Algorithms For Minimization Without Derivatives describes and analyzes some practical methods for finding approximate zeros and minima of functions.

Contents include the use of successive interpolation for finding simple zeros of a function and its derivatives; an algorithm with guaranteed convergence for finding a zero of a function; an algorithm with guaranteed convergence for finding a minimum of a function of one variable; global minimization given an upper bound on the second derivative; and a new algorithm for minimizing a function of several variables without calculating derivatives.

Among the Features

- Provides reliable FORTRAN and ALGOL computer programs.
- Offers many numerical examples.
- Presents methods that are easily used on computers. These methods require no derivatives, and only function values need to be computed.
- Discusses in detail a method for finding global minima or maxima of functions of one or more variables.

(continued from front flap)

- Contains a complete analysis of the rate of convergence of some commonly used methods for finding zeros and minima of functions.
- Proves convergence for most of the algorithms, and provides error bounds that allow for the effect of rounding errors.
- Includes a comprehensive and up-to-date bibliography.

Richard P. Brent is on the Research Staff of the IBM Thomas J. Watson Research Center at Yorktown Heights, New York. He received the Ph.D. from Stanford University. Dr. Brent is a member of the Society for Industrial and Applied Mathematics and the Association for Computing Machinery.

(continued on back flap)
Algorithms for Minimization Without Derivatives

Richard P. Brent

The ever-growing relevance of computers to our daily lives increases the importance of developing algorithms suitable for computer use. This outstanding text for graduate students and research workers proposes improvements to existing algorithms, extends their related mathematical theories, and offers details on new algorithms for approximating local and global minima. None of the algorithms discussed requires an evaluation of derivatives; all depend entirely on sequential function evaluation, a highly practical scenario in the frequent event of difficult-to-evaluate derivatives.

Topics include the use of successive interpolation for finding simple zeros of a function and its derivatives; an algorithm with guaranteed convergence for finding a minimum of a function of one variation; global minimization given an upper bound on the second derivative; and a new algorithm for minimizing a function of several variables without calculating derivatives.

Many numerical examples appear here, along with a complete analysis of the rate of convergence for most of the algorithms and error bounds that allow for the effect of rounding errors.


ALSO AVAILABLE

Computability and Unsolvability, Martin Davis. 288pp. 5⅛ x 8½. 61471-9
Theory of Approximation, N. I. Achieser. 317pp. 5⅛ x 8½. 67129-1

For current price information write to Dover Publications, or log on to www.doverpublications.com—and see every Dover book in print.

Free Dover Mathematics and Science Catalog (59065-8) available upon request.
DOVER BOOKS ON MATHEMATICS

HANDBOOK OF MATHEMATICAL FUNCTIONS, Milton Abramowitz and Irene A. Stegun. (61272-4) $32.95

THEORY OF APPROXIMATION, N. I. Achieser. (67129-1) $8.95

TENSOR ANALYSIS ON MANIFOLDS, Richard L. Bishop and Samuel I. Goldberg. (64039-6) $9.95

TABLES OF INDEFINITE INTEGRALS, G. Petit Bois. (60225-7) $8.95

VECTOR AND TENSOR ANALYSIS WITH APPLICATIONS, A. I. Borisenko and I. E. Tarapov. (63833-2) $10.95

THE HISTORY OF THE CALCULUS AND ITS CONCEPTUAL DEVELOPMENT, Carl B. Boyer. (60509-4) $9.95

THE QUALITATIVE THEORY OF ORDINARY DIFFERENTIAL EQUATIONS: AN INTRODUCTION, Fred Brauer and John A. Nohel. (63846-5) $12.95

PRINCIPLES OF STATISTICS, M. G. Bulmer. (63769-3) $9.95

THE THEORY OF SPINORS, Élie Cartan. (64078-1) $8.95

ADVANCED NUMBER THEORY, Harvey Cohn. (64023-X) $10.95

STATISTICS MANUAL, Edwin L. Crow, Francis Davis, and Margaret Maxfield. (60599-X) $8.95

FOURIER SERIES AND ORTHOGONAL FUNCTIONS, Harry F. Davis. (65973-9) $13.95

COMPUTABILITY AND UNSOLVABILITY, Martin Davis. (61471-9) $12.95

ASYMPTOTIC METHODS IN ANALYSIS, N. G. de Bruijn. (64221-6) $9.95

PROBLEMS IN GROUP THEORY, John D. Dixon. (61574-X) $10.95

THE MATHEMATICS OF GAMES OF STRATEGY, Melvin Dresher. (64216-X) $7.95

ASYMPTOTIC EXPANSIONS, A. Erdélyi. (60318-0) $6.95

COMPLEX VARIABLES: HARMONIC AND ANALYTIC FUNCTIONS, Francis J. Flanigan. (61388-7) $10.95

ON FORMALLY UNDECIDABLE PROPOSITIONS OF PRINCIPIA MATHEMATICA AND RELATED SYSTEMS, Kurt Gödel. (66980-7) $6.95

A HISTORY OF GREEK MATHEMATICS, Sir Thomas Heath. (24073-8, 24074-6) Two-volume set $29.90

PROBABILITY: ELEMENTS OF THE MATHEMATICAL THEORY, C. R. Heathcote. (41149-4) $8.95

INTRODUCTION TO NUMERICAL ANALYSIS, Francis B. Hildebrand. (65363-3) $16.95

METHODS OF APPLIED MATHEMATICS, Francis B. Hildebrand. (67092-3) $12.95

TOPOLOGY, John G. Hocking and Gail S. Young. (65676-4) $13.95

MATHEMATICS AND LOGIC, Mark Kac and Stanislaw M. Ulam. (67085-6) $7.95

MATHEMATICAL METHODS AND THEORY IN GAMES, PROGRAMMING, AND ECONOMICS, Samuel Karlin. (67020-1) $24.95

MATHEMATICAL FOUNDATIONS OF INFORMATION THEORY, A. I. Khinchin. (60434-9) $5.95

(continued on back flap)
CALCULUS REFRESHER, A. Albert Klafl. (20370-0) $10.95
PROBLEM BOOK IN THE THEORY OF FUNCTIONS, Konrad Knopp. (41451-5) $8.95
INTRODUCTORY REAL ANALYSIS, A. N. Kolmogorov and S. V. Fomin. (61226-0) $14.95
SPECIAL FUNCTIONS AND THEIR APPLICATIONS, N. N. Lebedev. (60624-4) $12.95
TENSORS, DIFFERENTIAL FORMS, AND VARIATIONAL PRINCIPLES, David Lovelock and Hanno Rund. (65840-6) $13.95
SURVEY OF MATRIX THEORY AND MATRIX INEQUALITIES, Marvin Marcus and Henryk Minc. (67102-X) $8.95
ABSTRACT ALGEBRA AND SOLUTION BY RADICALS, John E. and Margaret W. Maxfield. (67121-6) $9.95
FUNDAMENTAL CONCEPTS OF ALGEBRA, Bruce E. Meserve. (61470-0) $12.95
FUNDAMENTAL CONCEPTS OF GEOMETRY, Bruce E. Meserve. (63415-9) $11.95
FIFTY CHALLENGING PROBLEMS IN PROBABILITY WITH SOLUTIONS, Frederick Mosteller. (65355-2) $5.95
ELEMENTS OF THE TOPOLOGY OF PLANE SETS OF POINTS, M. H. A. Newman. (67037-6) $6.95
NUMBER THEORY AND ITS HISTORY, Oystein Ore. (65620-9) $12.95
MATRICES AND TRANSFORMATIONS, Anthony J. Pargett. (63634-8) $7.95
PROBABILITY THEORY: A CONCISE COURSE, Y. A. Rozanov. (63544-9) $8.95
ORDINARY DIFFERENTIAL EQUATIONS AND STABILITY THEORY: AN INTRODUCTION, David A. Sánchez. (63828-6) $6.95
LINEAR ALGEBRA, Georgi E. Shilov. (63518-X) $11.95
ESSENTIAL CALCULUS WITH APPLICATIONS, Richard A. Silverman. (66097-4) $11.95
A CONCISE HISTORY OF MATHEMATICS, Dirk J. Struik. (60255-9) $8.95
PROBLEMS IN PROBABILITY THEORY, MATHEMATICAL STATISTICS AND THEORY OF RANDOM FUNCTIONS, A. A. Sveshnikov. (63717-4) $14.95
TENSOR CALCULUS, J. L. Synge and A. Schild. (63612-7) $13.95
CALCULUS OF VARIATIONS WITH APPLICATIONS TO PHYSICS AND ENGINEERING, Robert Weinstock. (63069-2) $12.95
INTRODUCTION TO VECTOR AND TENSOR ANALYSIS, Robert C. Wrede. (61879-X) $12.95
DISTRIBUTION THEORY AND TRANSFORM ANALYSIS, A. H. Zemanian. (65479-6) $13.95

Paperbound unless otherwise indicated. Prices subject to change without notice. Available at your book dealer or online at www.doverpublications.com. Write for free catalogues to Dept. 23, Dover Publications, Inc., 31 East 2nd Street, Mineola, NY 11501. Please indicate field of interest. Each year Dover publishes over 400 books on fine art, music, crafts and needlework, antiques, languages, literature, children's books, chess, cookery, nature, anthropology, science, mathematics, and other areas.

Manufactured in the U.S.A.
WITHOUT DERIVATIVES

MINIMIZATION

ALGORITHMS FOR
INTRODUCTION
DIVERGED DIFFERENCES, AND INTEGRAL
SOME USEFUL RESULTS ON TAYLOR SERIES.

1

INTRODUCTION AND SUMMARY

Preface

CONTENTS
CONTENTS

PREFACE xi

1 INTRODUCTION AND SUMMARY 1

1.1 Introduction 1

1.2 Summary 4

2 SOME USEFUL RESULTS ON TAYLOR SERIES, DIVIDED DIFFERENCES, AND LAGRANGE INTERPOLATION 9

2.1 Introduction 9

2.2 Notation and definitions 10

2.3 Truncated Taylor series 11

2.4 Lagrange interpolation 12

2.5 Divided differences 13

2.6 Differentiating the error 15


3

THE USE OF SUCCESSIVE INTERPOLATION
FOR FINDING SIMPLE ZEROS OF A FUNCTION
AND ITS DERIVATIVES

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>19</td>
</tr>
<tr>
<td>3.2</td>
<td>The definition of order</td>
<td>21</td>
</tr>
<tr>
<td>3.3</td>
<td>Convergence to a zero</td>
<td>22</td>
</tr>
<tr>
<td>3.4</td>
<td>Superlinear convergence</td>
<td>24</td>
</tr>
<tr>
<td>3.5</td>
<td>Strict superlinear convergence</td>
<td>26</td>
</tr>
<tr>
<td>3.6</td>
<td>The exact order of convergence</td>
<td>29</td>
</tr>
<tr>
<td>3.7</td>
<td>Stronger results for $q = 1$ and $2$</td>
<td>34</td>
</tr>
<tr>
<td>3.8</td>
<td>Accelerating convergence</td>
<td>40</td>
</tr>
<tr>
<td>3.9</td>
<td>Some numerical examples</td>
<td>43</td>
</tr>
<tr>
<td>3.10</td>
<td>Summary</td>
<td>45</td>
</tr>
</tbody>
</table>

4

AN ALGORITHM WITH GUARANTEED
CONVERGENCE FOR FINDING A ZERO
OF A FUNCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>47</td>
</tr>
<tr>
<td>4.2</td>
<td>The algorithm</td>
<td>48</td>
</tr>
<tr>
<td>4.3</td>
<td>Convergence properties</td>
<td>53</td>
</tr>
<tr>
<td>4.4</td>
<td>Practical tests</td>
<td>54</td>
</tr>
<tr>
<td>4.5</td>
<td>Conclusion</td>
<td>56</td>
</tr>
<tr>
<td>4.6</td>
<td>ALGOL 60 procedures</td>
<td>58</td>
</tr>
</tbody>
</table>

5

AN ALGORITHM WITH GUARANTEED
CONVERGENCE FOR FINDING A MINIMUM
OF A FUNCTION OF ONE VARIABLE

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>61</td>
</tr>
<tr>
<td>5.2</td>
<td>Fundamental limitations because of rounding errors</td>
<td>63</td>
</tr>
<tr>
<td>5.3</td>
<td>Unimodality and $\delta$-unimodality</td>
<td>65</td>
</tr>
<tr>
<td>5.4</td>
<td>An algorithm analogous to Dekker's algorithm</td>
<td>72</td>
</tr>
<tr>
<td>5.5</td>
<td>Convergence properties</td>
<td>75</td>
</tr>
<tr>
<td>5.6</td>
<td>Practical tests</td>
<td>76</td>
</tr>
<tr>
<td>5.7</td>
<td>Conclusion</td>
<td>78</td>
</tr>
<tr>
<td>5.8</td>
<td>An ALGOL 60 procedure</td>
<td>79</td>
</tr>
</tbody>
</table>
GLOBAL MINIMIZATION GIVEN AN UPPER BOUND ON THE SECOND DERIVATIVE

6.1 Introduction 91
6.2 The basic theorems 64
6.3 An algorithm for global minimization 86
6.4 The rate of convergence in some special cases 97
6.5 A lower bound on the number of function evaluations required 100
6.6 Practical tests 103
6.7 Some extensions and generalizations 105
6.8 An algorithm for global minimization of a function of several variables 107
6.9 Summary and conclusions 111
6.10 ALGOL 60 procedures 112

A NEW ALGORITHM FOR MINIMIZING A FUNCTION OF SEVERAL VARIABLES WITHOUT CALCULATING DERIVATIVES

7.1 Introduction and survey of the literature 116
7.2 The effect of rounding errors 122
7.3 Powell's algorithm 124
7.4 The main modification 128
7.5 The resolution ridge problem 132
7.6 Some further details 135
7.7 Numerical results and comparison with other methods 137
7.8 Conclusion 154
7.9 An ALGOL W procedure and test program 155

BIBLIOGRAPHY 169

APPENDIX: FORTRAN subroutines 187

INDEX 193
PREFACE

The problem of finding numerical approximations to the zeros and extrema of functions, using hand computation, has a long history. Recently considerable progress has been made in the development of algorithms suitable for use on a digital computer. In this book we suggest improvements to some of these algorithms, extend the mathematical theory behind them, and describe some new algorithms for approximating local and global minima. The unifying thread is that all the algorithms considered depend entirely on sequential function evaluations; no evaluations of derivatives are required. Such algorithms are very useful if derivatives are difficult to evaluate, which is often true in practical problems.

An earlier version of this book appeared as Stanford University Report CS-71-198, Algorithms for finding zeros and extrema of functions without calculating derivatives, now out of print. This expanded version is published in the hope that it will interest graduate students and research workers in numerical analysis, computer science, and operations research.

I am greatly indebted to Professors G. E. Forsythe and G. H. Golub for their advice and encouragement during my stay at Stanford. Thanks are due to them and to Professors J. G. Herriot, F. W. Dorr, and C. B. Moier, both for their careful reading of various drafts and for many helpful suggestions. Dr. T. J. Rivlin suggested how to find bounds on polynomials (Chapter 6), and Dr. J. H. Wilkinson introduced me to Dekker’s algorithm (Chapter 4). Parts of Chapter 4 appeared in Brent (1971d), and are included in this book by kind permission of the Editor of The Computer Journal. Thanks go to
Professor F. Dorr and Dr. I. Sobel for their help in testing some of the algorithms; to Michael Malcolm, Michael Saunders, and Alan George for many interesting discussions; and to Phyllis Winkler for her nearly perfect typing. I am also grateful for the influence of my teachers V. Grenness, H. Smith, Dr. D. Faulkner, Dr. E. Strzelecki, Professors G. Preston, J. Miller, Z. Janko, R. Floyd, D. Knuth, G. Polya, and M. Schiffer.

Deepest thanks go to Erin Brent for her help in obtaining some of the numerical results, testing the algorithms, plotting graphs, reading proofs, and in many other ways.

Finally I wish to thank the Commonwealth Scientific and Industrial Research Organization, Australia, for its generous support during my stay at Stanford.

This work is dedicated to Oscar and Nancy Brent, who laid the foundations; and to George Forsythe, who guided the construction.

R. Brent
and as $N$ is the least positive $n$ such that $x_n \geq b$, this gives

$$N = \left\lceil \frac{\sqrt{\frac{M}{k}}}{\sqrt{k(b-a) + 1}} \right\rceil.$$  

(4.12) shows that $N$ is essentially proportional to $\sqrt{M}$.

**Diagram 4.1** A straight line

Two limiting cases of (4.12) are interesting. If $t$ is small and $k$ not too small, so that $k(b-a) \gg t$, then

$$N \approx \sqrt{\frac{M(b-a)}{2k}},$$

(4.13)

which is independent of $t$. (In this section we are neglecting the effect of rounding errors, but these should not be important if $t$ satisfies the weak condition (3.68).)

If $k$ is very small, so that $k(b-a) \ll t$, then (4.12) gives

$$N \approx \frac{b-a}{2\delta},$$

(4.14)

and the algorithm proceeds in steps of size about $2\delta$, where $\delta$ is given by (4.1).

**A parabola**

If the global minimum of $f$ occurs at an interior point $\mu$, then $f'(\mu) = 0$. If $f''(\mu) \neq 0$ we may analyze the behavior of the algorithm near $\mu$ by considering the parabolic approximation $f(\mu) + \frac{1}{2}f''(\mu)(x - \mu)^2$ to $f(x)$. Thus, suppose that

$$M > m > 0$$

(4.15)

and

$$f(x) = \frac{1}{2}m(x - \mu)^2 + t,$$

(4.16)