NUMERICAL METHODS FOR NON-LINEAR OPTIMIZATION

edited by
F. A. LOOTSMA

Numerical Methods for Non-linear Optimization

Conference sponsored by the

Science Research Council

University of Dundee, Scotland, 1971

Edited by

F. A. LOOTSMA

Philips Research Laboratories
Eindhoven, The Netherlands

1972

Academic Press **London · New York**

ACADEMIC PRESS INC. (LONDON) LTD.
24/28 Oval Road,
London NW1

United States Edition published by
ACADEMIC PRESS INC.
111 Fifth Avenue
New York, New York 10003

Library of Congress Catalog Card Number: 72-84446
ISBN: 0-12-455650-7

PRINTED IN GREAT BRITAIN BY
WILLIAM CLOWES & SONS LIMITED,
LONDON, COLCHESTER AND BECCLES

Contributors

BEALE, E. M. L., *Scientific Control Systems Ltd., Sanderson House, London, England*

BIGGS, M. C., *Numerical Optimization Centre, The Hatfield Polytechnic, Hatfield, Hertfordshire, England*

BOROWSKI, N., *Department of Electrical Engineering, The University of British Columbia, Vancouver, British Columbia, Canada*

BRANIN, F. H., JR., *Systems Development Division Laboratory, I.B.M. Corporation, Kingston, New York, U.S.A.*

BROYDEN, C. G., *Computing Centre, University of Essex, Colchester, Essex, England*

DAVIES, M., *Department of Mathematics, University of Surrey, Guildford, Surrey, England*

DENNIS, J. E., JR., *Department of Computer Science, Cornell University, Ithaca, New York, U.S.A.*

DIXON, L. C. W., *The Numerical Optimization Centre, The Hatfield Polytechnic, Hatfield, Hertfordshire, England*

ECKHARDT, U., *Zentralinstitut für Angewandte Mathematik, Kernforschungsanlage Jülich, Jülich, Germany*

FLETCHER, R., *Theoretical Physics Division, U.K.A.E.A. Research Group, Atomic Energy Research Establishment, Harwell, England*

GOLDFARB, D., *The City College of the City University of New York, New York, New York, U.S.A.*

GONÇALVES, A. S., *Universidade de Coimbra, Instituto de Matemática, Coimbra, Portugal*

GOULD, F. J., *Department of Statistics and Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill, North Carolina, U.S.A.*

HANSEN, P., *Institut d'Economie Scientifique et de Gestion, Lille, France*

HIMMELBLAU, D. M., *Department of Chemical Engineering, The University of Texas at Austin, Austin, Texas, U.S.A.*

HOO, S. K., *Systems Development Division Laboratory, I.B.M. Corporation, Kingston, New York, U.S.A.*

HUTCHINSON, D., *Department of Computational Science, The University of Leeds, Leeds, England*

JOHNSON, M. P., *Department of Mathematics, Mid-Essex Technical College, Chelmsford, Essex, England*

KREUSER, J., *University of Minnesota, Minneapolis, Minnesota, U.S.A.*

LILL, S. A., *Department of Computational and Statistical Science, University of Liverpool, Liverpool, England*

LONEY, S. T., *South of Scotland Electricity Board, Glasgow, Scotland*

LOOTSMA, F. A., *Philips Research Laboratories, Eindhoven, The Netherlands*

MCCORMICK, G. P., *George Washington University, Washington, D.C., U.S.A.*

MIFFLIN, R., *Department of Administrative Sciences, Yale University, New Haven, Connecticut, U.S.A.*

OSBORNE, M. R., *Australian National University, Canberra, Australia*

PARKINSON, J. M., *Department of Computational Science, The University of Leeds, Leeds, England*

POWELL, M. J. D., *Mathematics Branch, Theoretical Physics Division, U.K.A.E.A. Research Group, Atomic Energy Research Establishment, Harwell, England*

ROSEN, J. B., *University of Minnesota, Minneapolis, Minnesota, U.S.A.*

RYAN, D. M., *Australian National University, Canberra, Australia*

SARGENT, R. W. H., *Department of Chemical Engineering, Imperial College, London, England*

SCHRACK, G., *The University of British Columbia, Department of Electrical Engineering, Vancouver, British Columbia, Canada*

SEBASTIAN, D. J., *Department of Chemical Engineering, Imperial College, London, England*

UEING, U., *University of Bonn, Institut für Gesellschafts- und Wirtschaftswissenschaften, Bonn, Germany*

WHITTING, I. J., *Operational Research Department, The Gas Council, London, England*

Preface

Organization

This volume contains a collection of papers presented at the Conference on Numerical Methods for Non-linear Optimization, which was held at the University of Dundee (Scotland), from 28th June to 1st July 1971. The choice of Dundee was not a coincidence. The Science Research Council had generously decided to give considerable financial support to the University of Dundee for the purpose of supporting a symposium on the theory of Numerical Analysis during the period September 1970 to August 1971. Many pure and applied mathematicians, numerical analysts etc. visited Dundee, and two major gatherings were organized: a Research Conference on Numerical Analysis and the above named Conference (the idea originated from the Mathematics Department in Dundee), which fitted very well into the framework of the 'Numerical Analysis Year'.

The Dundee Conference on Numerical Methods for Non-linear Optimization was the successor to the Conference on Optimization (sponsored by the British Computer Society and the Institute of Mathematics and its Applications), held in March 1968 at the University of Keele (England). It was also confined to the area of non-linear, unconstrained and constrained optimization with the emphasis on the numerical aspects; methods for integer, linear programming have been left out of consideration. The papers of the Keele conference have been edited by R. Fletcher, and published by Academic Press, London, 1969, under the title "Optimization".

The Programme Committee for the Dundee Conference consisted of M. J. D. Powell (Atomic Energy Research Establishment, Harwell; Chairman of the Committee), G. A. Watson (University of Dundee; secretary of the Committee), M. R. Osborne (Australian National University, Canberra; visiting professor at the University of Dundee), and the editor (Philips Research Laboratories, Eindhoven, The Netherlands; temporarily at Mullard Research Laboratories, Redhill, Surrey). We have cooperated in a most pleasant atmosphere. An encouraging factor was the rapid and favourable response of the invited speakers: P. Wolfe (IBM, New York), R. W. H. Sargent (Imperial College, London), G. P. McCormick (The George Washington University, Washington), E. M. L. Beale (Scientific Control Systems, London), R. Fletcher (Atomic Energy Research Establishment, Harwell), J. Abadie (Electricité de France, Paris), and J. B. Rosen (University of Minnesota, Minneapolis); in addition M. R. Osborne and the editor agreed to present a survey on their favourite subject.

The programme was finally composed during a meeting in Dundee, on 5th and 6th May 1971. This completed the task of the Programme Committee. The local organization was further carried out by G. A. Watson assisted by many others too numerous to mention by name, both academic and secretarial staff and students of the Mathematics Department at the University of Dundee. They gave all the time and effort that was needed to ensure a well run conference. We thank all of them for the smooth organisation and for their excellent hospitality.

Outline of the Present Volume

Surveying the contents one will find that the contributions fall into a few, distinct categories.

First, there are the papers dealing with methods for unconstrained optimization. The theoretical aspects are further explored by M. J. D. Powell, J. B. Dennis, C. G. Broyden and M. P. Johnson, and E. M. L. Beale; the authors are mainly concerned with the successful variable-metric and conjugate-gradient methods. Numerical experiments to compare algorithms for unconstrained optimization are presented by R. W. H. Sargent and D. J. Sebastian, D. M. Himmelblau, J. M. Parkinson and D. Hutchinson, G. Schrack and N. Borowski, and L. C. W. Dixon. It is interesting to note that some of the authors devote considerable attention to the non-gradient methods. In the last few years this important class of methods has been somewhat neglected; at least, this is our impression if we inspect the leading journals and compare the number of papers on this subject with the abundant literature on gradient methods.

Problems of non-linear least squares and curve fitting are presented by M. R. Osborne, M. Davies and I. J. Whitting, and S. T. Loney. These problems have a wide range of applications, and the numerical methods for solving them are so closely connected with optimization methods that we gladly received the papers for presentation at the conference.

An appealing direction for future research is the design of methods to find global minima of problems that may have local, non-global minima. This volume contains three papers on this subject presented by G. P. McCormick, U. Ueing, and F. H. Branin and S. K. Hoo. Particularly the lively presentation by Branin should be mentioned here; it was one of the highlights of the conference.

Lastly, the reader will find a collection of papers on constrained optimization. Methods for quadratic programming are discussed by D. Goldfarb and A. S. Gonçalves, and P. Hansen presents a quadratic-programming method for the particular problem where the variables are restricted to the values 0 and 1. Methods tailored to problems with linear constraints are surveyed by R. Fletcher. Surprisingly enough, there is only one other paper, by J. B. Rosen and J. Kreuser, that deals extensively with methods of this kind. The Programme Committee expected more papers on this subject: in the literature these methods are often considered as effective tools if the (majority of the) constraints are linear. A broad treatise on complementary algorithms for programming problems with an infinite number of constraints is given by U. Eckhardt. Finally, keen interest is still being shown in the penalty-function approach for solving constrained minimization problems. This volume contains seven papers on this subject, presented by F. J. Gould, R. Mifflin, R. Fletcher, Shirley Lill, M. R. Osborne and D. M. Ryan, M. C. Biggs, and by the editor.

As in many other fields of scientific research there is an astonishing proliferation of names and terms. During the conference this was alluded to by F. H. Branin, who suddenly interrupted his talk to ask his audience a teasing question; an immediate answer was given by D. J. Wilde.

Branin: By the way, everybody seems to know who Newton is, but who is Raphson?

Wilde: He is Newton's programmer, I guess.

Branin: That is the best answer I have ever heard.

Presentation of Numerical Results

The literature on optimization has shown a regrettable lack of uniformity in the presentation of numerical results. When the number of function evaluations is recorded, it may include the gradient evaluations as well; it is not always obvious whether the derivatives were obtained by numerical differentiation or not; the accuracy required *a priori* is sometimes not mentioned so that one does not know what is understood by a solution to a given test problem; the execution time is often omitted, although this quantity might be of interest if a number of algorithms are compared on the same computer. The Programme Committee therefore suggested some guidelines to the contributors, and it is gratifying that many authors have followed the recommendations. Nevertheless, there are several inherent difficulties which should be mentioned here.

The number of function, gradient, and possibly Hessian-matrix evaluations, sometimes lumped together in the number of equivalent function evaluations, is a machine-independent measure of efficiency. However, it is an unsatisfactory yardstick for comparing optimization algorithms, since it does not accurately inform the reader of the total effort necessary to solve a given test problem. The gradient projection methods, for instance, will entail many time-consuming array manipulations which are often not negligible with respect to the function and gradient evaluations. Execution times and the computer used could be mentioned in order to provide a basis for overall comparison. And this is also a reasonable basis: the ultimate purpose of many investigations in this field is the development of algorithms that are faster than the existing ones. We cannot ignore the following problems, however.

First, computers with multi-programming and time-sharing facilities do not always inform the user of the execution time for the total job or, which is even more desirable, for specified parts of the job. Several contributors (Dixon, Biggs) have expressed their concern about this state of affairs, and we deplore a development that might destroy an important criterion for the comparison of algorithms.

Even if the computer used and the compiler are specified, a comparison of various optimization algorithms on different computers remains cumbersome. To our knowledge, only one such attempt, with the execution time as performance criterion, has been made. The study was carried out by Colville (see A. R. Colville, "A comparative study on non-linear programming codes". IBM NYSC Report No. 320-2949, June 1968), and his report gives an impressive list of specialists in the field of optimization cooperating in this project. We appreciate the attempt, but we have found many deficiencies in the report. The standard timing programme (the basis for the comparison) is lengthy and not very representative; basically, it consists of a number of arithmetic operations and array manipulations (in the proper balance?), but subroutine calls are practically missing. Furthermore, a detailed account of the number of function and derivative evaluations is not given, and the accuracy of the calculated solutions is sometimes rather poor.

Nevertheless, we feel that the comparison of optimization algorithms on the basis of execution times is very important. The papers by Himmelblau and Biggs, for instance, show that the difference in execution times is sometimes not very striking, even if there are considerable variations in the number of function and derivative evaluations. In those circumstances the properties of the computers and compilers cannot be neglected, and it is our conviction that significant progress in optimization methods can only be made if more attention is given to these aspects, as well as to the mathematical approach.

Acknowledgement

It is a pleasure to acknowledge the considerable support I have received from the referees. It is gratifying to record that hardly anyone who was asked to review a contribution refused to do so, notwithstanding their many professional commitments. I am also most grateful to the contributors for their coöperation and their willingness to follow the suggestions of the referees.

I wish to express my gratitude to the Board of Directors and to my col-

leagues at the Computer Departments of Mullard Research Laboratories (Redhill, Surrey, England) and Philips Research Laboratories (Eindhoven, The Netherlands). The time I spent on the Dundee Conference and the proceedings could not have been available without their acceptance that other tasks would have to be postponed.

I would like to dedicate the book to my wife Ricky for her whole-hearted support and encouragement.

Eindhoven, June 1972 F. A. LOOTSMA

Contents

Contributors v
Preface vii
1. Some Properties of the Variable Metric Algorithm
M. J. D. Powell 1
2. On Some Methods Based on Broyden's Secant Approximation to the Hessian
J. E. Dennis, Jr. 19
3. A Class of Rank-1 Optimization Algorithms
C. G. Broyden and M. P. Johnson 35
4. A Derivation of Conjugate Gradients
E. M. L. Beale 39
5. Numerical Experience with Algorithms for Unconstrained Minimization
R. W. H. Sargent and D. J. Sebastian 45
6. A Uniform Evaluation of Unconstrained Optimization Techniques
D. M. Himmelblau. 69
7. A Consideration of Non-gradient Algorithms for the Unconstrained Optimization of Functions of High Dimensionality
J. M. Parkinson and D. Hutchinson 99
8. An Investigation into the Efficiency of Variants on the Simplex Method
J. M. Parkinson and D. Hutchinson 115
9. An Experimental Comparison of Three Random Searches
G. Schrack and N. Borowski 137
10. The Choice of Step Length, a Crucial Factor in the Performance of Variable Metric Algorithms
L. C. W. Dixon 149
11. Some Aspects of Non-linear Least Squares Calculations
M. R. Osborne 171
12. A Modified Form of Levenberg's Correction
M. Davies and I. J. Whitting 191
13. A Dynamic Programming Algorithm for Load Duration Curve Fitting
S. T. Loney 203

14. Attempts to Calculate Global Solutions of Problems that may have Local Minima
G. P. McCormick 209
15. A Combinatorial Method to Compute a Global Solution of Certain Non-convex Optimization Problems
U. Ueing 223
16. A Method for Finding Multiple Extrema of a Function of *n* Variables
F. H. Branin, Jr. and S. K. Hoo 231
17. Extensions of Newton's Method and Simplex Methods for Solving Quadratic Programs
D. Goldfarb 239
18. A Primal-Dual Method for Quadratic Programming with Bounded Variables
A. S. Gonçalves 255
19. Quadratic Zero-One Programming by Implicit Enumeration
P. Hansen 265
20. Minimizing General Functions Subject to Linear Constraints
R. Fletcher 279
21. A Gradient Projection Algorithm for Non-linear Constraints
J. B. Rosen and J. Kreuser 297
22. Pseudo-Complementary Algorithms for Mathematical Programming
U. Eckhardt 301
23. A Survey of Methods for Solving Constrained Minimization Problems via Unconstrained Minimization
F. A. Lootsma 313
24. Non-linear Tolerance Programming
F. J. Gould 349
25. On the Convergence of the Logarithmic Barrier Function Method
R. Mifflin 367
26. A Class of Methods for Non-linear Programming. III: Rates of Convergence
R. Fletcher 371
27. Generalization of an Exact Method for Solving Equality Constrained Problems to deal with Inequality Constraints
S. A. Lill 383
28. A Hybrid Algorithm for Non-linear Programming
M. R. Osborne and D. M. Ryan 395
29. Constrained Minimization Using Recursive Equality Quadratic Programming
M. C. Biggs 411
Author Index 429
Subject Index 435

1. Some Properties of the Variable Metric Algorithm

M. J. D. POWELL
Mathematics Branch, Theoretical Physics Division,
U.K.A.E.A. Research Group, Atomic Energy Research Establishment,
Harwell, England

Summary

The variable metric algorithm for calculating the least value of a function $F(\mathbf{x})$ is usually successful in practice, but it has not been analysed theoretically, except when $F(\mathbf{x})$ is a uniformly convex function. Therefore, in this paper some preliminary theorems are given, that require only that a level set $\{\mathbf{x} | F(\mathbf{x}) \leqslant F(\mathbf{x}^{(1)})\}$ is bounded, and that $F(\mathbf{x})$ has bounded second derivatives. In this paper we are unable to show convergence in general, but an interesting corollary of the theorems is that we can prove convergence if $F(\mathbf{x})$ is convex, whereas the previous theorems depend on *uniform* convexity.

1. Introduction

To solve the problem of calculating on a computer the least value of a function $F(x_1, x_2, \ldots, x_n) = F(\mathbf{x})$, say, the variable metric algorithm (Davidon, 1959; Fletcher and Powell, 1963) is often used, and it is usually successful. This algorithm is described briefly in Section 2, and some of its properties, taken from previously published papers, are given in Section 3. However, there are no theorems that explain the success of the algorithm, except in the special case when $F(\mathbf{x})$ is a uniformly convex function. So his paper describes some preliminary results that depend on much less restrictive conditions on $F(\mathbf{x})$.

The algorithm is iterative, and, given a starting point $\mathbf{x}^{(1)}$, it generates a sequence of points $\mathbf{x}^{(k)}$ $(k = 1, 2, \ldots)$, that is intended to converge to the point at which $F(\mathbf{x})$ is least. The vector $\mathbf{g}^{(k)}$ is defined to be the gradient of $F(\mathbf{x})$ at $\mathbf{x}^{(k)}$.

The conditions on $F(\mathbf{x})$ that we impose are that the level set $\{\mathbf{x} | F(\mathbf{x}) \leqslant F(\mathbf{x}^{(1)})\}$ is bounded, and that $F(\mathbf{x})$ has bounded second derivatives. The notation $G(\mathbf{x})$ stands for the second derivative matrix of $F(\mathbf{x})$ at $\mathbf{x}$, and we let the bound be the inequality

$$\|G(\mathbf{x})\| \leqslant M. \tag{1}$$

Throughout this paper the vector norms are Euclidean, and the matrix norms are induced by the Euclidean vector norm.

Most of the given results are derived from the doubtful conjecture: 'There exist functions $F(\mathbf{x})$, satisfying the conditions of the last paragraph, for which the sequence of numbers $\|\mathbf{g}^{(k)}\|$ $(k = 1,2,\ldots)$ is bounded away from zero'. We would like to show that this conjecture is false, because then it would follow that the limit points of the sequence $\mathbf{x}^{(k)}$ $(k = 1,2,\ldots)$ include at least one stationary point of $F(\mathbf{x})$. We have used the term stationary point, instead of local minimum, because Wolfe (1971) has found a function for which the variable metric algorithm converges to a saddle point.

Therefore, in Section 4 we suppose that for some positive constant c the inequality $\|\mathbf{g}^{(k)}\| \geq c$ $(k = 1,2,\ldots)$ holds, and we deduce a number of consequences of this hypothesis. Some of these deductions are surprising, but unfortunately we have not been able to show that they are contradictory.

However, in Section 5 we note that the deductions are contradictory if we include the extra condition that $F(\mathbf{x})$ is convex. Thus we prove that the variable metric algorithm converges for convex functions, whereas the previous theorems (Powell, 1971) depend on *uniform* convexity.

In Section 6 a different method of analysis is used, and it is shown that if $F(\mathbf{x})$ is a function of only two variables, then the sequence $\mathbf{x}^{(k)}$ $(k = 1,2,\ldots)$ cannot converge to a point at which the gradient of $F(\mathbf{x})$ is non-zero.

Finally, in Section 7, there is a discussion of our theorems, and it is pointed out that they are relevant to a number of published algorithms (Dixon, 1971).

2. The Variable Metric Algorithm

The kth iteration of the algorithm calculates the point $\mathbf{x}^{(k+1)}$ from the point $\mathbf{x}^{(k)}$. It depends on a positive definite matrix $H^{(k)}$. To begin the first iteration, $H^{(1)}$ is frequently set to the unit matrix, but any other positive definite matrix may be used instead. As well as calculating $\mathbf{x}^{(k+1)}$, the kth iteration calculates the matrix $H^{(k+1)}$.

If $\mathbf{g}^{(k)} = \mathbf{0}$ the iterative process is terminated. Otherwise the point $\mathbf{x}^{(k+1)}$ is defined by the formula

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \lambda^{(k)} H^{(k)} \mathbf{g}^{(k)}, \tag{2}$$

where $\lambda^{(k)}$ is a multiplier. The value of $\lambda^{(k)}$ is obtained by considering the function of one variable

$$\phi(\lambda) = F(\mathbf{x}^{(k)} - \lambda H^{(k)} \mathbf{g}^{(k)}), \tag{3}$$

and ideally $\lambda^{(k)}$ should be the value of λ that minimizes $\phi(\lambda)$ subject to $\lambda \geq 0$. Because $H^{(k)}$ is positive definite, and because the level set $\{\mathbf{x} | F(\mathbf{x}) \leq F(\mathbf{x}^{(1)})\}$ is bounded, $\lambda^{(k)}$ is positive and finite. If it happens that more than one value of

λ minimizes $\phi(\lambda)$, we remove the ambiguity by letting $\lambda^{(k)}$ be the least positive value of λ that minimizes $\phi(\lambda)$.

However, in practice it is impossible to calculate $\lambda^{(k)}$ precisely, and it seems to be a good strategy to tolerate rather large errors in the value of $\lambda^{(k)}$. But in this paper we presume the ideal case where $\lambda^{(k)}$ does minimize $\phi(\lambda)$ subject to $\lambda \geqslant 0$.

The matrix $H^{(k+1)}$ is defined by the formula

$$H^{(k+1)} = H^{(k)} - \frac{H^{(k)}\boldsymbol{\gamma}^{(k)}\boldsymbol{\gamma}^{(k)T}H^{(k)}}{(\boldsymbol{\gamma}^{(k)}, H^{(k)}\boldsymbol{\gamma}^{(k)})} + \frac{\boldsymbol{\delta}^{(k)}\boldsymbol{\delta}^{(k)T}}{(\boldsymbol{\delta}^{(k)}, \boldsymbol{\gamma}^{(k)})}, \tag{4}$$

where $\boldsymbol{\delta}^{(k)}$ and $\boldsymbol{\gamma}^{(k)}$ are the vectors

$$\left.\begin{aligned} \boldsymbol{\delta}^{(k)} &= \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}, \\ \boldsymbol{\gamma}^{(k)} &= \mathbf{g}^{(k+1)} - \mathbf{g}^{(k)}, \end{aligned}\right\} \tag{5}$$

and where the superscript T distinguishes a row vector from a column vector. When equations (2) and (4) have been applied the $(k+1)$th iteration is begun.

3. Properties of the Variable Metric Algorithm

In this section we summarize the properties of the variable metric algorithm that have been published already, and that are required in later sections.

The numbers $(\boldsymbol{\delta}^{(k)}, \boldsymbol{\gamma}^{(k)})$ are positive, all the matrices $H^{(k)}$ $(k = 1,2,\ldots)$ are positive definite, and the definition of $\lambda^{(k)}$ implies the equation

$$(\mathbf{g}^{(k+1)}, \boldsymbol{\delta}^{(k)}) = 0, \qquad (k = 1, 2, \ldots), \tag{6}$$

(Fletcher and Powell, 1963). Therefore, because $\boldsymbol{\delta}^{(k)}$ is a positive multiple of $-H^{(k)}\mathbf{g}^{(k)}$, we have the identity

$$(\boldsymbol{\delta}^{(k)}, \boldsymbol{\gamma}^{(k)}) = \|\boldsymbol{\delta}^{(k)}\|(\mathbf{g}^{(k)}, H^{(k)}\mathbf{g}^{(k)})/\|H^{(k)}\mathbf{g}^{(k)}\|. \tag{7}$$

The determinants of the matrices $H^{(k)}$ satisfy the recurrence relation

$$\det(H^{(k+1)}) = \det(H^{(k)}) \frac{(\boldsymbol{\delta}^{(k)}, \boldsymbol{\gamma}^{(k)})}{(\boldsymbol{\gamma}^{(k)}, H^{(k)}\boldsymbol{\gamma}^{(k)})}, \tag{8}$$

(Pearson, 1969).

The equation

$$\frac{1}{(\mathbf{g}^{(k+1)}, H^{(k+1)}\mathbf{g}^{(k+1)})} = \frac{1}{(\mathbf{g}^{(k)}, H^{(k)}\mathbf{g}^{(k)})} + \frac{1}{(\mathbf{g}^{(k+1)}, H^{(k)}\mathbf{g}^{(k+1)})} \tag{9}$$

holds, and, because $H^{(k)}$ is positive definite, it implies that the sequence $(\mathbf{g}^{(k)}, H^{(k)}\mathbf{g}^{(k)})$ $(k = 1,2,\ldots)$ decreases strictly monotonically (Wolfe, 1969, private communication).

Defining $\Gamma^{(k)}$ to be the inverse of $H^{(k)}$, we have the recurrence relation

$$\Gamma^{(k+1)} = \left(I - \frac{\boldsymbol{\gamma}^{(k)}\boldsymbol{\delta}^{(k)T}}{(\boldsymbol{\delta}^{(k)}, \boldsymbol{\gamma}^{(k)})}\right) \Gamma^{(k)} \left(I - \frac{\boldsymbol{\delta}^{(k)}\boldsymbol{\gamma}^{(k)T}}{(\boldsymbol{\delta}^{(k)}, \boldsymbol{\gamma}^{(k)})}\right) + \frac{\boldsymbol{\gamma}^{(k)}\boldsymbol{\gamma}^{(k)T}}{(\boldsymbol{\delta}^{(k)}, \boldsymbol{\gamma}^{(k)})}, \tag{10}$$

(Fletcher, 1970).

The trace of $\Gamma^{(k+1)}$ is related to the trace of $\Gamma^{(k)}$ by the equation

$$\mathrm{Tr}(\Gamma^{(k+1)}) = \mathrm{Tr}(\Gamma^{(k)}) + \frac{\|\mathbf{g}^{(k+1)}\|^2 - \|\mathbf{g}^{(k)}\|^2}{(\mathbf{g}^{(k)}, H^{(k)}\mathbf{g}^{(k)})} + \frac{\|\boldsymbol{\gamma}^{(k)}\|^2}{(\boldsymbol{\delta}^{(k)}, \boldsymbol{\gamma}^{(k)})}, \tag{11}$$

(Powell, 1971).

If $F(\mathbf{x})$ is uniformly convex, then the sequence $\mathbf{x}^{(k)}$ $(k = 1,2,\ldots)$ converges to the point at which $F(\mathbf{x})$ is least (Powell, 1971).

The condition (1) implies the inequality

$$\|\boldsymbol{\gamma}^{(k)}\| \leqslant M\|\boldsymbol{\delta}^{(k)}\|, \qquad (k = 1,2,\ldots). \tag{12}$$

This result is not special to the variable metric algorithm, but it follows from the definition of a derivative. It is proved, for instance, in Section 3.2.3 of Ortega and Rheinboldt (1970).

4. Consequences of a Conjecture

Although the variable metric algorithm usually works well, in this section we consider the conjecture that for some functions $F(\mathbf{x})$ there exists a positive constant c such that

$$\|\mathbf{g}^{(k)}\| \geqslant c, \quad (k = 1,2,\ldots). \tag{13}$$

We deduce a number of lemmas from inequality (13) that I hoped would deny the conjecture, but the truth of the conjecture is still an open question.

To prove these lemmas we introduce two more definitions. We let Δ be the diameter of the level set $\{\mathbf{x} | F(\mathbf{x}) \leqslant F(\mathbf{x}^{(1)})\}$, so we have the inequality

$$\|\boldsymbol{\delta}^{(k)}\| \leqslant \Delta, \quad (k=1,2,\ldots), \tag{14}$$

and we let m be the bound

$$m = \sup \|\mathbf{g}^{(k)}\|, \qquad (k = 1,2,\ldots). \tag{15}$$

This bound exists because of condition (1) and because the points $\mathbf{x}^{(k)}$ belong to a bounded set.

Lemma 1

There exist positive constants, c_1 and c_2 say, such that the trace of $\Gamma^{(k+1)}$ is bounded by the inequalities

$$c_1 \sum_{j=1}^{k} \frac{\|\boldsymbol{\gamma}^{(j)}\|^2}{(\boldsymbol{\delta}^{(j)}, \boldsymbol{\gamma}^{(j)})} \leqslant \mathrm{Tr}(\Gamma^{(k+1)}) \leqslant c_2 \sum_{j=1}^{k} \frac{\|\boldsymbol{\gamma}^{(j)}\|^2}{(\boldsymbol{\delta}^{(j)}, \boldsymbol{\gamma}^{(j)})}. \tag{16}$$

Proof

Equations (9) and (11) imply the identity

$$\operatorname{Tr}(\Gamma^{(k+1)}) = \operatorname{Tr}(\Gamma^{(k)}) + \frac{\|\mathbf{g}^{(k+1)}\|^2 - c^2}{(\mathbf{g}^{(k+1)}, H^{(k+1)}\mathbf{g}^{(k+1)})} - \frac{\|\mathbf{g}^{(k)}\|^2 - c^2}{(\mathbf{g}^{(k)}, H^{(k)}\mathbf{g}^{(k)})}$$

$$- \frac{\|\mathbf{g}^{(k+1)}\|^2 - c^2}{(\mathbf{g}^{(k+1)}, H^{(k)}\mathbf{g}^{(k+1)})} + \frac{\|\boldsymbol{\gamma}^{(k)}\|^2}{(\boldsymbol{\delta}^{(k)}, \boldsymbol{\gamma}^{(k)})}, \tag{17}$$

and we apply this equation k times to express $\operatorname{Tr}(\Gamma^{(k+1)})$ in terms of $\operatorname{Tr}(\Gamma^{(1)})$. Thus we obtain the relation

$$\operatorname{Tr}(\Gamma^{(k+1)}) = \operatorname{Tr}(\Gamma^{(1)}) + \frac{\|\mathbf{g}^{(k+1)}\|^2 - c^2}{(\mathbf{g}^{(k+1)}, H^{(k+1)}\mathbf{g}^{(k+1)})} - \frac{\|\mathbf{g}^{(1)}\|^2 - c^2}{(\mathbf{g}^{(1)}, H^{(1)}\mathbf{g}^{(1)})}$$

$$+ \sum_{j=1}^{k} \left(- \frac{\|\mathbf{g}^{(j+1)}\|^2 - c^2}{(\mathbf{g}^{(j+1)}, H^{(j)}\mathbf{g}^{(j+1)})} + \frac{\|\boldsymbol{\gamma}^{(j)}\|^2}{(\boldsymbol{\delta}^{(j)}, \boldsymbol{\gamma}^{(j)})} \right)$$

$$< \operatorname{Tr}(\Gamma^{(1)}) + \frac{\|\mathbf{g}^{(k+1)}\|^2 - c^2}{(\mathbf{g}^{(k+1)}, H^{(k+1)}\mathbf{g}^{(k+1)})} + \sum_{j=1}^{k} \frac{\|\boldsymbol{\gamma}^{(j)}\|^2}{(\boldsymbol{\delta}^{(j)}, \boldsymbol{\gamma}^{(j)})}, \tag{18}$$

the last line being a consequence of inequality (13). Now because the trace of a matrix is equal to the sum of its eigenvalues, and because $\Gamma^{(k+1)}$ is positive definite, we infer the inequality

$$\frac{\|\mathbf{g}^{(k+1)}\|^2}{(\mathbf{g}^{(k+1)}, H^{(k+1)}\mathbf{g}^{(k+1)})} < \operatorname{Tr}(\Gamma^{(k+1)}), \tag{19}$$

and therefore expression (18) implies the bound

$$\operatorname{Tr}(\Gamma^{(k+1)}) < \operatorname{Tr}(\Gamma^{(1)}) + \frac{\|\mathbf{g}^{(k+1)}\|^2 - c^2}{\|\mathbf{g}^{(k+1)}\|^2} \operatorname{Tr}(\Gamma^{(k+1)}) + \sum_{j=1}^{k} \frac{\|\boldsymbol{\gamma}^{(j)}\|^2}{(\boldsymbol{\delta}^{(j)}, \boldsymbol{\gamma}^{(j)})}, \tag{20}$$

which is equivalent to the condition

$$\operatorname{Tr}(\Gamma^{(k+1)}) < \frac{\|\mathbf{g}^{(k+1)}\|^2}{c^2} \left(\operatorname{Tr}(\Gamma^{(1)}) + \sum_{j=1}^{k} \frac{\|\boldsymbol{\gamma}^{(j)}\|^2}{(\boldsymbol{\delta}^{(j)}, \boldsymbol{\gamma}^{(j)})} \right). \tag{21}$$

Because $\|\mathbf{g}^{(k+1)}\|$ is bounded above, the right-hand inequality of expression (16) is proved.