Frank H. Clarke

Optimization and Nonsmooth Analysis

CANADIAN MATHEMATICAL SOCIETY
SERIES OF
MONOGRAPHS AND ADVANCED TEXTS

A WILEY-INTERSCIENCE PUBLICATION

Optimization and Nonsmooth Analysis

FRANK H. CLARKE

Department of Mathematics
University of British Columbia
Vancouver, Canada

A Wiley-Interscience Publication

JOHN WILEY & SONS

New York • Chichester • Brisbane • Toronto • Singapore

Library of Congress Cataloging in Publication Data:

Clarke, Frank H.
Optimization and nonsmooth analysis.

(Canadian Mathematical Society series in mathematics)
"A Wiley-Interscience publication."
Includes bibliographical references and index.
1. Mathematical optimization. 2. Mathematical analysis. I. Title. II. Title: Nonsmooth analysis. III. Series.
QA402.5.C53 1983 519 83-1216
ISBN 0-471-87504-X

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Optimization and Nonsmooth Analysis

CANADIAN MATHEMATICAL SOCIETY SERIES OF MONOGRAPHS AND ADVANCED TEXTS

Monographies et Études de la Société Mathématique du Canada

Frank H. Clarke ■ Optimization and Nonsmooth Analysis

To my mother

Preface

"... nothing at all takes place in the universe in which some rule of maximum or minimum does not appear." So said Euler in the eighteenth century. The statement may strike us as extreme, yet it is undeniable that humankind's endeavors at least are usually associated with a quest for an optimum. This may serve to explain why there has been an enduring symbiosis between mathematical theories of optimization and the applications of mathematics, even though the forms of the problems (the "paradigms") being considered evolve in time.

The origins of analytic optimization lie in the classical calculus of variations and are intertwined with the development of the calculus. For this reason, perhaps, optimization theory has been slow to shed the strong smoothness (i.e., differentiability) hypotheses that were made at its inception. The attempts to weaken these smoothness requirements have often been ad hoc in nature and motivated by the needs of a discipline other than mathematics (e.g., engineering). In this book a general theory of nonsmooth analysis and geometry will be developed which, with its associated techniques, is capable of successful application to the spectrum of problems encountered in optimization. This leads not only to new results but to powerful versions of known ones. In consequence, the approach taken here is of interest even in the context of traditional, smooth problems in the calculus of variations, in optimal control, or in mathematical programming.

This book is meant to be useful to several types of readers. An effort is made to identify and focus upon the central issues in optimization, and results of considerable generality concerning these issues are presented. Because of this, although its scope is not encyclopedic, the work can serve as a reference text for those in any of the various fields that use optimization. An effort has been made to make the results accessible to those who are not expert in the subject. Thus the first chapter is devoted to an explanation and overview of the book's contents. Here and elsewhere the reader will find examples drawn from economics, engineering, mathematical physics, and various branches of analysis.

The reader who wishes not only to gain access to the main results in the book but also to follow all the proofs will require a graduate-level knowledge of real and functional analysis. With this prerequisite, an advanced course in optimization can be based upon the book. The remaining type of reader we have in mind is the expert, who will discover, we believe, interesting tools and techniques of nonsmooth analysis and optimization.

FRANK H. CLARKE

Vancouver, British Columbia
March 1983

Acknowledgments

Writing a book is a humbling experience. One learns (among other things) how much one depends upon others. The book could not have been written without the much appreciated support of the Killam Foundation, which awarded me a two-year Research Fellowship. I wish also to acknowledge the continued support of the Natural Sciences and Engineering Research Council of Canada. I would like to thank my colleagues at UBC and my co-workers in the field (most of them) for stimulating me over the years (wittingly or otherwise), and the Canadian Mathematical Society for choosing my book for its new series. Foremost in my personal mentions must be Terry Rockafellar who, for one thing, bears the blame for getting me interested in optimization in the first place. I am grateful also to the following colleagues, all of whom have also influenced the contents of the book: Jean-Pierre Aubin, Masako Darrough, Ivar Ekeland, Jean-Baptiste Hiriart-Urruty, Lucien Polak, Rodrigo Restrepo, Richard Vinter, Vera Zeidan. My thanks go to Gus Gassmann and Philip Loewen for courageously scouring a draft of the manuscript, to all the staff of the UBC Mathematics Department for their help and good cheer, and to the good people at John Wiley & Sons.

F. H. C.

A Note From the Publisher

We are proud to have been chosen as the Publisher for the Canadian Mathematical Society Series of Monographs and Advanced Texts and pleased that this book, *Optimization and Nonsmooth Analysis*, represents such an outstanding start to what we anticipate will become one of the finest series in mathematics. Publishing for the Canadian Mathematical Society represents the continuation of our long history of publishing for scientific societies, and in this, our 176th year, we once again affirm our commitment to serving the scientific community in Canada and throughout the World.

June 1983

Optimization and Nonsmooth Analysis

Contents

Chapter One

Introduction and Preview

In adolescence, I hated life and was continually on the verge of suicide, from which, however, I was restrained by the desire to know more mathematics.

BERTRAND RUSSELL, *The Conquest of Happiness*

Just as "nonlinear" is understood in mathematics to mean "not necessarily linear," we intend the term "nonsmooth" to refer to certain situations in which smoothness (differentiability) of the data is not necessarily postulated. One of the purposes of this book is to demonstrate that much of optimization and analysis which have evolved under traditional smoothness assumptions can be developed in a general nonsmooth setting; another purpose is to point out the benefits of doing so. We shall make the following points:

1. Nonsmooth phenomena in mathematics and optimization occur naturally and frequently, and there is a need to be able to deal with them. We are thus led to study differential properties of nondifferentiable functions.
2. There is a recent body of theory (*nonsmooth analysis*) and associated techniques which are well suited to this purpose.
3. The interest and the utility of the tools and methods of nonsmooth analysis and optimization are not confined to situations in which nonsmoothness is present.

Our complete argument in support of these contentions is the entirety of this book. In this chapter we get under way with a nontechnical overview of the theory and some of its applications. The final sections are devoted to placing in context the book's contributions to dynamic optimization (i.e., the calculus of variations and optimal control), the single largest topic with which it deals.

1.1 EXAMPLES IN NONSMOOTH ANALYSIS AND OPTIMIZATION

For purposes of exposition, it is convenient to define five categories of examples.

1.1.1 Existing Mathematical Constructs

The first example is familiar to anyone who has had to prepare a laboratory report for a physics or chemistry class. Suppose that a set of observed data points $(x_0, y_0), (x_1, y_1), \ldots, (x_N, y_N)$ in the x-y plane is given, and consider the problem of determining the straight line in the x-y plane that best fits the data. Assuming that the given data points do not all lie on a certain line (any lab instructor would be suspicious if they did), the notion of "best" must be defined, and any choice is arbitrary. For a given line $y = mx + b$, the error e_i at the ith data point (x_i, y_i) is defined to be $|mx_i + b - y_i|$. A common definition of best approximating line requires that the slope m and the intercept b minimize $\{\Sigma_{i=0}^N e_i^2\}^{1/2}$ over all m and b (or, equivalently, $\Sigma_{i=0}^N e_i^2$). On the face of it, it seems at least as natural to ask instead that the total error $\Sigma_{i=0}^N e_i$ be minimized. The characteristics of the resulting solution certainly differ. In Figure 1.1, for example, the dashed line represents the "least total error" solution (see Example 2.3.17), and the solid line represents the "least total square error" solution. Note that the former ignores the anomalous data point which presumably corresponds to a gross measurement error. The least squares solution, in contrast, is greatly affected by that point. One or the other of these solutions may be preferable; the point we wish to make is that the function $\Sigma_{i=0}^N e_i$ is nondifferentiable as a function of m and b. Thus the usual methods for minimizing differentiable functions would be inapplicable to this function, and different methods would have to be used. Of course, the reason that the least square definition is the common one is that it leads to the minimization of a smooth function of m and b.

The two functions being minimized above are actually special cases of the L^2 and L^1 norms. The differentiability (or otherwise) of norms and of other classes of functions has been and remains a central problem in functional analysis. One of the first results in this area is due to Banach, who characterized the continuous functions x on $[0, 1]$ at which the supremum norm

$$\|x\| := \max_{0 \leqslant t \leqslant 1} |x(t)|$$

is differentiable. (His result is rederived in Section 2.8.)

An interesting example of a nondifferentiable function is the *distance function* d_C of a nonempty closed subset C of R^n. This is the function defined by

$$d_C(x) := \min\{|x - c| : c \in C\},$$

where $|\cdot|$ refers to the Euclidean norm. (It is a consequence of the results of Section 2.5 that when C is convex, for example, d_C must fail to be differentiable at any point on the boundary of C.) The distance function has been a useful tool in the geometrical theory of Banach spaces; it will serve us as well, acting as a bridge between the analytic and the geometric concepts developed later. As an illustration, consider the natural attempt to define directions of tangency to C, at a point x lying in C, as the vectors v satisfying $d'_C(x; v) = 0$, where the notation refers to the customary one-sided directional derivative. Since such directional derivatives do not necessarily exist (unless extra smoothness or convexity hypotheses are made), this approach is only feasible (see Section 2.4) when an appropriate nonsmooth calculus exists.

As a final example in this category, consider the initial-value problem

$$\frac{d}{dt}x(t) = f(t, x(t)), \qquad x(0) = u.$$

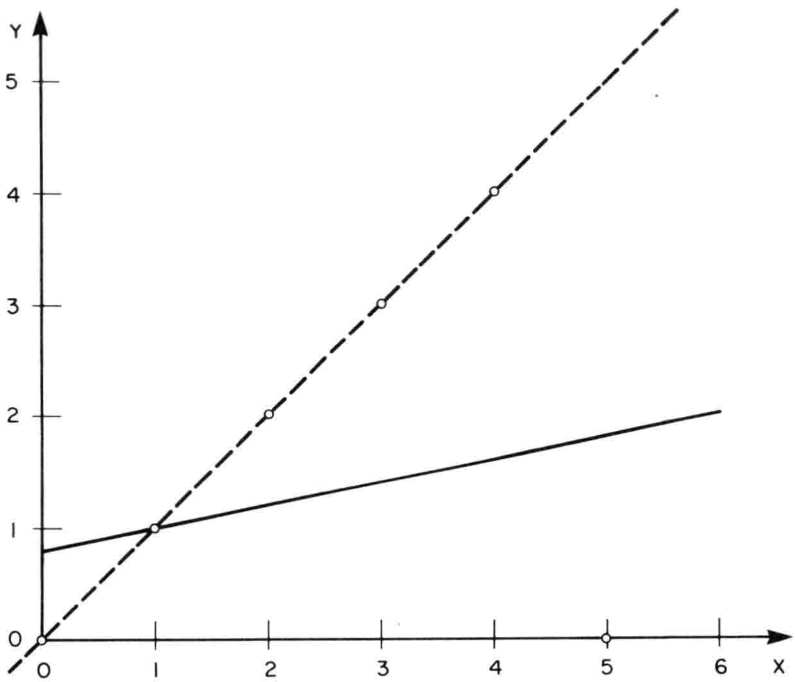

Figure 1.1 Two lines fitted to six data points (indicated by circles).

It is well known that the natural framework for studying existence and uniqueness of solutions is that of functions f which satisfy a Lipschitz condition in the x variable. It is desirable, then, to be able to study in this same framework the closely related issue of how solutions depend on the initial value u. The classical theory, however, hinges upon the resolvent, which is defined in terms of derivatives of f. This confines the analysis to smooth functions f. In Section 7.4 we extend the theory to the Lipschitz setting.

1.1.2 Direction Phenomena

Consider an elastic band whose upper end is fixed, and whose lower end is tied to a unit point mass. When the band is stretched a positive amount x, it exerts an upward (restoring) force proportional to x (Hooke's Law). When unstretched, no force is exerted. (This contrasts to a spring, which also exerts a restoring force when compressed.) If the mass is oscillating vertically, and if $x(t)$ measures the (positive or negative) amount by which the distance from the mass to the upper end of the band exceeds the natural (unstretched) length of the band, Newton's Law yields $\ddot{x}(t) = f(x(t))$, where f is given by

$$f(x) = \begin{cases} g - kx & \text{if } x \geqslant 0 \\ g & \text{if } x \leqslant 0. \end{cases}$$

(g is the acceleration of gravity, and k is the proportionality constant for Hooke's Law; friction and the weight of the band have been neglected.) Note that the function f is continuous but not differentiable at 0.

As another example, consider a flat solar panel in space. When the sun's rays meet its surface, the energy produced is proportional to $\cos\alpha$, where α is the (positive) angle of incidence (see Figure 1.2). When the panel's back is turned to the sun (i.e., when α exceeds $\pi/2$), no energy is produced. It follows then that the energy produced is proportional to the quantity $f(\alpha)$, where f is given by

$$f(\alpha) = \begin{cases} \cos\alpha & \text{if } \alpha \leqslant \pi/2 \\ 0 & \text{if } \alpha \geqslant \pi/2. \end{cases}$$

This again is a function that fails to be differentiable.

As a last illustration in this category, consider the electrical circuit of Figure 1.3 consisting of a diode, a capacitor, and an impressed voltage. A diode is a resistor whose resistance depends upon the direction of the current. If I is the current and V is the voltage drop across the diode, one has the following nonsmooth version of Ohm's Law:

$$I = \begin{cases} V/R_+ & \text{if } V \geqslant 0 \\ V/R_- & \text{if } V \leqslant 0, \end{cases}$$