Vladimir F. Dem'yanov
and
Leonid V. Vasil'ev

NONDIFFERENTIABLE OPTIMIZATION

Vladimir F. Dem'yanov
and Leonid V. Vasil'ev

NONDIFFERENTIABLE OPTIMIZATION

Optimization Software, Inc.
Publications Division, New York

Authors
Vladimir F. Dem'yanov
Department of Applied Mathematics
and Control Processes
Leningrad State University
Leningrad
USSR

Leonid V. Vasil'ev
Department of Mathematics
Leningrad Polytechnic Institute
Leningrad
USSR

Series Editor
A.V. Balakrishnan
School of Engineering
University of California
Los Angeles, CA 90024
USA

Translator
Prof. Tetsushi Sasagawa
Department of Mechanical
Engineering
Sophia University
7-1, Kioi-cho, Chiyoda-ku
Tokyo
Japan

Library of Congress Cataloging-in-Publication Data

Dem'ianov, V.F. (Vladimir Fedorovich), 1938-
Nondifferentiable optimization.

(Translations series in mathematics and engineering)
Includes bibliographies and index.
1. Operations research. 2. Mathematical
optimization. I. Vasil'ev, L.V. (Leonid Vasil'evich)
II. Title. III. Series.
T57.6.D4613 1985 519.8 85-18736
ISBN 0-911575-09-X

© 1985 by Optimization Software, Inc., Publications Division,
4 Park Avenue, New York, New York 10016. All rights reserved.
No part of this publication may be reproduced in any form or
by any means without written permission of the publisher.

Printed in the United States of America.

Worldwide Distribution Rights by Springer-Verlag New York Inc.,
175 Fifth Avenue, New York, New York 10010, USA and
Springer-Verlag Berlin Heidelberg New York Tokyo,
Heidelberg Platz 3, Berlin-Wilmersdorf-33, The Federal
Republic of Germany

ISBN 0-911575-09-X Optimization Software, Inc.
ISBN 0-387-90951-6 Springer-Verlag New York Berlin Heidelberg Tokyo
ISBN 3-540-90951-6 Springer-Verlag Berlin Heidelberg New York Tokyo

About the Authors

Vladimir F. Dem'yanov received his Ph.D. from Leningrad State University in 1964 and his D.Sci. degree from the Computing Center of the USSR Academy of Sciences in 1972. He was appointed Professor at the Applied Mathematics Department of Leningrad State University in 1970. His main interest is Nondifferentiable Optimization.

Leonid V. Vasil'ev received his Ph.D. from Leningrad State University in 1977. In 1980 he was appointed Professor at the Mathematics Department of the Leningrad Polytechnic Institute. His main interest is also Nondifferentiable Optimization.

Preface

Of recent coinage, the term "nondifferentiable optimization" (NDO) covers a spectrum of problems related to finding extremal values of nondifferentiable functions. Problems of minimizing nonsmooth functions arise in engineering applications as well as in mathematics proper. The Chebyshev approximation problem is an ample illustration of this. Without loss of generality, we shall consider only *minimization* problems.

Among nonsmooth minimization problems, minimax problems and convex problems have been studied extensively ([31], [36], [57], [110], [120]). Interest in NDO has been constantly growing in recent years (monographs: [30], [81], [127] and articles and papers: [14], [20], [87]-[89], [98], [130], [135], [140]-[142], [152], [153], [160], all dealing with various aspects of nonsmooth optimization).

For solving an arbitrary minimization problem, it is necessary to:

1. Study properties of the objective function, in particular, its differentiability and directional differentiability.

2. Establish necessary (and, if possible, sufficient) conditions for a global or local minimum.

3. Find the direction of descent (steepest or, simply, feasible--in appropriate sense).

4. Construct methods of successive approximation.

In this book, the minimization problems for nonsmooth functions of a finite number of variables are considered. Of fundamental importance are necessary conditions for an extremum (for example, [24], [45], [57], [73], [74], [103], [159], [163], [167], [168].

此为试读,需要完整PDF请访问: www.ertongbook.com

In the case of smooth functions, the importance of the concept of a gradient is well known. However, for nonsmooth functions, gradients do not exist. For a maximum function and a convex function, the subgradient plays a role similar to that of the gradient: with every point x_0 we associate a compact set $\partial f(x_0)$, which is called the subdifferential of the function $f(x)$ at the point x_0. Using the subdifferential, it is possible to:

1. Find the directional derivative of the function at the point x_0:

$$\frac{\partial f(x_0)}{\partial g} \equiv \lim_{\alpha\to+0} \alpha^{-1}[f(x_0 + \alpha g) - f(x_0)] = \max_{v\in\partial f(x_0)} (v,g) .$$

2. Verify necessary conditions for a minimum: for a point x^* to be a minimum point of the function $f(x)$ on E_n, it is necessary that

$$0 \in \partial f(x^*) .$$

3. Find the direction of steepest descent: if $0 \notin \partial f(x_0)$, then the direction

$$g(x_0) = -v(x_0) \| v(x_0)\|^{-1} ,$$

where $\| v(x_0)\| = \min_{v\in\partial f(x_0)} \| v \|$, $v(x_0) \in \partial f(x_0)$, is the direction of steepest descent of the function $f(x)$ at the point x_0.

Such an important role of the subdifferential has prompted an attempt to extend the concept of a subdifferential to Lipschitzian functions: F.H. Clarke [133], [134]; J. Warga [9], [168]; B.H. Pshenichnyj [104]; N.Z. Shor [126], [127]; A. Gol'dshtejn [139], [140], among others.

Using subdifferentials and subgradients, it is possible to construct several methods of successive approximation for minimizing convex functions, maximum functions, as well as other classes of functions ([30], [36], [91], [127], [149]-[151], [156], [170], [171]).

The problem of minimizing a smooth function $f(x)$ on the set

$$\Omega = \{x \in E_n | h_i(x) \leqslant 0 \quad \forall i \in 1:N\} ,$$

where $h_i(x)$ is a smooth function on E_n, is in fact a problem of NDO, because the set Ω can be represented as follows:

$$\Omega = \{x \in E_n | h(x) \leq 0\} ,$$

where $h(x) = \max_{i \in 1:N} h_i(x)$ is no longer a smooth function.

The objective of this book is a systematic exposition of the theory of optimization of nondifferentiable functions. In Chapter 1, the basic results from the theory of convex functions, convex sets, and point-to-set mappings are introduced. Much attention is paid to ε-subdifferentials and properties of ε-subdifferential mappings. Convex functions are essential not only because they constitute a large class of nonsmooth functions, but also because the tools of the theory of convex functions can be extended to more general classes of nonsmooth functions.

This concept of a convex function and of a maximum function is tied in with that of a directional derivative. Quite a few authors, among those cited above, do not use directional derivatives in their generalizations of the subdifferential. However, in optimization problems, the directional derivative is more natural, as well as more useful.

In Chapter 2, a new class of nondifferentiable functions, that is, the class of quasidifferentiable functions, is described. For such functions the concept of a quasidifferential, which is closely related to that of a directional derivative, plays a significant role. It appears that for each point there exists a pair of convex sets (quasidifferential). The quasidifferential is a generalization of the concept of a derivative (for smooth functions) and of a subdifferential (for convex functions).

The notion of quasidifferentials simplifies considerably the statement of necessary conditions for an extremum and the problem of finding the directions of steepest descent and ascent. The principal formulas of quasidifferential calculus, which is indeed a generalization of the classical quasidifferential calculus, are established next. The class of quasidifferentiable functions is a linear space closed with respect to all "differentiable" operations as well as operations of taking pointwise maxima and minima (while the class of convex functions is not a linear space but a convex cone). The concept of quasidifferentiable sets is a natural extension. A necessary condition for an extremum of a quasidifferentiable function on a quasidifferentiable set is established in terms of quasidifferentials, which essentially extends the class of problems which can be investigated analytically. For a large class of quasidifferentiable functions, it is possible to algorithmize the process of verifying necessary conditions, as well as the process of finding steepest descent or ascent directions. However, numerical techniques still need to be developed.

Chapters 3 and 4 are devoted to numerical methods for solving NDO problems, including minimization of convex functions and maximum functions. Successive approximation methods are classified as relaxation and non-relaxation methods. A method is called the relaxation method if the value of a function at each step is smaller than that at the preceding step. We discuss both classes of methods, but not the advantages of one versus the other, because the "dragon" of optimization is multiheaded and it takes a special sword to cut-off each head. Thus, the method of subgradient descent is simple to instrument but converges very slowly. Many methods depend on the aims and available means. Sometimes, it is possible to make a rough but quick approximation; in other cases, high accuracy may be needed and computational complexity is not a problem.

Most of these methods are "first-order" methods, since the first-order approximations (derivative, subgradient, subdifferential) are used. One might expect that a further development of the NDO theory will involve higher-order methods.

Some material is relegated to exercises. We do not consider stochastic procedures ([51], [83], [107], [117]); nor problems of game theory ([21], [58], [63], [64]) and those of multicriteria optimization ([195]), where NDO is needed.

Included in this book are the results of recent research in nonsmooth optimization, obtained at the Department of Applied Mathematics/Control Processes and at the Institute of Computational Mathematics of Leningrad State University.

Some results were reported at the Seventh and Eighth All-Union Summer Schools on Optimization at Shchukino (1977) and Shushenskoe (1979).

The authors are indebted to N.N. Moiseev for his encouragement and support throughout; and to A.B. Pevnyj and A.M. Rubinov for a careful reading of the manuscipt and constructive suggestions. Thanks are also due to E.F. Vojton, M.K. Gavurin, Yu.M. Ermol'ev, S.S. Kukateladze, V.N. Malozemov, B.N. Pshenichnyj, and V.M. Tikhomirov for their many helpful comments for improving the exposition.

Notation

$\inf \{f(x) \mid x \in A\}$ is shortened to $\inf\limits_{x \in A} f(x)$.

$$\Omega = \{v \in E_n \mid \exists \alpha_0 > 0 : x_0 + \alpha v \in A \ \forall \alpha \in [0, \alpha_0]\}$$

is interpreted as follows: Ω is a set of points $v \in E_n$, for which there exists an $\alpha_0 > 0$ such that $x_0 + \alpha v \in A$ for all $\alpha \in [0, \alpha_0]$.

The set of integers from p to q is denoted by p:q.

The lower and upper limits are denoted by $\underline{\lim}$ and $\overline{\lim}$, respectively.

The number of elements of a set A is denoted by $|A|$.

The symbol $\equiv$ implies "equal by definition."

The symbol ■ indicates the end of a proof.

Material which is used in the sequel is delineated by an asterisk.

Contents

此为试读,需要完整PDF请访问: www.ertongbook.com

Chapter 1

FUNDAMENTALS OF CONVEX ANALYSIS AND RELATED PROBLEMS

1. CONVEX SETS. CONVEX HULLS. SEPARATION THEOREM

1. In what follows we shall consider the n-dimensional Euclidean space E_n of vectors $x = (x^{(1)},\dots,x^{(n)})$. The space E_n is assumed to be linear. Let us introduce some notation:

$$0_n = 0 = (0,\dots,0) \in E_n \ ,$$

$$x_k = (x_k^{(1)};\dots,x_k^{(n)}) \ ,$$

$$(x_1,x_2) = \sum_{i=1}^{n} x_1^{(i)} x_2^{(i)} \ ,$$

$$||x|| = \sqrt{(x,x)} \ ,$$

$$x^2 = (x,x) \ .$$

The Cauchy-Buniakowski inequality

$$|(x_1,x_2)| \le ||x_1|| \ ||x_2||$$

is valid for all vectors $x_1,x_2 \in E_n$.

The vectors $x_1,\dots,x_r$ are said to be *linearly independent* if the equality $\sum_{k=1}^{r} \alpha_k x_k = 0$ implies that all coefficients α_k, $k \in 1:r$ are equal to zero.

If $r \ge n+1$, then the vectors $x_1,\dots x_r$ are *linearly dependent*, i.e., there exist scalars $\beta_1,\dots,\beta_r$ such that $\sum_{k=1}^{r} \beta_k^2 > 0$ (i.e., the β_k are not all equal to zero) and

$$\sum_{k=1}^{r} \beta_k x_k = 0 \tag{1.1}$$

If $r \geq n+2$, then we have the equality

$$\sum_{k=1}^{r} \beta_k = 0 \tag{1.2}$$

in addition to (1.1).

To prove this, we introduce the vectors

$$\bar{x}_k = (1, x_k^{(1)}, \ldots, x_k^{(n)}) \in E_{n+1}, \quad k \in 1:r, \quad r \geq n+2 .$$

Since any $n+2$ vectors in E_{n+1} are linearly dependent, there exist scalars β_k such that $\sum_{k=1}^{r} \beta_k^2 > 0$ and

$$\sum_{k=1}^{r} \beta_k \bar{x}_k = 0_{n+1} . \tag{1.3}$$

It follows from (1.3) that

$$\sum_{k=1}^{r} \beta_k x_k = 0_n , \qquad \sum_{k=1}^{r} \beta_k = 0$$

(here we have set the first component and each of the n remaining components equal to zero).

The set which contains no elements is said to be *empty* and is denoted by $\emptyset$.

Let

$$S_\delta(x_0) = \{x \in E_n \mid \|x-x_0\| \leq \delta\} , \qquad \delta > 0 .$$

The set $S_\delta(x_0)$ is said to be a δ-*neighborhood* of the point x_0. A point x_0 is said to be an *interior point* of a set G if there exists a $\delta > 0$ such that $S_\delta(x_0) \subset G$. We shall denote the set of interior points of a set G by $\operatorname{int} G$ (this set may be empty).

A set $G \subset E_n$ is said to be *open* if for any x_0 there exists a

$\delta > 0$ such that $S_\delta(x_0) \subset G$. It is obvious that $G = \text{int}\, G$ for any open set G.

A set of points x which may be represented in the form $x = \lim_{k\to\infty} x_k$, where $x_k \in G \quad \forall k \in 1:\infty$, is said to be the *closure* of a set $G \subset E_n$. We shall denote the closure of a set G by $\bar{G}$.

A set $G \subset E_n$ is said to be *closed* if $x_0 \in G$ follows from the relation $x_k \xrightarrow[k\to\infty]{} x_0$, $x_0 \in G \quad \forall k \in 1:\infty$. It is obvious that $G = \bar{G}$ for any closed set G.

A point x_0 is said to be a *boundary point* of a set $G \subset E_n$ if, for any $\delta > 0$, its δ-neighborhood $S_\delta(x_0)$ includes at least one point which does not belong to G and at least one point which does belong to G (here x_0 may not belong to G). We shall denote the set of boundary points of a set G by G_{fr}.

A set G is said to be *bounded* if there exists a real number $K < +\infty$ such that $||x|| \le K \quad \forall x \in G$.

A set G is said to be *unbounded* if for any $K > 0$ there exists an $x \in G$ such that $||x|| > K$.

It is obvious that the union, intersection, sum and difference of two bounded sets are again bounded sets.

The intersection of two sets of which at least one is bounded is a bounded set.

If A and B are closed sets, then their union and intersection are again closed sets. However, this property no longer holds for the sum, difference and algebraic difference.

EXAMPLE 1. Let

$$A = \{x = (x^{(1)}, x^{(2)}) \in E_2 \mid x^{(2)} \ge \frac{1}{x^{(1)}},\ x^{(1)} > 0\} ,$$

$$B = \{x = (x^{(1)}, x^{(2)}) \in E_2 \mid x^{(1)} = 0,\ x^{(2)} \le 0\} .$$

It is obvious that the sets A and B are closed but not bounded.

The set

$$C = A+B = \{x=(x^{(1)},x^{(2)}) \mid x^{(1)}>0,\ x^{(2)} \in (-\infty,\infty)\}$$

is not closed because we have

$$\bar{C} = \{x=(x^{(1)},x^{(2)}) \mid x^{(1)} \geq 0,\ x^{(2)} \in (-\infty,\infty)\} \neq C .$$

However, if the sets A and B are closed and at least one of them is bounded, then their sum (and algebraic difference) is also closed.

A set which has the property that, for every sequence constructed from its elements, we can select a convergent subsequence the limit of which belongs to the original set is said to be *compact*. It is well known that a set in E_n is compact iff it is closed and bounded.

DEFINITION 1. A set $\Omega \subset E_n$ is said to be *convex* if, in addition to two arbitrary points $x_1,x_2 \in \Omega$, the set contains the line segment connecting these points, i.e., $[x_1,x_2] \subset \Omega$, where

$$[x_1,x_2] = \{x \in E_n \mid x = \alpha x_1 + (1-\alpha)x_2,\ \alpha \in [0,1]\} .$$

A convex set Ω is said to be *strictly convex* if for any $x_1,x_2 \in \Omega$, $x_1 \neq x_2$, and any $\alpha \in (0,1)$ we have $x_\alpha = \alpha x_1 + (1-\alpha)x_2 \in \operatorname{int}\Omega$.

There exists another definition of a convex set.

DEFINITION 1*. A set $\Omega \subset E_n$ is said to be *convex* if, in addition to two arbitrary points x_1,x_2, the set includes the point $\frac{1}{2}(x_1+x_2)$, i.e., if, for any $x_1,x_2 \in \Omega$, the center point of the line segment connecting the points x_1,x_2 also belongs to Ω.

此为试读，需要完整PDF请访问：www.ertongbook.com