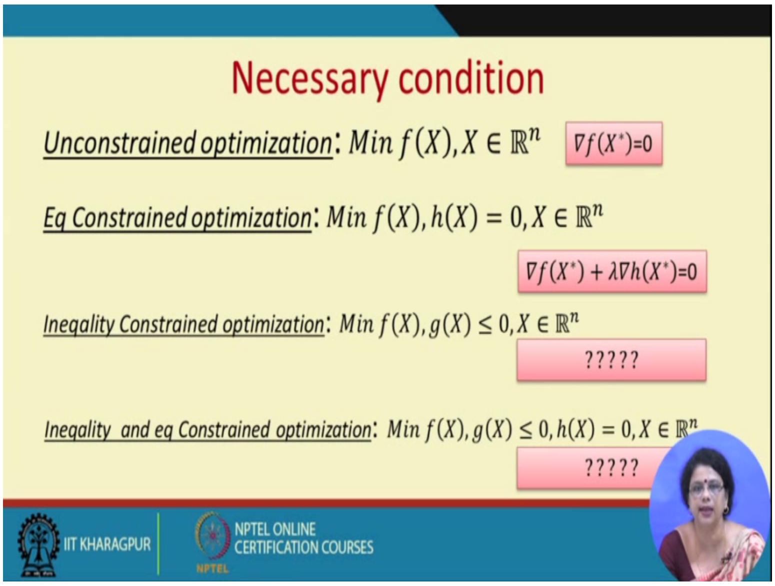
Constrained and Unconstrained Optimization
Prof. Debjani Chakraborty
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 46
Constrained NLP – I

In continuation to my previous class today I will deal with the non leaner programming problem with the inequality constraint. But, before to describe a with the inequality constraints, I will say few things few properties with the non-linear programming problem with equality constraints.

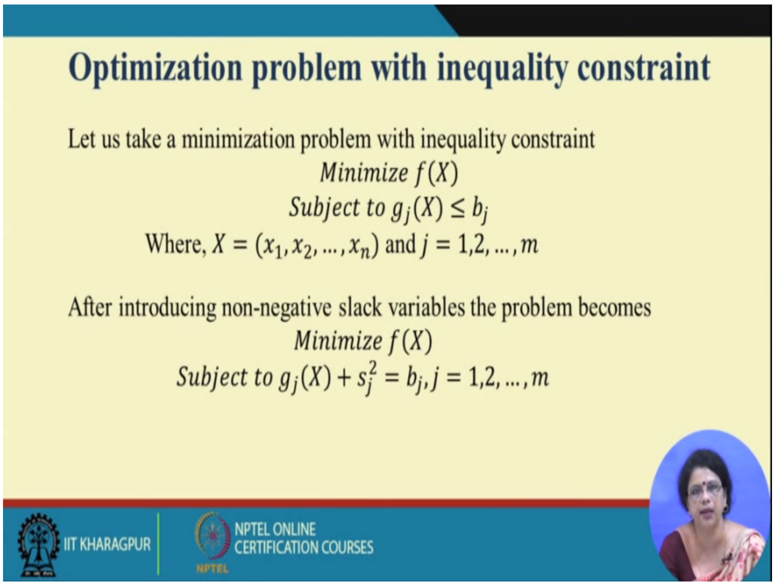
(Refer Slide Time: 00:45)



Now, we know that if we consider a unconstrained optimization problem minimization of f x, we know the necessary condition for this would be gradient of f X at the optimal point must be is equal to 0. Now this is for the unconstrained optimization, but if you consider the equality constraint optimization problem, we need to find out the lagrangian function. I explained you in the last class, in the lagrangian function by considering the lagrangian parameter. If we just go for the necessary condition then this is the necessary condition, grad f plus lag lagrangian multiply a lambda into grad of h X star equal to 0.

Now if we consider the inequality constraint instead of the equality constraint that minimization of f X g X less than is equal to 0. Where the here also the same that decision vector X is in the n dimensional space, now what should be the necessary condition for that. Now we have to find it out not only that; if we have the constraints, which are of the mixer type that we have the equality constraints as well as the inequality constraints, that is also we need to do what would be the necessary conditions for this problem. Now, in continuation to my previous class, where I dealt with the constraint equality constraint optimization problem.

(Refer Slide Time: 02:29)

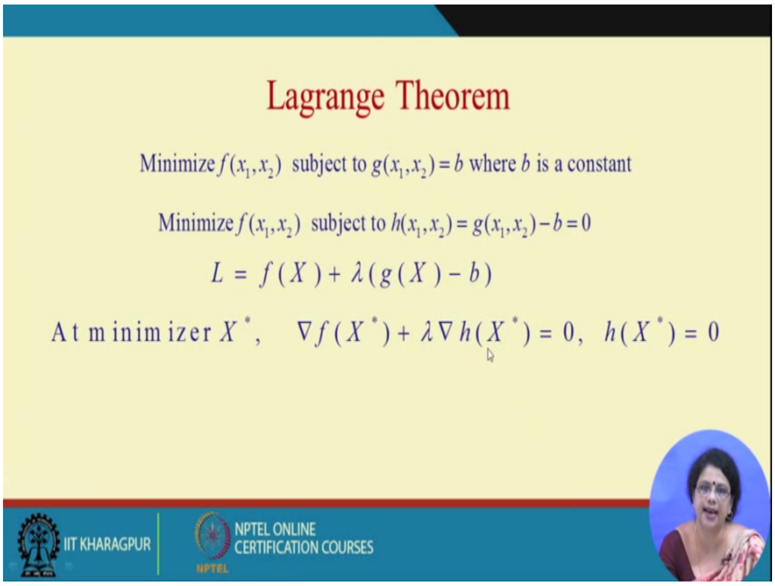


Now, we are here to find out what will happen when we will have the m number of inequality constraints. Now, as you have done in linear programming, whenever you are having any inequality constraint then either you introduce a slack variable or you introduce the surplus variable, depending on whether the inequality sign is less than equal to or greater than equal to.

Now, if the constraint is of less than equal to type, you introduce the constraint in introduce the slack variable, that why here we are con introducing a nonnegative slack variable that is s j, Since we are considering the slack variable as nonnegative that is why we are considering s j square. That is why we can make this problem as a equality type constraint. That is why you might be understanding that you must be guessing that we can handle this problem, just like our previous case where we dealt up with the equality type constraint and their non-linear objective function. we can do the same thing is it not we can construct the Lagrange function, we can introduce the Lagrange multiplier etc; now in that line there is a well known method that is called the Kuhn-tucker condition. Karush Kuhn-tucker condition, in short we are calling it as a KKT Condition.

There we are extending the Lagrange problem, Lagrange theorem in the inequality type constraints. Now, that is why we will construct the Lagrange function as well, but in state of introducing the Lagrange multiplier here we will introduce the KKT multiplier and the significance of KKT multiplier how to handle the problem with inequality constraint, where the slack variables are there all these things I will discuss.

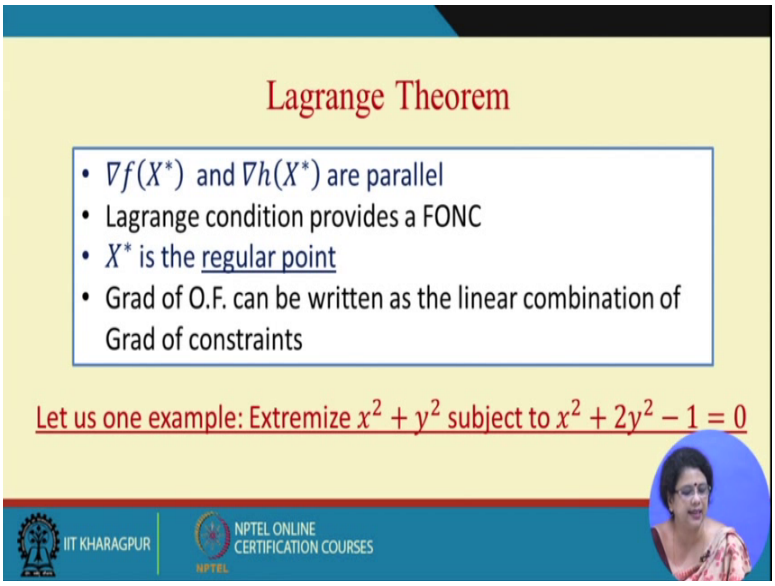
(Refer Slide Time: 04:44)



I will start discussing today only. Now let me revisit about the Lagrange theorem now in the Lagrange theorem, we know if we consider a 2 variable function which is objective as well as constraint. We are considering minimization of f X 1 X 2 subject to g X 1 X 2 is equal to b here the functions f and g only 1 constraint. I have considered and these are of non-linear type and b is a constant then we can make this one as minimization of f X 1 X 2 subject to h X 1 X 2 is equal to g minus b is equal to 0, there is know how to doing because just subtracting b from both side we can get it. That is why if we formulate a Lagrange function lambda l, in this way we can say that at the minimum point that should be the necessary condition that, grad of f plus lambda into grad of h at X star where X star is the minimizer of the optimization problem.

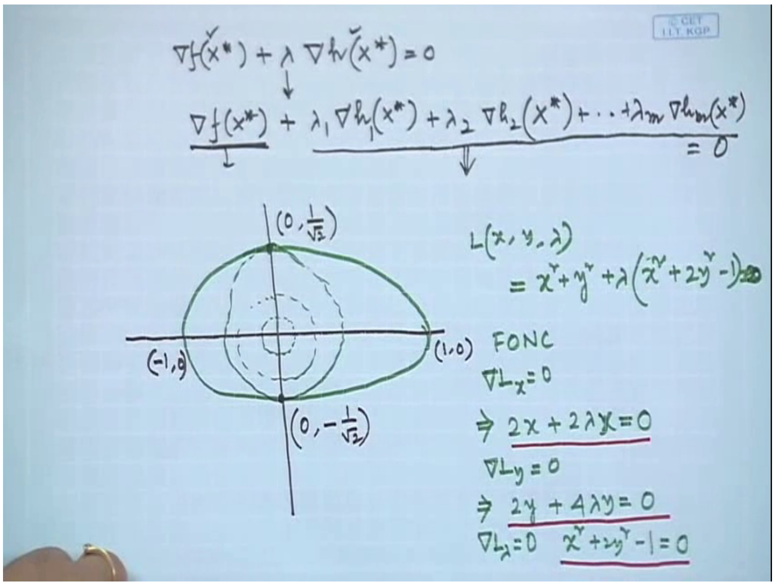
This is equal to 0 not only that we have to introduce, information about h as well that is why h X star must be is equal to 0 that was that has to be included.

(Refer Slide Time: 06:10)



Now there are few properties, what are the properties once we are declaring that grad of f X star is equal to just if I just write down the conditions once more then we can say that.

(Refer Slide Time: 06:28)



Grad of f X star must be plus lambda of grad of h X star is equal to 0. Now lambda is a constant can we not say from here that this is the vector this is the vector these 2 vectors are parallel, depending on the sign of lambda we can say whether this parallel vectors are directed in the same side or these are directed in the opposite side in the reverse direction. That part is clear now from the Lagrange condition, whatever condition we are getting that is the first order necessary condition, we do not know whether this condition is at all a sufficient condition or not.

Now, another point is that we are calling it X star as the minimizer of the non-linear programming problem, now this minimizer is being named as the regular point. Now what is the meaning of the regular point, we know the meaning of the feasible point feasible point is the point which is in the feasible space; that means, the point which satisfy all the constraint set, here also something related to that that the X star is regular point if it is in the feasible space, but it has certain other properties as well. I will discuss in the next. What is the regular point? now as I said that gradient of objective function is the parallel to gradient of constraint, but if we have more than 1 constraint then what we do, we write the grad of f X star e plus lambda 1 grad of h X star h 1 because if I consider more than 1 constraints together lambda 2 grad of h 2 X star up to, if there are n constraints lambda m grad of h m X star is equal to 0. That is why can we not say that, gradient of objective function can be written as, the linear combination of the gradient of constraints.

Therefore, we need to understand what is the meaning of it. What is the meaning graphically, once we have the individual single constraint we can see both are parallel, but if we have more number of constraints.

What is the meaning that the gradient of function is the linear combination of the gradient of constraints, all this meaning we need to understand one by one? For explaining those things, I am considering few examples one after another and there are few issues also that you know, whenever we are considering the non-linear programming problem even it is true for the linear programming problem. You must have been realized in linear programming problem that we are getting the feasible optimal point, but all the constraints are not active at the optimal stage, that is sure be few constraints are active few constraints are inactive.

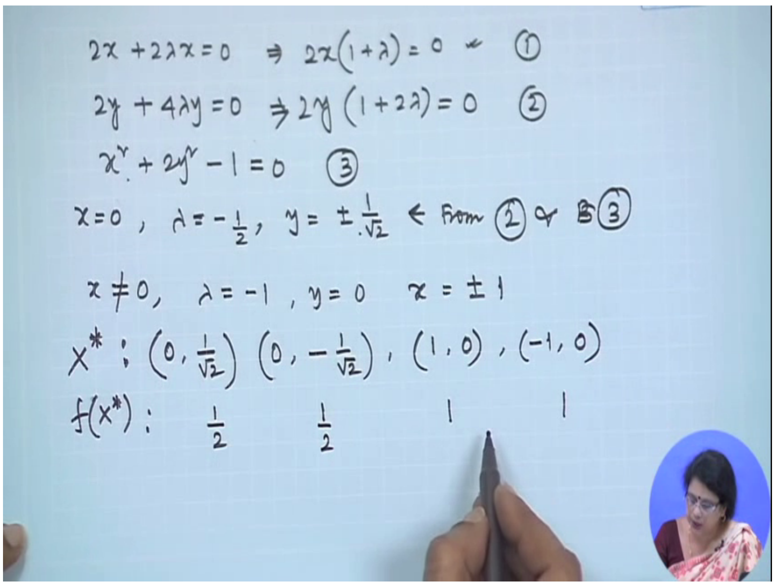
Here also we can just say that, we can get the optimal solution but is it possible for us so this Lagrange process, which constraints are active and which constraints are inactive and if the constraints are inactive then the question comes. Why should we include all those constraints in the set because unnecessary it is giving us more labor for computation, that is why all these issues to be addressed one by one. Let me start to explain to you all these issues one by one with few examples, let me consider the first example extremis X 1 plus y X 1 square plus y square X square plus y square subject to X square plus 2 y square minus 1 equal to 0. What you could see that objective function is a circle and constraint is an ellipse; let us try to see that what could be the solution for this problem. Now this is the X and y space, now if we just consider the ellipse drawing is not good enough, but this is the ellipse.

Now from the equation we can say that, this point is 1 0 this point is minus 1 0 and in the y since 2 y square is equal to 1 that is why it is 0 1 by root 2 and this is 0 minus 1 by root 2 this point right. Now this is the constraint set, only the constraint we can walk through the boundary of the ellipse because the constraint is of equality type. Feasible space is only the boundary of the ellipse, now if we consider that there is 1 circle X square plus y square that is why it is centered at 0 now circle is moving further all right. Now what should be the minimum value of the radius of the circle when it touches first the feasible space, certainly this is the point that is why that could be 1 optimal point and this is could be 1 optimal point.

We have to find out these 2 points are minimum point, that is why in doing all the calculations what we do we let us formulate we want to get all these things a the maximum points and the minimum points together by mathematically, that is why let us introduce the Lagrange function X y lambda this is equal to X square plus y square plus lambda into X square plus 2 y square minus 1 equal to 0 all right not 0 this is equal to. Now what is the first order necessary condition, the first order necessary condition should be grad of L with respect to X mast be is equal to 0. Which implies that 2 X plus 2 lambda y must be is equal to, I am sorry 2 X plus 2 lambda X must be is equal to 0 and this is a set this is another first order necessary condition with respect to y, partially differentiate if we do the L partially differentiation with respect to y then from here we are getting 2 y plus 4 lambda y equal to 0.

What else we are getting we are getting L lambda must be is equal to 0, it means that X square plus 2 y square minus 1 is equal to 0. Then what exactly we are getting we are getting 3 equations and 3 unknowns from here, if we just do the calculations let for this 1, let me write down once more the equations for you.

(Refer Slide Time: 15:02)



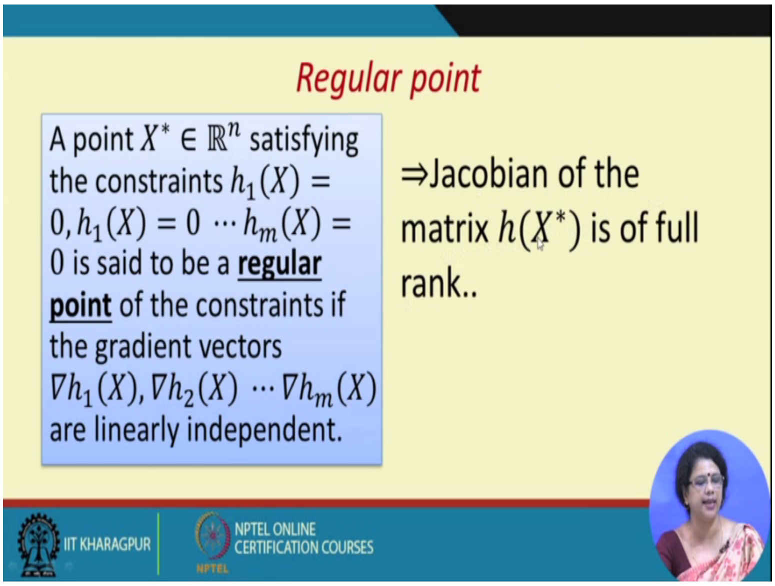
That is 2 X Plus 2 lambda X is equal to 0 implies 2 X 1 plus lambda is equal to 0 the next, 2 y plus four lambda y is equal to 0 this implies y 2 1 plus 2 lambda is equal to 0, what else we have X square plus 2 y square minus 1 is equal to 0. Look at the first, this one either X can be 0 or X cannot be 0 if X is equal to 0 then what we get if X equal to 0 then we get lambda is equal to minus 1 and from here what we get y is equal to plus minus 1 by root 2 from equation 3 this is one set all right. But if X not is equal to 0 then what you will get this f X equal to 0 then lambda cannot be is equal to minus 1. Then what should be the value for lambda then, if X equal to 0 from here we are getting the value for y is equal to plus minus 1 by root 2.

And from here we are getting the value for lambda s minus half, lambda is equal to minus half. This we are getting if we just considering 1 2 and 3 then from 2 and 3 sorry from 2 and 3 we get this condition and if X not is equal to 0 then 1 plus lambda must be is equal to 0, that is why we are getting lambda is equal to minus 1 and from here if lambda is equal to minus 1 then we are getting that y is equal to 0 here, if y is equal to 0 then X would be is equal to plus minus 1 that is the condition.

Now, from here how many points we are getting 4 points, what are the points these are the possible extreme points. The points are 0 1 by root 2 because from here 0 minus 1 by root 2, that is 1 0 because X plus minus 1 and minus 1 0 we are getting 4. Now what are the corresponding objective functions here if we just calculate then the objective function would be here it would be half at this point this is again half, this is again 1 and 1 all right. Then by looking at the values you could see that 1 0 and minus 1 0 we are getting the maximum values and half and here we are getting the minimum values if we just go back to the figure. We could see that 0 1 by root 2 and 0 1 by root 2 we are getting the minimum value and if the circle moves further and further when is it touches this 2 extreme points then we are getting the maximum value all right. This is the thing we are getting from here that is why we could see that there is only 1 constraint and that constraint is active constraint.

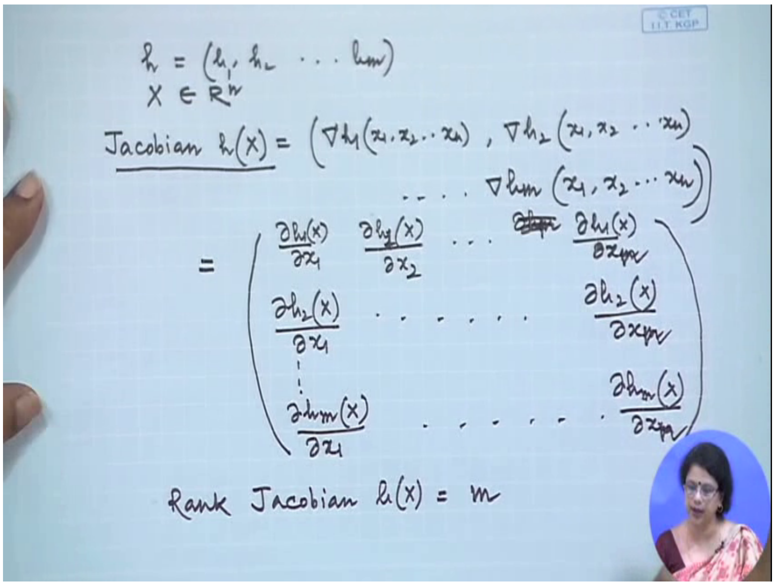
Now let me I will show you more examples, where I can show you that there are certain constraints, which are not at all active that is why up to this.

(Refer Slide Time: 19:11)



We have come crossed few new concept, one thing is that the definition of the regular point and we have to define it. What is the definition if we are having m number of constraints, a point which satisfying all m constraints that is the regular point of the constraint set, not only that at the regular point all the gradient of the constraint must be linearly independent all right. This is the definition not only it is in the feasible space for the non-linear programming problem know there, but also there is another condition gradient at the regular point this is these are all independent, in other way we can say that there is another property that the Jacobean matrix of h X star is of full rank, what is the meaning of that? I will tell you what is the Jacobean of h X star.

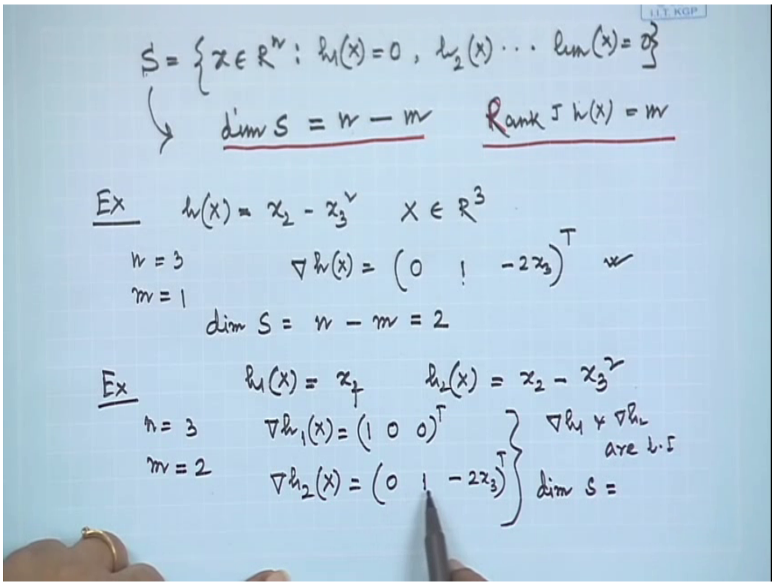
(Refer Slide Time: 20:10)



Now, we know h is equal to h 1, h 2 there are m number of constraints, add this is a function X h 1 h 2 h n these are all function X which belongs to rn all right. That is why m number of constraints n number of variables, now if we consider the Jacobean of h X this is equal to grad of h 1 X 1 X 2 X n grad of h 2 this is a vector, X 1 X 2 X n up to grad of h m X 1 X 2 X n this is the Jacobean, but why did we say that this is the matrix. That is a matrix because grad of h 1 when we have X 1 X 2 X n then, this is equal to we can write down the whole matrix as h 1 del h 1 X del X 1 del h 1 X del X 2 del h m this is 2 no sorry 1 and the last term is del h 1 divided by del X m. The second row del h 2 X del X 1 this way we will go del h 2 X m, you must be understanding that we are getting a matrix and what is the number of 4 rows and number of columns, you could get number of rows would be m and number of columns would be n.

Now it has been said that this must be of full rank; that means, always we will get, we should find out rank of Jacobean of h X must be is equal to m. This is the condition Jacobean of h X is equal to m X divided by del X m, this is X m X m these are all n because there are n number of variables these are not m all right. Now we need to find out that rank of Jacobean of h X is equal to m not only that we need to find out if we consider then what is the regular point.

(Refer Slide Time: 23:19)



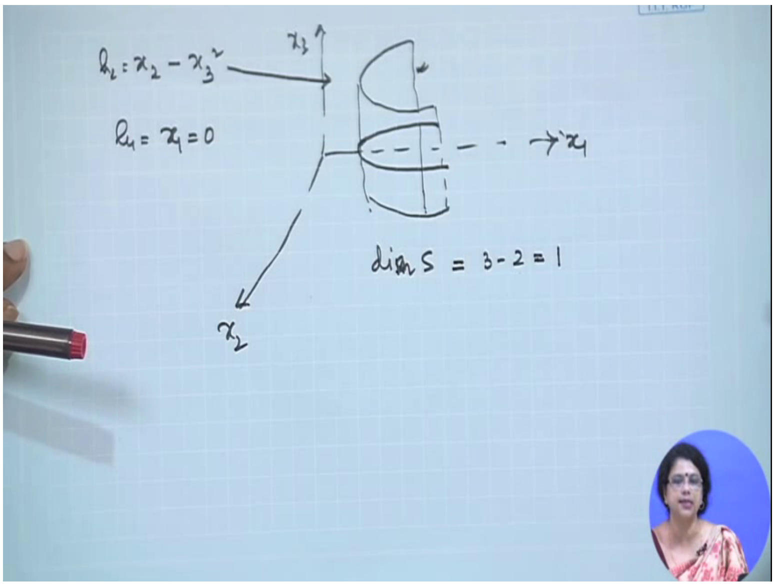
Then regular points are the points in r n such that, it satisfies the constraint set and we need to find out. That if this is the space we need to find out, if s is a space for it because we will get different X here, combination of X will form a space we need to find out that dimension of s is equal to that is a number of variable minus number of constraints.

That is why we need to prove 2 things here 1 thing is that dimension of s is equal to n by m and another thing is that rank of Jacobean of h X must be is equal to m. Let us consider few examples for it, both the examples we will consider one by one. The first example is that, we are getting h X is equal to there is only 1 constraint that h X is equal to X 2 minus X 3 square, it has been said that X is in r 3 only all right.

When we are dealing with the regular point we are only dealing with the constraints the constraint is not having X 1 that is true, but objective function must be having that and here we are considering here n is equal to 3 m is equal to 1, we have to find out dimension of s must be is equal to n minus m that is equal to 1 all right now if this is. So, if we just find out the Jacobean of h x, since only 1 that is why our we would not get we will get column vector with respect to X 1 it is 0, with respect to X 2 it is 1, with respect to X 3 it is minus 2 X 3 this we are getting grad of h must be is equal grad of h is this 1. What you are getting from here, we are getting that what is the dimension of this thing would be 2. That is why we can say that is the collection of all x, which satisfy the constraint. It has only 2 dim dimension 2 we can say the dimension of s is equal to n minus m is equal to 2.

Note the next example, we have considered here we are considering n is equal to 3, but m is equal to 2. we have 2 constraints together all right 1 constraint is h 1 X is equal to X 1 and h 2 X we have is equal to this is X 1. This is equal to X 2 minus X 3 square then from here we can get that grad of h 1 X is equal to 1 0 , grad of h 2 is equal to what is that 0 1 minus 2 X 3 all right. That is why from here what we can first of all we have to see that this is the regular point if we can see that grad of h 1 and grad of h 2 these are linearly independent, then only this is the regular point. That we can prove very easily from here that both are linearly independent 1 cannot be expressed in terms of the other 1, you can very easily you can find it out and not only that from here we can see dimension of s is equal to dimension of s is equal to that dimension. W is the dimension for this thing let me just explain to you.

(Refer Slide Time: 27:56)



Now, say we are having 3 coordinate axis no we 1 function is X 2 minus X 3 square. How it looks like this looks like this. Because this is a surface for us all right. Now this is the surface there is another constraint for you X 1 is equal to 0, we are having the this is X axis X 1 axis, this is X 2 axis and this is X 3 axis, 3 dimension this is your h 2 this is your h 1. If we just take h 1 X equal to 0, what exactly you are getting only one nothing else.

You are getting in 1 dimension, X 1 is just cutting here cross section because this is a plane X 1 equal to 0. Is a plane the cross section of the plane and this one will be only one dimension, that is why graphically also we can declare that dimension of s that the collection of regular points would satisfies both the in equations. In the dimension must be 3 minus 2 is equal to 1 all right, now this is all about the regular point now we will move to the inequality constraints in the next class.

Thank you very much.