G.L. Nemhauser et al., Eds., *Handbooks in OR & MS, Vol. 1*

Chapter I

A View of Unconstrained Optimization

J.E. Dennis Jr.

Department of Mathematical Sciences, Rice University, Houston, TX 77251, U.S.A.

Robert B. Schnabel

Department of Computer Science, Campus Box 430, University of Colorado, Boulder, CO 80309, U.S.A.

1. Preliminaries

1.1. *Introduction*

This chapter discusses the most basic nonlinear optimization problem in continuous variables, the *unconstrained optimization* problem. This is the problem of finding the minimizing point of a nonlinear function of n real variables, and it will be denoted

$$\text{given } f:\mathbb{R}^n \to \mathbb{R}, \quad \underset{x\in\mathbb{R}^n}{\text{minimize}}\, f(x)\,. \tag{1.1}$$

It will be assumed that $f(x)$ is at least twice continuously differentiable.

The basic methods for unconstrained optimization are most easily understood through their relation to the basic methods for a second nonlinear algebraic problem, the *nonlinear equations* problem. This is the problem of finding the simultaneous solution of n nonlinear equations in n unknowns, denoted

$$\text{given } F:\mathbb{R}^n \to \mathbb{R}^n\,, \quad \text{find } x_* \text{ for which } F(x_*) = 0 \tag{1.2}$$

where $F(x)$ is assumed to be at least once continuously differentiable. Therefore, this chapter also will discuss the nonlinear equations problem as necessary for understanding unconstrained optimization.

Unconstrained optimization problems arise in virtually all areas of science and engineering, and in many areas of the social sciences. In our experience, a

* Research supported by DOE DE-FG05-86ER25017, SDIO/IST/ARO DAAG-03-86-K-0113, and AFOSR 85-0243.

** Research supported by ARO contract DAAG 29-84-K-0140 and NSF grant CCR-8702403.

significant percentage of real-world unconstrained optimization problems are data fitting problems (see Section 6.2). The size of real-world unconstrained optimization problems is widely distributed, varying from small problems, say with n between 2 and 10, to large problems, say with n in the hundreds or thousands. In many cases, the objective function $f(x)$ is a computer routine that is expensive to evaluate, so that even small problems often are expensive and difficult to solve.

The user of an unconstrained optimization method is expected to provide the function $f(x)$ and a starting guess to the solution, x_0. The routine is expected to return an estimate of a *local minimizer* x_* of $f(x)$, the lowest point in some open subregion of $\mathbb{R}^n$. The user optionally may provide routines for evaluating the first and second partial derivatives of $f(x)$, but in most cases they are not provided and instead are approximated in various ways by the algorithm. Approximating these derivatives is one of the main challenges of creating unconstrained optimization methods. The other main challenge is to create methods that will converge to a local minimizer even if x_0 is far from any minimum point. This is referred to as the *global* phase of the method. The part of the method that converges quickly to x_*, once it is close to it, is referred to as the *local* phase of the method.

The emphasis of this chapter is on modern, efficient methods for solving unconstrained optimization problems, with particular attention given to the main areas of difficulty mentioned above. Since function evaluation so often is expensive, the primary measure of efficiency often is the number of function (and derivative) evaluations required. For problems with large numbers of variables, the number of arithmetic operations required by the method itself (aside from function evaluations) and the storage requirements of the method become increasingly important.

The remainder of the section reviews some basic mathematics underlying this area and the rest of continuous optimization. Section 2 discusses the basic local method for unconstrained optimization and nonlinear equations, Newton's method. Section 3 discusses various approaches to approximating derivatives when they aren't provided by the user. These include finite difference approximations of the derivatives, and secant approximations, less accurate but less expensive approximations that have proven to lead to more efficient algorithms for problems where function evaluation is expensive. We concentrate on the most successful secant method for unconstrained optimization, the 'BFGS' method, and as motivation we also cover the most successful secant method for nonlinear equations, Broyden's method. Sections 2 and 3 cover both small and large dimension problems, although the solution of small problems is better understood and therefore is discussed in more detail. Methods that are used when starting far from the solution, called global methods, are covered in Section 4. The two main approaches, line search methods and trust region methods, are both covered in some detail. In Section 5 we cover two important methods that do not fit conveniently in the Taylor series approach that underlies Sections 2 through 4. These are the Nelder–Mead simplex method,