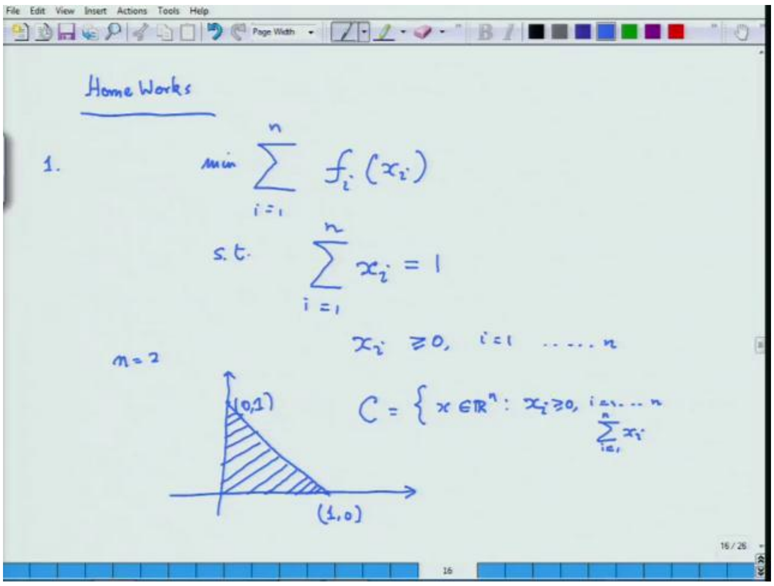
Foundation of Optimization
Prof. Dr. Joydeep Dutta
Department of Mathematics and statistics
Indian Institute of Technology, Kanpur

Lecture - 22

As we had said in the last class, in today's program, we are first going to discuss about some not really discuss, I am going to give you a homework. So, I am going to give you some problems for you to solve using the Karush-Kuhn-Tucker conditions or the Fritz John conditions that we have learnt here. I want to remind that, we would call a Fritz John conditions as a Karush-Kuhn-Tucker condition or Kuhn-Tucker conditions; if all the Fritz John multipliers associated with it it is normal.

Now, such a thing would happen, if the Mangarsarian from which constraint qualification holds, that is what we have seen in the last class, and also we have spoken about the fact that every linear program. We have proved that every linear, for every linear programming problem all the multipliers are normal multipliers, all the John multipliers are normal. Now, we are going to give some homework, especially I have picked them up from the book foundations of optimization by Osman Guler.

(Refer Slide Time: 01:21)



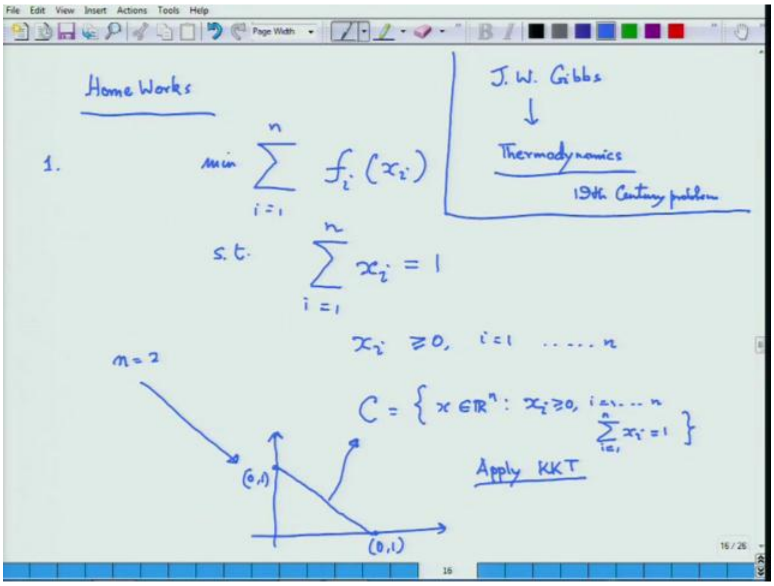
Let me list down the home works and after that we would start our study of Quasi Newton methods, which I have said mimics or, sorry which I have said needs to have, where one needs to have an understanding of constraint optimization. In order to really understand the convergence analysis of the Quasi Newton method, this is very important in algorithms, even if we do not run them and do not try to put them on the computer and check how they are behaving, which I personally do not do myself, because I have a theoretician.

So, it is important to have an understanding at least mathematically, how the sequence of iterates that we would generate through an algorithm behaves and many computer scientists would agree that, specially algorithm specialists. Algorithm specialists that they might not have really sat down on the machine, but somebody else have done the programming and trying to see how the things behave.

But, to mathematically predict, what would happen to a sequence of points that are generated by some algorithms that process is called the convergence analysis of that algorithm and those things are very important and very useful. When you are really talking about an algorithm that we had learnt such things, when we were talk, speaking about conjugate gradient methods and all those things the similar sort of approach, would be taken in the case of the Quasi Newton method. So, let us look at our first homework which says, so x is in r n each x i is in r such that, so this is called separable programming.

The full function can be written as a sum of some other functions, obviously I have here. So, if you have a function of two variables only that is n equal to 2, then the feasible set is easy to draw. So, this is 1 and 0 and this is 0 and 1 and we draw it here now, the feasible set would look like this. The interesting fact now is how do we now start solving this problem? Possibly, we would assume that, there is a solution, how will you first assume that? There is a solution. Now, the feasible set here, the set C is a sort of a all x in r n, for which all the components are non negative that it is, in r n plus and then, sorry I made a mistake.

(Refer Slide Time: 04:58)



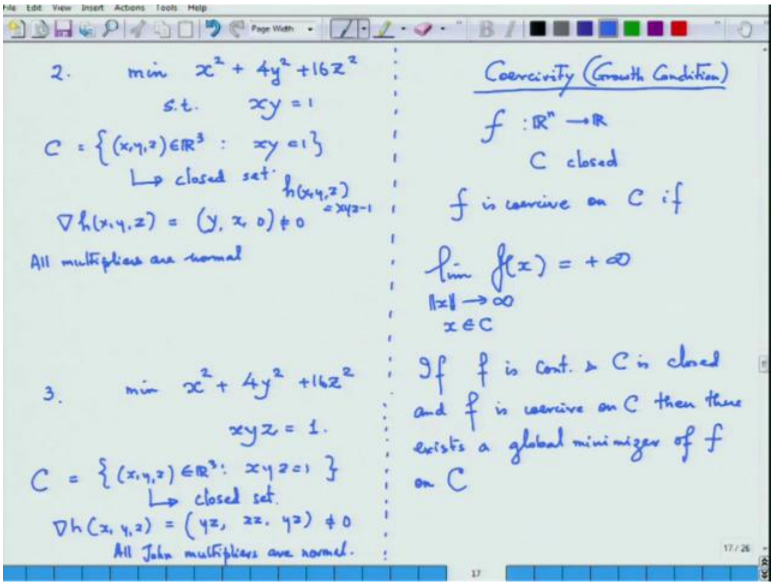
I think if that was less than equal to 1, so this is the sorry I did it. So, this is my feasible set C, this line passing through 1, 0 and 0, 1 and remaining only in the non negative part. When n is 2, this is your C, so first thing to observe at the that c is a convex compact set in r 2. So, this if we assume all the f i's are differentiable with, then you are assuming all of these are continuous and hence you are telling that, this is my feasible set, which is compact these are continuous functions, because they are differential and hence there will be a solution. Now, where will the solution lie is a different issue, this or it depends on the f i's. Suppose, I do not know the f I, I will try to at least have an qualitative feeling about the solution by applying the Karush-Kuhn-Tucker conditions.

So, apply KKT, now here one important piece of our learning would come into play this constraint and all these constraints here linear. So, the Karush-Kuhn-Tucker condition automatically hold means, all the John multipliers associated with this problem would always remain to be non would remain to be normal. Now, it is very important to have a little bit of look at the history of this problem, this is not a twentieth century or post world war, two problem.

It is a problem, which was known to the American physicist JW Gibbs, one of the greatest contributors in the field of thermodynamics, in the nineteen century. So, this problem that we see here was related thermodynamics and it came up in the study carried out by JW

Gibbs in the nineteen century. So, here you have a nineteen century problem to be solved by twentieth century method, so I leave this problem to you to have a look at.

(Refer Slide Time: 07:44)



Now, I will give you two problems, two different problems question number two and question number three, they will be slightly different and it is again important for you to figure out the solution. Now, it is more easier because here, but you have to guarantee that, what you get as solution? Because we have not yet spoken about second order conditions, so we cannot say that we applied the second order conditions, just by applying the Karush-Kuhn-Tucker conditions and getting some information can you from the problem structure say, what is your solution? That will be the question.

So, this is a problem in three dimensions right and subject to. Now, change the problem slightly to here, you have to proceed by applying the, we may proceed. Now, by applying, what we have learnt? Now, the question is, do we really have to look at the Karush-Kuhn-Tucker? Fritz John condition first and then look at the Karush-Kuhn-Tucker conditions, that is the question. So, when you are trying to solve this question, this is the first question should arise in your mind.

How is this problem? Right in the sense, that basically it is telling that x y is equal to 1. So, what does it mean? Does it mean that, I am able to solve this problem yes or no? Now, how do I know that there would be a global minimizer of this problem? How do I know that there is a global minimizer of this problem? Once these two are known, then I can start taking some steps to solve them here. We need the idea of coercivity, I have not spoken much about it or maybe I have not spoken at all. So, to know that these problems have a solution needs coercivity, because these things do not represent compact sets in r 3, they both represent close sets in r 3, but they do not represent compact sets.

So, here would come the notion of coercivity or it is some or it is also called the first order growth condition. So, it is how the function grows as the norm of the decision vector x becomes larger and larger decision vector x, y, z, here in this case. So, x vector x 1, x 2, x n becomes larger and larger now. So, if you are considering a function R n to R and you are considering a close set C f is coercive on C, if f of x the limit of this becomes plus infinity.

Whenever, norm of x goes to infinity with x remaining inside C such a thing would not happen, if a set s is bounded norm x is bounded. If the set C is bounded, you cannot have norm x going to plus infinity of course, with x remaining in C. So, these things become helpful, when the set C is actually not compact not bounded at least, so these facts then become helpful, so if this is there. So, if f is continuous and C is closed and f is coercive on C then there exists global minimizer of f on C.

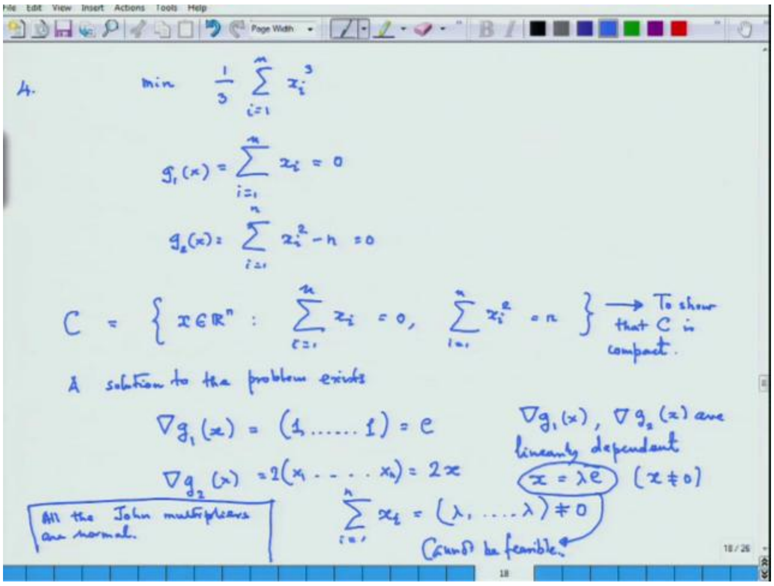
Now, I would give you an exercise; first at this set that the set C, x, y, z belonging to R 3 to 1 and this set C here. Now, your first step would be to show this is a closed set. So, I am just giving you some hints, so that you are enthused to take up your pen and paper and solve these problems. So, once you have that these two, then I would fairly be happy that, I have got the first step, because it is very important to see that this function as x y and z the norm of these. As these quantities go towards infinity this whole function, would go towards infinity. So, the coercivity condition is taken over. So, I am guaranteed that these have a global minimizer, now how to decide whether these problems are having normal or not normal right?

Now, you see one of the they, you have just inequality constraints one of the conditions, which will guarantee you always that there is a normal John multiplier is the linear independence of this inequality constant, x y minus 1 equal to 0, but this linear independence of the gradient of the inequality at the solution. Now, you have you can observe that 0, 0 cannot be an element of this set because 0, 0 is not equal to 1 0 is not equal to 1. So, both x and y here cannot be 0 z could be 0 does not matter. So, here the gradient of say this is your h. So, let me put h in this case as x y z minus 1. So, gradient of x y and z is actually y x and 0 and x and y both cannot be 0. So, this is a non 0 quantity none of them are 0. So, which once this is non 0 quantity, which means this.

Set is a linearly independent set trivially and then hence that linear independence of the gradient and the solution always guarantee the normal multiplier. So, all multipliers are normal, here also x, y, z, can all none of them can be 0, if one of them is 0, then this cannot be an element, here right the solution cannot. So, I need this only at the solution, so at the solution point means x and y none of them are 0 here. So, you have the, if x y is a x bar, y bar, z bar is a solution than x bar, y bar, z bar none of them are equal to 0 in this case. So, you have that grad of h, if I take this as my h. Now, x y z minus 1 equal to 0, then I will have here as and this is not a 0 vector.

So, here also all john multipliers are normal. So, we have three problems, we have some information about the behavior of the function. So, you start applying the Karush-Kuhn-Tucker conditions and try to see what you can get from them. We go to another interesting problem from Guler, which is pretty interesting by it is solution by it is whole structure and gives you a very nice answer, which you will maybe one of the answers would be discussed, so or maybe I will put them in the f a q's or in the notes.

(Refer Slide Time: 17:52)



So, this your function, so objective function this is the same one same sort of structure that we see in problem one, this is your problem number four and here there are some interesting constraints summation x i is 0. So, could some of the x i's are positive, some are negative and all those things sum of x i square is n. Now, very important question is, whether such a set of whether we have any a here we cannot talk about coercivity, because we have a cubing structure.

So, the cube structure that is a something is going to minus infinity some value of x i, because norm has to go to infinity. So, if x i is going to minus infinity x i q also goes to minus infinity. So, this is not a fair deal, actually to talk about coercivity here, so we have to decide by some other method this is a continuous function. So, my feasible set C, here it is very important to note that, this set is not only closed, but I would leave you as again as homework to show that C is compact.

Now, once you show that c is compact, you have actually told that there is already a solution to this problem, so a solution to this problem exists. Now, consider this as g 1 x and consider this as g 2 x. So, I am now going to ask whether these are linearly independent g 2 x means of course, this minus n rather I should write like this, whether these conditions are these constraints are linearly independent solutions, they are linearly independent at the solution. The gradient of this and this if the gradient of this and this form linearly independent set at the solution, then we are thoroughly sure that all the multipliers would be normal and that is exactly the way we need to now proceed. Now, let us compute the gradient this is and now whenever, these linearly independent, when x is a multiple of e when x is, so grad g 1 not linearly independent and linearly dependent.

We have to talk about it, we shall write grad g 1 and grad g 2 x or. So, x star suppose, x is not the solution are linearly dependent, then only we cannot say whether we will always have a normal multiplier and a abnormal multiplier. We can have anything, we can have, we will obviously have one abnormal, but we can also have normal. So, let us first check whether these are this is true.

So, in this particular case, let us observe that you have. So, when would this happen? So, if x is equal to some lambda times e, then you will have these are to be linearly independent. So, x is lambda time e some lambda then, is such an x feasible. Now, if I look at summation x i right, then this would be equal to lambda, but lambda cannot be 0 right, if x is 0 then lambda is of course, 0 then it is satisfying if x, but if all of the x cannot be 0, because then summation x i square cannot be equal to n, then n would be 0 which is not.

So, which means x is not 0, so all the lambdas, this lambda cannot be a 0 vector when e is not 0. So, this would not equal to 0. So, if x is equal to lambda e now, I am putting lambda around this is lambda, lambda, lambda, lambda, lambda. So, I am putting lambda in place of x i's, so which is violating the constraint. So, any such thing, this thing cannot be feasible this cannot be feasible. So, you see just from the problem conditions, we are been able to figure out whether all the John multipliers are normal or not. So, which means now we can simply apply, what is we will call the KKT condition right. Now, we will go to our next study the Quasi Newton method and which we had been postponing for quite some time in the next class.

(Refer Slide Time: 25:16)