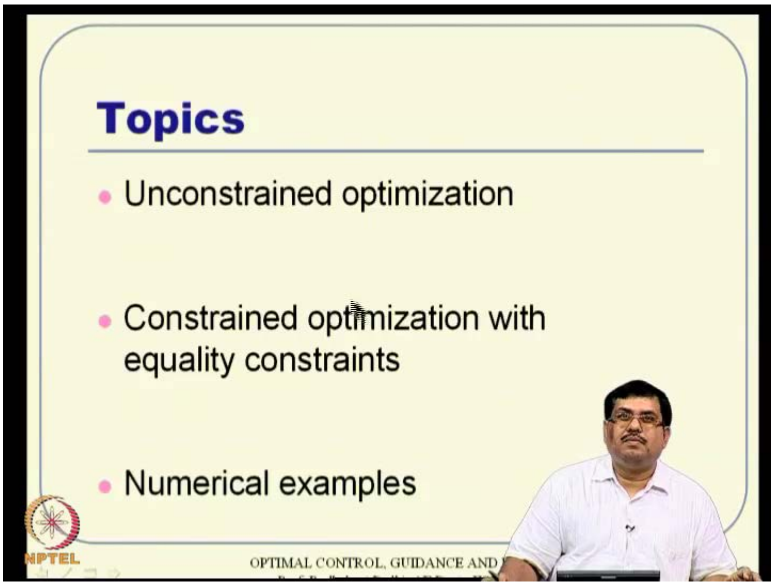
Optimal Control Guidance and Estimation
Prof. Radhakant Padhi
Department of Aerospace Engineering
Indian Institute of Science, Bangalore

Lecture No. # 05
An Overview of Static Optimization - II

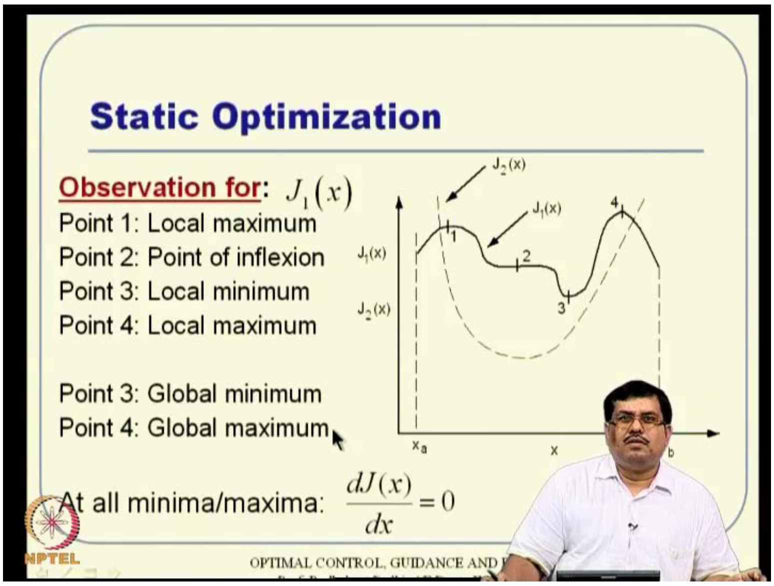
Hello everybody. Let us continue our lecture series in this course, optimal control guidance and estimation. We are here at lecture number 5, where we are talking about some sort of overview of static optimization and this is part 2 of the lecture actually. So, very quickly, I mean we can just see what all we discussed in the last lecture before coming to this one actually.

(Refer Slide Time: 00:38)



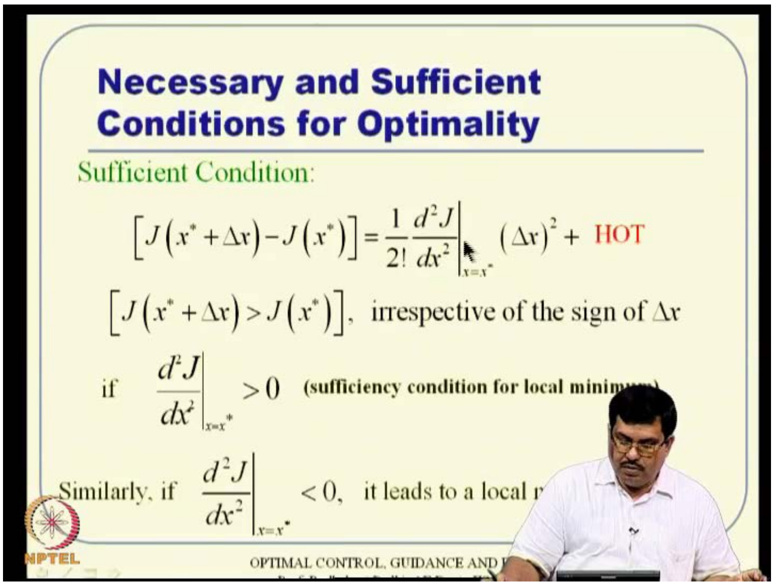
So, we talked about some sort of unconstrained optimization followed by some some constrained optimization with equality constraints and then, saw some numerical examples on the way.

(Refer Slide Time: 00:48)



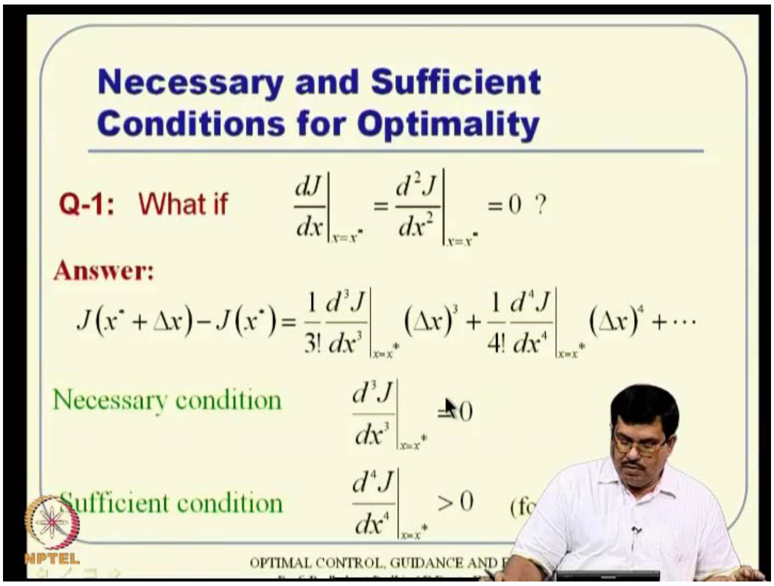
So, these are all what we discussed, I mean taking, getting motivated from a simple, I mean simple curve sort of thing and analyzing those points.

(Refer Slide Time: 01:08)



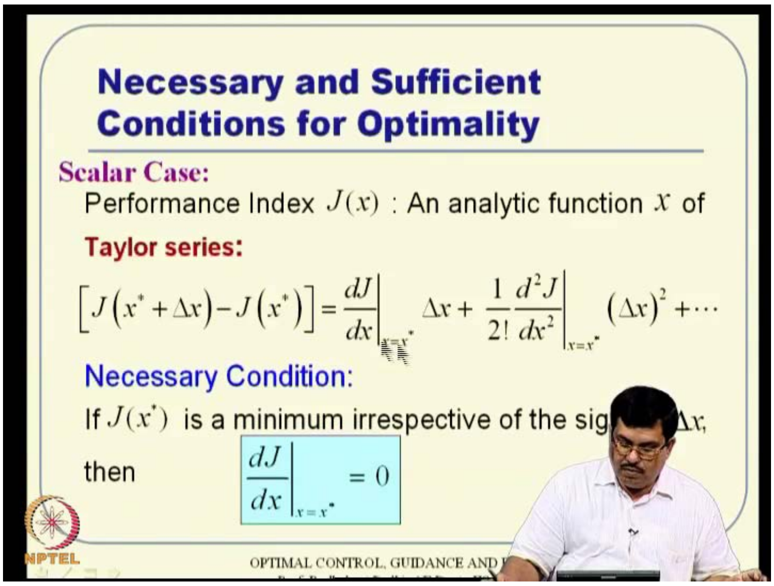
Then, finally analyzing whether the derivative is 0 or not. Then, it turns out after rigorous analysis through Tailor series that the first derivative needs to be 0. Then, we continued further with second derivative, third derivative and so on.

(Refer Slide Time: 01:16)



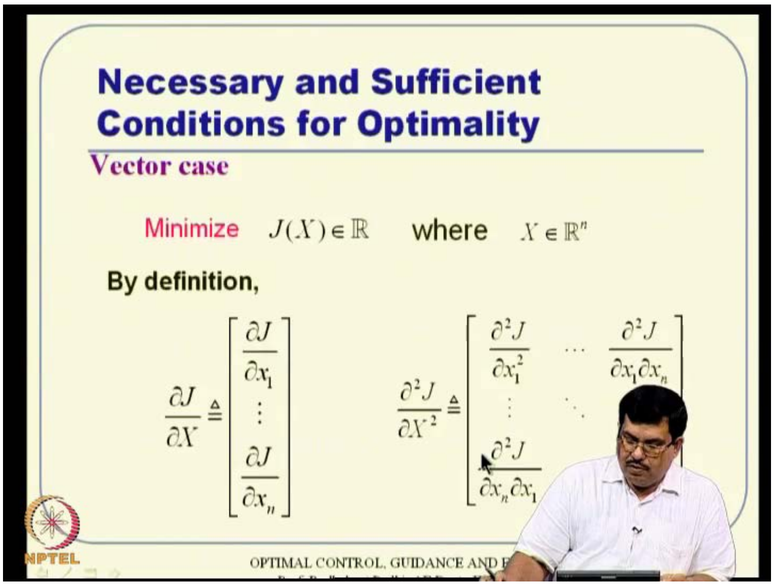
You just tends out that for for necessary conditions, all odd derivatives needs to be 0 and for sufficiency condition can be arrived at with even powers and all that, with the sine sensitivity all that. So, all these things we have discussed in the last lecture anyway.

(Refer Slide Time: 01:03)



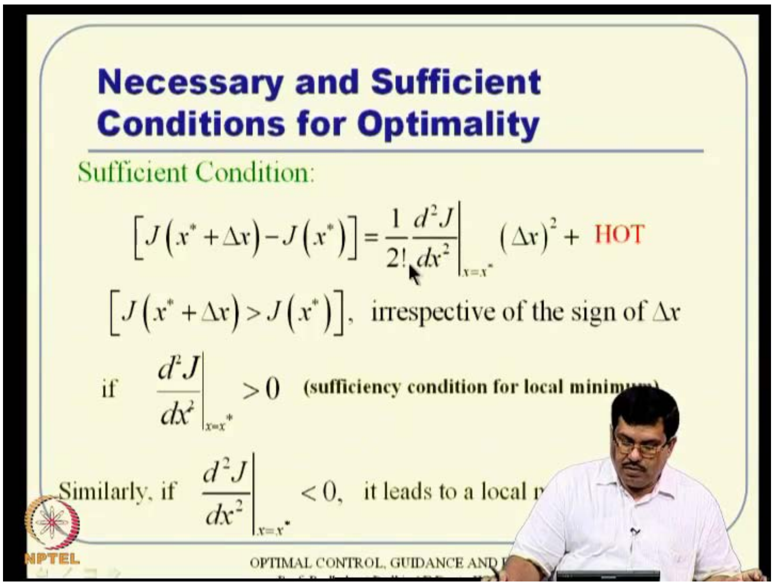
We saw some examples on the way and then, followed to this vector case where the decision variables are actually x 1 to x n number of variables.

(Refer Slide time: 01:28)



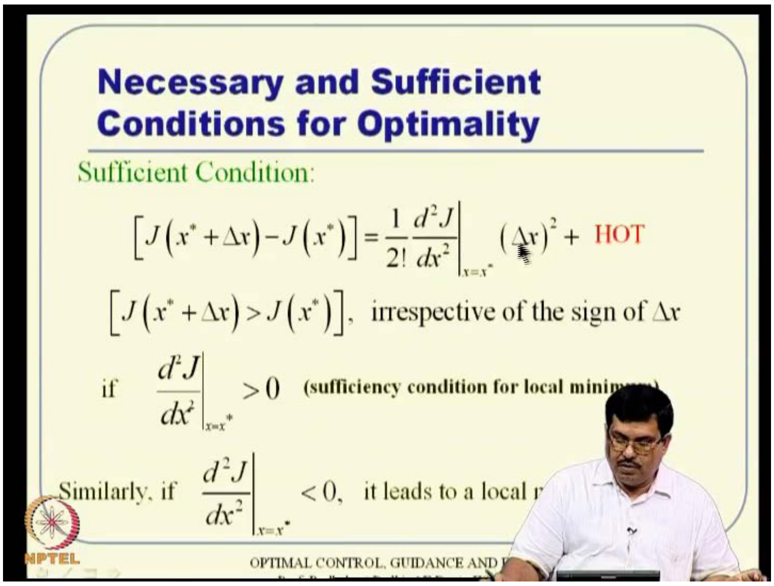
In that case, you will have to define some radiant vector and then, asymmetric things like that and again, they are using the same idea of Ruler series.

(Refer Slide Time: 01:07)



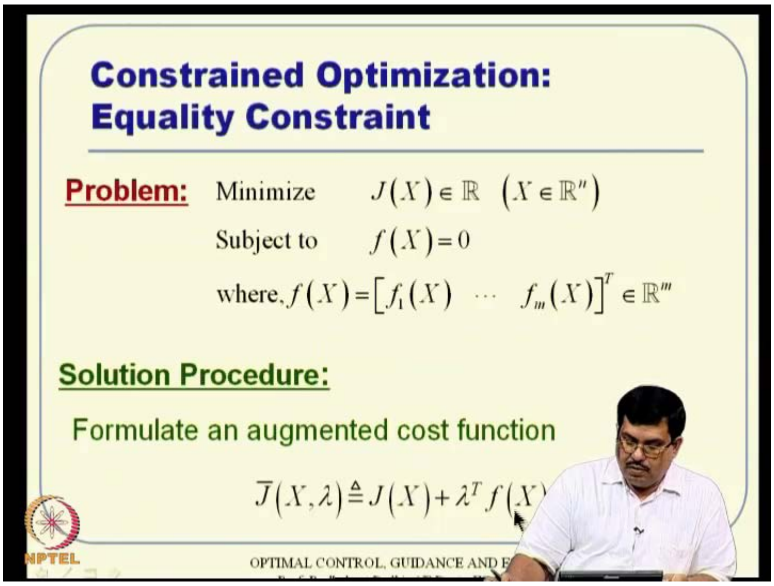
Then, we arrived at a conclusion that if the, so the necessary condition the radiant vector needs to be 0 and sufficiency condition can be arrived by looking at the positive definiteness or negative definiteness of this isometric actually and again, followed with some examples and all that it will appear.

(Refer Slide Time: 01:09)



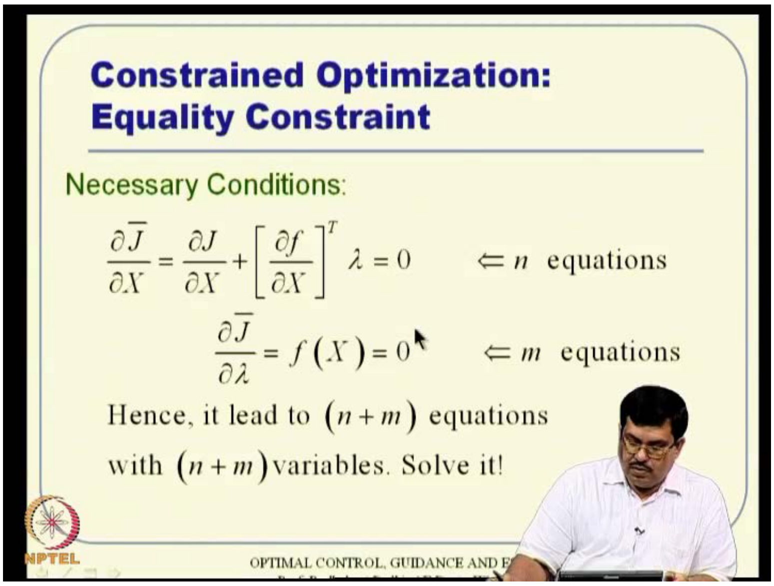
Then, we moved on to some constrained optimization with equality constraint and then, we told that this particular problem, the minimize J of x subject to the f of x equal to 0 is equivalent to minimizing this J bar of x and lambda, where both x and lambda are considered as p variable.

(Refer Slide Time: 02:09)



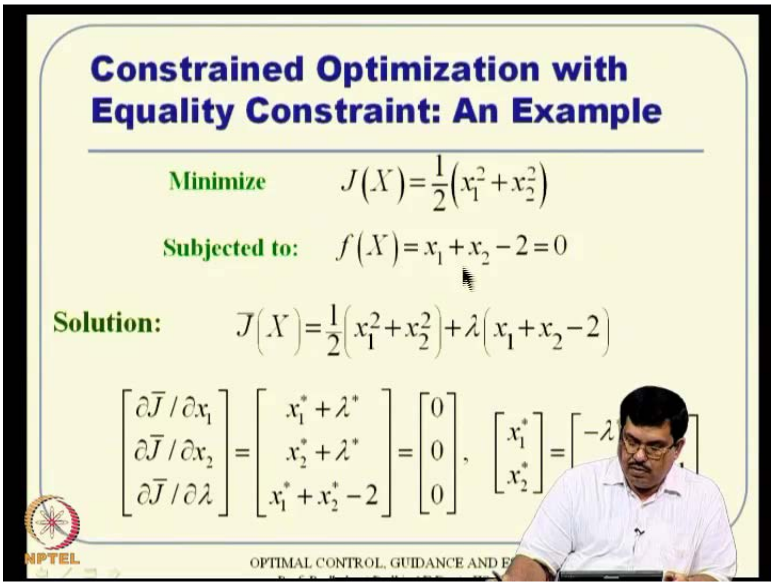
We constrained this function and then, use the analysis tools that you know for unconstrained optimization for J bar. With that in and then, connect with these 2 necessary condition again and went through the example problems.

(Refer Slide Time: 02:23)



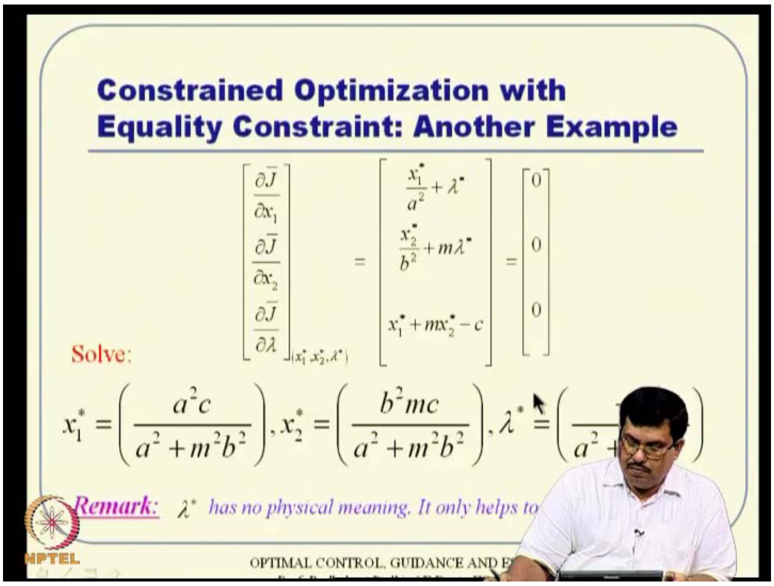
Finally, I mean we went some sufficiency condition as well and told that you have to construct this kind of a matrix. Then, see the determinant and interpret that is the function of only sigma's and analyze the the sine sensitivity of all these sigma's actually.

(Refer Slide Time: 02:29)



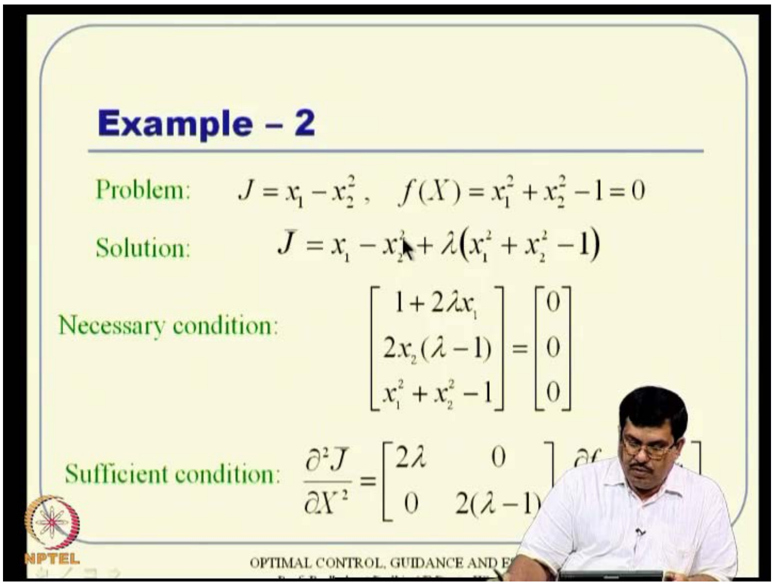
If all of them appear to be positive, it is the least minimum condition and all of them appear to be negative, then it is a maximum condition actually.

(Refer Slide Time: 02:28)



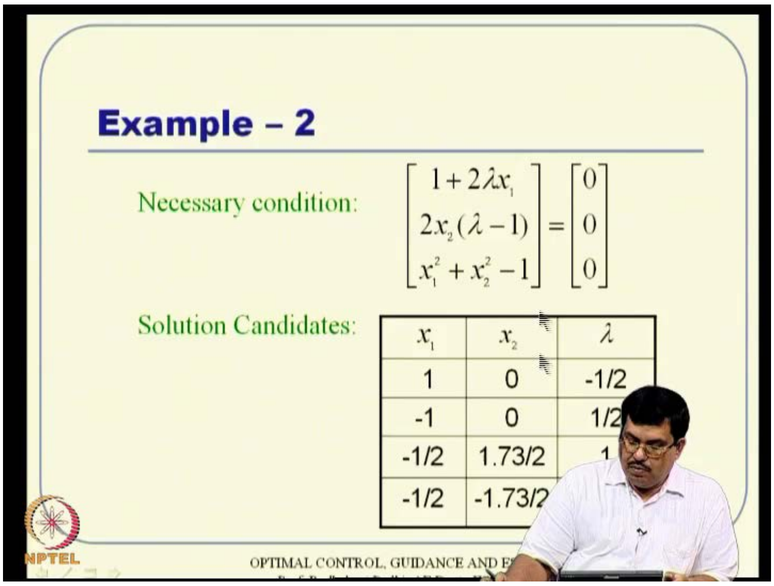
Again, we demonstrated that using some some examples and all that actually, all right. Then, we had a good example at the end, where took out this x 1 minus x 2 square and then, constraint values of a non-linear equation and it will ended up with some multiple cases.

(Refer Slide Time: 03:01)



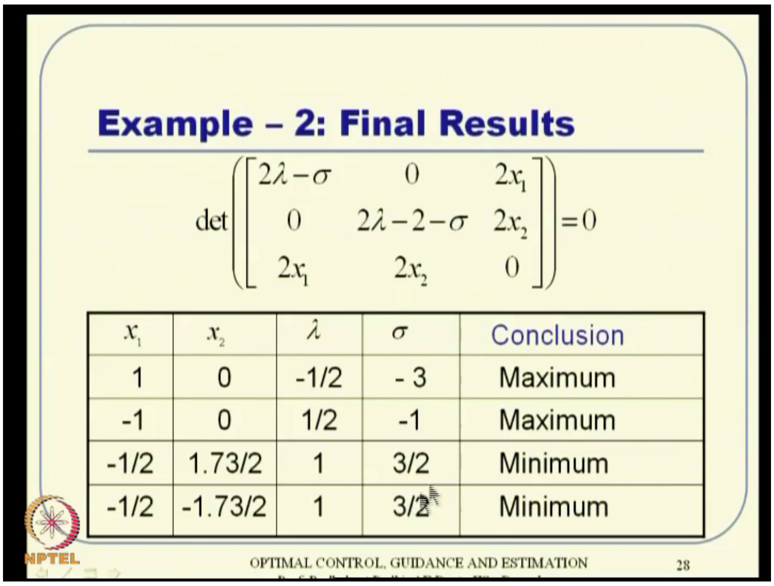
Depending on each other cases, we had analyzed this particular equation. Then, using that case one, what you got here? We got some some results that sigma equals to minus 3 and hence that leads to maximum and all that actually.

(Refer Slide Time: 03:10)



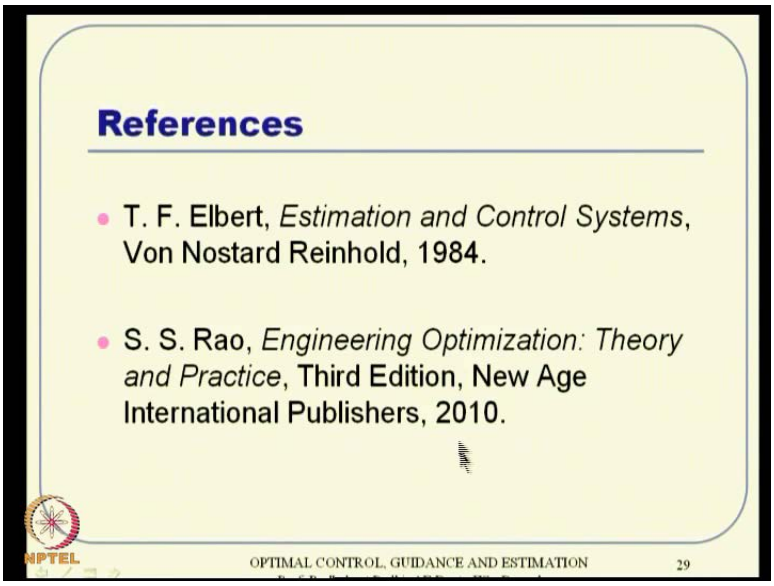
So, for considering and analyzing all these things for all the roots, we had a conclusion that these two will lead to maximum and these two lead to minimum. If we really want to have global maximum and global minimum ideas, then at this point, you have to evaluate the J, both both the points and whichever J appears to be maximum is a global maximum.

(Refer Slide Time: 03:24)



Then, these two you can evaluate the function J and whichever appears to minimum, appears to be a global minimum. So, that was the idea there actually. Then, we write some references and then, I suggested to look at some books and all that. So, that was our last lecture.

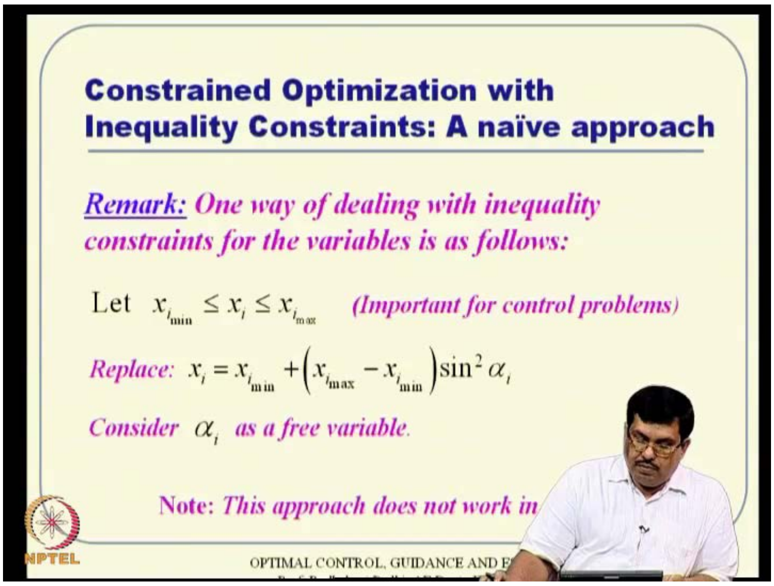
(Refer Slide Time: 03:48)



Now, coming to this particular lecture, we will talk about constraint optimization with mainly inequality constraints and also before moving on, we will also have a glimpse of, a small glimpse of numerical optimization techniques. Especially, this this stiffest designed method and this Newton's (()) like that usually we will see something. Those are the things that break one of any optimization techniques including dynamic optimization. That means optimal control areas. So, we have to, we will just have a glimpse of it.

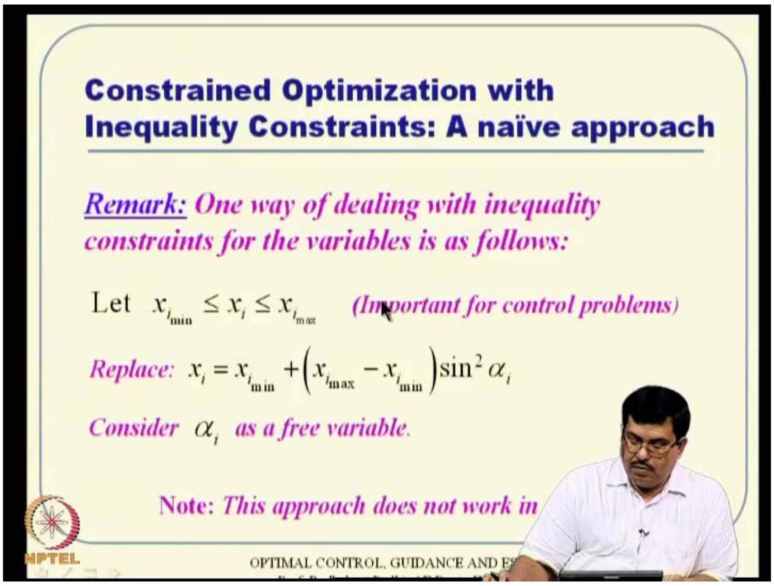
Well, again this is not a course on static optimization. So, we will not deal over there. We will not take a lot of time to kind of analyze various techniques and all that. At least good to have some idea of what is numerical optimization actually. Then, we will follow up with some numerical examples before winding up this lecture. All right. So, constrained optimization with inequality constraints, that is what we want to see first analytically and then, we will see some numerical things and all towards end of this lecture actually.

(Refer Slide Time: 04:58)



So, one very quick way of an inequality constraint approach or other I will call that as a naïve approach, say something like this. We will visualize a problem where if this variable, decision variable, all of them are constraint like this let say. So, all of x i or x 1, x 2, x 3 all that are the constraint between their corresponding minimum and maximum values. Then, how do you handle that problem?

(Refer Slide Time: 05:20)



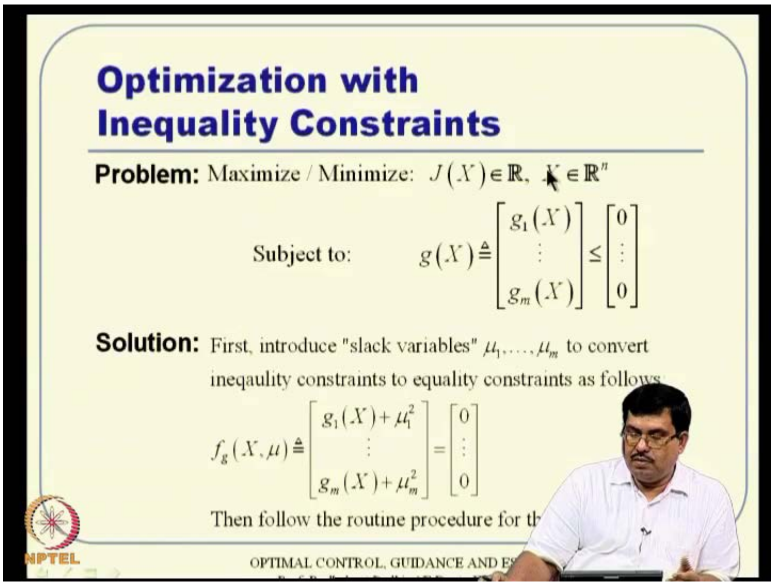
One easy idea or other make approach that turns out be, like ok you will construct some, I mean you just replace this x i, where something like this, where x i, I mean now you interpret alpha i as a free variable. That mean, if one no matter whatever is the value of alpha is, note that sine square alpha is always bounded between 0 and 1. So, it happens to be 0, then x i turn out to be minimum and if happens to be 1, then it turns out to be, these 2 will be cancelled out and it will turn to be maximum value.

So, no matter whatever the value is of alpha i, your x i will be bounded between these 2 things actually. So, that is the one idea, but also note that we have actually introduced a nominal transformation here and this x i and alpha i are not really linearly related. You have this human less problem of like so many local variables and so many values of alpha i, which will give it to that actually because sine square alpha i is not here. I mean, it is not uniquely valued and all I mean, it is not a unique function sort of things like that.

So, we will end up with a huge amount of problem here because most of the time, if you just do it and carry of numerical optimization which typically we have to do at the end of the day, then it is not a very good way of doing that. Also, a simple problem can get translated into a very complex problem through this early on transformation if you want to do it through analytical tools actually.

So, this is even though, it is a very lucrative approach and you can probably try out at the first go, It is not really universal solution that we are looking for actually. So, this approach in general does not work. If you can try it out, we do I mean in a given problem you can try it out. If it works out, nothing like that, but in general, it does not work out very well actually. So, we are in a process of pointing out for better approaches usually.

(Refer Slide Time: 07:09)



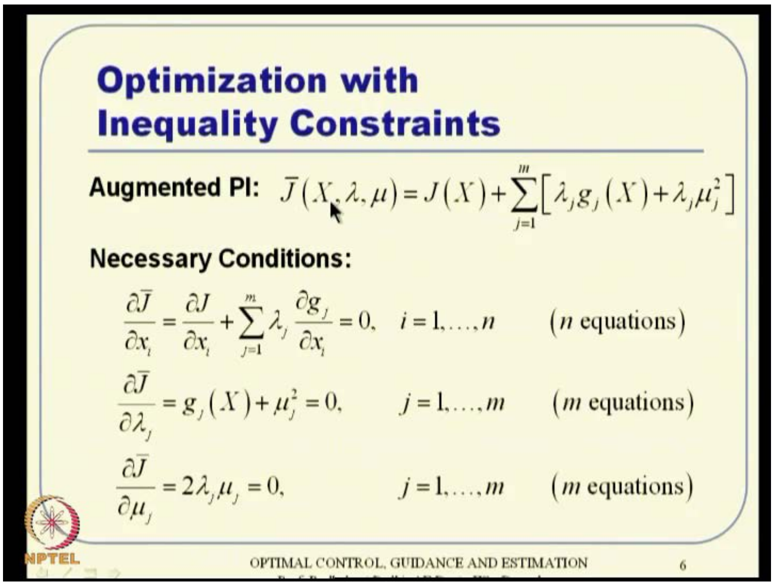
So, let us see that, I mean first the problem definition. What you really want here is minimize or maximize this J of x. Again, the cost function or objective function is always a scalar, but the decision number of free variables can be, I mean n dimension. That means, we have n number of free variables to choose from in a difference in combinations and all that, but this particular objective function is now subject to this, this constraint g 1 to g m less than equal to 0.

Now, remember this also inverse to this inequality constraint because we are considering less than equal to 0 here. If it equal to 0, then it all satisfy that. So, that means, what we discussed in the previous lecture in inequality constraints is a kind of embedded over here actually ok, but we have talking in more general thing. That mean, first first couple of equations inequality constraints. We are assured that constraints are not really equality; they are inequality constraints in a way actually ok.

So, how do you handle that in general? Now, consider I mean this particular problem that this is or less than equal to 0. So, the one approach is to introduce something called slack variables, ok mu 1 to mu m, all all are considered as slack variables to convert the inequality constraints to equality constraints because they are less than 0. So, I can always add a number there and just to make sure that we have positive quantities, I will added something like mu 1 square, mu 2 square up to mu m square actually ok.

So, if I do that, it turns out that it is is nothing, but inequality constraints problem actually. Then, the idea is to follow up the routine procedure for the equality constraints because we know how to handle the equality constraints analytically. We we can carry further and then, proceed further actually. However, not that is even though, it is less than equal to 0, not that we really do not know how far it is away from 0. So, these values of mu 1 to mu m is typically not known. I mean that even though symbolically you can write it that way. Using it for inferring something is not really very nice because you will not be aware of any idea of this number actually.

(Refer Slide Time: 09:45)

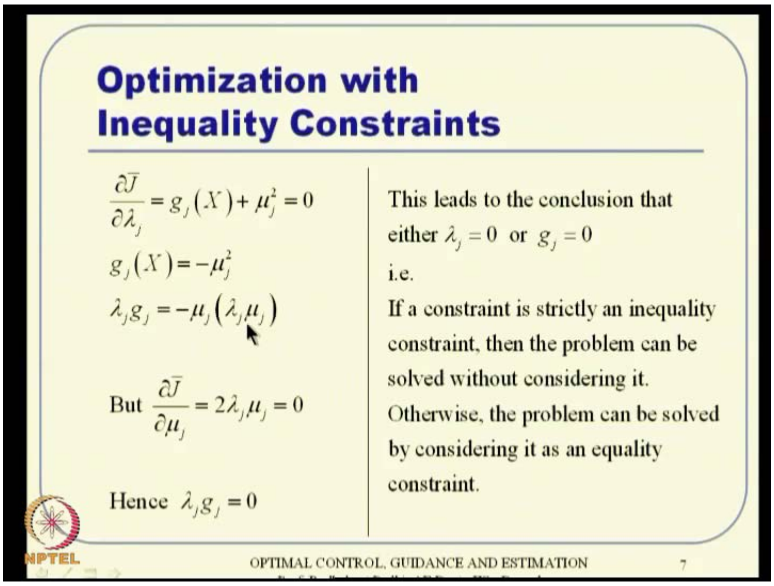


So, we want some conditions which are independent of at the final relations actually. So, let us see whether you can really do that or not actually. Anyway, so this is a problem of maximize or minimize J of x now subject to this constraint, which is equality constraint ok. So, this is what you are doing. So, now, if you look at this one and then, equality constraint, we know how to handle that. So, first to construct J bar and now J bar is a function of x and lambda, but also these mu is free variable, slack variable actually. So, this actually is a function of all these quantities, x lambda and mu ok.

So, this is constructed that way. So, as far a necessary condition is concerned, I mean concerned, we have to take take partial derivatives with respect to all these variables, x lambda and mu. X is n dimensional quantity whereas, lambda and mu you remember lambda and mu are given by dimension of the constraint equations and that is typically m basically. So, we will consider these n derivates here coming from x i, partial derivatives that has to be equal to 0 and then, finally followed by this del J bar by del lambda J which is like free variable of lambda considering that free variable lambda here here we talked about that actually ok.

So, this is the second equation and then, we have a free variable with mu and that mu turns out to be like this del J bar by del mu, that has to be equal to 0 in case. So, these conditions what you are getting n, n equations, m equations and m equations, that is n plus 2 m equations have to be solved to get that basically and we get the idea there actually. I mean we will get the solutions what we want to achieve. We have to utilize the equations now. How do you do that? So, let us analyze little more. So, this equation let us start with for second one.

(Refer Slide Time: 11:07)



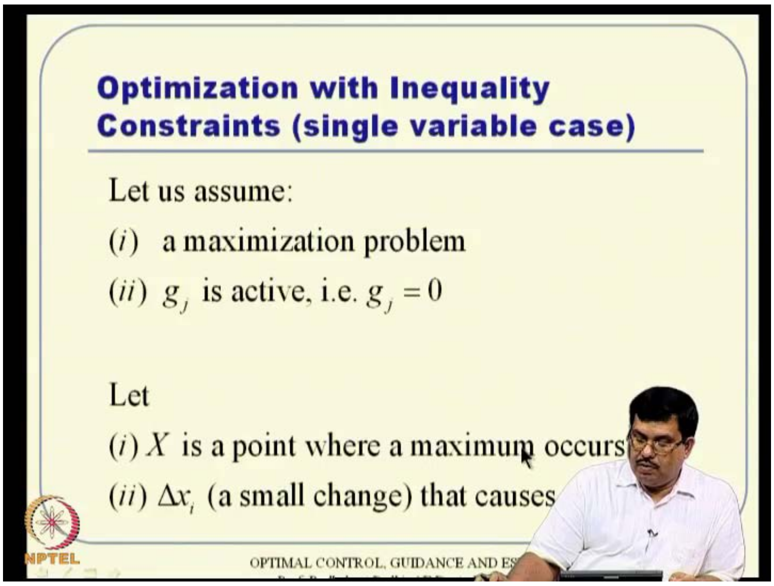
So, this turns out to be 0. Hence, g j of x, I can write it to a negative of u j square actually. Now, if I multiply those sides with lambda j, I can write it that way, but remember that this particular equation comes under now lambda j mu j equal to 0 now. So, what it tells us this this quantity 0. Hence, lambda j into g j has to be 0 ok. Just if if I take this equation and get g j x, the negative mu square mu j square, multiply both sides lambda j and then, use a third equation. What I am having here? So, once I get that, then it tends out that g j lambda j is nothing, but is 0.

So, this leads to the conclusion that either lambda j 0 or g j 0. Now, if lambda j is 0; that means, the constraint is kind of inequality I mean strictly inequality constraint, which means the constraint is not acting. That means, it is actually a kind of and you can visualize as if the constraint is not there actually, lambda j 0 you do not because that is coming through this constraint equation but because you have a constraint equation, you have that and then, it is coming through that. So, if lambda j is 0 for a particular thing, you can visualize very well as if the constraint is not there actually, but that is a mathematical interpretation actually.

So, if otherwise g j is equal to 0; that means, it is active constraint. Now, if we go back to here, the g j becomes equal to 0, it becomes an active constraint. So, if the problem is, I mean either the constraint g j is not there and it is as good as not there or it takes as a kind of active constraint actually. So, that is the conclusion here which makes a lot of sense actually. If I mean, if g j is strictly less than 0, then we can very well forget it actually and then, try to optimize only the (()) actually or anyway, this is the conclusion here. So, lambda j g j equal to 0, that leads to the conclusion that either lambda j 0 or g j 0. That means, putting in English words if a constraint is strictly an inequality constraint, then the problem can be solved without considering it. Otherwise, the problem can be solved by considering it as a strictly equality constraint. So, that is the conclusion here, but it also gives us a procedure to solve now.

Let us say, how to do that? Now, notice that this equation has nothing to do with the mu actually. This this equation is simplified to deal with lambda's and g's actually. So, really we do not need to know the value of mu now, all right. So, now to consider this this analysis further, let us talk about a single, I mean typical case, the typical maximization problem. I mean remember it is not minimization problem we are considering, we are considering the maximization problem and g j is active. That means g j is equal to 0 actually.

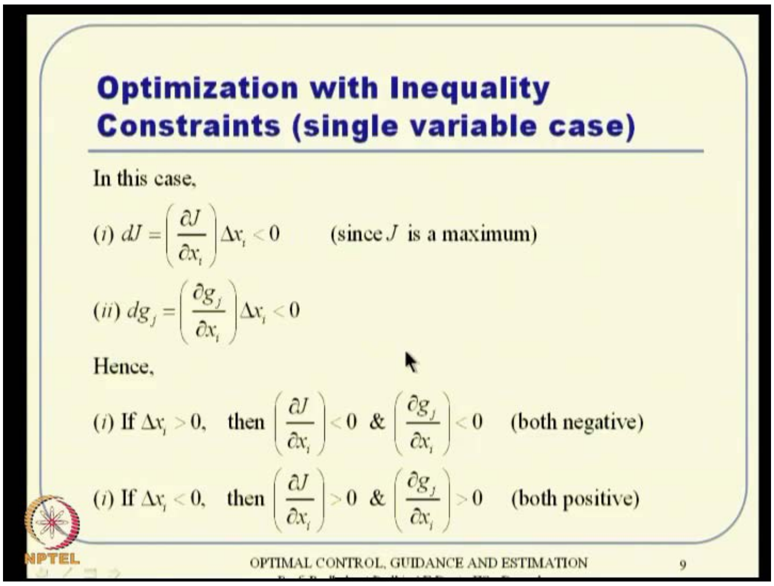
(Refer Slide Time: 13:42)



So, how do you conclude in that particular situation, what is happening there? Now, let us assume that x is a point where maximization or maximum occurs, maximum of the cost value occur and delta x i is a small change in x i direction that causes g j to be strictly negative. That means, the moment that I will go a little bit in that particular direction, my g j becomes inactive actually. In that situation only, we have got the optimization optimal point basically that happens, all right.

So, at that point the maximum occurs and it may in a particular direction of delta x i g j tends out to be strictly negative. That means, from active constraint, it becomes inactive constraint actually. So, that is the point where maximization occurs. Now, how do you handle? I mean how do or what do you infer from there actually?

(Refer Slide Time: 14:50)



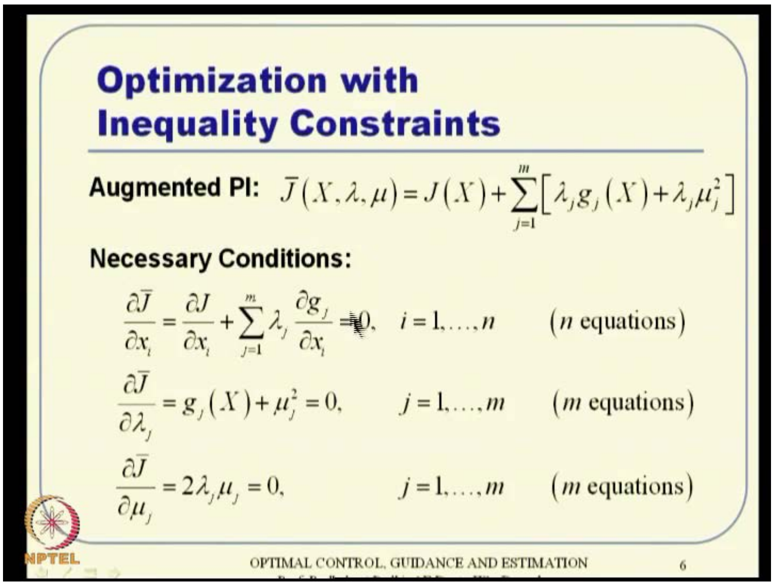
So, if we go back to d j, it tends out to be like that, del j by del x y into delta x y actually. Remember, it is g j is active actually. So, this is d j is nothing, but del j by del x i into delta x i. This is less than 0. Since, g j maximum point actually in that sense and then, because this particular direction, any direction I go from that particular thing, this has to decrease. That is a pre-optimization problem, but in a constraint optimization sense, it may not happen. We are not allowed to go to one side, but whatever direction we are allowed to go, if we go to any direction, any direction within that allowable zone, then we are suppose to decrease, I mean the cost function also decreases actually.

Then, this also becomes true because d j and d of g j are nothing, but something like this. In that particular direction turns out to be less than 0 from this inference actually. Thus, g j tends to be less than 0 actually here. So, from that it it turns out that this this condition is also true here. So, now, what you infer from this actually? So, if if these two conditions are true now, let us consider delta x i is a positive quantity. It can very well be negative quantity which is to be discussed here next.

If it is positive quantity, then what happens? We have to satisfy these conditions anyway. So, that means, del j by del x i and del g j by del x i has to be negative quantities. Then only, this this equation will be holding true. Now, if this happens to be a negative quantity, again opposite is true. Now, these, both of them have to be positive quantity now. I mean the noticed, I mean the thing to note here is, both of them has to be either negative or both of them has to be either positive. That is the conclusion for a maximization problem, remember. Now, if it is a minimization problem, this sign will alter. This will become positive here and then, their inferences has to both held to be like kind of an opposite sign each other and all that actually.

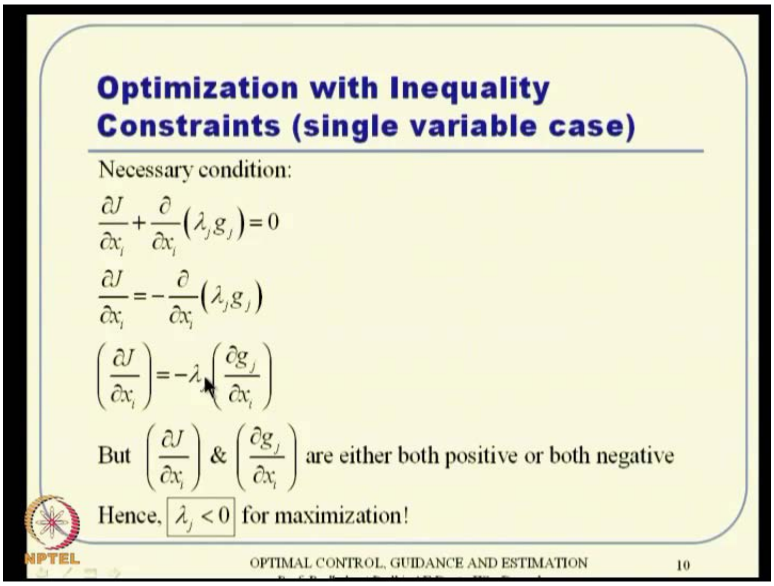
Now, let us not worry about that. We are considering a maximization problem and this is what happens. Both the things will take the same sign basically now. What do you conclude from there? Now, you will go back. So, for we have been dealing with these 2 equations. We never thought about that equation, I mean using that equation so far.

(Refer Slide Time: 17:12)



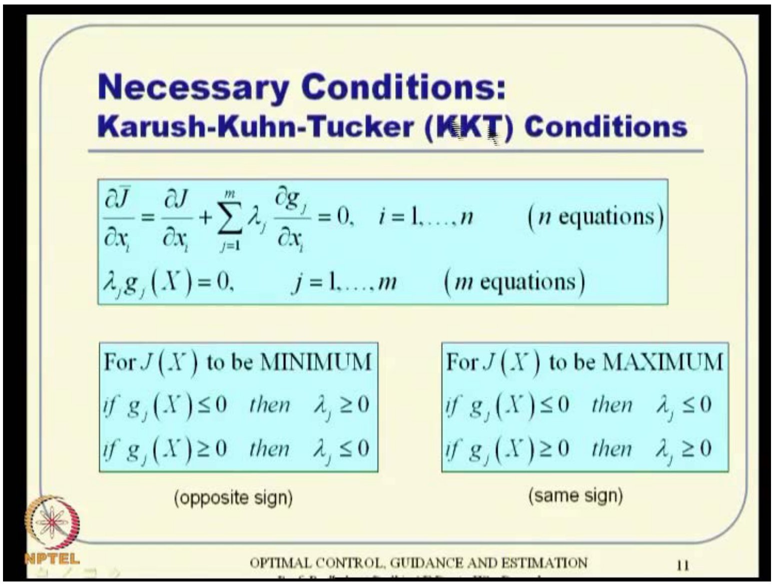
So, we will go back to that and see what you can get from there. Now, you remember this this quantity and this quantity has to be of same sign basically. That is what we have just discussed it actually. Now, going back to this, this is what we are getting here. That means, del j by del x i equal to that one. I will take that in other side of the story and because lambda j, I mean all that I can inform you here lambda j is not a function of x. So, I can take it out from there and it turns out that these gives us another constraint equation, that this partial derivate and this partial derivative are related to each other through that, but these two quantities are always having same sign and with a negative quantity here. That means lambda j has to be a negative. Remember this analysis told us that these two will always have the same sign, no matter whether the delta x i is positive or negative. So, if we have the same sign and this is a constraint equation with a negative sign, then lambda j has to be negative of each other, then only it will satisfy all that.

(Refer Slide Time: 17:24)



So, this is what it tells us. It tells us what if it is a maximization problem, right. I mean that is right, if that is a maximization problem, then these two, I mean this lambda j has to be a negative basically ultimately. So, this this kind of analysis you can do case by case and generalize things like that.

(Refer Slide Time: 18:24)



So, as summary sort of things, this this kind of analysis is called something like, earlier it was known as Kuhn Tucker condition and some books still follow that, but more popularly, now it is known as Karush Kuhn Tucker condition. The Karush Karush happens to be the independent results actually and then, Kuhn tucker divided these conditions. It become popular and later they realize that I think Kuhn-Tucker themselves or somebody else in the process of analysis, they realized that Karush is actually developed this condition as a part of the master pieces and then, the story goes on like that actually.

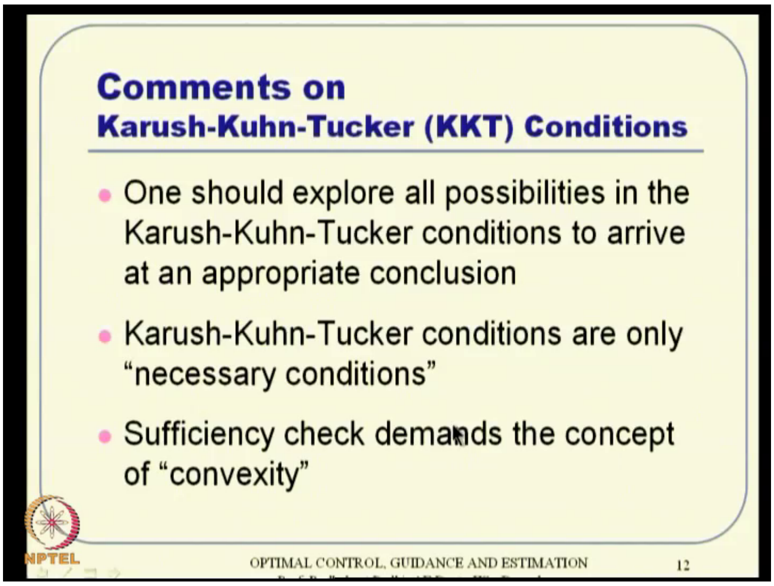
So, then they want to attach his name further and things like that and they prolonged here is Karush published his results. It became part of his thesis only. So, somewhat it was discovered, otherwise it would have been simply Kuhn Tucker condition all the time actually. So, I mean that also uses the clue that whenever you get something, try to publish. Please do not keep it yourself whether that actually anyway. So, this this is the condition here which is popularly known as KKT conditions now. So, Kuhn Tucker condition happens to be something like that.

So, what it tells us? It gives us these two conditions. That means, del j by del x i is to be equal to 0 and lambda j g j has to be equal to 0. These are the two equations, so using which we have to solve for various conditions for lambdas and all that actually. Then, for for j of x is to be, if it is a minimization problem and this constraint happens to be negative, I mean all less than equal to 0, then lambda j has to be positive. Remember these are (()). What we have analyzed this, actually maximization problem, moreover these happens to be like that. So, hence we will end up that condition. If it is this way, we will end up the opposite sign. If it is minimum, all these 2. So, depending on whether it is a minimum problem or maximum problem and depending on whether constant is negative or positive, we have the standard results getting on negative.

So, first of all your task is to first solve this thing, these two equations. Remember this equation do not include any mu's now. So, we have n plus m equation in terms of n plus m variables. It is possible to solve n variables from states, I mean rather very pre-variable x 1 to x n. I mean we will not talk about, it states actually in general.

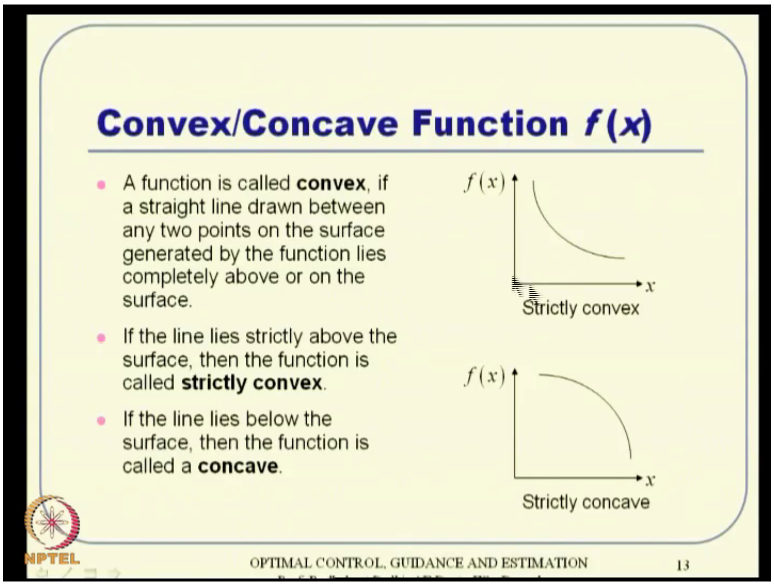
So, x 1 and x, x 1, x 2 up to x n is what we are talking as pre-variables and then, lambda's coming from because of the constant equation. So, these two sets of equations we can solve, get some values for x and lambda's and then, we will go away and try to see whether these conditions are satisfied or not, especially whether constant equation is satisfied and corresponding lambda is also satisfied or not actually.

(Refer Slide Time: 21:26)



Actually, these conditions are embedded like two set of conditions. One is the constant equation, where x l has to satisfied and then, the corresponding lambda's has to be satisfied that way also using that. That is the problem there. Actually for some comments on KKT conditions first of, I mean I already told you about that. That is all documented here. One should explore all possibilities of the KKT conditions to arrive at appropriate conclusion and also remember that KKT conditions are only necessary condition. They are not sufficient conditions at all and if you really want sufficiency's, then sufficiency's takes demand that some concepts of something called convexity in the function has to be convex and something like that actually.

(Refer Slide Time: 21:54)



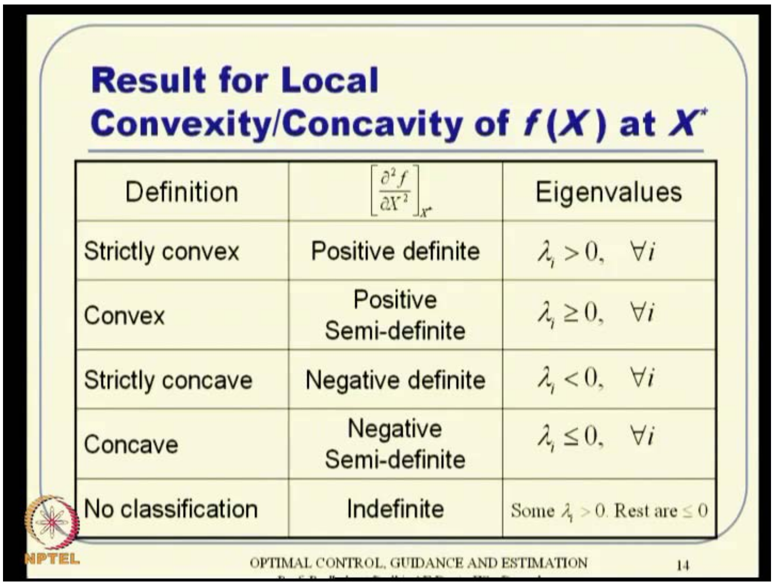
So, you see that, so the concave and convex function can be defined something like that. This particular, I mean in a scalar pre-variable sense if f of x can be given like that, then it is something like a convex function and if it is given like that, then it is a concave function. If it is strictly like this, it is strictly convex and strictly concave. The whole idea here is, if you join these two, any two points of this function, then this line is suppose to lie above this curve actually, ok. If we join any 2 lines here, these two lines, this line is completely lying below this function sort of thing actually. That is what is written here.

A function is called convex, if a straight line drawn between any two points on the surface generated by the functions lies completely above or on the surface actually. At a particular point when you need to tell that, it will all lie on the point. That is one single point only basically. If the line, I mean lies strictly above the surfaces on the other end, it is called a strictly, I mean if it is strictly lying above this, it is strictly, this is called strictly convex actually. Remember this. We are talking about two different points. If two points happens to be same, then it is a tangent point basically. They are not intersected in the other ok.

Just opposite case, if the line lies below the surfaces, it is called concave and if it is strictly below, it is I mean strictly concave and all that actually, but how do you do that? How do you analyze that with medically? That is more important actually. So, that is what this differential geometry and all that comes into picture and then, if the conclusion through calculus something like that actually ok.

You can always say these concavity and convexity are typically governed by the positive definitiveness and negative definitiveness minus of f of x actually. So, that means, del square f by del x square is contain come sort of curvature information. That curvature information will give us whether this is convex, concave and all that. The idea is like this. You you have a conducted solution x star, so that point you can evaluate this one and if it is a scalar variable, it is just a number, but if in general, it is a metric we are seen all that actually in previous lectures. So, these particulars matrix can be evaluated here and if it happens to be strictly positive definite, that means, all lambdas are strictly positive. Remember this is actually symmetric matrix guaranteed to be and all Eigen value's are real actually.

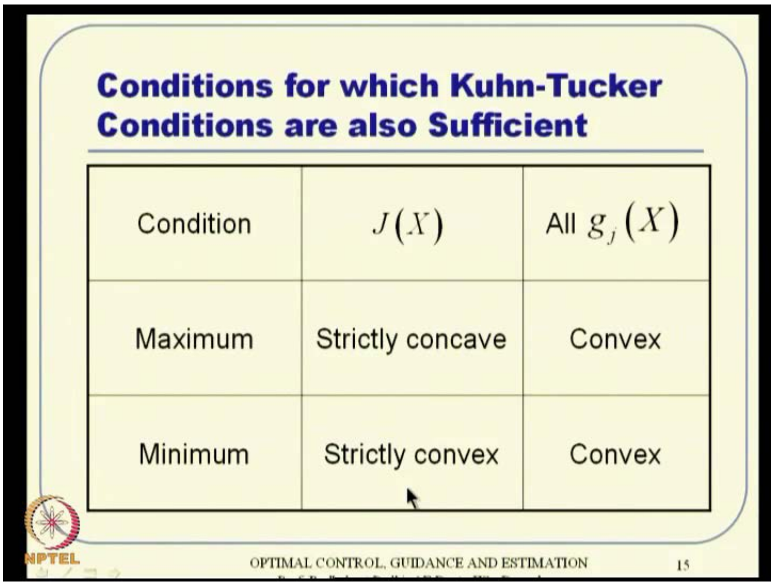
(Refer Slide Time: 23:27)


Result for Local Convexity/Concavity of $f(X)$ at X^*

Definition	$\left[\frac{\partial^2 f}{\partial X^2}\right]_{X^*}$	Eigenvalues
Strictly convex	Positive definite	$\lambda_i > 0, \quad \forall i$
Convex	Positive Semi-definite	$\lambda_i \geq 0, \quad \forall i$
Strictly concave	Negative definite	$\lambda_i < 0, \quad \forall i$
Concave	Negative Semi-definite	$\lambda_i \leq 0, \quad \forall i$
No classification	Indefinite	Some $\lambda_i > 0$. Rest are ≤ 0

So, if on top of that it happens to be positive number, then it is positive definite matrix and it is strictly convex function actually. Then, it is a positive semi-definite, negative negative definite, negative semi-definite, indefinite and all that, we can be inferred from that just by analyzing Eigen values of this matrix because it is a number after all. This matrix is evaluated at particular point extra, it is the number. So, you can evaluate all the Eigen values and depending on these conditions, whether all of them are positive, strictly positive things like that, you can infer whether the function is strictly convex, concave and all that at that particular point x star and that also has to be noted down actually and if you are some of them happens to be in positive, some happens to be negative, all that you can do talk anything about that and that is in this particular case typically called as indefinite function.

(Refer Slide Time: 25:12)



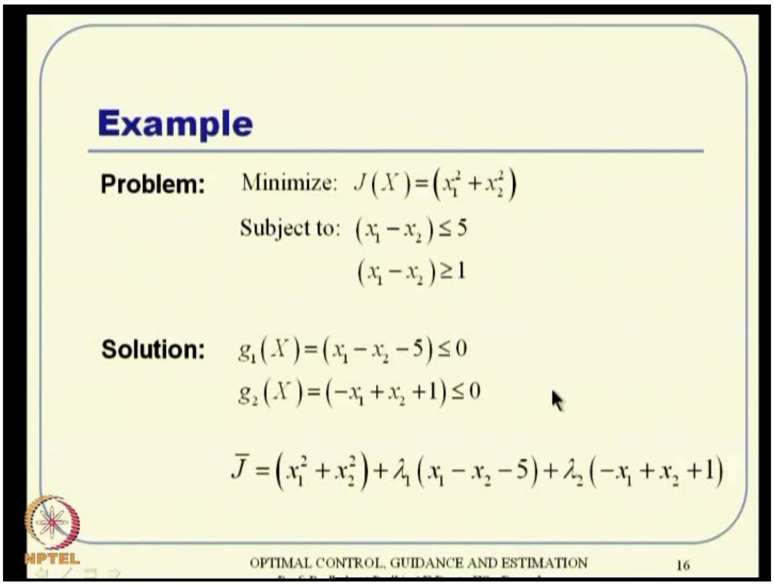
So, conditions for which these Kuhn Tucker conditions or rather Karush Kuhn Tucker conditions are also sufficient and tends to be like this actually. Why? In general, they are necessary, but in medicine if this kind of combinations are satisfied, then these whatever x star value we are talking about, that happens to be sufficient also basically ok.

So, what is that? Now, if you are looking for a maximization problem; that means your j of x has to be strictly concave and all g j's has to be convex, may not be strictly convex. While for a minimization problem, j of x has to be strictly convex, but g j has to be convex. So, this constant equation need not see, convex also includes linear equation. Linear equations are straight lines sort of representation in all all that way. Then, there there also included in the constant equation basically, but the cost function is to be strictly concave or strictly convex. That means, it has to be some sort of a, I mean quadratic fourth order problems for all that we talk for minimization or negative of the maximization and all that. Those will satisfy this kind of condition actually ok.

So, this is the summary actually. You you analyze all that. Let me go back to a little bit. You start from something like this. In a given problem, formulate these 2 equations and get some values of all x's and x sides and lambda j is and then, see whether these set of conditions are satisfied together for n condition that way. Then, if you have, then you I mean if these conditions are satisfied, you have these necessary conditions getting satisfied. Now, we will talk about sufficiency conditions. For sufficiency conditions, you have to see whether these two, j of x and g j of x is satisfied, this kind of thing. If satisfied, then it is both necessary in sufficient actually.

Obviously, it is time for an example to get some ideas clarified and all. So, let us talk about this kind of an example j of x reference to be x 1 square plus x 2 square, standard problem to minimize a quadratic function sort of thing. Subject to these 2 equations now, x 1 minus x 2 is less than equal to 5, but x 1 minus x 2 is also greater than equal to 1. That means, this value what you are looking at lies between 1 and 5 basically ok. So, in that segment have been whatever minimization we talk about can be derived analytically actually.

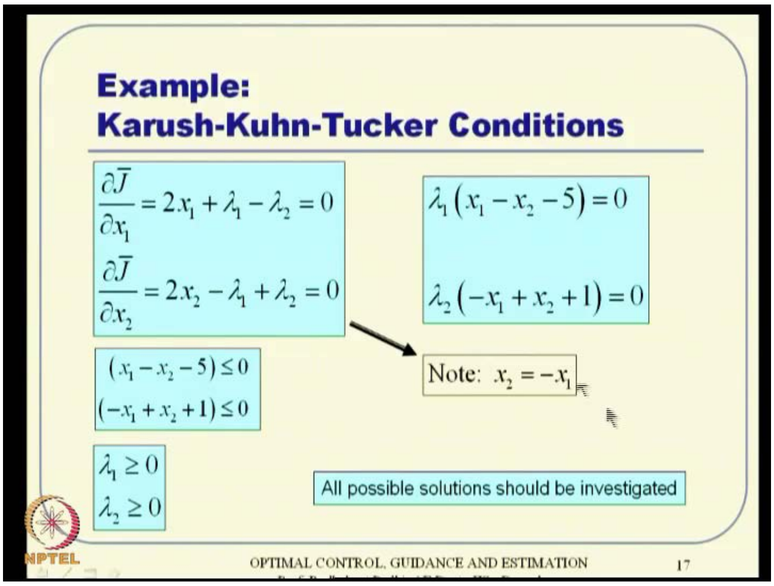
(Refer Slide Time: 27:05)



So, what is the first thing? First thing to note is what g j is. I mean what is g of g 1 and g 2 here and our analysis all tells that we have to do in a negative sense. I mean in a less than equality sense basically. So, first one is g j. So, we just talk x 1 minus x 2 minus 5 less than equal to 0. The second 1 we have to change sign, if you put a minus sign here it becomes I mean less than equal to minus 1. In that sense, it is minus x 1 plus x 2 and remember these minus 1 here, so it becomes plus x 1, sorry plus 1. So, this is g 2 of x has has to be less than equal to 0 actually ok.

So, now we have a standard form. Minimize this this function, this cross function subject to this g 1 less than equal to 0 and g 2 less than equal to 0 actually. That is what we are looking at. So, what is the standard thing? First we have to be formulate what is j bar and j bar is nothing, but j plus lambda transverse this this function g 1 g 2. That means lambda 1 into g 1 plus lambda 2 into g 2 that is j bar. J bar is j plus lambda 1 g 1 plus lambda 2 g 2. Now, it require the necessary conditions. First is all these partial derivatives has to be equal to 0.

(Refer Slide Time: 28:44)



So, partial derivatives in terms of x and in terms of lambdas, in all that we will see that, that see partially, ok let us go back to that. All that we are talking is partial derivatives with respect to x and then, directly we will go to this condition. This partial derivative with respect to lambda x are not necessary actually. Both condition have been derived and invaded here in in the process of deriving and all that and it has been invaded here actually. So, we do not need that.

All that we have to do is, partial derivatives with respect to these x sides and then, then use this condition directly. So, t his is what we are doing here. So, for j bar is there. So, partial derivatives with respect to x 1 and x 2 will give us this condition and then, lambda 1 g 1 equal to 0 and lambda 2 g 2 equal to 0. Now, what? How many equations? 4

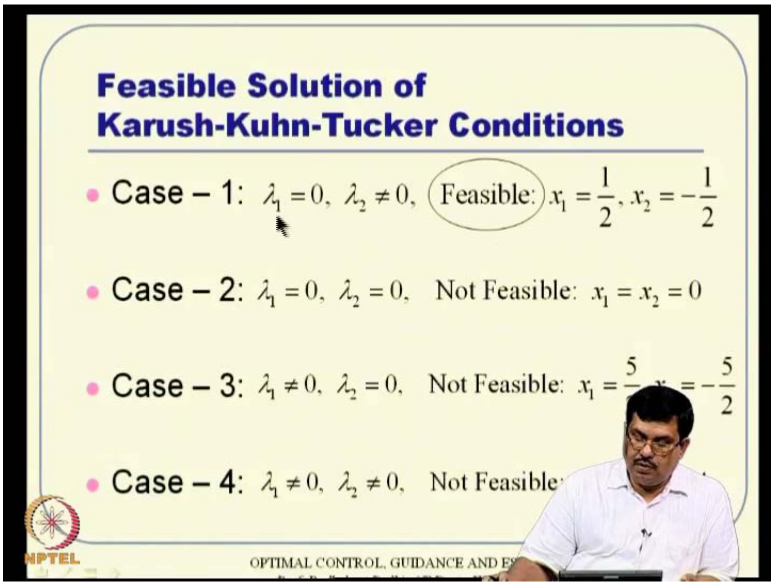
equations obviously. How many unknowns? Four unknowns x 1, x 2 and lambda 1, lambda 2.

So, this we will be able to solve this actually, but remember this is not a linear set of equation. Even though these equations are linear, these equations are certainly non-linear. You can think of them as bilinear and thing like that. Bilinear is also part of non-linear equations actually. So, this equation has to be solved and then, we will solve case by case and then, verify whether these, whatever candidate solution you are getting here, we will get multiple solution candidates and this satisfies these 4 equations simultaneously. If they satisfy their part of the necessary conditions and all, so they are candidate solutions. If they do not satisfy, in other words, one of the conditions not getting satisfied. That means that is not a possible solution actually ok.

So, that is what we are looking for here. Now also, very quickly you can observe that x 1 is equal to minus x i means x 2 equal to minus x 1 here by solve x 1. Here it is what about minus of lambda 1, lambda 2 by 2 and x 2 here is negative of that basically to solve that, so that x 2 is equal to minus x 1. That will quickly give us whether to keep a solution or reject a solution because if you know x 1 and x 2 and x 1 are just opposite signs, very quickly you can see whether it is happening or not. So, that does mean for sort of thing actually.

So, now if you go back and try to solve case by case and I certainly encourage you to solve it. So, first of all remember how you are going for solving it. Just look at the equation. We are getting either lambda 1 equal to 0 or this is equal to 0. Here it is either lambda 2 equal to 0 and that too equal to 0. With respect to these possible equations, we have to solve this actually.

(Refer Slide Time: 31:14)



So, either lambda 1 equal to 0 or lambda 2 not 0 or lamda 1, 0 lambda 2, 0, that is also possibility and lambda 1 not 0, lambda 2 not 0, that is the possibility and both of them is not 0 is also a possibility. In that case, both of them are not 0, then these two has to be equal to 0.

Now, in this situation, you solve it. Go ahead and solve it. In this particular case, this is the solution. This particular case that the solution think like that actually. Now, very quickly we can say remember this x 1, x 2 equal to minus x 1 right. So, this candidate is not actually, it does not satisfy that. I mean well, it satisfies in a trivial sense actually, but let me pick up another one. Well, I think this satisfies this. So, maybe we will not be able to do this, but suppose it is a different number here instead of 0, then it is very easy to kind of utilize this equation and then (()), but any way that is not a major point actually.

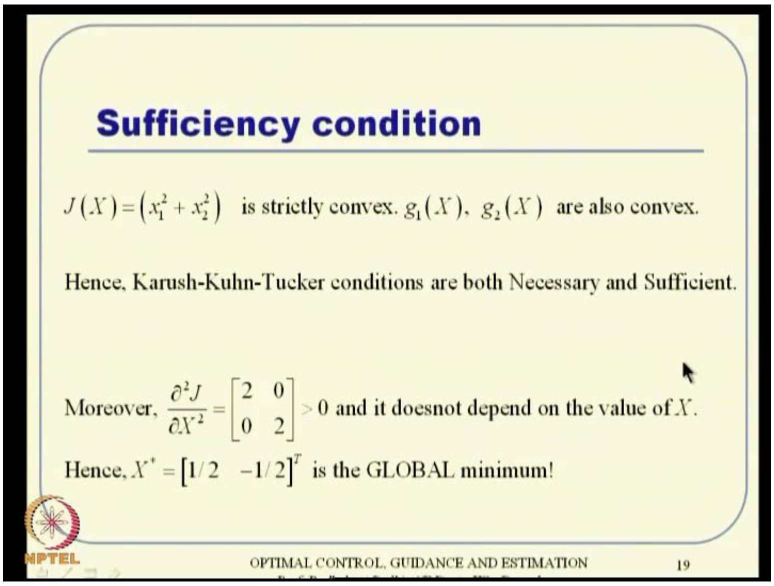
So, you have some values here, lambda 1, 0 lambda 2 not 0 like that and corresponding x 1, x 2. Now, go back to this equation and try to see whether that is satisfied or not actually. Now, what is happening here lambda 1 equal to 0, lambda 2 not equal to 0. Lambda 1 equal to 0 is acceptable ok. So, that is not a problem. So, lambda 2 not equal to 0 is also a possible candidate because lambda 2 can be strictly greater than 0. So, that is not a point actually. So, see this but this condition looking for is primarily evaluate this set of constraint actually. This condition has been taken care from ther also basically. So let us look at these two equations first.

So, if you have x 1 half and x2 minus half, if you put x 1 half and x2 minus half and this happens to be minus 5, so this is less than 0. This is satisfied. This one, second condition half and minus half, so this is first half and what is that, sorry. X 1 Is half and x 2 is minus half basically, so x 2 is minus half you put here and half here and all that actually. So, this minus half, minus half, minus one, plus one is 0 actually. So, this also satisfies. So, in that situation, it is acceptable. So, this a feasible condition actually, but it turns out that all other cases if you, I mean if you just through lambda 1 equal to 0, lambda 2 equal to 0 and things like that, then these two equation is to satisfy in all. We see that this kind is not feasible and similarly, all these are not feasible.

So, for example if I take this minus 5 by 2 plus 5 by 2 minus 5 by 2, now I suggest you to do it yourself basically because you not spend too much of time here, so that all the feasible condition is take s by k and try to analyze whether all these conditions are satisfied or not. Very very quickly you will see that only one condition it will satisfy. All others will not satisfy for some reason or other actually. So, that means, this is the feasible condition x 1 half and x 2 minus half or other things are not a feasible conditions actually.

Now, because this is the only condition that we are getting satisfied, you can always tell the sufficient condition is there in general and that is the intuitive feeling, but if you really want to find out, whether it is really sufficient or not, we have to go through that. So, that means, j of x equal to x 1 square plus x 2 square. Obviously, it is strictly convex function. Then, if you take second derivative, it will happen to be a diagonal matrix with 2 2 in the diagonal. So, 2 in the first diagonal and 2 in the second diagonal of matrix.

(Refer Slide Time: 34:13)



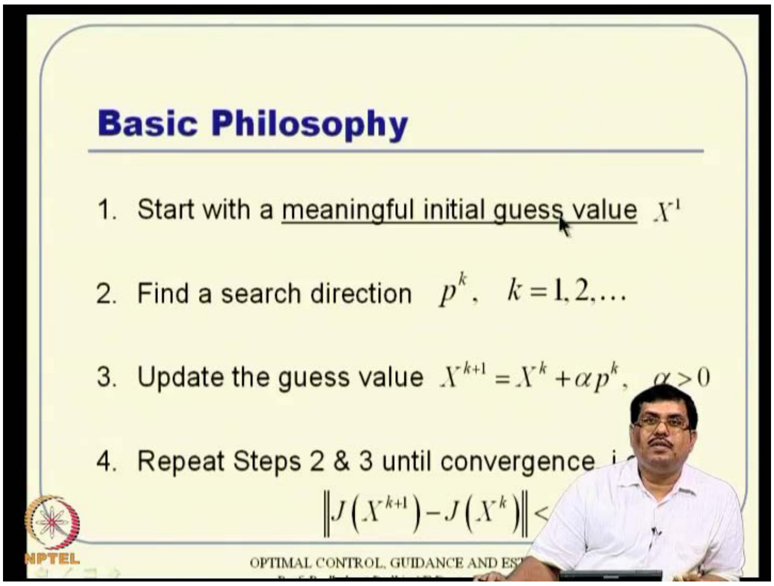
So, that means, Eigen values are 2, 2 which are both positive. It is strictly convex actually and g 1 and g 2, if you see this, this g 1 and g 2 whatever you are talking about is the linear equation. Hence, they are, I mean convex actually and we do not need strictly convexity, just convex is good enough. Hence, this condition there is Kohl Tucker conditions are both necessary and sufficient as well actually. So, that means, this particular case that we are talking about to whatever we are here, this I mean this solution candidate and under that you are talking about is both necessary and sufficient.

Hence, it is the candidate solution in both sense, but is it truly global and that also happens to be true because if this particular function what we are looking at, del square 0 del x square is not a function of x star. That means whatever your x star is what you are talking here, this second value condition what we are talking, I mean what is coming out here is just like diagonal matrix 2, 2 in the diagonal that is not a function of x star. That means, no matter whatever is x star, this condition is guaranteed to be satisfied. So, that is how we can conclude that. So, that is how the analysis proceeds something clearly actually.

So, this is the general is all about this, kind of analytical tools and all that. Remember you started with very simple here. So, just a scalar objective function of a simple looking form subject to just, I mean 2 equations with 2 and 3 variables. (()) to solve it using pen and paper, you will require a longer algebra to get all those conditions and in general, it turns out that is not a very good approach of doing complicated problems actually, some complex and complicated problem because it does not give us a very good analytical hold in a way that algebraic can be so involved that it may not be possible to do this kind of nice analysis in general. So, people always looked at this numerical optimization approaches and over the period of time, it feel as matured quite a lot. That means, here plus several several numerical algorithms now which are quite efficient actually and also, this this computer technology has seen this lot of improvement over, I mean the last couple of the years. So, here huge amount of computing power as well actually.

Of course, they are not sufficient in optimal controls and in a way 2 point boundary value problem and always see to that different class of problem form together, but having said that advancement of computer technology always help actually no matter what. So, utilizing all these powers together and since, then the people started utilizing this this online solution of this optimal control problem and all that. We will see all that later, but here we will see some basic ideas of what is this this static optimatization through numerical methods actually.

(Refer Slide Time: 37:44)



So, what is the basic philosophy here? Basic philosophy is the kind of we do not know the final solution. We do not want to go there at once actually. What we start is some sort of a meaningful initial guess and trivial, I will underline this statement because it is a critical thing in practice actually. As far as algorithm development is concern, it may not look very important, but in any any practice, the initial guess value or initial guess solution and all that has to be done carefully. It has to be done in a meaningful sense. In other words, it may not be an optimal solution, but it has to be as close to optimal as possible in general or if we talk about a control solution, although it has to be kind of a stable solution to begin with, anything like that actually ok.

So, if you do that, then your iteration and all becomes lesser, your computers load becomes lesser, your convergence properties became better, things like that actually ok. So, what happens here is you start with meaningful initial guess. Let us denote that as x superscript 1 actually and then, we have to all that matters is to find as such direction because the direction number of such direction and infinity really we can go any any direction you want. So, we have to find a meaningful such direction and let us denote that is p k and p suppose to k and k can be like 1, 2, 3, 4 anything actually.

Now, once you find its good such direction then you have to go to go something and some some step you have to to go or some some in that direction you have to move something. In other words, our initial value or whatever is the updated value has to move a little bit in that particular direction. That means, x k plus 1 can be updated this way. X k plus alpha time p k. Remember p k is of same dimensional state actually or again and again talk about state because x is a state variable in general, but if in a 3 variable sense, whatever is a 3 variable, in that particular direction, it has to move actually.

So, alpha I mean and typically alpha has to be a positive number because you do not want to go in the opposite direction. You you got a direction and you have to go in the same direction basically and you keep on doing this and repeat this procedure 2 and 3 until convergence. You have to keep that. Why it is necessary? Because this search direction happens to be very local. That mean the direction does not remain the same direction as you take a very step in that particular direction, that the direction of further design and all that actually happens to be different. That is the necessity of doing repeated calculations of this step and this step fluently.

If the direction remains to be unique; that means, you really do not have to search other direction and all that. You do not need to do all that. However, because it is very local property, you have to keep on doing this iterate research and finally, you will derive a convergence thing and convergence can be checked something like that and the cross function does not improve further basically. So, j of x k and j of x k plus 1, both are roughly say each other subject to some pre-selected tolerance value and thing like that. Sometimes people use relative things also. They divide it by norm of j x k and that has to be some percentage quantity now and that percentage happens to be that side let us say, less than 0.01 and you can stop actually .

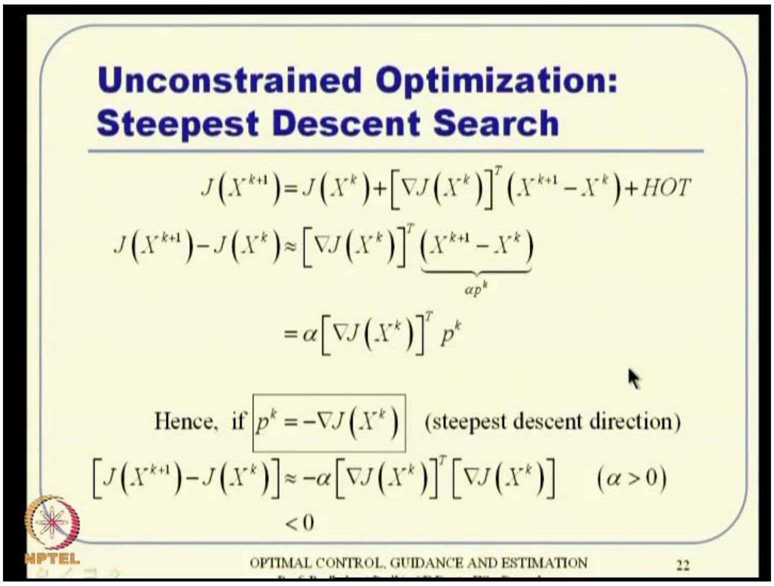
So, well, there are various ways of doing this but one way of doing that is, we just look at like there is no further improvement whether I take x k or x k plus 1 actually. That means, x k and x k plus 1, all kind of giving me the same value and also another way of getting convergence is to just to see these 2 values together along with this actually. You evaluate this and also make sure that these 2 points are not separated by the large distance. It may show it can happen if the function happens to be like let us say, 2 optimums like that and like minimum optimums and all of the same value, which is again typically very very rare, but it can happen now.

For example, like if I just can show something when the function happens to be something like this, then these two are of the same values actually. If x and f of j all over, it happens to be the same value and then these two values are almost the same, but the distance is large actually. Then, you have to select which one is your optimal point and all that actually. All those kind of situations happens with very rare case. No need to worry so much of that kind of thing actually.

So, all that it happens is ok I will continue the process as long as I do not see some improvement and all that. Whenever I see this property (()) ok this converse value and all that. This is the whole idea there. We start with a meaningful guess value, find some search direction, go a little baby step and repeat the procedure again and again until convergence exits. Now, the whole idea is how do you find those p k. That is the more important thing actually.

Let us say that part of analysis, so again back to that tailor series out of ideas here j of x k plus 1 can be expanded around j of x k and this is the relationship what we have actually, j of x k x k plus 1 is nothing, but j of x k plus gradient of j transposed and this error plus I had added terms actually. Again neglecting added terms, you can approximately make it equal. This minus this equal to that and hence, you can solve this or otherwise, x k plus 1 minus x k. Remember, these 2 are nothing, but alpha p k. So, I can put this alpha p k here and alpha is a constant number.

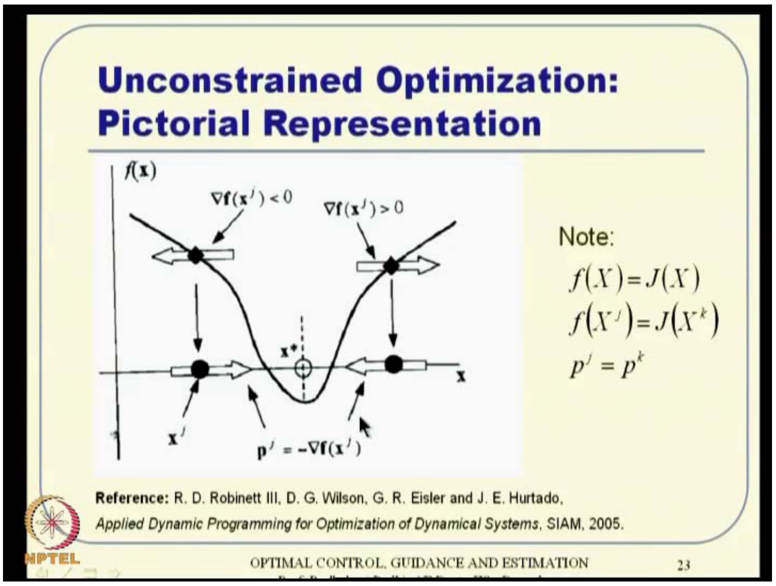
(Refer Slide Time: 42:40)



So, I can have it here and hence that p k if I get this direction actually nothing, but that. Why? Because if I select this p k, it has to be exactly same what I am having here. Then, this would never this, whatever j of x k plus 1minus j of x k will be a quadrative quantity with a negative sign basically. Now, this is again a negative sign, remember that. If I happen to select this, then it happens to be a negative square sort of things and hence, this is guaranteed to be less than 0. That is the whole idea of this actually. What it gives us? It gives us p k as a negative gradient of j i means, negative of gradient j basically. So, that is the direction of steepest descent actually.

So, if you happen to take it that way, then you will go along the steepest descent direction of that thing actually. So, this this will guarantee that these two would different j of x k plus 1 minus j of x k is strictly less than 0. That is the whole idea why you select these actually. Once you select this, what is the algorithm there? We have to just, once you select this, this is I mean this is the algorithm there, such k plus 1 is nothing, but x k plus alpha times p k, where p k happens to be that actually. Let us say that and you can proceed further in that pictorial that way.

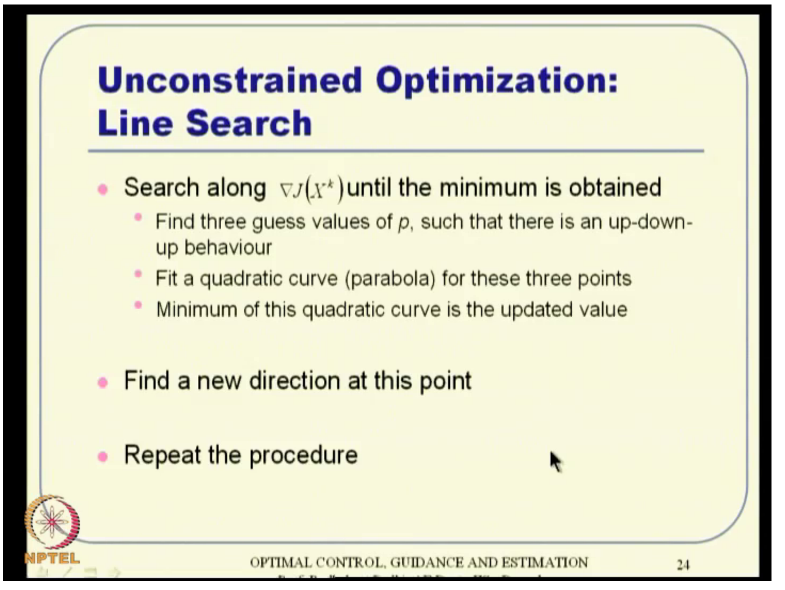
(Refer Slide Time: 44:21)



So, pictorially steepest speaking one, if it is the x k is a scalar quantity, then it reference suppose it happens to here guess value is somewhere here, then the gradient thing is somewhat negative and remember this this function if locally it is negative and hence, you have to move in an opposite direction of that. So, you will move a little bit towards that direction and if it happens to be here, then here the gradient is positive. In this case gradient is slow, so it is here, the gradient is positive. So, you move in the opposite direction again.

So, if you start here, you move in this direction. You start here and you move that direction. Ultimately, you converge somewhere where the gradient is 0 basically. So, that is all that and I have taken this this this kind of couple of diagrams from this particular book about there are some little bit different notations here. Whatever you see as there f is nothing, but what you consider as j and and similar things actually and small j and there k is out of thing.

(Refer Slide Time: 45:39)

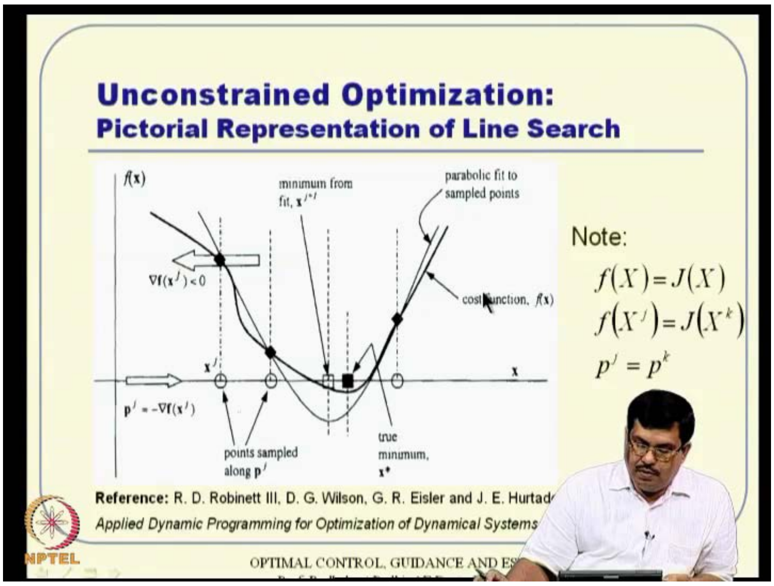


So, just make sure that the notations are not forgotten in each other, but otherwise the consumption explanations are taken from this place anyway. So, now, coming to the some some idea what is called as line search. The whole idea here is that we have got this this particular direction, so why taking a small baby step and forgetting this thing actually. You can continue searching in the direction as long as the function keeps on decreasing. That is the whole idea there actually. We keep on searching, expand, little more, little more, little more, then think greater and keep on doing that.

In other words, make this p k, I mean the travel along the direction of p k that means alpha is not large as possible actually and it will guarantee that the cost function keeps on decreasing. I mean it is at some point of time if it keeps on increasing, then there is no point in taking that enlarger actually. You will go up to that point where that function keeps on decreasing actually. That is the whole idea of line search actually.

So, how do you do that? The concept again is easy. What you have to let me explain this. First you have to find some solution here, some some value and then, we have found a direction anyways. Using that direction, you find two values now. This is a point where the function has a value and you got a direction.

(Refer Slide Time: 46:31)



So, using that direction you have another value and do not stop there. You keep on doing that and find another value, where the function is greater than this value. That means, greater, lesser and greater sort of thing actually and if it happens, then you can actually think that ok the function has to take a turn to go there actually, right. The function keeps on, this is the true function anyway that you want to optimize.

So, let us say this function keeps on decreasing, decreasing, decreasing and at some point of time, it has to increase anyway before this function is greater than that actually. That means, if it is a greater, lesser, greater sort of idea, then I can actually fit a parabolic a curve between the two curves, between the three points because parabolic has three variables a 0 plus a 1 x plus 1 sort of things actually, but I need 3 points, I got 3 points and I will fit a quadrative function using these 3 points and then, find out the minimum point of the quadrative function using the close (()) solution. That now we know that actually. Any parabolic function, if it is a 0 plus a 1x plus a 2x square and then, minimum of that particular function we know very well actually. So, you will go up to that point and you will say ok now, here is a quadrate solution and then, remember this function takes sort like that the true minimum somewhere lies somewhere here.

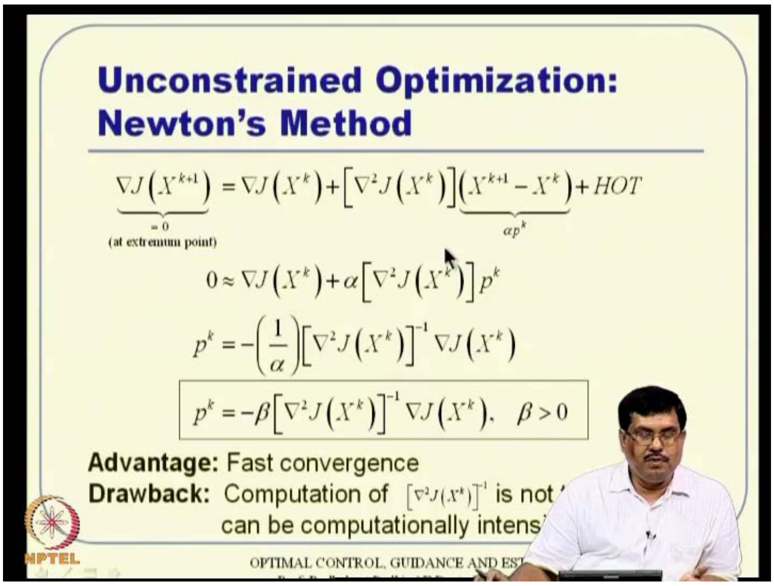
So, by simply doing a little more math in the background, what we are doing here is actually going in 1 step very close to that optimum 1 point actually. That is the whole idea of doing that and if if you apply this algorithm to, I mean couple of times you may converge actually because it has a very fast converging property. So, that is the whole idea of line search. So, telling in English words. First of all, find three guess values of p, such that there is a up down of behavior. Then, fit a quadratic curve and find the minimum value of the quadratic curve. I mean, it is not the minimum value of the actual function actually.

We got a quadratic curve which approximately represents the behavior. So, find the minimum values of this quadratic curve and then, update the value to that particular point actually. At that point again, you will find a new direction and things like that. So, you may repeat the procedure actually. Remember, the direction there is a very local property, so we may not able to deal with the same direction anymore actually. Once you go there, then you can find a new direction and continue the same procedure like that actually. That is how this this steepest descent and line search method is quite popular actually. One easy way of mechanization and a very intuitive procedure as well actually, all right.

Now, coming to some of the method what is called as a kind of Newton's method and all that. It goes a little bit further and remember the first derivative or first gradient, I mean the the gradient of j has to be equal to 0, right. That is what we know from either necessary condition of any pre-optimization that the gradient has to be equal to 0. So, if it is non 0, we will try to make it 0. I mean that is the whole idea there actually. So, again we will go back to the function expansion, but we will not expand the function pursue. We will expand the gradient of the function here. So, it is actually a little bit more intelligent way of looking at the problem actually. Thus, we know that the gradient has to be ultimately equal to 0.

So, we will analyze the gradient of the function and expand it further using that tailor series actually. Remember, it is a vector function now, but vector function can be still be expanded in terms of Taylor series any way. So, use this vector function and you expand it using Taylor series around the point x k. Then, what it gives us? It gives us the first term plus the second term and all that and we know, second term itself talks about (()) matrix now ok.

(Refer Slide Time: 49:19)



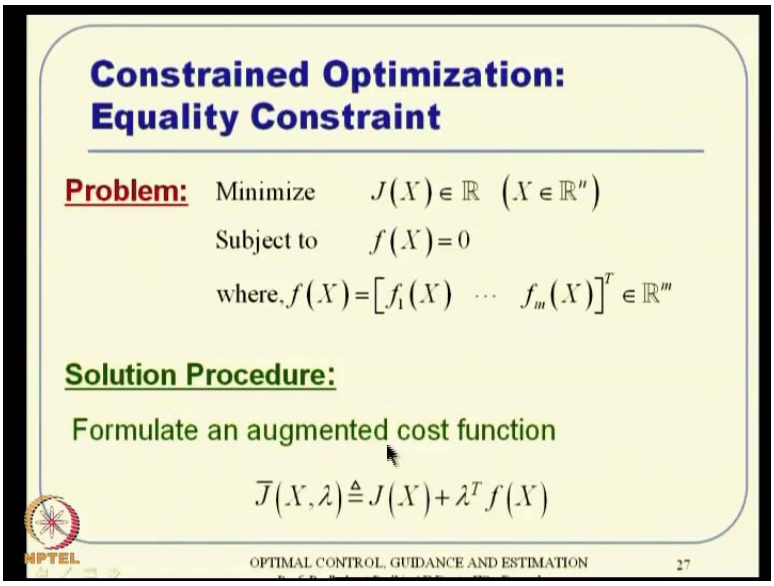
So, obviously, this gives us some sort of alpha p k interpretation this part of the thing. What is our aim? Our aim is to make it equal to 0. So, however, we will find the next k plus 1 in such a way that this gradient value del dot j x x k plus 1 happens to be 0. That is the whole idea actually. So, if this happens to be 0 and I neglect this arrow dot terms, then this requires sum is approximately true and from here, we will directly solve what is p k actually. So, p k gives this this direction, source direction. We can just solve it. You just take it to this side; it will give negative value and all that. Then, this inverse and which is happens to be I over alpha and you define 1 over alpha is beta, then this gives you some sort of a direction basically.

So, again this, once you get the direction, you can bring in the concept of this this line are something look like that the lines are since independent of this kind of ideas there actually. So, you you find a direction utilizing the fact that necessary conditions tells us the gradient has to be 0 actually, but remember, this is all true for pre-optimization thing. The moment you just concert of optimization, this gradient vector has normally be 0 actually. Now, we have to talk about gradient vector of an augmented state variable, I mean augmented cross function and that has to be equal to 0 and think like that.

Anyway, going like this is a very, I think in a kind of intelligent way of doing that. Again, Newton has always been intelligent actually for whatever it is. He is a great mathematician and try to kind of simplify many things actually. So, this is all the way going from Newton's kind of ideas. It goes like that. So, we can search find such direction very quickly using this thing. The advantage is that it gives us a fast convergence and the drawback associated, drawbacks happens to be this one because this is a matrix now and you have to compute a inverse of a matrix and matrix inversion is not computationally simple in general.

The dimension is lost. Then, inversion needs some certain amount of computational before going actually. So, that is the thing, but nowadays, that is not a major issue. We can talk about matrix inversion rather easily actually because both numerical algorithm are available as well as computers are fast. So, using both the, both of that I think this not a very major issue unless this dimension is really huge, something like 1000, 5000 and all that. Then, this becomes a big dimensional problem and all that, but in our control synthesis problems and all, it will not run in that kind of issues unless otherwise you talk about flexible system, we will important dimensional system and all that which in this course will not talk actually.

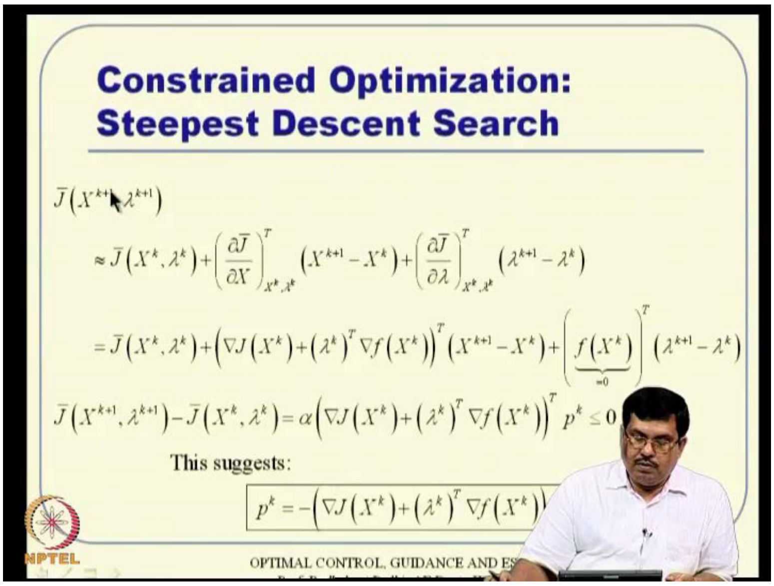
(Refer Slide Time: 53:04)



Now, coming back to the same idea how do handle this this equality constraints. Now, again as I told you, formulate this cross function. They are equal to this one. They try to interpret this as a pre-optimization with respect to x and lambda. So, carry out the same algebra as you carried out the Tailor series with that same thing, but now you get before lambda variable as well. So, now if you talk about that, it is a error in x and error in lambda as well.

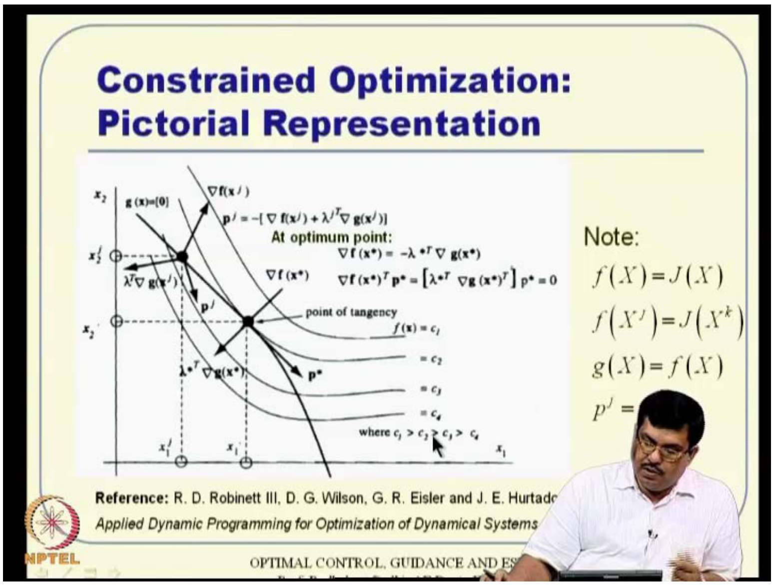
Now, this particular thing is evaluated at that and you can expand that j bar. J bar is nothing, but that actually is a function of j and lambda transpose j. So, I can j bar I can put it and write it actually ok. That means, j bar of x k plus 1 lambda k plus 1 minus of that if I do it here, it will give something like this actually. Why? Because ultimately this this particular thing is nothing, but a constraint equation.

(Refer Slide Time: 53:17)



This is 0. So, this is not counted actually. This term does not contribute. This is gone and you are left out with something like this actually. So, what does it suggest? Again we need to decrease this and this is something like alpha is a first n number, this into this. So, obviously, p k has to negative of that actually. Whatever you see here if you just select a negative of that quantity, then you are guaranteed to get a square sort of thing. That is the whole idea here. That is how you get the p k, but now p k is a linear combination of this gradient as well as that gradient and coupled with this lambda k and all that actually.

(Refer Slide Time: 54:34)



So, this is shown as geometrical interpretations and things like that. So, in general again this notational incompatibility will be there, but ignoring that for a second and I think that you can correlate that rather easily. So, del t tends to be something like that way. So, what happens here if you just look at the expression here, the search direction is negative direction of these two gradient vectors. First, you take this two gradient vectors, formulate a linear combination of that. That will give you some direction actually and negative of that direction happens to be source direction actually. Ok.

Now, another interpretation how long will it continue? It will continue until this all happens to be 0, very close to 0 or 0. Then, if it happens to be 0, then what is the interpretation? Interpretation is this one is negative of that actually. That means, these 2 quantities (()) and (()) turn out be aligned actually. So, this is what you will see here actually. Lambda transpose this (()) whatever direction is that and delta j, they happen the negative of each other. That is how it is actually.

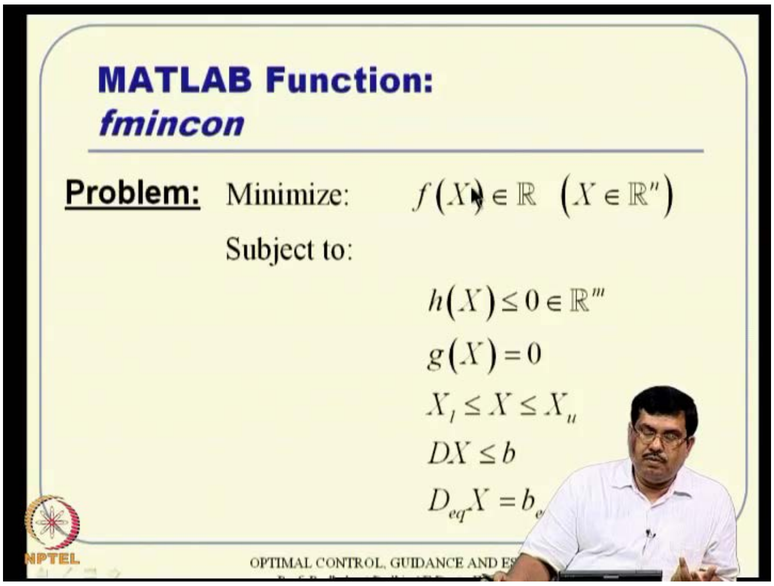
So, if you do that, they tends out at the optimum point, these 2 vectors are the line, but otherwise you conjugate direction sort of thing. This idea is something called conjugate direction and all that. So, you typically have a conjugate direction, but yet the optimal point these 2 vectors will get aligned and opposite direction sort of ideas and all that. So, what it tells you can draw this contour flow. So, that this c 1 happens to be greater than c 2 happens to be greater than c 3 and all that.

If you start at any point, this point still gives us some sort of conjugate. These vectors are not aligned yet and you continue a search up to a point, where these 2 vectors are aligned. Now, when it will be aligned? When it is a tangent vector to this particular curve actually. So, if you kind of proceed to that up to that point, where this direction if I talk about gradient of f. What about the direction happens to be a normal, I will mean with respect to the tangent vector. So, they are geometric interpretation and all which which is numerically, they are much more meaningful, but this particular course you can just remember that p k happens to be like this and you can keep on continuing yourself until this quantity goes to 0 basically. P k itself is a 0; that means, there is no such no more search direction actually ok.

Again more of that explanation you can see in this particular book. It is a typically good book for dynamical systems. That means, it talks about dynamic programming in details actually. Now, before I stop this lecture this I will also want to introduce you to this Mincom function of the MATLAB, which generalize, which invades all these in much much more this various numerical scheme, conjugate gradient direction, all sort of things actually and it talks about a generic way of solving a minimization problem, f of x is a scalar which can account for all sort of constraint. You can talk about less than equal to 0, you can talk about quarterly equal to 0, you can talk about states of the 3 variables getting constraint both sides, you can consider some linear equation getting constraint in equality sense as well as equality sense.

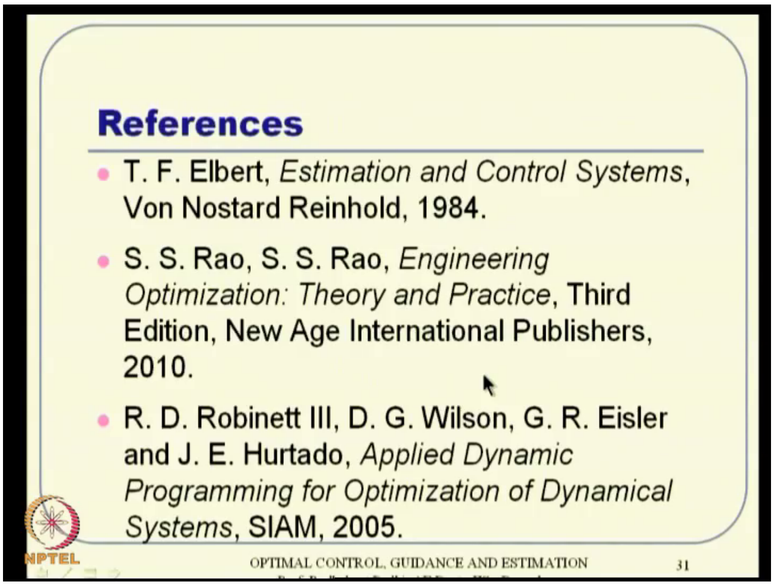
So, all sorts of constraints that you can think of is all invaded here with respect to this particular function to minimal. So, I also, I mean I will not talk too many details of what MATLAB itself has a help function actually. If you go through that help, they have documentation. There are additional documentation available and this particular function happens to be very popular worldwide actually.

(Refer Slide Time: 57:02)



So, it is not an extremely efficient function to use in online application as far as control thing is concerned. No matter what, it gives you a lot of these constraints getting invaded here. It is a very powerful technique and various algorithms also you can select within that actually. So, using that you can, I suggest that you solve some of these examples that we have discussed in the part of this lecture using this function also and see whether you are getting there or not. That will give you some practice and some hold on that to know the function and I suggest that all of you should become very comfortable using this function actually.

(Refer Slide Time: 58:37)



So, with that I think I will stop this lecture. Well, before that again these are references and anybody wants to do further, you can study further actually all right. So, let me stop this lecture here. Thank you.