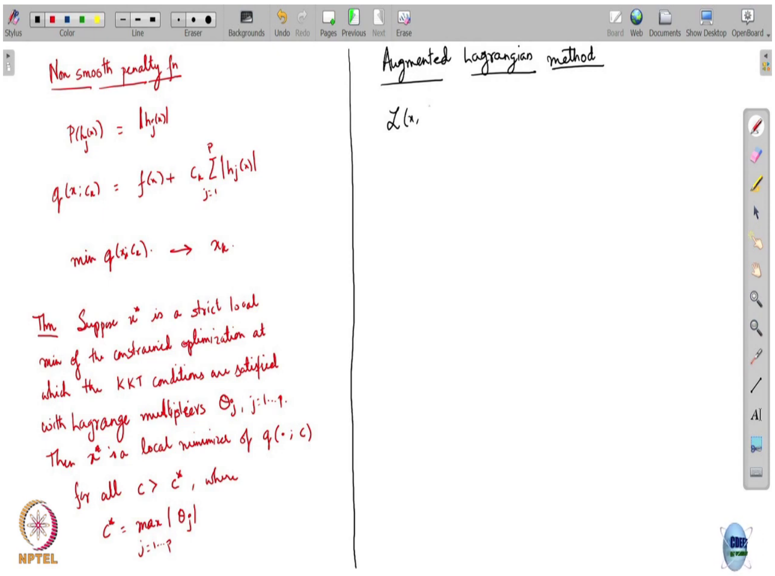
Optimization from Fundamentals
Prof. Ankur Kulkarni
Department of Systems and Control Engineering
Indian Institute of Technology, Bombay

Lecture - 21B
Augmented Lagrangian Methods

(Refer Slide Time: 00:21)

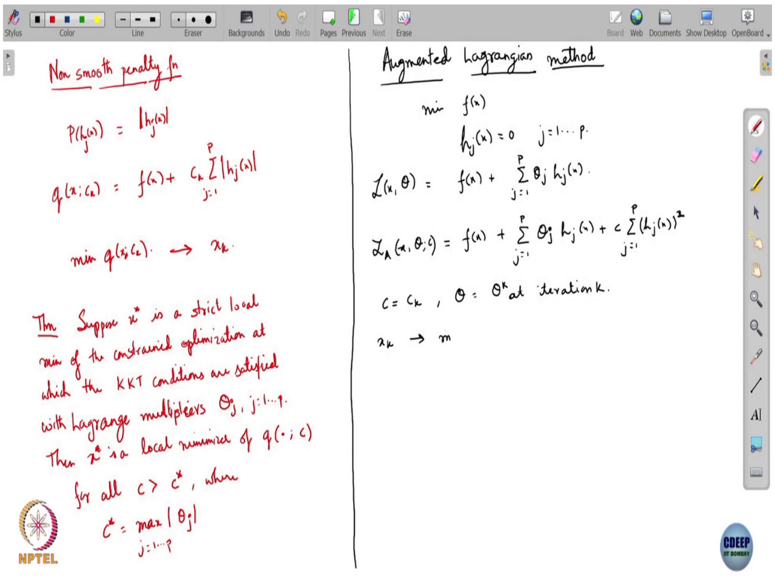


Now, another way around this is what I will now talk about. It is what are called Augmented Lagrangian Methods. Now, the augmented Lagrangian method does something slightly different in the sense that it is not just looking for a minimize minimization of a penalized problem, it is doing that it is looking for the minimization of a penalized problem, but then it is not just doing this blindly. It is minimizing, it is looking at a certain type of penalized problem it minimizes that penalized problem.

But then the penalty associated with it is not going to be increase blindly rather it is going to be tuned very carefully to the way we are getting solutions of, solutions from this particular penalized problem. So, the effort is to eventually start mimicking these KKT conditions of the constrained optimization problem.

And at the same time making sure that you get feasibility at a finite value you know even for finite values of the penalty parameter, alright. So, let me give you I will explain to you now what the augmented Lagrangian method is, ok. So, recall what we know about what we had used as what is called the Lagrangian. The Lagrangian was a function of x and the Lagrange multiplier is associated with it. So, I am going to look at it problem where the where we; so, let me first write out the optimization problem.

(Refer Slide Time: 02:13)



So, let us first look at a problem where with simply an equality with only an equality constraints. So, you are minimizing a function f subject to; say for simplicity let us have these only these equality constraints h j of x equal to 0 for j equals 1 to P, alright.

Now, the Lagrangian is defined in this way. So, the Lagrangian was L of x and here if these are Lagrange multipliers theta, so we define the Lagrangian as f plus summation theta j h j of x. The augmented Lagrangian is this particular quantity. The augmented Lagrangian is L of x comma theta which is f of x plus.

So, this was summation from 1 to P, is again summation from 1 to P. So, you have continue to have this particular thing theta j, j of x plus you also include in this your quadratic penalty and outside you can put a penalty parameter c, alright. So, you have your quadratic penalty added on top of the what is the usual Lagrangian that is why this is called an augmented Lagrangian. So, you are augmenting the Lagrangian with this additional term, right.

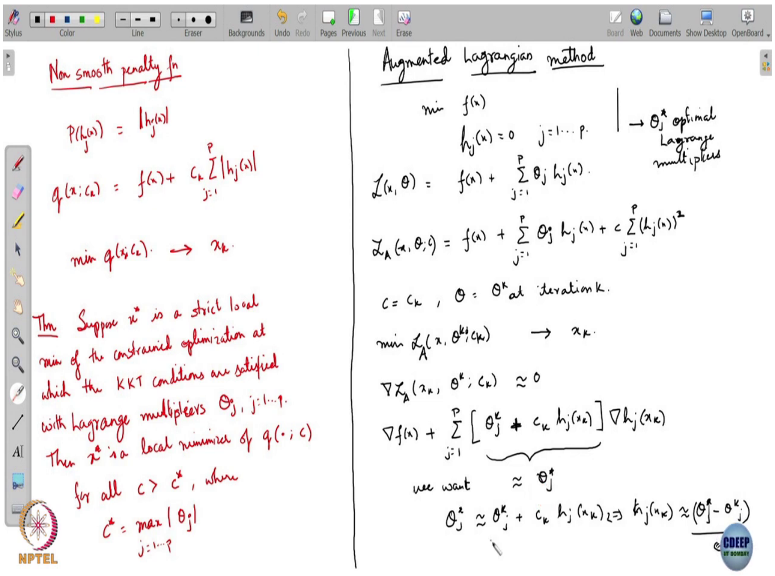
Now, what we would, what we would do is that, the algorithm would what it would do now is you can notice that you do not have just the primal variable x, but you also have the dual variable or the Lagrange multiplier theta present here. So, it will continue to adjust x and theta at each step. So, but and at the same time it will and it will keep increasing c k as well. So, actually to make to make the dependence on c explicit let me write this as this given c, alright, ok.

So, let us suppose we take the; so, what it the whole idea is to start to have an the, is to adjust your x and theta in such a way that you start with an initial estimate of x, you start with an initial estimate of theta and then as the iterations go along you start refining your estimates.

And the goal is to get, and is to refine them in that in such a direction that at the, that eventually at convergence you would end up the minimizing this augmented Lagrangian would amount to the same as minimizing the true Lagrangian and hence solving the KKT conditions, alright. So, let me explain this, ok.

So, suppose we have I have for sake of arguments suppose we have fixed to motivate this suppose we have fixed c at c k. So, c equals, suppose I fix c equal to some c k and I fix theta equal to some theta some theta k, ok at iteration k. And now if we end up, if at iteration k. Now, with this suppose I minimize the augmented Lagrangian, alright. So, then that would make, that if I minimize the augmented Lagrangian to get a value x k, the x k is obtained.

(Refer Slide Time: 06:20)



So, you minimize L of x comma the theta k given c k, alright. You minimize this; you know and by minimizing this you get x k, alright. Now, this would mean sorry; the augmented Lagrangian. So, now, this would mean that the gradient of the augmented Lagrangian with respect to x must be close to 0 or exactly 0. If you have come to the exact solution, so we would need that at x k this theta k comma c k should be approximately equal to 0.

You calculate the gradient of the augmented Lagrangian that turns out to be gradient of f, gradient of f plus summation j goes from 1 to P theta k j minus c sorry, plus c k times h j of x k, the whole thing times gradient of h j evaluated at x k. Now, if you look at, if you compare this with your KKT conditions what you would want is that eventually this term here which is in the bracket this should start resembling, we want that this should start resembling your optimal Lagrange multiplier, ok.

So, suppose this is, so suppose theta j, so suppose theta j star are the optimal Lagrange multipliers, then in that case we would want this quantity the one that is that I have put an under brace below that quantity should start resembling or start approaching theta j star, right.

Now, now I can rearrange this and write this in the following way. So, I write this as theta j star to be approximately equal to theta k j plus c k times h j of x k or equivalently I say h j of x k is approximately equal to theta j star minus theta k j, the whole divided by c k. Now, what this means is that if somehow I am able to get the here this is the consequence of this little calculation is that it tells me that if I am somehow able to get this term here, the numerator to be close, close to 0.

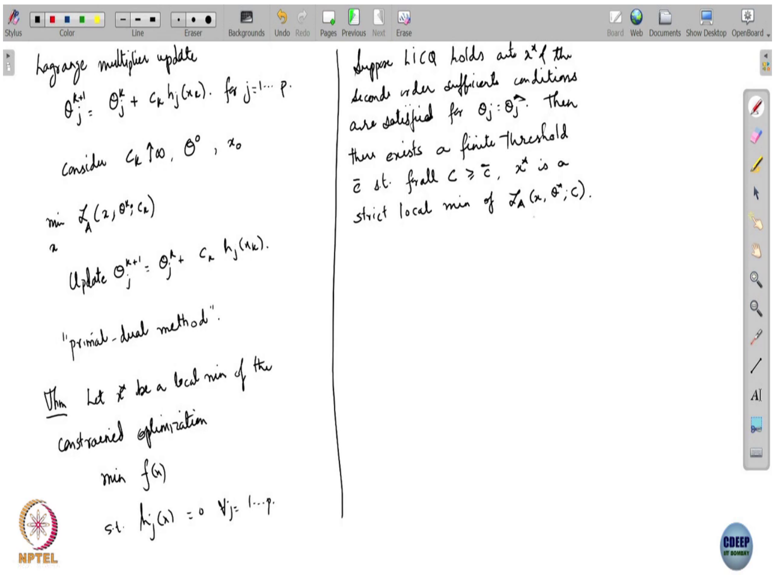
That is means if I have my, if I have my Lagrange multiplier correctly figured out; that means, if theta k is close to theta star then even when c k is finite I should still be able to make this get close become, make my h j of h j of x k close to 0. That means, that should approach feasibility even for finite values of c k.

Recall now this is exactly what we were trying to achieve when I when we introduce this this non-differentiable penalty function. With the non-differentiable penalty function we were able to do this by keeping, we were able to use the finite value of the penalty parameter and yet get to the true solution at and get to a feasible solution. Here is, simply something similar is happening. With the finite, even with the finite value of the penalty parameter we are able to get feasibility, right.

Now, this particular equation also gives us a hint on how exactly should one update, how exactly one should one update the penalty, how; sorry the value of the theta k, right. So, what we did right now was we fixed the value of c k, we fixed the value of theta k and we said we let us minimize the augmented Lagrangian to get is to get x k.

And then, what we are saying is well if we were at nearly at the true solution the way it should behave is that this should become equal to theta this should become equal to the next to the next, to the optimal Lagrange multiplier.

(Refer Slide Time: 12:09)



So, thinking from this what we the way we update now are Lagrange multiplier. So, the Lagrange multiplier update, Lagrange multiplier update is simply this that, I do theta k plus 1 j equals theta k j plus c k times h j of x k; and does this for all j going from 1 to P, right. So, this is basically the idea.

So, the overall augmented Lagrangian method is then to do the following. You let you consider, you consider a sequence of penalty parameters consider c k that increases, an increasing sequence of penalty parameters. And so, this can I have taken this to go to infinity, but it is basically simply an increasing sequence, alright.

So, you choose an increasing sequence of penalty parameters for, and I start off with some value of theta, start off with some value of theta 0 which is your initial guess about the Lagrange multiplier. Start off with some value of x 0 which is your initial guess about the primal solution.

And then you iterate by doing the following. You minimize be, at each iteration you minimize, this overall x and alongside after once you get the x you also update theta k plus 1 as theta k for the jth constraint plus c k h j of x k. And you do this now in infinite loop.

So, what one is doing effectively is one is doing by doing these 2 things, you are increasing the penalty for a parameter and you are simultaneously adjusting the primal and the dual variables. So, this sort of a method is what is called a method which, this sort of a method is what is an example of a primal dual method.

Why is this a primal dual method? Because it is not just simply searching in the primal space. It is searching in the primal space and also at the same time, using the using what it is learnt from the primal space to also inform or update the current guess about the dual variables.

So, effectively it is you, if you think about it is a operating in some sense simultaneously in the primal dual space though priority is in some sense given to the primal space and the dual space is treated almost as a parameter. So, one of the big lessons in optimization is that optimization is best seen you know through neither the primal space nor the dual space, but rather the jointly the primal dual space. When we did, when we did a study of optimization of duality in optimization.

For example, our way of addressing that problem was through the cost constraint pairs, right. That could be attained by an optimization problem. So, that object was an object that lied jointly in the objective and the constraint space. So, analogous to that in the decision space is the, that is in the values space, so in the decision space the analogous thing to do is look at primal variables as well as simultaneously the dual variables.

So, this the augmented Lagrangian method is so effective because of this, because one, because at it is heart it is trying to play with both variables at once. So, and is therefore, attacking the problem in the in you know using all the levers available. Although, it is not a full blown primal dual method.

A full blown primal dual method is an example of a full blown primal dual method would be an interior point method that is the method I will talk to you about next, but this method come is sort of a stepping stone towards interior point methods, right.

So, the main theorem that we can, we have for this, for an augmented Lagrangian method is that if essentially if. So, let x star be a local minimum of the constrained optimization problem, constrained optimization minimize f subject to h j of x equal to 0 for all j going from 1 to P.

Suppose L i C Q holds at x star and second order sufficient conditions are satisfied for theta j equal to some theta j star, then there exists a threshold or let us say a finite threshold c bar such that for all c greater than equal to c bar, x star is a strict local minimizer of.

So, it is a strict local minimizer of this augmented Lagrangian, ok. So, from this theorem; so, this is now not a complete theorem of convergence of the augmented Lagrangian method that is, that takes a little bit more to state its quite a mouthful to state that. But essentially from here what we are seeing is that if I fix the Lagrange multiplier at the optimal value then it is enough for me to keep a finite value of the, finite value of the penalty parameter.

And with that finite value of the penalty parameter I can I my original the solution of my original problem can also be obtained by minimizing the augmented Lagrangian function. Now, what one can do is from build on this and then work with not the exact value of the Lagrange multiplier.

But rather value of the Lagrange multiplier that is approaching the optimal. And then from there you get you one can conclude the convergence of the convergence of the augmented Lagrangian method to a KKT point of, to a KKT point of the constrained optimization problem, alright, ok.

So, this basically brings us to a close on this particular topic of augmented Lagrangian methods. But there is an, before I move to interior point methods, there is another a type of method which I wanted to highlight and wanted to teach you about which is what is called a cutting plane method.

Now, this sort of method actually relies on convexity and makes you know very very clever use of the convex analytic geometry of the problem. So, it is not a very general method, but it is probably, it is a very elegant method, ok. So, I would say that let; so, I thought I would just talk to you about it.