Nonlinear Programming

Theory, Algorithms, and Applications

GARTH P. McCORMICK
The George Washington University
Washington, D.C.

A Wiley-Interscience Publication
JOHN WILEY & SONS
New York • Chichester • Brisbane • Toronto • Singapore

Copyright © 1983 by John Wiley & Sons, Inc.

All rights reserved. Published simultaneously in Canada.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, John Wiley & Sons, Inc.

Library of Congress Cataloging in Publication Data:

McCormick, Garth P.
Nonlinear programming.

"A Wiley-Interscience publication."
Includes index.
1. Nonlinear programming. I. Title.
T57.8.M39 1982 519.7'6 82-16046
ISBN 0-471-09309-2

Printed in the United States of America

10 9 8 7 6 5 4 3 2

To Ginny first,

then to

Jessica

Rachel

Callie

and John

Preface

The primary purpose of this book is the exposition of algorithms for solving nonlinear programming (optimization) problems. This book also contains the theory of nonlinear optimization, mostly concerned with the characteristics of optimal points. An important subsidiary aim is the presentation of some optimization models of real world problems that can be solved by nonlinear programming methodology. After working in the field of algorithm development for several years, I was disturbed by the number of bad mathematical models of serious problems that were considered in the literature. In 1972 I initiated a course, "Applications of Linear and Nonlinear Optimization Theory," in the Department of Operations Research at the George Washington University in order to teach mathematical modeling, that is, how to do it and its limitations. This is the origin of the discussion in Chapter 1 of characteristics of real world problems that prevent their representation as optimization models. The students are given a structured way in which to analyze application papers and to assess their relevance in solving the underlying problems. The material in Chapter 1, the models that precede each major section of the book, and papers gathered from the published literature constitute the subject matter of the applications course.

The bulk of the material can be used for a traditional course in nonlinear programming. There is a rigorous development of first- and second-order optimality conditions. Algorithms are presented for optimization problems in order of increasing complexity, that is, for single-variable problems, unconstrained n-variable problems, linearly constrained problems, and, finally, nonlinearly constrained problems. For the latter two categories, the components of the algorithms are presented in general and then it is shown how classical methods fit into the general framework. This is a departure from the usual way of presenting algorithms—by author, with each subalgorithmic detail treated as though it were intrinsic to the method, whereas in fact it is usually just a replaceable part.

Some of the newer algorithms, multiplier (augmented Lagrangian) methods, exact penalty function methods, and sequential linearization techniques are explained with the use of an idealized exact penalty function created by the movement of a particle under different forces. It is my contention that many algorithms can be motivated through the use of simple physical models, and this approach is used in the development of Newton's method for unconstrained and constrained optimization problems.

This book contains several topics not usually covered in books on nonlinear programming. The future of nonlinear programming lies in the integration of numerical techniques for solving systems of equations. Some preliminary attempts to do this and the necessary background for understanding the current literature are contained here.

There is no lack of algorithms for solving nonlinear programming problems, as evidenced by the material in this book and in its references. Probably the single most important thing, in addition to the primitive nature of the use of matrix methods, preventing the automatic solution of nonlinear programming problems is the lack of a computationally oriented way of representing nonlinear functions. Chapter 3 is a summary of a new approach using "factorable" functions to provide the interface between computer-coded algorithms and the algebraic representation of nonlinear programming problems. This new approach has a great variety of implications for other aspects of nonlinear programming. One of these is evidenced in Chapter 18 on solving nonconvex programming problems.

The field of nonlinear programming is vast, and no attempt has been made to include all the developments to date or in progress. It would require a book of several volumes to do this. In some cases, long proofs of important theorems are omitted. Many of the topics covered are included because of my special research interests. Basic or pioneering works are mentioned at the ends of chapters, where a synthesis of the many contributing works is contained.

Much of the research that contributed to this book was sponsored by the Army Research Office, the Office of Naval Research, the Air Force Office of Scientific Research, the Department of Energy, and the National Bureau of Standards. Their help is appreciated.

Many individuals have contributed to portions of the book, and their contributions are noted at the appropriate places. I wish to thank Professor William H. Marlow, Director of the Institute for Management Science and Engineering, for creating an environment in which basic and applied research could be carried out.

Professor Stephen M. Robinson was particularly helpful in the early stages of this book: the material relating to the classical Newton's method in Chapter 7 and several other contributions are due to him and are greatly appreciated.

GARTH P. MCCORMICK

Washington, D.C.
December 1982

此为试读，需要完整PDF请访问：www.ertongbook.com

Use of Cross Referencing in This Book

Within each section important items, such as theorems, equations, and definitions, are numbered sequentially. The numbers of the chapter and section appear at the top of each page; thus 6.3 at the top of a page indicates that the material appearing there is from Section 3 of Chapter 6. When reference is made to an item within the same section, only the item number is used. If the referenced item is in the same chapter but in a different section, both the section and item numbers are used; thus **[7.5]** means item 5 of Section 7 of the same chapter. When the item referenced is in a different chapter, all three numbers are used; thus equation **[2.7.5]** indicates equation 5 of Section 7 of Chapter 2.

Notation

x	$(x^1, \ldots, x^n)^T$, an $n \times 1$ vector of variables
E^n	Euclidean n space
$f'(x_0)$	The $1 \times n$ derivative vector of the function f evaluated at x_0, that is, the $1 \times n$ vector whose jth component is $\partial f(x_0)/\partial x^j$
$\nabla f(x_0)$	The $n \times 1$ gradient vector of f at x_0, that is, the $n \times 1$ vector whose jth component is $\partial f(x_0)/\partial x^j$ [note that $f'(x_0) = \nabla f(x_0)^T$]
$f''(x_0)$	The $n \times n$ Hessian matrix of f at x_0, that is, the matrix whose i, jth element is $\partial^2 f(x_0)/\partial x^i \, \partial x^j$
$\nabla^2 f(x_0)$	The $n \times n$ Hessian matrix defined above
$A^{\#}$	A pseudo-inverse of the matrix A $(AA^{\#}A = A)$
A^{+}	The Penrose–Moore generalized inverse of the matrix A

Contents

I BASICS

Chapter 1 The Nature of Optimization Problems 3

1.1 General Remarks, 3
1.2 Characteristics of Optimization Problems, 9
1.3 Mathematical Structure of Optimization Problems, 21
Problems, 24

Chapter 2 Analytical Background 28

2.1 Fundamentals of Analysis, 28
2.2 Convexity, 38
2.3 Optimality Conditions for Unconstrained Problems, 56
2.4 Matrix Theory, 60
Problems, 71

Chapter 3 Factorable Functions 76

3.1 Computationally Oriented Way of Representing Nonlinear Functions, 76
3.2 First and Second Derivatives, 78
3.3 Use in Matrix Methods, 84
3.4 Other Uses of Factorable Functions, 87
Problems, 88

II UNCONSTRAINED PROBLEMS

Chapter 4 Unconstrained Optimization Models 93

4.1 Parameter Estimation, 93

4.2 Application to Solutions of Nonlinear Equations, 104
Problems, 106

Chapter 5 Minimizing a Function of a Single Variable **109**

5.1 Newton's Method, 109
5.2 Methods Using First Derivatives, 116
5.3 Methods That Use Only Function Values, 120
5.4 Modified Method of Golden Sections, 125
Problems, 129

Chapter 6 General Convergence Theory for Unconstrained Minimization Algorithms **132**

6.1 Step-Size Procedures, 132
Problems, 141

Chapter 7 Newton's Method With Variations **143**

7.1 A Cauchy–Newton Approach to Unconstrained Minimization, 143
7.2 Local Convergence and Rate of Convergence Results for Newton's Method, 147
7.3 Global Convergence Using Directions of Negative Curvature, 149
7.4 General Algorithm for Computing an Estimate of the Positive Part and Directions of Nonpositive and Negative Curvature When the Hessian Matrix Is Given in Dyadic Form, 153
7.5 Linear Algebra Required for UU^T Updating, 157
7.6 Specific Realization of an Algorithm for Computing the Positive Part and Directions of Nonpositive Curvature, 160
7.7 Convergence of the Modified Newton Method, 161
7.8 Survey of Modified Newton Algorithms, 163
Problems, 167

Chapter 8 Conjugate Direction Algorithms **169**

8.1 Motivation and General Statement, 169
8.2 The Conjugate Gradient Method, 171
8.3 The Projection Method, 176
Problems, 178

Chapter 9 Quasi-Newton Methods 180

9.1 The Secant Method for Minimizing Functions of Several Variables, 180
9.2 Methods That Save Old Information, 185
9.3 Single-Step Convergence Rate, 199
Problems, 202

III OPTIMALITY CONDITIONS FOR CONSTRAINED PROBLEMS

Chapter 10 First- and Second-Order Optimality Conditions 207

10.1 Preliminary Results, 207
10.2 First-Order Necessary Conditions, 210
10.3 Second-Order Necessary Conditions, 218
10.4 Second-Order Sufficiency Conditions, 222
10.5 Nonrigorous Development of the Optimality Conditions, 227
Problems, 236

Chapter 11 Applications of Optimality Conditions 240

11.1 Algorithmic Motivation, 240
11.2 Rate of Convergence of Algorithms, 243
11.3 Sensitivity Analysis, 244
Problems, 249

IV LINEARLY CONSTRAINED PROBLEMS

Chapter 12 Models with Linear Constraints 253

12.1 Chemical Equilibrium, 253
12.2 An Inventory Problem, 254

Chapter 13 Variable-Reduction Algorithms 259

13.1 Introduction, 259
13.2 Null-Space Computation, 264
13.3 Subspace Determination, 270
13.4 Termination Strategies, 272
13.5 General Algorithm for the Linearly Constrained Problem, 276
13.6 Matrix Methods for Linearly Constrained Problems, 282
Problems, 289

V NONLINEARLY CONSTRAINED PROBLEMS

Chapter 14 Models with Nonlinear Constraints **297**

14.1 Engineering Design, 297
14.2 Water Pollution Control, 299

Chapter 15 Direct Algorithms for Nonlinearly Constrained Problems **308**

15.1 Introduction, 308
15.2 Derivation of a Natural Arc Method, 309
15.3 Variable-Reduction Methods for the Nonlinear Equality-Constrained Problem, 314
15.4 Complete General Algorithm for the Nonlinearly Constrained Problem, 316
Problems, 327

Chapter 16 Sequential Unconstrained Minimization Techniques **329**

16.1 Logarithmic Barrier Function Method, 329
16.2 Quadratic Loss Penalty Function Method, 333
16.3 Extrapolation and Duality, 335
16.4 Convergence Theory for Barrier and Penalty Functions, 342
16.5 An Idealized Exact Penalty Function, 346
16.6 Exact Penalty Function Methods, 358
16.7 Multiplier Methods, 366
Problems, 370

Chapter 17 Sequential Constraint Linearization Techniques **374**

17.1 Derivation of a Class of Sequential Linearly Constrained Problems, 374
17.2 Extensions to Inequality Constraints, 381
Problems, 383

VI OTHER TOPICS

Chapter 18 Obtaining Global Solutions **387**

18.1 Introduction, 387
18.2 Outline of the Global Solution Method, 388
18.3 Procedure for Creating Convex Underestimating Programs, 389
18.4 Local/Global Verification, 403
Problems, 414

此为试读，需要完整PDF请访问：www.ertongbook.com

Chapter 19 Geometric Programming **417**

19.1 Primal and Dual Geometric Programming Problems, 417
Problems, 424

References **425**

Author Index **435**

Subject Index **439**

I

Basics

Chapter 1 has enough material to provide the basis for a course in the applications of nonlinear programming. An introduction to simple nonlinear programming problems using as a basis the classical diet problem begins the book in Section 1.1. Section 1.2 contains a discussion of the characteristics that must be present in real world problems in order for them to be legitimately modeled as optimization problems. In Section 1.3 is an elementary discussion of the mathematical characteristics of optimization problems. With this background a moderately sophisticated reader should be able to understand published nonlinear programming models and much of the algorithmic and theoretical material in this book.

In order to provide the reader with a complete understanding of the proofs of theorems and some of the more complicated algorithms, Chapter 2 contains the necessary mathematical background: functional analysis, convexity, and linear algebra.

Chapter 3 contains a discussion of factorable functions. This is a new topic oriented toward providing a computationally valuable way to represent complicated functions of several variables.

1

The Nature of Optimization Problems

1.1 GENERAL REMARKS

The study of nonlinear programming begins with algebra. It begins with the notion that it is useful to express a relationship between measurable attributes symbolically rather than writing down all or some of the possible associated values. The relationship between the area of a triangle and its height and base can be expressed by the examples in Table 1.1.1.

Use of algebra simplifies this task by providing a shorthand way of expressing the relationship. Let A denote the area of a triangle, let x denote the base length, and y the height. Then $A = xy/2$.

This is an example of an explicit functional relationship. Such relationships are useful as models of the real world and are helpful in making decisions. The question of how many pounds of grass seed to buy for a triangular plot with a base of 20 ft and a height of 30 ft can be answered by computing the area and knowing the recommended coverage per square foot.

Even this simple mathematical model of the situation is subject to criticism. No plot of ground is actually triangular and slope considerations have not been taken into account. The uneven surface of the ground may require more grass seed in some places, less in others—better to get an experienced gardener, let him look at it, and buy the quantity he suggests. In the absence of an experienced gardener, though, most people would accept the basic model as an aid in deciding on how much grass seed to buy.

Indeed, all mathematical models can be challenged. It is not difficult to present arguments stating that models oversimplify the complexity of the real situation, both with respect to the operating laws of cause and effect and also their inclusiveness; still it is useful in many cases to abstract the important elements of a real world problem and model them. The hard question for the model builder is whether or not the *important* elements of a real world problem can be quantified.

Mathematical model building is not an activity that is carried on by just a few abstractly oriented persons. Most people implicitly develop mathematical

Table 1.1.1 Associated Values of the Area, Height, and Base Length of a Triangle

Length of Base (in.)	Height of Triangle (in.)	Area of Triangle ($in.^2$)
0.2	7.0	0.7
5.6	10.0	28.0
1.25	4.4	2.75

models and make decisions based on rather simple assumptions. Questions of how long to cook a roast, how much fertilizer to apply to a crop, how long it will take to drive a car from point A to point B, at what speed to take a picture, and the number of calories to eat in a day are all answered in some form by the use of a mathematical model.

One difference between a good and a bad model is whether or not the proper functional relationships are used. Consideration of the rate of heat transfer and the shape of a roast implies that the cooking time should not vary directly with the number of pounds. The average speed of a car on a trip is not the average of the speeds on its segments. A crop's yield per acre as a function of the fertilizer applied follows an exponential form.

Mathematical models arise from a variety of sources and needs. Tolstoy resorted to the theory of integral calculus to explain how the aggregate of Russian souls could combine to defeat Napoleon.

Mathematical model building has often come under attack, but like the praise of its advocates, the attacks are many times undifferentiated. Obviously in some circumstances model building is relevant, and working with a model can help one predict, suggest controls, or simply visualize phenomena that would not be possible to understand intuitively from the mass of detail in the real world situation. In other cases the models are constructed (it is sometimes felt) to avoid looking at the relevant complexities; these models fail to describe the real world situation accurately. There are many situations that are impossible to model and where functional relationships do not exist. In this book critical questions will be asked whenever models are presented, whether or not development of the model is warranted. The reader should be critical, but should also look for situations where models are warranted.

As a simplification, it can be stated that optimization problems are posed in one of two modes: the descriptive or the prescriptive (normative). As a simple example of the former and to illustrate in concrete terms what an optimization problem is, consider the following.

[1] Example. Among all the triangles with a base of 8 in. and an area of at least 12 $in.^2$, find that one with the smallest perimeter.

In setting up an optimization problem the "variables" or "unknowns" of the problem must be specified. Often there is a choice among the quantities to consider as variables. In this example enough variables are needed to completely specify a triangle. Since the base is fixed, knowing the values of any two other independent quantities of the triangle will specify it. Let x denote one side of the triangle and let y denote the height. These are the *variables* of the optimization problem. Other quantities that are fixed inputs to the problem are called *parameters*.

There are an infinite number of values of x and y for which the associated triangle has an area of at least 12 in.^{2}. It is intuitively clear that for some values, call them $\bar{x}$ and $\bar{y}$, the perimeter of the triangle is smallest. The requirement that the triangle have an area of at least 12 in.^{2} is stated algebraically as $8y/2 \geqq 12$ or, more simply, $y \geqq 3$. From Figure 1.1.1 arises another natural constraint, that $x \geqq y$ hold.

A functional relationship must be derived by relating the problem variables and parameters to the perimeter of the triangle. From two applications of the Pythagorean theorem (referring again to Figure 1.1.1) the length of the third side is

$$\left\{ y^2 + \left[8 - (x^2 - y^2)^{1/2}\right]^2\right\}^{1/2}.$$

The perimeter to be minimized is then

$$8 + x + \left\{ y^2 + \left[8 - (x^2 - y^2)^{1/2}\right]^2\right\}^{1/2}.$$

Combining this information results in the following optimization problem: find values $(\bar{x}, \bar{y})$ that

[2] $$\underset{(x,y)}{\text{minimize}} f(x, y) = 8 + x + \left\{ y^2 + \left[8 - (x^2 - y^2)^{1/2}\right]^2\right\}^{1/2},$$

subject to the constraints that

[3] $$y \geqq 3$$

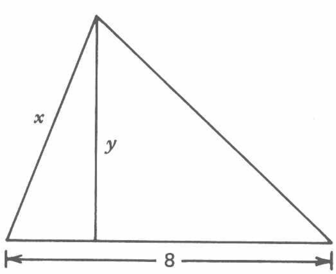

Figure 1.1.1 Geometry of triangle example.

此为试读，需要完整PDF请访问：www.ertongbook.com