Principles of Optimization Theory

C R Bector
S Chandra
J Dutta

Alpha Science

Principles of

Optimization Theory

C R BECTOR
S CHANDRA
J DUTTA

Alpha Science International Ltd.
Harrow, U.K.

C R Bector
Asper School of Business
Faculty of Management
University of Manitoba
Winnipeg, Canada

S Chandra
Department of Mathematics
India Institute of Technology Delhi
New Delhi, India

J Dutta
Department of Mathematics
India Institute of Technology Kanpur
Kanpur, India

Alpha Science International Ltd.
Hygeia Building, 66 College Road,
Harrow, Middlesex HA1 1BE, U.K.

Printed from the camera-ready copy provided by the Authors

ISBN 1-84265-166-8

Printed in India

Principles of

Optimization Theory

DEDICATED
TO
Late Professor J. N. Kapur
Professor B. D. Craven
and
Professor B. Mond

Foreword

Optimization is at the heart of applied mathematics. Scientists seem to agree that nature's behavior is always optimal in one way or another, so that, for instance, theoretical physics is classically based on the so called variational principles, which are a precise mathematical formulation of this fundamental idea. More recently, in the twentieth century economic theory also laid its foundations on the basis that economic agents optimize: consumers maximize their utility functions, firms are profit maximizers, Thus optimization has the foremost importance to help us understand the laws of nature and even how society behaves. Besides its prominent role in theoretical science, optimization is an essential tool for the formulation and solution of many problems arising in applied sciences like, e.g., engineering, finance, management science or operations research. Last but not least, optimization is interesting in itself from a purely mathematical point of view, as it contains beautiful theories and results and offers us challenging open problems of a theoretical nature as well as of an algorithmic or numerical character.

Despite optimization problems having been present in mathematics since very early times, optimization theory has been settled as a solid and autonomous field only in relatively recent times. A systematic study of optimization theory began in the eighteenth century, when the calculus of variations started to develop and to be used to formulate the fundamental principles of physics. The calculus of variations and the theory of optimal control theory, its modern extension, is nowadays studied in the framework of infinite dimensional function spaces, and so they are out of the scope of this book, as it only deals with spaces of finite dimension. This allows the authors to avoid the difficulties and technicalities that infinite dimensions bear, thus making the book more accessible to the beginner, who can later on take profit of the insight acquired in the more simple finite dimensional setting to study the much more complex infinite dimensional one. Moreover, finite dimensional spaces are the natural framework for mathematical programming, the branch of optimization that started to develop in the twentieth century tossed by its applications to operations research and economics.

This book is written by three very active researchers, whose important contributions to optimization theory mostly concentrate in the topics of generalized convexity and nonsmooth analysis. However they have only included the essential tools for optimization that these subfields provide, thus avoiding misleading the beginners by presenting more specialized concepts and techniques that are unnecessary at this introductory level. In this way they have succeeded to make a friendly presentation, which constitutes an excellent starting point for further reading. As reflected by the title, the book deals with the fundamental principles: convex analysis, convex optimization, optimality conditions and duality theory. But it also contains two chapters on more specialized topics, namely, nonsmooth optimization and monotone maps, which make the book more advanced than most standard introductory textbooks. A particularly interesting feature of these chapters is that they include some material on generalized convexity and generalized monotonicity. For instance, in the chapter on nonsmooth analysis the authors relate the notions of quasiconvexity and pseudoconvexity, in the case of locally Lipschitz functions, to Clarke generalized gradients. This is nicely complemented in the last chapter, where the concepts of quasimonotonicity and pseudomonocity for operators are introduced and used to characterize quasiconvexity and pseudoconvexity, respectively, in terms of gradients.

Undoubtedly, this book will be very useful for graduate students in applied mathematics, economics, engineering and operations research and, more generally, for anybody wishing to learn the essential principles of optimization theory. It will also constitute a helpful reference for users of optimization and researchers in this field. I therefore welcome this nice contribution to the existing literature on optimization theory.

Juan Enrique Martinez-Legaz
Barcelona, Spain

Preface

Optimization is one of the most important branches of modern applied mathematics. Various problems arising in the areas of engineering design, economic theory, physics, chemistry, business and management science call for minimizing or maximizing functions. Optimization can be thus described as the art of minimizing or maximizing functions. Naturally there are two major aspects of optimization namely theory and computation. The importance of optimization in various applications has been possible by the development of robust algorithms. However these algorithms are in effect generated from a beautiful mathematical theory of optimization. In this book we deal with various aspects of optimization theory. In the last three decades of the 20th century there has been rapid progress in optimization theory. A large number of good books and monographs have been written in various areas like convex optimization, nonsmooth optimization, semi-infinite optimization, etc, and some of them have even become classics. However to the best of our knowledge there are hardly any text which covers the fundamentals of these various aspects of optimization theory in a compact way so that an advanced masters degree student or a beginning graduate student gets a good feel about the broad area. Keeping this in view we decided to write a text on optimization theory which covers most of the important areas yet present them in a compact fashion. However in doing so we did not intend to compile a loose string of definitions, lemmas and theorems but rather take a more conversational style in order to emphasize the fundamental facts. On the other hand we also make a sincere effort to blend some ideas from current research into the text so that a more research oriented reader gets some ideas of the main areas of research.

We begin by discussing some basic facts regarding the existence of optimal points and then move on to the fundamentals of convex analysis which indeed lies at the heart of optimization theory. In this book we have laid a great importance on optimality conditions in both smooth and nonsmooth settings. It is important even for the application oriented reader to have a good idea of optimality conditions since they are pivotal in developing good algorithms to solve optimization problems. We also discuss the other important aspects like duality theory and monotone maps.

There are several starred section in this book. The beginner can skip those sections during the first reading without hampering the logical flow and come back to these sections at a later stage. There are lot of exercises scattered throughout the book and most of them extend some of the important ideas. We encourage the reader to try them out in order to get a better understanding. The reader can also find a set of problems after the end of each chapter beginning with Chapter 2.

We believe that this book will be helpful to M.Sc. and Ph.d students in the area of mathematics, economics and operations research as well as advanced B. Tech and Ph.d students in engineering sciences in the Indian Universities. However we also feel that this book can be used a text or a reference in a graduate course in optimization theory in any university in the world.

Though we have made our best efforts to make the presentation simple and error free some errors may still remain and we hold ourselves responsible for that. We would be grateful if readers can intimidate us regarding the errors by emailing us at jdutta@iitk.ac.in(email address of J. Dutta).

We would feel happy if the students for whom this book is intended finds it useful.

C R Bector (Winnipeg)
S Chandra (New Delhi)
J Dutta (Kanpur)

Acknowledgements

In the long process of writing this book we have been encouraged and helped by many individuals. We would first like to thank Professor J. E. Martinez-Legaz for writing a forward to this book. We are also thankful to Professor A. M. Rubinov for providing some constructive suggestions and Professor Boris Mordukhovich for providing some of his early works in variational analysis. We are grateful to Dr. C. S. Lalitha, Dr. A. Mehra, S. Sandhu and A. Bhattacharya for reading parts of the manuscripts and suggesting various improvements. We would also like to thank Professor K. Deb for providing some interesting problems on optimality conditions. Further we thank Reshma Khemchandani for drawing all the figures in this book and we are also thankful to M. Sajid for his tremendous help during the final preparation of the manuscript in LATEX. We also thank our respective departments for providing all the facilities while this book was written. Most importantly we are grateful to our families for their patience during this period.

List of Symbols

$\mathbb{R}^n$: n-dimensional Euclidean Space.

$\langle .,. \rangle$: inner product or dot product.

$\|x\|$: norm of a vector x.

∇f : gradient of a function f.

$\nabla^2 f$: Hessian of a function f.

$\partial f(x)$: Subdifferential of a convex function f at x.

$\partial^c f(x)$: The Clarke subdifferential of f at x.

$f'(x, v)$: The directional derivative of f at x in the direction v.

$B_r(x)$: An open ball of radius r centered at x.

$B_1(0)$: The unit ball.

Contents

List of Figures

Chapter 1

INTRODUCTION AND BASIC FACTS

1.1 Optimization : A Little History

Optimization theory is one of the most lively and exciting branches of modern mathematics. There is hardly any area of human progress where optimization has not made itself useful. This can be summed up in the words of the great mathematician L. Euler who once remarked the following-Nothing takes place in the world whose meaning is not that of some maximum or minimum. The sophisticated theory of optimization that we see today has its root on antiquity. There had been a wealth of geometrical and mechanical problems which are in fact optimization problems but has been solved by using geometrical knowledge since a general framework for solving optimization was gradually becoming clear in the later half of the 17th century. For example, we provide below, the following ancient optimization problem.

Let A and B be two points on the same side of a line l. Find a point D on l such that the sum of the distance of A and B from D is minimum.

There is in fact a legend associated with one of the oldest problems in optimization theory which dates back to the 9th century. This problem, known as the **Dido's Problem**, derives its name from the Phoenician princess Dido. Fleeing from the persecution from her brother Dido set off westward along the Mediterranean shore in search of haven. A certain point, which is now called the bay of Tunis, caught her attention. She struck a deal with the local leader

Yarb for the sale of land. She made a small demand of having as much land as can be encircled in a bulls hide. Dido then cut a bull's hide into narrow strips, tied them together and encircled a large piece of land. On it she built a fortress and the city of Carthage where she died a martyr's death.

The problem of Dido can be rephrased in the language of modern mathematics as follows. *Among all plane closed curves of a given length find the one that encloses the maximum area.* This problem also known as the *isoperimetric problem* , has a simple answer, that such a curve is a circle. However the methods to solve this and other similar problems in a general fashion evolved only in the 18th and 19th century in the form of a subject called the *Calculus of Variations* . In the ancient days geometry was indeed the main tools to solve this sort of problems. Thus, in the ancient days each problem was solved by a specific method. By the mid 17th century it was clear that one needed to have a general method to solve a maximum and minimum problem. This need arose simultaneously with the development of the notion of a function. It was Fermat who first developed the still used general method. He gave a method which inherently used the notion of the derivative that was still not formally discovered. Though Fermat presented the case for polynomial functions, yet his method can be used for any differentiable functions. In modern language if a function $f : \mathbb{R} \to \mathbb{R}$ achieves a maximum or minimum at x_0 then $f'(x_0) = 0$. It was around 1684 Leibniz established the necessary condition for a minimum and maximum and also developed the method to distinguish between a maximum and a minimum via second derivatives. *Acta Eruditorum* is the first scientific journal of the world which started publications in 1682. In 1696 the famous Swiss mathematician Johann Bernoulli posed his famous *Brachistochrone Problem* in Acta Eruditorum that can be stated as follows. *Given two points A and B in the vertical plane. Assume that a particle is allowed to move under its own gravity from A to B. What should be the curve along which the particle should move so as to reach B from A in the shortest time ?*. This problem even attracted the genius of Isaac Newton. Bernoulli gave this problem to his student Leonard Euler who solved the problem by developing a general scheme that led to the rise of a subject called the *Calculus of Variations* that has been applied to many physical and engineering problems. In 1759 a young French mathematician called J. L. Lagrange gave a completely different and novel approach to solve the problems in Calculus of Variations. This method, now known as the *Lagrange Multiplier rule*, is one of the main cornerstones of optimization theory. We devote large part of this small monograph to the study of Lagrangian Multiplier rules for various classes of optimization problems.

It was during the World War II that a new chapter was written in the history of optimization with the development of *Linear Programming*. George B. Dantzig developed his famous *Simplex Method* for solving the linear programming problem during the mid 1940's. Though initially applied for war problems it was soon found that Linear Programming was very useful to solve problems in economics, business and engineering sciences. In 1951 H. W. Kuhn and Albert. W. Tucker developed the Lagrangian multiplier rule for convex and other nonlinear programming problems which also involved inequality constraints. The Kuhn-Tucker optimality conditions became very useful and important in developing algorithms for solving convex and other nonlinear programming problems with differentiable functions. It was later found that in 1939 W. Karush developed optimality conditions similar to that of Kuhn and Tucker. Thus those conditions are now known as the Karush-Kuhn-Tucker conditions. The development of Convex Analysis as a subject in itself is another important factor responsible for the development of the subject of optimization. W. Fenchel, J. J. Moreau and R. T. Rockafellar are the key contributors to the development of convex analysis and convex optimization. A radically new development of optimization theory took place with a growing realization that many applied problems lack the important property of differentiability. In fact many convex progamming problems belong to this category. In early 1970's F. H. Clarke coined the term *Nonsmooth Optimization* to identify the part of optimization theory that deals with non-differentiable problems and he himself made a path-breaking contribution in this area. After the second World War, apart from the theoretical aspect of optimization, there has been tremendous development in the field of computational methods for solving optimization problems. This development was definitely fueled by the development of the simplex method for linear programming. The emphasis shifted to non-linear programming by the development of a powerful method for unconstrained optimization due to W. C. Davidon , a physicist at the Argonne National Laboratory, U.S.A. in the late 1950s. Though the work of Davidon was rejected initially and it remained a research report yet gave rise to large number of work in this area and was finally published in the first issue of the *SIAM Journal of Optimization*. The work of Davidon was further carried forward by M. J. D. Powell and R. Fletcher leading to the rise of powerful Quasi-Newton Methods. See Fletcher [19] for more historical development of the practical methods of optimization. In 1984 N. Karmarkar developed a polynomial time algorithm for solving a linear programming problem and once again turned the attention many researchers to linear programming. Karmarkar's work gave rise to the powerful interior point methods which are at present one of the most powerful computational tools of