Springer Optimization and Its Applications 59

Messaoud Bounkhel

Regularity Concepts in Nonsmooth Analysis

Theory and Applications

Springer

Springer Optimization and Its Applications

VOLUME 59

Aims and Scope
Optimization has been expanding in all directions at an astonishing rate during the last few decades. New algorithmic and theoretical techniques have been developed, the diffusion into other disciplines has proceeded at a rapid pace, and our knowledge of all aspects of the field has grown even more profound. At the same time, one of the most striking trends in optimization is the constantly increasing emphasis on the interdisciplinary nature of the field. Optimization has been a basic tool in all areas of applied mathematics, engineering, medicine, economics, and other sciences.

The series *Springer Optimization and Its Applications* publishes undergraduate and graduate textbooks, monographs and state-of-the-art expository work that focus on algorithms for solving optimization problems and also study applications involving such problems. Some of the topics covered include nonlinear optimization (convex and nonconvex), network flow problems, stochastic optimization, optimal control, discrete optimization, multi-objective programming, description of software packages, approximation techniques and heuristic approaches.

For further volumes:
http://www.springer.com/series/7393

Messaoud Bounkhel

Regularity Concepts in Nonsmooth Analysis

Theory and Applications

Springer

Messaoud Bounkhel
Department of Mathematics
College of Science
King Saud University
11451 Riyadh
Saudi Arabia
bounkhel@ksu.edu.sa

ISSN 1931-6828
ISBN 978-1-4614-1018-8 e-ISBN 978-1-4614-1019-5
DOI 10.1007/978-1-4614-1019-5
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011938686

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To the soul of my dear father, Ahmed, who died in 1994 and who always believed in me and encouraged me to study mathematics.

To my dear mother Fatma, who has always stood behind all my successes.

To my wife Leila, my children Saged, Kawtar, and Yakine, and to my brothers and sisters and all the members of my family.

Preface

The term *nonsmooth analysis theory* had been used in the 1970s by F. Clarke when he studied and applied the differential properties of functions and sets that are not differentiable in the usual sense. Since Clarke's work, the field of nonsmooth analysis theory has known a considerable expansion, namely with the appearence of an important concept which is the concept of "*regularity* " (regularity of functions and regularity of sets). The primary motivation for introducing regularity notions is to obtain equalities in calculus rules involving various constructs in nonsmooth analysis. The first notion of regularity appeared in Clarke's work (in the 1970s) to ensure equality form in the calculus rules of the Clarke subdifferential for Lipschitz continuous functions.

Many investigators (Rockafellar, Mordukhovich, Thibault, Poliquin et al.) have since then introduced and used many other notions of regularity in the development of nonsmooth analysis theory.

In the last decades, regularity concepts played an increasing role in the applications of nonsmooth analysis such as differential inclusions, optimization, variational inequalities, as well as in nonsmooth analysis itself. Consequently, it is becoming more and more desirable to introduce regularity, at an early stage of study, to graduate students and young researchers in order to familiarize them with the basic concepts and their applications. This book is devoted to the study of various regularity notions in nonsmooth analysis and their applications. To the best of my knowledge, the present work is the first thorough study of the regularity of functions, sets, and multifunctions as well as their important applications to differential inclusions and variational inequalities.

This book is divided into three parts. In the first part, we present an accessible and thorough introduction to nonsmooth analysis theory. Main concepts and some useful results are stated and illustrated through examples and exercises.

In Part II, the most important and recent results of various regularity concepts of sets, functions, and set-valued mappings, in nonsmooth analysis theory are presented. These results include some that have been demonstrated in different works that were published either singly (see [39, 44, 45, 48]), or in collaboration with Thibault (see [58–63]).

Part III contains six chapters, each of which addresses a different application of nonsmooth analysis theory. These applications are the fruit of research that I conducted either singly (see [42, 43]) or in collaboration with various researchers in the field (see [53–55, 58, 64]).

Batna, Algeria Messaoud Bounkhel

Acknowledgements

My gratitude goes first and foremost to my dear mother Fatma for supporting me during the long years of preparation of this project, which started in 2001, and without whose moral help and encouragement this book would not have seen light. I am also indebted to my wife Leila for her interest in my book and for sharing with me the long hours of work that it took.

I extend my appreciation to the Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia, for funding the publication of this work.

After my family members and King Saud University, my gratitude goes to Boris Mordukhovich who has encouraged me at all stages of this project and who has advised me to write a book on regularity concepts in nonsmooth analysis theory and has lent me his full support to achieve such a project. His suggestions on the first versions of the book, over the years of preparation, have always been most pertinent and valuable. Similarly, I am very grateful to Alex Kruger for his accurate remarks and suggestions concerning the preliminary version of the book. Special thanks are addressed to Lionel Thibault, my doctoral thesis adviser, from whom I learned nonsmooth analysis theory and much more.

I finally would like to thank my dear friend Mustapha Bouchareb, a teacher and writer by trade, for taking the time to thoroughly read and correct the English of the preliminary version of the book and for patiently answering a tremendous amount of questions on the English language.

Contents

List of Figures

Part I
Nonsmooth Analysis Theory

Chapter 1
Nonsmooth Concepts

1.1 Introduction

This book assumes a basic knowledge of topological vector space and functional analysis. Moreover, we recall in this section several concepts and fundamental preliminaries which will be used in what follows. The following notation is used throughout this book.

X is a real topological vector space or a real normed vector space or a Banach space with norm $\|\cdot\|$ and $\mathbf{H}$ is a real Hilbert space. The inner product between elements of $\mathbf{H}$ is denoted by $\langle \cdot, \cdot \rangle$, the same notation is also employed for the pairing between X and its topological dual space X^* (the space of continuous linear functionals defined on X). The closed unit ball in X or $\mathbf{H}$ centered at some point $\bar{x}$ and with radius $r > 0$ is denoted by $B(\bar{x}, r)$. For $\bar{x} = 0$ and $r = 1$ we will use the standard notation $\mathbf{B}$ instead of $B(0,1)$. The notation $\mathbf{B}_*$ is used for the closed unit ball in X^* centered at the origin and with radius 1. Whenever needed, we use the notation $\mathbf{B}_Z$ for the closed unit ball centered at the origin of a given normed vector space Z. We will denote by $\mathscr{N}(\bar{x})$ the set of all neighborhoods of $\bar{x}$. For a given set S, the following expressions: $\operatorname{int} S, \operatorname{cl} S, \operatorname{bd} S$, signify the interior, closure, and boundary of S, respectively.

Definition 1.1. Let X be a real vector space. A set S is said to be *convex* provided that for every pair of element (x,y) of S the segment $[x,y] = \{\alpha y + (1-\alpha)x : \alpha \in [0,1]\}$ is contained in S. The *convex hull* of a nonconvex set S is defined as the intersection of all the sets containing S. It is denoted by $co\, S$ and has the following characterization:

$$co\, S = \left\{ \sum_{i=1}^{n} \alpha_i x_i : \quad n \in \mathbf{N}, \ \sum_{i=1}^{n} \alpha_i = 1, \ \alpha_i \geq 0, \ x_i \in S \right\}.$$

The closure of $co\, S$ is called the *closed convex hull* and denoted by $\overline{co}\, S$.

M. Bounkhel, *Regularity Concepts in Nonsmooth Analysis: Theory and Applications*, Springer Optimization and Its Applications 59, DOI 10.1007/978-1-4614-1019-5_1,

Definition 1.2. Let f be an extended real valued function, i.e., $f : X \to \mathbf{R} \cup \{+\infty\}$. We call the sets

$$\operatorname{dom} f = \{x \in X : \ f(x) < +\infty\} \text{ and } \operatorname{epi} f = \{(x,r) \in X \times \mathbf{R} : \ f(x) \leq r\},$$

the *effective domain* of f and the *epigraph* of f, respectively.

1. f is said to be a *convex function* on an open convex set $\Omega \subset X$ provided that

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y), \text{ for all } x, y \in \Omega, \text{ and all } \alpha \in [0,1].$$

   When Ω is the whole space X we will say that f is convex.
2. f is said to be *lower semicontinuous* (in short l.s.c.) at some point $\bar{x}$ in $\operatorname{dom} f$ provided that

$$f(\bar{x}) \leq \liminf_{x \to \bar{x}} f(x).$$

   We will say that f is l.s.c. on X if it is l.s.c. at any point of X.

Exercise 1.1.

1. Prove that f is l.s.c. on X if and only if its epigraph $\operatorname{epi} f$ is closed in $X \times \mathbf{R}$.
2. Prove that f is l.s.c. on X if and only if the r-level set $\{x \in X : \ f(x) \leq r\}$ is closed for any $r \in \mathbf{R}$.
3. Prove that f is convex if and only if its epigraph $\operatorname{epi} f$ is convex. As a consequence the effective domain of convex functions is always convex.
4. Prove that the convexity of f implies the convexity of all the r-level sets. Prove by giving a counter example that the converse in the last question is not true in general.

1.2 From Derivatives to Subdifferentials

In this section, we begin with some classical concepts of differentiability (directional, Gâteaux, and Fréchet) and we will try via optimization problems to explain the evolution of the concept of differentiability from the Fréchet derivative to the generalized gradient concept (also called Clarke subdifferential).

Let X be a real topological vector space, $f : X \to \mathbf{R} \cup \{+\infty\}$ be an extended real valued function and $\bar{x} \in \operatorname{dom} f$.

1. The *directional derivative* of f at $\bar{x}$ in the direction $v \in X$ is given by

$$f'(\bar{x}; v) = \lim_{\delta \downarrow 0} \delta^{-1} \left[f(\bar{x} + \delta v) - f(\bar{x}) \right], \tag{1.1}$$

   when the limit exists.

2. We say that f is *Gâteaux differentiable* at $\bar{x}$ provided $f'(\bar{x};v)$ exists for all $v \in X$ and $f'(\bar{x};\cdot)$ is linear continuous, that is, there exists an element (necessarily unique) $f'_G(\bar{x}) \in X^*$ (called the *Gâteaux derivative*) satisfying

$$\langle f'_G(\bar{x}), v \rangle = f'(\bar{x};v), \text{ for all } v \in X. \tag{1.2}$$

3. If the convergence in (1.1) is uniform with respect to v in bounded subsets of X, we say that f is *Fréchet differentiable* at $\bar{x}$, and we write $f'(\bar{x})$ instead of $f'_G(\bar{x})$.

Remark 1.1.

1. A function may admit a directional derivative $f'(\bar{x};v)$ at $\bar{x}$ in every direction $v \in X$, but fails to admit a Gâteaux derivative $f'_G(\bar{x})$ at $\bar{x}$. For example, let X be a Banach space, $f(x) = \|x\|$, and $\bar{x} = 0$. This function has a directional derivative $f'(\bar{x};v)$ for every direction $v \in X$ and $f'(\bar{x};v) = \|v\|$, while the Gâteaux derivative of this function at $\bar{x}$ does not exist because the function $v \mapsto f'(\bar{x};v) = \|v\|$ is not linear.
2. The Fréchet and Gâteaux differentiability concepts are not equivalent in general even in finite dimensional cases. It is not hard to check that Fréchet differentiability at a point implies its continuity at that point, which is not the case for Gâteaux differentiability. For example, a l.s.c. function f (which is not necessarily continuous) may have a Gâteaux derivative f'_G at a point of discontinuity.
3. If X is a normed vector space and f is a locally Lipschitz, that is, for any point $\bar{x} \in X$ there exists some neighborhood V of $\bar{x}$ and some constant $L > 0$ such that

$$|f(x) - f(y)| \le L\|y - x\|, \text{ for all } x, y \in V,$$

then the two above concepts are equivalent.

1.2.1 Unconstrained Minimization Problems

In most situations in optimization, we begin by considering the following abstract minimization problem: minimize $f(x)$ subject to $x \in S$ where $f : S \to \mathbf{R}$ is defined on S which is a subset of a real vector space X. If we redefine the function f so that $f(x) = +\infty$ for $x \notin S$, then minimizing f over S is equivalent to minimizing the new f over all of X. So, no generality is lost in this paragraph if we restrict our attention to the case where $S = X$. Let $f : X \to \mathbf{R}$ be a function and $\bar{x}$ be a point in X. Thus, let us consider the following unconstrained minimization problem:

$$\text{(UP)} \qquad \begin{cases} \text{Minimize} \quad f(x) \\ \text{subject to } x \in X. \end{cases}$$

Definition 1.3. We will say that

1. f has a local minimum at $\bar{x}$ if and only if there exists a neighborhood V of $\bar{x}$ such that $f(\bar{x}) \leq f(x)$, for all $x \in V$.
2. f has a global minimum at $\bar{x}$ over X if and only if $f(\bar{x}) \leq f(x)$, for all $x \in X$.

Assume that f is Gâteaux differentiable at $\bar{x} \in X$.

Fact 1. If f has a local minimum at $\bar{x}$, then there exists some $\varepsilon > 0$ such that

$$\left\langle f'_{\mathrm{G}}(\bar{x}), x - \bar{x} \right\rangle \geq 0, \text{ for all } x \in \bar{x} + \varepsilon \mathbf{B}. \tag{1.3}$$

Proof. Assume that f has a local minimum at $\bar{x}$, then there exists some $\alpha > 0$ such that

$$f(\bar{x}) \leq f(x), \quad \text{for all } x \in \bar{x} + \alpha \mathbf{B}. \tag{1.4}$$

Fix $\varepsilon \in (0, \alpha)$ and $\delta \in (0, \frac{\alpha}{\varepsilon})$, and fix any $x \in \bar{x} + \varepsilon \mathbf{B}$. Hence,

$$\bar{x} + \delta(x - \bar{x}) \in \bar{x} + \delta \varepsilon \mathbf{B} \subset \bar{x} + \alpha \mathbf{B}$$

and so we get by (1.4)

$$f(\bar{x} + \delta(x - \bar{x})) - f(\bar{x}) \geq 0,$$

for all $\delta \in (0, \frac{\alpha}{\varepsilon})$ and for all $x \in \bar{x} + \varepsilon \mathbf{B}$. Therefore, as f is Gâteaux differentiable at $\bar{x}$, the limit

$$\lim_{\delta \downarrow 0} \delta^{-1} \left[f(\bar{x} + \delta(x - \bar{x})) - f(\bar{x}) \right]$$

exists and so $\left\langle f'_{\mathrm{G}}(\bar{x}), x - \bar{x} \right\rangle \geq 0$ for all $x \in \bar{x} + \varepsilon \mathbf{B}$. □

Exercise 1.2.

1. Prove that the converse in Fact 1 is not true in general. This ensures that (1.3) is only a necessary optimality condition for (UP).
2. Prove that (1.3) is equivalent to

$$f'_{\mathrm{G}}(\bar{x}) = 0. \tag{1.5}$$

Assume now that the function f is not Gâteaux differentiable and f is convex. Take for instance $f(x) = \|x\|$. For this function, $f'_{\mathrm{G}}(0)$ does not exist and it is clear that f has a global minimum over X at $\bar{x} = 0$. But we cannot make use of Fact 1 to derive necessary optimality conditions like relations (1.3) or (1.5) for problem (UP), because f is not Gâteaux differentiable at $\bar{x}$. So it is a natural question to ask what could replace f'_{G} in those relations? One could think of making use of the directional derivative instead of the Gâteaux derivative as follows:

$$f'(\bar{x}; v) = 0, \text{ for all } v \in X. \tag{1.6}$$

However, the relation (1.6) does not hold for the above function, although $\bar{x}$ is a global minimum. Indeed, we can check that $f'(\bar{x}; v) = \|v\|$, for all $v \in X$ and so

$f'(\bar{x};v) = 0$ only for $v = 0$ and $f'(\bar{x};v) \neq 0$ for every $v \neq 0$. Therefore, we have to propose, something else to replace f'_{G} which is the subdifferential of f that we define below.

Definition 1.4. Let f be a convex continuous function on X and let $\bar{x} \in X$. We define the subdifferential of f at $\bar{x}$ as follows

$$\partial^{\mathrm{conv}} f(\bar{x}) = \{\zeta \in X^* : \langle \zeta, v \rangle \leq f'(\bar{x};v), \text{ for all } v \in X\}. \tag{1.7}$$

Exercise 1.3. For every convex continuous function f, every $x \in X$, and every direction $v \in X$ one has:

1. The function $\delta \mapsto \delta^{-1}[f(x+\delta v) - f(x)]$ is nondecreasing for δ small enough.
2. The directional derivative $f'(\bar{x};v)$ exists and is positively homogeneous and subadditive on X with respect to v.
3. $$\partial^{\mathrm{conv}} f(\bar{x}) = \{\zeta \in X^* : \langle \zeta, x - \bar{x} \rangle \leq f(x) - f(\bar{x}), \text{ for all } x \in X\}. \tag{1.8}$$
4. Calculus rules:
$$\partial^{\mathrm{conv}}(f+g)(\bar{x}) = \partial^{\mathrm{conv}} f(\bar{x}) + \partial^{\mathrm{conv}} g(\bar{x}) \text{ and } \partial^{\mathrm{conv}}(\alpha f)(\bar{x}) = \alpha \partial^{\mathrm{conv}} f(\bar{x}), \tag{1.9}$$
whenever $\alpha \in \mathbf{R}$ and g is a convex continuous function on X.

Using the subdifferential concept we can derive an analogue to Fact 1 for convex continuous functions, i.e., necessary optimality conditions.

Proposition 1.1. *Let f be a convex continuous function on X and let $\bar{x} \in X$. If f has a local minimum over X at $\bar{x}$, then*

$$0 \in \partial^{\mathrm{conv}} f(\bar{x}). \tag{1.10}$$

Proof. It follows the same lines as in the proof of Fact 1, by using the definition of the subdifferential in (1.7) or it follows directly from (1.8). □

In fact, for convex continuous functions we have a stronger version of Fact 1. Indeed, we can prove that (1.10) is a necessary and sufficient optimality condition for (UP). Further, any local minimum is a global minimum.

Proposition 1.2. *Let f be a convex continuous function on X and let $\bar{x} \in X$. The relation (1.10) is equivalent to each one of the following assertions:*

1. *f has a local minimum over X at $\bar{x}$;*
2. *f has a global minimum over X at $\bar{x}$.*

Proof. It follows from the relation (1.8). □

Proposition 1.3. *If f is a convex continuous and Gâteaux differentiable function at $\bar{x}$, then $\partial^{\mathrm{conv}} f(\bar{x}) = \{f'_{\mathrm{G}}(\bar{x})\}$ and so the relation (1.5) becomes a necessary and sufficient optimality condition for (UP).*

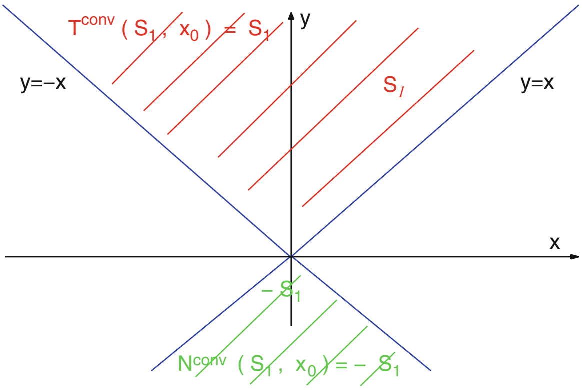

Fig. 1.1 Tangent and normal cones to convex sets

Proof. Let ζ be any element of $\partial^{\text{conv}} f(\bar{x})$. Then, $\langle \zeta, v\rangle \leq f'(\bar{x}; v)$ for all $v \in X$. On the other hand, by the Gâteaux differentiability of f at $\bar{x}$ one has $f'(\bar{x}; v) = \langle f'_{\text{G}}(\bar{x}), v\rangle$ for all $v \in X$. Consequently, we get $\langle \zeta, v\rangle \leq \langle f'_{\text{G}}(\bar{x}), v\rangle$, for all $v \in X$, which ensures that $\zeta = f'_{\text{G}}(\bar{x})$ and so $\partial^{\text{conv}} f(\bar{x}) = \{f'_{\text{G}}(\bar{x})\}$. The second part of the proposition follows from Proposition 1.2 and the first part of this proposition. □

1.2.2 Constrained Minimization Problems

Consider now the following constrained minimization problem:

$$\text{(CP)} \qquad \begin{cases} \text{Minimize} \quad f(x) \\ \text{subject to } x \in S, \end{cases}$$

where f is a convex continuous function and S is a closed convex set in X. First, we define the tangent cone and the normal cone for closed convex sets by

$$T^{\text{conv}}(S; \bar{x}) = \text{cl}\,[\mathbf{R}_+(S - \bar{x})] = \text{cl}\,\{\lambda(s - \bar{x}) : \lambda \geq 0, s \in S\}$$

and $N^{\text{conv}}(S; \bar{x})$ is the negative polar cone[1] of $T^{\text{conv}}(S; \bar{x})$, i.e.,

$$N^{\text{conv}}(S; \bar{x}) = \{\zeta \in X^* : \langle \zeta, v\rangle, \text{ for all } v \in T^{\text{conv}}(S; \bar{x})\}.$$

[1]For a closed nonempty set $L \subset X$, the negative polar of L is denoted by L^0 and defined as

$$L^0 = \{\zeta \in X^* : \langle \zeta, v\rangle, \text{ for all } v \in L\}.$$

Example 1.1. Let $S_1 = \{(x,y) \in \mathbf{R}^2 : y \geq |x|\}$ and $\bar{x} = (0,0)$ (see Fig. 1.1). This set is a closed convex cone and

$$T^{\text{conv}}(S_1;\bar{x}) = \text{cl}\,[\mathbf{R}_+(S_1 - \bar{x})] = \text{cl}\,[\mathbf{R}_+(S_1)] = \text{cl}\,[S_1] = S_1$$

and

$$N^{\text{conv}}(S_1;\bar{x}) = -S_1 = \{(x,y) \in \mathbf{R}^2 : y \leq -|x|\}.$$

Exercise 1.4. Prove the following assertions for closed convex sets S and $\bar{x} \in S$:

1. $N^{\text{conv}}(S;\bar{x}) = \{\zeta \in X : \langle \zeta, x - \bar{x} \rangle \leq 0, \text{ for all } x \in S\}$.
2. The distance function d_S is convex if and only if S is convex.
3. $\partial^{\text{conv}} d_S(\bar{x}) = N^{\text{conv}}(S;\bar{x}) \cap \mathbf{B}_*$.
4. $T^{\text{conv}}(S;\bar{x})$ is a closed convex cone containing the vector zero.

Exercise 1.5. Prove the following:

1. Every l.s.c. convex function is continuous over $int(\text{dom} f)$ the interior of the effective domain of f.
2. Assume that X is a real normed vector space. Every convex function which is finite on an open convex set Ω and bounded around some point $\bar{x} \in \Omega$, is locally Lipschitz on Ω.
3. For any closed subset S of X, and f is Lipschitz with ratio $k > 0$ on an open convex set Ω containing S, then any global minimum $\bar{x}$ of f over S is a global minimum of the function $f + kd_S$ over the whole space X.

We derive in the following proposition a necessary and sufficient optimality condition for (CP).

Proposition 1.4. *Let f be a convex continuous function on a closed convex set S and let $\bar{x} \in int(S)$. Then the following assertions are equivalent:*

1. *f has a local minimum over S at $\bar{x}$, i.e., there exists a neighborhood V of $\bar{x}$ such that $f(\bar{x}) \leq f(x)$, for all $x \in S \cap V$;*
2. *f has a global minimum over S at $\bar{x}$, i.e., $f(\bar{x}) \leq f(x)$, for all $x \in S$;*
3.
$$0 \in \partial^{\text{conv}} f(\bar{x}) + N^{\text{conv}}(S;\bar{x}).$$

Proof. The implication $(1) \Rightarrow (2)$ is left to the reader as an exercise. We prove the implication $(2) \Rightarrow (3)$. Assume that f has a global minimum over S at $\bar{x}$. First, by the second part of Exercise 1.5, f is locally Lipschitz at $\bar{x}$ with some constant $k > 0$. Then by the third part of Exercise 1.5 the function $f + kd_S$ has a global minimum over X at $\bar{x}$, that is, $(f + kd_S)(\bar{x}) \leq (f + kd_S)(x)$ for all $x \in X$. This ensures by (1.9), (1.10) and the third part of Exercise 1.4, that $0 \in \partial^{\text{conv}}(f + kd_S)(x) = \partial^{\text{conv}} f(\bar{x}) + k\partial^{\text{conv}} d_S(\bar{x}) \subset \partial^{\text{conv}} f(\bar{x}) + N^{\text{conv}}(S;\bar{x})$. The converse $(3) \Rightarrow (2)$ follows directly from the characterization of the normal cone in the first part of Exercise 1.4. □

Assume now that the function f is neither convex nor Gâteaux differentiable. In this case, the directional derivative $f'(\bar{x};v)$ does not exist necessarily. Take for instance $f(x) = -\|x\|$ or $f(x) = x^2 \sin(1/x)$, for $x \neq 0$ and $f(0) = 0$, and take $\bar{x} = 0$. Even if $f'(\bar{x};v)$ exists, it may not preserve its important properties cited in Exercise 1.3. Consequently, the subdifferential $\partial^{\text{conv}} f$ loses almost all of its properties, and in particular relation (1.8) as well as the characterization of the global minimum given in Propositions 1.1 and 1.4. Thus, it would be interesting to ask *what could possibly replace both the Gâteaux derivative (for Gâteaux differentiable functions) in Fact 1 and the subdifferential (for convex continuous functions) in Propositions 1.1 and 1.4*. The answer to this question was given by Clarke in [86] when he introduced a generalized gradient (also known as the Clarke subdifferential) for nondifferentiable nonconvex functions and developed a new theory that he called *Nonsmooth Analysis Theory*. Our primary goal in this book is to focus upon this theory and its applications.

1.3 Subdifferentials

In this section, we will assume that X is a normed vector space and $f : X \to \mathbf{R}$ is a locally Lipschitz function at $\bar{x} \in X$ with ratio $k > 0$.

1.3.1 The Generalized Gradient (Clarke Subdifferential)

We have seen that for convex continuous functions the subdifferential was defined in terms of the directional derivative $f'(\bar{x},\cdot)$ (see Definition 1.4). Following the same idea, we define the the generalized gradient by using a new concept of directional differentiability because, as we have mentioned in the end of the previous section, the directional derivative $f'(\bar{x},\cdot)$ loses almost all of its properties and it is not the appropriate directional derivative that can be used to define the generalized gradient (Clarke subdifferential). The new concept of directional derivative is called *the generalized directional derivative* (also known as *Clarke directional derivative*) and is defined by

$$f^0(\bar{x};v) = \limsup_{\substack{x \to \bar{x} \\ t \downarrow 0}} t^{-1}[f(x+tv) - f(x)]. \tag{1.11}$$

The generalized gradient (Clarke subdifferential) of f at $\bar{x}$ is defined then as

$$\partial^{\text{C}} f(\bar{x}) = \{\zeta \in X^* : \langle \zeta, v \rangle \leq f^0(\bar{x};v), \text{ for all } v \in X\}. \tag{1.12}$$

The following proposition summarizes the most important properties of the generalized directional derivative and the generalized gradient for locally Lipschitz functions.

Proposition 1.5.

1. *The function* $v \mapsto f^0(\bar{x};v)$ *is finite, positively homogeneous, subadditive, and satisfies*
$$|f^0(\bar{x};v)| \le k\|v\|, \text{ for all } v \in X. \tag{1.13}$$
2. $(f+g)^0(\bar{x};v) \le f^0(\bar{x};v) + g^0(\bar{x};v)$, *where* g *is a locally Lipschitz function at* $\bar{x}$.
3. *Sum rules:*
$$\partial^C(f+g)(\bar{x}) \subset \partial^C f(\bar{x}) + \partial^C g(\bar{x}),$$
where g *is a locally Lipschitz function at* $\bar{x}$.
4. *For every* $\alpha \in \mathbf{R}$ *one has* $(\alpha f)^0(\bar{x};v) = \alpha f^0(\bar{x};v)$ *and hence* $\partial^{\mathrm{C}}(\alpha f)(\bar{x}) = \alpha \partial^{\mathrm{C}} f(\bar{x})$.
5. *If* f *has a local minimum or maximum at* $\bar{x}$, *then* $0 \in \partial^{\mathrm{C}} f(\bar{x})$.
6. *The generalized gradient* $\partial^{\mathrm{C}} f(\bar{x})$ *is a nonempty, convex,* w^**-compact subset in* X^* *and satisfies* $\partial^{\mathrm{C}} f(\bar{x}) \subset k\mathbf{B}_*$.
7. *If* x_n *and* ζ_n *are two sequences in* X *and* X^* *respectively such that* $\zeta_n \in \partial^{\mathrm{C}} f(x_n)$ *and* x_n *strongly converges to* x *and* ζ_n w^**-converges to* ζ, *then we have* $\zeta \in \partial^{\mathrm{C}} f(x)$.
8. *Mean Value Theorem: If* f *is locally Lipschitz on an open neighborhood containing the segment* $[x,y]$, *then there exists* $z \in [x,y]$ *and* $\xi \in \partial^{\mathrm{C}} f(z)$ *satisfying*
$$f(y) - f(x) = \langle \xi, y-x \rangle.$$
9. *Chain rule: Let* $F : \mathbf{H} \to \mathbf{R}^n$ *be locally Lipschitz*[2] *at* $\bar{x}$ *and let* $g : \mathbf{R}^n \to \mathbf{R}$ *be locally Lipschitz at* $F(\bar{x})$. *Then the function* $g \circ F$ *is locally Lipschitz at* $\bar{x}$ *and*
$$\partial^{\mathrm{C}}(g \circ F)(\bar{x}) \subset \overline{co}\{\partial^{\mathrm{C}}(\langle \xi, F(\cdot)\rangle)(\bar{x}) : \ \xi \in \partial^{\mathrm{C}} g(F(\bar{x}))\}.$$
10. *Pointwise maximum rule: Let* f *be a pointwise maximum of a finite number of locally Lipschitz functions at* $\bar{x}$, *that is,* $f(x) = \max_{1 \le n \le N} f_n(x)$ *with each* f_n *locally Lipschitz at* $\bar{x}$. *Then* f *is locally Lipschitz at* $\bar{x}$ *and satisfies*
$$\partial^{\mathrm{C}} f(\bar{x}) \subset co\{\partial^{\mathrm{C}} f_n(\bar{x}) : n \in I(\bar{x})\},$$
where $I(\bar{x})$ *denotes the set of indices* n *for which* $f(\bar{x}) = f_n(\bar{x})$.

Proof.

1. By the local Lipschitz property of f at $\bar{x}$, we get for $t > 0$ small enough and for x sufficiently close to $\bar{x}$
$$|t^{-1}[f(x+tv) - f(x)]| \le k\|v\|, \text{ for all } v \in X.$$

[2] $F : \mathbf{H} \to \mathbf{R}^n$ is locally Lipschitz at $\bar{x}$ means that $F = (f_1, f_2, \ldots, f_n)$ and each $f_i : \mathbf{H} \to \mathbf{R}$ $(i = 1, 2, \ldots, n)$ is locally Lipschitz at $\bar{x}$.