Numerical Optimization
Prof. Shirish K. Shevade
Department of Computer Science and Automation
Indian Institute of Science, Bangalore

Lecture - 41
Summary

Now, let me summarize the topics that we studied in this course. So, we started with the introduction to optimization problem where we saw that optimization has lot of application in science and engineering. So, many problems in physics can be solved using optimization techniques. Similarly, many problems in engineering can also be solved using some of the optimization techniques. Now, most of the practical optimization problems are constraint optimization problems. And as we saw in this last few lectures that the constraint optimization problems can be solved using unconstrained optimization ideas. So, after see some mathematical background on matrices, vector spaces, we moved on to study convex functions. In many applications, convex sets and convex functions play an important role and one could point about the convex function is that there is no problem of local minima. So, we saw that for convex function every local minimum is the global minimum and the set of solutions form of convex sets.

So, we discussed about the convex sets, convex functions and how to check the convexity of twice differentiable functions and then we studied the theory of unconstrained optimization. In particular, we discussed about the first order necessary conditions and second order necessary and sufficient conditions. We also saw that for unconstrained minimization of convex function, the first order necessary condition of optimality are also sufficient. Now, after studying this theory of unconstrained optimization view, the studied some of the methods those could be used for solving unconstrained optimization problems. So, we started with the (( )) method and then studied Newton method, quasi Newton method and conjugate gradient method.

(( )) method use is only the derivative information, while the Newton method you just the second order information. Now, we also saw that Newton method requires the inversion of the hessian matrix to be computed at every iteration and which is computationally expensive. And further, Newton method may not converges from any starting point. So, then we discussed about quasi Newton's methods which approximate the inverse of the hessian by a positive definite matrix. So, the quasi Newton methods do not use the second order information directly, but for quadratic functions we sure that when the iterations terminate the hessian of the original objective function would be the inversion of the hessian of the original objective function would be obtained at the at the termination.

We also studied conjugate gradient methods in this course. So, we saw that for all the optimization methods, it was important to ensure different conditions to achieve global conversion. In particular, if we start from in an arbitrary point and make sure that there is at least a constant decrease in the objective function at every iteration and the sufficient angle decent conditions is satisfied then the optimization algorithm would move closer to the optimal point or in the limit it will go in the vicinity of the optimal point. So, we saw this is global convergence theorem and also we discussed about the different rates at which different optimization algorithms converges to the optimal solution.

We then moved on to study the theory of constraint optimization. In particular, we considered a general linear program of the type minimize the effects subject to the inequality consume inequality constraints and some equality constraints. And we saw the (( )) KKT condition for the optimality of the given optimization problem. So, we also saw the first order and second order necessary and sufficient conditions for an optimality for an optimal point for given constraints optimization problem. Similar to the unconstrained optimization problems, for the convex programming problems, where the objective function is to minimize a convex, what the objective is to minimize the convex function subject to convex constraint set.

We saw that the first order KKT conditions are necessary and sufficient, but this may not hold for general non-linear programming problem. We also saw that when we talk about the constraint convex programming problems, the first order KKT conditions are sufficient only when the letters conditions are satisfied. So, for other non-linear programs under the regularity assumptions, we saw that the KKT conditions are sufficient; the second order KKT conditions are satisfied. We then discussed about the duality theory; duality is a very important concept in optimization, and we saw that many problems which are difficult to solve, can be solve very easily (( )) by solving the dual problems. In particular, we saw some examples where the number of variables is very large and the number of constraints is small,. So, in such cases it may be a good idea to convert the original primal problem into the dual problem, and then solve the dual problem.

Now, we also saw some conditions under which the solution, the optimal solution of a primal problem and the optimal solutions of a dual problem are same. So, in such cases, if we are not going to lose anything by solving the dual problems, and if the dual problem is easier to solve than the primal problem. It may be a good idea, to solve the dual problem. The dual function also have some advantages, again particular the dual problem is the convex programming problem, irrespective of the original primal problem. So, when we discuss the theory of convex optimization, we saw that there is no problem of local minima as far as convex optimization is concerned.

So, one can utilize this fact and the fact that sometimes that dual programs are easier to solve to write the,. So, these facts can be use to convert the original primal problem to the dual problem and then solve it and under certain special conditions the optimal primal and the optimal dual values are the same. We then studied linear programming,. So, in particular we saw how to use KKT condition to design and algorithm to solve a linear program. Now, the simplex method, which has been very popular to solve linear programs was also discussed. So, we use the KKT conditions to find out the incoming and outgoing basic variables, incoming non-basic and outgoing basic variables at every point of iteration. One important thing about linear programs is that the solution lies at the extreme point of the feasible regions. So, it is enough to search through the set of extreme points to get a solution, and simplex method gives us the systematic way of moving from one vertex to another and find the find the optimal solution.

Now, the simplex method has some drawbacks like in some cases it may have to search through all possible vertexes to find the solution, and the number of vertexes could be exponentially large in number. And therefore, simplex method of simplex algorithm is not a polynomial type algorithm. We then saw some of the interior point methods that could be used to solve linear program, in particular we saw (( )) scaling method and (( )) methods. So, Kermerkar methods you just a projective transformation to convert the original linear program into an equivalent program, and then certain assumptions that equivalent program can be treated as a linear program. And the idea of this transformation is to get reasonable improvement in the objective function. And if this procedure is repeated at every iteration, try to transform the given point x, k into the voice phase to get y, k such that y, k is close to the interior of the close to the center of the feasible region. And then take a step in the (( )) directions and this procedure is repeated Kermerkar showed that that this procedure will have a polynomial type algorithm, which is much better than (( )) ellipsoidal method.

We also saw variant of Kermerkar's method and that is the (( )) scaling method where the transformation use this different than the one use by Kermerkar. Although this (( )) scaling method is simple and easy to implement, it is not a polynomial type algorithm, but it was very simple method to study. So, we saw that method. After studying linear programming and then the associated dual problem, we studied some of the algorithm for solving constrained optimization problems. Now, there is plenty of techniques to solve constrained optimization methods and we have chosen only a few of them in this course. To techniques that we discussed where active set method, (( )) function methods, then augmented Lagrangian method and (( ) method.

Active set methods are very simple and easily to implement. So, the idea is that at every iteration given a feasible point x, k, we find out the set of constrained which are active and solve the optimization problem with respective to those active constrains. Now, the solution of this optimization problem may be either within the feasible region or may be away from the feasible region. So, if it is within the same, if it is within the feasible region, we find out the new set of active constrained. Otherwise if the solution was outside the feasible region we ensure that the, the new point is broad back to the feasible set and then at that point the set of active constrains is found and the same procedure is repeated. Now, if the objective function is convex then one can show that if every time we are going to decrease objective function then this procedure will terminate in finite number of iterations, because there are only finite combinations of active constrains for a given problems.

We also saw the penalty function and the barrier function method. The idea is that instead of minimizing the problem effects subject to some constraint set one can define an auxiliary function using either a penalty function or a barrier function and solve an unconstraint optimization problems. Now, the the idea is to use a sequence of auxiliary functions and solve the corresponding unconstrained problems at every iteration. And if we repeat this procedure, you will get sequence x k of optimal excess for every objective for every auxiliary function and this x ks would converse to x star as the iterations progress. So, that is the idea of penalty function methods or barrier function methods.

The barrier function methods are similar to the interior point methods, which start with of feasible points and then ensure that at every iteration the point remains feasible. So, at no point of time the barrier function methods, let iteration point x k go beyond the feasible region. On the other hand, the penalty function methods, start the with point which needs not to be feasible, and then slowly bring it towards the feasible region, as the iteration progress. Now, both these methods, do the depend on the parameter see and as see go to infinity this methods will face some numerical difficulties. So, we discuss the idea of augmented Langragian methods; now this methods use the perturbed optimization problem, and estimate the optimal value of the Langragian multipliers at every iteration.

So, the idea is very simple, and can also be extended to the inequality constraint problems. By writing the inequality constraints as equality constraints using a new set of variables. So, if you write the inequality constraints as equality constraints and then solve the problem, the augmented langrage methods are very useful. And we also saw finally, the cutting plane methods, which are very useful and we solve the approximate dual programs, because of original dual problems use the infinitely constraint linear program. Now, all these methods are very popular and the choice of the methods depends on the problem at hand. So, this complete the discussion in this course, I hope this course contents were very useful.

Thank you, also would like to thank the NPTEL staff for their assistance. So, thank you very much.