Algorithms and Application for Special Classes
of Nonlinear Least Squares Problems

Stephen J. Wright

A Nonlinear Least Squares problem is an optimization problem for which the objective function to be minimized has the form

$$F(x) = \sum_{i=1}^{m} f_i^2(x) \ , \quad x \in R^n \ , \quad m \geq n \ .$$

The structure of $F(x)$ allows the problem to be solved using Newton-type methods, in which the Hessian is efficiently approximated from the first derivatives, rather than explicitly calculated. In most of the work which has appeared in recent years, it has been assumed that the f_i's are twice continuously differentiable, that there are no explicit constraints on the variables, and that the Jacobian is small enough to be held in core and factorized. The first part of this **thesis** presents an overview of some of the algorithms which have been used on this general problem.

Special types of nonlinear least squares problems are then examined, in particular

(i)   problems for which there are general linear inequality constraints on x ,

(ii) problems in which the functions f_i may be nonsmooth

(that is have discontinuous first partial derivatives),
and

(iii)   problems for which the Jacobian matrix

$$J(x) = \left[\frac{\partial f_i}{\partial x_j}\right]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

is sparse;   that is, most of the f_i's are functions of
only a small subset of the variables.

Chapter 2 deals with the first of the above classes. Two algorithms based on a trust region implementation of the Levenberg-Marquardt method are presented. An active set strategy is used to deal with the constraints. At each iteration, estimates are made of the effect of relaxing each of the currently active constraints. At most one active constraint can be relaxed at each iteration, subject to an antizigzagging strategy. Convergence results are presented, and some test examples are discussed.

The case in which $F(x)$ is nonsmooth is dealt with in Chapter 3. We present a bundle algorithm based on the method of Mifflin [1] for general nonsmooth functions. The search direction is modified by taking account of the first-derivative curvature term of the Hessian, and a damping term is used in place of a line search parameter. Computational aspects of the algorithm are discussed, and convergence results (based on those of Mifflin) are given.

Chapter 4 discusses algorithms for the sparse problem. An inexact Levenberg-Marquardt method is proposed. This method makes use of the iterative Paige and Saunders algorithm LSQR [2], to solve the linearized subproblem which arises at each iteration. Convergence is demonstrated, and it is shown that under certain conditions, the method retains the quadratic convergence property of Levenberg-Marquardt in the zero-residual case. Computational results are given.

In the final chapter, we consider the geophysical problem of inverting seismic arrival time data, in order to accurately locate earthquake coordinates and to obtain density information about the earth's crust in

the vicinity of the recording stations. This can be formulated as a sparse nonlinear least squares problem, and is solved using a modification of the algorithm of Chapter 4. Numerical results are obtained for two actual data sets, and are compared with the results obtained from a program currently in use at the US Geological Survey.

The appendices contain listings of the FORTRAN programs which were used to implement the algorithms of Chapters 2, 3, and 4.

References

[1]   R. Mifflin, "An algorithm for constrained optimization with semismooth functions", *Math. Oper. Res.* **2** (1977), 191-207.

[2]   C.C. Paige and M.A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares", *ACM Trans. Math. Software* **8** (1982), 43-71.

Department of Mathematics,
University of Queensland,
St Lucia,
Queensland 4067,
Australia.