feasible direction we should not be able to reduce the function value in this direction function value does reduce, but that is not feasible similarly along this direction if the grad f had component in this direction that is if grad f were like this then it would have a component along this direction in that case the function would be reducing this direction, but this direction is the not feasible. So, this is the space with lambda h therefore, the grad h.

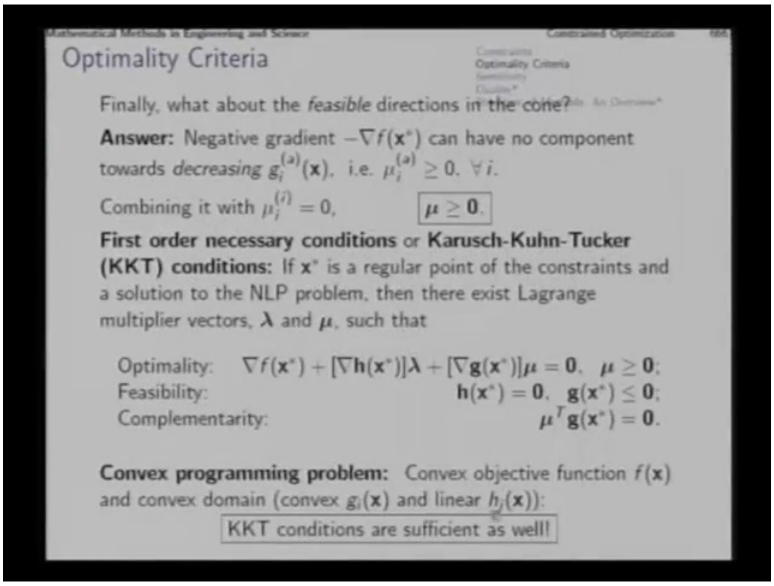
So, therefore, lambda could be positive or negative. So, negative gradient can have a positive component along grad h or a negative component grad h does not mater, but consider the situation is grad g at this point since g is at active constraint; that means, the value of g is 0 here this is the direction of grad h grad g; that means, here it is positive here it is negative now if the negative gradient has a component along this direction.

That means, along this direction it is possible to reduce the function value because negative gradient has a component along this direction, but that does not harm that does not stop this point from being a local minimum because this is direction in which even if the function value does decrease it does not matter because that direction is not feasible in this direction g will become positive grad g in this direction g is 0 here g is positive here on the other hand if minus grad f like this having a component along negative grad g that is in the grad direction if it as a negative components; that means, it is feasible this at this point g is 0 in this direction g will become negative; that means, the point will be feasible the constraint will be satisfied. So, this point is feasible.

So, if negative grad g negative gradient of the function minus grad f as a component in this direction that is as an negative component along grad g then this point cannot be local minimum because since it is a component of minus gradient negative gradient of the function along this direction function could be decrease and if this is a direction which is in the direction of negative gradient of g as well; that means, it is feasible that mean this direction will be feasible direction and this an direction at the same time a feasible direction in which the function value can be decreased.

So, if that happens then this point a local minimum so; that means, for this point to be a local minimum the components along the negative gradient of inequality constraints must be positive that is this is allowed bu; this is not allowed. So, mu cannot be negative this is what we get when we consider the feasible directions in the cone.

(Refer Slide Time: 33:52)



Describe by the active inequality constraint. So, negative gradient can have no component towards decreasing g; that means, mu should be all non negative ok.

So, corresponding to active constraints mu should be non negative and the in act for the inactive constraint we already know that there are all 0. So, together we can write this. So, all the conditions that we have collected till now if we summarized then we get what is called the first order necessary conditions or KKT conditions Karusch Kuhn Tucker conditions and the summary is this if x star is a regular point of the constraints and a solution to the optimization problem minimize f subject to g less than equal to 0 and h equal to 0 then there exist lambdas and mu's that is Lagrange multiplier vectors such that this and this that is grad f plus grad h lambda plus grad g mu is 0 with mu non negative which is which comes from optimality; optimality requirement and from the feasibility requirement which is part of the problem statement itself that is h x star 0 g x star is negative or 0 and complimentary condition is this.

This condition could as i well be written as that is for all i mu i g i that x star should be no sorry 0 that is complimentary condition that is mu's and g's are complimentary to each other that is if g is on 0, then mu must 0 that is for inactive constraints and if mu is non zero then g must be 0 that is mu can be non zero only for active constraints.

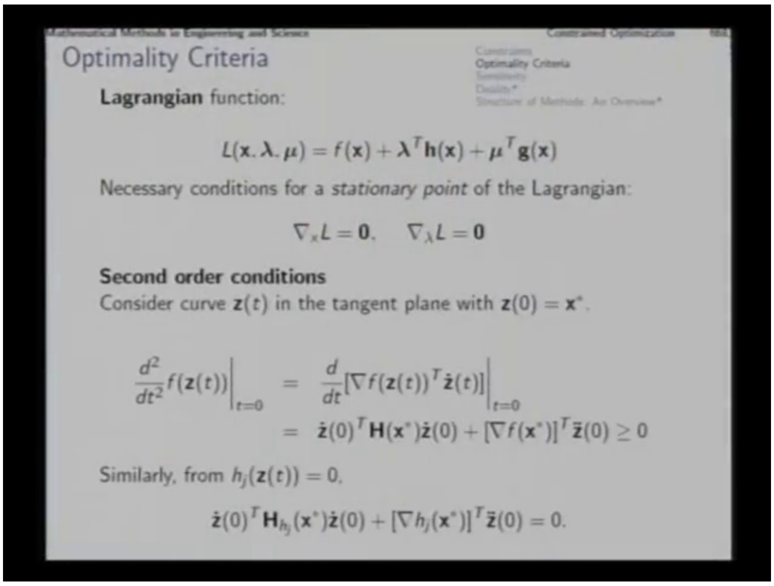
So, that is why this is called complimentary conditions. So, you find that here you have got n equations number of variables here you have got another bunch of equation number of equality constraints and here you have got another bunch of equation which are the which is the number of inequality constraints same number of unknown you have in x you have the number of ray unknown which is equal to n the dimension of the problem and you have got that many lambdas as many equitation's here and you have got that many mus as many equations here. So, the number of equations this, this, this and number of unknowns x lambda mu is same apart from that there are a few inequalities mu's are greater than 0 g are less than equal to 0.

So, any point x star with suitable lambdas and mu's that satisfy all these inequality and equality requirements they are said to be those point are said to be KKT points or corresponding to third points so; that means, they satisfy the first order necessary conditions that does not means that they are local minima that only means that if a point is local minimum then that satisfies all these conditions which some suitable lambda mu values there is one class of problems called convex programming problems in which the objective function is convex and the domain is also convex characterized by convex g i and linear h x that is equality constraint functions are linear and inequality constraint function are all convex.

So, that describe a convex domain. So, if you are trying to minimize a convex function in a convex domain then that is if your problem is a convex programming problem then this KKT conditions are not only necessary, but also sufficient, but for a general problem these are not sufficient for a general problem you need to consider the second order conditions also to be certain that a point satisfying this constrained conditions is a local minimum point.

So, the second order condition is a little complicated we will just make a brief over view of it and try to understand what it means rather than going in to the detailed derivation and before that we defined what is called the Lagrangian function.

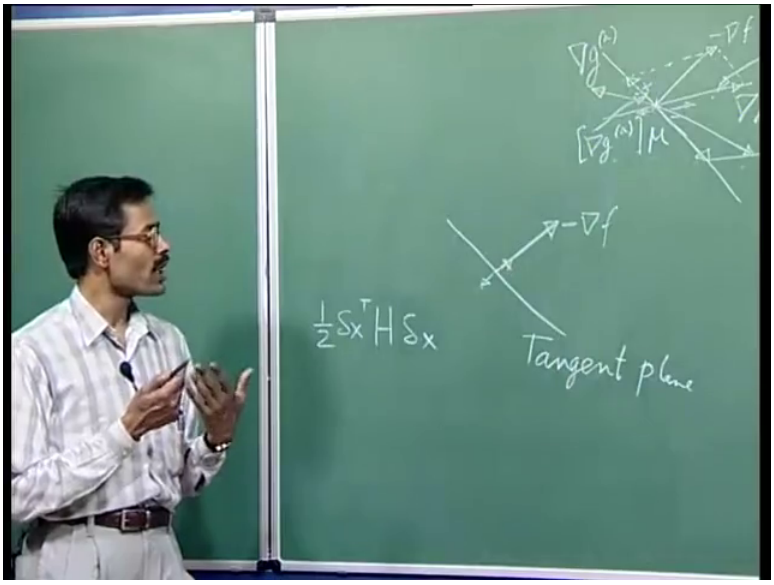
(Refer Slide Time: 38:39)



This function describe in with the help of the objective function the constraint functions from equality constraints and inequality constraints along with their Lagrange multipliers lambdas and mu's this function is called the Lagrangian of the problem and this is why these lambdas and mu's are called Lagrangian multipliers corresponding to equality constraints and inequality constraints. So, you will find that the first order optimality conditions that we worked out this; this equal to 0 this is essentially the gradient of this which respect to x this is gradient of x plus grad h lambda plus grad g mu that is what we had here.

So, gradient of lambda equal to 0 is the optimality condition. So, necessary condition for a stationary point of the Lagrangian also would be the same that is this and derivative with respect lambda will be actually given like this which will be nothing, but h x equal to 0 that is the feasibility condition equality condition equality constraint itself. Now the idea of the second order necessary and insufficient conditions is essentially the analysis of the second order change in the tangent plane.

(Refer Slide Time: 40:15)



So, we again re correct the tangent plane discussion that we had earlier it is the tangent plane then our previous discussion tell us that minus grad f must be perpendicular to it must be orthogonal to it because a component of negative gradient along the tangent plane would give us a first order change along the tangent plane which means that we could decrease function along the tangent plane along the feasible direction. So, since that is not allowed we are already sure that negative gradient of the function must be orthogonal to the tangent plane so; that means, along the tangent plane from this point there will be no possibility of a first order decrease no first order change because first order change of the function value in the tangent plane is 0 around this point what about the second order change.

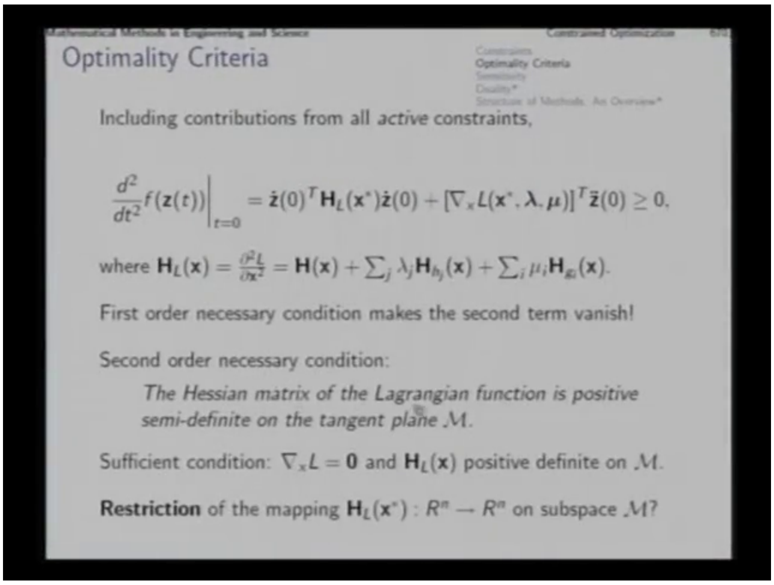
So, as you know the second order change is given by this. So, when we try to analyze the second order change we say that we do not care if there is a second order change which is in this direction or in this direction because in this direction we have already completed the first order analyzes and this points are this directions are anyway not feasible. So, for grad h say; this as this can have a component along grad h and a grad g for a matter. So, along grad h this movement itself is not feasible. So, we do not bother about that along grad g movement is feasible in one direction, but along that direction since mu itself is positive; that means, the first order change itself is positive so; that means, the first order done dominating the tailor series will not live any scope for the second order to be perceptible in the immediate neighborhood.

So, the second order analysis we need to conduct only on the tangent plane then we say that if we take a small movement small move if we make a small move along a tangent plane then how the function value is going to change up to second order level first order change along tangent plane is 0 anyway. So, what is the second order change if that second order change is nonnegative then we say that is the condition that is the necessary condition for the current point to be a local minimum point on the other hand if the second order change is positive then we would say that it is sufficient to ensure along with the KKT conditions if would be sufficient to ensure that is the current point is a local minimum point.

So, the necessary condition sufficient condition will be the same as the positive semi definite and positive define nature of this hessian matrix, but not in all directions only on that tangent plane; that means, the hessian should be in now here hessian that would be involved and that actually is the result of a little complicated analysis which we are omitting the hessian that we will get involve and that actually is the result of a little complicated analysis because which we are omitting the hessian that will get in involve here is actually not the hessian of the original function the objective function, but hessian of the Lagrangian.

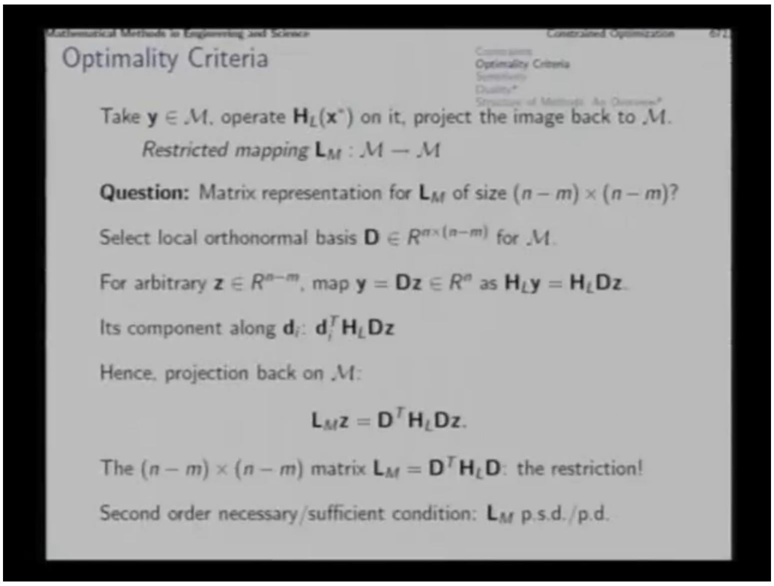
So, we say that the effect of the hessian of the Lagrangian on the tangent plane should be like a positive definite matrix the outside the tangent plane that is normal to the tangent plane even if it displaced a negative Eigenvalue that does not harm.

(Refer Slide Time: 44:16)



So; that means, the condition is that the hessian matrix of the Lagrangian function is positive semi-definite on the tangent plane m orthogonal to the tangent plane even if it do not behave in a positive semi definite fashion even if it behaves with an indefiniteness, it does not matter the only requirement is on the tangent plane. So, this is necessary conditions positive semi definiteness of the restriction of hessian on the tangent plane sufficient condition is that it is positive definite. So, if we want to analyze that a hessian matrix is positive definite on a particular subspace then for that.

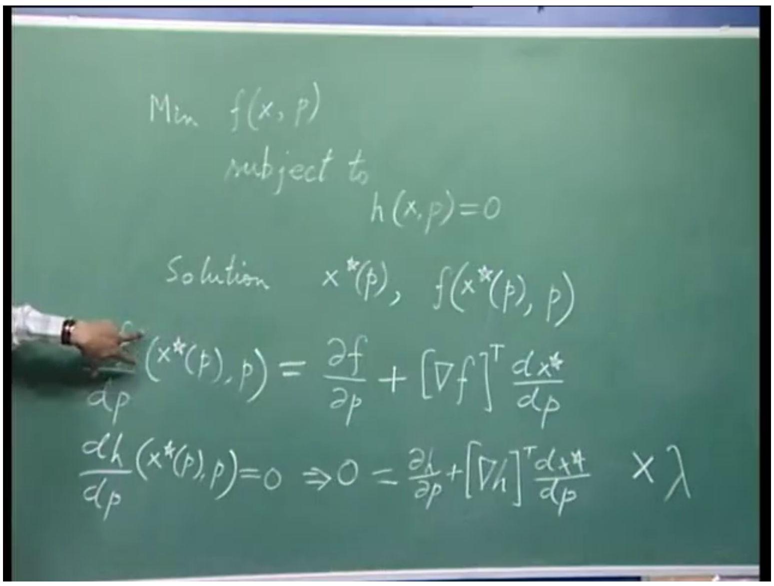
(Refer Slide Time: 45:04)



We can construct this matrix that is if d is an orthogonal; orthonormal basis of the tangent plane and h l is the hessian of the Lagrangian function then d transpose HLD will be a mapping will be a symmetric mapping within the tangent plane that is it maps the tangent plane to itself and we can examine the positive definiteness of this matrix which is smaller n minus m by n minus m the other subspace the orthogonal complimentary subspace of m dimensions is removed out of it. So, we can consider positive definition of this.

So, this is the second order condition. So, along with KKT condition this being positive definite is sufficient for the current point to be a local minimum for all problems even for non convex problems now long back in the first lecture of optimization on optimization I mention to you that at the beginning of the optimization process some of the variables of the problem can be considered constraint for the analysis and frozen those are called the parameters after the solution is found typically we would like to examine whether freezing their value was a good idea that is we would like to examine the sensitivity of the solution to those parameters. So, how do you how do we analyze the sensitivity consider.

(Refer Slide Time: 46:48)



Considered this NLP problem non-linear programming problem for simplicity I have kept only equality constraints and not inequality constraints, but the them applies for inequality constraints also again for simplicity I have considered only one parameter which is kept fixed now suppose for solving this problem in the beginning we assigned a value 2 p and hence solve the problem and then we got the solution as x star a point in the solution space and the corresponding function value.

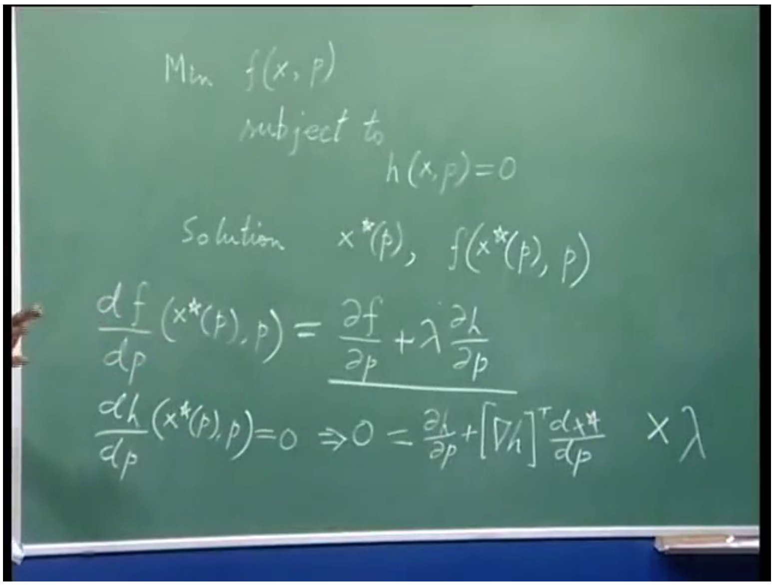
Now, note this when we find the function value we find it at that x stars which means that with that constraint value p in the beginning consider another important issue that as we gave a particular value as we assigned a particular value to p, we got this optimal point we got this minimum point if we had given another value of p, we would have got a different minimum point, if we had continuously varied p that is p equal to 1 p equal to 1.01 p equal to 1.02 p equal to 1.03, then continuously we would get some x star some other x star some other x star some other x star that way, we can considered that this x star is actually a function of the p that we give right here as variables of the problem x and p are independent variable, but x star is the optimum point which has been arrived at through a long process of optimization after assigning the value of p. So, as we keep on changing the value of p the resulting optimal point will keep on changing. So, that way x star is actually a function of p and therefore, when we finally, evaluate the function we will be actually evaluating this right.

Now, we want to find out if we change the little bit then how this would change a more importantly how the corresponding function value would be change which we want to minimize that is we want to find out d f by d p at what rate the change of p affect the change of f you know how to find the total derivatives of this kind of function f will depend on p in 2 ways one directly and the other through x star p. So, the total derivative would be this partially derivative plus a derivative reflecting the dependence through x star dependence on p through x star.

So, that will be grad f that is derivative with respect to x multiplied with how x star itself changes with p right. Now given the function here it would not be difficult for you to find out this, but then how to find out this yes you could solve the same optimization problem for another p and then get this, but there is a simpler way to solve for it and for that you note this h x p equal to 0 at the solution point x star p, this is satisfied right if you make a small difference small change in p and then for the entire problem then you will get another x star p, which is also satisfied the new h corresponding to the new value of p; that means, whatever changes in p you made h of x p still remain 0 because it will be the mu x stars will be feasible which respect to the h x to the mu h define with the help of the mu p right.

So, this will still satisfied; that means, as you change p h remains 0; that means, d h by dp of this is 0 always. So, that will mean 0 equal to derivative of this we construct in this same manner and now what you can do is that multiplied this lower equation with lambda and add to the upper equation on this side 0 will be multiplied with lambda and add to this; that means, on the left side there will be no change on the right side, there will be some change what will that change lambda into this will be added here and here what will be added d x star by d p is constraint is common here will be added grad f plus lambda grad h. So, grad h plus lambda grad h is the first order necessary condition for the point x star to be minimum.

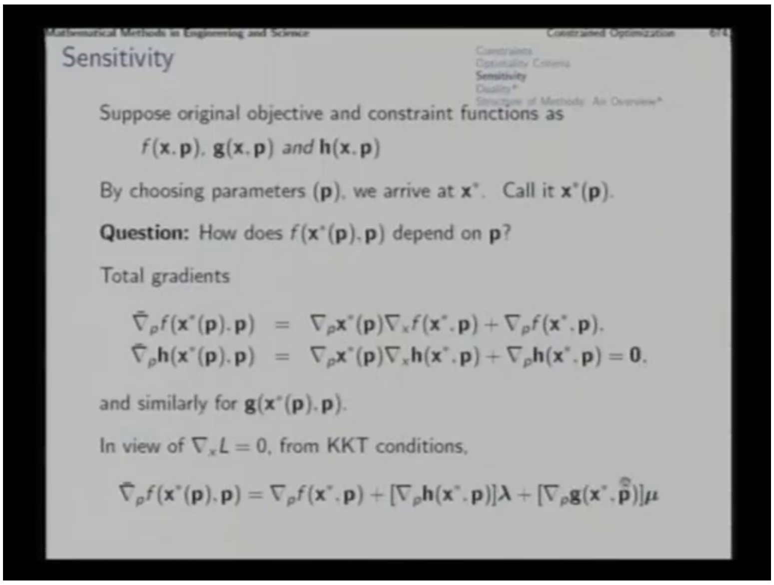
(Refer Slide Time: 53:39)



That equal to 0 is the condition; that means, lambda times this added to this we will make it 0 and lambda times this added to this will make it grad f plus lambda grad h that is del f by del p plus lambda del h by del p. So, here you find that analyzing the sensitivity is actually not that difficult.

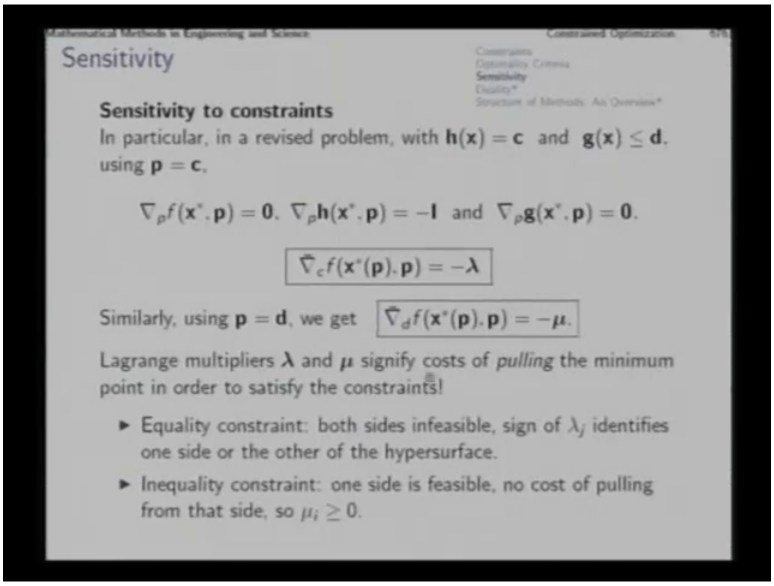
So, you can construct this partial derivatives from the given functions and analyze the sensitivity of the problem of the solution to the parameter p.

(Refer Slide Time: 54:14)



These things when formulized in terms of large number constraints and large number of parameters give these long conditions similarly you can check the sensitivity you can assuming the sensitivity with respect to constraints also.

(Refer Slide Time: 54:32)



And if you do that then you find the sensitivity of the functions of the solutions to the constraint is just given by lambda mu and that way you can say Lagrange multiplier lambda and mu signify costs of pulling the minimum point in order to satisfy the constraints that is lambda and mu Lagrange multiplier are cost for satisfying the constraint beyond this we will discuss in the next lecture in which we will consider duality and study the structure of non-linear optimization method.

Thank you.