Lecture Notes in Computer Science 769

Edited by G. Goos and J. Hartmanis

J. L. Nazareth

The Newton-Cauchy Framework

A Unified Approach to Unconstrained Nonlinear Minimization

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Series Editors

Gerhard Goos
Universität Karlsruhe
Postfach 69 80
Vincenz-Priessnitz-Straße 1
D-76131 Karlsruhe, Germany

Juris Hartmanis
Cornell University
Department of Computer Science
4130 Upson Hall
Ithaca, NY 14853, USA

Author

J. L. Nazareth
Department of Pure and Applied Mathematics, Washington State University
Pullman, WA 99164-3113, USA

CR Subject Classification (1991): G.1.6, J.1, J.2, J.6
1991 Mathematics Subject Classification: 49Mxx, 65Kxx, 90-08, 90Cxx

ISBN 3-540-57671-1 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-57671-1 Springer-Verlag New York Berlin Heidelberg

Printed in Germany

Typesetting: Camera-ready by author
45/3140-543210 - Printed on acid-free paper

Preface

Minimizing a nonlinear, multidimensional function $f(x)$, $x \in \Re^n$, where f is smooth and not necessarily convex, is a central problem of computational optimization. An understanding of the methods used to solve it is essential for anyone interested in computation for several reasons:

- Nonlinear minimization problems frequently arise in practice, their solutions either being of immediate interest or required at intermediate stages of a more complex calculation.

- Closely related problems, in particular, solving systems of nonlinear equations or nonlinear least squares data-fitting problems, can be posed as nonlinear minimization problems; alternatively, techniques used in minimization methods can be suitably adapted or specialized to solve such problems.

- Nonlinear ordinary differential equations, partial differential equations or optimal control problems, which are defined over function spaces, must eventually be discretized for solution on a computer. This leads, in turn, to finite-dimensional nonlinear equation-solving or minimization problems.

- Unconstrained minimization techniques form the backbone of methods for solving *constrained* minimization problems. In this regard, it is worth noting that the addition of constraints can *sometimes* render an optimization problem computationally *more* tractable. As an extreme case, suppose f is to be minimized subject to a set of $n - 1$ independent linear equality constraints along with finite lower and upper bounds on the n variables. Then the problem is equivalent to a unidimensional minimization on a line segment.

- Recent interior-point techniques for linear programming (LP), which now supplement Dantzig's simplex method, have moved computational LP away from its traditional base in combinatorial programming, and repositioned it, alongside linearly constrained nonlinear optimization, in the transition region between unconstrained nonlinear minimization on the one hand and nonlinearly constrained optimization on the other. Again, techniques of unconstrained minimization or related techniques of nonlinear equation-solving play a key role in these new interior-point LP methods.

Two classical methods for minimizing a nonlinear function are *Cauchy's method*, which uses a search direction of steepest descent, and *Newton's method*, which uses a search direction derived from a local quadratic approximating model obtained, in turn, from the Taylor expansion. More recently, during the digital computer era, there have been two further algorithmic breakthroughs. The *conjugate gradient method (CG)*, proposed by Hestenes and Stiefel in 1952, and the *variable metric method*, developed by Davidon during the years 1955-1959. The CG method was originally proposed for minimizing a strictly convex quadratic function or, equivalently, for solving a positive definite symmetric system of linear equations, and it was straightforwardly adapted to general nonlinear minimization by Fletcher and Reeves in 1964. Nowadays, CG-related is a generic name for a class of methods that require limited computer storage. The variable metric method was clarified and brought to the attention of the optimization community by Fletcher and Powell in 1963, and subsequently broadened and relabelled under the generic name *quasi-Newton*, incorporating the contributions of numerous researchers. Interestingly enough, Davidon's original paper developed the seminal underlying ideas in the setting of non-quadratic problems, but the important clarification and promulgation of Fletcher and Powell, which gained wide acceptance for the variable metric method, placed considerable emphasis on its theoretical properties on quadratic functions. Thus the historical development of *both* modern breakthroughs of computational nonlinear optimization occurred at the interface with computational linear algebra. More recently, computational nonlinear minimization has broken loose from these early moorings, and its methods can now be formulated quite independently of the historical context.

The literature on unconstrained nonlinear minimization is vast and there are several useful expository texts that discuss individual methods in detail, including, for example, Avriel [1976], Bazaraa, Sherali and Shetty [1993], Dennis and Schnabel [1983], Fletcher [1980], Gill, Murray and Wright [1981] and Luenberger [1984]. *Noticeably lacking, however, is a treatment that reveals the essential unity of the subject.* This is the central concern of our research monograph, namely, to explore the relationships between the main methods, to develop a unifying Newton/Cauchy framework and to point out its rich wealth of algorithmic implications. The monograph also makes a contribution to clarifying the notation of the subject, currently full of conflicts and contradictions, as well as the terminology of quasi-Newton methods, which currently resembles 'alphabet soup', to quote Dennis and Schnabel [1983]. We concentrate on *basic conceptual methods* rather than on algorithmic variants and implementational details, and in the interests of brevity, we try to keep to a minimum the duplication of material that is widely available in the literature and the texts previously cited.

The monograph consists of six main chapters organized as follows:

Chapter 1 introduces the conjugate gradient and quasi-Newton methods within the context of their historical development, namely, convex quadratic minimization. The methods assume a particularly elegant form in this setting and have interesting algebraic properties. To neglect this aspect is to turn one's back on some of most attractive results of the subject, as well as the important interface with computational linear algebra. Our development follows a logical and orderly progression of ideas, highlighting the small set of basic principles that give rise to the methods. However, we do not then, according to historical precedent, generalize the methods in patchwork fashion so that they can be applied to nonquadratic problems. Our objectives in this chapter are much more limited, namely, *to motivate the methods and demonstrate their properties in a particular simple setting.*

In Chapter 2, we begin afresh in the setting of general nonlinear functions. Our starting point for development is the classical steepest-descent *metric-based* method of Cauchy. We follow an orderly progression of ideas leading, in turn, to Davidon's variable metric, the conjugate-gradient metric and the Newton metric, and highlighting the relationship between them.

In Chapter 3, our starting point for development is the classical *model-based* Newton's method. Again, a logical train of ideas lead, in succession, to the quasi-Newton model, the CG-related model and the Cauchy model.

Computational unconstrained nonlinear optimization comes to life from a study of the interplay between the metric-based (Chapter 2) and model-based (Chapter 3) points of view, with the motivating development of Chapter 1 in the background to lend added dimension. This is the topic of Chapter 4, which ties together the preceding three chapters, develops the Newton/Cauchy framework and indicates its rich array of algorithmic implications.

Chapter 5 discusses the basic conditions for establishing global convergence of *implementable algorithms* derived from the Newton/Cauchy framework and the important idea of *hierarchical implementation* of optimization methods.

Finally, Chapter 6 overviews nonlinear unconstrained optimization technology, namely, the wide variety of implementable mathematical and numerical algorithms that can be derived from the basic conceptual methods of the Newton/Cauchy framework.

Chapters 1, 2 and 3 are each largely self-contained and develop all the necessary expository detail. On the other hand, Chapters 4, 5 and 6 are more concise and collectively provide a free-standing structured guide to the unconstrained optimization literature, to which they make extensive reference. They also highlight some important unexplored territory.

The monograph assumes the following background on the part of the reader:

- basic multivariate calculus; see, for example, Rudin [1976],

- the basic factorizations of computational linear algebra; see, for example, Golub and Van Loan [1989] or Watkins [1991],

- elementary convex analysis, in particular, the characterization of convex sets and differentiable convex functions; see, for example, Avriel [1976] or Luenberger [1984],

- the basic algorithms for *one-dimensional* minimization and solving a nonlinear equation of a *single* variable; see any introductory numerical analysis text, for example, Kahaner, Moler and Nash [1989],

- and the basic theoretical characterization of optimal points, in particular, definitions of local and global optima, necessary and sufficient optimality conditions for smooth unconstrained functions and some basic exposure to the Lagrangian/Karush-Kuhn-Tucker optimality conditions for constrained optimization. These topics are widely treated in several excellent texts, for example, Avriel [1976], Luenberger [1984], or Zangwill [1969], and their duplication here would be pointless. In this regard, it is also worth noting explicitly that as a result of the repositioning of linear programming mentioned previously, it is increasingly likely that newcomers to optimization algorithms will be introduced to the subject via the methods of unconstrained nonlinear minimization rather than the more traditional combinatorial-based linear programming and the simplex method, and that the next generation of textbooks on computational linear and nonlinear optimization will adopt the following pattern of exposition: 1. Optimality conditions for both unconstrained and constrained optimization. 2. Unconstrained minimization methods. 3. Methods for linearly constrained problems including linear programming. 4. Methods for nonlinearly constrained problems.

The monograph is addressed to a broad spectrum of practitioners, researchers, instructors, and students, and we hope that it proves to be both useful and a refreshing new prespective on computational nonlinear optimization. For pedagogical purposes, it can be used to *supplement* one of the optimization texts cited earlier. It can also be used as the primary text of a graduate research seminar course when the instructor fills in background material as needed, fleshes out Chapters 4 through 6 through assigned readings in the optimization literature, and encourages students to explore uncharted territory via appropriate research projects.

Finally, it is a pleasure to thank various people who have contributed directly or indirectly to this effort. My thanks to Bill Davidon, whose algorithmic genius has always been a source of inspiration. I thank my Optimization Group colleagues Bob Mifflin and Kuruppu Ariyawansa with whom I share many research interests, and also Mike Kallaher, former chairman of WSU's Department of Pure and Applied (and, in all but name, Computational) Mathematics, whose enlightened approach has given the department a secure balance on the three equal legs of the mathematical tripod. My affiliation at the University of Washington has been invaluable, and I have profited from interaction, in particular, with Jim Burke, Alan Goldstein, Bob O'Malley, Terry Rockafellar and Paul Tseng. I thank Brian Smith, Laura Kustiner, and Miguel Gomez - students in the Fall '92 graduate course on computational nonlinear optimization, who provided useful feedback when some of this material was formulated and presented.

And finally, on a personal note, I thank my wife Abigail for her encouragement when the task of writing grew burdensome, as it always does, even for a relatively short monograph such as this.

Bainbridge Is., WA, 1993 J.L. Nazareth

Table of Contents