MODERN ANALYTIC AND COMPUTATIONAL METHODS IN SCIENCE AND MATHEMATICS

A Group of Monographs and Advanced Textbooks

Editor: RICHARD BELLMAN, University of Southern California

Already Published:

1. R. E. Bellman, R. E. Kalaba, and Marcia C. Prestrud, Invariant Imbedding and Radiative Transfer in Slabs of Finite Thickness, 1963
2. R. E. Bellman, Harriet H. Kagiwada, R. E. Kalaba, and Marcia C. Prestrud, Invariant Imbedding and Time-Dependent Transport Processes, 1964
3. R. E. Bellman and R. E. Kalaba, Quasilinearization and Nonlinear Boundary-Value Problems, 1965
4. R. E. Bellman, R. E. Kalaba, and Jo Ann Lockett, Numerical Inversion of the Laplace Transform: Applications to Biology, Economics, Engineering, and Physics, 1966
5. S. G. Mikhlin and K. L. Smolitskiy, Approximate Methods for Solution of Differential and Integral Equations, 1967
6. R. N. Adams and E. D. Denman, Wave Propagation and Turbulent Media, 1966
7. R. L. Stratonovich, Conditional Markov Processes and Their Application to the Theory of Optimal Control, 1968
8. A. G. Ivakhnenko and V. G. Lapa, Cybernetics and Forecasting Techniques, 1967
9. G. A. Chebotarev, Analytical and Numerical Methods of Celestial Mechanics, 1967
10. S. F. Feshchenko, N. I. Shkil', and L. D. Nikolenko, Asymptotic Methods in the Theory of Linear Differential Equations, 1967
12. J. Kowalik and M. R. Osborne, Methods for Unconstrained Optimization Problems, 1968
13. R. E. Larson, State Increment Dynamic Programming, 1968

In Preparation:

11. A. G. Butkovskii, Optimal Control Theory for Distributed Parameters
14. Sidney J. Yakowitz, Mathematics of Adaptive Control Processes
15. S. K. Srinivasan, Stochastic Theory and Cascade Processes
16. D. U. von Rosenberg, Methods for the Numerical Solution of Partial Differentiation Equations

R. B. Banerji, Theory of Problem Solving: An Approach to Artificial Intelligence

R. Lattès and J.-L. Lions, The Method of Quasi-Reversibility: Applications to Partial Differential Equations

MODERN ANALYTIC AND COMPUTATIONAL METHODS IN SCIENCE AND MATHEMATICS

METHODES MODERNES D'ANALYSE ET DE COMPUTATION EN SCIENCE ET MATHÉMATIQUE

NEUE ANALYTISCHE UND NUMERISCHE METHODEN IN DER WISSENSCHAFT UND DER MATHEMATIK

НОВЫЕ АНАЛИТИЧЕСКИЕ И БЫЧИСЛИТЕЛЬНЫЕ МЕТЕОЛЫ В НАУКЕ И МАТЕМАТИКЕ

Editor

RICHARD BELLMAN, UNIVERSITY OF SOUTHERN CALIFORNIA

Methods for Unconstrained Optimization Problems

by

J. KOWALIK
Computing Department
The Boeing Company, Renton, Washington

AND

M. R. OSBORNE
Computer Centre
The Australian National University
Canberra, A.C.T., Australia

American Elsevier Publishing Company, Inc.
New York 1968

AMERICAN ELSEVIER PUBLISHING COMPANY, INC.
52 Vanderbilt Avenue
New York, N.Y. 10017

ELSEVIER PUBLISHING CO. LTD.
Barking, Essex, England

ELSEVIER PUBLISHING COMPANY
335 Jan Van Galenstraat
P.O. Box 211, Amsterdam, The Netherlands

Standard Book Number 444-00041-0

Library of Congress Card Number 68-24805

PRINTED IN THE UNITED STATES OF AMERICA

Chapter 6

NUMERICAL RESULTS

6.1. INTRODUCTION

In this chapter we present the results of numerical experiments made using a selection from among the methods described in the preceding chapters. The plan of this chapter is as follows. We summarize the test problems in Section 2, consider some questions relating to the problem of scaling in Section 3, and describe the techniques used for minimization along a line in Section 4. The major comparisons are broken up in the same way as the chapters in which the relevant methods are described. Thus we consider direct search methods in Section 5, conjugate gradient methods in Section 6, and methods for minimizing sums of squares in Section 7. In Section 8 we illustrate the use of several of the minimization techniques for solving the sequence of unconstrained problems generated by the SUMT transformation.

In tabulating the results we have generally estimated the work required by giving the number of function calls. In methods which use derivatives it should be noted that the function value and the partial derivatives with respect to each independent variable are evaluated at each function call, and in several published comparisons of available techniques (for example Box [1]) what is given is the number of function calls multiplied by $(n + 1)$ where n is the number of independent variables. We have not followed this practice, and we suggest that the reader weigh the number of function calls using his knowledge of the function involved. We feel this is preferable to assuming that the calculation of each partial derivative can be equated with the function calculation.

The calculations have been carried out in single precision arithmetic on an IBM 360/50 computer. The programs are written in FORTRAN and PL/1.

6.2. THE SAMPLE PROBLEMS

Here we summarize briefly the problems on which the optimization techniques have been tested. Most have been taken from the literature but the Enzyme and Watson functions are new. These turn out to be somewhat more difficult than the other problems (a version of the Watson problem with 9 variables proving particularly difficult), and could prove useful additions to the comparatively limited number of test problems currently available.

(i) *Rosenbrock* [2]:

$$f = 100(x_2 - x_1^2)^2 + (1 - x_1)^2. \tag{6.2.1}$$

This is perhaps the most widely used test function. Here f has a minimum of 0 at $\mathbf{x} = (1, 1)$. There is a steep valley along the parabola $x_2 = x_1^2$.

(ii) *Cube* [3]:

$$f = 100(x_2 - x_1^3)^2 + (1 - x_1)^2. \tag{6.2.2}$$

This is a variant on the Rosenbrock function, and has similar properties.

(iii) *Beale* [4]:

$$f = \sum_{i=1}^{3} (C_i - x_1(1 - x_2^i))^2, \tag{6.2.3}$$

where $C_1 = 1.5$, $C_2 = 2.25$, $C_3 = 2.625$. This function has a narrow curving valley approaching $x_2 = 1$, and has minimum 0 at $\mathbf{x} = (3, .5)$.

(iv) *Box* [1]:

$$f = \sum_{y} \{(e^{-x_1 y} - e^{-x_2 y}) - x_3(e^{-y} - e^{-10y})\}^2, \tag{6.2.4}$$

where $y = 0.1(.1)1$. Box has described this function and the two-dimensional one obtained by setting $x_3 = 1$. In the latter case a contour plot shows a strongly asymmetric curved valley. Here the desired minimum is $f = 0$ at $\mathbf{x} = (1, 10, 1)$.

(v) *Enzyme* [5]:

$$f = \sum_{i=1}^{11} \left(V_i - \frac{x_1(y_i^2 + x_2 y_i)}{y_i^2 + x_3 y_i + x_4}\right)^2. \tag{6.2.5}$$

Table 6.2.1 gives the values of V_i, y_i, $x = 1, 2, \ldots, 11$. The minimum is $f = 3.075 \times 10^{-4}$ at $\mathbf{x} = (.1928, .1916, .1234, .1362)$. The origin of this problem has been described in detail by Kowalik and Morrison [6].

(vi) *Watson*:

$$f = \sum_{i=1}^{30} \left\{ \sum_{j=1}^{m} (j - 1) x_j y_i^{j-2} - \left(\sum_{j=1}^{m} x_j y_i^{j-1} \right)^2 - 1 \right\}^2 + x_1^2, \tag{6.2.6}$$

where $y_i = (i - 1)/29$. In this problem an attempt is made to approximate to the solution of the differential equation

$$\frac{dz}{dx} - z^2 = 1, \qquad z(0) = 0, \tag{6.2.7}$$

Table 6.2.1

i	V_i	y_i
1	.1957	4
2	.1947	2
3	.1735	1
4	.1600	.5
5	.0844	.25
6	.0627	.167
7	.0456	.125
8	.0342	.1
9	.0323	.0823
10	.0235	.0714
11	.0246	.0625

in $0 \leqslant x \leqslant 1$ by a polynomial of degree m by minimizing the sum of squares of the residuals at selected points. For the points indicated, the solution for $m = 6$ is $f = 2.288 \times 10^{-3}$ at $\mathbf{x} = (-.016, 1.012, -.233, 1.260, -1.513, .993)$. This is not an example for which polynomial approximation is suitable, so that the comparative difficulty of this problem could have been anticipated.

(vii) *Rosen-Suzuki* [7]:

$$f = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4 \tag{6.2.8}$$

subject to

$$-x_1^2 - x_2^2 - x_3^2 - x_4^2 - x_1 + x_2 - x_3 + x_4 + 8 \geqslant 0, \tag{6.2.9}$$

$$-x_1^2 - 2x_2^2 - x_3^2 - 2x_4^2 + x_1 + x_4 + 10 \geqslant 0, \tag{6.2.10}$$

$$-2x_1^2 - x_2^2 - x_3^2 - 2x_1 + x_2 + x_4 + 5 \geqslant 0. \tag{6.2.11}$$

This function has a minimum $f = -44$ at $\mathbf{x} = (0, 1, 2, -1)$. The constraint (6.2.11) is active.

(viii) *Beale* [8]:

$$f = 9 - 8x_1 - 6x_2 - 4x_3 + 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 \tag{6.2.12}$$

subject to

$$x_1 \geqslant 0, \qquad x_2 \geqslant 0, \qquad x_3 \geqslant 0 \tag{6.2.13}$$

and

$$x_1 + x_2 + 2x_3 \leqslant 3. \tag{6.2.14}$$

The solution is $f = \frac{1}{9}$ at $\mathbf{x} = (\frac{4}{3}, \frac{7}{9}, \frac{4}{9})$, and the constraint (6.2.14) is active.

problems may be tackled but also because the transformed problems can be expected to be severe tests of the methods applied to solve them. The reason for this is that if the constrained problem has its solution on the boundary of the feasible region then the convergence of the transformed objective functions will not, in general, be uniform. Also, as the active constraint is approached the assumption that the objective function can be adequately approximated by a quadratic form in a neighborhood of the minimum will at best be true only in a very small neighborhood. However, there is also the advantage that the minimum for $r = r_k$ will be a very good approximation to the minimum for $r = r_{k+1}$ as r becomes small.

We can argue that conjugate direction methods should work well because when r is not small they should function normally, and when $r \to 0$ they should still work as descent methods applied to find a solution from a very good first approximation. In our experiments we have tried the Davidon and Powell (conjugate direction) methods, and we have also used the Davies, Swann, and Campey (or DSC) method which we described in outline in Chapter 3, Section 8. This is a descent method that does not use conjugate directions. We give in Table 6.8.1 the results of applying the Davidon method to the Rosen-Suzuki function, and we note that similar results have been obtained with the other methods. The final number of function calls for both of them was 900. The results for the Beale functions are given in Tables 6.8.2 to 6.8.4. We note the good performance of the DSC method. This would tend to confirm our suggestion that ultimately it is the descent nature of the computation rather than any special property of conjugate directions that is important here.

These problems do little more than indicate the feasibility of the SUMT approach to mathematical programming. However there is a growing literature of successful applications [14].

REFERENCES

1. M. J. Box, A comparison of Several Current Optimization Methods and the Use of Transformations in Constrained Problems, *Computer J.*, *9*(1966), 67.
2. H. H. Rosenbrock, An Automatic Method for Finding the Greatest or Least Value of a Function, *Computer J.*, *3*(1960), 175.
3. A. Leon, A Comparison among Eight Known Optimizing Procedures, in *Recent Advances in Optimization Techniques*, A. Lavi and T. P. Vogl, eds., John Wiley & Sons, New York, 1960, p. 33.
4. E. M. L. Beale, On an Iterative Method for Finding a Local Minimum of a Function of More Than One Variable, *Statistical Techniques Research Group*, Princeton University, Tech. Report 25, 1958.

5. J. Kowalik, A Note on Nonlinear Regression Analysis, *Austral. Computer J.*, *1*(1967), 51.
6. J. Kowalik and J. F. Morrison, Analysis of Kinetic Data for Allosteric Enzyme Reactions as a Nonlinear Regression Problem, *Math. Biosciences*, *2*(1968), 57–66.
7. J. B. Rosen and S. Suzuki, Construction of Nonlinear Programming Test Problems, *Comm. ACM*, *8*(1965), 113.
8. E. M. L. Beale, Numerical Methods, in *Nonlinear Programming*, *J. Abadie*, ed., North-Holland Publishing Co., Amsterdam, 1967, p. 150.
9. M. J. D. Powell, An Efficient Method for Finding the Minimum of a Function without Calculating Derivatives, *Computer J.*, *7*(1964), 155.
10. R. Fletcher and C. M. Reeves, Function Minimization by Conjugate Gradients, *Computer J.*, *7*(1964), 151.
11. J. Kowalik, Nonlinear Programming Procedures and Design Optimization, *Acta Polytech. Scandinav., Math. Comp. Mach. Ser.*, No. 13 (1966).
12. J. A. Nelder and R. Mead, A Simplex Method for Function Minimization, *Computer J.*, *7*(1965), 308.
13. R. Fletcher, Certification of Algorithm 251 [E4], *Comm. ACM*, *9*(1966), 686.
14. For applications of the SUMT transformation see reference 11 and also:

    G. P. McCormick, W. C. Nylander, and A. V. Fiacco, *Computer Program Implementing the Sequential Unconstrained Minimization Technique for Nonlinear Programming*, Research Analysis Corporation, Tech. Paper RAC-TP-151, McLean, Va., April, 1965.

    A. Charnes, The SUMT Method for Convex Programming: Some Discussion and Experience, in *Recent Advances in Optimization Techniques*, A. Lavi and T. P. Vogl, eds. John Wiley & Sons, New York, 1996, p. 215.

    B. C. Rush, J. Brodien, and G. P. McCormick, A Nonlinear Programming Model for Launch Vehicle Design and Costing, *Operations Res.*, *15* (1967), 185.

    A. V. Fiacco, G. P. McCormick, and W. C. Nylander, Nonlinear Programming, Duality and Shadow Pricing by Sequential Unconstrained Optimization, paper presented to First World Econometric Congress, Rome, September 14, 1967.

    D. Kavlie, J. Kowalik, S. Lund, and J. Moe, Design Optimization Using a General Nonlinear Programming Method, *European Shipbuilding*, *4*(1966), 1.

    J. Moe and S. Lund, Cost and Weight Minimization of Structures with Special Emphasis on Longitudinal Strength Members of Tankers, paper presented to Summer Meeting of the Royal Institution of Naval Architects and Koninklijk Instituut van Ingenieurs, Holland, 1967.