APPENDIX B

Objective Functions

When a new method for unconstrained minimization of a function of several variables is devised, it is customary for its author to give numerical results which are obtained by applying the method to a selection of well known objective functions. These functions have been constructed in such a way as to present various computational difficulties. The following is a list of some of the more frequently used test functions, together with an initial estimate $\mathbf{x}^{(0)}$ of the minimizer $\mathbf{x}^*$, and a brief description, where appropriate, of the computational difficulties presented by the optimization problem. Problems 1 and 8–12 are used to illustrate the behaviour of the methods for the least squares problem which is described in Chapter 7.

1. Rosenbrock (1960)

$$f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

$$\mathbf{x}^{(0)} = [-1.2, 1.0]^T \quad \mathbf{x}^* = [1, 1]^T \quad f(\mathbf{x}^*) = 0$$

This is a steep-sided sharply curving valley, the bottom of which follows the parabolic curve $x_2 = x_1^2$. A descent method must in general take very short steps in order to negotiate the sharp bend near $[0, 0]^T$ and must therefore provide new search directions $\mathbf{p}$ very frequently.

2. Wood's function, quoted by Colville (1968)

$$f(\mathbf{x}) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_3^2 - x_4)^2$$
$$+ 10.1\{(x_2 - 1)^2 + (x_4 - 1)^2\} + 19.8(x_2 - 1)(x_4 - 1)$$

$$\mathbf{x}^{(0)} = [-3, -1, -3, -1]^T \quad \mathbf{x}^* = [1, 1, 1, 1]^T \quad f(\mathbf{x}^*) = 0$$

This objective function has a saddle point near $[-1, 1, -1, 1]^T$ and the sequences $\{\mathbf{x}^{(k)}\}$ generated by most methods tend, at least initially, to stay near this point.

3. Miele and Cantrell (1969b)

$$f(\mathbf{x}) = (\exp(x_1) - x_2)^4 + 100(x_2 - x_3)^6 + \{\tan(x_3 - x_4)\}^4 + x_1^8$$

$$\mathbf{x}^{(0)} = [1, 2, 2, 2]^T \quad \mathbf{x}^* = [0, 1, 1, 1]^T \quad f(\mathbf{x}^*) = 0$$

This is a severely nonlinear function in all variables.

4. Powell (1962)

$$f(\mathbf{x}) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$

$$\mathbf{x}^{(0)} = [3, -1, 0, 1]^T \quad \mathbf{x}^* = [0, 0, 0, 0]^T \quad f(\mathbf{x}^*) = 0$$

For this problem the Hessian matrix of f at $\mathbf{x}^*$ is singular

5. Fletcher and Powell (1963)

$$f(\mathbf{x}) = 100\{(x_3 - 10\theta)^2 + (r - 1)^2\} + x_3^2$$

where

$$r = |(x_1^2 + x_2^2)^{1/2}|$$

and

$$\theta = \begin{cases} \dfrac{1}{2\pi}\tan^{-1}\left(\dfrac{x_2}{x_1}\right) & (x_1 > 0) \\[2ex] \dfrac{1}{2\pi}\tan^{-1}\left(\dfrac{x_2}{x_1}\right) + \dfrac{1}{2} & (x_1 < 0) \end{cases}$$

$$\mathbf{x}^{(0)} = [-1, 0, 0]^T \quad \mathbf{x}^* = [1, 0, 0]^T \quad f(\mathbf{x}^*) = 0$$

This is a steep-sided valley which follows a helical path. It presents the same kind of difficulties as Problem 1.

6. Dixon (1973b)

$$f(\mathbf{x}) = (1 - x_1)^2 + (1 - x_{10})^2 + \sum_{i=1}^{9} (x_i^2 - x_{i+1})^2$$

$$\mathbf{x}^{(0)} = [-2, \ldots, -2]^T \quad \mathbf{x}^* = [1, \ldots, 1]^T \quad f(\mathbf{x}^*) = 0$$

This objective function has 10 variables and provides some indication of the behaviour of optimization methods as the number of variables increases.

7. Powell (1966)

$$f(\mathbf{x}) = x_1^4 + x_1x_2 + (1 + x_2)^2$$

$$\mathbf{x}^{(0)} = [0, 0]^T$$

For this problem,

$$\min_{\alpha} f(\mathbf{x}^{(0)} - \alpha\mathbf{G}^{(0)-1}\mathbf{g}^{(0)}) = f(\mathbf{x}^{(0)})$$

so that no progress can be made beyond $\mathbf{x}^{(0)}$ by using Newton's method together with an exact line search.

8. Hartley (1961)

$$S(\mathbf{x}) = \sum_{i=1}^{m} \{Y(\mathbf{t}_i; \mathbf{x}) - y_i\}^2$$

where

$$Y(\mathbf{t}_i; \mathbf{x}) = x_1 + x_2 \exp(t_i x_3)$$

and the t_i and the y_i are given in the following table.

i	t_i	y_i
1	-5	127
2	-3	151
3	1	421
4	3	460
5	5	426
6	-1	379

For this problem, $m = 6, n = 3$ and

$$\mathbf{x}^{(0)} = [580, -180, -0.16]^T \quad S^{(0)} \approx 0.3 \times 10^5$$

$$\mathbf{x}^* = [523.3, -157.0, -0.1994]^T \quad S^* \approx 0.13 \times 10^5$$

This is a least squares problem in which S^* is non-zero.

9. Meyer and Roth (1972)

$$S(\mathbf{x}) = \sum_{i=1}^{m} \{Y(\mathbf{t}_i;\mathbf{x}) - y_i\}^2$$

in which

$$Y(\mathbf{t}_i;\mathbf{x}) = \frac{x_1 x_3 t_{1i}}{(1 + x_1 t_{1i} + x_2 t_{2i})}$$

and the $\mathbf{t}_i$ and the y_i are given in the following table.

i	t_{1i}	t_{2i}	y_i
1	1.0	1.0	0.126
2	2.0	1.0	0.219
3	1.0	2.0	0.076
4	2.0	2.0	0.126
5	0.1	0.0	0.186

For this problem, $m = 5, n = 3$ and

$$\mathbf{x}^{(0)} = [10.39, 48.83, 0.74]^T \quad S^{(0)} \approx 0.365 \times 10^{-1}$$

$$\mathbf{x}^* = [3.13, 15.16, 0.78]^T \quad S^* \approx 0.4 \times 10^{-4}$$

This least squares problem was originally presented by Box and Hunter (1965). As in Problem 8, S^* is non-zero.

10. Meyer and Roth (1972)

$$S(\mathbf{x}) = \sum_{i=1}^{m} \{Y(t_i;\mathbf{x}) - y_i\}^2$$

where

$$Y(t_i;\mathbf{x}) = x_1 \exp\left\{\frac{x_2}{(t_i + x_3)}\right\}$$

and the t_i and the y_i are given in the following table.

i	t_i	y_i
1	50	34 780
2	55	28 610
3	60	23 650
4	65	19 630
5	70	16 370
6	75	13 720
7	80	11 540
8	85	9 744
9	90	8 261
10	95	7 030
11	100	6 005
12	105	5 147
13	110	4 427
14	115	3 820
15	120	3 307
16	125	2 872

For this problem, $m = 16, n = 3$ and

$$\mathbf{x}^{(0)} = [0.02, 4000, 250]^T \quad S^{(0)} \approx 0.17 \times 10^{10}$$

$$\mathbf{x}^* = [0.0056, 6181.4, 345.2]^T \quad S^* \approx 88$$

The data for this least squares problem give the resistance of a thermistor as a function of temperature. The problem defeats many least squares fitting procedures.

11. Pereyra (1967)

$$S(\mathbf{x}) = \sum_{i=1}^{m} \{Y(\mathbf{t}_i;\mathbf{x}) - y_i\}^2$$

where

$$Y(\mathbf{t}_i;\mathbf{x}) = x_2 \sin(t_i x_1) + x_3$$

The data are obtained from

$$y_i = \sin t_i \quad (i = 1, \ldots, 13)$$

where the values of the t_i are given in the following table.

i	t_i	i	t_i
1	0.105	7	1.1
2	0.25	8	1.25
3	0.4	9	1.35
4	0.55	10	1.45
5	0.7	11	1.55
6	0.9	12	1.57
		13	1.6

For this problem, $m = 13, n = 3$, and

$$\mathbf{x}^{(0)} = [0.9, 0.9, 0.1]^T \qquad S^{(0)} \approx 0.16 \times 10^{-1}$$

$$\mathbf{x}^* = [1.0, 1.0, 0.0]^T \qquad S^* = 0.0$$

12. Meyer and Roth (1972)

$$S(\mathbf{x}) = \sum_{i=1}^{m} \{Y(\mathbf{t}_i;\mathbf{x}) - y_i\}^2$$

where

$$Y(\mathbf{t}_i;\mathbf{x}) = x_3\{\exp(-x_1 t_{1i}) + \exp(-x_2 t_{2i})\}$$

The values of y_i are computed from $Y(\mathbf{t}_i;\mathbf{x}^*)$ with the values of t_{1i} and t_{2i} given in the following table.

i	t_{1i}	t_{2i}
1	0.0	0.0
2	0.6	0.4
3	0.6	1.0
4	1.4	1.4
5	2.6	1.4
6	3.2	1.6
7	0.8	2.0
8	1.6	2.2
9	2.6	2.2
10	4.0	2.2
11	1.2	2.6
12	2.0	2.6
13	4.6	2.8
14	3.2	3.0
15	1.6	3.2
16	4.2	3.4
17	2.0	3.8
18	3.2	3.8
19	2.8	4.2
20	4.2	4.2
21	5.4	4.4
22	5.6	4.8
23	3.2	5.0

For this problem, $m = 23, n = 3$ and

$$\mathbf{x}^{(0)} = [12.0, 1.0, 25.0]^T \qquad S^{(0)} \approx 216.0$$

$$\mathbf{x}^* = [14.3, 1.5, 20.1]^T \qquad S^* = 0.0$$

The double exponential formula is notoriously difficult to fit to a set of data.

APPENDIX C

The Cholesky Decomposition of a Symmetric Positive Definite Matrix

In this appendix we prove Theorem 3.4.1 which for convenience is restated.

Theorem 3.4.1

If $\mathbf{A}$ is an $n \times n$ symmetric positive definite matrix, then there exists a lower triangular matrix $\mathbf{L}$ such that $\mathbf{A} = \mathbf{L}\mathbf{L}^T$.

Proof

The theorem is trivially true for $n = 1$ because a 1×1 positive definite matrix is equivalent to a positive real number, which certainly has a square root.

Let $\mathbf{A}_i$ be the $i \times i$ principal minor matrix of $\mathbf{A}$ (i.e. the matrix formed from the first i rows and the first i columns of $\mathbf{A}$). Since $\mathbf{A}$ is positive definite, $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ $(\forall \mathbf{x} \neq \mathbf{0})$. In particular, $\mathbf{x}_i^T\mathbf{A}\mathbf{x}_i > 0$, where $\mathbf{x}_i$ is given by

$$\mathbf{x}_i = [x_{1i}, \ldots, x_{ii}, 0, \ldots, 0]^T$$

in which the first i elements x_{ji} $(j = 1, \ldots, i)$ of $\mathbf{x}_i$ are non-zero.

Now $\mathbf{A}$ may be partitioned according to

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}_i & \mathbf{P}_i \\ \hline \mathbf{Q}_i & \mathbf{R}_i \end{array}\right]$$

where $\mathbf{P}_i$, $\mathbf{Q}_i$, and $\mathbf{R}_i$ are $i \times (n-i)$, $(n-i) \times i$, and $(n-i) \times (n-i)$ matrices respectively, and $1 \leqslant i \leqslant n-1$. So

$$\mathbf{x}_i^T\mathbf{A}\mathbf{x}_i = \mathbf{y}_i^T\mathbf{A}_i\mathbf{y}_i$$

where

$$\mathbf{y}_i = [x_{1i}, \ldots, x_{ii}]^T$$

Hence

$$\mathbf{y}_i^T\mathbf{A}_i\mathbf{y}_i > 0 \quad (\forall \mathbf{y}_i \neq \mathbf{0})$$

and so $\mathbf{A}_i$ is positive definite. Clearly this holds for $i = 1, \ldots, n$ so all principal minor matrices $\mathbf{A}_i$ of a positive definite matrix $\mathbf{A}$ are themselves positive definite.

Suppose that there is a lower triangular matrix $\mathbf{L}_{i-1}$ such that

$$\mathbf{A}_{i-1} = \mathbf{L}_{i-1}\mathbf{L}_{i-1}^T \tag{C.1}$$

for some $i \geqslant 2$. We shall show that this implies the existence of a lower triangular matrix $\mathbf{L}_i$ such that

$$\mathbf{A}_i = \mathbf{L}_i\mathbf{L}_i^T \tag{C.2}$$

A Criterion for the Convergence of a Sequence in a Compact Set of $\{R^n, \|\cdot\|\}$

In the proof of Theorem 7.7.2 it is stated that if the sequence $\{\mathbf{x}^{(k)}\}$ has only one limit point then we are entitled to assert that $\{\mathbf{x}^{(k)}\}$ is convergent. This is a consequence of a theorem which, it would appear, is seldom proved in textbooks of analysis, although often quoted in various forms.

Theorem J.1

If 1. $\{R^n, \|\cdot\|\}$ is a normed linear space;
2. $S \subset R^n$ is compact;
3. $\{\mathbf{x}^{(k)}\}$ is a sequence in S;
4. $\mathbf{x}^*$ is the unique limit point of $\{\mathbf{x}^{(k)}\}$,

then $\{\mathbf{x}^{(k)}\}$ converges to $\mathbf{x}^*$ as $k \to \infty$.

Proof

Suppose that the sequence $\{\mathbf{x}^{(k)}\}$ does not converge to $\mathbf{x}^*$. Then there is a subsequence $\{\mathbf{x}^{(k_i)}\}$ of $\{\mathbf{x}^{(k)}\}$ and a real number $\epsilon > 0$ such that

$$\|\mathbf{x}^{(k_i)} - \mathbf{x}^*\| > \epsilon \ (\forall i)$$

Therefore *no* subsequence of $\{\mathbf{x}^{(k_i)}\}$ converges to $\mathbf{x}^*$. Therefore by Hypothesis 4, no subsequence of $\{\mathbf{x}^{(k_i)}\}$ converges. But by Hypothesis 3, $\{\mathbf{x}^{(k_i)}\}$ lies in S, and so by Hypothesis 2 and Definition 1.3.6, $\{\mathbf{x}^{(k_i)}\}$ must contain a subsequence $\{\mathbf{x}^{(k_j)}\}$ which is convergent in S. Also, since $\{\mathbf{x}^{(k_j)}\}$ is a subsequence of $\{\mathbf{x}^{(k)}\}$, then $\{\mathbf{x}^{(k_j)}\}$ must converge to $\mathbf{x}^*$ by Hypothesis 4. This contradicts the fact that no subsequence of $\{\mathbf{x}^{(k_i)}\}$ converges to $\mathbf{x}^*$. Therefore $\{\mathbf{x}^{(k)}\}$ converges to $\mathbf{x}^*$. $\square$

For a far more sophisticated treatment of the convergence of descent methods and, in particular, of the ideas presented in Section 7.7 the reader should consult Chapter 14 of the book by Ortega and Rheinboldt (1970).

References

Bard, Y. (1968) On a Numerical Instability of Davidon-like Methods, *Mathematics of Computation* **22**, 665–666.

Bard, Y. (1970) Comparison of Gradient Methods for the Solution of Nonlinear Parameter Estimation Problems, *S.I.A.M. Journal of Numerical Analysis* 7, 157–186.

Bard, Y. (1974) *Nonlinear Parameter Estimation*, Academic Press.

Biggs, M.C. (1970) Computational Experience with Variable Metric Methods for Function Minimization *The Hatfield Polytechnic, Numerical Optimization Centre, T.R.* 7.

Box, G.E.P. and Hunter, W.G. (1965) The Experimental Study of Physical Mechanisms, *Technometrics* 7, 23–42.

Bracken, J. and McCormick, G.P. (1968) *Selected Applications of Nonlinear Programming*, John Wiley and Sons Inc.

Brent, R.P. (1973) *Algorithms for Minimization Without Derivatives*, Prentice-Hall Inc.

Brown, K.M. and Dennis, J.E. (1972) Derivative Free Analogues of the Levenberg-Marquardt and Gauss Algorithms for Nonlinear Least Squares Approximation, *Numerische Mathematik* **18**, 289–297.

Broyden, C.G. (1967) Quasi-Newton Methods and their Application to Function Minimization, *Mathematics of Computation* **21**, 368–381.

Broyden, C.G. (1970) The Convergence of a Class of Double-Rank Minimization Algorithms: 1. General Considerations, *Journal of the Institute of Mathematics and its Applications* **6**, 76–90.

Burkill, J.C. and Burkill, H. (1970) *A Second Course in Mathematical Analysis*, Cambridge University Press.

Burley, D.M. (1974) *Studies in Optimization*, International Textbook Company Ltd.

Businger, P. and Golub, G. (1965) Linear Least Squares Solutions by Householder Transformations, *Numerische Mathematik* 7, 269–276.

Colville, A.R. (1968) A Comparative Study of Non-Linear Programming Codes, *IBM New York Scientific Centre T.R.* 320–2925.

Davidon, W.C. (1959) Variable Metric Method for Minimization, *Research and Development Report ANL*-5990, US Atomic Energy Commission, Argonne National Laboratories.

Davidon, W.C. (1968) Variance Algorithm for Minimization, *The Computer Journal* **10**, 406–410.

Davies, D., Swann, W.H. and Campey, I.G. (1964) Report on the Development of a New Direct Search Method of Optimization, *ICI Ltd, Central Instrument Laboratory Research Note* 64/3. (Author Swann, W.H.).

Dixon, L.C.W. and Biggs, M.C. (1970) MEANDER – A Newton Based Procedure for n-Dimensional Function Minimization, *The Hatfield Polytechnic, Numerical Optimization Centre, T.R.* **9**.

Dixon, L.C.W. (1971a) Variable Metric Algorithms: Necessary and Sufficient Conditions for Identical Behaviour on Non-quadratic Functions, *The Hatfield Polytechnic, Numerical Optimization Centre, T.R.* **26**.

Dixon, L.C.W. (1971b) The Choice of Steplength a Crucial Factor in the Performance of Variable Metric Algorithms, *The Hatfield Polytechnic, Numerical Optimization Centre, T.R.* **27**.

Dixon, L.C.W. (1972) Conjugate Gradient Algorithms: Quadratic Termination Without Linear Searches, *The Hatfield Polytechnic, Numerical Optimization Centre, T.R.* **38**.

Dixon, L.C.W. (1973a) Nonlinear Optimization: A Survey of the State of the Art, *The Hatfield Polytechnic, Numerical Optimization Centre, T.R.* **42**.

Dixon, L.C.W. (1973b) Conjugate Directions without Linear Searches, *Journal of the Institute of Mathematics and its Applications* **11**, 317–328.

Dixon, L.C.W. (1976) *Optimization in Action*, Academic Press Inc.

Dennis, J.E. and Moré, J.J. (1977) Quasi-Newton Methods, Motivation and Theory, *S.I.A.M. Review* **19**, 46–89.

Fiacco, A.V. and McCormick, G.P. (1968) *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley and Sons Inc.

Flanagan, P.D., Vitale, P.A. and Mendelsohn, J. (1969) A Numerical Investigation of Several One-dimensional Search Procedures in Nonlinear Regression Problems, *Technometrics* **11**, 265–284.

Fletcher, R. and Powell, M.J.D. (1963) A Rapidly Convergent Descent Method for Minimization, *The Computer Journal* **6**, 163–168.

Fletcher, R. and Reeves, C.M. (1964) Function Minimization by Conjugate Gradients, *The Computer Journal* 7, 149–153.

Fletcher, R. (1970) A New Approach to Variable Metric Algorithms, *The Computer Journal* **13**, 317–322.

Fletcher, R. (1971a) A Survey of Algorithms for Unconstrained Optimization, *United Kingdom Atomic Energy Authority Research Group Report T.P.* 456.

Fletcher, R. (1971b) A Modified Marquardt Subroutine for Non-linear Least Squares, *United Kingdom Atomic Energy Authority Research Group Report R*-6799.

Fletcher, R. (1972a) A Fortran Subroutine for Minimization by the Method of Conjugate Gradients, *United Kingdom Atomic Energy Authority Research Group Report R*-7073.

Fletcher, R. (1972b) Fortran Subroutines for Minimization by Quasi-Newton Methods, *United Kingdom Atomic Energy Authority Research Group Report R*-7125.

Fletcher, R. and Freeman, T.L. (1975) A Modified Newton Method for Minimization, *University of Dundee, Department of Mathematics, Numerical Analysis Report* 7.

Forsythe, G.E. and Moler, C.B. (1967) *Computer Solution of Linear Algebraic Systems*, Prentice-Hall Inc.

Franklin, J.N. (1968) *Matrix Theory*, Prentice-Hall Inc.

Gill, P.E. and Murray, W. (1971) Quasi-Newton Methods for Unconstrained Optimization, *National Physical Laboratory Report Maths* 97.

Gill, P.E. and Murray, W. (1972a) Two Methods for the Solution of Linearly Constrained and Unconstrained Optimization Problems, *National Physical Laboratory Report NAC* 25.

Gill, P.E. and Murray, W. (1972b) Quasi-Newton Methods for Unconstrained Optimization, *Journal of the Institute of Mathematics and its Applications* 9, 91–108.

Gill, P.E. and Murray, W. (1974a) Safeguarded Steplength Algorithms for Optimization Using Descent Methods, *National Physical Laboratory Report NAC* 37.

Gill, P.E. and Murray, W. (1974b) Newton-type Methods for Unconstrained and Linearly Constrained Optimization, *Mathematical Programming* 7, 311–350.

Gill, P.E. and Murray, W. (1976) Algorithms for the Solution of the Non-linear Least-squares Problem, *National Physical Laboratory Report NAC* 71.

Gill, P.E., Murray, W. and Picken, S.M. (1972) The Implementation of Two Modified Newton Algorithms for Unconstrained Optimization, *National Physical Laboratory Report NAC* 24.

Gill, P.E., Murray, W. and Pitfield, R.A. (1972) The Implementation of Two Revised Quasi-Newton Algorithms for Unconstrained Optimization, *National Physical Laboratory Report NAC* 11.

Goldfold, S.M., Quandt, R.E. and Trotter, H.F. (1966) Maximization by Quadratic Hill-climbing, *Econometrics* **34**, 541–551.

Goldstein, A.A. (1962) Cauchy's Method of Minimization, *Numerische Mathematik* **4**, 146–150.

Goldstein, A.A. (1965) On Steepest Descent, *S.I.A.M. Journal of Control, Series A* **3**, 147–151.

Goldstein, A.A. and Price, J.F. (1967) An Effective Algorithm for Minimization, *Numerische Mathematik* **10**, 184–189.

Golub, G. (1965) Numerical Methods for Solving Linear Least Squares Problems, *Numerische Mathematik* 7, 206–216.

Greenstadt, J. (1967) On the Relative Efficiencies of Gradient Methods, *Mathematics of Computation* **21**, 360–367.

Hadley, G. (1969) *Linear Algebra*, Addison-Wesley Publishing Company Inc.

Hartley, H.O. (1961) The Modified Gauss-Newton Method for the Fitting of Nonlinear Regression Functions by Least Squares, *Technometrics* **3**, 269–280.

Hebden, M.D. (1973) An Algorithm for Minimization Using Exact Second Derivatives, *United Kingdom Atomic Energy Authority Research Group Report T.P.* 515.

Hestenes, M.R. and Steifel, E. (1952) Methods of Conjugate Gradients for Solving Linear Systems, *Journal of Research of the National Bureau of Standards* **49**, 409–436.

Himmelblau, D.M. (1972) *Applied Nonlinear Programming*, McGraw-Hill Book Company.

Hooke, R. and Jeeves, T.A. (1961) Direct Search Solution of Numerical and Statistical Problems, *Journal of the Association for Computing Machinery* **8**, 212–229.

Huang, H.Y. (1970) Unified Approach to Quadratically Convergent Algorithms for Function Minimization, *Journal of Optimization Theory and Applications* **5**, 405–423.

Huang, H.Y. and Levy, A.V. (1970) Numerical Experiments on Quadratically Convergent Algorithms for Function Minimization, *Journal of Optimization Theory and Applications* **6**, 269–282.

Kowalik, J. and Osborne, M.R. (1968) *Methods for Unconstrained Optimization Problems*, American Elsevier Publishing Company Inc.

Lawson, C.L. and Hanson, R.J. (1974) *Solving Least Squares Problems*, Prentice-Hall Inc.

Levenberg, K. (1944) A Method for the Solution of Certain Non-linear Problems in Least Squares, *Quarterly Journal of Applied Mathematics* **2**, 164–168.

Mangasarian, O. (1969) *Nonlinear Programming*, McGraw-Hill Book Company.

Marquardt, D.W. (1963) An algorithm for Least-squares Estimation of Nonlinear Parameters, *Journal of the Society for Industrial and Applied Mathematics* **11**, 431–441.

McCormick, G.P. (1969) *See* Pearson, J.D. (1969).

Meyer, R.R. (1970) Theoretical and Computational Aspects of Nonlinear Regression, *Technical Paper P*-1819, *Shell Development Company, Emeryville, California.*

Meyer, R.R. and Roth, P.M. (1972) Modified Damped Least Squares: An Algorithm for Non-linear Estimation, *Journal of the Institute of Mathematics and its Applications* 9, 218–233.

Miele, A. and Cantrell, J.W. (1969a) Study on a Memory Gradient Method for the Minimization of Functions, *Journal of Optimization Theory and Applications* **3**, 459–470.

Miele, A. and Cantrell, J.W. (1969b) Memory Gradient Method for the Minimization of Functions, *Symposium on Optimization, Nice.*

Murray, W. (1972) Second Derivative Methods, in *Numerical Methods for Unconstrained Optimization* (Murray, W., Ed.), Academic Press Inc.

Murtagh, B.A. and Sargent, R.W.H. (1969) A Constrained Minimization Method with Quadratic Convergence, in *Optimization* (Fletcher, R., Ed.), Academic Press Inc.

Nash, J.C. (1977) Minimizing a Non-linear Sum of Squares Function on a Small Computer, *Journal of the Institute of Mathematics and its Applications* **19**, 231–237.

Nelder, J.A. and Mead, R. (1965) A Simplex Method for Function Minimization, *The Computer Journal* 7, 308–313.

Noble, B. (1969) *Applied Linear Algebra*, Prentice-Hall Inc.

Ortega, J.M. and Rheinboldt, W.C. (1970) *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press Inc.

Parkinson, J.M. and Hutchinson, D. (1972) An Investigation into the Efficiency of Variants on the Simplex Method, in *Numerical Methods for Non-linear Optimization* (Lootsma, F.A., Ed.), Academic Press Inc.

Pearson, J.D. (1969) Variable Metric Methods of Minimization, *The Computer Journal* **12**, 171–178.

Peckham, G. (1970) A New Method for Minimizing a Sum of Squares Without Calculating Gradients, *The Computer Journal* **13**, 418–420.

Pereyra, V. (1967) Iterative Methods for Solving Nonlinear Least Squares Problems, *S.I.A.M. Journal of Numerical Analysis* **4**, 27–36.

Phillips, G.M. and Taylor, P. (1973) *Theory and Applications of Numerical Analysis*, Academic Press Inc.

Polak, E. and Ribiere, G. (1969) Note sur la Convergence de Methodes de Directions Conjugees, *Revue Francaise Inform. Rech. Operation.* **16-R1**, 35–43.

Powell, M.J.D. (1962) An Iterative Method for Finding Stationary Values of a Function of Several Variables, *The Computer Journal* **5**, 147–151.

Powell, M.J.D. (1964) An Efficient Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivatives, *The Computer Journal* 7, 155–162.

Powell, M.J.D. (1965) A Method for Minimizing a Sum of Squares of Non-linear Functions Without Calculating Derivatives, *The Computer Journal* 7, 303–307.

Powell, M.J.D. (1966) Minimization of Functions of Several Variables, in *Numerical Analysis* (Walsh, J., Ed.), Academic Press Inc.
Powell, M.J.D. (1969) Rank One Methods for Unconstrained Optimization, *United Kingdom Atomic Energy Authority Research Group Report T.P.* 372.
Powell, M.J.D. (1971) Recent Advances in Unconstrained Optimization, *Mathematical Programming* **1**, 26–57.
Powell, M.J.D. (1974) Introduction to Constrained Optimization, in *Numerical Methods for Constrained Optimization* (Gill, P.E. and Murray, W., Eds.), Academic Press Inc.
Rhead, D.G. (1971) Some Numerical Experiments on Zangwill's Method for Unconstrained Minimization, *Working Paper ICSI* 319, *University of London.*
Rosenbrock, H.H. (1960) An Automatic Method for Finding the Greatest or Least Value of a Function, *The Computer Journal* **3**, 175–184.
Russell, D.L. (1970) *Optimization Theory*, W.A. Benjamin Inc.
Schubert, L.K. (1970) Modification of a Quasi-Newton Method for Nonlinear Equations with a Sparse Jacobian, *Mathematics of Computation* **24**, 27–30.
Shanno, D.F. (1970a) Conditioning of Quasi-Newton Methods for Function Minimization, *Mathematics of Computation* **24**, 647–656.
Shanno, D.F. (1970b) An Accelerated Gradient Projection Method for Linearly Constrained Nonlinear Estimation, *S.I.A.M. Journal of Applied Mathematics* **18**, 322–334.
Sorenson, H.W. (1969) Comparison of Some Conjugate Direction Procedures for Function Minimization, *Journal of the Franklin Institute* **288**, 421–441.
Smith Jr., F.B. and Shanno, D.F. (1971) An Improved Marquardt Procedure for Non-linear Regressions, *Technometrics* **13**, 63–74.
Spendley, W. (1969) Nonlinear Least Squares Fitting using a Modified Simplex Minimization Method, in *Optimization* (Fletcher, R., Ed.), Academic Press Inc.
Spendley, W., Hext, G.R. and Himsworth, F.R. (1962) Sequential Applications of Simplex Designs in Optimization and Evolutionary Operation, *Technometrics* **4**, 441–461.
Stewart, G.W. (1967) A Modification of Davidson's Minimization Method to Accept Difference Approximations to Derivatives, *Journal of the Association for Computing Machinery* **14**, 72–83.
Topkis, D.M. and Veinott, A.F. (1967) On the Convergence of Some Feasible Direction Algorithms for Nonlinear Programming, *S.I.A.M. Journal of Control* **5**, 268–279.
Wilkinson, J.H. (1963) *Rounding Errors in Algebraic Processes, HMSO*.
Wilkinson, J.H. (1965) *The Algebraic Eigenvalue Problem*, Oxford University Press.
Wolfe, P. (1969) Convergence Conditions for Ascent Methods, *S.I.A.M. Review* **11**, 226–235.
Wolfe, P. (1971) Convergence Conditions for Ascent Methods II: Some Corrections, *S.I.A.M. Review* **13**, 185–188.
Zangwill, W.I. (1967) Minimizing a Function Without Calculating Derivatives, *The Computer Journal* **10**, 293–296.
Zangwill, W.I. (1969) *Nonlinear Programming: A Unified Approach*, Prentice-Hall Inc.

Index