RECENT ADVANCES IN UNCONSTRAINED OPTIMIZATION *

M.J.D. POWELL
A.E.R.E., Harwell, Berks., Great Britain

Received 5 January 1971

We survey the development of algorithms and theory for the unconstrained optimization problem during the years 1967-1970. Therefore (except for one remark) the material is taken from papers that have already been published. This exception is an explanation of some numerical difficulties that can occur when using Davidon's (1959) variable metric algorithm.

1. Introduction

The unconstrained optimization problem is to calculate the least value of a given function, \( F(\mathbf{x}) \) say, where \( \mathbf{x} \) denotes a column vector of real variables. We let \( n \) be the number of variables, so the components of \( \mathbf{x} \) are \( \{x_1, x_2, ..., x_n\} \). The term "unconstrained" implies that the value of each variable can be any real number.

Many computer algorithms have been proposed for solving this problem, and the most convenient to use are those that require one to provide a computer subroutine which calculates only the numerical value of \( F(\mathbf{x}) \) for any value of \( \mathbf{x} \). However most algorithms require the subroutine to calculate some derivatives of \( F(\mathbf{x}) \) also. Thus the objective function is specified. Then each algorithm employs some strategy, based on calculated function values, that tries to adjust the variables \( \{x_1, x_2, ..., x_n\} \) to values that give the least value of \( F(\mathbf{x}) \).

In this paper we consider the developments in unconstrained optimization that have been published since 1967. The year 1967 was chosen because the books by Wilde and Beightler [1], Kowalik and Osborne [2] and Box, Davies and Swann [3], and also the review paper by Powell [4], together give a good survey of earlier work.

* This paper was presented at the 7th Mathematical Programming Symposium 1970, The Hague, The Netherlands.
Recent advances in unconstrained optimization

It has happened in the last three years that most published papers concern the case where the user has to program the calculation of $F(x)$ and its first derivative, $g(x)$ say, and this work has given the subject of gradient algorithms a nice structure, which forms the main part of this paper, beginning in section 4. In section 2 we treat algorithms that require no derivatives, and in section 3 we treat algorithms that require second derivatives.

The backbone of the structure of gradient algorithms is given in section 4, where we consider the very elegant theory due to Huang [5], which has developed from a number of recent papers. The original variable metric algorithm (Davidon [6]; Fletcher and Powell [7]) is the best known special case of Huang’s theory, and it is discussed in section 5, because some new properties of this method have been found recently. Then in section 6 some other special cases of Huang’s theory are considered, and one of these is especially interesting, because it may supersede the original variable metric method.

However the methods of section 6, like Davidon’s [6] algorithm, require an awkward subproblem to be solved many times, namely the minimization of a function of one variable. Therefore some methods have been proposed that avoid this subproblem, and two of these are described in section 7. A good practical alternative is to accept a very rough estimate of the solution of this subproblem, and this approach is considered in section 8.

A different general approach to gradient algorithms was proposed by Greenstadt [8], based on variational methods. It is outlined in section 9, and we note that it is relevant to a number of useful algorithms, including the new one proposed by Powell [9].

In the final section of this paper an opinion is offered on future research in unconstrained optimization.

2. Methods without derivatives

Many “direct search methods” for unconstrained optimization were proposed between 1960 and 1965, and they are described by Kowalik and Osborne [2] and by Box et al. [3]. However, in the last three years relatively few papers have been published on algorithms that do not require the calculation of any derivatives. The main subjects of these recent papers are (i) Estimating derivatives by numerical differences in