Nonlinear Programming Methods.S3
Primal Methods

In solving a nonlinear program, primal methods work on the original problem directly by searching the feasible region for an optimal solution. Each point generated in the process is feasible and the value of the objective function constantly decreases. These methods have three significant advantages: (1) if they terminate before confirming optimality (which is very often the case with all procedures), the current point is feasible; (2) if they generate a convergent sequence, it can usually be shown that the limit point of that sequence must be at least a local minimum; (3) they do not rely on special structure, such as convexity, so they are quite general. Notable disadvantages are that they require a phase 1 procedure to obtain an initial feasible point and that they are all plagued, particularly when the problem constraints are nonlinear, with computational difficulties arising from the need to remain within the feasible region as the algorithm progresses. The convergence rates of primal methods are competitive with those of other procedures, and for problems with linear constraints, they are often among the most efficient.

Primal methods, often called *feasible direction methods*, embody the same philosophy as the techniques of unconstrained minimization but are designed to deal with inequality constraints. Briefly, the idea is to pick a starting point satisfying the constraints and to find a direction such that (*i*) a small move in that direction remains feasible, and (*ii*) the objective function improves. One then moves a finite distance in the determined direction, obtaining a new and better point. The process is repeated until no direction satisfying both (*i*) and (*ii*) can be found. In general, the terminal point is a constrained local (but not necessarily global) minimum of the problem. A direction satisfying both (*i*) and (*ii*) is called a *usable feasible direction*. There are many ways of choosing such directions, hence many different primal methods. We now present a popular one based on linear programming.

Zoutendijk's Procedure

Once again, we consider problem (23) with constraint set is $S = \{\mathbf{x} : g_i(\mathbf{x}) \leq 0, i = 1,\ldots,m\}$. Assume that a starting point $\mathbf{x}^0$ S is available. The problem is to choose a vector $\mathbf{d}$ whose direction is both usable and feasible. Let $g_i(\mathbf{x}^0) = 0$, i I, where the indices in I correspond to the binding constraints at $\mathbf{x}^0$. For feasible direction $\mathbf{d}$, a small move along this vector beginning at the point $\mathbf{x}^0$ makes no binding constraints negative, i.e.,

$$\frac{d}{dt} g_i(\mathbf{x}^0+t\mathbf{d})\Big|_{t=0} = \quad g_i(\mathbf{x}^0)^{\mathrm{T}}\mathbf{d} \leq 0, \quad i \quad I$$

A usable feasible vector has the additional property that

$$\frac{d}{dt} f(\mathbf{x}^0+t\mathbf{d})\Big|_{t=0} \quad = \quad f(\mathbf{x}^0)^T\mathbf{d} < 0$$

Therefore the function initially decreases along the vector. In searching for a "best" vector **d** along which to move, one could choose that feasible vector minimizing $f(\mathbf{x}^0)^T\mathbf{d}$. If some of the binding constraints were nonlinear, however, this could lead to certain difficulties. In particular, starting at $\mathbf{x}^0$ the feasible direction $\mathbf{d}^0$ that minimizes $f(\mathbf{x}^0)^T\mathbf{d}$ is the projection of $-\ f(\mathbf{x}^0)$ onto the tangent plane generated by the binding constraints at $\mathbf{x}^0$. Because the constraint surface is curved, movement along $\mathbf{d}^0$ for any finite distance violates the constraint. Thus a recovery move must be made to return to the feasible region. Repetitions of the procedure lead to inefficient zigzagging. As a consequence, when looking for a locally best direction it is wise to choose one that, in addition to decreasing f, also moves away from the boundaries of the nonlinear constraints. The expectation is that this will avoid zigzagging. Such a direction is the solution of the following problem.

$$\text{Minimize} \quad \xi \tag{28a}$$

$$\text{subject to} \quad g_i(\mathbf{x}^0)^T\mathbf{d} - \phi_i\xi \le 0, \quad i \quad I \tag{28b}$$

$$f(\mathbf{x}^0)^T\mathbf{d} - \xi \le 0 \tag{28c}$$

$$\mathbf{d}^T\mathbf{d} = 1 \tag{28d}$$

where $0 \le \phi_i \le 1$ is selected by the user. If all $\phi_i = 1$, then any vector $(\mathbf{d}, \xi)$ satisfying (28b) - (28c) with $\xi < 0$ is a usable feasible direction. That with minimum ξ value is a best direction which simultaneously makes $f(\mathbf{x}^0)^T\mathbf{d}$ and $g_i(\mathbf{x}^0)^T\mathbf{d}$ as negative as possible; i.e., steers away from the nonlinear constraint boundaries. Other values of ϕ_i enable one to emphasize certain constraint boundaries relative to others. Equation (28d) is a normalization requirement ensuring that ξ is finite. If it were not included and a vector $(\mathbf{d}, \xi)$ existed satisfying (28b) - (28c) with ξ negative, then ξ could be made to approach $-\infty$, since (28b) - (28c) are not homogeneous. Other normalizations, such as $|d_j| \le 1$ for all j, are also possible.

Because the vectors f and g_i are evaluated at a fixed point $\mathbf{x}^0$, the above direction-finding problem is almost linear, the only nonlinearity being (28d). Zoutendijk showed that this constraint can be handled by a modified version of the simplex method so problem (28) may be solved with reasonable efficiency. Note that if some of the constraints in the original NLP (1) were given as equalities, the algorithm would have to be modified slightly.

Of course, once a direction has been determined, the step size must still be found. This problem may be dealt with in almost the same manner as in the unconstrained case. It is still desirable to minimize the objective function along the vector **d**, but now no constraint may be violated. Thus t is determined to minimize $f(\mathbf{x}^k + t\mathbf{d}^k)$ subject to the constraint $\mathbf{x}^k + t\mathbf{d}^k \quad S$. Any of the techniques discussed in Section 11.6 can be used. A new point is thus determined and the direction-finding problem is re-solved. If at some point the minimum $\zeta \geq 0$, then there is no feasible direction satisfying $\quad f(\mathbf{x}^0)^{\mathrm{T}}\mathbf{d} < 0$ and the procedure terminates. The final point will generally be a local minimum of the problem. Zoutendijk showed that for convex programs the procedure converges to the global minimum.