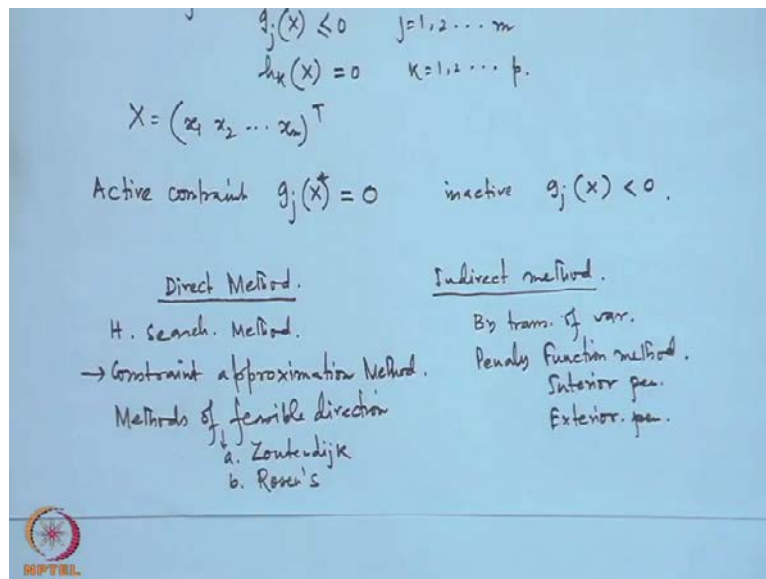


Optimization
Prof. Debjani Chakraborty
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 32
Nonlinear Programming - Constrained Optimization Techniques

(Refer Slide Time: 00:24)



Now, today we are dealing with constrained optimization problem. This is general form of constrained optimization problem, find the decision vector X ? Where we want to minimize that minimizes the function $f(X)$ subject to the constraint $g_j(X)$ less than equal to 0, $h_k(X)$ is equal to 0, and the decision vector is having n components, decision variable vector. Now, here we have considered the m number of inequality constraints, p number of equality constraints. Now, this constraint set divides the whole design space into 2 parts; one part is the feasible space that is the decision space; where the constraints are being satisfies. And the other space by the constraints at least 1 of the constraint is violated that is the infeasible space.

Now, in the these in the decision space that is in the feasible space; however target is to find that X which minimizes $f(X)$. Now, there are, since these are functions we have considered the functions $f(X)$ $g_j(X)$ and h_k these are all non-linear in nature. And due to the inherent complexity of the non-linear functions, is this not so easy to find out the optimal solution in the decision space. Now, in this connection let me tell you few things

that is in the feasible space, if some constraint particular j is being satisfied with the equality side. Then it is said to be the active constraint rather if the optimal solution X^* is on the some of the constraint, then corresponding constraint is said to be the active constraint otherwise the constraint is inactive.

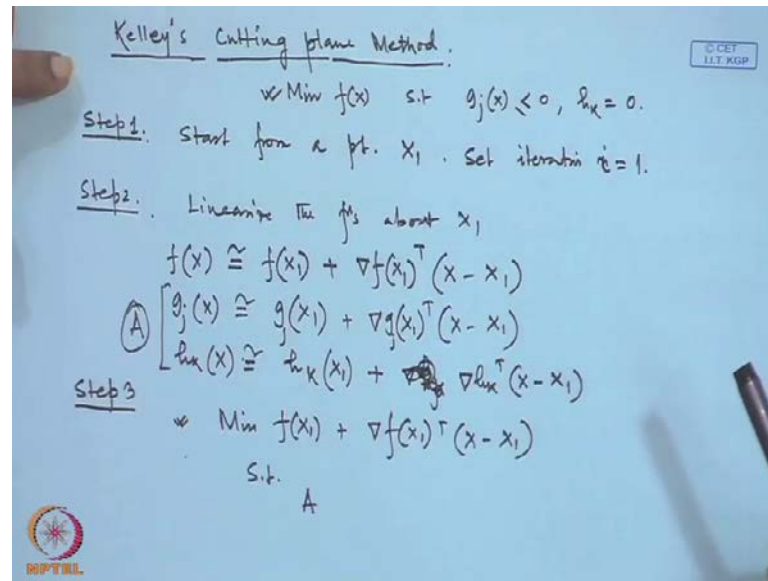
That means that in the decisions process that constraint is not taking part for achieving the optimal solution that is the idea of the active constraint. And the inactive constraint for the inactive constraint; thus, it is being satisfied with the lesser than sign that is the idea. Now, we know how to solve it. For solving this kind of problem as we know the first technique is that KKT conditions through, the KKT condition introducing the Lagrange multipliers we can solve this problem other than this there is a set of methods are available.

The methods can be categorized into 2 parts; one is the direct method and another one is the indirect method. Direct method considers the problem the given non-linear problem in an explicit manner whereas, in the indirect method the given non-linear problem is being solved by sequence of unconstrained minimization problem; for this minimization non-linear problem and that is the indirect method. Direct methods, there are there is a list of direct methods are available.

One is that heuristic search method few of them I will discuss today. Constraint approximation method; that is the cutting plane method suggested by Kelley. And there is another set of methods are available that is the method of feasible direction. In this set of methods there are 2 methods are available; one is that Zoutendijk feasible direction method and second one is the Rosen's gradient projection method. And in the indirect method there is a also list of indirect methods are there by transformation of variable we can solve the given non-linear programming problem.

And, other than this there is a set of penalty function method. In the penalty function method there are 2 kinds of penalty function methods; one is the interior penalty function method and another one is the exterior penalty function method. Thus, we will discuss both the types direct method and indirect method in detail. Let me start with the direct method today.

(Refer Slide Time: 05:30)



Let me start with the constraint approximation method; that is the in other name we can just name it as a Kelley's cutting plane method. This method is very nice method in the same spach the objective, the basic philosophy of this method is that the non-linear functions $f(X)$, $g_j(X)$ and $h_k(X)$ these are in the non-linear form. Now, it is all the non-linear functions are approximated with the linear functions. And a sequence of linear programming problems are being solved. And at the end will reach to the optimal solution of the original problem that is the idea.

Now, for this thing let me consider a problem again minimization of $f(X)$ subject to $g_j(X)$ less than 0 and $h_k(X)$ is equal to 0. Now, the step one; is that this is a sequential process we will just talk we will start from a point X_1 . The point X_1 need not be feasible point. Now, and let us set the iteration number as i is equal to 1. Now, the step 2; is at this point we will just linearise that functions $f(X)$, $g_j(X)$ and $h_k(X)$ with the Taylor series approximation. Once, we are doing that what is happening that we are having the non-linear feasible space; rather the feasible space bounded by the non-linear functions.

Now, we are approximating this non-linear feasible space with the linearly bounded feasible space. And once we are considering point X_1 which is not even the feasible point need not to be feasible point. Then the linear feasible space rather than feasible space bounded by the linear functions that would be the outside of the original non-linear

feasible space that is the idea. So, let me just elaborate the process little more than you can understand it. What we do? We will just linearize the functions at X_1 using Taylor series approximation. Thus, we can write $f(X)$ can be approximated equal to $f(X_1)$ plus $\text{del } f(X_1) T(X \text{ minus } X_1)$.

Similarly, we can write $g_j(X)$ is equal to approximately equal to $g_j(X_1)$, $g_j(X_1)$ plus $\text{grad of } g_j(X_1) T, (X \text{ minus } X_1)$. Similarly, h_k can be written in the similar manner that is the with the Taylor series approximation we are just writing it this is $h_k \text{ grad of } h_k(X \text{ minus } X_1)$. Now, once we are doing that; that means, let me here we have ignored the higher order terms. Because we have to ignore, because we wanted to linearise, we wanted to have the linear functions. If we have, if we want to approximate with the quadratic function. Then we will check the second term as well that is the idea.

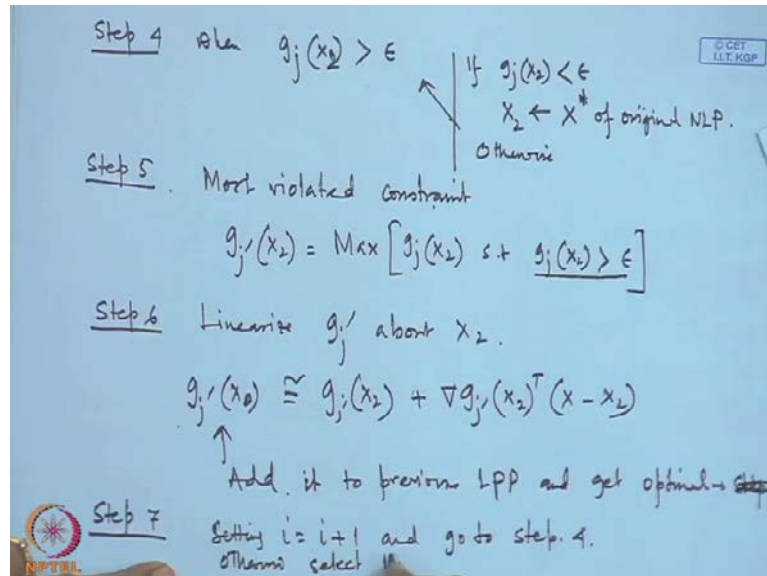
Now, if this is so; then we can reformulate the reformulate the given problem. This problem can be reformulated again with minimization of $f(X_1)$ plus $\text{grad of } f(X_1) T, (X \text{ minus } X_1)$. Rather, let me just do it subject to this set say this is the set A all right. And once this is so it is very clear this is a linear programming problem. And we can solve this linear programming problem and we will get the optimal solution. Now, you see whatever optimal solution we will get for this problem that optimal solution would be the out would be the outside of the feasible space of the original problem.

Once again, I am repeating the thing you see we have taken a feasible space; the original feasible space is bounded by the non-linear functions. Now, we have a approximated each functions with a linear functions that is why since, the point above the point X_1 , X_1 is in the outside the feasible space. Thus, the approximated linear programming problem will have the feasible space which should be the outside the feasible space of the original problem. Thus, as we know that always the optimal solution for the linear programming problem would be in the boundary. And it is in the extreme that is why whatever optimal solution we will achieve here; that would be the point that is that lie outside the feasible space of the original problem that I wanted to say.

Now, once we are getting the optimal solution that optimal solution may satisfy all the constraints or may not satisfy all the constraints. If it is satisfies all the constraints then that optimal solution can be considered as a optimal solution of the original problem. That means whatever optimal solution we have achieved from this linear programming

problem; that is within the feasible space of the original problem. But it is not happening as such because X_1 is not really at the feasible point.

(Refer Slide Time: 12:00)



Now, if it is not being satisfied then we have to adopt the next process that is why let me move to the step 4; when $g_j(X_1)$ is greater than very small number epsilon. Because in general we know that $g_j(X_1)$ is lesser than equal to 0; that is why if $f(X_1)$ is satisfied is not X_1 . Let me consider the optimal solution we are getting out of it that is X_2 that is why we will just substitute X_2 in all inequality constraints. And if we see that it is satisfies then $g_j(X_2)$ must be lesser than equal to 0. But if it is not lesser than 0; that means, g_j some j , $g_j(X_2)$ is greater than epsilon.

Then, in that case that problem cannot be considered as a optimal solution; that is why we can write down the other logic as well if $g_j(X_2)$ is less than epsilon. Then X_2 would be the optimal solution of original problem. We can approximate original NLP; otherwise, go to this step that is the idea. Once, this is so then if we see some $g_j(X)$ are not satisfies or violated; then we will calculate, we will just see. In step 5; most violated constraint we have to identify that. Mathematically, we can say that we will identify that j which is maximum of all $g_j(X_2)$. And such that $g_j(X_2)$ is greater than epsilon; that means, we are considering those constraints which are violated. And we will consider the most violated constraint by considering the max of these.

Once, we are getting that then we will. In the next step we will linearise the corresponding $g_j X$ prime about X^2 . That means, we will get another linear constraint of this. Thus, it will become $g_j X$ prime X^2 would be X would be approximately equal to $g_j X^2$ prime X^2 plus grad of g_j prime (X^2) and $(X$ minus $X^2)$. And we will add this constraint in the previous linear programming problem. That means, what we are just reconsidering the whole feasible space. The original linear feasible space was A .

Now, we are adding another constraint this constraint; that means, the feasible space is being reconstructed. And we will get an optimal solution for that linear programming problem as well. And we will repeat the process in the similar manner we will check whether all the constraints are being satisfied are not. If these are being satisfied we can declare the next optimal solution of the original NLP. Otherwise, we have to repeat from step 4 to step 6 again and again. And we have to identify in each case which constraint is being violated. In that way we will just proceed and we will get the optimal solution at the end. Thus, we can say take next step 7; if we just want to complete this iteration process previous LPP and get optimal. Then go to step 4 otherwise go to step 4 how by setting i is equal to i plus 1 otherwise select the optimal.

Now, what is the advantage of this process is that if I just look at the whole process. Once again, if I just want to summarize the process; that the advantage is that this is like a ((Refer Time: 17:15)) cutting plane method. We are using for integer programming problem here also the same thing; we are just adding the linear constraints one after another original feasible space. And the advantage is that we can solve, we can use the simplex algorithm rather the dual simplex method. In each step we are adding one constraint at each time. And we can use the dual complex very easily. And we will get the optimal solution of the original problem; which is very complicated may be that is the non-linear programming problem.

And, there is a disadvantage as well the disadvantage is that. And this method works only for convex programming problem. Because for the convex programming problem only we can ensure the convergence of the process otherwise in many cases it may diverge. And another disadvantage we can just point it out that is that since we are adding the constraints again and again. There is a chance to become the problem as a big one. And though we can solve it with linear programming problem that is not a problem for us.

(Refer Slide Time: 18:43)

Example Min $f(x_1, x_2) = x_1 - x_2$
subject to,
 $g(x_1, x_2) = 3x_1^2 - 2x_1x_2 + x_2^2 - 1 \leq 0$
 $x_1^* = 0, x_2^* = 1, f_{\min} = -1$

Step 1 $-2 \leq x_1 \leq 2 \quad -2 \leq x_2 \leq 2.$
Min $f = x_1 - x_2$ s.t.
 $x_1^* = \begin{pmatrix} -2 \\ 2 \end{pmatrix} \quad f_{x_1^*} = -4.$

Step 2 Linearise g about x_1^* .
 $\nabla g = 6x_1 - 2x_2$
 $-2x_1 + 2x_2$

© CET
I.I.T. KGP

NPTL

Let us apply this method for a numerical problem; then it is it would be clear to you. Let us consider a linear program, a non-linear programming problem; that is minimize $f(x_1, x_2)$ is equal to $x_1 - x_2$. We have considered the linear objective function that does not matter we can consider non-linear as well. And there is only 1 constraint g that is this problem I have taken from the optimization book written by S. S. Rao this is the constraint for us and it is. And if we just solve one, if you want to solve it in otherwise in by using the Kuhn tucker conditions. We will see the optimal solution will be x_1^* would be is equal to 0, x_2^* would be is equal to 1 and f_{\min} would be minus 1.

Now, we will use the cutting plane method. And we will see how the problem can be solved. Now, if we see the feasible space for this; the feasible space may be unbounded. And this we want to bound it that is why in the step 1; let us consider another constraint that is x_1 is in between minus 2 and 2. And x_2 in between minus 2 and 2 these are given. Now, first what we will do, we will solve minimize f that is $x_1 - x_2$, subject to these 2 constraints. And we will see that we will get optimal solution that would be starting point of our problem x_1^* would be is equal to (minus 2, 2). And we will get the f value at this point $f_{x_1^*}$ is equal to minus 4.

Now, once this is so we will just in the next step. We will linearise g about x_1^* that can be done very easily. First we will find out the grad of g ; grad of g would be is equal

to 6×1 minus 2×2 and minus 2×1 plus 2×2 . And we will consider this grad of g at the point minus 2, 2 and minus 2 and 2.

(Refer Slide Time: 21:40)

The image shows a whiteboard with handwritten mathematical work. At the top, the function is approximated as $g(x_1, x_2) \approx g(x_1) + \nabla g(x_1)^T (x - x_1)$, which simplifies to $\approx -16x_1 + 8x_2 - 25 \leq 0$. Below this, the word "Step 3" is written. The problem is stated as "Min $x_1 - x_2$ " with the constraint "Sub to, $-16x_1 + 8x_2 - 25 \leq 0$ " and " $-2 \leq x_1, x_2 \leq 2$ ". The optimal solution is given as $x_2^* = \begin{cases} -0.5625 \\ 2 \end{cases}$ and the functional value is $f = -2.56$. At the bottom, the subgradient is denoted as $g_j(x_2^*)$. There are logos for "IIT KGP" and "IITBTL" on the board.

And, ultimately we can linearise it with the function $g(x_1, x_2)$ would be approximately equal to that is $g(x_1)^*$ plus grad of $g(x_1)$ and $(x - x_1)$. If we just do it we will get these value as approximately equal to minus $16x_1$ plus $8x_2$ minus 25. And since this is the approximated linear function of the given non-linear function. Let us consider the next step with this constraint. Thus, the next step we will consider the next linear programming problem as minimization of $x_1 - x_2$, subject to this constraints plus 25 less than equal to 0; and x_1 and x_2 both minus 2, 2.

Because we wanted to bound the feasible space that is why we have consider this constraint. Not only that this constraint helps us in finding out the initial starting point even. Now, we will solve it. And once we will solve it, we will get the next optimal solution x_2^* that would be is equal to minus 0.5625 and 2. Where, the functional value would be is equal to minus 2.56 previously it was minus 4. Then next we could see the functional value has decreased that is why if we go further iteration there is a chance to converse the process. But if we just see then we will see that again that here the value of $g(x_2^*)$ that is very high it is given 23. If I just summarize the whole process in the table we can see it.

(Refer Slide Time: 23:54)

Iteration No. i	New Linearized Constraint	Sol ⁿ of the Approximating LP Problem X_{i+1}	$f(X_{i+1})$	$g_i(X_{i+1})$
1.	$-2 \leq x_1 \leq 2$ and $-2 \leq x_2 \leq 2$	$(-2.0, 2.0)$	-4.000	23.000
2.	$-16.0x_1 + 8.0x_2 - 25.0 \leq 0$	$(-0.563, 2.000)$	-2.563	6.199
3.	$-7.375x_1 + 5.125x_2 - 8.199 \leq 0$	$(0.279, 2.000)$	-1.722	2.120
4.	$-2.332x_1 + 3.444x_2 - 4.120 \leq 0$	$(-0.530, 0.838)$	-1.367	1.431
5.	$-4.853x_1 + 2.735x_2 - 3.431 \leq 0$	$(-0.053, 1.160)$	-1.213	0.478
6.	$-2.629x_1 + 2.427x_2 - 2.478 \leq 0$	$(0.427, 1.485)$	-1.058	0.484
7.	$-0.411x_1 + 2.117x_2 - 2.484 \leq 0$	$(0.171, 1.207)$	-1.036	0.132
8.	$-1.390x_1 + 2.072x_2 - 2.132 \leq 0$	$(0.018, 1.040)$	-1.023	0.047
9.	$-1.972x_1 + 2.045x_2 - 2.047 \leq 0$	$(-0.166, 0.840)$	-1.007	0.068
10.	$-2.678x_1 + 2.013x_2 - 2.068 \leq 0$	$(-0.073, 0.930)$	-1.003	0.017

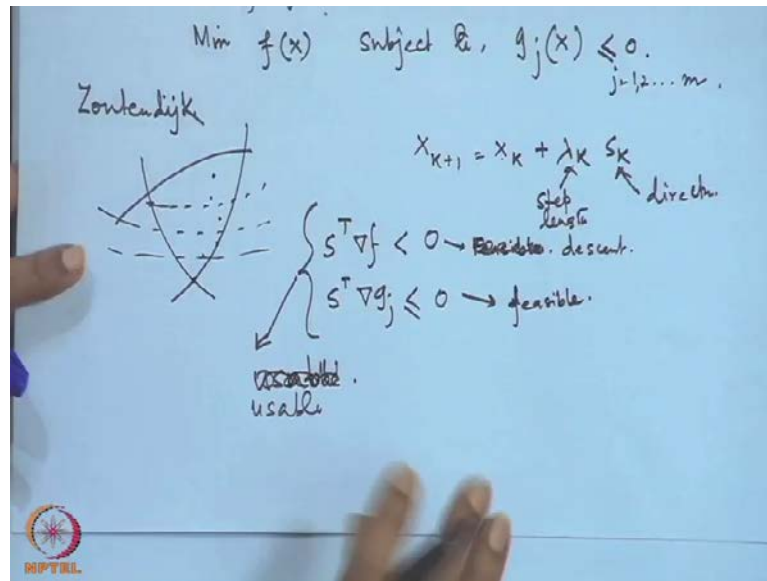
This is the initial state for us as we have considered minus 2, 2. And here the g_j value was 23. Now, in the next step we have considered these constraint we got this optimal solution; we could see the functional values minus 2 as I will showed you. And then we will see again the g_j value is coming very high that is 6. That means, we are still this value is very high that is why we will we have to consider we have to linearise these g_j with respect to this point again. Otherwise, we will be very far from the boundary of the feasible region original the given problem.

That means, once we are linearising g_j with respect to this point we will get this constraint. And in the next problem we will add this constraint again; that means, we will add this constraint here itself. And once we are adding this constraint we will get another point that would be 0.2 and 2 functional value is again decreasing g_j value is decreasing again and again. Thus, we will just repeat the process again and again 10 times. We could see that functional value is converging to minus 1. And g_j value is almost 0 that is why that is quite acceptable. And as we say in the original if we just apply the K K T condition there also we could see the f mean value was coming minus 1. And the corresponding optimal solution was coming 0, 1.

Here also the same thing that we are getting the optimal solution as 0, 1 almost 1. That is why this could be taken as the approximate optimal solution of the original non-linear programming problem. And that is all about the starting the Kelley's cutting plane

method. Through, which we are getting the sequential linear programming problem. And we can solve the problem with the dual simplex technique. And very nice this method is nice. In that sense, that we need not tackle the non-linear functions term very easily with the linear functions only we are getting the solutions. But the disadvantage as I said that it is applicable for the convex programming problem. For the non convex problem it is not easy to consider.

(Refer Slide Time: 26:32)



Now, let me go to the next method that is the method of feasible direction. This is another direct method because we are considering the problem. We are addressing the constraint original constraint say it original problem. Thus, this is not an indirect method. But how? What is the basic philosophy of this method? Let me just first summarize it then I will go to algorithm of this method. The method states that minimization of $f(X)$. Let me consider the constraints $g_j(X)$ that is the non-linear inequality constraints; where, f and g are non-linear functions. In the similar manner there are m number of inequality constraints.

Now, the method of feasible direction there is a beauty of this method is that this is also a searching technique we will start from a point. And we will iteratively we will proceed in such a manner that functional value is decreasing in each iteration. Not only that we will remain in the feasible space that is the idea of the method of feasible direction. And the mathematician Zoutendijk invented this method. And this method is very robust method.

And this method is applicable very much. And we are using this method here in there for non-linear programming problem.

Let me just elaborate that method for you. Now, as we have seen in the k th condition that is we start from a point X_k ; then if this point is a regular point rather the feasible point. Then that feasible; if I want to move to the feasible point then the idea is that we have infinite number of directions through which we can do. Because as we have seen for the unconstrained optimization technique. We have adopted this process we started from X_k . And we are moving to X_{k+1} with the considering λ_k as the step length.

And, this is the direction; that means, from X_k point if this is a point for us. We are moving to the next point this is specific step length. And there is a specific direction; that means, X_k is going to X_{k+1} . Since, this is X_k we can proceed in different way. In infinite number of directions we are having in which direction should I move? Zoutendijk has suggested some technique he has said just moved through the usable direction. Then we will reach to a better we will adopt a better that is the idea. Now, let me tell you what is that? As we know that if I move from a point X_k to the next point X_{k+1} with the direction S_k ; then we can say that S is a feasible direction.

If you remember we have done that in our k th condition when we have described k th condition in detail. We have said that this direction would be this dot product that is if this is the feasible direction S . And if the $\text{grad } f$ is the gradient of the objective function always it makes the obtuse angle. That means, lesser than 0 not only that we have also consider to remember that $S^T \text{grad } f$ is always less than is equal to 0 . That also makes the obtuse angle for the minimization problem.

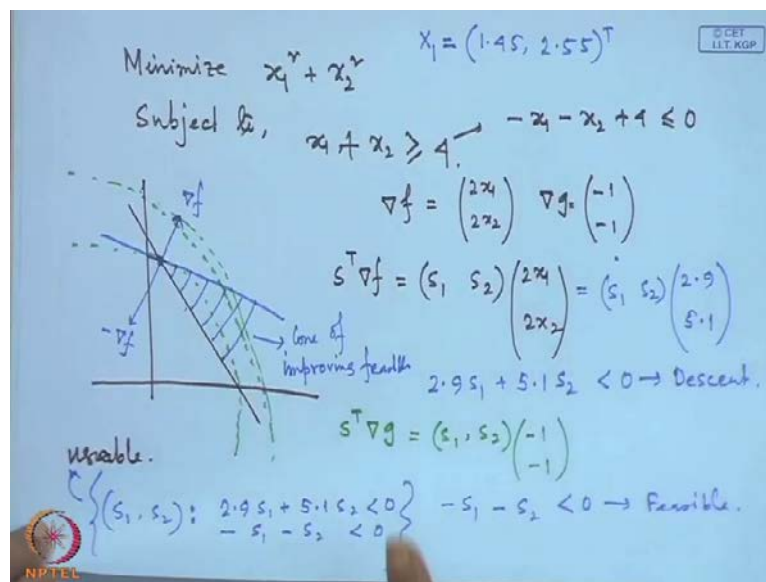
Now, Zoutendijk he said that this direction is being this is the feasible direction. If I move with this condition this is a feasible direction. And if I move with this condition that no sorry, this is not the feasible direction that is the descent direction. That means, for the minimization problem if I consider the direction as S . And if we consider $S^T \text{grad } f$ if it is giving the less than value; that means, if it is making the obtuse angle. That if the direction makes the obtuse angle with the gradient of the objective function that is the descent direction. That means that direction gives me better functional value for the minimization problem; that means, the functional value improves. We are getting the

better minimum in this direction not only that if S is the direction; if we consider $S^T \nabla g_j$ where j can run from 1 to m .

Now, this direction is the feasible direction; then it makes again the obtuse angle. And this is called the feasible direction. And he has suggested that if any direction consider both the conditions together; that means, any direction x which is as well as feasible; then that direction can be said as usable direction. Thus, according to Zoutendijk he has suggested that always move in direction where the direction is usable all right. Let me just whatever philosophy said. Let me summarize whole thing mathematically.

Now, how far we can move that is a next question. Because if I move by considering both the conditions together then from this point, I will go to this point, then I can go to this point, then I will go to the boundary of the feasible space. In this way I will move. But we know if this is the objective function for us. And optimal will come here that is why and that is the KKT point. And at KKT point, we will see that this will all satisfy with the equality sign that part we are just considering in the next.

(Refer Slide Time: 33:18)



Before doing that thing let me consider a minimization problem. And let us try to visualize what is the feasible direction? What is the meaning of that? Plus x_2 greater than equal to 4. If it is draw this one then we can see that this is the circular region. And this is a line for us and this is the circle. And if I just move further this is a circular region do not mistake it. Then the optimal will come here is the optimal solution all

right. Because we have considered the constraint $x_1 + x_2 \geq 4$; that means, this region, now minimization problem.

Now, here mathematically we can say that let me say which direction e_t is grad of f it is $2 \times 1, 2 \times 2$ grad of g equal to minus 1, minus 1. If we considered the function g as $\text{minus } x_1 \text{ minus } x_2 \text{ plus } 4 \text{ less than equal to } 0$. Because I was trying to explain the methodology with $g \text{ lesser than equal to } 0$ that is why grad g would be minus 1, minus 1 this case. If this is so then we will say that this is the descent directions for us. And how what it would be if I can consider the s_1, s_2 is the descent direction for us and how it would be if I consider the vector S_1, S_2 is the direction; then it would be $2 \times 1, 2 \times 2$.

Let us start the process with an initial point $x_1 (1.45 \ 2.55)$ this is very much satisfying the constraint this is the point for us. Then we can see that $S^T \text{ grad } a$ for this point it is coming as $(S_1, S_2), (2.9 \text{ and this is } 5.1)$. Now, we can say that $2.9 S_1 \text{ plus } 5.1 S_2 \text{ less than } 0$. It is the feasible it is the descent direction; that means, if we move from this in this direction then that would gives us the descent direction. Now, what is that direction at all? Now, this is the gradient of f rather this is the minus grad f . Because the functional value is decreasing further and further. And if I just take a tangent here that tangent the half space of that; that means, this region would be the descent direction for f that is the idea.

Now, let us move to the next what is the feasible direction; that means, we will calculate $S^T \text{ grad } g$. This is coming $(S_1, S_2), (\text{minus } 1, \text{minus } 1)$; that means, we can say that $\text{minus } S_1, \text{minus } S_2 \text{ lesser than } 0$ this is the feasible direction. Now, what is the usable direction? Then usable direction is that direction; that means, that set of (S_1, S_2) which satisfies both the conditions together that is $2.9 S_1 \text{ plus } 5.1 S_2 \text{ lesser than } 0$ that is the descent direction. And $\text{minus } S_1, \text{minus } S_2 \text{ less than } 0$; that is a feasible direction, which direction it could be? This is this direction is the sorry, this direction is the usable direction give as said by the Zoutendijk all right. What is this direction then? This direction would be nothing but this cone. That is why this cone is being said as the cone of improving feasible direction clay.

(Refer Slide Time: 38:51)

Find S , $s^T \nabla f < 0$
 $s^T \nabla g_j \leq 0$.
Artificial. $\alpha \rightarrow \alpha = \text{Max} [s^T \nabla f, s^T \nabla g_j]_{x_1}$
Find S Minimize α
s.t. $s^T \nabla f \leq \alpha$
 $s^T \nabla g_j \leq \alpha \forall j$
 $\rightarrow -1 \leq s_i \leq 1$ $(s^T s = 1)$

Now, whatever thing I said numerically same thing we will just translate in mathematical form that we will see that we will find S ? Forgetting the usable direction such that $S^T \nabla f$ is less than 0, $S^T \nabla g$ is less than equal to 0 all right. Now, if I consider 1 artificial variable. But how to solve it mathematically? That is why we are formulating a linear we are formulating a problem optimization problem. So that we can consider the we can very easily we can consider the direction. Artificial variable α such that α is equal to max of $S^T \nabla f$ and $S^T \nabla g_j$ these are all g_j at point X_1 . Because we are X_1 is our starting point all right. Now, you see all the values are negative values $S^T \nabla f$ is negative, $S^T \nabla g$ is negative that is why α is negative here.

Once, α is negative what is our target? Our target is the process has been designed in such a way that if we just move from a point X_1 ; then we will move to that direction. So that if I just move in this direction; then there is a chance that if I take a λ such that I will be out of the feasible space that λ should not be that. So that I am out of the feasible space always I want to be in the feasible space. That is why select suitable λ . So that we will remain in the next if this is X_2 for us.

Then, we will remain within the feasible space; that is why we will try to remain far from the boundary of the constraint that is the idea. That is why we are trying to maximize α ; that means, we wanted to say that we will reduce α till it becomes negative. Thus, we can formulate a problem as minimize α ; sorry, minimize α such that S

$\nabla^T g \leq \alpha$. Because α is maximum of these 2 and $\nabla^T g_j \leq \alpha$ for all j is equal to 1 to n here also for all j all right. This is the problem; that means, that find S ? Such that this is being satisfied that is why finding out the visible direction is not difficult for us. Because that direction we will search which will satisfy this optimization problem.

Let us use this one after doing certain modification. Now, you see one thing is that we are interested about the direction S not about the magnitude S is it not. Because this is the direction we are looking for, not looking for the magnitude of the moving direction. That is why always we will try to normalize S by putting $\|S\| = 1$; rather we can say that $S^T S = 1$. That is normalization process we can do it other if we consider i -th component. How many components of A should be, A should be have n number of components. That is because we are moving in the design space where n number of that is n dimensional space.

Because the number of decision variables are n that is why we wanted to consider $S^T S = 1$, otherwise we may consider S_i in between minus 1 and 1. Then the advantage of considering this constraint is that we can bound this feasible region. Otherwise, this space may be unbounded for certain case all right or else we can consider this or we can consider this is not being considered. Because always it will give you a quadratic function; but it will give you a linear function all right that is why we are considering this.

(Refer Slide Time: 43:22)

$$\begin{array}{l}
 \text{Min } \alpha \\
 \text{s.t. } s^T \nabla f \leq \alpha. \\
 s^T \nabla g_j \leq \alpha \quad \forall j \\
 -1 \leq s_i \leq 1
 \end{array}
 \quad \left| \quad
 \begin{array}{l}
 \beta = -\alpha \quad \text{Find } S. \\
 \text{Min } -\beta \\
 \text{s.t. } s^T \nabla f + \beta \leq 0 \\
 s^T \nabla g_j + \theta_j \beta \leq 0 \quad \forall j \\
 -1 \leq s_i \leq 1. \\
 \theta_j > 0 \\
 \theta_j \rightarrow \text{Push factor } \theta_j = 1
 \end{array}$$

$$b_j = s_j - 1 \quad 0 \leq b_j \leq 2$$

Now, we are doing certain modification for this problem. Let me just write down this modification in the next. Minimization alpha subject to $S^T \text{grad } f \leq \alpha$, $S^T \text{grad } g_j \leq \alpha$, for all j is equal to 1 to m . And we have considered the i -th component all i -th all components of S is in between 1. Now, we are modifying this problem by considering another variable beta is equal to minus alpha. If we just do that we are getting another problem. Just see minimization of minus beta subject to $S^T \text{grad } f + \beta \leq 0$; $S^T \text{grad } g_j + \beta \leq 0$; and S_i in between minus 1 and 1.

What is the advantage of considering that. Because you see this is a linear, these are all the linear functions what is that what are the decision variables? Decision variables are S_i because we wanted to find X, S ? That is the is the feasible direction. Now, alpha is the negative value. That is why we wanted to impose that the decision variable that is one of that is beta that must be positive. That is why this can be considered very nicely that beta is positive here. Now, there is another modification Zoutendijk suggested that for the constraint set the tone of the improving feasible direction he wanted to capture the angle in between that. That is why he has introduced another term in the constraint set that is θ_j , for all j . And this θ_j is being named as push factor rather this is capturing that is angle between the angle of the cone.

Now, in general theta is always a positive. Because if theta j is negative is not possible because it is always it is acute angle for us. Not only that theta j can be if it is equal to 0 then it is tangent altogether. And always we are considering theta j greater than 0. And by default we are considering the theta j equal to 1 in each case. Otherwise, we can consider it as a variable, we can consider as a. Now, for finding out feasible direction as a as a whole we can say this is the optimization problem we have to solve at each iteration.

Now, this can be further be modified with another modification let us see. So that we will get a function where S i is not even negative value. We wanted to have a positive value here. again because here S i is coming negative. And we know for the linear programming problem the simplex algorithm works only if non negativity constraint we have to ensure that. That is why let us consider another variable t i is equal to S i minus 1 all right. Then since S i is in between minus 1 and 1, t i would be in between 0 and 2 all right. No, rather it is 1 minus S, t i would be is equal to, S i would be is equal to t i plus 1 all right in the similar way. Now, if this is so if we just write down this problem again this problem is being formulated in this manner.

(Refer Slide Time: 47:39)

Find (t_1, t_2, \dots, t_n) which

Minimize $-\beta$

Subject to:

$$t_1 \frac{\partial g_1}{\partial x_1} + t_2 \frac{\partial g_2}{\partial x_2} + \dots + t_n \frac{\partial g_n}{\partial x_n} + \theta_j / \beta \leq \sum_{i=1}^n \frac{\partial g_i}{\partial x_i}$$

$$\forall j = 1, 2, \dots, m.$$

$$t_1 \frac{\partial f}{\partial x_1} + t_2 \frac{\partial f}{\partial x_2} + \dots + t_n \frac{\partial f}{\partial x_n} \leq \sum_{i=1}^n \frac{\partial f}{\partial x_i}$$

$$0 \leq t_i \leq 2.$$

$$\beta > 0$$

Find t_1, t_2, t_n in place of S_1, S_2, S_n which minimizes minus beta subject to. If we just elaborate here we are putting t_i is equal to S_i plus 1. If I just put S_i plus 1 then we are considering as $t_1, \frac{\partial g_1}{\partial x_1}$ by $\frac{\partial x_1}{\partial x_1}$ plus $t_2, \frac{\partial g_2}{\partial x_2}$ by $\frac{\partial x_2}{\partial x_2}$, t_n now g_1 . Because

first constraint we are considering $x_n + \theta_j \beta_j \leq \sum_{i=1}^n x_i$ is equal to 1 to n, Δg_j by Δx_i all right. Because we have considered S_i is equal to $t_i - 1$. Now, this is for all j that is why let me put j here, for all j from 1 to m . And, for the objective function we are considering 1 constraint only less than equal to $\sum_{i=1}^n x_i$ is equal to 1 to n, Δf by Δx_i . And t_i is in between 0 and 2 what else β_j is greater than 0. This is the linear programming problem right. That is why that would be considered as the usable direction that is all about the usable direction.

(Refer Slide Time: 49:40)

Zoutendijk's method of feasible direction:

STEP 1: Start with initial feasible point x_1
Evaluate $f(x_i)$ and $g_j(x_i) \forall j$. Set $i=1$

STEP 2: If $g_j(x_i) < 0$ (i.e. interior pt.)
feasible direction $s_i = -\nabla f(x_i)$
Normalise s_i and go to step 5
If atleast one $g_j(x_i) = 0$ go to step 3.

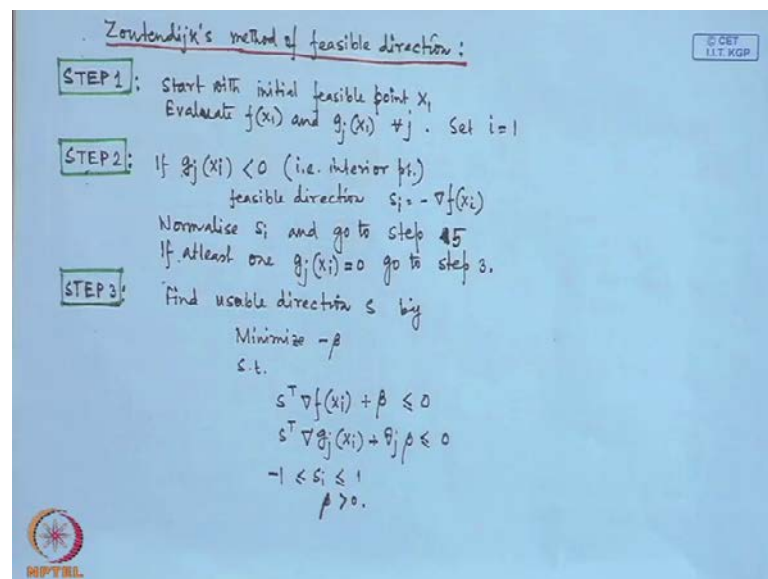
STEP 3: Find usable direction s by
Minimize $-\beta$
s.t.

$$s^T \nabla f(x_i) + \beta \leq 0$$

$$s^T \nabla g_j(x_i) + \theta_j \beta \leq 0$$

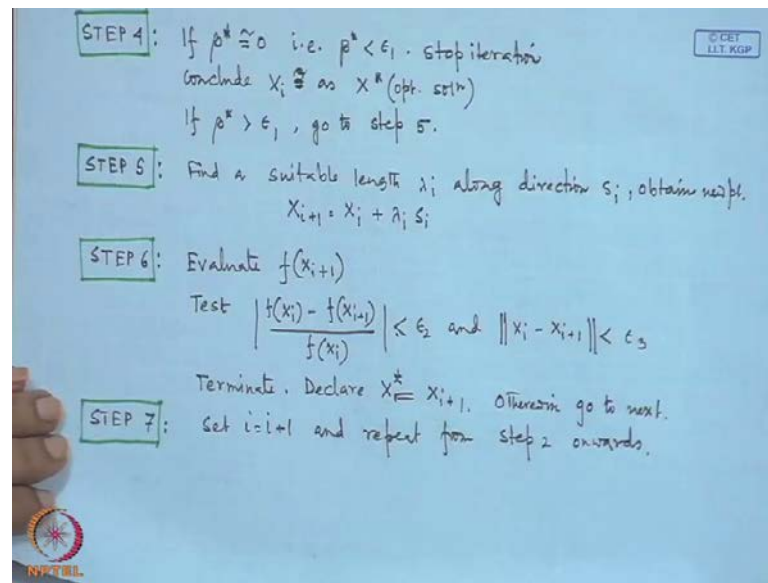
$$-1 \leq s_i \leq 1$$

$$\beta > 0.$$



If I go back to the methodology that is the Zoutendijk method. For feasible direction this is the process for us start with an initial point X_1 ; evaluate $f(X_1)$, $g_j(X_1)$. And, if $g_j(X_1)$ is lesser than 0; that means, we are in the interior point. And in the feasible direction S_i would be is equal to minus grad of f . And we have to normalize S_i and we have to go the step 5 if we are in the interior point. But if we are in the boundary point then we have to find out a feasible direction by this process. That is minimization of minus beta $S^T \text{grad } f + \beta \leq 0$ in this way we are finding out the. There is a another constraint that is the beta is greater than 0 that is we are getting the usable direction alright.

(Refer Slide Time: 50:40)



Now, whatever direction we are getting if this is almost equal to 0 then we can stop our iteration; that means, there is no further possibility for improvement. But if we see that it is not equal to 0 rather beta star is less than epsilon. Epsilon is the very small positive value. Then we will stop iteration we will declare that the corresponding X_i as the optimal solution. If it is not so then we need to go the next step. Because we have to find out a suitable direction; because how we will find out lambda i? And S_i in such a way that we can move to the next point $i + 1$. And we will do the process how we will find out the functional value at this point.

And, if the difference of the functional value rather the rate of difference is very small. And if X_i s are very close the norm that is because these are all the vectors these are very close to each other. Then we will declare that one as the optimal solution otherwise we have to move to the we will we have to go to the next; if it is not optimal. Because we have to repeat the process again and again; we will see whether X_2 is in the boundary of the feasible space or not. If it is in the boundary of the feasible space something is there, if it is not in the boundary of the feasible space, it is in the boundary of the feasible space; then we have to repeat this process from step 2 again. But if it is in the interior point we can proceed further with this step length and we will move to the next.

(Refer Slide Time: 52:31)

Minimize $f(x_1, x_2) = x_1^2 + x_2^2 - 4x_1 - 4x_2 + 8$
 Subject to,
 $g(x_1, x_2) = x_1 + 2x_2 - 4 \leq 0$.
 Start with $x_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$.
STEP 1 Here $g(x_1) < 0$, Thus $S_1 = -\nabla f(x_1) = \begin{Bmatrix} 4 \\ 4 \end{Bmatrix}$
 After normalization $S_1 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$.
STEP 2 New point $x_2 = x_1 + \lambda_1 S_1 = \begin{Bmatrix} \lambda_1 \\ \lambda_1 \end{Bmatrix}$
 $f(x_2) = 2\lambda^2 - 8\lambda + 8$.
 \rightarrow at $\lambda = 2$ $\frac{df}{d\lambda} = 0$ $g_1(x_2) = 0$
 But $x_2 = \begin{Bmatrix} 2 \\ 2 \end{Bmatrix}$, $g_1(x_2)$ is violated.
 $g_1(x_2) > 0$

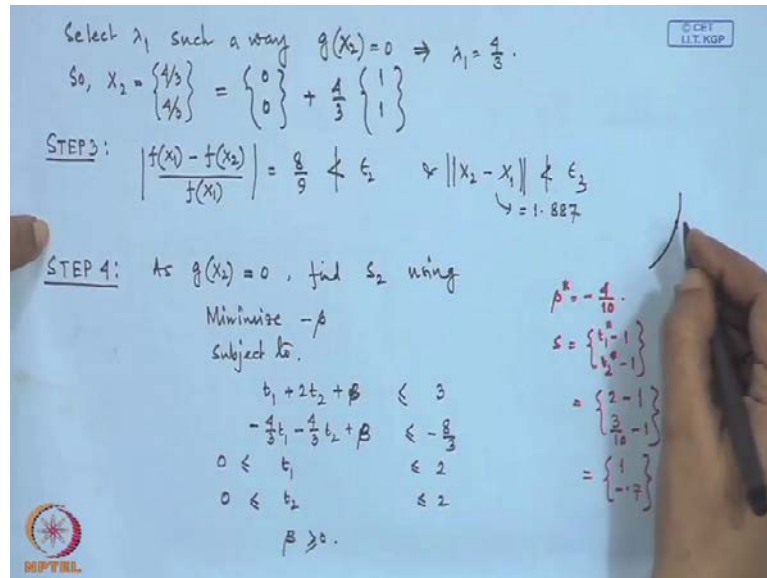
And, I have prepared one problem for you to understand in a better way. Just look at this is the problem for us $f(x_1, x_2)$. And this is a non-linear x_1 square plus x_2 square minus $4x_1$, minus $4x_2$ plus h and g is this one. Now, we are starting the process from $0, 0$; since, $0, 0, 0, 0$ the point is very much in the feasible space not only that the point is in the interior point it is not the boundary point. Because with this point it is not being satisfied with equality sign. If we can consider the starting point as $2, 1$ then certainly it is a boundary point and we have to do some processing.

So that we are not out of the feasible space; otherwise, we can remain with the next. Just you see what we are doing as I said $g(x_1)$ is lesser than 0 and we are getting the next direction as $1, 1$. Because we are considering the descent direction of f that is $4, 4$ we are normalizing it and we are getting the direction $1, 1$. And our next task is to find out which one we are trying to move from x_1 to x_2 in such a way that x_2 is equal to x_1 plus λS_1 . S_1 is this one, x_1 is given $0, 0$ what is λ ? We have to just find out in the next.

For doing that as we did it for the unconstrained optimization problem here also the same thing. We will do the same we will just find out the function f and we will see for which value of λ f is minimum. That is why we are getting λ is equal to 2 then f is minimum. But if λ is equal to 2 then the next point x_2 would be is equal to $2, 2$. Just look at the constraint if it is $2, 2$; then I will be out of the feasible space; that means, g_1

is violated that is not accepted. That is why what should be done we have to reconstruct lambda 1, we have to suitably select lambda 1 in such a way that g 1(X 2) will be at the most on the boundary all right. But here g 1 (X 2) is coming greater than 0 it is not accepted. That is why we will select lambda in such a way that it is in the boundary.

(Refer Slide Time: 55:02)



If this is so very easily we can select by considering g(X 2) by substituting X 2 in g we will see the lambda 1 value is coming 4 by 3. That is why we are saying X 2 is equal to 4 by 3, 4 by 3 which is nothing but X 1 plus lambda 1, plus S 1 all right. Now, we will select whether X 2 is optimal or not. Now, we see the difference of f (X 1) and f (X 2) that is not very low, that is very high, that is not acceptable. That is why X 2 is not can be cannot be considered not only that norm of difference X 2 minus X 1 that is also very high that is not very small value.

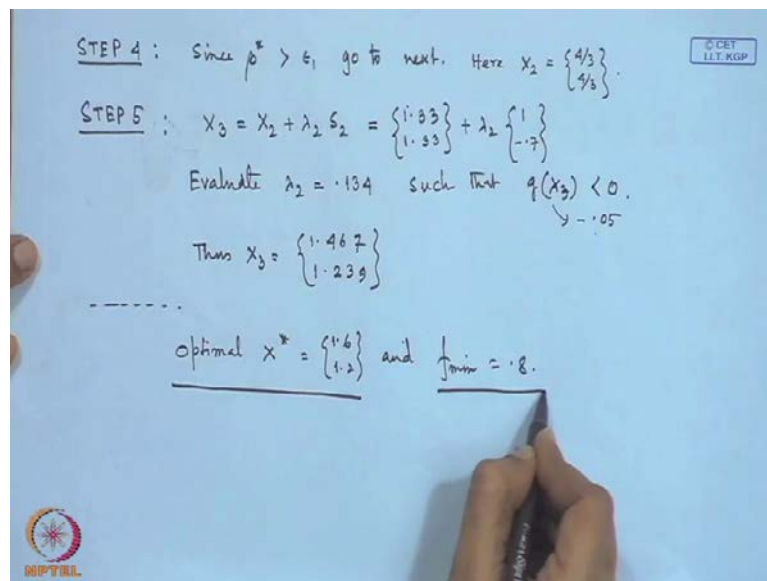
Here, we are considering epsilon 2 and epsilon 3 as a pre assigned very small positive value may be 0.0001 and 0.0002 even we can consider the same value. That is why you see not only once we are reaching to X 2. X 2 is in the boundary on the boundary of the feasible space because it is satisfying the constraint with the equality sign. That is why we have to select the next process; the next S 2 that is the next direction in such a way when I should not be out of the feasible space; that is the check for us.

Because if I you see I am in the boundary there is every chance that I will be out of the feasible space. That is why always select that direction that is usable direction; that

means, that lies within the direction the cone of improving feasible direction. That is why in the next case what I am doing? I am selecting the usable direction that problem with that linear programming problem. As I explained you before that is minimized minus beta subject to the t_1, t_2 ; I just showed you the that is the this one.

This is the whole linear programming problem writing again for this one and we are getting these linear programming problem. And we are getting the beta star is equal to minus 4 by 10. Now, once beta star is equal to minus 4 by 10 as we have considered t is equal to $S_i + 1$ that is why S could be t_i, t_1 star minus 1, t star from here we are getting t_2 star we are getting from 3 by 10. That is why my next that is S_2 could be 1 minus 0.7; that means, this is the usable direction in the next. That is why from S_2 I can move in this direction.

(Refer Slide Time: 57:56)



Again the problem is still there. What should be the step length? That is why we have to add on to the next process that is lambda 2. We have to find out lambda 2 in such a way that I will remain within the feasible space. Not only that we are selecting the descent direction. And that is the usable direction that is why we are selecting lambda 2 in a suitable manner. So that I remained with the use in the usable direction. If I just repeat the process as we did for in step 2 the similar way we are selecting lambda 2 as 0.1 3 4.

Then, we see that $g(X_3)$ is within the feasible space and for this X_3 . That is why this lambda 2 is quite acceptable for us that is why we can move to the next. In this way we

just proceed further and further after few steps we can see the optimal solution is coming this way. And we will get the functional value as 0.8. And this is the method where we are using the method of feasible direction. There is another alternative method is also there that is the Rosen's gradient projection method.

Now, today I am not going to detail that method. But either of these 2 methods we can use it. The beauty of this method is that we can very nicely we can adopt a search iterative process in such a way that always we will remain within the feasible space. And that this method is very popular method that is why I discuss this one with this today I am concluding. In the next I will discuss few methods on indirect searching technique for constraint optimization problem that is penalty function method.

Thank you very much for today.