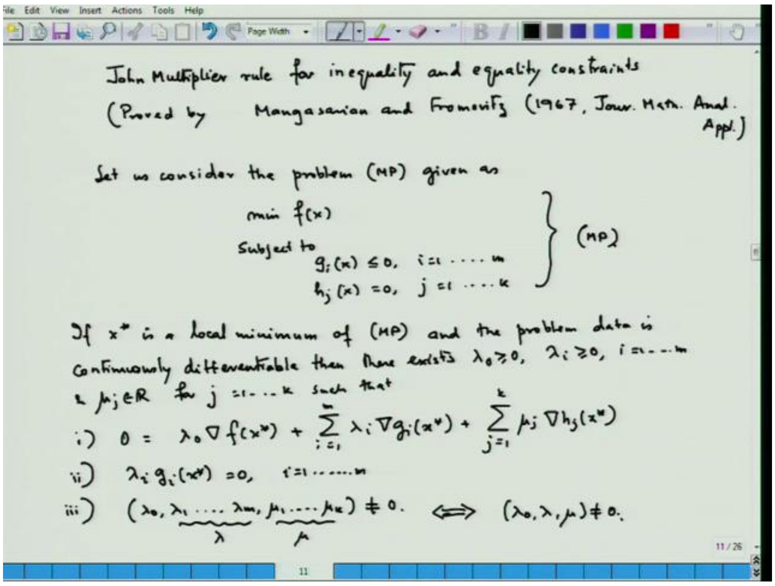
**Foundation of Optimization**
**Prof. Dr. Joydeep Dutta**
**Department of Mathematics and Statistics**
**Indian Institute of Technology, Kanpur**

**Lecture - 19**

Now in the last class, we had shown some example of the Lagrange multiplier rule and you see that we have been able to prove that lambda naught is equal 1. We would like to continue with the Lagrange multiplier rule, but I just want to with examples, but I want to recollect.

(Refer Slide Time: 00:35)

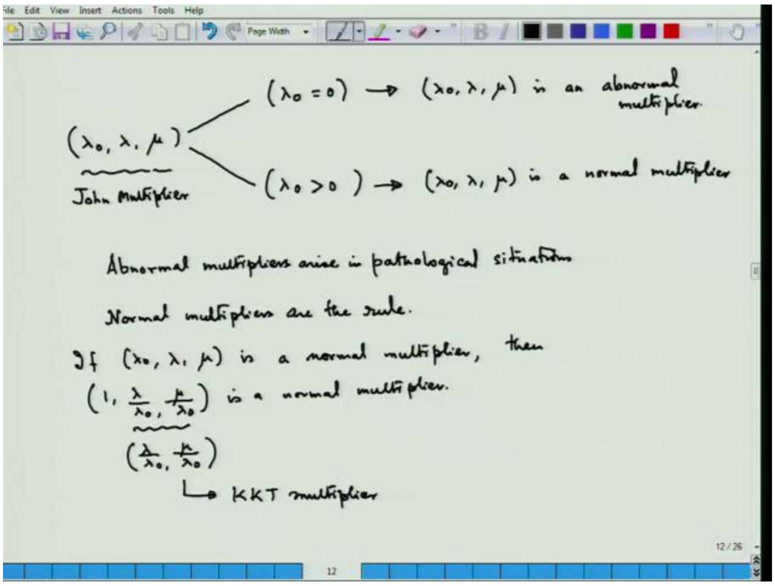



That the John multiplier rules for both equality and an inequality constraint was first given by Mangasarian and Fromovitz. Now, I will have inequality and equality constraints. So, this was proved by (( )) Mangasarian, Fromovitz; and Fromovitz in a paper in 1967 possibly one of the most fundamental papers of the subject; a classic, published in 1967 in the journal of math analysis and applications. And what he proved was the following. So, let us consider the problem MP, the one which we had earlier done, but with equality constraints. So, we are considering the problem MP, now we will have equality constraints also.

If x star, this is my MP. If x star is a local minimum of MP and I should better put it in a standard form is a local minimum. And, all these are continuously differentiable and the problem data is continuously differentiable that is f, g i and h j all are continuously differentiable; g i for all i, h j i for all j and the problem data is continuously differentiable. This was a big advancement for optimization, this idea of combining everything and shows some sort of unity in the Lagrange multiplier principle continuously differentiable; then, there exists. Look here I do not have just differentiability, but continuous differentiability. Essentially, to handle the equality constraints here you have to bring in the idea of implicit function theorem where you need continuous differentiability and once you want to combine everything you will need a continuous differentiability on all of them.

So, then there exists lambda naught greater than equal to 0; lambda i greater than equal to 0; where i equal to 1 to m. And lambda mu j element of R; for j is equal to 1 to k, such that, 0 lambda i. Number ii as the complimentary slackness condition associated with the inequality constraints. Number iii is the most crucial condition. So, if I consider this I write as a vector lambda, if I write as a vector mu; then, this thing can be equivalently written as lambda naught, lambda mu not equal to 0.

(Refer Slide Time: 06:47)




Now, we can classify these multipliers, which are we are going to call the john multipliers. So, this is what we will refer to as the john multiplier; you can also refer it as the Lagrange multiplier. So, there are two cases which can arise which is important to us at least; lambda naught is equal to 0 and lambda naught greater than 0 or without loss of generality lambda naught is equal to 1. Let us just take lambda naught greater than 0 times. So, in this case we say that lambda naught, lambda mu is an abnormal multiplier. In this case we say that lambda naught, lambda mu is a normal multiplier. In his very recent book one of the greatest optimization theorist of our times, Francis Clark has mentioned that if you look at the John multiplier this lambda naught is usually greater than 0; most of the time. At least from the problems current setting lambda naught greater than equal to 0 comes out automatically.

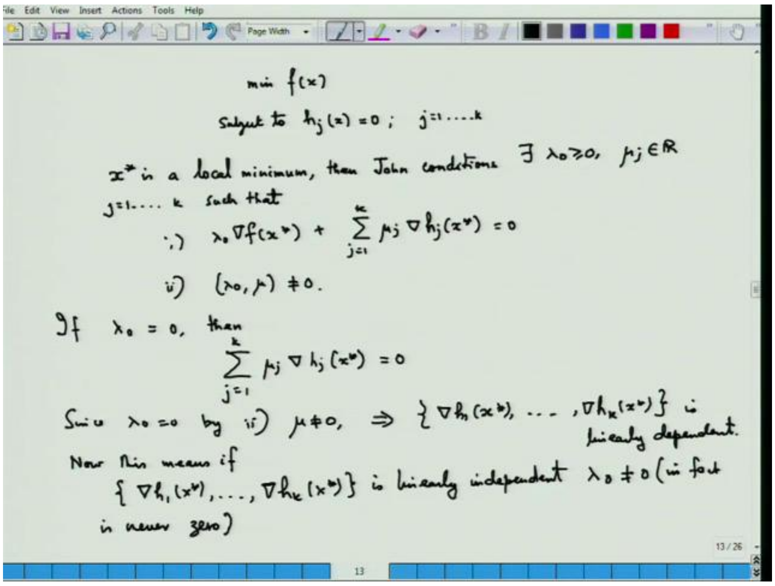
Now, what is more important to understand at this stage is that except in very pathological situations, we will always get lambda naught equal to 0. So, abnormal multipliers usually arise in certain pathological situation and one of the situation is; for example, where you have only one element in the feasible set, that is a pathological situation. So, this comes only in pathological abnormal multipliers arise in pathological situation.

Normal multipliers are the rule, abnormal multipliers are the exception. Now, it is very very important to realize this following fact is that normal multipliers are sometimes also referred to as the Karush Kuhn tucker multipliers. So, if lambda naught, lambda Mu is a normal multiplier, then is a normal multiplier. Usually in this particular case, this setting; that is lambda by lambda naught and Mu by lambda naught; this vector is also called the KKT multiplier, which is linked with the celebrated KKT conditions. But as we have said to go the KKT conditions, you need to impose certain things on the constraints. But I want to reassert you that it is the fritz john view point which is basically the true view point in the, proper view point in optimality conditions; because it gives you a lot of information. The very important information that it gives you whether you bother about the extra conditions on the constraints or not; the normal multipliers are the rule in the game.

This is something extremely fundamental and has to be kept in mind. Straight jump to the idea of additional conditions on the constraints and to the Karush Kuhn tucker conditions, might take away from you a much more richer view point of optimality conditions. So, here we will concentrate largely on the fritz john view point and we will show that even looking at the Fritz John conditions, we can get a condition which will ensure that not a single multiplier is abnormal.

Now, you tell me the reason for having abnormal multipliers? The reason that we do not want abnormal multipliers is that because abnormal multipliers take off the role of the objective function in the optimality conditions further abnormal multipliers can arise when a point is not really optimal, but can just satisfies the john conditions. So, there are certain bad issues with the abnormal multipliers but let me tell you abnormal multipliers are reality. They can even arise when you have local optimizer of a problem even at that sort of point a normal multiplier an abnormal multiplier can arise. For example, I will show you a simple situation, how a normal multiplier can and how an abnormal multiplier can arise.

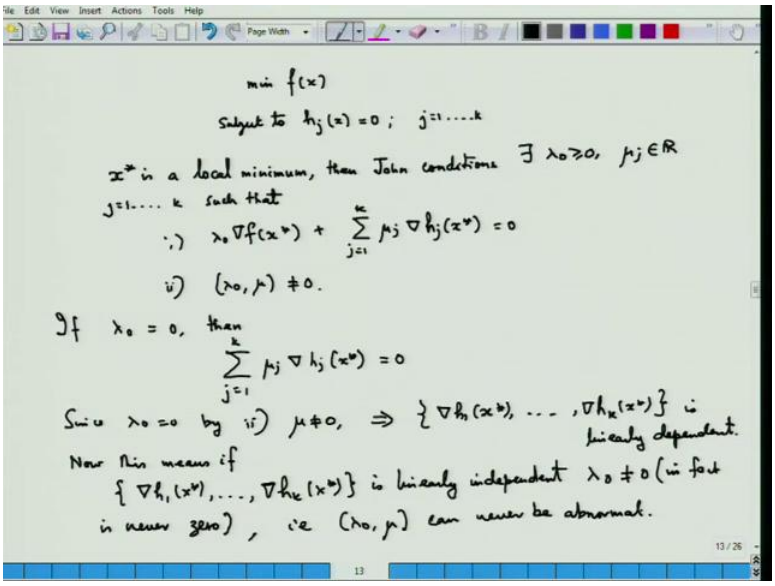
(Refer Slide Time: 13:07)




Now, look at this problem, very simple problem minimize f (x). So, we have inequality constraints. Now, by say x star is a local minima. Then by the fritz john condition, by the john conditions there exists lambda naught and mu j element of R, j moving from 1 to k, such that so, j is equal to 1 to a; (( )) that is, what is a condition. Now, of course, we are assuming continuous differentiability. Now, what is important to know, what happens when lambda naught is equal to 0? Then, since, lambda naught is equal to 0, by condition number ii Mu is not equal to 0; implying that the set h (k) is linearly dependent.

Now, this means if summation, if this set is linearly independent this cannot be greater than 0. This means if sorry I am writing m it should be k is linear independent lambda naught is not equal to 0 in fact is never 0. Now, suppose I have a situation where I have all these vectors to be linearly independent; then, I know that lambda naught is not equal to 0. Now, I repose this problem in slightly different way.

(Refer Slide Time: 17:17)

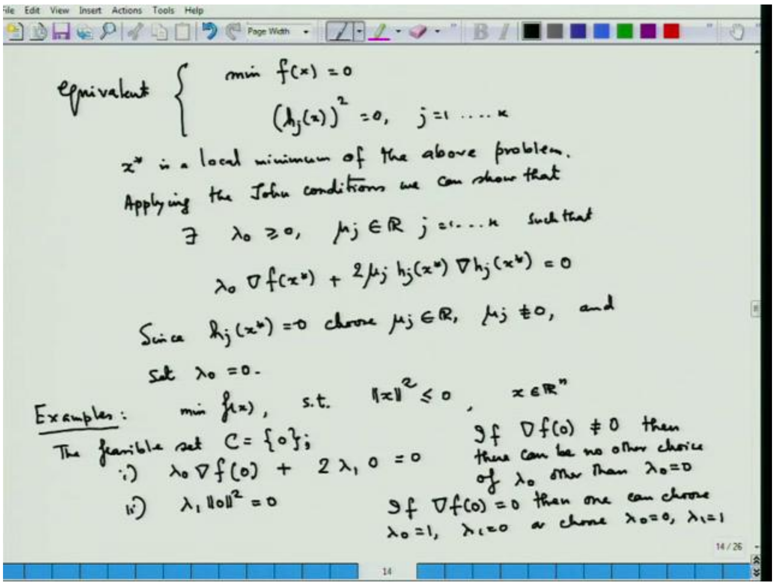


I will pose in equivalent the following equivalent way minimize. So, any x which is satisfying this square of anything is 0; so the object must itself be 0. If the object is not 0, how can the square be 0? So, this real number, the square of this real number is 0; this must itself be 0. So, these and these, so this problem and this problem are equivalent problems. Now, let me write down the Fritz John condition. So, x star is a local minimum of the above problem; this problem is equivalent to the previous problem.

So, applying the john conditions again; so applying the john conditions again, what we have is the following. We have that there exists lambda naught greater than 0 and Mu j element of R; such that, lambda naught. Now, you see since, h j x star is equal to 0; choose mu j element of R and mu j not equal to 0 and set lambda naught equal to 0. So what you have shown that if I pose the same problem which is nice and never have an abnormal multiplier, the same problem if it is posed in a different way; it will give us at least one abnormal multiplier. I think which we do not want.

So, this point of view of viewing the optimality conditions to abnormal and normal multipliers, leads to a better view than just jumping to certain conditions which will guarantee the Karush Kuhn Tucker conditions. So, if you have (( )) in this problem, the previous problem, it tells you that the Lagrange multiplier always is valid, the john conditions is always valid with normal multipliers. It can never have abnormal multipliers. That is the multiplier can never be abnormal; that is, lambda naught Mu can never be abnormal. We will speak more about this question of abnormality and we will give some examples now.

(Refer Slide Time: 21:06)



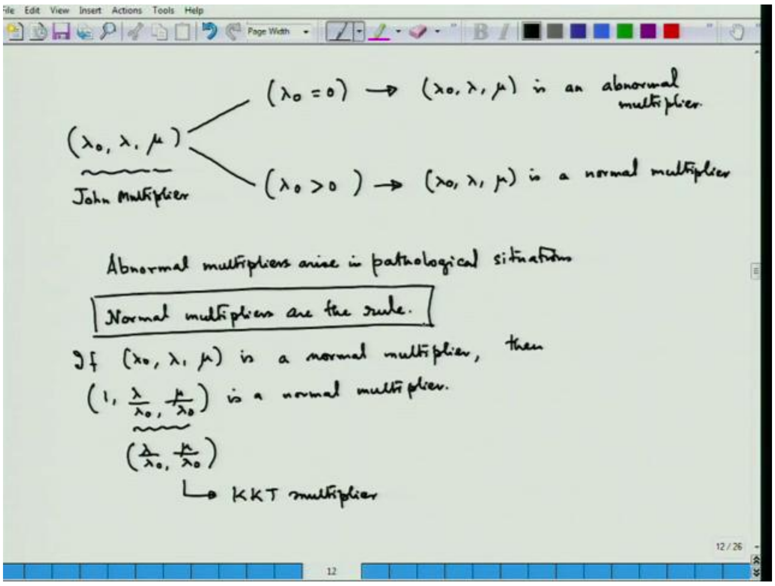
So, we will now show by an example, where the objective function; the structure of the objective function and the pathology of the pathological situation that we will get by the feasible set, just having one element. We will show that an abnormal multiplier can exist and co-exist with a normal multiplier. So, if you take the set of all fritz john multipliers which is a cone without 0; that cone can contain normal multipliers as well as abnormal multipliers.

So, what we really need is to show that there is at least one normal multiplier. In some cases there will be no other choice other than abnormal multipliers. For example, if you want to look at this function minimize f (x), where f is say differentiable continuously differentiable; such that now you see that the only feasible set is the origin. So, the minimum is of course achieved at 0. So, this means what? This means the following. Now, you can write, I have no particular a on x on f x. So, x will always put 0, which is anyway 0; it is obvious.

Now, look at this condition. If grad f naught is not equal to 0. See this is anyway 0; if this is not equal to 0, not a 0 factor; you cannot put lambda naught to be greater than 0. Then, there is no other choice, there can be no other choice of lambda naught lambda naught other than lambda naught is equal to 0. There cannot be any other lambda naught greater than equal to 0 other than 0 which will satisfy this; if this is true. Now, if grad f (0) equal to 0; which can be quite a frequent case.

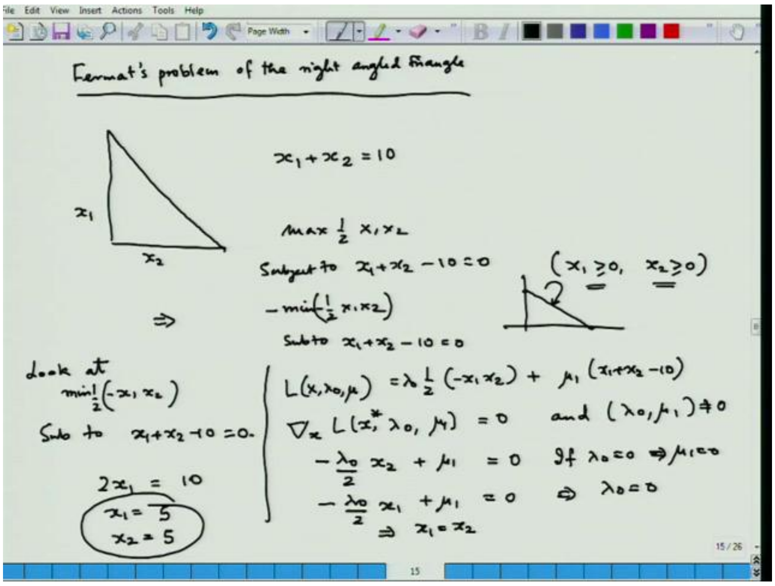
Then, if grad f (0) is equal to 0, then of course, you can have lambda naught; this is a very pathological situation, grad f 0 equal to 0. Then, here there is no other they all the multipliers are norm abnormal if you have lambda grad f not equal to 0. And grad f not is equal to 0, then one can choose lambda naught to be equal to 1 and lambda 1 is equal to 0; or choose lambda naught equal to 0 and lambda 1equal to 1. That is your both are normal multiplier and abnormal multiplier coexisting side by side for the, this kind of problem. So, only in pathological, very bad situations or these sort of wrong posing and a bad ill posing of the problems; that this becomes 0.

(Refer Slide Time: 25:34)




In general, as we see this is the rule, this is something extremely fundamental and you should remember this; normal multipliers are the rule. We will again give more example to substantiate such a claim that we have made. Of course, you cannot prove it but you can show that most examples such a thing would actually work. So, let us give some more examples from Brinkhuis and Tikhomirov book which had not only showed and recommended a book; which I want most of the readers, who are sincerely interested in knowing about optimization theory should read this book.

(Refer Slide Time: 26:03)




So, here we will describe the Fermat's problem, geometrical Fermats problem. So, the Fermats problem of the right angled triangle is like this. So, you have right angled triangle one side is x 1, one side is x 2. I expect the sum to be 10 something; so, there could be many such combinations of x 1 and x 2, which will give me x1 plus x2 is 10. Now, find that combination which will give me a triangle with maximum area; that is problem is max of find max of half x 1, x 2 subject to. Now, you can pose this problem equivalently as minus mean of minus half, minus 10 equal to 0. So, it is enough for us to just look at.

So, the optimality conditions would be how. Now, you see how are we sure that it would have a solution. Of course, we are expecting x 1 and x 2 to be strictly greater than 0; that we at least know that there is a lower bound, right. That is something what we need to have, need to think about; you have to first decide how do you know that there is a solution to this. Now, if I restrict x 1 and x 2 to be greater than equal to 0; then if x 1 and x 2 these are restricted to be greater than equal to 0. If you take greater than equal to then, you are sure that there is a solution because that set would become a compact set; x1 plus x2 10, basically it will become something like this. Now, how do I on that this compact set this will have a solution, but the solution could be in the boundary, could be in the interior. But essentially, I want something in the interior because in the boundary one of them would be 0. Because the solution in the boundary it does not make sense then it will not be a triangle.

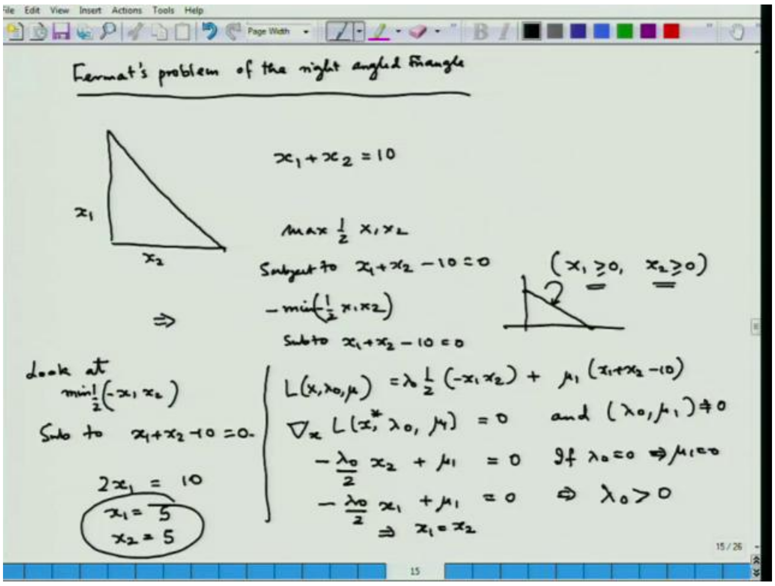
So, the solution must lie, if it lies even if I take like this; if the solution of this problem with this additional restriction if it lies it must lie in the interior. There will be a solution of minimizing this over this set, which is this. But if a solution actually lies, if I can minimize or maximize whatever and then the solution would have to lie in the interior; it cannot lie in any of the boundary points. Because then if one of them is 0, then the triangle it does not mean the triangle has no meaning. So, my solution would be meaningful if x1 is strictly greater than 0 and x 2 is strictly greater than 0; solution would there will be a solution. So, then I apply the Lagrange multiplier rule on this; my Lagrangian L x lambda is half of lambda naught; sorry lambda naught lambda, it is a lambda naught lambda. You can take lambda or Mu, the equality constraints you should not just keep On taking mu, because we are getting habituated to that. There would be something like this, you really have to figure out this, check out. So that would give you first we take with this, lambda naught minus lambda naught by 2 x2 Mu naught.

Not mu naught sorry mu, maybe mu 1 is better because I think I am making symbolical mistakes. And you know lambda naught cannot be 0. If I put lambda naught equal to 0; from this equations I have mu 1 equal to 0; so, lambda 1, mu 1 both cannot be 0. Simultaneously; so because there would be some of such a local minimum, which is say x star if you want. So, if there is a local minimum x star, then corresponding into by fritz john conditions there is, there are multipliers like this which will give me this equal to 0. Now, see how it how it always comes out form the problem condition; that lambda naught cannot be 0. So, this is not a pathological problem with one element in the feasible set; this is a very natural problem. So, the normal multiplier is the rule. So, if lambda naught equals to 0 it implies that mu 1, mu 1 equal to 0 which cannot be true. Because here we should always have this condition from the john conditions; so, this implies that lambda naught is not equal to 0.

So, how do you figure out? Basically, what you will get from here because you can put that everything in both of them are equal to minus mu and so these two are equal. So, what you will get is x1 is equal to x 2. So, if x 1 is equal to x 2, you will know that 2 x 1 is equal to 10 or x 1 is equal to 5 and x 2 is equal to 5. And since, there is only unique solution; so, this is the minimizer of this. So, the maximum value is half into 5 into 5; so, that is 25 by 2.

So, this I would like to end today’s lecture and tomorrow we will get on with more of such problems. It is very important to practice such problems to get in your mind the feeling that contrary to the popular view, which takes the KKT point of view, to the study of optimality conditions. It is the john multiplier rule, which is essentially fundamental to the study of modern optimality conditions or the Lagrange type multiplier rule like this sort of multiplier; this is the john multiplier rule. So, it is the john multiplier rule which gives you more information about the problem itself than a KKT would give you. Of course, you will here, get a KKT condition; naturally, I can put lambda naught equal to 1. But the problem is it is so beautifully comes out.

(Refer Slide Time: 34:56)




Sorry, sorry here I have written lambda naught equal to 0, no that is a mistake. I wanted to I had said lambda naught strictly greater than 0. Because if lambda naught equal to 0, this will be equal to 0, this will be validate. So, which means that you see from this is a nice problem which says that from the very basic structure of the problem, lambda naught becomes strictly greater than 0. So, normal multipliers are once again the rule and we are going to really show you by few more examples then the normal multipliers are rule. So, if we just know fritz john condition you can tell a lot of things about optimization then by getting down by, bogged down by the fact that you have to impose certain conditions and get constraint qualifications.

Now of course, in the hind sight you can say, I can now put here 5 5. Now, when you do not know the solution; see for equality constraint, one of the constraint qualifications or one of the conditions which you can impose to guarantee the this equal to 0 is this. But x star has to be known for you to check the constraint qualification. Here you see, you now know that if I put 5, 5; the gradient it will become 1, 1 and it will not be 0. So, that is linear independence. So, this is only known once you know the solution, but while computing the solution you have already figured out that lambda naught cannot be 0. If there is a solution exists then lambda naught any way would not be 0. So, the john conditions, the problem itself is guaranteeing you lambda naught is strictly greater than 0; so, normal multiplier is a rule. So, immediate rush to see whether there is a constraint qualification and whether KKT conditions are satisfied, will in many cases misguide you in understanding the problem itself.

Thank you very much.