Lecture XIII
Algorithms for Nonlinear Constraints

I. General Objective

A. The general objective is to maximize or minimize a nonlinear objective function subject to nonlinear constraints. Gill, Murray, and Wright lay out the basic problems:

$$NEP \min_x f(x)$$
$$st\ c_i(x) = 0 \quad i = 1,..t$$

and

$$NIP \min_x f(x)$$
$$st \quad c_i(x) \geq 0 \quad i = 1,\ldots t$$

B. These optimization problems are much more difficult than the equality or inequality constraints posed in the preceding sections due to the difficulties involved in maintaining feasiblity.

1. By the formulation of the null-space in the linear equality scenario, it was always possible to guarantee that x_{k+1} was feasible given that x_k was feasible.
2. Similarly, expansion to linear inequality constraints only added caveats to the step length algorithm and checks on whether a constraint could be deleted.
3. However, the solution of nonlinear constraints may be difficult, if not impossible, without the incorporation of a nonlinear objective function.

C. I want to discuss three different methodologies for optimization under nonlinear constraints. The first two are the use of penalty and barrier functions. The third procedure is the projected augmented Lagrangian method.

II. Penalty Functions

A. The penalty and barrier function procedures are very similar. The primary concept of both procedures is to append an additional term onto the objective function which imposes a cost for violating the constraint.

B. Taking the Penalty function first, assume that we want to optimize

$$\min_x x^2$$
$$st\ x - 1 = 0$$

Granted that the constraint is linear, and could be solved using a projected quasi-Newton algorithm, a quadratic penalty function for this problem could be specified as

$$P_Q(x, r) = x^2 + \frac{r}{2}(x-1)^2$$

where the ρ term yields a penalty for violating the constraint. Solving this problem while increasing the penalty parameter yields an optimal solution with the constraint imposed.

C. Algorithm DP (Model algorithm with a differentiable penalty function)
   - DP1. [Check termination conditions.] If x_k satisfies the optimality conditions, terminate with the current solution.
   - DP2. [Minimize the penalty function.] Using x_k as the starting point, execute an algorithm to solve the unconstrained subproblem

     $$\min_x P_Q(x, p_k)$$

     subject to safeguard conditions to protect against unboundedness. Let x_{k+1} be the estimate of this subproblem.
   - DP3. [Increase the penalty parameter.] Set ρ_{k+1} to a larger value than ρ_k. and go back to step DP1.

II. Barrier Function

A. The barrier function works well when the constraint can be evaluated either above or below its constrained value. Specifically, if the nonlinear constraint is an equality, then a sufficient penalty parameter can be derived to force that the constraint be close to the right hand side. However, in certain cases, you may want to guarantee feasibility as a minimum condition. For this type of problem the barrier method may be preferred.

B. Focusing on the logarithmic barrier function, we create a subproblem similar to that generated in the penalty parameter framework. Specifically

$$B(x,r) = f(x) - r\sum_{i=1}^{m} \ln(c_i(x))$$

where r is the barrier parameter and the ln(.) term in the maximization problem guarantees that the $c_i(x)$ will be greater than zero.

III. Projected Augmented Lagrangian

A. Minos 5.1 uses the projected augmented Lagrangian algorithm to optimize problems involving nonlinear constraints.

B. This algorithm as instituted in Minos solves a sequence of subproblems. Each subproblem, or major iteration, solves a linearly constrained minimization (maximization) problem. The constraints for this linear subproblem are the linear constraints plus the linearlized nonlinear constraints. Then like the straightforward penalty parameter method, the penalty can be increased to approach the nonlinear constraints with an arbitrary level of precision.

1. Mathematically, the general problem can be written

$$\begin{aligned}\min_x \quad & f(x) + c'x + d'y - l'(G(x) - G(x_k)) + \frac{1}{2} r(G(x) - G(x_k))'(G(x) - G(x_k))\\ st \quad & G(x) + A_1 y = b_1\\ & A_2 x + A_3 y = b_2\\ & l \le \begin{pmatrix} x \\ y \end{pmatrix} \le u\end{aligned}$$

2. In this problem, the major iteration involves the linearlization of the constraints around the point x_k.

a. The first step is to linearlize the nonlinear constraints around the starting point of the major iteration:

$$G(x) \approx G(x_k) + \nabla_x G(x_k)(x - x_k)$$
$$= G(x_k) + \nabla_x G(x_k)\, x - \nabla_x G(x_k) x_k$$

b. Given this relationship, the linear constraints can then be written as:

$$\begin{pmatrix} \nabla_x G(x_k) & A_1 \\ A_2 & A_3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_1 + \nabla_k G(x_k)\, x_k - G(x_k) \\ b_2 \end{pmatrix}$$

Ignoring the slackness conditions