Generalized Convexity, Nonsmooth Variational Inequalities, and Nonsmooth Optimization

Q. H. Ansari, C. S. Lalitha, and M. Mehta

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Generalized Convexity, Nonsmooth Variational Inequalities, and Nonsmooth Optimization

Generalized Convexity, Nonsmooth Variational Inequalities, and Nonsmooth Optimization

Q. H. Ansari
Aligarh Muslim University
India

C. S. Lalitha
University of Delhi South Campus
India

M. Mehta
Satyawati College, University of Delhi
India

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20130531

International Standard Book Number-13: 978-1-4398-6821-8 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

© 2014 by Taylor & Francis Group, LLC

Contents

Preface

The subjects of generalized convexity, generalized monotonicity, and variational inequalities have evoked a lot of interest in recent times. While there are several texts addressing these singularly, there was a need for a text that combined the three concepts and brought out the finer nuances of these. Also, the topic of variational inequalities defined by a bifunction is new and has not been dealt by any of the books in the past.

The significance of studying variational inequalities lies in the fact that it allows one to deepen the understanding of various classes of problems like system of nonlinear equations, optimization problems, complementarity problems, and fixed point problems. The variational inequality theory not only provides us with a tool for formulating a variety of equilibrium problems but also provides us with algorithms for computational purposes.

This text is aimed primarily at postgraduates and those involved in research, but there is sufficient elementary material for undergraduate courses. The results have been presented in a simple and lucid way without compromising the rigor of the subject. To help the reader understand the theory presented, each chapter includes several examples and counterexamples.

The book is divided into two parts. The former part deals with generalized convexity and generalized monotonicity. In this part we investigate the notions of convexity and generalized convexity for both the differentiable and the nondifferentiable case. For the nondifferentiable case these notions have been introduced in terms of a bifunction as well as in terms of the Clarke subdifferential.

The latter part of the book provides insight into variational inequalities and optimization problems both in smooth and nonsmooth settings. In this part we investigate existence and uniqueness criteria for a variational inequality, study gap function associated with it, and also discuss numerical methods to solve it. Characterizations of solution sets of an optimization problem or a variational inequality is another aspect we ponder upon in this book. The study is further extended to variational inequalities defined by a bifunction or set-valued version given in terms of the Clarke subdifferential.

Chapter 1 commences with the study of convex sets and convex functions. Various characterizations and properties of convex and generalized convex functions have been studied. These functions are used to obtain sufficiency conditions that are necessary, namely, the classical Fermat theorem or Kuhn-Tucker conditions in nonlinear programming. The chapter concludes with the notions of subgradients and subdifferentials for nondifferentiable convex functions.

Chapter 2 deals with different kinds of generalized directional derivatives, namely, the Gâteaux derivative, Dini directional derivative, Dini-Hadamard directional derivative, and Clarke directional derivative. These derivatives share a very important property, namely, positive homogeneity as a function of the direction. Besides this property several other useful properties and calculus rules for these derivatives are discussed. The subdifferentials defined by means of these derivatives and mean value theorems in terms of these subdifferentials are presented.

In Chapter 3, we unify the generalized derivatives by considering a bifunction and present the various notions of convexity and generalized convexity in terms of this bifunction. This notion of generalized convexity encompasses most of the existing definitions involving the generalized derivatives. Apart from these topics, characterizations for pseudolinearity in terms of bifunctions are presented. Discussions on generalized convexity and pseudolinearity in terms of the Clarke subdifferential are also included.

The role of monotonicity in variational inequality theory is the same as that of convexity in optimization theory, and hence it becomes imperative to study monotonicity. Chapter 4 concentrates on monotonicity and generalized monotonicity defined for a bifunction and set-valued maps. The chapter elucidates at length the relation between generalized convexity and the corresponding generalized monotonicity. Further, we briefly broach the topic of pseudoaffine bifunctions and show that they are closely related to pseudolinear functions defined by the bifunction.

Chapter 5 includes an extensive discussion on existence and uniqueness results for variational inequalities, gap functions, and methods for solving variational inequalities including the projection method and auxiliary principle method. Gap functions play a crucial role in linking a variational inequality with an optimization problems. This facilitates the use of numerical methods, available for optimization problem, to derive solutions of a variational inequality. In this chapter we have also dealt with characterizations of solution sets of optimization problems and variational inequality problems.

In Chapter 6, we study a variational inequality involving a bifunction, which pertains to an optimization problem involving a nonsmooth objective. We present some existence criteria for the nonsmooth variational inequality, study-related gap functions, and provide saddle point characterizations.

The objective of Chapter 7 is to study the characterizations of the solution sets of optimization problems having convex or pseudolinear objective function in terms of a bifunction and pseudoaffine variational inequalities when one of the solutions is known.

Chapter 8 delves into generalized variational inequalities given in terms of a set-valued map. Relations of a generalized variational inequality with complementarity and fixed point problems involving set-valued maps have also been explored. Besides establishing the existence results for these variational inequalities, we have also derived characterizations involving a generalized gap function. The study is further extended to a generalized variational inequality involving the Clarke subdifferential and the chapter concludes with characterizations of solution sets of an optimization problem involving a generalized pseudolinear function.

This book would not have come to life without the support and engagement of some key people and it would be in order for us to recognize them for their contribution.

We would like to express our gratitude to our friends and colleagues: Prof. Zafar Ahsan, Prof. S. Al-Homidan, Prof. F. Giannessi, Prof. G. M. Lee, Prof. X.-Q. Yang, and Prof. J.-C. Yao for their kind words of encouragement and inputs at different times during the preparation of this book.

We would like to convey our special thanks to Miss Aastha Sharma, Commissioning Editor at Taylor & Francis, India, for taking keen interest in publishing this book.

Last, but not least, we wish to thank our families for their infinite patience, encouragement, and forbearance. This work would not have seen the light of day had it not been for their unflinching support during this project.

We welcome any opinions, suggestions, and added information that will improve future editions and help readers in the future.

Benefits for readers will be the best reward for the authors.

Q. H. Ansari, C. S. Lalitha,
and M. Mehta

Symbols

$\forall$	for all
$\langle .,. \rangle$	inner product in $\mathbb{R}^n$
$\|\cdot\|$	norm on $\mathbb{R}^n$
$\exists$	there exists
$\mathbf{0}$	zero element in the vector space $\mathbb{R}^n$
2^K	family of all subsets of a set K
$A^\top$	transpose of a matrix A
$[A]$	linear hull of a set A
$\mathrm{Aff}(A)$	affine hull of a set A
$\mathrm{argmin} f$	set of all minima of a function f
$\mathbb{B}_r(x)$	open ball with center at x and radius r
$\mathbb{B}_r[x]$	closed ball with center at x and radius r
$\mathrm{b}(A)$	boundary of a set A
$\mathrm{cl}(A)$	closure of a set A
$\mathrm{co}(A)$	convex hull of a set A
$\mathrm{C}(A)$	conic hull of a set A
$\mathrm{dom}(f)$	domain of a function f
$\mathrm{Dom}(F)$	domain of a set-valued map F
$\mathrm{epi}(f)$	epigraph of a function f
$\nabla f(x)$	gradient of a function f at x
$\nabla^2 f(x)$	Hessian matrix of a function f at x
$f'(x;d)$	directional derivative of f at x in the direction d
$f'_+(x;d)$	right-sided directional derivative of f at x in the direction d
$f'_-(x;d)$	left-sided directional derivative of f at x in the direction d
$f^C(x;d)$	Clarke directional derivative of f at x in the direction d
$f^D(x;d)$	Dini upper directional derivative of f at x in the direction d
$f_D(x;d)$	Dini lower directional derivative of f at x in the direction d
$f^{DH}(x;d)$	Dini-Hadamard upper directional derivative of f at x in the direction d
$f_{DH}(x;d)$	Dini-Hadamard lower directional derivative of f at x in the direction d
$f^G(x;d)$	Gâteaux directional derivative of f at x in the direction d
$f^{MP}(x;d)$	Michel-Penot directional derivative of f at x in the direction d
$f^R(x;d)$	Rockafellar upper subderivative of f at x in the direction d
$f^{wR}(x;d)$	upper weak Rockafellar derivative of f at x in the direction d
$f_{wR}(x;d)$	lower weak Rockafellar derivative of f at x in the direction d
$f^A(x;d)$	adjacent derivative or upper derivative of f at x in the direction d

$\partial f(x)$	subdifferential of a function f at x
$\partial^D f(x)$	Dini upper subdifferential of a function f at x
$\partial_D f(x)$	Dini lower subdifferential of a function f at x
$\partial^C f(x)$	Clarke subdifferential of a convex function f at x
$\partial^{MP} f(x)$	Michel-Penot subdifferential of a function f at x
graph(f)	graph of a function f
hyp(f)	hypograph of a function f
H^+	upper closed half-space
H^{++}	upper open half-space
H^-	lower closed half-space
H^{--}	lower open half-space
int(A)	interior of a set A
$J(f)(x)$	Jacobian matrix of a function f at x
K^*	dual cone of the cone K
$L(f, \alpha)$	lower level set of a function f at level α
$\mathbb{N}$	set of all natural numbers
$N_K(x)$	normal cone to a set K at x
P_K	projection operator
$P_K(x)$	projection of x onto a set K
$\mathbb{R}$	set of all real numbers
$\mathbb{R}_+$	set of all nonnegative real numbers
$\mathbb{R}^n$	n-dimensional Euclidean space
$\mathbb{R}^n_+$	$\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_i \geq 0$ for all $i\}$
$U(f, \alpha)$	upper level set of a function f at level α

Acronyms

AVIP auxiliary variational inequality problem

CP complementarity problem

FPP fixed point problem

GCP generalized complementarity problem

GVIP generalized variational inequality problem

GMVIP generalized Minty variational inequality problem

inf infimum

max maximum

min minimum

MP minimization problem

MVI Minty variational inequality

MVIP Minty variational inequality problem

NCP nonlinear complementarity problem

OP optimization problem

sup supremum

SVFPP set-valued fixed point problem

SVI Stampacchia variational inequality

SVIP Stampacchia variational inequality problem

VI variational inequalities

VIP variational inequality problem

WGMVIP weak generalized Minty variational inequality problem

WGVIP weak generalized variational inequality problem

Part I

Generalized Convexity and Generalized Monotonicity

Chapter 1

Elements of Convex Analysis

1.1 Introduction

Convex functions play a vital role in almost all the branches of mathematics as well as other areas such as science, economics, and engineering. The main reason for this being that they are very well suited to extremum problems as necessary conditions for the existence of a minimum also become sufficient in the presence of convexity. Convex functions were introduced in the beginning of the 20th century by Jensen [111] and more than forty years later a thorough study of conjugate functions was initiated by Fenchel [76, 77]. The lecture notes by Fenchel [77] led to the classic book *Convex Analysis* by Rockafellar [182].

However, since not all real life problems can be formulated as a convex model it becomes a necessity to extend the study to deal with nonconvex problems. Some of the nonconvex functions studied in literature preserve certain properties of convex functions, which in turn help to study the optimality conditions. One of the well-known classes of functions, occurring in various fields such as economics, engineering, management science, probability theory, and various applied sciences, is the class of quasiconvex functions. Although in most of the literature de Finetti [68] is mentioned as the first author to introduce quasiconvex functions, these functions were previously considered by von Neumann [168] and by Popoviciu [177] independently. Another class of functions for which necessary optimality conditions also become sufficient is the class of pseudoconvex functions. These functions were introduced by Mangasarian [152]. The class of pseudoconvex functions includes the class of all differentiable convex functions and is included in the class of all differentiable quasiconvex functions.

The first volume [189] devoted exclusively to generalized convexity was published in 1981 followed by the monograph *Generalized Concavity* by Avriel, Diewert, Schaible, and Zang [28]. Both volumes maintain the close relationship between the mathematics of generalized convex functions and its relevance to applications, especially in economic theory, which described the field from the beginning. Since then, several books and volumes have already appeared in the literature on convex functions and their generalizations with applications in different branches of science, engineering, management, social sciences, etc.

[26, 27, 28, 30, 32, 36, 48, 58, 69, 77, 86, 90, 96, 105, 106, 110, 125, 138, 150, 153, 157, 171, 180, 181, 182, 185, 189, 190].

1.2 Preliminaries and Basic Concepts

Throughout the chapter, $\mathbf{0}$ will be considered as the zero vector in the corresponding vector space. We denote by $\mathbb{R}^n$ the n-dimensional Euclidean space whose norm $\|\cdot\|$ and inner product $\langle .,. \rangle$ are defined by

$$\|x\| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \quad \text{and} \quad \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i,$$

for all $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. The open (respectively, closed) ball with center at x and radius r is denoted by

$$\mathbb{B}_r(x) = \{y \in \mathbb{R}^n : \|x - y\| < r\}$$

$$\big(\text{respectively, } \mathbb{B}_r[x] = \{y \in \mathbb{R}^n : \|x - y\| \leq r\}\big).$$

Definition 1.1. A subset S of $\mathbb{R}^n$ is said to be a *neighborhood* of $x \in \mathbb{R}^n$ if there exists an open ball centered at x and contained in S.

For a given point $x \in \mathbb{R}^n$ and an $\varepsilon > 0$, the open ball $\mathbb{B}_\varepsilon(x)$ is also called an *ε-neighborhood of x*.

Definition 1.2. Let S be a nonempty subset of $\mathbb{R}^n$. A point $x \in S$ is called an *interior point of S* if there is an ε-neighborhood of x that is contained in S, that is, there exists an $\varepsilon > 0$ such that $\|y - x\| < \varepsilon$ implies $y \in S$.

The set of all the interior points of S is called the *interior of S* and it is denoted by $\text{int}(S)$. In other words, $\text{int}(S) = \{x \in S : \mathbb{B}_\varepsilon(x) \subseteq S \text{ for some } \varepsilon > 0\}$.

Definition 1.3. A subset S of $\mathbb{R}^n$ is said to be *open* if $S = \text{int}(S)$.

Definition 1.4. Let S be a nonempty subset of $\mathbb{R}^n$. A point $x \in \mathbb{R}^n$ is called a *limit point of S* if for each $\varepsilon > 0$, $S \cap (\mathbb{B}_\varepsilon(x) \setminus \{x\}) \neq \emptyset$, that is, if every ε-neighborhood of x contains at least one point of S other than x.

The union of S and the set of all the limit points of S is called the *closure of S* and it is denoted by $\text{cl}(S)$. In other words, $\text{cl}(S) = \{x \in \mathbb{R}^n : S \cap (\mathbb{B}_\varepsilon(x) \setminus \{x\}) \neq \emptyset \text{ for all } \varepsilon > 0\}$.

Clearly, $\text{cl}\,(\mathbb{B}_\varepsilon(x)) = \mathbb{B}_\varepsilon[x]$.

Definition 1.5. A subset S of $\mathbb{R}^n$ is said to be *closed* if $S = \text{cl}(S)$.

Definition 1.6. Let S be a subset of $\mathbb{R}^n$. A point $x \in \mathbb{R}^n$ is said to be a *boundary point of* S if for each $\varepsilon > 0$, $\mathbb{B}_\varepsilon(x)$ contains points of S as well points not belonging to S.

The set of all the boundary points of S is called the *boundary of* S and it is denoted by $\mathrm{b}(S)$.

Definition 1.7. A subset S of $\mathbb{R}^n$ is said to be *bounded* if it can be contained within a ball of finite radius.

Definition 1.8. A subset S of $\mathbb{R}^n$ is said to be *compact* if it is closed and bounded.

Remark 1.1. It is well known that a subset S of $\mathbb{R}^n$ is compact if and only if every sequence in S has a convergent subsequence with limit in S.

In optimization one often encounters functions taking values $+\infty$ or $-\infty$. For instance, consider a constrained minimization problem with objective function f defined on $\mathbb{R}^n$ subject to the constraint $x \in K$, where K is a subset of $\mathbb{R}^n$. This problem is equivalent to an unconstrained problem with objective function $\hat{f}$, where $\hat{f}$ takes value $f(x)$ for $x \in K$ and value $+\infty$ for points outside K. In a way some sort of penalty is being imposed for violating the constraints. Also, while dealing with the directional derivatives, parametric optimization problems, or conjugate functions one has to deal with extended real-valued functions. So, it becomes imperative to define algebraic operations involving extended reals. Addition has been extended to $\mathbb{R} \cup \{\pm\infty\}$ as follows

$$(+\infty) + r = +\infty, \quad \text{for all} \;\; r \in \mathbb{R} \cup \{\pm\infty\},$$

$$(-\infty) + r = -\infty, \quad \text{for all} \;\; r \in \mathbb{R} \cup \{-\infty\}.$$

This convention provides the following equivalence:

$$r + s \geq 0 \quad \text{if and only if} \quad r \geq -s \quad \text{for all } r, s \in \mathbb{R} \cup \{\pm\infty\}. \tag{1.1}$$

Also, we assume that $0 \times (\pm\infty) = 0$.

Definition 1.9. A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be

(a) *positively homogeneous* if for all $x \in \mathbb{R}^n$ and all $r \geq 0$, $f(rx) = rf(x)$;

(b) *subadditive* if

$$f(x + y) \leq f(x) + f(y), \quad \text{for all } x, y \in \mathbb{R}^n;$$

(c) *sublinear* if it is positively homogeneous and subadditive;

(d) *subodd* if for all $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $f(x) \geq -f(-x)$.

Every real-valued odd function is subodd. It can be seen that the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is subodd but it is neither odd nor subadditive.

Remark 1.2. (a) By using relation (1.1), f is subodd if and only if $f(x) + f(-x) \geq 0$, for all $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.

(b) If f is sublinear and is not constant with value $-\infty$ such that $f(0) \geq 0$, then f is subodd.

Definition 1.10. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be an extended real-valued function.

(a) The *effective domain* of f is defined as

$$\mathrm{dom}(f) := \{x \in \mathbb{R}^n : f(x) < +\infty\}.$$

(b) The function f is called *proper* if $f(x) < +\infty$ for at least one $x \in \mathbb{R}^n$ and $f(x) > -\infty$ for all $x \in \mathbb{R}^n$.

(c) The *graph* of f is defined as

$$\mathrm{graph}(f) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : y = f(x)\}.$$

(d) The *epigraph* of f is defined as

$$\mathrm{epi}(f) := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq \alpha\}.$$

(e) The *hypograph* of f is defined as

$$\mathrm{hyp}(f) := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : f(x) \geq \alpha\}.$$

(f) The *lower level set* of f at level $\alpha \in \mathbb{R}$ is defined as

$$L(f, \alpha) := \{x \in \mathbb{R}^n : f(x) \leq \alpha\}.$$

(g) The *upper level set* of f at level $\alpha \in \mathbb{R}$ is defined as

$$U(f, \alpha) := \{x \in \mathbb{R}^n : f(x) \geq \alpha\}.$$

The epigraph (hypograph) is thus a subset of $\mathbb{R}^{n+1}$ that consists of all the points of $\mathbb{R}^{n+1}$ lying on or above (on or below) the graph of f. From the above definitions, we have

$$(x, \alpha) \in \mathrm{epi}(f) \text{ if and only if } x \in L(f, \alpha),$$

and

$$(x, \alpha) \in \mathrm{hyp}(f) \text{ if and only if } x \in U(f, \alpha).$$

Definition 1.11. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be

(a) *bounded above* if there exists a real number M such that $f(x) \leq M$, for all $x \in \mathbb{R}^n$;

(b) *bounded below* if there exists a real number m such that $f(x) \geq m$, for all $x \in \mathbb{R}^n$;

(c) *bounded* if it is bounded above as well as bounded below.

For $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$, we write

$$\inf f := \inf\{f(x) : x \in \mathbb{R}^n\},$$

$$\text{argmin} f := \text{argmin}\{f(x) : x \in \mathbb{R}^n\} := \{x \in \mathbb{R}^n : f(x) = \inf f\}.$$

Let us recall that

$$\liminf_{y \to x} f(y) := \sup_{r>0} \inf_{y \in \mathbb{B}_r(x)} f(y),$$

and

$$\limsup_{y \to x} f(y) := \inf_{r>0} \sup_{y \in \mathbb{B}_r(x)} f(y).$$

These can be characterized by

$$\liminf_{y \to x} f(y) = \min\{\alpha \in \mathbb{R} \cup \{\pm\infty\} : \exists\, y_m \to x \text{ with } f(y_m) \to \alpha\},$$

and

$$\limsup_{y \to x} f(y) = \max\{\alpha \in \mathbb{R} \cup \{\pm\infty\} : \exists\, y_m \to x \text{ with } f(y_m) \to \alpha\},$$

respectively. It is easy to observe that the following relations hold:

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

(b) $\liminf_{y \to x} (-f(y)) = -\limsup_{y \to x} f(y)$.

(c) $\limsup_{y \to x} (-f(y)) = -\liminf_{y \to x} f(y)$.

(d) $\liminf_{y \to x} f(y) = \limsup_{y \to x} f(y)$ if and only if $\lim_{y \to x} f(y)$ exists.

Theorem 1.1. Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be bounded in a neighborhood of $y \in \mathbb{R}^n$. The following assertions hold.

(a) $\limsup_{y \to x} (f(y) + g(y)) \leq \limsup_{y \to x} f(y) + \limsup_{y \to x} g(y)$.

(b) $\liminf_{y \to x} (f(y) + g(y)) \geq \liminf_{y \to x} f(y) + \liminf_{y \to x} g(y)$.

(c) If $\lim_{y \to x} g(y)$ exists, then

$$\limsup_{y \to x} (f(y) + g(y)) = \limsup_{y \to x} f(y) + \lim_{y \to x} g(y)$$

and

$$\liminf_{y \to x} (f(y) + g(y)) = \liminf_{y \to x} f(y) + \lim_{y \to x} g(y).$$

Definition 1.12. A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be *lower semicontinuous* at a point $x \in \mathbb{R}^n$ if $f(x) \leq \liminf_{m\to\infty} f(x_m)$ whenever $x_m \to x$ as $m \to \infty$. f is said to be *lower semicontinuous* on $\mathbb{R}^n$ if it is lower semicontinuous at each point of $\mathbb{R}^n$.

A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be *upper semicontinuous* at a point $x \in \mathbb{R}^n$ if $f(x) \geq \limsup_{m\to\infty} f(x_m)$ whenever $x_m \to x$ as $m \to \infty$. f is said to be *upper semicontinuous* on $\mathbb{R}^n$ if it is upper semicontinuous at each point of $\mathbb{R}^n$.

Remark 1.3. A function $f : \mathbb{R}^n \to \mathbb{R}$ is lower (respectively, upper) semicontinuous on $\mathbb{R}^n$ if and only if the lower level set $L(f, \alpha)$ (respectively, the upper level set $U(f, \alpha)$) is closed in $\mathbb{R}^n$ for all $\alpha \in \mathbb{R}$. Also, f is lower (respectively, upper) semicontinuous on $\mathbb{R}^n$ if and only if the epigraph epi(f) (respectively, hypograph hyp(f)) is closed. Equivalently, f is lower (respectively, upper) semicontinuous on $\mathbb{R}^n$ if and only if the set $\{x \in \mathbb{R}^n : f(x) > \alpha\}$ (respectively, the set $\{x \in \mathbb{R}^n : f(x) < \alpha\}$) is open in $\mathbb{R}^n$ for all $\alpha \in \mathbb{R}$.

Definition 1.13. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *differentiable* at $x \in \mathbb{R}^n$ if there exists a vector $\nabla f(x)$, called the gradient, and a function $\alpha : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(y) = f(x) + \langle \nabla f(x), y - x\rangle + \|y - x\|\alpha(y - x), \quad \text{for all } y \in \mathbb{R}^n,$$

where $\lim_{y\to x} \alpha(y - x) = 0$.

If f is differentiable, then

$$f(x + \lambda v) = f(x) + \lambda\langle \nabla f(x), v\rangle + o(\lambda), \quad \text{for all } x + \lambda v \in \mathbb{R}^n,$$

where $\lim_{\lambda\to 0} \dfrac{o(\lambda)}{\lambda} = 0$.

The *gradient* of f at $x = (x_1, x_2, \ldots, x_n)$ is a vector in $\mathbb{R}^n$ given by

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n}\right).$$

Definition 1.14. An $n \times n$ symmetric matrix M of real numbers is said to be *positive semidefinite* if $\langle y, My\rangle \geq 0$ for all $y \in \mathbb{R}^n$. It is called *positive definite* if $\langle y, My\rangle > 0$ for all $y \neq \mathbf{0}$.

Definition 1.15. Let $f = (f_1, \ldots, f_\ell) : \mathbb{R}^n \to \mathbb{R}^\ell$ be a vector-valued function such that the partial derivative $\dfrac{\partial f_i(x)}{\partial x_j}$ of f_i with respect to x_j exists for $i = 1, 2, \ldots, \ell$ and $j = 1, 2, \ldots, n$. Then the *Jacobian matrix* $J(f)(x)$ is given by

$$J(f)(x) = \begin{bmatrix} \dfrac{\partial f_1(x)}{\partial x_1} & \cdots & \dfrac{\partial f_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_\ell(x)}{\partial x_1} & \cdots & \dfrac{\partial f_\ell(x)}{\partial x_n} \end{bmatrix},$$

where $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Definition 1.16. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *twice differentiable* at $x \in \mathbb{R}^n$ if there exist a vector $\nabla f(x)$ and an $n \times n$ symmetric matrix $\nabla^2 f(x)$, called the *Hessian matrix*, and a function $\alpha : \mathbb{R}^n \to \mathbb{R}$ such that

$$\begin{aligned} f(y) &= f(x) + \langle \nabla f(x), y - x\rangle + \langle y - x, \nabla^2 f(x)(y - x)\rangle + \|y - x\|^2 \alpha(y - x), \\ &\quad \text{for all } y \in \mathbb{R}^n, \end{aligned}$$

where $\lim_{y\to x} \alpha(y - x) = 0$.

If f is twice differentiable, then

$$f(x+\lambda v) = f(x)+\lambda\langle\nabla f(x), v\rangle+\lambda^2\langle v, \nabla^2 f(x)v\rangle+o(\lambda^2), \quad \text{for all } x+\lambda v \in \mathbb{R}^n,$$

where $\lim_{\lambda\to 0} \frac{o(\lambda^2)}{\lambda^2} = 0$.

The *Hessian matrix* of f at $x = (x_1, x_2, \ldots, x_n)$ is given by

$$\nabla^2 f(x) \equiv H(x) = \begin{bmatrix} \dfrac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}.$$

1.3 Convex Sets

Let x and y be two different points in $\mathbb{R}^n$. The set $L = \{z \in \mathbb{R}^n : z = \lambda x + (1 - \lambda)y \text{ for all } \lambda \in \mathbb{R}\}$ is the *line* through x and y.

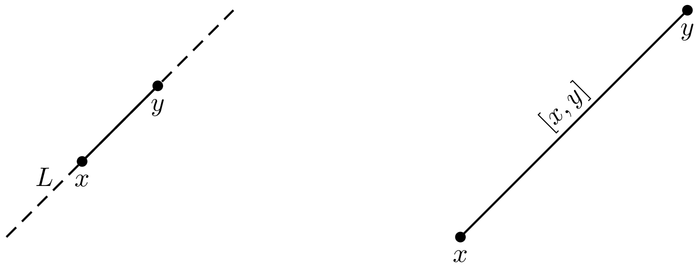
The set $[x, y] := \{z \in \mathbb{R}^n : z = \lambda x + (1 - \lambda)y \text{ for } 0 \leq \lambda \leq 1\}$ is the *line segment* with the endpoints x and y. Similarly, we have the sets

$$\begin{aligned} [x, y[\ &:= \{z \in \mathbb{R}^n : z = \lambda x + (1 - \lambda)y \text{ for all } 0 < \lambda \leq 1\}, \\ ]x, y] &:= \{z \in \mathbb{R}^n : z = \lambda x + (1 - \lambda)y \text{ for all } 0 \leq \lambda < 1\}, \\ ]x, y[\ &:= \{z \in \mathbb{R}^n : z = \lambda x + (1 - \lambda)y \text{ for all } 0 < \lambda < 1\}. \end{aligned}$$

Definition 1.17. A subset W of $\mathbb{R}^n$ is said to be a *subspace* if for all $x, y \in W$ and $\lambda, \mu \in \mathbb{R}$, we have $\lambda x + \mu y \in W$.

Definition 1.18. A subset M of $\mathbb{R}^n$ is said to be an *affine set* if for all $x, y \in M$ and $\lambda, \mu \in \mathbb{R}$ such that $\lambda + \mu = 1$, we have $\lambda x + \mu y \in M$, that is, for all $x, y \in M$ and $\lambda \in \mathbb{R}$, we have $\lambda x + (1 - \lambda)y \in M$.

Thus a subset M of $\mathbb{R}^n$ is an affine set if it contains the whole line through any two of its points.



A line through x and y A line segment with endpoints x and y

FIGURE 1.1: A line and a line segment

Definition 1.19. A subset K of $\mathbb{R}^n$ is said to be a *convex set* if for all $x, y \in K$ and $\lambda, \mu \geq 0$ such that $\lambda + \mu = 1$, we have $\lambda x + \mu y \in K$, that is, for all $x, y \in K$ and $\lambda \in [0, 1]$, we have $\lambda x + (1 - \lambda) y \in K$.

Geometrically speaking, a subset K of $\mathbb{R}^n$ is convex if it contains the line segment joining any two of its points.

In $\mathbb{R}^2$, lines, line segments, circular discs, elliptical discs, and interiors of triangles are all convex.

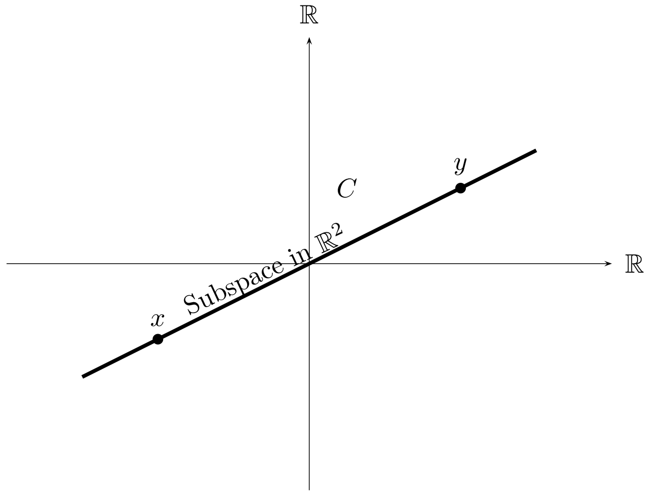


FIGURE 1.2: A subspace in $\mathbb{R}^2$: The line through $\mathbf{0}$, x and y

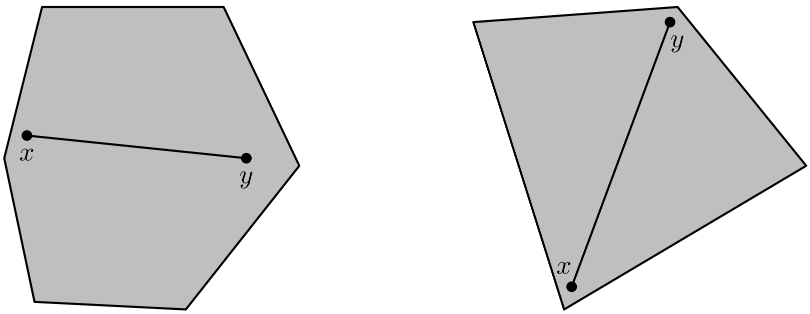


FIGURE 1.3: Convex sets

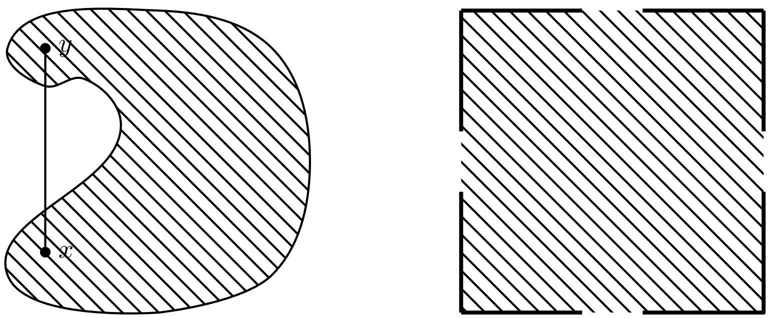


FIGURE 1.4: Nonconvex sets

It follows directly from the definition that the intersection of convex sets, the Cartesian product of convex sets, the image and inverse image of a convex set under a linear transformation, and the interior and the closure of a convex set are convex. In particular, if A and B are convex sets and $\alpha \in \mathbb{R}$, then $A + B = \{x + y : x \in A,\ y \in B\}$ and $\alpha A = \{\alpha x : x \in A\}$ are convex.

Definition 1.20. A subset C of $\mathbb{R}^n$ is said to be a *cone* if for all $x \in C$ and $\lambda \geq 0$, we have $\lambda x \in C$.

A subset C of $\mathbb{R}^n$ is said to be a *convex cone* if it is convex and a cone, that is, for all $x, y \in C$ and $\lambda, \mu \geq 0$, $\lambda x + \mu y \in C$.

Also, a cone C is said to be *pointed* if $C \cap (-C) = \{\mathbf{0}\}$.

Geometrically speaking, a subset C of $\mathbb{R}^n$ is a convex cone if it is the pie slice with apex at origin and edges passing through x and y, for all $x, y \in C$.

Remark 1.4. It is clear from the above definitions that every subspace is an

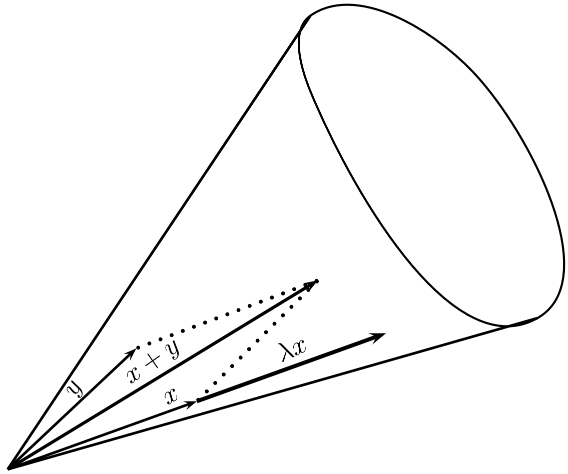


FIGURE 1.5: A convex cone

affine set as well as a convex cone, and every affine set and every convex cone are convex. But the converse of these statements are not true in general.

Evidently, the empty set, a singleton set, and the whole space $\mathbb{R}^n$ are all both affine and convex.

Definition 1.21. Let C be a closed convex pointed cone in $\mathbb{R}^n$. The *dual cone* C^* of C is defined by

$$C^* = \{y \in \mathbb{R}^n : \langle y, x\rangle \geq 0 \text{ for all } x \in C\}.$$

Geometrically, C^* consists of all those vectors that make a nonobtuse angle with every vector in C.

Definition 1.22. Given $x_1, x_2, \dots, x_m \in \mathbb{R}^n$, a vector $x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m$ is called

(a) *a linear combination* of $x_1, x_2, \dots, x_m$ if $\lambda_i \in \mathbb{R}$ for all $i = 1, 2, \dots, m$;

(b) *an affine combination* of $x_1, x_2, \dots, x_m$ if $\lambda_i \in \mathbb{R}$ for all $i = 1, 2, \dots, m$ and $\displaystyle\sum_{i=1}^{m} \lambda_i = 1$;

(c) *a convex combination* of $x_1, x_2, \dots, x_m$ if $\lambda_i \geq 0$ for all $i = 1, 2, \dots, m$ and $\displaystyle\sum_{i=1}^{m} \lambda_i = 1$;

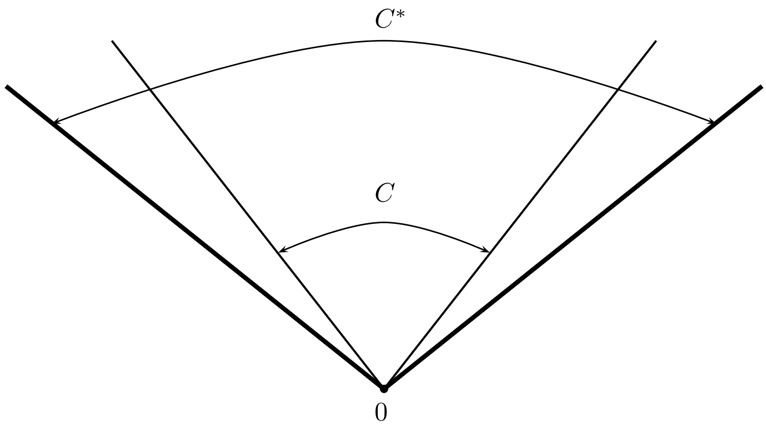


FIGURE 1.6: A cone and its dual

(d) *a cone combination* of $x_1, x_2, \ldots, x_m$ if $\lambda_i \geq 0$ for all $i = 1, 2, \ldots, m$.

Remark 1.5. An affine combination is a linear combination when the sum of the coefficients is 1.

Definition 1.23. Let $S = \{x_1, x_2, \ldots, x_m\}$ be a subset of $\mathbb{R}^n$. Then the set of all cone combinations of elements of S is referred to as the *polyhedral cone* generated by $x_1, x_2, \ldots, x_m$.

Definition 1.24. The vectors $x_1, x_2, \ldots, x_m$ in $\mathbb{R}^n$ are said to be *affinely independent* if none is an affine combination of the others. In other words, the vectors $x_1, x_2, \ldots, x_m$ are affinely independent if there exist scalars $\lambda_i \in \mathbb{R}$ for all $i = 1, 2, \ldots, m$ with $\sum_{i=1}^m \lambda_i = 1$ for which $\sum_{i=1}^m \lambda_i x_i = 0$.

The following theorem shows that a set K is a subspace, affine, convex, or cone according to whether it is closed under linear, affine, convex, or cone combination, respectively, of points of K.

Theorem 1.2. A subset K of $\mathbb{R}^n$ is convex (respectively, subspace, affine, convex cone) if and only if every convex (respectively, linear, affine, cone) combination of points of K lies in it.

Proof. Since a set that contains all convex combinations of its points is obviously convex, we only consider K is convex and prove that it contains any convex combination of its points, that is, if K is convex and $x_i \in K$, $\lambda_i \geq 0$ for all $i = 1, 2, \ldots, m$ and $\sum_{i=1}^m \lambda_i = 1$, then we have to show that $\sum_{i=1}^m \lambda_i x_i \in K$. We prove this by induction on the number of points m of K occurring in a convex combination. If $m = 1$, the assertion is simply $x_1 \in K$, and thus, the result is evidently true. If $m = 2$, then $\lambda_1 x_1 + \lambda_2 x_2 \in K$ for $\lambda_i \geq 0$, $i = 1, 2$, $\sum_{i=1}^{2} \lambda_i = 1$, holds because K is convex. Now suppose that the result is true for m, and we establish the result for $m+1$. If $\lambda_{m+1} = 1$, the result holds because $\lambda_i = 0$ for $i = 1, 2, \ldots, m$ and the result is true for $m = 1$. If $\lambda_{m+1} \neq 1$, we have

$$\begin{aligned}
\sum_{i=1}^{m+1} \lambda_i x_i &= \sum_{i=1}^{m} \lambda_i x_i + \lambda_{m+1} x_{m+1} \\
&= \sum_{i=1}^{m} (1 - \lambda_{m+1}) \frac{\lambda_i x_i}{1 - \lambda_{m+1}} + \lambda_{m+1} x_{m+1} \\
&= (1 - \lambda_{m+1}) \sum_{i=1}^{m} \frac{\lambda_i}{1 - \lambda_{m+1}} x_i + \lambda_{m+1} x_{m+1} \\
&= (1 - \lambda_{m+1}) \sum_{i=1}^{m} \mu_i x_i + \lambda_{m+1} x_{m+1},
\end{aligned} \tag{1.2}$$

where $\mu_i = \dfrac{\lambda_i}{(1 - \lambda_{m+1})}$, $i = 1, 2, \ldots, m$. But then $\mu_i \geq 0$ for $i = 1, 2, \ldots, m$ and

$$\sum_{i=1}^{m} \mu_i = \frac{\sum_{i=1}^{m} \lambda_i}{1 - \lambda_{m+1}} = \frac{1 - \lambda_{m+1}}{1 - \lambda_{m+1}} = 1,$$

as the result is true for m points, $y = \sum_{i=1}^{m} \mu_i x_i \in K$. Hence, from (1.2) and the convexity of K, we have

$$\sum_{i=1}^{m+1} \lambda_i x_i = (1 - \lambda_{m+1}) y + \lambda_{m+1} x_{m+1} \in K.$$

The proofs for the other cases related to a subspace, an affine set, and a convex cone follow exactly on the same pattern. □

Definition 1.25. The intersection of all the convex sets (respectively, subspaces, affine sets) containing a given set $S \subseteq \mathbb{R}^n$ is called the *convex hull* (respectively, *linear hull, affine hull*) of S and is denoted by $\mathrm{co}(S)$ (respectively, $[S]$, $\mathrm{Aff}(S)$). Similarly, the intersection of all the convex cones containing S is called the *conic hull* of S and it is denoted by $C(S)$.

By Theorem 1.2, the convex (respectively, affine, conic) hull is a convex set (respectively, affine set, convex cone). In fact, $\mathrm{co}(S)$ (respectively, $\mathrm{Aff}(S)$, $C(S)$) is the smallest convex set (respectively, affine set, convex cone) containing S.

Theorem 1.3. Let S be a nonempty subset of $\mathbb{R}^n$. Then, $x \in \mathrm{co}(S)$ if and only if there exist x_i in S, $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$, for some positive integer m

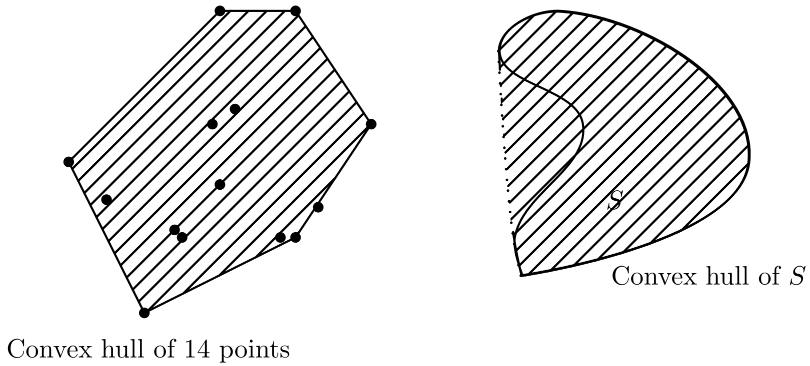


FIGURE 1.7: Convex hull

with $\sum_{i=1}^{m} \lambda_i = 1$ such that $x = \sum_{i=1}^{m} \lambda_i x_i$. In other words, S is convex if and only if any convex combination of points from S is again a point of S.

Proof. Let x be a convex combination of some points of S. Since $\text{co}(S)$ is a convex set containing S, therefore, from Theorem 1.2, every convex combination of its points lies in it, hence, $x \in \text{co}(S)$.

Conversely, let $K(S)$ be the set of all convex combinations of elements of S. It is easily shown that the set

$$K(S) = \left\{ \sum_{i=1}^{m} \lambda_i x_i : x_i \in S,\ \lambda_i \geq 0,\ i = 1, 2, \ldots, m,\ \ \sum_{i=1}^{m} \lambda_i = 1, m \geq 1 \right\},$$

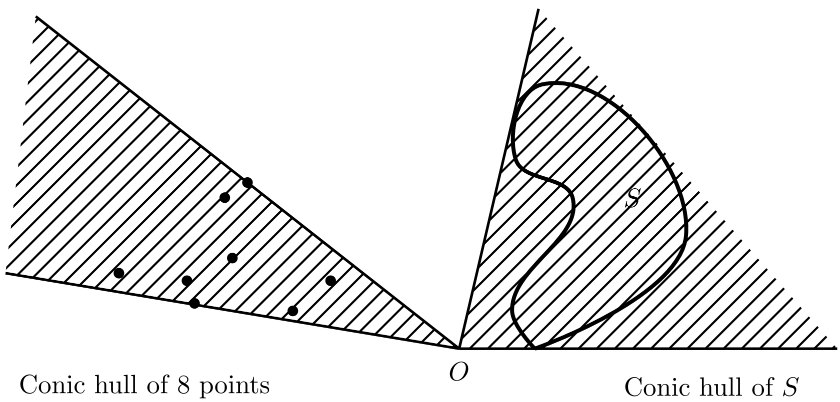


FIGURE 1.8: Conic hull

is convex. Namely, consider $y = \sum_{i=1}^{m} \lambda_i y_i$ and $z = \sum_{j=1}^{\ell} \mu_j z_j$ where $y_i \in S$, $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$, $\sum_{i=1}^{m} \lambda_i = 1$ and $z_j \in S$, $\mu_j \geq 0$, $j = 1, 2, \ldots, \ell$, $\sum_{j=1}^{\ell} \mu_j = 1$, and let $0 \leq \lambda \leq 1$. Then

$$\lambda y + (1-\lambda) z = \sum_{i=1}^{m} \lambda \lambda_i y_i + \sum_{j=1}^{\ell} (1-\lambda) \mu_j z_j,$$

where $\lambda \lambda_i \geq 0$, $i = 1, 2, \ldots, m$, $(1-\lambda)\mu_j \geq 0$, $j = 1, 2, \ldots, \ell$ and

$$\sum_{i=1}^{m} \lambda \lambda_i + \sum_{j=1}^{\ell} (1-\lambda) \mu_j = \lambda \sum_{i=1}^{m} \lambda_i + (1-\lambda) \sum_{j=1}^{\ell} \mu_j = \lambda + (1-\lambda) = 1.$$

Also, this set of convex combinations contains S since each x in S can be written as $x = 1 \cdot x$. By the definition of $\mathrm{co}(S)$, as the intersection of all convex supersets of S, we deduce that $\mathrm{co}(S)$ is contained in $K(S)$.

Thus, the convex hull of S is the set of all (finite) convex combinations from within S. □

As in Theorem 1.3, the subspace $[S]$ spanned by a nonempty subset S of $\mathbb{R}^n$ consists of all (finite) linear combinations of points of S.

The above result holds for affine sets and convex cones as well.

Corollary 1.1. (a) A subset S of $\mathbb{R}^n$ is convex if and only if $S = \mathrm{co}(S)$.

(b) A subset S of $\mathbb{R}^n$ is affine if and only if $S = \mathrm{Aff}(S)$.

(c) A subset S of $\mathbb{R}^n$ is a convex cone if and only if $S = C(S)$.

(d) A subset S of $\mathbb{R}^n$ is a subspace if and only if $S = [S]$.

We remark that the convex hull of a bounded (respectively, open) subset of $\mathbb{R}^n$ is bounded (respectively, open). Also, the convex hull of a compact subset of $\mathbb{R}^n$ is compact. For the proof, we refer to other studies [86, 170, 171, 182]. However, the convex hull of a closed subset of $\mathbb{R}^n$ is not necessarily closed. The convex hull of the closed set

$$S = \{(x_1, x_2) : x_2 \geq |x_1|\} \cup \{(x_1, x_2) : x_2 \geq 1\},$$

given by

$$\mathrm{co}(S) = \{(x_1, x_2) : x_2 > 0\} \cup \{(0, 0)\},$$

is not a closed set.

1.4 Hyperplanes

A hyperplane in an n-dimensional space is a flat subset of dimension $n-1$, which separates the whole space into two half-spaces. It is a generalization of a plane in $\mathbb{R}^3$.

Definition 1.26. Given a nonzero vector $a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, the set $H = \{x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n : \langle a, x\rangle = a_1x_1 + a_2x_2 + \cdots + a_nx_n = \lambda\}$ is called a *hyperplane* in $\mathbb{R}^n$. The vector a is called a *normal* to the hyperplane H.

Every other normal to H is either a positive or a negative scalar multiple of a.

An $(n-1)$-dimensional affine set in $\mathbb{R}^n$ is a hyperplane.
In $\mathbb{R}^2$ the hyperplanes are the straight lines and in $\mathbb{R}^3$ such sets are planes.

A good interpretation of this is that every hyperplane has "two sides," like one's picture of a line in $\mathbb{R}^2$ or a plane in $\mathbb{R}^3$.

It is clear that every affine subset of $\mathbb{R}^n$ is an intersection of a finite collection of hyperplanes.

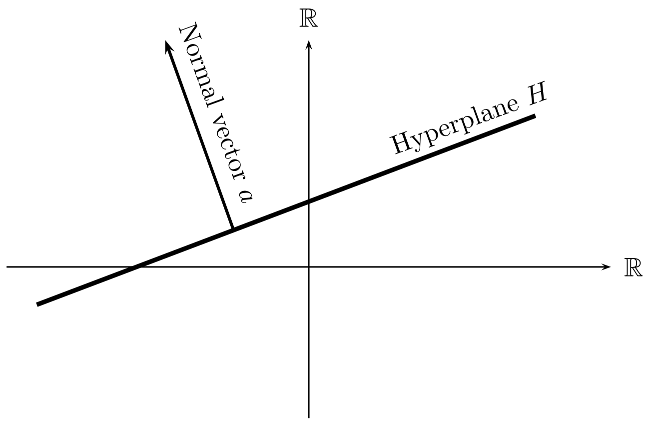


FIGURE 1.9: A hyperplane and a normal vector

Definition 1.27. Let a be a nonzero vector in $\mathbb{R}^n$ and λ be a scalar.

(a) $H^{++} = \{x \in \mathbb{R}^n : \langle a, x\rangle > \lambda\}$ is called *upper open half-space.*

(b) $H^{--} = \{x \in \mathbb{R}^n : \langle a, x\rangle < \lambda\}$ is called *lower open half-space.*

(c) $H^+ = \{x \in \mathbb{R}^n : \langle a, x\rangle \geq \lambda\}$ is called *upper closed half-space.*

(d) $H^- = \{x \in \mathbb{R}^n : \langle a, x\rangle \leq \lambda\}$ is called *lower closed half-space.*

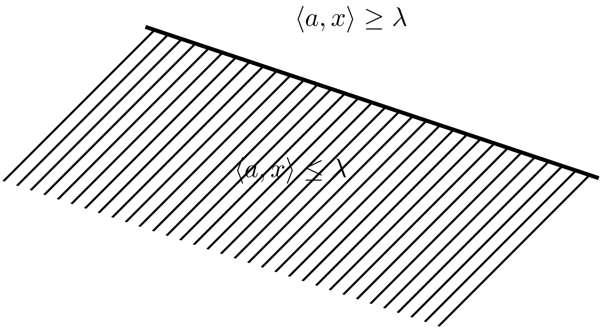


FIGURE 1.10: Closed half-spaces

Remark 1.6. (a) It is clear that the half-spaces are convex sets.

(b) The half-spaces and hyperplane partition $\mathbb{R}^n$ into disjoint sets as

$$\mathbb{R}^n = H^{--} \cup H \cup H^{++}.$$

Also,

$$\mathbb{R}^n = H^- \cup H^+.$$

(c) Note that any point in $\mathbb{R}^n$ lies in H^+, in H^-, or in both.

Definition 1.28. Let S be a nonempty set in $\mathbb{R}^n$ and $\bar{x} \in \mathrm{b}(S)$. A hyperplane $H = \{x \in \mathbb{R}^n : \langle a, x-\bar{x}\rangle = 0,\ a \neq \mathbf{0},\ a \in \mathbb{R}^n\}$ is called a *supporting hyperplane to S at $\bar{x}$* if either

$$S \subseteq H^+, \quad \text{that is,} \quad \langle a, x-\bar{x}\rangle \geq 0, \quad \text{for all } x \in S,$$

or

$$S \subseteq H^-, \quad \text{that is,} \quad \langle a, x-\bar{x}\rangle \leq 0, \quad \text{for all } x \in S.$$

H is called a *proper supporting hyperplane* to S at $\bar{x}$ if, in addition to the aforementioned properties, it satisfies $S \not\subseteq H$.

Remark 1.7. Definition 1.28 can be stated equivalently as follows: The hyperplane $H = \{x \in \mathbb{R}^n : \langle a, x-\bar{x}\rangle = 0\}$ is a supporting hyperplane of S at $\bar{x} \in \mathrm{b}(S)$ if

$$\langle a, \bar{x}\rangle = \inf\{\langle a, x\rangle : x \in S\},$$

or else

$$\langle a, \bar{x}\rangle = \sup\{\langle a, x\rangle : x \in S\}.$$

Definition 1.29. Let S be a nonempty subset of $\mathbb{R}^n$. A hyperplane $H = \{x \in \mathbb{R}^n : \langle a, x\rangle = \lambda,\ a \neq \mathbf{0},\ a \in \mathbb{R}^n,\ \lambda \in \mathbb{R}\}$ is called a *supporting hyperplane of* S if either $S \subseteq H^+$ (or $S \subseteq H^-$) and $\mathrm{cl}(S) \cap H \neq \emptyset$.

H is called a *proper supporting hyperplane* of S if, in addition to the said properties, we have $\mathrm{cl}(S) \cap H \neq S$.

If $\bar{x} \in \mathrm{cl}(S) \cap H$, then H supports the set S at $\bar{x}$. It is obvious that a hyperplane H may support a set S at several distinct points.

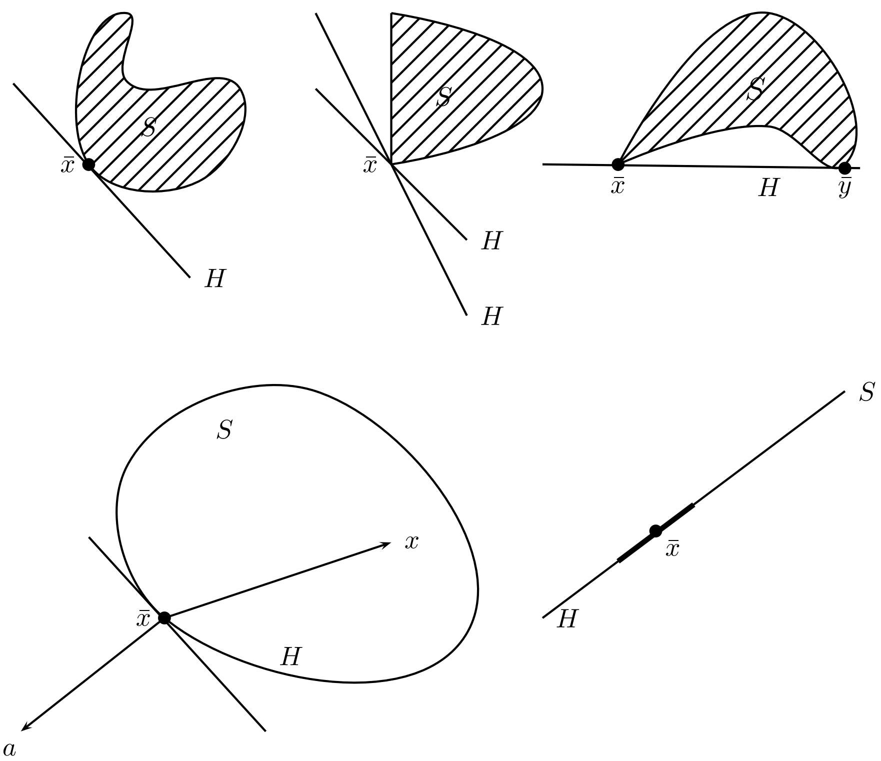


FIGURE 1.11: Supporting hyperplanes

The following result says that $\inf\{\langle a, x\rangle : x \in S\} = \lambda$ if the hyperplane $H = \{x \in \mathbb{R}^n : \langle a, x\rangle = \lambda,\ a \neq \mathbf{0},\ a \in \mathbb{R}^n,\ \lambda \in \mathbb{R}\}$ supports S and S is contained in H^+.

Theorem 1.4. Let $H = \{x \in \mathbb{R}^n : \langle a, x\rangle = \lambda,\ a \neq \mathbf{0},\ a \in \mathbb{R}^n,\ \lambda \in \mathbb{R}\}$ be a supporting hyperplane of a nonempty set S in $\mathbb{R}^n$ such that $S \subseteq H^+$. Then, $\inf\{\langle a, x\rangle : x \in S\} = \lambda$.

Proof. Since $S \subseteq H^+$, we have $\langle a, x \rangle \geq \lambda$ for all $x \in S$, that is, $\inf\{\langle a, x \rangle : x \in S\} \geq \lambda$. If the equality does not hold, then there exists an $\varepsilon > 0$ such that $\langle a, x \rangle \geq \lambda + \varepsilon$ for all $x \in S$.

But by the definition of a supporting hyperplane, $\mathrm{cl}(S) \cap H \neq \emptyset$, and therefore, there exists $y \in \mathrm{cl}(S)$ such that $\langle a, y \rangle = \lambda$. Since $y \in \mathrm{cl}(S)$, there exists a sequence $\{x_m\}$ in S such that $x_m \to y$, and so, $\langle a, y \rangle = \lim_{m \to \infty} \langle a, x_m \rangle \geq \lambda + \varepsilon$, a contradiction. □

Corollary 1.2. If $H = \{x \in \mathbb{R}^n : \langle a, x \rangle = \lambda,\ a \neq \mathbf{0},\ a \in \mathbb{R}^n,\ \lambda \in \mathbb{R}\}$ is a supporting hyperplane for S such that $S \subseteq H^-$, then $\sup\{\langle a, x \rangle : x \in S\} = \lambda$.

Definition 1.30. Let K be a nonempty closed subset of $\mathbb{R}^n$ and $y \in \mathbb{R}^n$. A point $\bar{x} \in K$ is said to be the *projection of* y *on* K or *best approximation of* y *on* K, denoted by $\bar{x} = P_K(y)$, if

$$\|y - \bar{x}\| = \min_{x \in K} \|y - x\|.$$

Remark 1.8. If $y \in K$, then the projection is unique and $\bar{x} = y$. We note that the projection of y on K may not always exist (for example, if K is open) and when it exists it may not be unique (for example, if $K = \{x \in \mathbb{R}^2 : \|x\| \geq 1\}$ and y is the origin). However, under closedness and convexity assumptions the following assertion holds.

Theorem 1.5. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and y a point in $\mathbb{R}^n$ with $y \notin K$. Then, there exists a unique point $\bar{x} \in K$ such that

$$\|y - \bar{x}\| = \min_{x \in K} \|y - x\|. \tag{1.3}$$

Also, the unique point $\bar{x}$ satisfies the following inequality:

$$\langle y - \bar{x}, x - \bar{x} \rangle \leq 0, \quad \text{for all } x \in K. \tag{1.4}$$

Proof. We first establish the equivalence between (1.3) and (1.4). Assume that $\bar{x} \in K$ satisfies (1.3). Let us fix $x \in K$ and $\lambda \in\ ]0,1[$. Since K is convex, we have $(1-\lambda)\bar{x} + \lambda x \in K$. Then from (1.3), we obtain

$$\|y - x\|^2 - \|y - \bar{x}\|^2 \geq 0, \quad \text{for all } x \in K,$$

or

$$\begin{aligned} \frac{1}{\lambda}\left(\|y - [(1-\lambda)\bar{x} + \lambda x]\|^2 - \|y - \bar{x}\|^2\right) &= \frac{1}{\lambda}\left(\|(y - \bar{x}) - \lambda(x - \bar{x})\|^2 - \right. \\ &\quad \left. \|y - \bar{x}\|^2\right) \geq 0. \end{aligned}$$

Simplifying this expression, we obtain

$$\lambda\|x - \bar{x}\|^2 - 2\langle y - \bar{x}, x - \bar{x} \rangle \geq 0.$$

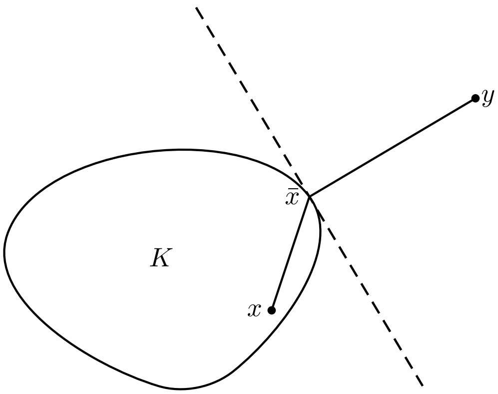


FIGURE 1.12: The projection of a point y onto K

Letting $\lambda \to 0^+$, we obtain the inequality (1.4).

Conversely, suppose that $\bar{x} \in K$ satisfies inequality (1.4). Let us fix $x \in K$, then from inequality (1.4), we get

$$\begin{aligned} 0 &\geq \langle y - \bar{x}, x - \bar{x} \rangle \\ &= \langle y - \bar{x}, (y - \bar{x}) - (y - x) \rangle \\ &= \|y - \bar{x}\|^2 - \langle y - \bar{x}, y - x \rangle. \end{aligned}$$

By means of the Cauchy-Schwarz inequality, we have

$$\|y - \bar{x}\|^2 \leq \langle y - \bar{x}, y - x \rangle \leq \|y - \bar{x}\| \; \|y - x\|.$$

Therefore, $\|y - \bar{x}\| \leq \|y - x\|$.

Finally, we prove the existence of a minimizing point $\bar{x}$ satisfying (1.3). Choose a closed ball $\mathrm{cl}\,(\mathbb{B}_{\alpha+1}(y))$, where $\alpha = \inf\limits_{x \in K} \|y - x\|$. As $K \cap \mathrm{cl}\,(\mathbb{B}_{\alpha+1}(y))$ is compact, the continuous function $\|y - x\|$ has its minimum at a point $\bar{x}$. It is also evident that $\|y - \bar{x}\| = \alpha$.

Now we verify the uniqueness of the point $\bar{x}$. Suppose that $\hat{x}$ is another point satisfying the inequality (1.4). Then, choosing $x = \hat{x}$ in the inequality (1.4) and applying the same relation to $\hat{x}$ and choosing $x = \bar{x}$, it yields

$$\langle y - \bar{x}, \hat{x} - \bar{x} \rangle \leq 0 \quad \text{and} \quad \langle y - \hat{x}, \bar{x} - \hat{x} \rangle \leq 0.$$

By adding these two inequalities, we obtain

$$\langle \bar{x} - \hat{x}, \bar{x} - \hat{x} \rangle = \|\bar{x} - \hat{x}\|^2 \leq 0,$$

which implies that $\hat{x} = \bar{x}$. □

Remark 1.9. The inequality (1.4) shows that $y - \bar{x}$ and $x - \bar{x}$ subtend a nonacute angle between them. The projection $P_K(y)$ of y on K can be interpreted as the result of applying to y the operator $P_K : \mathbb{R}^n \to K$, which is called *projection operator*. Note that $P_K(x) = x$ for all $x \in K$.

Corollary 1.3. Let K be a nonempty closed convex subset of $\mathbb{R}^n$. Then, for all $x, y \in \mathbb{R}^n$,

$$\|P_K(x) - P_K(y)\| \leq \|x - y\|, \tag{1.5}$$

that is, the projection operator P_K is nonexpansive. In particular, P_K is continuous on K.

Proof. Given $x, y \in \mathbb{R}^n$, let $u = P_K(x)$ and $v = P_K(y)$. Then, $u, v \in K$, and by (1.4), we obtain

$$\langle x - u, z - u \rangle \leq 0, \quad \text{for all } z \in K,$$

$$\langle y - v, z - v \rangle \leq 0, \quad \text{for all } z \in K.$$

Choosing $z = v$ in the first inequality and $z = u$ in the second and adding the resultants, we obtain

$$\|u - v\|^2 \leq \langle x - y, u - v \rangle \leq \|x - y\| \, \|u - v\|, \tag{1.6}$$

that is, $\|u - v\| \leq \|x - y\|$, which holds even when $\|u - v\| = 0$. $\square$

The geometric interpretation of the nonexpansivity of P_K is given in the following figure. We observe that if strict inequality holds in (1.5), then the projection operator P_K reduces the distance. However, if the equality holds in (1.5) then the distance is conserved.

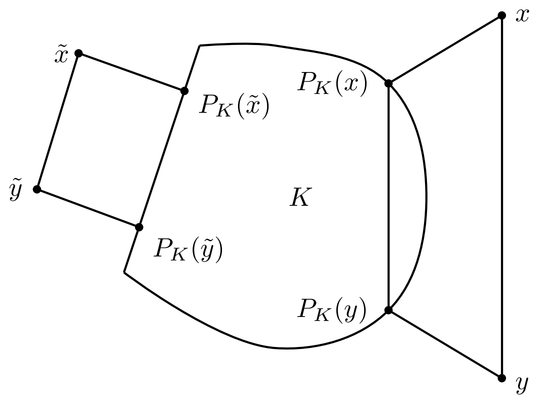


FIGURE 1.13: The nonexpansiveness of the projection operator

Remark 1.10. If K is a closed convex set and y is a point outside the set K, then we can find a hyperplane such that the convex set K lies in one of the half-spaces generated by the hyperplane and point y lies in the other open half-space.

Theorem 1.6. If K is a nonempty closed convex subset of $\mathbb{R}^n$ and $y \notin K$, then there exists a nonzero vector $a \in \mathbb{R}^n$ such that $\inf_{x \in K} \langle a, x \rangle > \langle a, y \rangle$.

Proof. By Theorem 1.5, there exists a unique projection $\bar{x} \in K$ of y on K such that for all $x \in K$,

$$\langle y - \bar{x}, x - \bar{x} \rangle \leq 0,$$

that is,

$$\langle x, y - \bar{x} \rangle \leq \langle \bar{x}, y - \bar{x} \rangle. \tag{1.7}$$

Note that $y \neq \bar{x}$, and so,

$$0 < \|y - \bar{x}\|^2 = \langle y - \bar{x}, y - \bar{x} \rangle = \langle y, y - \bar{x} \rangle - \langle \bar{x}, y - \bar{x} \rangle. \tag{1.8}$$

Combining the inequalities (1.7) and (1.8), we obtain

$$\langle x, y - \bar{x} \rangle \leq \langle \bar{x}, y - \bar{x} \rangle < \langle y, y - \bar{x} \rangle, \quad \text{for all } x \in K.$$

Letting $\bar{x} - y = a$, it follows that

$$\langle a, x \rangle \geq \langle a, \bar{x} \rangle > \langle a, y \rangle, \quad \text{for all } x \in K,$$

and the desired result follows. □

The following theorem is a consequence of the previous result and may be called "theorem of the supporting hyperplane."

Theorem 1.7. If y is a point on the boundary of a nonempty convex subset K of $\mathbb{R}^n$, then there exists a nonzero vector $a \in \mathbb{R}^n$ such that $\inf_{x \in K} \langle a, x \rangle = \langle a, y \rangle$, that is, there exists a hyperplane $H = \{x \in \mathbb{R}^n : \langle a, x - y \rangle = 0\}$ that supports K at y.

Proof. Since $y \in \mathrm{b}(K)$, there exists a sequence $\{x_m\}$ outside $\mathrm{cl}(K)$ such that $x_m \to y$. From the previous theorem, there exists a sequence $\{a_m\} \subseteq S = \{a \in \mathbb{R}^n : \|a\| = 1\}$ such that $\inf_{x \in K} \langle a_m, x \rangle > \langle a_m, x_m \rangle$ for each m, and so, $\langle a_m, x \rangle > \langle a_m, x_m \rangle$ for all m and all $x \in K$. Since S is compact, there exists a converging subsequence $a_{m_i} \to a$, and therefore, we have

$$\langle a, x \rangle = \lim_{i \to \infty} \langle a_{m_i}, x \rangle \geq \lim_{i \to \infty} \langle a_{m_i}, x_{m_i} \rangle = \langle a, y \rangle, \quad \text{for all } x \in K.$$

This completes the proof. □

Remark 1.11. The preceding result says that at each boundary point of a convex set there is a supporting hyperplane passing through it.

1.5 Convex Functions

This section deals with the notion of a convex function and some of its properties.

Definition 1.31. Let K be a nonempty convex subset of $\mathbb{R}^n$. A function $f : K \to \mathbb{R}$ is said to be

(a) *convex* if for all $x, y \in K$ and all $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y);$$

(b) *strictly convex* if for all $x, y \in K$, $x \neq y$ and all $\lambda \in\ ]0,1[$,

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y).$$

A function $f : K \to \mathbb{R}$ is said to be (*strictly*) *concave* if $-f$ is (strictly) convex.

Geometrically speaking, a function $f : K \to \mathbb{R}$ defined on a convex subset K of $\mathbb{R}^n$ is convex if the line segment joining any two points on the graph of the function lies on or above the portion of the graph between these points. Similarly, f is concave if the line segment joining any two points on the graph of the function lies on or below the portion of the graph between these points. Also, a function for which the line segment joining any two points on the graph of the function lies strictly above the portion of the graph between these points is referred to as strictly convex function.

Some of the examples of convex functions defined on $\mathbb{R}$ are $f(x) = e^x$, $f(x) = x$, $f(x) = |x|$, $f(x) = \max\{0, x\}$. The functions $f(x) = -\log x$ and $f(x) = x^\alpha$ for $\alpha < 0$, $\alpha > 1$ are strictly convex defined on the interval $]0, \infty[$. Clearly, every strictly convex function is convex but the converse may not be true. For example, the function $f(x) = x$ defined on $\mathbb{R}$ is not strictly convex. The function $f(x) = |x + x^3|$ is a nondifferentiable strictly convex function on $\mathbb{R}$.

Proposition 1.1. A function $f : K \to \mathbb{R}$ defined on a nonempty convex subset K of $\mathbb{R}^n$ is convex if and only if its epigraph is a convex set.

Proof. Let f be a convex function on K, and let (x, α), $(y, \beta) \in \mathrm{epi}(f)$, then

$$f(x) \leq \alpha \quad \text{and} \quad f(y) \leq \beta. \tag{1.9}$$

Since f is a convex function, we have for $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \leq \lambda\alpha + (1-\lambda)\beta,$$

that is, $\lambda(x,\alpha)+(1-\lambda)(y,\beta)\in \mathrm{epi}(f)$ for all $\lambda\in[0,1]$.

Conversely, as $(x,f(x))$, $(y,f(y))\in \mathrm{epi}(f)$ and $\mathrm{epi}(f)$ is a convex set, it follows that

$$\lambda(x,f(x))+(1-\lambda)(y,f(y))\in \mathrm{epi}(f),\quad \text{for all } \lambda\in[0,1],$$

which implies that $f(\lambda x+(1-\lambda)y)\le \lambda f(x)+(1-\lambda)f(y)$ for all $\lambda\in[0,1]$. $\square$

Remark 1.12. It can be easily seen that a function $f: K\to\mathbb{R}$ defined on a convex subset K of $\mathbb{R}^n$ is concave if and only if its hypograph is a convex set.

If f is a convex function defined on K, then for every $\alpha\in\mathbb{R}$ the lower level set $L(f,\alpha)=\{x\in K: f(x)\le\alpha\}$ is also convex. However, the converse is not true. For example, the function $f(x)=x^3$ defined on $\mathbb{R}$ is not convex, but the lower level set $L(f,\alpha)=\{x\in\mathbb{R}: x\le\alpha^{1/3}\}$ is convex for all $\alpha\in\mathbb{R}$.

Theorem 1.8. Let K be a nonempty convex subset of $\mathbb{R}^n$. A function $f: K\to\mathbb{R}$ is convex if and only if for all $x_1,x_2,\ldots,x_m\in K$ and $\lambda_i\in[0,1]$, $i=1,2,\ldots,m$ with $\sum_{i=1}^m\lambda_i=1$,

$$f\left(\sum_{i=1}^m \lambda_i x_i\right)\le\sum_{i=1}^m\lambda_i f(x_i). \tag{1.10}$$

The inequality (1.10) is called *Jensen's inequality.*

Proof. Suppose that the Jensen's inequality (1.10) holds. Then trivially, f is convex.

Conversely, we assume that the function f is convex. Then, we show that the Jensen's inequality (1.10) holds. We prove it by induction on m. For $m=1$ and $m=2$, the inequality (1.10) trivially holds. Assume that the inequality (1.10) holds for m. We shall prove the result for $m+1$. If $\lambda_{m+1}=1$, the result holds because $\lambda_i=0$, for $i=1,2,\ldots,m$ and the result is true for $m=1$. If $\lambda_{m+1}\ne 1$, we have

$$\begin{aligned}
f\left(\sum_{i=1}^{m+1}\lambda_i x_i\right) &= f\left(\sum_{i=1}^{m}\lambda_i x_i+\lambda_{m+1}x_{m+1}\right)\\
&= f\left(\sum_{i=1}^{m}(1-\lambda_{m+1})\frac{\lambda_i x_i}{1-\lambda_{m+1}}+\lambda_{m+1}x_{m+1}\right)\\
&= f\left((1-\lambda_{m+1})\sum_{i=1}^{m}\frac{\lambda_i}{1-\lambda_{m+1}}x_i+\lambda_{m+1}x_{m+1}\right)\\
&= f\left((1-\lambda_{m+1})\sum_{i=1}^{m}\mu_i x_i+\lambda_{m+1}x_{m+1}\right)
\end{aligned}$$

$$\leq \quad (1-\lambda_{m+1}) f\left(\sum_{i=1}^{m} \mu_i x_i\right) + \lambda_{m+1} f(x_{m+1})$$
$$\leq \quad (1-\lambda_{m+1}) \sum_{i=1}^{m} \mu_i f(x_i) + \lambda_{m+1} f(x_{m+1}),$$

where $\mu_i = \dfrac{\lambda_i}{(1-\lambda_{m+1})}$, $i = 1, 2, \ldots, m$ with $\mu_i \geq 0$ for $i = 1, 2, \ldots, m$ and

$$\sum_{i=1}^{m} \mu_i = \frac{\sum_{i=1}^{m} \lambda_i}{1-\lambda_{m+1}} = \frac{1-\lambda_{m+1}}{1-\lambda_{m+1}} = 1.$$

This completes the proof. □

Theorem 1.9. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a function. Then, f is convex if and only if for every $x \in K$, there exists a vector $\xi \in \mathbb{R}^n$ such that

$$\langle \xi, y - x \rangle \leq f(y) - f(x), \quad \text{for all } y \in K. \tag{1.11}$$

Proof. Let f be convex and $x \in K$. Then, epi(f) is convex and the point $(x, f(x))$ belongs to the boundary of epi(f). By Theorem 1.7, there exists a nonzero vector $(\xi_0, \mu) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\langle (\xi_0, \mu), (y, \alpha) \rangle \geq \langle (\xi_0, \mu), (x, f(x)) \rangle, \quad \text{for all } (y, \alpha) \in \text{epi}(f),$$

equivalently,

$$\langle \xi_0, y \rangle + \mu\alpha \geq \langle \xi_0, x \rangle + \mu f(x), \quad \text{for all } (y, \alpha) \in \text{epi}(f),$$

that is,

$$\langle \xi_0, y - x \rangle \geq \mu(f(x) - \alpha), \quad \text{for all } (y, \alpha) \in \text{epi}(f). \tag{1.12}$$

If $\mu = 0$, then $\langle \xi_0, y - x \rangle \geq 0$ for all $y \in K$. Since x belongs to the open set K, there exists $\delta > 0$ such that $x - \delta\xi_0 \in K$, and hence,

$$\langle \xi_0, x - \delta\xi_0 - x \rangle \geq 0, \quad \text{equivalently,} \quad \delta\langle \xi_0, \xi_0 \rangle \leq 0.$$

Therefore, $\delta\|\xi_0\|^2 \leq 0$, and hence, $\xi_0 = \mathbf{0}$. Thus, $(\xi_0, \mu) = (\mathbf{0}, 0)$, contradicting the fact that (ξ_0, μ) is a nonzero vector.

If $\mu < 0$, then it is possible to take α sufficiently large in order to have

$$\langle \xi_0, y - x \rangle < \mu(f(x) - \alpha),$$

in contradiction to the inequality (1.12).

Therefore, $\mu > 0$. Choose $\alpha = f(y)$ and $\xi = -\xi_0/\mu$. Then from the inequality (1.12), we obtain

$$\langle \xi, y - x \rangle \leq f(y) - f(x).$$

Conversely, let $x, y \in K$ and $\lambda \in\ ]0,1[$. Since K is convex, $\lambda x + (1-\lambda)y \in K$. By the hypothesis, we obtain

$$\langle \xi, \lambda x + (1-\lambda)y - x\rangle \le f(\lambda x + (1-\lambda)y) - f(x), \tag{1.13}$$

and

$$\langle \xi, \lambda x + (1-\lambda)y - y\rangle \le f(\lambda x + (1-\lambda)y) - f(y). \tag{1.14}$$

Multiplying inequality (1.13) by λ and inequality (1.14) by $(1-\lambda)$, and then adding the resultants, we get

$$f(\lambda x + (1-\lambda y)) \le \lambda f(x) + (1-\lambda) f(y).$$

Hence, f is convex. □

Corollary 1.4. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a function. Then, f is strictly convex if and only if for every $x \in K$, there exists a vector $\xi \in \mathbb{R}^n$ such that

$$\langle \xi, y - x\rangle < f(y) - f(x), \quad \text{for all } y \in K,\ y \ne x.$$

Proof. Since every strictly convex function is convex, by Theorem 1.9, there exists a vector $\xi \in \mathbb{R}^n$ such that

$$\langle \xi, y - x\rangle \le f(y) - f(x), \quad \text{for all } y \in K. \tag{1.15}$$

Suppose, on the contrary, there exists $\hat{x} \ne x$ such that $\langle \xi, \hat{x} - x\rangle = f(\hat{x}) - f(x)$. Then by strict convexity of f for all $\lambda \in\ ]0,1[$, we have

$$f(\lambda x + (1-\lambda)\hat{x}) < \lambda f(x) + (1-\lambda) f(\hat{x}) = f(x) + (1-\lambda)\langle \xi, \hat{x} - x\rangle. \tag{1.16}$$

Letting $y = \lambda x + (1-\lambda)\hat{x}$ in (1.15), we get

$$f(\lambda x + (1-\lambda)\hat{x}) \ge f(x) + (1-\lambda)\langle \xi, \hat{x} - x\rangle,$$

which contradicts (1.16).

Converse follows as in Theorem 1.9. □

Remark 1.13. If in Theorem 1.9, the set K is not open then the sufficient part of this theorem may not be true. That is, if for every $x \in K$, there exists a vector $\xi \in \mathbb{R}^n$ such that

$$\langle \xi, y - x\rangle \le f(y) - f(x), \quad \text{for all } y \in K,$$

then f is not necessarily convex.

Indeed, consider the set $K = \{(x_1, x_2) \in \mathbb{R}^2 : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1\}$ and the function $f : K \to \mathbb{R}$ defined by

$$f(x_1, x_2) = \begin{cases} 0, & \text{if } 0 \le x_1 \le 1,\ 0 < x_2 \le 1, \\ \frac{1}{4} - \left(x_1 - \frac{1}{2}\right)^2, & \text{if } 0 \le x_1 \le 1,\ x_2 = 0. \end{cases}$$

Then, for each point in the interior of K, $\xi = \mathbf{0}$ satisfies the inequality (1.11). However, f is not convex on K since $\operatorname{epi}(f)$ is clearly not a convex set.

The next result provides a characterization for a differentiable convex function.

Theorem 1.10. [19, 153] Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a differentiable function. Then,

(a) f is convex if and only if for all $x, y \in K$,

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x). \tag{1.17}$$

(b) f is strictly convex if and only if the inequality is strict in (1.17) for $x \neq y$.

Proof. (a) If f is a convex function, then for all $\lambda \in [0, 1]$

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y).$$

For $\lambda > 0$, we have

$$\frac{f((1-\lambda)x + \lambda y) - f(x)}{\lambda} \leq f(y) - f(x),$$

which on taking limit $\lambda \to 0^+$ leads to (1.17) as f is a differentiable function.

Conversely, let $\lambda \in [0, 1]$ and $u, v \in K$. On taking $x = (1-\lambda)u + \lambda v$ and $y = u$ in (1.17), we have

$$\lambda \langle \nabla f((1-\lambda)u + \lambda v), u - v \rangle \leq f(u) - f((1-\lambda)u + \lambda v). \tag{1.18}$$

Similarly, on taking $x = (1-\lambda)u + \lambda v$ and $y = v$ in (1.17), we have

$$-(1-\lambda)\langle \nabla f((1-\lambda)u + \lambda v), u - v \rangle \leq f(v) - f((1-\lambda)u + \lambda v). \tag{1.19}$$

Multiplying inequality (1.18) by $(1-\lambda)$ and inequality (1.19) by λ, and then adding the resultants, we obtain

$$f((1-\lambda)u + \lambda v) \leq (1-\lambda)f(u) + \lambda f(v).$$

(b) Suppose that f is strictly convex and $x, y \in K$ be such that $x \neq y$. Since f is convex, the inequality (1.17) holds. We need to show the inequality is strict. Suppose on the contrary that

$$\langle \nabla f(x), y - x \rangle = f(y) - f(x).$$

Then for $\lambda \in \;]0, 1[$, we have

$$f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y) = f(x) + \lambda \langle \nabla f(x), y - x \rangle.$$

Let $z = (1-\lambda)x + \lambda y$, then $z \in K$ and the above inequality can be written as

$$f(z) < f(x) + \langle \nabla f(x), z - x \rangle,$$

which contradicts the inequality (1.17). Proof of the converse part follows as given for the convex case. □

A mathematical model for a classical optimization problem can be constructed by specifying a constraint set K, which consists of the available decisions x, and a cost or objective function $f(x)$ that maps each $x \in K$ into a scalar and represents a measure of undesirability of choosing the decision x. Hence, a general form of a minimization problem (MP) is as follows:

$$\text{(MP)} \qquad \min f(x) \quad \text{subject to } x \in K,$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and the constraint set K is a nonempty subset of $\mathbb{R}^n$. A point $x \in K$ is called a *feasible solution* of (MP) and the set K is called the *feasible set.* If $\bar{x} \in K$ and $f(\bar{x}) \leq f(y)$ for all $y \in K$, then $\bar{x}$ is called a *minimum* or a *global minimum* of (MP). If $\bar{x} \in K$ and if there exists an ε-neighborhood $\mathbb{B}_\varepsilon(\bar{x})$ of $\bar{x}$ such that $f(\bar{x}) \leq f(y)$ for all $y \in K \cap \mathbb{B}_\varepsilon(\bar{x})$, then $\bar{x}$ is called a *local minimum* of (MP). Similarly, if $\bar{x} \in K$ and if $f(\bar{x}) < f(y)$ for all $y \in K \cap \mathbb{B}_\varepsilon(\bar{x})$, then $\bar{x}$ is called a *strict local minimum* of (MP).

Theorem 1.11. Let K be a nonempty compact subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a lower semicontinuous function. Then, the solution set of the minimization problem (MP) is nonempty and compact.

Proof. Let $\{x_m\} \subseteq K$ be a minimizing sequence, that is, $\lim\limits_{m\to\infty} f(x_m) = \inf\{f(x) : x \in K\}$. Since K is compact, there exists a subsequence $\{x_{m_k}\}$ converging to some $x \in K$, and since f is lower semicontinuous, it follows that $f(x) \leq \liminf\limits_{k\to\infty} f(x_{m_k}) = L$, say. Since $f(x) \geq L$, the result follows. □

A similar result for the existence of the maximum of f over K can be established by assuming f to be upper semicontinuous on the compact set K.

We now have some important properties of (strictly) convex functions. One of the most important properties of a differentiable convex function is that every point at which the gradient vanishes is a global minimum.

Theorem 1.12. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a differentiable (strictly) convex function. If $\nabla f(x) = \mathbf{0}$, then x is a (unique) global minimum of f over K; that is, every stationary point of f is a (unique) global minimum of f over K.

Proof. Let x be a stationary point of f, that is, $\nabla f(x) = \mathbf{0}$. Since f is a differentiable convex function, by Theorem 1.10(a), we have

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x), \quad \text{for all } y \in K,$$

which by stationarity condition yields $f(x) \leq f(y)$ for all $y \in K$. The uniqueness of the global minimum for the strictly convex case follows by Theorem 1.10(b). □

We observe here that a stationary point of a function can be a global minimum without the function being convex. For example, $x = 0$ is a stationary point of the nonconvex function $f(x) = 1 - e^{-x^2}$ defined on $\mathbb{R}$.

Theorem 1.13. Every local minimum of a (strictly) convex function $f : K \to \mathbb{R}$ defined on a convex set $K \subseteq \mathbb{R}^n$ is a (unique) global minimum of f over K. Further, the set of points at which a convex function attains its global minimum on K is a convex set.

Proof. Let $x \in K$ be a local minimum of f over K. Then, there exists $\delta > 0$ such that

$$f(x) \leq f(y), \quad \text{for all } y \in K \cap \mathbb{B}_\delta(x). \tag{1.20}$$

Suppose, on the contrary, x is not a global minimum over K. Then, there exists $x_0 \in K$, $x_0 \neq x$ such that $f(x_0) < f(x)$. By convexity of f, we have for $\lambda \in \ ]0, 1]$,

$$f((1-\lambda)x + \lambda x_0) \leq (1-\lambda)f(x) + \lambda f(x_0) < f(x).$$

This contradicts inequality (1.20) since $(1-\lambda)x + \lambda x_0 \in K \cap \mathbb{B}_\delta(x)$ for λ sufficiently small.

If f is a strictly convex function, we need to show that x is a unique minimum. On the contrary, suppose there exists $\hat{x} \in K$, $\hat{x} \neq x$ such that $\hat{x}$ is also a global minimum of f over K. Then, it is clear that $f(x) = f(\hat{x})$. By strict convexity of f, we have for $\lambda \in \ ]0, 1[$

$$f((1-\lambda)x + \lambda\hat{x}) < (1-\lambda)f(x) + \lambda f(\hat{x}) = f(x).$$

This contradicts the fact that x is a global minimum of f over K.

Let $S = \{x \in K : f(x) \leq f(y) \text{ for all } y \in K\}$ be the set of points at which f attains its global minimum. If $x_1, x_2 \in S$, then $f(x_1) \leq f(y)$ and $f(x_2) \leq f(y)$ for every $y \in K$. By convexity of f, for all $\lambda \in [0, 1]$ and for every $y \in K$, we have

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2) \leq (1-\lambda)f(y) + \lambda f(y) = f(y),$$

which implies that $(1-\lambda)x_1 + \lambda x_2 \in S$. Hence, S is convex. $\square$

Definition 1.32. Let K be a nonempty subset of $\mathbb{R}^n$ and $x \in K$ be a given point. A function $f : K \to \mathbb{R}$ is said to be *locally Lipschitz around* x if for some $k > 0$

$$|f(y) - f(z)| \leq k\|y - z\|, \quad \text{for all } y, z \in N(x) \cap K, \tag{1.21}$$

where $N(x)$ is a neighborhood of x. The constant k here is called the *Lipschitz constant* and it varies as the point x varies.

The function f is said to be *Lipschitz continuous* on K if there exists a real number $k > 0$ such that inequality (1.21) holds for all $y, z \in K$.

The class of Lipschitz continuous functions is quite large. It is invariant under usual operations of sum, product, and quotient.

It is clear that every Lipschitz continuous function is continuous. Also, every convex function is not only continuous but also locally Lipschitz in the interior of its domain.

Theorem 1.14. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a convex function and x be an interior point of K. Then, f is locally Lipschitz at x.

Proof. Without loss of generality, we may assume that $x = \mathbf{0}$. We divide the proof into three steps.

(i) We prove that the function f is bounded above in a neighborhood of $x = \mathbf{0}$.

Choose a system of $(m+1)$ affinely independent vectors $a_1, a_2, \ldots, a_{m+1} \in \mathbb{R}^n$ so small that the set $U := \operatorname{int}(\operatorname{co}\{a_1, a_2, \ldots, a_{m+1}\})$ contains $\mathbf{0}$ and is contained in the domain of f. Set $k := \max\{f(a_1), f(a_2), \ldots, f(a_{m+1})\}$. Then, any $y \in U$ can be expressed as a convex combination of $a_1, a_2, \ldots, a_{m+1}$ as $y = \sum_{i=1}^{m+1} \lambda_i a_i$ with $\lambda_i \geq 0$ for all $i = 1, 2, \ldots, m+1$ and $\sum_{i=1}^{m+1} \lambda_i = 1$. Then by convexity of f, we have

$$f(y) \leq \sum_{i=1}^{m+1} \lambda_i f(a_i) \leq k.$$

(ii) We show that f is bounded below in a neighborhood of $x = \mathbf{0}$.

Choose a positive number δ so small that $\mathbb{B}_{2\delta}(\mathbf{0}) \subseteq U$. For each $y \in \mathbb{B}_{2\delta}(\mathbf{0})$, we have $-y \in \mathbb{B}_{2\delta}(\mathbf{0})$. Since $\mathbf{0} = (y + (-y))/2$, by convexity of f, we have

$$f(\mathbf{0}) \leq \frac{1}{2} f(y) + \frac{1}{2} f(-y) \leq \frac{1}{2} f(y) + \frac{1}{2} k.$$

It follows that f is bounded below by $2f(\mathbf{0}) - k$ on the set $\mathbb{B}_{2\delta}(\mathbf{0})$. Hence, in view of (i), f is bounded near $x = \mathbf{0}$.

(iii) Finally, we prove that f is Lipschitz continuous on $\mathbb{B}_{\delta}(\mathbf{0})$.

Let β be a bound of $|f(y)|$ on $\mathbb{B}_{2\delta}(\mathbf{0})$. Let x_1 and x_2 be two arbitrary distinct points of the set $\mathbb{B}_{\delta}(\mathbf{0})$. Then,

$$x_3 := x_2 + \frac{\delta}{\|x_2 - x_1\|}(x_2 - x_1) \in \mathbb{B}_{2\delta}(\mathbf{0}).$$

Solving for x_2 yields

$$x_2 = \frac{\delta}{\|x_2 - x_1\| + \delta} x_1 + \frac{\|x_2 - x_1\|}{\|x_2 - x_1\| + \delta} x_3.$$

By convexity of f, we have

$$f(x_2) \leq \frac{\delta}{\|x_2 - x_1\| + \delta} f(x_1) + \frac{\|x_2 - x_1\|}{\|x_2 - x_1\| + \delta} f(x_3),$$

and thus,

$$f(x_2) - f(x_1) \leq \frac{\|x_2 - x_1\|}{\|x_2 - x_1\| + \delta} \left(f(x_3) - f(x_1)\right) \leq \gamma \|x_2 - x_1\|,$$

where $\gamma = (2k)/\delta$ is a constant independent of x_1 and x_2. By interchanging the roles of x_1 and x_2, we obtain the Lipschitz property of f on $\mathbb{B}_{\delta}(\mathbf{0})$. $\square$

Remark 1.14. It is clear from Theorem 1.14 that a convex function is continuous in the interior of its domain. However, a convex function need not be continuous. For example, consider the function $f : [-1, 1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2, & \text{if } -1 \leq x < 1, \\ 3, & \text{if } x = 1. \end{cases}$$

Here, f is a convex function on $[-1, 1]$ but not continuous at the boundary point $x = 1$.

The following theorem provides some properties of convex functions. The proof of the theorem is quite trivial, and hence omitted.

Theorem 1.15. Let K be a nonempty convex subset of $\mathbb{R}^n$.

(a) If $f_1, f_2 : K \to \mathbb{R}$ are two convex functions, then $f_1 + f_2$ is a convex function on K.

(b) If $f : K \to \mathbb{R}$ is a convex function and $\alpha \geq 0$, then αf is a convex function on K.

(c) For each $i = 1, 2, \ldots, m$, if $f_i : K \to \mathbb{R}$ is a convex function and $\alpha_i \geq 0$, then $\sum_{i=1}^m \alpha_i f_i$ is a convex function on K. Further, if at least one of the functions f_i is strictly convex with the corresponding $\alpha_i > 0$, then $\sum_{i=1}^m \alpha_i f_i$ is strictly convex on K.

Theorem 1.16. Let K be a nonempty convex subset of $\mathbb{R}^n$. For each $i = 1, 2, \ldots, m$, if $f_i : K \to \mathbb{R}$ is a convex function, then $\max\{f_1, f_2, \ldots, f_m\}$ is also a convex function on K.

Proof. Since f_i is a convex function for each $i = 1, 2, \ldots, m$, $\operatorname{epi}(f_i)$ is a convex set for each $i = 1, 2, \ldots, m$. If $f = \max\{f_1, f_2, \ldots, f_m\}$, then it can be seen that the $\operatorname{epi}(f) = \bigcap_{i=1}^m \operatorname{epi}(f_i)$. Since the intersection of convex sets is convex, it follows that $\operatorname{epi}(f)$ is a convex set, and hence $f = \max\{f_1, f_2, \ldots, f_m\}$ is a convex function. $\square$

Now, we give a characterization for differentiable convex functions.

Theorem 1.17. [77, 153] Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a differentiable function. Then,

(a) f is convex if and only if for all $x, y \in K$,

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0;$$

(b) f is strictly convex if and only if for all $x, y \in K$, $x \neq y$,

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle > 0.$$

Proof. Let f be a differentiable convex function, then by Theorem 1.10(a), we have

$$\langle \nabla f(x), y-x \rangle \leq f(y) - f(x), \quad \text{for all } x, y \in K.$$

By interchanging the roles of x and y, we have

$$\langle \nabla f(y), x-y \rangle \leq f(x) - f(y), \quad \text{for all } x, y \in K.$$

On adding the above inequalities we get the conclusion.

Conversely, by mean value theorem, for all $x, y \in K$, there exists $z = (1-\lambda)x + \lambda y$ for some $\lambda \in \ ]0,1[$ such that

$$\begin{aligned} f(y) - f(x) &= \langle \nabla f(z), y-x \rangle = (1/\lambda)\langle \nabla f(z), z-x \rangle \\ &\geq (1/\lambda)\langle \nabla f(x), z-x \rangle = \langle \nabla f(x), y-x \rangle, \end{aligned}$$

where the above inequality is obtained on using the given hypothesis. Hence, by Theorem 1.10(a), f is a convex function.

The proof for the strict convex case follows on the same lines by using Theorem 1.10(b). $\square$

The following example illustrates Theorem 1.17.

Example 1.1. The function $f(x) = x_1^2 + x_2^2$, where $x = (x_1, x_2) \in \mathbb{R}^2$ is a convex function on $\mathbb{R}^2$ and $\nabla f(x) = 2(x_1, x_2)$. For $x, y \in \mathbb{R}^2$,

$$\begin{aligned} \langle \nabla f(y) - \nabla f(x), y-x \rangle &= \langle 2(y_1 - x_1, y_2 - x_2), (y_1 - x_1, y_2 - x_2) \rangle \\ &= 2(y_1 - x_1)^2 + 2(y_2 - x_2)^2 \geq 0. \end{aligned}$$

The next theorem provides a characterization for a twice differentiable convex function in terms of its Hessian matrix. However, for a strictly convex function we have only a one way implication.

Theorem 1.18. Let K be a nonempty open convex subset of $\mathbb{R}^n$.

(a) A twice differentiable function $f : K \to \mathbb{R}$ is convex if and only if its Hessian matrix $H(x) = \nabla^2 f(x)$ is positive semidefinite for every $x \in K$.

(b) If the Hessian matrix $H(x)$ of a twice differentiable function $f : K \to \mathbb{R}$ is positive definite for every $x \in K$, then f is strictly convex.

Proof. (a) By the Taylor expansion of f, we have for $x, y \in K$,

$$f(y) = f(x) + \langle \nabla f(x), y-x \rangle + (1/2)\langle y-x, H(x+\lambda(y-x))(y-x) \rangle, \quad (1.22)$$

for some $\lambda \in \ ]0,1[$. If the Hessian matrix is positive semidefinite, we have

$$\langle y-x, H(x+\lambda(y-x))(y-x) \rangle \geq 0,$$

which implies that $f(y) - f(x) \geq \langle \nabla f(x), y-x \rangle$, that is, f is a convex function.

Conversely, suppose that the Hessian matrix is not positive semidefinite at some point $x \in K$, then there exists $y \in K$ such that $\langle y-x, H(x)(y-x)\rangle < 0$. By continuity of the Hessian matrix, we may select y so close to x that for all $\lambda \in [0,1]$, $\langle y-x, H(x+\lambda(y-x))(y-x)\rangle < 0$. In view of relation (1.22), it follows that $f(y) - f(x) < \langle \nabla f(x), y-x\rangle$, which contradicts the assumption that f is a convex function.

(b) If $H(x)$ is positive definite for every $x \in K$, then it can be easily established that f is strictly convex by using relation (1.22) and Theorem 1.10(b). □

For the function considered in Example 1.1, the Hessian matrix

$$H(x) = \nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

is positive semidefinite as $\langle y, H(x)y\rangle = 2\left(y_1^2 + y_2^2\right) \geq 0$ for all $y \in \mathbb{R}^2$.

The converse of Theorem 1.18(b) is not true. For example, the function $f(x) = (x_1 - 1)^4 + x_2^2$ defined on $\mathbb{R}^2$ is strictly convex but the Hessian matrix

$$H(x) = \nabla^2 f(x) = \begin{bmatrix} 12(x_1-1)^2 & 0 \\ 0 & 2 \end{bmatrix}$$

is not positive definite for $x = (1,1)$ as $\langle y, H(x)y\rangle = 12(x_1-1)^2 y_1^2 + 2y_2^2 = 0$ for $y = (1,0) \in \mathbb{R}^2$.

The following definition of strong convexity is a stronger version of strict convexity.

Definition 1.33. A function $f : K \to \mathbb{R}$ defined on a nonempty convex subset K of $\mathbb{R}^n$ is said to be *strongly convex with modulus* ρ if there exists a real number $\rho > 0$ such that for all $x, y \in K$ and $\lambda \in [0,1]$, we have

$$f((1-\lambda)x + \lambda y) + \rho\lambda(1-\lambda)\|y-x\|^2 \leq (1-\lambda)f(x) + \lambda f(y). \tag{1.23}$$

Clearly, every strongly convex function is strictly convex. However, the converse is not true. For example, the function $f(x) = x^4$ defined on $\mathbb{R}$ is strictly convex but it is not strongly convex as there is no $\rho > 0$ that satisfies the inequality (1.23) for every $x, y \in K$.

The following result provides a characterization for a differentiable strongly convex function.

Theorem 1.19. A differentiable function $f : K \to \mathbb{R}$ defined on a nonempty open convex subset K of $\mathbb{R}^n$ is strongly convex with modulus ρ if and only if

$$\langle \nabla f(x), y-x\rangle + \rho\|y-x\|^2 \leq f(y) - f(x), \quad \text{for all } x, y \in K. \tag{1.24}$$

Proof. If f is a strongly convex function with modulus ρ, then for all $\lambda \in [0,1]$, we have

$$f((1-\lambda)x + \lambda y) + \rho\lambda(1-\lambda)\|y-x\|^2 \leq (1-\lambda)f(x) + \lambda f(y).$$

For $\lambda > 0$, we have

$$\frac{f((1-\lambda)x+\lambda y)-f(x)}{\lambda}+\rho(1-\lambda)\|y-x\|^2 \leq f(y)-f(x),$$

which on taking limit $\lambda \to 0^+$ leads to (1.24) as f is a differentiable function.

Conversely, let $\lambda \in [0,1]$ and $u, v \in K$. By taking $x = (1-\lambda)u + \lambda v$ and $y = u$ in (1.24), we obtain

$$\lambda\langle \nabla f((1-\lambda)u+\lambda v), u-v\rangle + \rho\lambda^2\|u-v\|^2 \leq f(u) - f((1-\lambda)u+\lambda v). \quad (1.25)$$

Similarly, by taking $x = (1-\lambda)u + \lambda v$ and $y = v$ in (1.24), we obtain

$$-(1-\lambda)\langle \nabla f((1-\lambda)u+\lambda v), u-v\rangle + \rho(1-\lambda)^2\|u-v\|^2 \leq f(v) - f((1-\lambda)u+\lambda v). \quad (1.26)$$

Multiplying inequality (1.25) by $(1-\lambda)$ and inequality (1.26) by λ and then adding the resultants, we get

$$f((1-\lambda)u+\lambda v) + \rho\lambda(1-\lambda)\|u-v\|^2 \leq (1-\lambda)f(u) + \lambda f(v).$$

Hence, f is strongly convex. □

The following result was established by Zhu and Marcotte [214].

Theorem 1.20. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be continuously differentiable. If the gradient ∇f is Lipschitz continuous over K with Lipschitz constant k, then

$$\langle \nabla f(x), y-x\rangle + \frac{k}{2}\|y-x\|^2 \geq f(y) - f(x), \quad \text{for all } x, y \in K. \quad (1.27)$$

Proof. For all $x, y \in K$, define a function $\varphi : [0,1] \to \mathbb{R}$ by

$$\varphi(t) = f((1-t)x+ty) - f(x) - t\langle \nabla f(x), y-x\rangle, \quad \text{for all } t \in [0,1].$$

Then, we have

$$\varphi'(t) = \langle \nabla f((1-t)x+ty) - \nabla f(x), y-x\rangle.$$

Using the fundamental theorem of integral calculus and the fact $\phi(0) = 0$, we have

$$\begin{aligned}
\varphi(1) &= f(y) - f(x) - \langle \nabla f(x), y-x\rangle \\
&= \varphi(0) + \int_0^1 \varphi'(t)dt \\
&\leq \int_0^1 |\varphi'(t)|\, dt \\
&\leq \int_0^1 k\|t(y-x)\| \cdot \|y-x\| dt \\
&= \frac{k}{2}\|y-x\|^2.
\end{aligned}$$

□

In Theorem 1.18(b), it was observed that positive definiteness of the Hessian matrix of a function is a sufficient condition for the function to be strictly convex. The next theorem illustrates that positive definiteness of the Hessian matrix is a necessary condition for the strong convexity of a function.

Theorem 1.21. If $f : K \to \mathbb{R}$ is a twice differentiable strongly convex function with modulus ρ on a nonempty open convex subset K of $\mathbb{R}^n$, then its Hessian matrix $H(x) = \nabla^2 f(x)$ is positive definite for every $x \in K$.

Proof. By the Taylor expansion of f, for all $x \in K$, $\mathbf{0} \neq v \in \mathbb{R}^n$ and $\lambda > 0$, we have

$$f(x + \lambda v) = f(x) + \lambda \langle \nabla f(x), v \rangle + (1/2)\lambda^2 \langle v, H(x)v \rangle + o(\lambda^2).$$

Since f is strongly convex with modulus ρ, we have

$$\lambda \langle \nabla f(x), v \rangle + \rho\lambda^2 \|v\|^2 \leq f(x + \lambda v) - f(x),$$

which implies

$$(1/2)\lambda^2 \langle v, H(x)v \rangle + o(\lambda^2) \geq \rho\lambda^2 \|v\|^2.$$

On dividing by λ^2 and taking limit as $\lambda \to 0^+$, we obtain

$$(1/2)\langle v, H(x)v \rangle \geq \rho\|v\|^2 > 0.$$

This completes the proof. □

The converse of Theorem 1.21 may not be true. For example, the Hessian matrix of the function $f(x) = e^x$ defined on $\mathbb{R}$ is positive definite but the function is not strongly convex.

We now give a complete characterization for strongly convex functions.

Theorem 1.22. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A twice differentiable function $f : K \to \mathbb{R}$ is strongly convex with modulus ρ if and only if $H(x) - 2\rho I$ is positive semidefinite for every $x \in K$, where I is the identity matrix.

Proof. By the Taylor expansion of f, for all $x, y \in K$, we have

$$f(y) = f(x) + \langle \nabla f(x), y - x \rangle + (1/2)\langle y - x, H(x + \lambda(y - x))(y - x) \rangle,$$

for some $\lambda \in\ ]0, 1[$. If $H(x) - 2\rho I$ is positive semidefinite, then

$$\langle y - x, (H(x + \lambda(y - x)) - 2\rho I)(y - x) \rangle \geq 0,$$

which implies that $f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle + \rho\|y - x\|^2$.

The converse part can be proved on the lines of the proof of Theorem 1.18(a). □

Theorem 1.23. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A differentiable function $f : K \to \mathbb{R}$ is strongly convex with modulus ρ if and only if the function $f - \rho\|\cdot\|^2$ is convex.

Proof. Observe that for every $x, y \in K$ and $\lambda \in [0, 1]$, the fact that

$$f((1-\lambda)x+\lambda y)-\rho\|(1-\lambda)x+\lambda y\|^2 \leq (1-\lambda)[f(x)-\rho\|x\|^2]+\lambda[f(y)-\rho\|y\|^2],$$

holds if and only if

$$f((1-\lambda)x+\lambda y)-\rho[(1-\lambda)^2\|x\|^2+\lambda^2\|y\|^2+2\lambda(1-\lambda)\langle x,y\rangle] \\ \leq (1-\lambda)[f(x)-\rho\|x\|^2]+\lambda[f(y)-\rho\|y\|^2],$$

which on dividing by $\lambda > 0$ and taking limit leads to

$$\langle \nabla f(x), y-x\rangle + \rho\|y-x\|^2 \leq f(y)-f(x).$$

This completes the proof. □

As pointed out earlier the lower level sets of a convex function are convex. However, for a strongly convex function defined on a closed set the lower level sets are not only convex but compact also.

Theorem 1.24. Let K be a nonempty closed convex subset of $\mathbb{R}^n$. If $f : K \to \mathbb{R}$ is a strongly convex function with modulus ρ, then for every $\alpha \in \mathbb{R}$, the lower level set $L(f, \alpha)$ is a convex compact subset of K.

Proof. If f is a strongly convex function, then clearly it is convex. Hence the lower level set $L(f, \alpha)$ is convex for every $\alpha \in \mathbb{R}$. Since K is a closed set, it is clear that $L(f, \alpha)$ is a closed subset of K. It remains to show that $L(f, \alpha)$ is bounded for every α. If $L(f, \alpha)$ is empty for some α, then the conclusion follows. Assume that $L(f, \alpha)$ is a nonempty set. Choose $x \in L(f, \alpha)$. Suppose, on the contrary, that $L(f, \alpha)$ is not a bounded set. Then, there exists a sequence $\{x_m\}$ in $L(f, \alpha)$ such that $\|x_m\| \to +\infty$. By the strong convexity of f, we have

$$\begin{array}{rcl} \alpha - f(x) & \geq & f(x_m) - f(x) \\ & \geq & \langle \nabla f(x), x_m - x\rangle + \rho\|x_m - x\|^2 \\ & \geq & -\|\nabla f(x)\|\ \|x_m - x\| + \rho\|x_m - x\|^2, \end{array}$$

where the last inequality follows using the Cauchy-Schwarz inequality. For $x_m \neq x$, we have

$$\frac{\alpha - f(x)}{\|x_m - x\|^2} + \frac{\|\nabla f(x)\|}{\|x_m - x\|} \geq \rho.$$

On taking limit as $m \to +\infty$, we have $\rho \leq 0$, which is a contradiction. □

1.6 Generalized Convex Functions

We have seen that the convexity of the lower level sets is only a necessary condition for the convexity of a function but not sufficient. A generalization of convex functions, namely, quasiconvex functions, can be characterized completely in terms of the convexity of the lower level sets.

Definition 1.34. Let K be a nonempty convex subset of $\mathbb{R}^n$. A function $f : K \to \mathbb{R}$ is said to be

(a) *quasiconvex* if for all $x, y \in K$ and all $\lambda \in [0, 1]$,

$$f(\lambda x + (1 - \lambda)y) \leq \max \{f(x), f(y)\} ;$$

(b) *strictly quasiconvex* if for all $x, y \in K$, $x \neq y$ and all $\lambda \in\ ]0, 1[$,

$$f(\lambda x + (1 - \lambda)y) < \max \{f(x), f(y)\} ;$$

(c) *semistrictly quasiconvex* if for all $x, y \in K$, $f(x) \neq f(y)$ and all $\lambda \in\ ]0, 1[$,

$$f(\lambda x + (1 - \lambda)y) < \max \{f(x), f(y)\} .$$

A function $f : K \to \mathbb{R}$ is said to be (*strictly, semistrictly*) *quasiconcave* if $-f$ is (strictly, semistrictly) quasiconvex.

There is a word of caution regarding the definitions of strictly and semistrictly quasiconvex functions. A strictly quasiconvex function was referred to as a strongly quasiconvex function by Avriel [26] and a semistrictly quasiconvex function was referred to as a strictly quasiconvex function by others [26, 30, 153].

It is obvious that every (strictly) convex function is (strictly) quasiconvex but the converse is not necessarily true. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$ is a (strictly) quasiconvex function but not a (strictly) convex function. Also, a convex function is semistrictly quasiconvex but the converse may not be true. Again, we observe that $f(x) = x^3$ is semistrictly quasiconvex but not convex. Also, it must be pointed out here that strict quasiconvexity is not a generalization of convexity as a constant function is convex but not strictly quasiconvex. From the definitions it is obvious that a strictly quasiconvex function is quasiconvex. However, the fact that the converse is not always true is evident as the greatest integer function $f(x) = [x]$ is quasiconvex but not strictly quasiconvex on $\mathbb{R}$.

The following result shows that convexity of the lower level sets is a necessary and sufficient condition for the quasiconvexity of a function.

Theorem 1.25. Let K be a nonempty convex subset of $\mathbb{R}^n$. A function $f : K \to \mathbb{R}$ is quasiconvex if and only if the lower level sets $L(f, \alpha)$ are convex for all $\alpha \in \mathbb{R}$.

Proof. Let f be a quasiconvex function and for $\alpha \in \mathbb{R}$, let $x, y \in L(f, \alpha)$. Then, $f(x) \leq \alpha$ and $f(y) \leq \alpha$. Since f is a quasiconvex function, it follows that for all $\lambda \in [0, 1]$,

$$f((1-\lambda)x + \lambda y) \leq \max\{f(x), f(y)\} \leq \alpha,$$

that is, $(1-\lambda)x + \lambda y \in L(f, \alpha)$ for all $\lambda \in [0, 1]$. Hence, $L(f, \alpha)$ is convex.

Conversely, let $x, y \in K$ and $\bar{\alpha} = \max\{f(x), f(y)\}$. Then, $x, y \in L(f, \bar{\alpha})$, and by convexity of $L(f, \bar{\alpha})$, we have $(1-\lambda)x + \lambda y \in L(f, \bar{\alpha})$ for all $\lambda \in [0, 1]$. Thus for all $\lambda \in [0, 1]$,

$$f((1-\lambda)x + \lambda y) \leq \bar{\alpha} = \max\{f(x), f(y)\}.$$

□

Theorem 1.26. [19, 153] Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a differentiable function. Then, f is quasiconvex if and only if for all $x, y \in K$,

$$f(y) \leq f(x) \quad \Rightarrow \quad \langle \nabla f(x), y - x \rangle \leq 0. \tag{1.28}$$

Proof. Let f be a quasiconvex function. Then, for all $x, y \in K$ and all $\lambda \in [0, 1]$,

$$f(y) \leq f(x) \quad \Rightarrow \quad f((1-\lambda)x + \lambda y) \leq f(x).$$

For $\lambda > 0$, we have

$$\frac{f((1-\lambda)x + \lambda y) - f(x)}{\lambda} \leq 0,$$

which on taking limit $\lambda \to 0^+$ leads to (1.28) as f is a differentiable function.

Conversely, let $x, y \in K$ and $f(y) \leq f(x)$. It is enough to show that for all $\lambda \in\ ]0, 1[$, $f((1-\lambda)x + \lambda y) \leq f(x)$.

Define $F : [0, 1] \to \mathbb{R}$ by $F(\lambda) = f((1-\lambda)x + \lambda y)$. Clearly, $F(1) \leq F(0)$. We need to show $F(\lambda) \leq F(0)$ for all $\lambda \in\ ]0, 1[$. Suppose on the contrary that there exists $\lambda_0 \in\ ]0, 1[$ such that $F(\lambda_0) > F(0)$ and the derivative $F'(\lambda_0) < 0$. Then,

$$F'(\lambda_0) = \langle \nabla f(z_0), y - x \rangle < 0, \tag{1.29}$$

where $z_0 = (1-\lambda_0)x + \lambda_0 y$. Since $F(\lambda_0) > F(0)$, it follows that $f(z_0) > f(x)$ which by (1.28) further implies that $\langle \nabla f(z_0), x - z_0 \rangle \leq 0$. Rearranging the terms, we have $\lambda_0 \langle \nabla f(z_0), x - y \rangle \leq 0$ which implies that $\langle \nabla f(z_0), y - x \rangle \geq 0$, a contradiction to (1.29). □

The following interesting theorem due to Newman [169] and Crouzeix [56] gives sufficient conditions for a real-valued quasiconvex function to be convex.

Theorem 1.27 (Newman-Crouzeix)**.** Let K be a nonempty open convex subset of $\mathbb{R}$ and $f : K \to \mathbb{R}$ be a quasiconvex positively homogeneous function such that one of the following conditions hold:

(i) $f(x) < 0$ for all $x \in K$;

(ii) $f(x) \geq 0$ for all $x \in K$.

Then, f is convex.

We state the following well-known property of the maximization of a quasiconvex function over a polyhedron.

Theorem 1.28. [156] Let K be a compact polyhedron in $\mathbb{R}^n$. If a quasiconvex function $f : K \to \mathbb{R}$ attains a maximum over K, then it attains the maximum at some extreme point of K.

The following theorem provides a condition that ensures the strict quasiconvexity of a quasiconvex function.

Theorem 1.29. A quasiconvex function $f : K \to \mathbb{R}$ on a nonempty convex subset K of $\mathbb{R}^n$ is strictly quasiconvex if it is not constant on the line segment joining any two distinct points of K.

Proof. Let f be quasiconvex such that it is not constant on the line segment joining any two points of K. Suppose, on the contrary, f is not strictly quasiconvex. Then, there exist $x, y \in K$ such that $f(z) \geq \max\{f(x), f(y)\}$, where $z = (1-\lambda)x + \lambda y$ for some $\lambda \in \ ]0, 1[$. Since f is not constant on $[x, z]$ and $[z, y]$, it follows that there exist $u \in [x, z]$ and $v \in [z, y]$ such that $f(u) < f(z)$ and $f(v) < f(z)$. Thus, $f(z) > \max\{f(u), f(v)\}$, which contradicts the quasiconvexity of f as $z \in [u, v]$. $\square$

It is clear that every strictly quasiconvex function is semistrictly quasiconvex, but the converse may not be true.

Example 1.2. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0, \end{cases}$$

is semistrictly quasiconvex, but it is neither strictly quasiconvex nor quasiconvex.

The next theorem gives sufficient conditions for a semistrictly quasiconvex function to be quasiconvex.

Theorem 1.30. Every lower semicontinuous semistrictly quasiconvex function on a convex set is quasiconvex.

Proof. Let f be a semistrictly quasiconvex function defined on a convex subset K of $\mathbb{R}^n$, then for all $x, y \in K$, $f(x) \neq f(y)$ and $\lambda \in \,]0,1[$, we have

$$f((1-\lambda)x + \lambda y) < \max\{f(x), f(y)\}.$$

It remains to show that if $f(x) = f(y)$ and $\lambda \in \,]0,1[$, then

$$f((1-\lambda)x + \lambda y) \leq \max\{f(x), f(y)\}.$$

On the contrary, let $f(z) > f(x)$ for some $z \in \,]x,y[$. Then, $z \in \Omega := \{z \in \,]x,y[\; : f(z) > f(x)\}$. Since f is a lower semicontinuous function, the set Ω is open. Thus, there exists $z_0 \in \,]x,z[$ such that $z_0 \in \Omega$. Since $z, z_0 \in \Omega$, by semistrict quasiconvexity of f, we have

$$f(x) < f(z) \quad \Rightarrow \quad f(z_0) < f(z),$$

and

$$f(y) < f(z_0) \quad \Rightarrow \quad f(z) < f(z_0),$$

which is a contradiction. □

Invoking Theorem 1.25, the following theorem can easily be proved giving the same arguments as in the proof of Theorem 1.16.

Theorem 1.31. For each $i = 1, 2, \ldots, m$, let $f_i : K \to \mathbb{R}$ be a quasiconvex function on a nonempty convex subset K of $\mathbb{R}^n$. Then, $\max\{f_1, f_2, \ldots, f_m\}$ is also a quasiconvex function on K.

We have seen that a local minimum of a convex function over a convex set K is a global minimum over K. The function $f(x) = [x]$ is quasiconvex on $K = [0,3]$ but $x = 3/2$ is a local minimum and not a global minimum on K. However such a result holds for quasiconvex functions if the local minimum is strict.

Theorem 1.32. Every strict local minimum of a quasiconvex function $f : K \to \mathbb{R}$ on a nonempty convex subset K of $\mathbb{R}^n$ is a strict global minimum of f over K. Moreover, the set of points at which f attains its global minimum on K is a convex set.

Proof. Let $x \in K$ be a strict local minimum of f over K. Then, there exists $\delta > 0$ such that

$$f(x) < f(y), \quad \text{for all } y \in K \cap \mathbb{B}_\delta(x). \tag{1.30}$$

Suppose, on the contrary, x is not a strict global minimum over K. Then, there exists $x_0 \in K$, $x_0 \neq x$ such that $f(x_0) \leq f(x)$. By quasiconvexity of f, for all $\lambda \in \,]0,1]$, we have

$$f((1-\lambda)x + \lambda x_0) \leq f(x).$$

This contradicts (1.30) since $(1-\lambda)x + \lambda x_0 \in K \cap \mathbb{B}_\delta(x)$ for λ sufficiently small.

Let f attain its global minimum value, say α, that is, $\alpha = \min\{f(y) : y \in K\}$. The set of points at which f attains its global minimum on K is $L(f, \alpha)$, which by quasiconvexity of f is a convex set. □

A stationary point of a quasiconvex function defined on a nonempty convex subset K of $\mathbb{R}^n$ is not necessarily a global minimum over K. For instance, $x = 0$ is a stationary point of $f(x) = x^3$ defined on $\mathbb{R}$, but it is not a global minimum over $\mathbb{R}$. The sum of two quasiconvex functions defined on a convex set K may not be a quasiconvex function on K. For example, $f_1(x) = e^x - x^2$ and $f_2(x) = -e^x$ are both quasiconvex functions, but the sum $f_1 + f_2$ is not a quasiconvex function.

A strictly quasiconvex function either does not have a minimum or it achieves minimum at the most at one point.

Theorem 1.33. A strictly quasiconvex function $f : K \to \mathbb{R}$ on a nonempty convex subset K of $\mathbb{R}^n$ attains its minimum on K at not more than one point.

Proof. Assume, on the contrary, that f attains global minimum at two distinct points u and $v \in K$. Then,

$$f(u) \leq f(y) \quad \text{and} \quad f(v) \leq f(y), \quad \text{for all } y \in K. \tag{1.31}$$

Taking $y = v$ in the first inequality and $y = u$ in the second inequality, we get $f(u) = f(v)$. By strict quasiconvexity of f, we have for all $\lambda \in \;]0, 1[$, $f((1 - \lambda)u + \lambda v) < f(u)$, which contradicts (1.31). □

Remark 1.15. From Theorem 1.33, it is clear that if f is strictly quasiconvex on a nonempty convex subset K of $\mathbb{R}^n$, then it is not constant on the line segment joining any two distinct points of K.

We have observed that a convex function is semistrictly quasiconvex. Moreover, like a convex function, a semistrictly convex function also enjoys the following property.

Theorem 1.34. Every local minimum of a semistrictly quasiconvex function f on a nonempty convex set K of $\mathbb{R}^n$ is a global minimum of f over K.

Proof. It follows exactly on the lines of the proof of Theorem 1.13. □

Remark 1.16. Unlike quasiconvex functions, the set of points at which a semistrictly quasiconvex function attains its global minimum is not a convex set. This fact is justified by the function considered in Example 1.2.

We have seen that a stationary point of a quasiconvex function f need not be a global minimum of f over its domain. But this is an important property one is looking for when one is required to identify global minimum points of a minimization problem. In this attempt a class of generalized convex functions preserving this property was introduced by Mangasarian [152], namely, pseudoconvex functions. These functions are basically defined for differentiable functions even though the concept exists in literature for nondifferentiable functions as well [27].

Definition 1.35. Let K be a nonempty open subset of $\mathbb{R}^n$. A differentiable function $f : K \to \mathbb{R}$ is said to be

(a) *pseudoconvex* if for all $x, y \in K$,

$$\langle \nabla f(x), y - x \rangle \geq 0 \quad \Rightarrow \quad f(y) \geq f(x);$$

(b) *strictly pseudoconvex* if for all $x, y \in K$, $x \neq y$,

$$\langle \nabla f(x), y - x \rangle \geq 0 \quad \Rightarrow \quad f(y) > f(x).$$

A function f is (strictly) pseudoconcave if $-f$ is (strictly) pseudoconvex.

Comparing with the characterization for convex functions it is clear that if K is an open convex set then every differentiable convex function on K is pseudoconvex on K. The converse implication is, however, not always true. For instance, the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x + x^3$ is pseudoconvex but not convex on $\mathbb{R}$.

Theorem 1.35. Let K be a nonempty open convex subset of $\mathbb{R}^n$. If $f : K \to \mathbb{R}$ is pseudoconvex, then it is semistrictly quasiconvex, and hence quasiconvex.

Proof. Assume, on the contrary, that f is not semistrictly quasiconvex. Then, there exist $x, y \in K$ and $\bar{z} \in \ ]x, y[$ such that

$$f(x) < f(y) \quad \text{and} \quad f(\bar{z}) \geq f(y). \tag{1.32}$$

Also, there exists $z_0 \in \ ]x, y[$ such that

$$f(z_0) \geq f(z), \quad \text{for all } z \in [x, y]. \tag{1.33}$$

Since $z_0 \in \ ]x, y[$ and maximizes f over $[x, y]$, it follows that $\nabla f(z_0) = 0$. Thus, we have $\langle \nabla f(z_0), x - z_0 \rangle = 0$, which by pseudoconvexity of f implies that $f(x) \geq f(z_0)$. This along with (1.32) implies that $f(y) > f(z_0)$ which contradicts (1.33). Thus, f is a semistrictly quasiconvex function. Since a differentiable function is continuous, it follows from Theorem 1.30 that f is a quasiconvex function. $\square$

Remark 1.17. The converse of Theorem 1.35 is not necessarily true. For example, the function $f(x) = x^3$ defined on $\mathbb{R}$ is quasiconvex but not pseudoconvex because for $x = 0$ and $y = -1$, $\langle \nabla f(x), y - x \rangle = 0$ and $f(y) < f(x)$.

In Theorem 1.12, we saw that every stationary point of a convex function is its global minimum. The next theorem shows that this property is also shared by the class of pseudoconvex functions.

Theorem 1.36. Let K be a nonempty open subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a pseudoconvex function. Then, every stationary point of f is a global minimum of f over K.

Proof. Let x be a stationary point of f, that is, $\nabla f(x) = 0$. Since f is a pseudoconvex function, for every $y \in K$, we have

$$\langle \nabla f(x), y - x\rangle \geq 0 \quad \Rightarrow \quad f(y) \geq f(x),$$

which by stationarity condition yields $f(x) \leq f(y)$ for all $y \in K$. Hence, $x \in K$ is a global minimum of f over K. □

Next we present conditions under which a quasiconvex function is a pseudoconvex function.

Theorem 1.37. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a quasiconvex function. Then, f is pseudoconvex if and only if every stationary point is a global minimum of f over K.

Proof. Let f be a pseudoconvex function. Then by Theorem 1.36, every stationary point is a global minimum of f over K. We prove the converse by contradiction. Suppose that every stationary point of f is a global minimum, but f is not pseudoconvex. Then, there exists a pair of points $x, y \in K$ such that $f(y) < f(x)$ and $\langle \nabla f(x), y - x\rangle \geq 0$. It is clear that $\nabla f(x) \neq \mathbf{0}$, otherwise by the hypothesis x must be a global minimum, which is not so. Since f is quasiconvex and $f(y) < f(x)$ implies $\langle \nabla f(x), y - x\rangle \leq 0$. Thus, $\langle \nabla f(x), y - x\rangle = 0$. As f is a differentiable function hence a continuous function, there exists a real number $t > 0$ such that $f(y + t\nabla f(x)) < f(x)$. The quasiconvexity of f implies

$$\frac{f(x + \lambda(y + t\nabla f(x) - x)) - f(x)}{\lambda} \leq 0, \quad \text{for all } \lambda \in \ ]0, 1].$$

Taking limit as λ tends to zero, it follows that $\langle \nabla f(x), y + t\nabla f(x) - x\rangle \leq 0$. Since $\langle \nabla f(x), y - x\rangle = 0$ and $t > 0$, it follows that $\langle \nabla f(x), \nabla f(x)\rangle \leq 0$, which is a contradiction as $\nabla f(x) \neq \mathbf{0}$. □

Corollary 1.5. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a function such that $\nabla f(x) \neq \mathbf{0}$ for every $x \in K$. Then, f is pseudoconvex if and only if it is quasiconvex.

Proof. Follows from Theorems 1.35 and 1.37. □

Remark 1.18. The sum of two pseudoconvex functions defined on an open subset K of $\mathbb{R}^n$ may not be a pseudoconvex function. For example, $f_1(x) = -x$ and $f_2(x) = x + x^3$ are both pseudoconvex functions, but the sum $f_1 + f_2$ is not a pseudoconvex function.

Martos [157] defined a class of functions called pseudomonotonic functions, which includes those functions that are both pseudoconvex as well as pseudoconcave. Many authors extended the study of this class of functions, which were later termed as pseudolinear functions. Here we study pseudolinearity for the differentiable case.

Definition 1.36. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A differentiable function $f : K \to \mathbb{R}$ is said to be *pseudolinear* if it is both pseudoconvex and pseudoconcave.

Some of the examples of pseudolinear function defined on $\mathbb{R}$ are e^x, $x + x^3$, and $\tan^{-1} x$.

We present certain characterizations of pseudolinear functions given by Chew and Choo [44].

Theorem 1.38. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a differentiable function. Then the following statements are equivalent:

(a) f is a pseudolinear function.

(b) For any $x, y \in K$, $\langle \nabla f(x), y - x \rangle = 0$ if and only if $f(x) = f(y)$.

(c) There exists a real-valued function p defined on $K \times K$ such that for any $x, y \in K$,

$$p(x, y) > 0 \quad \text{and} \quad f(y) = f(x) + p(x, y)\langle \nabla f(x), y - x \rangle.$$

Proof. (a) $\Rightarrow$ (b): Suppose that f is pseudolinear. Let $x, y \in K$ be such that $f(x) = f(y)$. By Theorem 1.35, it follows that f is both quasiconvex and quasiconcave, and hence by Theorem 1.26 it follows that $\langle \nabla f(x), y - x \rangle = 0$. The other way implication follows by simply using the definition of pseudolinearity.

(b) $\Rightarrow$ (c): If $\langle \nabla f(x), y - x \rangle = 0$ for some $x, y \in K$, then by the hypothesis $f(x) = f(y)$, and hence we define $p(x, y) = 1$. If $\langle \nabla f(x), y - x \rangle \neq 0$, then by the hypothesis $f(x) \neq f(y)$, and hence, we define

$$p(x, y) = \frac{f(y) - f(x)}{\langle \nabla f(x), y - x \rangle}.$$

We show that $p(x, y) > 0$. Let $f(y) > f(x)$. If $f((1 - \lambda)x + \lambda y) \leq f(x)$ for some $\lambda \in\ ]0, 1[$, then by continuity of f, there exists $\bar{\lambda} \in\ ]0, 1[$ such that $\bar{\lambda} \geq \lambda$ and $f((1 - \bar{\lambda})x + \bar{\lambda} y) = f(x)$. By hypothesis, we have

$$\langle \nabla f(x), y - x \rangle = \frac{1}{\bar{\lambda}} \langle \nabla f(x), (1 - \bar{\lambda})x + \bar{\lambda} y - x \rangle = 0,$$

which is a contradiction as $f(x) \neq f(y)$. Thus, $f((1 - \lambda)x + \lambda y) > f(x)$ for every $\lambda \in\ ]0, 1[$. Hence, $\langle \nabla f(x), y - x \rangle \geq 0$. But as $\langle \nabla f(x), y - x \rangle \neq 0$, it follows that $\langle \nabla f(x), y - x \rangle > 0$. So, we conclude that $p(x, y) > 0$. Similarly, we can show that $p(x, y) > 0$ if $f(y) < f(x)$.

(c) $\Rightarrow$ (a): The proof follows trivially. $\square$

The function p obtained in the above theorem is called the *proportional function* of f.

Theorem 1.39. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a continuously differentiable function. Then, f is pseudolinear if and only if for any $x, y \in K$,

$$\langle \nabla f(x), y-x \rangle = 0 \quad \Rightarrow \quad f(x) = f((1-\lambda)x+\lambda y), \quad \text{for all } \lambda \in [0,1]. \tag{1.34}$$

Proof. Let f be pseudolinear on K. Let $x, y \in K$ be such that $\langle \nabla f(x), y-x \rangle = 0$, then

$$\langle \nabla f(x), (1-\lambda)x + \lambda y - x \rangle = \lambda \langle \nabla f(x), y - x \rangle = 0.$$

Hence, by Theorem 1.38, it follows that $f(x) = f((1-\lambda)x + \lambda y)$ for all $\lambda \in [0,1]$.

Conversely, suppose that the implication (1.34) holds and assume the contrary that f is not pseudolinear on K. Then, by Theorem 1.38, there exist $x, y \in K$ such that either $f(x) - f(y)$ or $\langle \nabla f(x), y - x \rangle$ is zero but not both. Since implication (1.34) holds, it follows that $f(x) = f(y)$ but $\langle \nabla f(x), y-x \rangle \neq 0$. Without loss of generality, we assume that $\langle \nabla f(x), y - x \rangle > 0$. Define $g : [0,1] \to \mathbb{R}$ as

$$g(\lambda) = f((1-\lambda)x + \lambda y).$$

Then g is a continuously differentiable function such that $g(0) = g(1)$ and $g'(0) > 0$. Thus, g assumes a local maximum at some point $\bar{\lambda} \in\ ]0,1[$. Therefore, for $z = (1-\bar{\lambda})x + \bar{\lambda} y$,

$$0 = g'(\bar{\lambda}) = \langle \nabla f(z), y - x \rangle = \frac{1}{\bar{\lambda}} \langle \nabla f(z), z - x \rangle.$$

Thus, by implication (1.34), we have $f(z) = f((1-\lambda)z + \lambda x)$ for all $\lambda \in [0,1]$, which further implies that $\langle \nabla f(x), z - x \rangle = 0$, that is, $\langle \nabla f(x), y - x \rangle = 0$, which is a contradiction. $\square$

Chew and Choo [44] pointed out that a real-valued differentiable pseudolinear function defined on an interval must be either a constant function or a strictly monotonic function whose derivative does not vanish at any point of the interval.

An interesting class of pseudolinear functions is that of linear fractional functions of the form

$$f(x) = \frac{\langle a, x \rangle + c}{\langle b, x \rangle + d},$$

where $a, b, x \in \mathbb{R}^n$, $c, d \in \mathbb{R}$ and $\langle b, x \rangle + d$ is nonzero and keeps the same sign for every $x \in K$. If for $x, y \in K$, we set

$$p(x,y) = \frac{\langle b, x \rangle + d}{\langle b, y \rangle + d},$$

then,

$$\nabla f(x) = \frac{1}{(\langle b, x\rangle + d)^2}\left[(\langle b, x\rangle + d)a - (\langle a, x\rangle + c)b\right].$$

It can be seen that $f(y) = f(x) + p(x, y)\langle \nabla f(x), y - x\rangle$, and hence, f is pseudolinear.

Several books and monographs on convex analysis provide further details on convex functions, generalized convex functions, and their properties [19, 28, 30, 86, 153, 180].

1.7 Optimality Criteria

In this section, we study necessary and sufficient optimality conditions for unconstrained and constrained minimization problems. Necessary optimality criteria help to identify points that are not optimal. We call the minimization problem (MP), introduced in Section 1.5, as the unconstrained minimization problem (UMP) when $K = \mathbb{R}^n$ and is defined as follows:

$$\text{(UMP)} \qquad \min f(x)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable function.

We first give the following necessary optimality criteria for the unconstrained minimization problem (UMP).

Theorem 1.40. If x is a (local) minimum of (UMP) then $\nabla f(x) = \mathbf{0}$.

Proof. If x is a local minimum of (UMP), then there exists $\varepsilon > 0$ such that

$$f(x) \le f(y), \quad \text{for all } y \in \mathbb{B}_\varepsilon(x).$$

Consider $d \in \mathbb{R}^n$ with $\|d\| = 1$. Then, $x + \lambda d \in \mathbb{B}_\varepsilon(x)$, for $0 < \lambda < \varepsilon$. Hence,

$$f(x) \le f(x + \lambda d), \quad \text{for all } \lambda,\ 0 < \lambda < \varepsilon.$$

Using the differentiability of f, we have

$$\lambda\langle \nabla f(x), d\rangle + o(\lambda) \ge 0.$$

Dividing by $\lambda > 0$ and taking limit as $\lambda \to 0^+$, we get

$$\langle \nabla f(x), d\rangle \ge 0, \quad \text{for all } d \in \mathbb{R}^n,\ \|d\| = 1.$$

Consider any $\mathbf{0} \ne u \in \mathbb{R}^n$ and set $d = u/\|u\|$. We then have

$$\langle \nabla f(x), u\rangle \ge 0, \quad \text{for all } \mathbf{0} \ne u \in \mathbb{R}^n,$$

which implies that $\nabla f(x) = \mathbf{0}$. □

We recall (Theorem 1.12 and Theorem 1.36) that one of the most important properties possessed by convex as well pseudoconvex functions is that every stationary point is a global minimum. Thus, it follows that the converse of Theorem 1.40 is true for these functions.

Theorem 1.41. If $\nabla f(x) = \mathbf{0}$ and f is either a convex or a pseudoconvex function on $\mathbb{R}^n$, then x is a minimum of (UMP).

We present optimality criterion for a constrained optimization problem when the constraint set is given explicitly by inequality constraints.

Consider the optimization problem

$$\text{(CMP)} \qquad \min f(x), \quad \text{subject to } g_i(x) \leq 0,\ i = 1, 2, \ldots m,$$

where f and g_i, $i = 1, 2, \ldots, m$, are real-valued differentiable functions defined on $\mathbb{R}^n$.

We mention the Fritz John necessary optimality criterion for the problem (CMP) in terms of the functions involved. Before that we state the Gordon-type theorem of alternative, which would be required in the proof of the Fritz John theorem. For proof one may refer to Rockafellar [182].

Theorem 1.42. Let $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$, be convex functions and let K be a nonempty convex subset of $\mathbb{R}^n$. Then one and only one of the following systems can have a solution but not both:

(i) $f_i(x) < 0$, $x \in K$;

(ii) there exists $\mathbf{0} \neq \lambda \in \mathbb{R}^m_+$ such that $\sum_{i=1}^m \lambda_i f_i(x) \geq 0$, for all $x \in K$.

We present the Fritz John optimality criterion for (CMP).

Theorem 1.43. If x is a (local) minimum of (CMP), then there exist $\lambda_0 \in \mathbb{R}$, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m$ such that

$$\lambda_0 \nabla f(x) + \sum_{i=1}^m \lambda_i \nabla g_i(x) = 0, \tag{1.35}$$

$$\lambda_i g_i(x) = 0,\ i = 1, 2, \ldots, m, \tag{1.36}$$

$$\lambda_0 \geq 0,\ \lambda_i \geq 0,\ i = 1, 2, \ldots, m,\ (\lambda_0, \lambda) \neq \mathbf{0}. \tag{1.37}$$

Proof. Consider the following system

(i) $\langle \nabla f(x), d \rangle < 0$,

(ii) $g_i(x) + \langle \nabla g_i(x), d \rangle < 0$, for all $i = 1, 2, \ldots, m$.

We will first prove that the above system has no solution $d \in \mathbb{R}^n$. On the contrary suppose that $d \in \mathbb{R}^n$ solves the above system. By differentiability of g_i, we have for $i = 1, 2, \ldots, m$ and $0 < \lambda < 1$,

$$g_i(x + \lambda d) = g_i(x) + \lambda \langle \nabla g_i(x), d \rangle + o(\lambda),$$

where $\lim_{\lambda \to 0} \frac{o(\lambda)}{\lambda} = 0$. The above equation yields that

$$g_i(x + \lambda d) = (1 - \lambda) g_i(x) + \lambda \left(g_i(x) + \langle \nabla g_i(x), d \rangle \right) + o(\lambda).$$

As x is a solution to (CMP) therefore, we have $g_i(x) \leq 0$ for all $i = 1, 2, \ldots, m$, and the above equation becomes

$$g_i(x + \lambda d) \leq \lambda \left(g_i(x) + \langle \nabla g_i(x), d \rangle \right) + o(\lambda).$$

Now from $g_i(x) + \langle \nabla g_i(x), d \rangle < 0$ for all $i = 1, 2, \ldots, m$, we see that for λ sufficiently small,

$$(g_i(x) + \langle \nabla g_i(x), d \rangle) + \frac{o(\lambda)}{\lambda} < 0.$$

This implies that $\frac{g_i(x + \lambda d)}{\lambda} < 0$. Thus, for $0 < \lambda < 1$ sufficiently small, we have $x + \lambda d$ is a feasible solution of (CMP). Also, $\langle \nabla f(x), d \rangle < 0$ implies that $f(x + \lambda d) < f(x)$ for $0 < \lambda < 1$ sufficiently small. This is a contradiction to the fact that x is a (local) minimum of (CMP). Thus, the system of inequalities considered above has no solution $d \in \mathbb{R}^n$. Invoking Theorem 1.42, there exist scalars $\lambda_0 \geq 0$, $\lambda_i \geq 0$, $i = 1, 2, \ldots m$, not all zero such that for every $d \in \mathbb{R}^n$,

$$\lambda_0 \langle \nabla f(x), d \rangle + \sum_{i=1}^{m} \lambda_i \langle \nabla g_i(x), d \rangle + \sum_{i=1}^{m} \lambda_i g_i(x) \geq 0. \tag{1.38}$$

Taking $d = \mathbf{0}$ in the above expression, we get $\sum_{i=1}^{m} \lambda_i g_i(x) \geq 0$. Also, the inequalities $g_i(x) \leq 0$ and $\lambda_i \geq 0$, imply that $\sum_{i=1}^{m} \lambda_i g_i(x) \leq 0$. Hence, we have $\sum_{i=1}^{m} \lambda_i g_i(x) = 0$. Therefore, $\lambda_i g_i(x) = 0$, for all $i = 1, 2, \ldots, m$. Thus, from inequality (1.38), we have for $d \in \mathbb{R}^n$

$$\lambda_0 \langle \nabla f(x), d \rangle + \sum_{i=1}^{m} \lambda_i \langle \nabla g_i(x), d \rangle \geq 0,$$

which implies that (1.35) holds. □

Conditions (1.35) to (1.37) are termed *Fritz John conditions* and condition (1.36) is termed as *complementary slackness condition.*

Remark 1.19. Let $I = \{i : g_i(x) = 0\}$ denote the collection of indices for which g_i are active, then condition (1.36) ensures that $\lambda_i = 0$ for $i \notin I$. Then, the Fritz John conditions can be modified as

$$\lambda_0 \nabla f(x) + \textstyle\sum_{i \in I} \lambda_i \nabla g_i(x) = 0,$$
$$\lambda_0 \geq 0, \ \lambda_i \geq 0, \ i \in I, \ (\lambda_0, \lambda_i) \neq \mathbf{0}.$$

Remark 1.20. In condition (1.35) one can either have $\lambda_0 = 0$ or $\lambda_0 \neq 0$. In the situation where $\lambda_0 = 0$, no information can be inferred from the objective function as condition (1.35) reduces to $\sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0$. Such cases are referred to as degenerate or abnormal. The necessary conditions in such cases do not take into account the properties of the function, but only the geometry of constraints. In other words, Fritz John conditions are relevant only when $\lambda_0 > 0$. If $\lambda_0 \neq 0$, it can be observed from (1.35) that

$$-\nabla f(x) = \sum_{i=1}^{m} \lambda_i^0 \nabla g_i(x),$$

where $\lambda_i^0 = \dfrac{\lambda_i}{\lambda_0}$.

We discuss the conditions that ensure $\lambda_0 > 0$ in Fritz John conditions. Restriction imposed on the constraints to ensure that λ_0 is nonzero corresponding to any objective function is generally termed as constraint qualification. If λ_0 is nonzero, one can choose $\lambda_0 = 1$ and the Fritz John conditions (1.35) to (1.37) can be reformulated as

$$\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0, \tag{1.39}$$

$$\lambda_i g_i(x) = 0, \quad i = 1, 2, \ldots, m, \tag{1.40}$$

$$\lambda_i \geq 0, \quad i = 1, 2, \ldots, m. \tag{1.41}$$

Conditions (1.39) to (1.41) are referred to as *Kuhn-Tucker conditions*. These conditions are also referred as *Karush-Kuhn-Tucker conditions* to give credit to Karush [127] who had obtained these conditions in his thesis in 1939, much before it was given by Kuhn and Tucker [140] in 1951. The scalars satisfying Kuhn-Tucker conditions are referred to as *Lagrange multipliers*.

Remark 1.21. We give the geometrical interpretation of the Kuhn-Tucker conditions.

From condition (1.39), we have

$$-\nabla f(x) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x),$$

and since condition (1.40) ensures that $\lambda_i = 0$ for $i \notin I$, it follows that

$$-\nabla f(x) = \sum_{i \in I} \lambda_i \nabla g_i(x) \in \left\{ \sum_{i \in I} \lambda_i \nabla g_i(x) : \lambda_i \geq 0,\ i \in I \right\}.$$

This implies that $-\nabla f(x)$ belongs to the polyhedral cone generated by the gradients of active constraints at x.

A number of constraint qualifications exist which ensure the existence of Lagrange multipliers. One of the constraint qualifications is due to Slater [193].

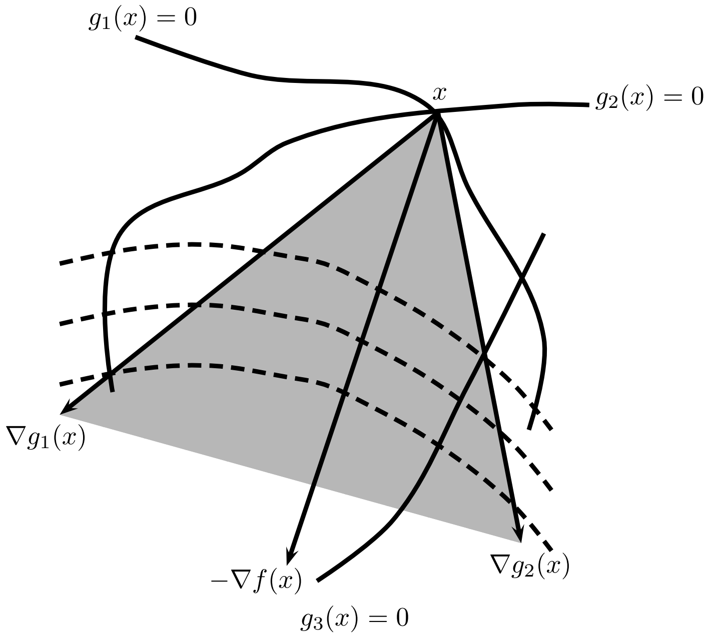


FIGURE 1.14: $-\nabla f(x)$ lies in the cone generated by $\nabla g_1(x)$ and $\nabla g_2(x)$, where g_1 and g_2 are active at the minimum point x

Definition 1.37. The *Slater's constraint qualification* holds if each g_i is convex (or pseudoconvex) and there exists a feasible point $\hat{x}$ such that $g_i(\hat{x}) < 0$, $i = 1, 2, \ldots, m$.

The following result is the Kuhn-Tucker optimality criterion.

Theorem 1.44. If x is a (local) minimum of (CMP) and if each g_i is convex (or pseudoconvex) and Slater's constraint qualification holds, then there exists $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m$ satisfying the conditions (1.39) - (1.41).

Proof. By Theorem 1.43, conditions (1.35) to (1.37) hold. We only need to show that $\lambda_0 \neq 0$. Suppose on the contrary that $\lambda_0 = 0$, then from (1.35) we have $\sum_{i=1}^{m} \lambda_i \nabla g_i(x) = \mathbf{0}$, which further implies that

$$\sum_{i \in I} \lambda_i \nabla g_i(x) = \mathbf{0}, \tag{1.42}$$

where $\lambda_i \geq 0$, $i \in I$ are all not all zero. Since Slater's constraint qualification holds, there exists a feasible point $\hat{x}$ such that $g_i(\hat{x}) < 0$, $i = 1, 2, \ldots, m$. As for $i \in I$, $g_i(\hat{x}) < g_i(x)$, we have by convexity (or pseudoconvexity) of g_i,

$\langle \nabla g_i(x), \hat{x} - x \rangle < 0$. Also since $\lambda_i \geq 0$, $i \in I$ are all not all zero, it follows on multiplying by λ_i and summing over $i \in I$, we have

$$\left\langle \sum_{i \in I} \lambda_i \nabla g_i(x), \hat{x} - x \right\rangle < 0,$$

which contradicts (1.42). □

We have the following sufficient optimality criteria.

Theorem 1.45. If x is a feasible solution of (CMP) and there exists $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m$ satisfying the conditions (1.39) to (1.41), then x is a minimum of (CMP) if f and g_i, $i = 1, 2, \ldots, m$ are convex functions on $\mathbb{R}^n$.

Proof. Suppose on the contrary that x is not a minimum of (CMP), then there exists a feasible point $\hat{x}$ such that $f(\hat{x}) < f(x)$. By convexity of f, it follows that $\langle \nabla f(x), \hat{x} - x \rangle < 0$, which by condition (1.39) yields

$$\left\langle \sum_{i=1}^{m} \lambda_i \nabla g_i(x), \hat{x} - x \right\rangle > 0.$$

Since by (1.40) $\lambda_i = 0$, for $i \notin I$, we have

$$\left\langle \sum_{i \in I} \lambda_i \nabla g_i(x), \hat{x} - x \right\rangle > 0. \tag{1.43}$$

As $g_i(\hat{x}) \leq g_i(x)$ for $i \in I$, by convexity of g_i, it is obvious that $\langle \nabla g_i(x), \hat{x} - x \rangle \leq 0$, which on multiplying by λ_i and summing over $i \in I$ yields a contradiction to (1.43). □

Finally, we present Theorem 1.45 when f is pseudoconvex and each g_i is quasiconvex. Proof follows on the lines of Theorem 1.45 and is hence omitted.

Theorem 1.46. If x is a feasible solution of (CMP) and there exists $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m$ satisfying the conditions (1.39) to (1.41), then x is a minimum of (CMP) if f is pseudoconvex and each g_i, $i = 1, 2, \ldots, m$ is quasiconvex on $\mathbb{R}^n$.

1.8 Subgradients and Subdifferentials

For differentiable functions, characterizations of convexity are known in terms of its gradient. However, while dealing with nondifferentiable convex functions the notion of subgradient is used instead of the gradient.

The study of subgradients plays a crucial role in convex optimization as the computation of the subgradients can help minimize (almost) any convex function. Also, nondifferentiable convex functions can be characterized via the monotonicity of the subdifferential map.

Using the convention for addition given in Section 1.2, we extend the definition of convexity to extended real-valued functions. A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be *convex* if for all $x, y \in \mathbb{R}^n$ and all $\lambda \in [0, 1]$,

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Definition 1.38. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a convex function and x be a point where f is finite. Then a vector $\xi \in \mathbb{R}^n$ is called a *subgradient* of f at $x \in K$ if

$$\langle \xi, y - x\rangle \leq f(y) - f(x), \quad \text{for all } y \in \mathbb{R}^n. \tag{1.44}$$

Similarly, if f is concave, then a vector $\xi \in \mathbb{R}^n$ is called a *supergradient* of f at a point $x \in K$, where f is finite, if

$$\langle \xi, y - x\rangle \geq f(y) - f(x), \quad \text{for all } y \in \mathbb{R}^n.$$

The set of all the subgradients of a convex function f at a point x, where f is finite, is called *subdifferential* of f at x, and it is denoted by $\partial f(x)$, that is,

$$\partial f(x) = \{\xi \in \mathbb{R}^n : \langle \xi, y - x\rangle \leq f(y) - f(x), \text{ for all } y \in \mathbb{R}^n\}.$$

A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be *subdifferentiable* at a point x, where f is finite, if $\partial f(x) \neq \emptyset$.

The subdifferential $\partial f(x)$ of a function f at a point x is an image of a set-valued map $\partial f : \mathbb{R} \to 2^{\mathbb{R}^n}$. This set-valued map is called the *subdifferential mapping*. For details on set-valued maps see Appendix A.

Definition 1.39. Let K be a nonempty closed convex subset of $\mathbb{R}^n$. A vector $v \in \mathbb{R}^n$ is said to be *normal* to K at $x \in K$ if $\langle v, y - x\rangle \leq 0$ for all $y \in K$. The set of all such vectors is called the *normal cone* to K at x, and it is denoted by $N(K, x)$.

Remark 1.22. (a) In view of Theorem 1.9, if $f : K \to \mathbb{R}$ is a convex function defined on an open convex subset K of $\mathbb{R}^n$, then for every $x \in K$, there exists a vector $\xi \in \mathbb{R}^n$ such that

$$\langle \xi, y - x\rangle \leq f(y) - f(x), \quad \text{for all } y \in K. \tag{1.45}$$

We observe that the convex function f can be extended to a convex function defined on $\mathbb{R}^n$, by assigning value $+\infty$ for points outside K. We define an extended real-valued function $\hat{f} : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ as

$$\hat{f}(x) = \begin{cases} f(x), & \text{if } x \in K, \\ +\infty, & \text{if } x \notin K. \end{cases}$$

Then, from inequality (1.45), we have

$$\langle \xi, y - x \rangle \leq \hat{f}(y) - \hat{f}(x), \quad \text{for all } y \in \mathbb{R}^n.$$

This implies that $\partial \hat{f}(x) \neq \emptyset$ for every $x \in K$, that is, $\partial f(x) \neq \emptyset$ for every $x \in K$.

(b) It can be easily seen that the set $\partial f(x)$ is nonempty, closed, and convex if $x \in \text{int}(K)$, see Bertsekas et al. [32, Proposition 4.2.1].

(c) The subdifferential $\partial f(x)$ is a singleton set if and only if f is differentiable at x and in this case $\partial f(x) = \{\nabla f(x)\}$.

(d) We see that ξ is a subgradient of f at a point x if and only if $(\xi, -1) \in \mathbb{R}^{n+1}$ is a normal vector to a supporting hyperplane for epi(f) at $(x, f(x))$. Indeed,

$$\begin{aligned}
& \xi \text{ is a subgradient of } f \\
\Leftrightarrow \quad & \text{inequality (1.44) holds} \\
\Leftrightarrow \quad & \langle (\xi, -1), (y - x, \alpha - f(x)) \rangle \leq 0, \quad \text{for all } \alpha \geq f(y) \\
\Leftrightarrow \quad & (\xi, -1) \in N\left(\text{epi}(f), (x, f(x))\right).
\end{aligned}$$

Remark 1.23. (a) Let $]a, b[$ be an open interval and $f : \ ]a, b[\ \to \mathbb{R}$ be a convex function. Then the left derivative f'_- and the right derivative f'_+ both exist and are nondecreasing. Also, f is subdifferentiable at every point $x \in \ ]a, b[$ and $\partial f(x) = \left[f'_-(x), f'_+(x)\right]$.

(b) The Euclidean norm $f(x) = \|x\|$ for all $x \in \mathbb{R}^n$ is not differentiable at $x = \mathbf{0}$, but it is subdifferentiable and

$$\partial f(\mathbf{0}) = \{\xi \in \mathbb{R}^n : \|y\| \geq \langle \xi, y \rangle \text{ for all } y\},$$

that is, the closed unit ball.

The following theorems follow directly from the definition of the subdifferential.

Theorem 1.47. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a convex function and $x \in \mathbb{R}^n$ be a point where f is finite. Then, x minimizes f over K if and only if $\mathbf{0} \in \partial f(x)$.

Theorem 1.48. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a convex function, and $\alpha > 0$. Then for any $x \in K$, $\partial(\alpha f)(x) = \alpha \partial f(x)$.

Section 2.6 of the next chapter deals with Dini, Clarke, and other kinds of subdifferentials, which are more general than the subdifferential discussed in this section. The following theorems are particular cases of the results given in Section 2.6, therefore, we omit the proof. However, a direct proof of the following theorems can be found in the literature [32, 105, 106, 190].

Theorem 1.49. Let K be a nonempty open convex subset of $\mathbb{R}^n$, $f_i : K \to \mathbb{R}$, $i = 1, 2, \ldots m$ be convex functions. Then for every $x \in K$,

$$\partial f_1(x) + \cdots + \partial f_m(x) = \partial(f_1 + \cdots + f_m)(x).$$

We mention the mean value theorem for subdifferentiable functions whose proof can be found in Hiriart Urruty and Lemaréchal [105, Theorem 2.3.3].

Theorem 1.50 (Mean Value Theorem)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. For any two distinct points $x, y \in \mathbb{R}^n$, there exists $\lambda \in\ ]0,1[$ and $\xi \in \partial f(x_\lambda)$, where $x_\lambda = (1-\lambda)x + \lambda y$, such that

$$f(y) - f(x) = \langle \xi, y - x \rangle, \tag{1.46}$$

that is,

$$f(y) - f(x) \in \bigcup_{\lambda \in\ ]0,1[} \{\langle \partial f(x_\lambda), y - x \rangle\}.$$

The mean value theorem can also be given in the following integral form.

Theorem 1.51. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. For any $x, y \in \mathbb{R}^n$,

$$f(y) - f(x) = \int_0^1 \langle \partial f(x_\lambda), y - x \rangle d\lambda. \tag{1.47}$$

Remark 1.24. The relation (1.47) means that if $\{\xi_\lambda : \lambda \in [0,1]\}$ is any selection of subgradients of f on the line segment $[x, y]$, that is, $\xi_\lambda \in \partial f(x_\lambda)$ for all $\lambda \in [0,1]$, then

$$\int_0^1 \langle \xi_\lambda, y - x \rangle d\lambda$$

is independent of the selection and its value is $f(y) - f(x)$.

The following result provides the necessary and sufficient condition for a convex function to be strongly convex.

Theorem 1.52. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a convex function. Then, f is strongly convex with modulus $\rho > 0$ if and only if for all $x, y \in K$,

$$\langle \xi, y - x \rangle + \rho \|y - x\|^2 \le f(y) - f(x), \quad \text{for all } \xi \in \partial f(x), \tag{1.48}$$

or equivalently,

$$\langle \zeta - \xi, y - x \rangle \ge \sigma \|y - x\|^2, \quad \text{for all } \xi \in \partial f(x) \text{ and } \zeta \in \partial f(y), \tag{1.49}$$

where $\sigma = 2\rho$.

Proof. For any $x, y \in K$ and $\lambda \in \;]0,1[$, let $x_\lambda = \lambda y + (1-\lambda)x = x + \lambda(y-x)$. We first prove that the inequalities (1.48) and (1.49) are equivalent.

From inequality (1.48), we obtain

$$\langle \xi, y-x\rangle + \rho\|y-x\|^2 \leq f(y) - f(x), \quad \text{for all } \xi \in \partial f(x),$$

and

$$\langle \zeta, x-y\rangle + \rho\|y-x\|^2 \leq f(x) - f(y), \quad \text{for all } \zeta \in \partial f(y).$$

Adding the above two inequalities, we obtain the inequality (1.49).

Conversely, let $\varphi : \mathbb{R} \to \mathbb{R}$ be a function defined by $\varphi(\lambda) = f(x_\lambda)$. Then by applying Theorem 1.51 to φ, we have

$$f(y) - f(x) = \varphi(1) - \varphi(0) = \int_0^1 \langle \xi_\lambda, y-x\rangle d\lambda, \tag{1.50}$$

where $\xi_\lambda \in \partial f(x_\lambda)$ for all $\lambda \in [0,1]$. Take $\xi \in \partial f(x)$ arbitrary and apply (1.49), we obtain

$$\langle \xi_\lambda - \xi, x_\lambda - x\rangle \geq \sigma\|x_\lambda - x\|^2.$$

By using the value of x_λ, we have

$$\lambda\langle \xi_\lambda, y-x\rangle \geq \lambda\langle \xi, y-x\rangle + \sigma\lambda^2\|y-x\|^2.$$

For $\lambda \in \;]0,1]$, we have

$$\langle \xi_\lambda, y-x\rangle \geq \langle \xi, y-x\rangle + \sigma\lambda\|y-x\|^2.$$

Integrating both the sides with respect to λ between the limits 0 and 1 and using (1.50), we obtain relation (1.48).

We next show that the inequality (1.48) implies the strong convexity of f. By substituting $x = x_\lambda$ in inequality (1.48), we obtain

$$\langle \xi, y-x_\lambda\rangle + \rho\|y-x_\lambda\|^2 \leq f(y) - f(x_\lambda), \quad \text{for all } \xi \in \partial f(x_\lambda),$$

that is,

$$(1-\lambda)\langle \xi, y-x\rangle + \rho(1-\lambda)^2\|y-x\|^2 \leq f(y) - f(x_\lambda), \quad \text{for all } \xi \in \partial f(x_\lambda). \tag{1.51}$$

Similarly, we have

$$\lambda\langle \xi, x-y\rangle + \rho\lambda^2\|x-y\|^2 \leq f(x) - f(x_\lambda), \quad \text{for all } \xi \in \partial f(x_\lambda). \tag{1.52}$$

Multiplying inequality (1.51) by λ and inequality (1.52) by $1-\lambda$, and then adding the resultants, we obtain

$$(1-\lambda)f(x) + \lambda f(y) \geq f(x_\lambda) + \rho\|y-x\|^2\left(\lambda(1-\lambda)^2 + \lambda^2(1-\lambda)\right),$$

that is,

$$f((1-\lambda)x + \lambda y) + \rho\lambda(1-\lambda)\|y-x\|^2 \leq (1-\lambda)f(x) + \lambda f(y).$$

Hence, f is strongly convex with modulus ρ.

Finally, we prove that strong convexity of f implies inequality (1.48). The inequality (1.23) can be re-written as

$$\frac{f(x_\lambda) - f(x)}{\lambda} + \rho(1-\lambda)\|y-x\|^2 \leq f(y) - f(x).$$

As

$$\frac{f(x_\lambda) - f(x)}{\lambda} \geq \langle \xi, y-x \rangle \quad \text{for all} \quad \xi \in \partial f(x)$$

we obtain equality (1.48) on taking limit $\lambda \to 0^+$. □

Remark 1.25. A set-valued mapping ∂f that satisfies the inequality (1.49) is called *strongly monotone* with constant $\sigma > 0$.

Analogously to the previous theorem, we can establish the following theorem, which gives the necessary and sufficient condition for a convex function to be strictly convex in terms of the subdifferential.

Theorem 1.53. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a convex function. Then f is strictly convex if and only if for all $x, y \in K$, $x \neq y$,

$$\langle \xi, y-x \rangle < f(y) - f(x), \quad \text{for all } \xi \in \partial f(x), \tag{1.53}$$

or equivalently,

$$\langle \zeta - \xi, y-x \rangle > 0, \quad \text{for all } \xi \in \partial f(x) \text{ and } \zeta \in \partial f(y). \tag{1.54}$$

Remark 1.26. A set-valued mapping ∂f that satisfies the inequality (1.54) is called *strictly monotone*.

The next results present some properties of the subdifferential set-valued mapping ∂f of a convex function $f : \mathbb{R}^n \to \mathbb{R}$.

Theorem 1.54. The subdifferential mapping $\partial f : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ of a convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is monotone in the sense that for all $x, y \in \mathbb{R}^n$,

$$\langle \xi - \zeta, x-y \rangle \geq 0, \quad \text{for all } \xi \in \partial f(x) \text{ and } \zeta \in \partial f(y). \tag{1.55}$$

Proof. From inequality (1.44), we have

$$\langle \xi, y-x \rangle \leq f(y) - f(x), \quad \text{for all } \xi \in \partial f(x),$$

and

$$\langle \zeta, x-y \rangle \leq f(x) - f(y), \quad \text{for all } \zeta \in \partial f(y).$$

Adding these inequalities, we obtain the result. □

Theorem 1.55. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then the graph $\text{Graph}(\partial f) = \{(x,\xi) \in \mathbb{R}^n \times \mathbb{R}^n : \xi \in \partial f(x)\}$ of the subdifferential set-valued mapping ∂f is closed in $\mathbb{R}^n \times \mathbb{R}^n$.

Proof. Let $\{(x_m, \xi_m)\}$ be a sequence in Graph(∂f) converging to $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$. Then for all $y \in \mathbb{R}^n$,

$$\langle \xi_m, y - x_m \rangle \leq f(y) - f(x_m), \quad \text{for all } m.$$

Since every convex function is continuous in the interior of its domain and the scalar product is continuous, taking limit as $m \to \infty$, we obtain

$$\langle \xi, y - x \rangle \leq f(y) - f(x),$$

and hence, $(x, \xi) \in$ Graph(∂f). Thus, Graph(∂f) is closed. □

Theorem 1.56. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. The set-valued mapping ∂f is locally bounded in the sense that the image $\partial f(B)$ of a bounded subset B of $\mathbb{R}^n$ is a bounded set in $\mathbb{R}^n$.

Proof. Let $x \in B$ and $\mathbf{0} \neq \xi \in \partial f(x)$ be arbitrary. Taking $y = x + \frac{\xi}{\|\xi\|}$ in inequality (1.44), we get

$$\left\langle \xi, \frac{\xi}{\|\xi\|} \right\rangle \leq f\left(x + \frac{\xi}{\|\xi\|}\right) - f(x),$$

equivalently,

$$\|\xi\| \leq f\left(x + \frac{\xi}{\|\xi\|}\right) - f(x).$$

By Theorem 1.14, f is Lipschitz continuous on the bounded set $\mathbb{B}_1(B) = \{\mathbb{B}_1(x) : x \in B\}$, and hence

$$\begin{aligned} \|\xi\| &\leq \left| f\left(x + \frac{\xi}{\|\xi\|}\right) - f(x) \right| \\ &\leq k, \end{aligned}$$

where k is the Lipschitz constant of f. Thus, $\partial f(B)$ is bounded. □

We already know that $\partial f(x)$ is compact and convex, however, we derive this result from the previous two theorems.

Corollary 1.6. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then $\partial f(B)$ is a compact subset of $\mathbb{R}^n$ if B is a compact in $\mathbb{R}^n$.

Proof. Let $\{x_m\}$ be a sequence in a compact set B with a subsequence $\{x_{m_k}\}$ converging to x, say. Take $\xi_m \in \partial f(x_m)$ and extract the subsequence $\{\xi_{m_k}\}$. From Theorem 1.56, a subsequence of $\{\xi_{m_k}\}$ converges to some $\xi \in \mathbb{R}^n$, and from Theorem 1.55, $\xi \in \partial f(x)$. Hence, $\partial f(B)$ is compact. □

Since the subdifferential ∂f of a convex function $f : \mathbb{R}^n \to \mathbb{R}$ is a compact and convex valued mapping, we have the following result.

Theorem 1.57. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then, the subdifferential mapping $\partial f : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ is upper semicontinuous at every point $x \in \mathbb{R}^n$ in the sense that for every $\varepsilon > 0$, there exists $\delta > 0$ such that $\partial f(y) \subseteq \mathbb{B}_\varepsilon(\partial f(x))$ for all $y \in \mathbb{B}_\delta(x)$, where $\mathbb{B}_\varepsilon(\partial f(x)) = \partial f(x) + \mathbb{B}_\varepsilon(\mathbf{0})$.

Proof. Assume to the contrary that for some x, there exist $\varepsilon > 0$ and a sequence $\{(x_m, \xi_m)\}$ such that

$$\begin{gathered} x_m \to x \quad \text{for } m \to \infty \quad \text{and} \\ \xi_m \in \partial f(x_m),\ \xi_m \notin \mathbb{B}_\varepsilon(\partial f(x)), \quad \text{for } m = 1, 2, \ldots. \end{gathered} \tag{1.56}$$

By Theorem 1.56, $\{\xi_m\}$ is a bounded sequence, and hence, a subsequence of the sequence $\{\xi_m\}$ converges to ξ, say. By Theorem 1.55, $\xi \in \partial f(x)$. This contradicts (1.56). $\square$

Chapter 2

Generalized Derivatives and Generalized Subdifferentials

2.1 Introduction

We often come across different kinds of real-world problems that can be written in the form of an optimization problem where the objective function is not necessarily differentiable in the classical sense. Therefore, several kinds of derivatives have been introduced and studied in the literature. In this chapter, we study some of these derivatives, namely, directional derivatives, Gâteaux derivatives, Dini derivatives, Dini-Hadamard derivatives, and Clarke derivatives. We present some fundamental properties of these derivatives. The mean value theorem, which is one of the most important results from calculus, has also been presented in terms of different kinds of derivatives.

2.2 Directional Derivatives

In this section, we discuss directional derivatives of an extended real-valued function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$. Some basic properties of these derivatives are also presented.

Definition 2.1. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a function and $x \in \mathbb{R}^n$ be a point where f is finite.

(a) The *right-sided directional derivative* of f at x in the direction $d \in \mathbb{R}^n$ is defined by

$$f'_+(x; d) = \lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t},$$

if the limit exists, finite or not.

(b) The *left-sided directional derivative* of f at x in the direction $d \in \mathbb{R}^n$ is defined by

$$f'_-(x; d) = \lim_{t \to 0^-} \frac{f(x + td) - f(x)}{t},$$

if the limit exists, finite or not.

For $d = \mathbf{0}$ the zero vector in $\mathbb{R}^n$, both $f'_+(x;\mathbf{0})$ and $f'_-(x;\mathbf{0})$ are defined to be zero.

Since

$$f'_+(x;-d) = \lim_{t\to 0^+} \frac{f(x-td)-f(x)}{t} = \lim_{\tau\to 0^-} \frac{f(x+\tau d)-f(x)}{-\tau} = -f'_-(x;d),$$

we have

$$-f'_+(x;-d) = f'_-(x;d).$$

If $f'_+(x;d)$ exists and $f'_+(x;d) = f'_-(x;d)$, then it is called the *directional derivative* of f at x in the direction d. Thus, the *directional derivative* of f at x in the direction $d \in \mathbb{R}^n$ is defined by

$$f'(x;d) = \lim_{t\to 0} \frac{f(x+td)-f(x)}{t},$$

provided the limit exists, finite or not.

Remark 2.1. (a) If $f'(x;d)$ exists, then $f'(x;-d) = -f'(x;d)$.

(b) If f is differentiable at the points where it is finite, then for each $x \in \mathbb{R}^n$, where f is finite, and each nonzero vector $d \in \mathbb{R}^n$, the function f has a directional derivative at x in the direction d and is given by

$$f'(x;d) = \sum_{i=1}^{n} d_i \frac{\partial f(x)}{\partial x_i} = \langle \nabla f(x), d\rangle.$$

In particular, if $d = (0,0,\ldots,0,1,0,\ldots,0,0) = e_i$, where 1 is at the ith place, then $f'(x;e_i) = \dfrac{\partial f(x)}{\partial x_i}$ is the partial derivative of f with respect to x_i.

(c) For the function f defined on $\mathbb{R}$, that is, $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$, we denote $f'_+(x;1)$ and $f'_-(x;1)$ by $f'_+(x)$ and $f'_-(x)$, respectively.

Example 2.1. Let $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ be a function defined by

$$f(x) = \begin{cases} +\infty, & \text{if } x < 0, \\ 2, & \text{if } x = 0, \\ x, & \text{if } 0 < x \le 1, \\ x^2, & \text{if } 1 < x \le 2, \\ +\infty, & \text{if } x > 2. \end{cases}$$

Then, using the convention given in Chapter 1, we have

$$f'_+(x;1) = f'_+(x) = \begin{cases} +\infty, & \text{if } x < 0, \\ -\infty, & \text{if } x = 0, \\ 1, & \text{if } 0 < x < 1, \\ 2x, & \text{if } 1 \le x < 2, \\ +\infty, & \text{if } x \ge 2, \end{cases}$$

and

$$f'_-(x;1) = f'_-(x) = \begin{cases} +\infty, & \text{if } x < 0, \\ -\infty, & \text{if } x = 0, \\ 1, & \text{if } 0 < x \leq 1, \\ 2x, & \text{if } 1 < x \leq 2, \\ +\infty, & \text{if } x > 2. \end{cases}$$

We see that $f'_+(x) = f'_-(x)$ for $0 \leq x < 1$ and $1 < x < 2$, and $f'_+(x) > f'_-(x)$ for $x = 1$ and $x = 2$.

The following result ensures the existence of $f'_+(x;d)$ and $f'_-(x;d)$ when f is a convex function.

Theorem 2.1. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be an extended real-valued convex function and x be a point in $\mathbb{R}^n$ where f is finite. Then, for each direction $d \in \mathbb{R}^n$, the ratio $\dfrac{f(x+td) - f(x)}{t}$ is a nondecreasing function of $t > 0$, $f'_+(x;d)$ and $f'_-(x;d)$ exist for every direction d and

$$f'_+(x;d) = \inf_{t>0} \frac{f(x+td) - f(x)}{t}, \tag{2.1}$$

and

$$f'_-(x;d) = \sup_{t<0} \frac{f(x+td) - f(x)}{t}. \tag{2.2}$$

Moreover, $f'_+(x;d)$ is a convex and positively homogeneous function of d with

$$f'_-(x;d) \leq f'_+(x;d). \tag{2.3}$$

Proof. Let $x \in \mathbb{R}^n$ be any point such that $f(x)$ is finite. Let $0 < t_1 < t_2$. Then, by convexity of f, we have

$$\begin{aligned} f(x + t_1 d) &= f\left(\frac{t_1}{t_2}(x + t_2 d) + \left(1 - \frac{t_1}{t_2}\right) x\right) \\ &\leq \frac{t_1}{t_2} f(x + t_2 d) + \left(1 - \frac{t_1}{t_2}\right) f(x). \end{aligned}$$

This inequality implies that

$$\frac{f(x + t_1 d) - f(x)}{t_1} \leq \frac{f(x + t_2 d) - f(x)}{t_2}.$$

Thus, the ratio $\dfrac{f(x+td) - f(x)}{t}$ is a nondecreasing function of t for $t > 0$.

For given $t > 0$, by the convexity of f, we have

$$\begin{aligned} f(x) &= f\left(\frac{t}{1+t}(x - d) + \frac{1}{1+t}(x + td)\right) \\ &\leq \frac{t}{1+t} f(x - d) + \frac{1}{1+t} f(x + td) \\ &= \frac{1}{1+t}\left(t f(x - d) + f(x + td)\right). \end{aligned}$$

It follows that $(1+t)f(x) \leq tf(x-d) + f(x+td)$, and so,

$$\frac{f(x+td)-f(x)}{t} \geq f(x) - f(x-d).$$

Hence the decreasing sequence of values $\dfrac{f(x+td)-f(x)}{t}$, as $t \to 0^+$, is bounded below by the constant $f(x)-f(x-d)$. Thus, the limit in the definition of $f'_+(x;d)$ exists and is given by

$$f'_+(x;d) = \lim_{t\to 0^+} \frac{f(x+td)-f(x)}{t} = \inf_{t>0} \frac{f(x+td)-f(x)}{t}.$$

Since $f'_+(x;d)$ exists in every direction d, the equality $-f'_+(x;-d) = f'_-(x;d)$ implies that $f'_-(x;d)$ exists in every direction d.

The relation (2.2) can be established on the lines of the proof given to derive (2.1).

Let $\lambda > 0$ be a real number. Then,

$$f'_+(x;\lambda d) = \lim_{\lambda t\to 0^+} \frac{\lambda(f(x+\lambda td)-f(x))}{\lambda t} = \lambda f'_+(x;d).$$

Hence $f'_+(x;\cdot)$ is positively homogeneous.

Similarly, we can show that $f'_-(x;\cdot)$ is also positively homogeneous.

Next, we show that $f'_+(x;\cdot)$ is convex. Let $d_1, d_2 \in \mathbb{R}^n$ and $\lambda_1, \lambda_2 \geq 0$ be such that $\lambda_1 + \lambda_2 = 1$. From the convexity of f, we have

$$\begin{aligned} f\left(x + t(\lambda_1 d_1 + \lambda_2 d_2)\right) \quad - \quad & f(x) \\ = \quad & f\left((\lambda_1+\lambda_2)x + t(\lambda_1 d_1 + \lambda_2 d_2)\right) - (\lambda_1+\lambda_2)f(x) \\ = \quad & f\left(\lambda_1(x+td_1) + \lambda_2(x+td_2)\right) - \lambda_1 f(x) - \lambda_2 f(x) \\ \leq \quad & \lambda_1 f(x+td_1) + \lambda_2 f(x+td_2) - \lambda_1 f(x) - \lambda_2 f(x) \\ = \quad & \lambda_1\left(f(x+td_1) - f(x)\right) + \lambda_2\left(f(x+td_2) - f(x)\right) \end{aligned}$$

for all sufficiently small t. Dividing by $t > 0$ and letting $t \to 0^+$, we obtain

$$f'_+(x;\lambda_1 d_1 + \lambda_2 d_2) \leq \lambda_1 f'_+(x;d_1) + \lambda_2 f'_+(x;d_2).$$

Hence $f'_+(x;d)$ is convex in d.

By subadditivity of $f'_+(x;d)$ in d with $f'_+(x;d) < +\infty$ and $f'_+(x;-d) < +\infty$, we obtain

$$f'_+(x;d) + f'_+(x;-d) \geq f'_+(x;\mathbf{0}) = 0,$$

and thus,

$$f'_+(x;d) \geq -f'_+(x;-d) = f'_-(x;d).$$

If $f'_+(x;d) = +\infty$ or $f'_+(x;-d) = +\infty$, then the inequality (2.3) holds trivially. $\square$

Example 2.2. Let $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ be defined as

$$f(x) = \begin{cases} 0, & \text{if } x = 0, \\ +\infty, & \text{if } x \neq 0. \end{cases}$$

Then,

$$f'_-(0; d) = \begin{cases} 0, & \text{if } d = 0, \\ -\infty, & \text{if } d \neq 0, \end{cases}$$

which is a concave function of d. However,

$$f'_+(0; d) = \begin{cases} 0, & \text{if } d = 0, \\ +\infty, & \text{if } d \neq 0, \end{cases}$$

is a convex function of d.

Then next theorem provides a relation between the directional derivative of a function f and a sequence of functions converging continuously to f. We first give the notion of continuous convergence.

Definition 2.2. Let K be nonempty subset of $\mathbb{R}^n$. A sequence of functions $\{f_m\}$, where $f_m : K \to \mathbb{R}$, is said to *converge continuously* to a function $f : K \to \mathbb{R}$ if for every $x \in K$ and every sequence $\{x_m\}$ in K converging to x, $\{f_m(x_m)\}$ converges to $f(x)$.

Theorem 2.2. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a convex function. Let $\{f_m\}$ be a sequence of convex functions $f_m : K \to \mathbb{R}$ such that it converges continuously to f. Then for all $x \in K$, $d \in \mathbb{R}^n$, and any sequence $\{x_m\}$ in K and $\{d_m\}$ in $\mathbb{R}^n$ converging to x and d, respectively, we have

$$\limsup_{m\to\infty} (f_m)'_+(x_m; d_m) \leq f'_+(x; d).$$

Proof. From the definition of right-sided directional derivative, it follows that for any $\varepsilon > 0$, there exists sufficiently small $t > 0$ such that

$$\frac{f(x + td) - f(x)}{t} < f'_+(x; d) + \varepsilon.$$

Therefore, for all sufficiently large m, we have, using (2.1),

$$(f_m)'_+(x_m; d_m) \leq \frac{f_m(x_m + td_m) - f_m(x_m)}{t} < f'_+(x; d) + 2\varepsilon.$$

By taking the limit as $m \to \infty$, we obtain

$$\limsup_{m\to\infty} (f_m)'_+(x_m; d_m) \leq f'_+(x; d) + 2\varepsilon.$$

Since ε was arbitrary, it is true for all $\varepsilon > 0$, and thus,

$$\limsup_{m\to\infty} (f_m)'_+(x_m; d_m) \leq f'_+(x; d).$$

This completes the proof. □

Remark 2.2. If $f_m = f$ for all m, then Theorem 2.2 says that $f'(x;d)$ is upper semicontinuous as a function of (x, d).

The following result shows that the subgradient of a convex function can be characterized in terms of its directional derivative.

Theorem 2.3. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued convex function and $x \in \mathbb{R}^n$ be such that $f(x)$ is finite. Then, a vector $\xi \in \mathbb{R}^n$ is a subgradient of f at x if and only if

$$f'_+(x;d) \geq \langle \xi, d\rangle, \tag{2.4}$$

for every direction $d \in \mathbb{R}^n$.

Proof. Let ξ be a subgradient of f at x. Then,

$$f(y) \geq f(x) + \langle \xi, y - x\rangle, \quad \text{for all } y \in \mathbb{R}^n. \tag{2.5}$$

Setting $y = x + td$, where $t > 0$, in (2.5), we have

$$\frac{f(x+td) - f(x)}{t} \geq \langle \xi, d\rangle, \quad \text{for all } d \in \mathbb{R}^n. \tag{2.6}$$

Since the difference quotient decreases to $f'_+(x;d)$ as $t \to 0^+$, (2.4) holds.

Conversely, assume that (2.4) holds for every $d \in \mathbb{R}^n$. Then (2.6) holds by the same argument as above, and consequently, (2.5) holds. Thus, ξ is a subgradient of f at x. $\square$

From Theorems 2.1 and 2.3, we have the following results.

Corollary 2.1. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued convex function and $x \in \mathbb{R}^n$ be such that $f(x)$ is finite. Then,

$$f(y) \geq f(x) + f'_+(x; y - x), \quad \text{for all } y \in \mathbb{R}^n.$$

In particular, if f is differentiable at x, then

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle, \quad \text{for all } y \in \mathbb{R}^n.$$

Corollary 2.2. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued convex function and $x, y \in \mathbb{R}^n$ be such that $f(x)$ and $f(y)$ are finite. Then,

$$f'_+(y; y - x) \geq f'_+(x; y - x), \tag{2.7}$$

and

$$f'_-(y; y - x) \geq f'_-(x; y - x). \tag{2.8}$$

In particular, if f is differentiable at x and y, then

$$\langle \nabla f(y) - \nabla f(x), y - x\rangle \geq 0. \tag{2.9}$$

Proof. From Corollary 2.1, we have

$$f(y) \geq f(x) + f'_+(x; y - x), \tag{2.10}$$

and

$$f(x) \geq f(y) + f'_+(y; x - y). \tag{2.11}$$

By adding inequalities (2.10) and (2.11), we obtain

$$-f'_+(y; x - y) \geq f'_+(x; y - x).$$

Since $-f'_+(x; -d) = f'_-(x; d)$, by using inequality (2.3), we get

$$f'_+(y; y - x) \geq f'_-(y; y - x) = -f'_+(y; x - y) \geq f'_+(x; y - x).$$

Hence, the inequality (2.7) holds. Similarly, we can establish the inequality (2.8). The inequality (2.9) holds using Remark 2.1(b). □

We present a mean value theorem for directional derivatives.

Theorem 2.4. Let K be a nonempty open subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a differentiable function. If the line segment joining x and $x + d$ lies in K, then there exists $s \in \ ]0, 1[$ such that

$$f(x + d) - f(x) = f'(x + sd; d). \tag{2.12}$$

Proof. Since K is an open subset of $\mathbb{R}^n$, we can select an open interval I of real numbers, which contains the numbers 0 and 1, such that $x + \lambda d$ belongs to K for all $\lambda \in I$. For all $\lambda \in I$, define

$$\varphi(\lambda) = f(x + \lambda d).$$

Then,

$$\begin{aligned} \varphi'(\lambda) &= \lim_{\tau \to 0} \frac{\varphi(\lambda + \tau) - \varphi(\lambda)}{\tau} \\ &= \lim_{\tau \to 0} \frac{f(x + \lambda d + \tau d) - f(x + \lambda d)}{\tau} \\ &= f'(x + \lambda d; d). \end{aligned} \tag{2.13}$$

By applying the mean value theorem for real-valued functions of one variable to the restriction of the function $\varphi : I \to \mathbb{R}$ to the closed interval $[0, 1]$, we obtain

$$\varphi(1) - \varphi(0) = \varphi'(s), \quad \text{for some } s \in \ ]0, 1[.$$

By using (2.13) and the definition of $\varphi : [0, 1] \to \mathbb{R}$, we obtain (2.12). □

For the differentiable function we have the following result which follows from Theorem 2.4.

Corollary 2.3. If in Theorem 2.4 f is a differentiable function, then there exists $s \in \ ]0, 1[$ such that $f(x + d) - f(x) = \langle \nabla f(x + sd), d \rangle$.

2.3 Gâteaux Derivatives

We define the Gâteaux derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$, which is a generalization of the directional derivative.

Definition 2.3. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *Gâteaux differentiable* at a point $x \in \mathbb{R}^n$ if the directional derivative $f'(x; d)$ exists for all $d \in \mathbb{R}^n$, that is, if

$$f^G(x; d) = \lim_{t \to 0} \frac{f(x + td) - f(x)}{t} \tag{2.14}$$

exists for all $d \in \mathbb{R}^n$.

A function $f^G : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is called the *Gâteaux derivative* of f at $x \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^n$ if f is Gâteaux differentiable at $x \in \mathbb{R}^n$.

$f^G(x; d)$ is called the *value of the Gâteaux derivative* of f at x in the direction d.

The relation (2.14) is equivalent to the following relation

$$\lim_{t \to 0} \left| \frac{f(x + td) - f(x)}{t} - f^G(x; d) \right| = 0. \tag{2.15}$$

We establish that the Gâteaux derivative is unique.

Proposition 2.1. The Gâteaux derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ is unique provided it exists.

Proof. Assume that there exist two operators f^G and f^{*G} that satisfy (2.15). Then, for all $x \in \mathbb{R}^n$ and all $d \in \mathbb{R}^n$, and for sufficiently small t, we have

$$\begin{aligned}
\left| f^G(x; d) - f^{*G}(x; d) \right| &= \left| \left(\frac{f(x + td) - f(x)}{t} - f^G(x; d) \right) \right. \\
&\qquad \left. - \left(\frac{f(x + td) - f(x)}{t} - f^{*G}(x; d) \right) \right| \\
&\leq \left| \frac{f(x + td) - f(d)}{t} - f^G(x; d) \right| \\
&\qquad + \left| \frac{f(x + td) - f(d)}{t} - f^{*G}(x; d) \right| \\
&\to 0 \text{ as } t \to 0.
\end{aligned}$$

Therefore, $\left| f^G(x; d) - f^{*G}(x; d) \right| = 0$ for all $x \in \mathbb{R}^n$ and all $d \in \mathbb{R}^n$. Hence, $f^G(x; d) = f^{*G}(x; d)$, and thus, $f^G \equiv f^{*G}$. ☐

Remark 2.3. As we have seen in Remark 2.1(a), if $f^G(x; d)$ exists, then $f^G(x; -d) = -f^G(x; d)$.

If $f : \mathbb{R}^n \to \mathbb{R}$ has continuous partial derivatives of order 1, then f is

Gâteaux differentiable at $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and in the direction $d = (d_1, d_2, \ldots, d_n) \in \mathbb{R}^n$, and it is given by

$$f^G(x; d) = \sum_{k=1}^{n} \frac{\partial f(x)}{\partial x_k} d_k.$$

For a fixed $a \in \mathbb{R}^n$, the value of the Gâteaux derivative at a is

$$f^G(a; d) = \sum_{k=1}^{n} \frac{\partial f(x)}{\partial x_k} d_k \Bigg|_{x=a},$$

and the Gâteaux derivative f^G is a bounded linear operator from $\mathbb{R}^n \times \mathbb{R}^n$ into $\mathbb{R}$. $f^G(a; d)$ can also be written in the form of an inner product as

$$f^G(a; d) = \langle y, d \rangle,$$

where

$$y = \left(\frac{\partial f(a)}{\partial x_1}, \frac{\partial f(a)}{\partial x_2}, \ldots, \frac{\partial f(a)}{\partial x_n} \right).$$

Theorem 2.5. Let K be a nonempty open convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a convex function. If f is Gâteaux differentiable at $x \in K$, then $f^G(x; d)$ is linear in d. Conversely, if $f'_+(x; d)$ is linear in d, then f is Gâteaux differentiable at x.

Proof. Let f be Gâteaux differentiable at $x \in K$, then for all $d \in \mathbb{R}^n$

$$-f'_+(x; -d) = f'_-(x; d) = f'_+(x; d).$$

Therefore, for all $d, u \in \mathbb{R}^n$, we have

$$\begin{aligned} f'_+(x; d) + f'_+(x; u) &\geq f'_+(x; d + u) \\ &= -f'_+(x; -(d + u)) \\ &\geq -f'_+(x; -d) - f'_+(x; -u) \\ &= f'_+(x; d) + f'_+(x; u), \end{aligned}$$

and thus,

$$f'_+(x; d + u) = f'_+(x; d) + f'_+(x; u).$$

Since $f^G(x; d) = f'_+(x; d) = f'_-(x; d)$, we have

$$f^G(x; d + u) = f^G(x; d) + f^G(x; u).$$

For $\alpha \in \mathbb{R}$ with $\alpha \neq 0$, we have

$$\begin{aligned} f^G(x; \alpha d) &= \lim_{\alpha t \to 0} \frac{\alpha(f(x + t\alpha d) - f(x))}{\alpha t} \\ &= \alpha f^G(x; d). \end{aligned}$$

Hence $f^G(x;d)$ is linear in d.

Conversely, assume that $f'_+(x;d)$ is linear in d. Then,

$$0 = f'_+(x;d-d) = f'_+(x;d) + f'_+(x;-d).$$

Therefore, for all $d \in \mathbb{R}^n$, we have

$$f'_-(x;d) = -f'_+(x;-d) = f'_+(x;d).$$

Thus, f is Gâteaux differentiable at x. □

Remark 2.4. (a) A nonconvex function $f : \mathbb{R}^n \to \mathbb{R}$ may be Gâteaux differentiable at a point but the Gâteaux derivative may not be linear at that point. For example, consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{x_1^2 x_2}{x_1^2+x_2^2}, & \text{if } x \neq (0,0), \\ 0, & \text{if } x = (0,0), \end{cases}$$

where $x = (x_1, x_2)$. For $d = (d_1, d_2) \neq (0,0)$ and $t \neq 0$, we have

$$\frac{f((0,0)+t(d_1,d_2)) - f(0,0)}{t} = \frac{d_1^2 d_2}{d_1^2 + d_2^2}.$$

Then,

$$f^G((0,0);d) = \lim_{t \to 0} \frac{f((0,0)+t(d_1,d_2)) - f(0,0)}{t} = \frac{d_1^2 d_2}{d_1^2 + d_2^2}.$$

Therefore, f is Gâteaux differentiable at $(0,0)$, but $f^G((0,0);d)$ is not linear in d.

(b) For a real-valued function f defined on $\mathbb{R}^n$, the partial derivatives may exist at a point, but f may not be Gâteaux differentiable at that point. For example, consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{x_1 x_2}{x_1^2+x_2^2}, & \text{if } x \neq (0,0), \\ 0, & \text{if } x = (0,0), \end{cases}$$

where $x = (x_1, x_2)$. For $d = (d_1, d_2) \neq (0,0)$ and $t \neq 0$, we have

$$\frac{f((0,0)+t(d_1,d_2)) - f(0,0)}{t} = \frac{d_1 d_2}{t(d_1^2 + d_2^2)}.$$

Then,

$$\lim_{t \to 0} \frac{f((0,0)+t(d_1,d_2)) - f(0,0)}{t} = \lim_{t \to 0} \frac{d_1 d_2}{t(d_1^2 + d_2^2)},$$

exists only if $d = (d_1, 0)$ or $d = (0, d_2)$. That is, $f^G(\mathbf{0};\mathbf{0})$ does not exist but $\dfrac{\partial f(0,0)}{\partial x_1} = 0 = \dfrac{\partial f(0,0)}{\partial x_2}$, where $\mathbf{0} = (0,0)$ is the zero vector in $\mathbb{R}^2$.

(c) The existence, linearity, and continuity of $f^G(x; d)$ in d do not imply the continuity of the function f. For example, consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{x_1^3}{x_2}, & \text{if } x_1 \neq 0 \text{ and } x_2 \neq 0, \\ 0, & \text{if } x_1 = 0 \text{ or } x_2 = 0, \end{cases}$$

where $x = (x_1, x_2)$. Then,

$$f^G((0,0); d) = \lim_{t \to 0} \frac{t^3 d_1^3}{t^2 d_2} = 0,$$

for all $d = (d_1, d_2) \in \mathbb{R}^2$ with $(d_1, d_2) \neq (0,0)$. Thus, $f^G(\mathbf{0}; d)$ exists and it is continuous and linear in d, but f is discontinuous at $(0,0)$. The function f is Gâteaux differentiable but not continuous. Hence a Gâteaux differentiable function is not necessarily continuous.

(d) The Gâteaux derivative $f^G(x; d)$ of a function f is positively homogeneous in the second argument, that is, $f^G(x; rd) = rf^G(x; d)$ for all $r > 0$. But, as we have seen in part (a), in general, $f^G(x; d)$ is not linear in d.

The following theorem shows that the partial derivatives and Gâteaux derivative are the same if the function f defined on $\mathbb{R}^n$ is convex.

Theorem 2.6. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a convex function. If the partial derivatives of f at $x \in K$ exist, then f is Gâteaux differentiable at x.

Proof. Suppose that the partial derivatives of f at $x \in K$ exist. Then, the Gâteaux derivative of f at x is the linear functional

$$f^G(x; d) = \sum_{k=1}^{n} \frac{\partial f(x)}{\partial x_k} d_k, \quad \text{for } d = (d_1, d_2, \ldots d_n) \in \mathbb{R}^n.$$

For each fixed $x \in K$, define a function $g : K \to \mathbb{R}$ by

$$g(d) = f(x + d) - f(x) - f^G(x; d).$$

Then, g is convex and $\dfrac{\partial g(0)}{\partial x_k} = 0$ for all $k = 1, 2, \ldots, n$, since the partial derivatives of f exist at x. Now, if $\{e_1, e_2, \ldots, e_n\}$ is the standard basis for $\mathbb{R}^n$ then by the convexity of g, we have for $\lambda \neq 0$

$$\begin{aligned} g(\lambda d) &= g\left(\lambda \sum_{k=1}^{n} d_k e_k\right) \\ &\leq \frac{1}{n} \sum_{k=1}^{n} g\left(n\lambda d_k e_k\right) \\ &= \lambda \sum_{k=1}^{n} \frac{g(n\lambda d_k e_k)}{n\lambda}. \end{aligned}$$

So,

$$\frac{g(\lambda d)}{\lambda} \leq \sum_{k=1}^{n} \frac{g(n\lambda d_k e_k)}{n\lambda}, \quad \text{for } \lambda > 0,$$

and

$$\frac{g(\lambda d)}{\lambda} \geq \sum_{k=1}^{n} \frac{g(n\lambda d_k e_k)}{n\lambda}, \quad \text{for } \lambda < 0.$$

Since

$$\lim_{\lambda \to 0} \frac{g(n\lambda d_k e_k)}{n\lambda} = \frac{\partial g(0)}{\partial d_k} = 0, \quad \text{for all } k = 1, 2, \ldots, n,$$

we have

$$\lim_{\lambda \to 0} \frac{g(\lambda d)}{\lambda} = 0,$$

and so, f is Gâteaux differentiable at x. □

We now mention a mean value theorem in terms of the Gâteaux derivative whose proof is similar to the proof of Theorem 2.4, and, therefore, we omit it.

Theorem 2.7. Let K be a nonempty open subset of $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be Gâteaux differentiable with Gâteaux derivative $f^G(x; d)$ at $x \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^n$. Then for any points $x \in \mathbb{R}^n$ and $x + d \in \mathbb{R}^n$, there exists $s \in \ ]0, 1[$ such that

$$f(x + d) - f(x) = f^G(x + sd; d).$$

2.4 Dini and Dini-Hadamard Derivatives

For nonconvex functions, the limit in the definitions of directional derivatives and Gâteaux derivatives may not exist. Therefore, other kinds of generalized directional derivatives were developed that are useful in the applications for nonconvex cases.

Definition 2.4. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a function and x be a point where f is finite.

(a) The *Dini upper directional derivative* at $x \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^n$ is defined by

$$\begin{aligned} f^D(x; d) &= \limsup_{t \to 0^+} \frac{f(x + td) - f(x)}{t} \\ &= \inf_{s>0} \sup_{0<t<s} \frac{f(x + td) - f(x)}{t}. \end{aligned}$$

(b) The *Dini lower directional derivative* at $x \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^n$ is defined by

$$\begin{aligned} f_D(x;d) &= \liminf_{t \to 0^+} \frac{f(x+td) - f(x)}{t} \\ &= \sup_{s>0} \inf_{0<t<s} \frac{f(x+td) - f(x)}{t}. \end{aligned}$$

(c) Similarly, we define

$$f_-^D(x;d) = \limsup_{t \to 0^-} \frac{f(x+td) - f(x)}{t}$$

and

$$f_D^-(x;d) = \liminf_{t \to 0^-} \frac{f(x+td) - f(x)}{t}.$$

These definitions can also be given for functions defined on a convex subset K of $\mathbb{R}^n$ at a point $x \in K$ in the direction $d \in \mathbb{R}^n$, provided $x + td \in K$ for sufficiently small $t > 0$.

Since for each $d \in \mathbb{R}^n$, we have

$$f_-^D(x;d) = -f_D(x;-d) \quad \text{and} \quad f_D^-(x;d) = -f^D(x;-d),$$

therefore, it is quite obvious that it is enough to deal only with the Dini upper directional derivative and the Dini lower directional derivative.

Also, it can be observed that $-f_D(x;d) = (-f)^D(x;d)$ for all $x \in \mathbb{R}^n$ and $d \in \mathbb{R}^n$.

For a function f defined on $\mathbb{R}$, that is, $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$, we denote $f^D(x;1)$, $f_D(x;1)$, $f_-^D(x;1)$; and $f_D^-(x;1)$ by $f^D(x)$, $f_D(x)$, $f_-^D(x)$, and $f_D^-(x)$, respectively. In other words, we write $f^D(x)$ in place of $f^D(x;d)$ if the directional vector d is the scalar 1, that is,

$$f^D(x) = \limsup_{t \to x^+} \frac{f(t) - f(x)}{t - x},$$

and

$$f_D(x) = \liminf_{t \to x^+} \frac{f(t) - f(x)}{t - x}.$$

By using Definition 2.4, it can be easily seen that for any $x, d \in \mathbb{R}^n$, we have

$$f_D(x;d) \leq f^D(x;d),$$

and, of course, when we get the equality in the above inequality, we obtain the right-sided directional derivative $f'_+(x;d)$. However, this equality does not ensure the convexity of the directional derivative, which is very important in certain applications. We further note that if the Dini upper and Dini lower directional derivatives in a direction d are finite at a given point, then the function is continuous at that point along the direction d. But the converse need not be true in general. For example, consider the real-valued function $f(x) = \sqrt{|x|}$ for all $x \in \mathbb{R}$. Then f is continuous, but its Dini upper and Dini lower directional derivatives at $x = 0$ in the direction $d = 1$ and $d = -1$ are infinite. One of the most important features of the Dini upper and Dini lower directional derivatives is that they always exist even when the function is discontinuous, although they are not necessarily finite.

We observe that the function $f(x) = |x|$ does not have a derivative at the point $x = 0$ but does have one-sided derivatives $f'_+(0) = 1$ and $f'_-(0) = -1$. The following example shows that a continuous function may not have even a one-sided directional derivative at a point but it may have the Dini derivative at that point. For further details, we refer to Thomson et al. [200].

Example 2.3. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} |x| \left|\cos\left(\frac{1}{x}\right)\right|, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Since $\left|\cos\left(\frac{1}{x}\right)\right| \leq 1$ for all $x \neq 0$,

$$\lim_{x \to 0} f(x) = 0 = f(0),$$

so f is continuous at $x = 0$. It is clear that f is also continuous at all the other points of $\mathbb{R}$, so f is a continuous function.

The oscillatory behavior of f is such that the sets

$$\left\{x : \left|\cos\left(\frac{1}{x}\right)\right| = 1\right\} \quad \text{and} \quad \left\{x : \left|\cos\left(\frac{1}{x}\right)\right| = 0\right\}$$

both have zero as a two-sided limit point. Thus, each of the sets

$$\{x : f(x) = |x|\} \quad \text{and} \quad \{x : f(x) = 0\}$$

has zero as a two-sided limit point. Inspection of the difference quotient reveals that

$$\limsup_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} = 1, \quad \text{while} \quad \liminf_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} = 0,$$

so $f'_+(x)$ does not exist at $x = 0$. Similarly, $f'_-(0)$ does not exist.

However,

$$f^D(0) = 1, \quad f_D(0) = 0, \quad f_-^D(0) = 0, \quad f_D^-(0) = -1,$$

and elsewhere $f'(x)$ exists and all the four Dini derivatives have the same value.

In the next example, the given function is not continuous.

Example 2.4. Consider the discontinuous function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0, & \text{if } x \text{ is rational}, \\ 1, & \text{if } x \text{ is irrational}. \end{cases}$$

Then, at every rational x,

$$f^D(x) = \infty, \quad f_D(x) = 0, \quad f_-^D(x) = \infty, \quad f_D^-(x) = 0.$$

For x irrational, there are similar values for the Dini derivatives.

The following result presents some elementary properties and calculus rules for Dini upper (Dini lower) directional derivatives. The proof follows directly from the definition of Dini upper (Dini lower) directional derivatives, therefore, we omit it. For further study, we refer to [69, 86, 87, 88, 89, 96, 113, 159].

Theorem 2.8. Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be real-valued functions. The following assertions hold:

(a) Homogeneity: $f^D(x; d)$ is positively homogeneous in d, that is, for all $r > 0$ we have $f^D(x; rd) = rf^D(x; d)$.

(b) Scalar multiple: For $r > 0$, $(rf)^D(x; d) = rf^D(x; d)$, and for $r < 0$, $(rf)^D(x; d) = rf_D(x; d)$.

(c) Sum rule: $(f + g)^D(x; d) \leq f^D(x; d) + g^D(x; d)$, provided that the sum on the right-hand side exists.

(d) Product rule: $(fg)^D(x; d) \leq [g(x)f]^D(x; d) + [f(x)g]^D(x; d)$, provided that the sum on the right-hand side exists, the functions f and g are continuous at x, and that one of the following conditions is satisfied: $f(x) \neq 0$; $g(x) \neq 0$; $f^D(x; d)$ is finite; and $g^D(x; d)$ is finite.

(e) Quotient rule: $\left(\frac{f}{g}\right)^D(x; d) \leq \dfrac{[g(x)f]^D(x; d) + [-f(x)g]^D(x; d)}{[g(x)]^2}$, provided that the expression on the right-hand side exists and the function g is continuous at x.

If, in addition, the functions f and g are directionally differentiable at x, then the inequalities in the last three assertions become equalities.

Properties and calculus rules for Dini lower directional derivatives can be obtained in a similar manner. The next result shows that Dini upper and Dini lower directional derivatives are convenient tools for characterizing an extremum of a function.

Theorem 2.9. For a function $f : \mathbb{R}^n \to \mathbb{R}$, the following assertions hold:

(a) If $f(x) \leq f(x + td)$ (respectively, $f(x) \geq f(x + td)$) for all $t > 0$ sufficiently small, then $f_D(x; d) \geq 0$ (respectively, $f^D(x; d) \leq 0$). In particular, if f is directionally differentiable at x, and $f(x) \leq f(x + td)$ (respectively, $f(x) \geq f(x + td)$) for all $t > 0$ sufficiently small, then $f'(x; d) \geq 0$ (respectively, $f'(x; d) \leq 0$).

(b) If $f^D(x+td;d) \geq 0$ for all $x, d \in \mathbb{R}^n$ and $t \in\ ]0,1[$ and if the function $t \mapsto f(x+td)$ is continuous on $[0,1]$, then $f(x) \leq f(x+d)$.

Proof. (a) It follows directly from the definitions of Dini lower and Dini upper directional derivatives.

(b) Assume the contrary that $f(x) > f(x+d)$ for some $x, d \in \mathbb{R}^n$. Consider the function $g : [0,1] \to \mathbb{R}$ defined by

$$g(t) = f(x+td) - f(x) + t(f(x) - f(x+d)).$$

Then, g is continuous on $[0,1]$ and $g(0) = g(1) = 0$. Thus, there exists $\hat{t} \in [0,1[$ such that g has a maximum value at $\hat{t}$. Set $y := x + \hat{t}d$. Then,

$$g(\hat{t}) \geq g(\hat{t}+t), \quad \text{for all } t \in [0, 1-\hat{t}],$$

and hence,

$$f(y+td) - f(y) \leq t(f(x+d) - f(x)), \quad \text{for all } t > 0 \text{ sufficiently small.}$$

Dividing the above inequality by t and taking the limit superior as $t \to 0^+$, we obtain

$$f^D(y;d) = \limsup_{t \to 0^+} \frac{f(y+td) - f(y)}{t} \leq f(x+d) - f(x) < 0,$$

a contradiction to the hypothesis. This completes the proof. □

We next prove the following mean value theorem for upper semicontinuous functions of one variable given in Diewert [69].

Theorem 2.10 (Diewert's Mean Value Theorem)**.** If $f : [a,b] \to \mathbb{R}$ is an upper semicontinuous function, then there exists $c \in [a,b[$ such that

$$f_D(c) \leq f^D(c) \leq \frac{f(b) - f(a)}{b-a}.$$

Proof. Let $\beta = \dfrac{f(b) - f(a)}{b-a}$. Define a function $g : [0,1] \to \mathbb{R}$ by

$$g(t) = f(t) - \beta t.$$

Then as f is upper semicontinuous on $[a,b]$, it follows that g is also upper semicontinuous on the compact set $[a,b]$. By Berge's maximum theorem [31], there exists a point $c \in [a,b]$ such that g attains its maximum at c. Then, $g(c) \geq g(t)$ for all $t \in [a,b]$, that is,

$$f(t) - f(c) \leq \beta(t-c), \quad \text{for all } t \in [a,b].$$

If $c \in [a,b)$, then from the above inequality, we get

$$\frac{f(t) - f(c)}{t-c} \leq \beta, \quad \text{for all } t \in\ ]c,b].$$

The above inequality yields that $f_D(c) \leq f^D(c) \leq \beta$. □

We next state the mean value theorem for lower semicontinuous functions.

Corollary 2.4. If $f : [a,b] \to \mathbb{R}$ is a lower semicontinuous function, then there exists $c \in [a,b[$ such that

$$f^D(c) \geq f_D(c) \geq \frac{f(b)-f(a)}{b-a}.$$

Proof. The proof follows by giving the same argument as in Theorem 2.10. □

Combining Theorem 2.10 and Corollary 2.4, we have the following mean value theorem for continuous functions.

Corollary 2.5. If $f : [a,b] \to \mathbb{R}$ is a continuous function, then there exist $c, d \in [a,b[$ such that

$$f^D(c)(b-a) \leq f(b)-f(a) \leq f_D(d)(b-a).$$

Corollary 2.6. Let $f : [a,b] \to \mathbb{R}$ be an upper semicontinuous function. If for all $c \in [a,b[$, $f^D(c) \geq 0$ (respectively, $f^D(c) \leq 0$), then f is a nondecreasing (respectively, nonincreasing) function on $[a,b]$.

Proof. Let us assume that $f^D(c) \geq 0$ for all $c \in [a,b[$. Assume to the contrary that f is strictly decreasing on $[a,b]$. Then there exist $t_1, t_2 \in [a,b[$ such that $t_1 < t_2$ and $f(t_1) > f(t_2)$. By Theorem 2.10, there exists $t \in [t_1,t_2[$ such that

$$f^D(t) \leq \frac{f(t_2)-f(t_1)}{t_2-t_1},$$

which implies that $f^D(t) < 0$, which is a contradiction to the hypothesis.

By using Corollary 2.4, it can similarly be shown that f is nonincreasing on $[a,b]$ if $f^D(c) \leq 0$ for all $c \in [a,b[$. □

Definition 2.5. Let K be a nonempty convex subset of $\mathbb{R}^n$. A function $f : K \to \mathbb{R}$ is said to be *radially upper* (*lower*) *semicontinuous* on K (also known as *upper* (*lower*) *hemicontinuous* on K) if for every pair of distinct points $x, y \in K$, the function f is upper (lower) semicontinuous on the line segment $[x,y]$.

Also, the function f is said to be *radially continuous* on K (also known as *hemicontinuous* on K) if it is both radially upper semicontinuous and radially lower semicontinuous on K.

Example 2.5. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined as

$$f(x,y) = \begin{cases} \frac{2x^2y}{x^4+y^2}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

Then we observe that f is continuous at every point of $\mathbb{R}^2$ except at $(0,0)$. If we approach x along the path $y = mx^2$, then

$$\lim_{(x,y)\to(0,0)} f(x,y) = \frac{2m}{1+m^2},$$

which is different for different values of m. However, f is radially continuous, because if we approach $(0,0)$ along the line $y = mx$, then

$$\lim_{(x,y)\to(0,0)} f(x,y) = \frac{2mx}{x^2+m^2} = 0 = f(0,0).$$

We now give Diewert's mean value theorem for functions defined on a convex subset of $\mathbb{R}^n$.

Theorem 2.11 (Diewert's Mean Value Theorem)**.** Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a function. Then for every pair of distinct points x and y in K, the following assertions hold.

(a) If f is radially upper semicontinuous on K, then there exists $w \in [x,y[$ such that

$$f^D(w; y-x) \leq f(y) - f(x).$$

(b) If f is radially lower semicontinuous on K, then there exists $v \in [x,y[$ such that

$$f_D(v; y-x) \geq f(y) - f(x).$$

(c) If f is radially continuous on K, then there exist $w, v \in [x,y[$ such that

$$f^D(w; y-x) \leq f(y) - f(x) \leq f_D(v; y-x).$$

Moreover, if the Dini upper directional derivative $f^D(w; y-x)$ is continuous in w on the line segment $[x,y[$, then there exists a point $u \in [x,y[$ such that

$$f^D(u; y-x) = f(y) - f(x).$$

Proof. Let x and y be two distinct points in K. Define a function $g : [0,1] \to \mathbb{R}$ by

$$g(t) = f(x + t(y-x)).$$

(a) If f is radially upper semicontinuous on K, then g is an upper semicontinuous function on $[0,1]$. By Theorem 2.10, there exists $\hat{t} \in [0,1[$ such that

$$g^D(\hat{t}) \leq g(1) - g(0).$$

If we set $w = x + \hat{t}(y-x)$, then we have

$$g^D(\hat{t}) = f^D(w; y-x),$$

which leads to the result.

(b) Proof follows by using Corollary 2.4.

(c) Proof follows by using Corollary 2.5. If $f^D(w; y-x)$ is continuous in w on the line segment $[x,y[$, then by the intermediate value theorem, there exists a point $u \in [x,y[$ such that

$$f^D(u; y-x) = f(y) - f(x).$$

□

Remark 2.5. The continuity of $f^D(\cdot; y-x)$ in the second part of Theorem 2.11(c) cannot be dropped. For example, consider the function $f(x) = |x|$ defined on $\mathbb{R}$. Then it is directionally differentiable everywhere. For $x = -1$ and $y = 1$, we have

$$f'(w; y-x) = \begin{cases} -2, & \text{if } w < 0, \\ 2, & \text{if } w \geq 0, \end{cases}$$

which is discontinuous at $w = 0$. There is no u between x and y such that $0 = f(y) - f(x) = f'(u; y-x)$.

Lipschitz functions are continuous, but are not always directionally differentiable. For example, consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0, & \text{if } x \in \;]-\infty, 0] \cup [1, \infty[, \\ -2x + \frac{2}{3^n}, & \text{if } x \in \left[\frac{2}{3^{n+1}}, \frac{1}{3^n}\right[, \\ 2x - \frac{2}{3^{n+1}}, & \text{if } x \in \left[\frac{1}{3^{n+1}}, \frac{2}{3^{n+1}}\right[, \end{cases}$$

for $n = 0, 1, 2, \ldots$. Then f is Lipschitz on $\mathbb{R}$ with Lipschitz constant $k = 2$. However, for $x = 0$ and $d = 1$ we have $f^D(x; d) = 1$ and $f_D(x; d) = 0$, which shows that f is not directionally differentiable at x.

A generalization of the Dini (upper and lower) directional derivative is the Dini-Hadamard (upper and lower) directional derivative.

Definition 2.6. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a function and $x \in \mathbb{R}^n$ be a point where f is finite.

(a) The *Dini-Hadamard upper directional derivative* of f at x in the direction $d \in \mathbb{R}^n$ is defined by

$$f^{DH}(x; d) = \limsup_{\substack{v \to d \\ t \to 0^+}} \frac{f(x + tv) - f(x)}{t}.$$

(b) The *Dini-Hadamard lower directional derivative* of f at x in the direction $d \in \mathbb{R}^n$ is defined by

$$f_{DH}(x; d) = \liminf_{\substack{v \to d \\ t \to 0^+}} \frac{f(x + tv) - f(x)}{t}.$$

If $f^{DH}(x; d) = f_{DH}(x; d)$, then we denote it by $f^{DH*}(x; d)$, that is,

$$f^{DH*}(x; d) = \lim_{\substack{v \to d \\ t \to 0^+}} \frac{f(x + tv) - f(x)}{t}.$$

From the definitions of Dini upper (lower) directional derivative and Dini-Hadamard upper (lower) directional derivative, we can easily obtain the following relations:

$$(-f)^{DH}(x;d) = -f_{DH}(x;d), \quad (-f)_{DH}(x;d) = -f^{DH}(x;d),$$
$$f_{DH}(x;d) \leq f_D(x;d) \leq f^D(x;d) \leq f^{DH}(x;d). \tag{2.16}$$

In the following example the Dini (upper and lower) directional derivative and Dini-Hadamard (upper and lower) directional derivative are different.

Example 2.6. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function defined by

$$f(x_1, x_2) = \begin{cases} 0, & \text{if } x_2 = 0, \\ x_1 + x_2, & \text{if } x_2 \neq 0. \end{cases}$$

Let $x = (0,0)$ and $d = e_1 = (1,0)$. Then we can easily calculate that

$$f^{DH}(x;d) = 1 \quad \text{and} \quad f^D(x;d) = 0.$$

The following result shows that the Dini upper (lower) directional derivative and the Dini-Hadamard upper (lower) directional derivative at a point coincide if f is locally Lipschitz around that point.

Theorem 2.12. Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz around a point $x \in \mathbb{R}^n$. Then for every $d \in \mathbb{R}^n$,

$$f^{DH}(x;d) = f^D(x;d) \quad \text{and} \quad f_{DH}(x;d) = f_D(x;d).$$

Proof. Let $N(x)$ be a neighborhood of x and f be Lipschitz continuous on $N(x)$ with Lipschitz constant k. Let $d \in \mathbb{R}^n$ be arbitrary. Then there exist $\delta > 0$ and $\tau > 0$ such that for all $v \in \mathbb{R}^n$ and $t \in \mathbb{R}$ satisfying the conditions $\|v - d\| < \delta$ and $0 < t < \tau$, we have $x + td$, $x + tv \in N(x)$. Consequently,

$$|f(x + tv) - f(x + td)| \leq kt\|v - d\|.$$

Therefore,

$$\limsup_{\substack{v \to d \\ t \to 0^+}} \frac{f(x + td) - f(x + tv)}{t} = 0.$$

By applying the properties of lim sup, we obtain

$$\begin{aligned} f^{DH}(x;d) &= \limsup_{\substack{v \to d \\ t \to 0^+}} \frac{f(x + tv) - f(x)}{t} \\ &\leq \limsup_{\substack{v \to d \\ t \to 0^+}} \frac{f(x + td) - f(x)}{t} + \limsup_{\substack{v \to d \\ t \to 0^+}} \frac{f(x + tv) - f(x + td)}{t} \\ &= f^D(x;d). \end{aligned}$$

Since $f^D(x;d) \leq f^{DH}(x;d)$, we have the equality.

Similarly, we can prove that $f_{DH}(x;d) = f_D(x;d)$. $\square$

Theorem 2.13. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Then $f_{DH}(x; d)$ (respectively, $f^{DH}(x; d)$) is a lower (respectively, upper) semicontinuous function in d.

Proof. To prove that $f_{DH}(x; d)$ is a lower semicontinuous function in d is equivalent to show that the upper level set $M(x, \alpha) = \{d \in \mathbb{R}^n : f_{DH}(x; d) > \alpha\}$ is open for all $\alpha \in \mathbb{R}$.

By the definition of the Dini-Hadamard lower directional derivative, we have

$$f_{DH}(x; d) = \lim_{\substack{\beta \to 0^+ \\ \tau \to 0^+}} \varphi_d(\beta, \tau),$$

where

$$\varphi_d(\beta, \tau) = \inf_{\substack{\|v-d\|<\beta \\ 0<t<\tau}} \frac{f(x + tv) - f(x)}{t}.$$

By the definition of $\varphi_d(\beta, \tau)$ we have $\varphi_d(\beta, \tau)$ is nonincreasing, and hence

$$\varphi_d(\beta, \tau) \leq f_{DH}(x; d), \qquad \text{for all positive numbers } \beta \text{ and } \tau.$$

Assume that $M(x, \alpha) \neq \emptyset$ and let $d \in M(x, \alpha)$ be arbitrary. Hence, there exist positive numbers β_0 and τ_0 such that $\alpha < \varphi_d(\beta_0, \tau_0)$.

Let $\mathbb{B}_\alpha(d)$ be the open ball with center at d and radius α. Then we show that $\mathbb{B}_\alpha(d) \subseteq M(x, \alpha)$, which will prove that $M(x, \alpha)$ is open. Let $w \in \mathbb{B}_\alpha(d)$ be an arbitrary element. Consider an open ball $\mathbb{B}_s(w)$ with center at w and radius s contained in $\mathbb{B}_\alpha(d)$. Then,

$$\alpha < \varphi_d(\beta_0, \tau_0) \leq \varphi_w(s, \tau_0) \leq f_{DH}(x; w).$$

Hence, $w \in M(x, \alpha)$. Therefore, $\mathbb{B}_\alpha(d) \subseteq M(x, \alpha)$.

Similarly, we can prove that $f^{DH}(x; d)$ is upper semicontinuous in d. $\square$

The following result provides an optimality condition for an unconstrained optimization problem.

Theorem 2.14. If $x \in \mathbb{R}^n$ is a local minimum of the function $f : \mathbb{R}^n \to \mathbb{R}$, then $f_{DH}(x; d) \geq 0$ for all $d \in \mathbb{R}^n$.

Proof. Since $x \in \mathbb{R}^n$ is a local minimum point of the function f, then for each $d \in \mathbb{R}^n$, each v in the neighborhood of d and each $t > 0$ sufficiently small, we have

$$\frac{f(x + tv) - f(x)}{t} \geq 0.$$

Thus,

$$f_{DH}(x; d) = \liminf_{\substack{v \to d \\ t \to 0^+}} \frac{f(x + tv) - f(x)}{t} \geq 0.$$

This completes the proof. $\square$

In view of relation (2.16) and Theorem 2.14, we have the following corollary.

Corollary 2.7. If $x \in \mathbb{R}^n$ is a local minimum of the function $f : \mathbb{R}^n \to \mathbb{R}$, then $f^D(x; d) \geq 0$ for all $d \in \mathbb{R}^n$.

For details on the calculus rules for Dini-Hadamard derivatives, we refer to Schirotzek [190].

2.5 Clarke and Other Types of Derivatives

In this section, we study different types of directional derivatives, including the Clarke directional derivative and Michel-Penot directional derivative.

Definition 2.7. Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz around $x \in \mathbb{R}^n$ and d be any other vector in $\mathbb{R}^n$.

(a) The *Clarke directional derivative* of f at x in the direction d is defined by

$$\begin{aligned} f^C(x; d) &= \limsup_{\substack{y \to x \\ t \to 0^+}} \frac{f(y + td) - f(y)}{t} \\ &= \inf_{\varepsilon, \delta > 0} \sup_{\substack{0 < t < \delta \\ 0 < \|y - x\| < \varepsilon}} \frac{f(y + td) - f(y)}{t}. \end{aligned}$$

(b) The *Michel-Penot directional derivative* of f at x in the direction d is defined by

$$f^{MP}(x; d) = \sup_{y \in \mathbb{R}^n} \limsup_{t \to 0^+} \frac{f(x + ty + td) - f(x + ty)}{t}.$$

Proposition 2.2. Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz around $x \in \mathbb{R}^n$ with Lipschitz constant $k > 0$. Then,

(a) the functions $d \mapsto f^C(x; d)$ and $d \mapsto f^{MP}(x; d)$ are finite, sublinear on $\mathbb{R}^n$, and satisfies

$$\left|f^{DH}(x; d)\right| \leq \left|f^{MP}(x; d)\right| \leq \left|f^C(x; d)\right| \leq k\|d\|; \tag{2.17}$$

(b) the functions $d \mapsto f^C(x; d)$ and $d \mapsto f^{MP}(x; d)$ are Lipschitz continuous with Lipschitz constant k;

(c) $f^C(x; d)$ is upper semicontinuous as a function of $(x; d)$;

(d) $f^C(x; -d) = (-f)^C(x; d)$ and $f^{MP}(x; -d) = (-f)^{MP}(x; d)$.

Proof. (a) To begin with we prove the last inequality in (2.17). Let us fix $d \in \mathbb{R}^n$. Since f is locally Lipschitz around x, we have

$$\frac{1}{t}\left|f(y+td)-f(y)\right| \leq \frac{1}{t}k\|td\| = k\|d\|,$$

whenever $\|y-x\|$ and $t>0$ are small. Hence $\left|f^C(x;d)\right| \leq k\|d\|$.

We next establish the second inequality in (2.17). Let us fix $d \in \mathbb{R}^n$ and let $\varepsilon > 0$ be given. For each $z \in \mathbb{R}^n$, there exists $\delta(z) > 0$ such that

$$\frac{1}{t}\left|f(x+td+tz)-f(x+tz)\right| < \left|f^C(x;d)\right| + \varepsilon, \quad \text{for all } t \in \;]0,\delta(z)[.$$

Therefore,

$$\limsup_{t\to 0^+}\frac{1}{t}\left|f(x+td+tz)-f(x+tz)\right| \leq \left|f^C(x;d)\right| + \varepsilon, \quad \text{for all } z \in \mathbb{R}^n,$$

and hence, $\left|f^{MP}(x;d)\right| \leq \left|f^C(x;d)\right| + \varepsilon$. By letting $\varepsilon \to 0^+$, we obtain the assertion.

Finally, we prove the first inequality in (2.17). Let us fix $d \in \mathbb{R}^n$ and let $\varepsilon > 0$ be given. For all sufficiently small $t > 0$ and all $z \in \mathbb{R}^n$ such that $\|d-z\|$ is sufficiently small, we have

$$\begin{aligned}
& \frac{1}{t}\left|f(x+tz)-f(x)\right| \\
\leq\; & \frac{1}{t}\left|f(x+td+t(z-d))-f(x+t(z-d))\right| + \frac{1}{t}\left|f(x+t(z-d))-f(x)\right| \\
\leq\; & \frac{1}{t}\left|f(x+td+t(z-d))-f(x+t(z-d))\right| + k\|z-d\| \\
\leq\; & \left|f^{MP}(x;d)\right| + \varepsilon + k\|z-d\|.
\end{aligned}$$

By letting $t \to 0^+$ and $z \to d$, and finally $\varepsilon \to 0^+$, we obtain the assertion.

We now give arguments to show that the function $d \mapsto f^C(x;d)$ is sublinear. By the definition of the Clarke directional derivative, it is clear that the function $d \mapsto f^C(x;d)$ is positively homogeneous as for all $\lambda > 0$. Thus we have

$$\begin{aligned}
f^C(x;\lambda d) &= \limsup_{\substack{y\to x\\ t\to 0^+}} \frac{f(y+t\lambda d)-f(y)}{t} \\
&= \lambda \limsup_{\substack{y\to x\\ t\lambda\to 0^+}} \frac{f(y+t\lambda d)-f(y)}{t\lambda} \\
&= \lambda f^C(x;d).
\end{aligned}$$

It remains to show that the function $d \mapsto f^C(x;d)$ is subadditive. Let

$d_1, d_2 \in \mathbb{R}^n$. We have

$$\begin{aligned} f^C(x; d_1 + d_2) &= \limsup_{\substack{y \to x \\ t \to 0^+}} \frac{f(y + t(d_1 + d_2)) - f(y)}{t} \\ &\leq \limsup_{\substack{y \to x \\ t \to 0^+}} \frac{f(y + t(d_1 + d_2)) - f(y + td_2)}{t} \\ &\quad + \limsup_{\substack{y \to x \\ t \to 0^+}} \frac{f(y + td_2) - f(y)}{t}. \end{aligned}$$

Let $\hat{y} := y + td_2$. Then $\hat{y} \to x$ as $y \to x$ and $t \to 0^+$, hence,

$$f^C(x; d_1 + d_2) \leq f^C(x; d_1) + f^C(x; d_2).$$

We prove that the function $d \mapsto f^{MP}(x; d)$ is sublinear. As in the case of the Clarke directional derivative, it is easy to see that the function $d \mapsto f^{MP}(x; d)$ is positively homogeneous. So, we only prove that the function $d \mapsto f^{MP}(x; d)$ is subadditive. Let $d_1, d_2 \in \mathbb{R}^n$ and let $\varepsilon > 0$ be given. For all $t > 0$ sufficiently small, we obtain

$$\frac{1}{t} \left(f(x + t(d_1 + d_2) + tz) - f(x + t(d_2 + z)) \right) \leq f^{MP}(x; d_1) + \frac{\varepsilon}{2}, \text{ for all } z \in \mathbb{R}^n,$$

and

$$\frac{1}{t} \left(f(x + td_2 + tz) - f(x + tz) \right) \leq f^{MP}(x; d_2) + \frac{\varepsilon}{2}, \text{ for all } z \in \mathbb{R}^n.$$

Adding these two inequalities, we get

$$\frac{1}{t} \left(f(x + t(d_1 + d_2) + tz) - f(x + tz) \right) \leq f^{MP}(x; d_1) + f^{MP}(x; d_2) + \varepsilon,$$

for all $z \in \mathbb{R}^n$. Therefore, $f^{MP}(x; d_1 + d_2) \leq f^{MP}(x; d_1) + f^{MP}(x; d_2)$.

(b) Let $d_1, d_2 \in \mathbb{R}^n$. If $t > 0$ is small and y is close to x, then

$$\begin{aligned} f(y + td_1) - f(y) &= (f(y + td_2) - f(y)) + (f(y + td_1) - f(y + td_2)) \\ &\leq (f(y + td_2) - f(y)) + kt\|d_1 - d_2\|, \end{aligned}$$

and therefore,

$$f^C(x; d_1) \leq f^C(x; d_2) + k\|d_1 - d_2\|.$$

By an analogous estimate with d_1 and d_2 interchanged, we obtain

$$\left| f^C(x; d_1) - f^C(x; d_2) \right| \leq k\|d_1 - d_2\|.$$

Therefore, the function $d \mapsto f^C(x; d)$ is Lipschitz continuous with Lipschitz constant k.

Analogously, we can easily verify that the function $d \mapsto f^{MP}(x; d)$ is Lipschitz continuous with Lipschitz constant k.

(c) Let $\{x_m\}$ and $\{d_m\}$ be sequences in $\mathbb{R}^n$ such that $x_m \to x \in \mathbb{R}^n$ and $d_m \to d \in \mathbb{R}^n$. By the definition of the Clarke directional derivative, for each m, there exist $z_m \in \mathbb{R}^n$ and $t_m > 0$ such that $\|z_m - x_m\| + t_m < \frac{1}{m}$ and

$$\begin{aligned} f^C(x_m; d_m) - \frac{1}{m} &\leq \frac{f(z_m + t_m d_m) - f(z_m)}{t_m} \\ &= \frac{f(z_m + t_m d) - f(z_m)}{t_m} + \frac{f(z_m + t_m d_m) - f(z_m + t_m d)}{t_m} \\ &\leq \frac{f(z_m + t_m d) - f(z_m)}{t_m} + k\|d_m - d\|, \end{aligned}$$

where $k > 0$ is the Lipschitz constant of f around x. By letting $m \to \infty$, the definition of upper limit gives

$$\limsup_{m \to \infty} f^C(x_m; d_m) \leq f^C(x; d).$$

Hence $f^C(x; d)$ is upper semicontinuous as a function of $(x; d)$.

(d) We have

$$\begin{aligned} f^C(x; -d) &= \limsup_{\substack{y \to x \\ t \to 0^+}} \frac{f(y - td) - f(y)}{t} \\ &= \limsup_{\substack{\bar{y} \to x \\ t \to 0^+}} \frac{(-f)(\bar{y} + td) - (-f)(\bar{y})}{t} \\ &= (-f)^C(x; d). \end{aligned}$$

In this connection, $\bar{y}$ stands for $y - td$.

For each $z \in \mathbb{R}^n$, we have

$$\begin{aligned} &\limsup_{t \to 0^+} \frac{f(x - td + tz) - f(x + tz)}{t} \\ = \; &\limsup_{t \to 0^+} \frac{(-f)(x + td + t(z - d)) - (-f)(x + t(z - d))}{t}. \end{aligned}$$

By taking the supremum over z and $z - d$, respectively, we obtain $f^{MP}(x; -d) = (-f)^{MP}(x; d)$. □

Remark 2.6. Since positive homogeneity and subadditivity imply convexity, it follows that $f^C(x; d)$ is a convex function of d.

Remark 2.7. It is clear from Proposition 2.2.2 in Clarke [48] that if $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz around x and has a Gâteaux derivative $f^G(x; d)$, then $f^G(x; d) \leq f^C(x; d)$ for all d.

Remark 2.8. In general, the existence of a Gâteaux derivative does not imply the existence of a finite Clarke directional derivative. Conversely, the existence of a Clarke directional derivative does not imply the existence of a Gâteaux derivative.

Theorem 2.15. If $f : \mathbb{R}^n \to \mathbb{R}$ is convex and locally Lipschitz around x, then $f^{MP}(x; d) = f^C(x; d) = f^{DH*}(x; d)$ for all $d \in \mathbb{R}^n$.

Proof. Let $\delta > 0$ be any fixed number. Then $f^C(x; d)$ can be written as

$$f^C(x; d) = \lim_{\varepsilon\to 0^+} \sup_{\substack{\|z-x\|<\varepsilon\delta \\ 0<t<\varepsilon}} \frac{f(z + td) - f(z)}{t}.$$

Since f is convex, by Theorem 2.1, the function

$$t \mapsto \frac{f(z + td) - f(z)}{t}$$

is nondecreasing, hence

$$f^C(x; d) = \lim_{\varepsilon\to 0^+} \sup_{\|z-x\|<\varepsilon\delta} \frac{f(z + \varepsilon d) - f(z)}{\varepsilon}. \tag{2.18}$$

By Lipschitz condition, for any $z \in \mathbb{B}_{\varepsilon\delta}(x)$, we have

$$\begin{aligned} & \left| \frac{f(z + \varepsilon d) - f(z)}{\varepsilon} - \frac{f(x + \varepsilon d) - f(x)}{\varepsilon} \right| \\ \leq\ & \frac{1}{\varepsilon} \left(|f(z + \varepsilon d) - f(x + \varepsilon d)| + |f(x) - f(z)| \right) \\ \leq\ & 2\delta k, \end{aligned}$$

where k is the Lipschitz constant. Therefore,

$$\begin{aligned} f^C(x; d) &\leq \lim_{\varepsilon\to 0^+} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon} + 2\delta k \\ &= f'_+(x; d) + 2\delta k = f^{DH*}(x; d) + 2\delta k, \end{aligned}$$

where the last equality holds because of Theorem 2.12. Since δ was arbitrary, it follows that $f^C(x; d) \leq f^{DH*}(x; d)$. Since $f^{DH*}(x; d) \leq f^C(x; d)$ always holds, we obtain $f^C(x; d) = f^{DH*}(x; d)$.

Now we prove that $f^{MP}(x; d) = f^{DH*}(x; d)$ for all $d \in \mathbb{R}^n$. Indeed, $f^D(x; d) = f_D(x; d) = f'_+(x; d)$ as f is a convex function defined on $\mathbb{R}^n$, and for all $d \in \mathbb{R}^n$,

$$\begin{aligned} f^{MP}(x; d) &= \sup_{y\in\mathbb{R}^n} \limsup_{t\to 0^+} \frac{f(x + ty + td) - f(x + ty)}{t} \\ &\leq \sup_{y\in\mathbb{R}^n} \left(\limsup_{t\to 0^+} \frac{f(x + ty + td) - f(x)}{t} \right. \\ &\qquad \left. + \limsup_{t\to 0^+} \frac{f(x) - f(x + ty)}{t} \right) \\ &= \sup_{y\in\mathbb{R}^n} \left(f'_+(x; y + d) - f'_+(x; y) \right) \\ &\leq f'_+(x; d), \end{aligned}$$

where the last inequality holds because $f'_+(x;\cdot)$ is subadditive. Since $f'_+(x;d) \leq f^{MP}(x;d)$ and in view of Theorem 2.12, we have $f^{DH*}(x;d) = f'_+(x;d) = f^{MP}(x;d)$. □

The following simple example further clarifies the relationship among Dini, Clarke, and Gâteaux directional derivatives.

Example 2.7. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} ax, & \text{if } x \geq 0, \\ bx, & \text{if } x < 0, \end{cases}$$

where $a, b \in \mathbb{R}$. Then, we can easily calculate that

$$f^G(0;d) = f_D(0;d) = f^D(0;d) = \begin{cases} ad, & \text{if } d \geq 0, \\ bd, & \text{if } d < 0, \end{cases}$$

and

$$f^C(0;d) = \begin{cases} \max\{ad, bd\}, & \text{if } d \geq 0, \\ \min\{ad, bd\}, & \text{if } d < 0. \end{cases}$$

Besides the directional derivatives studied so far there are many more in literature. We study here yet another generalized directional derivative defined as

$$f^*(x;d) = \limsup_{\substack{t \to 0^+ \\ v \to d \\ y \to x}} \frac{f(y+tv) - f(y)}{t},$$

and

$$f_*(x;d) = \liminf_{\substack{t \to 0^+ \\ v \to d \\ y \to x}} \frac{f(y+tv) - f(y)}{t}.$$

Theorem 2.16. If $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz around x with Lipschitz constant k, then

$$f^*(x;d) = \limsup_{\substack{t \to 0^+ \\ y \to x}} \frac{f(y+td) - f(y)}{t} = f^C(x;d),$$

and

$$f_*(x;d) = \liminf_{\substack{t \to 0^+ \\ y \to x}} \frac{f(y+td) - f(y)}{t}.$$

Proof. Since

$$\begin{aligned} \left| \frac{f(y+tv) - f(y)}{t} - \frac{f(y+td) - f(y)}{t} \right| &= \left| \frac{f(y+tv) - f(y+td)}{t} \right| \\ &\leq k\|v - d\| \to 0 \end{aligned}$$

as $y \to x$, $v \to d$, and $t \to 0^+$. We get the result. □

Definition 2.8. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, $x \in \mathbb{R}^n$, and d be a vector in $\mathbb{R}^n$. The *Rockafellar upper subderivative* of f at x in the direction d is defined by

$$\begin{aligned} f^R(x;d) &= \limsup_{\substack{t\to 0+\\ y\to x}} \inf_{v\to d} \frac{f(y+tv)-f(y)}{t} \\ &= \sup_{\varepsilon>0} \limsup_{\substack{t\to 0+\\ y\to x}} \inf_{\|v-d\|<\varepsilon} \frac{f(y+tv)-f(y)}{t}. \end{aligned}$$

Theorem 2.17. The directional derivatives $f^*(x;d)$ and $f^R(x;d)$ are convex in d.

Proof. We first prove that the function $d \mapsto f^*(x;d)$ is convex. For that, let $d_1, d_2 \in \mathbb{R}^n$ and $\{x_m\} \in \mathbb{R}^n$, $\{t_m\} \in \mathbb{R}_+$, $\{v_m\} \in \mathbb{R}^n$ be sequences with $x_m \to x$, $t_m \to 0$, $v_m \to d_1 + d_2$ such that

$$\limsup_{m\to\infty} \frac{f(x_m+t_m v_m)-f(x_m)}{t_m} = f^*(x; d_1+d_2).$$

Then $\{v_m\}$ can be decomposed in two sequences $\{u_m\}$ and $\{w_m\}$ such that $u_m + w_m = v_m$ for all m and $u_m \to d_1$, $w_m \to d_2$. We obtain

$$\begin{aligned} f^*(x;d_1+d_2) &= \limsup_{m\to\infty} \frac{f(x_m+t_m(u_m+w_m))-f(x_m)}{t_m} \\ &\leq \limsup_{m\to\infty} \frac{f(x_m+t_m w_m+t_m u_m)-f(x_m+t_m w_m)}{t_m} \\ &\quad + \limsup_{m\to\infty} \frac{f(x_m+t_m w_m)-f(x_m)}{t_m} \\ &\leq f^*(x;d_1)+f^*(x;d_2). \end{aligned}$$

Hence $f^*(x;\cdot)$ is subadditive. It is easy to see that $f^*(x;\cdot)$ is positively homogeneous. Hence, $f^*(x;\cdot)$ is convex.

Finally, we prove that the function $d \mapsto f^R(x;d)$ is convex. For that, let $d_1, d_2 \in \mathbb{R}^n$, $\varepsilon > 0$ and $\{x_m\} \in \mathbb{R}^n$, $\{t_m\} \in \mathbb{R}$ with $t_m > 0$ be sequences with $x_m \to x$ and $t_m \to 0^+$. By the definition of $f^R(x;d_2)$, there exists a sequence $\{w_m\} \in \mathbb{B}_{\varepsilon/2}(d_2)$ such that

$$\limsup_{m\to\infty} \frac{f(x_m+t_m w_m)-f(x_m)}{t_m} \leq f^R(x;d_2).$$

Since $x_m + t_m w_m \to x$, by the definition of $f^R(x;d_1)$, we can find a sequence $\{u_m\} \in \mathbb{B}_{\varepsilon/2}(d_1)$ such that

$$\limsup_{m\to\infty} \frac{f(x_m+t_m w_m+t_m u_m)-f(x_m+t_m w_m)}{t_m} \leq f^R(x;d_1).$$

Thus, with $d_m = u_m + w_m$, $m \in \mathbb{N}$, we have found a sequence $\{d_m\} \in \mathbb{B}_{\varepsilon/2}(d_1 + d_2)$ with

$$\begin{aligned}
& \limsup_{m\to\infty} \frac{f(x_m + t_m d_m) - f(x_m)}{t_m} \\
\leq \; & \limsup_{m\to\infty} \frac{f(x_m + t_m w_m + t_m u_m) - f(x_m + t_m w_m)}{t_m} \\
& + \limsup_{m\to\infty} \frac{f(x_m + t_m w_m) - f(x_m)}{t_m} \\
\leq \; & f^R(x; d_1) + f^R(x; d_2).
\end{aligned}$$

Since the above inequality is true for each $\varepsilon > 0$, each sequence $x_m \to x$ and each sequence $t_m \to 0^+$, we get

$$\begin{aligned}
f^R(x; d_1 + d_2) &= \sup_{\varepsilon>0} \limsup_{\substack{t\to 0^+ \\ y\to x \\ v\to d_1+d_2}} \inf_{\|v-d_1-d_2\|<\varepsilon} \frac{f(y+tv) - f(y)}{t} \\
&\leq f^R(x; d_1) + f^R(x; d_2).
\end{aligned}$$

Hence, $f^R(x; \cdot)$ is subadditive. It is easy to see that $f^R(x; \cdot)$ is positively homogeneous. Hence, $f^R(x; \cdot)$ is convex. □

Several other kinds of derivatives are defined in the literature [82, 83, 86, 103, 104, 116, 160, 183, 184, 185, 190].

2.6 Dini and Clarke Subdifferentials

The Clarke subdifferential theory was a breakthrough in convex analysis. It extended the subdifferential theory of convex functions to nonconvex functions and found applications in several fields. The Michel-Penot subdifferential is a subclass of the Clarke subdifferential and is the same as the Clarke subdifferential at regular points. These subdifferentials coincide with the usual subdifferential if the function is convex.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz around $x \in \mathbb{R}^n$. The Clarke and Michel-Penot subdifferentials of f at x are denoted by $\partial^C f(x)$ and $\partial^{MP} f(x)$, respectively, and are defined by

$$\partial^C f(x) = \left\{ \xi \in \mathbb{R}^n : \langle \xi, d \rangle \leq f^C(x; d), \text{ for all } d \in \mathbb{R}^n \right\},$$

and

$$\partial^{MP} f(x) = \left\{ \xi \in \mathbb{R}^n : \langle \xi, d \rangle \leq f^{MP}(x; d), \text{ for all } d \in \mathbb{R}^n \right\}.$$

Analogously, we define the *Dini lower subdifferential* $\partial_D f(x)$ and *Dini upper subdifferential* $\partial^D f(x)$ of f at x as follows:

$$\partial_D f(x) = \{\xi \in \mathbb{R}^n : \langle \xi, d\rangle \leq f_D(x;d), \text{ for all } d \in \mathbb{R}^n\},$$

and

$$\partial^D f(x) = \left\{\xi \in \mathbb{R}^n : \langle \xi, d\rangle \leq f^D(x;d), \text{ for all } d \in \mathbb{R}^n\right\}.$$

The elements of the respective subdifferentials are called *subgradients*.

Theorem 2.18. Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz around x with Lipschitz constant k. The following assertions hold.

(a) The subdifferentials $\partial^C f(x)$ and $\partial^{MP} f(x)$ are nonempty, convex, and compact, and satisfy

$$\partial^{MP} f(x) \subseteq \partial^C f(x) \subseteq \mathbb{B}_k[\mathbf{0}],$$

that is, for all $\xi \in \partial^{MP} f(x)$ or $\xi \in \partial^C f(x)$, $\|\xi\| \leq k$.

(b) For any $\lambda \in \mathbb{R}$, $\partial^C(\lambda f)(x) = \lambda \partial^C f(x)$ and $\partial^{MP}(\lambda f)(x) = \lambda \partial^{MP} f(x)$.

(c) For every $d \in \mathbb{R}^n$, we have

$$f^C(x;d) = \max\left\{\langle \xi, d\rangle : \xi \in \partial^C f(x)\right\},$$

$$f^{MP}(x;d) = \max\left\{\langle \xi, d\rangle : \xi \in \partial^{MP} f(x)\right\}.$$

Proof. (a) From the definitions of the Clarke derivative and Michel-Penot derivative, it follows that $f^{MP}(x;d) \leq f^C(x;d)$, and thus, we get $\partial^{MP} f(x) \subseteq \partial^C f(x)$.

We prove the last part. Since f is locally Lipschitz around x, there exists $\varepsilon > 0$ such that

$$|f(y) - f(z)| \leq k\|y - z\|, \quad \text{for all } y, z \in \mathbb{B}_\varepsilon(x).$$

Thus, if $y \in \mathbb{B}_\varepsilon(x)$ and $\xi \in \partial^C f(x)$, then $\langle \xi, z - y\rangle \leq f(z) - f(y) \leq k\|z - y\|$, and therefore, $\|\xi\| \leq k$ for all $\xi \in \partial^C f(x)$ and all $y \in \mathbb{B}_\varepsilon(x)$.

(b) Let $\lambda \geq 0$. Since $(\lambda f)^C(x;\cdot) = \lambda f^C(x;\cdot)$, we get $\partial^C(\lambda f)(x) = \lambda \partial^C f(x)$.

If we verify that the formula holds for $\lambda = -1$, then the formula is true for all $\lambda \in \mathbb{R}$. So, we let $\lambda = -1$. Then,

$$\begin{aligned} \partial^C(-f)(x) &= \left\{\xi \in \mathbb{R}^n : \langle \xi, d\rangle \leq (-f)^C(x;d) \text{ for all } d \in \mathbb{R}^n\right\} \\ &= \left\{\xi \in \mathbb{R}^n : \langle \xi, -z\rangle \leq f^C(x;z) \text{ for all } z \in \mathbb{R}^n\right\} \\ &= -\partial^C f(x). \end{aligned}$$

Similarly, we can prove that $\partial^{MP}(\lambda f)(x) = \lambda \partial^{MP} f(x)$.

For the proof of part (c), we refer to Schirotzek [190]. □

Analogous to Theorem 2.18(c), we have the following result.

Theorem 2.19. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. For every $d \in \mathbb{R}^n$, we have

$$f_D(x; d) = \max\{\langle \xi, d\rangle : \xi \in \partial_D f(x)\},$$

$$f^D(x; d) = \max\left\{\langle \xi, d\rangle : \xi \in \partial^D f(x)\right\}.$$

Remark 2.9. We mention that, besides the lower semicontinuity, the sublinearity of $f^C(x;\cdot)$ and $f^{MP}(x;\cdot)$ ensures the nonemptiness of the respective subdifferential. However, the Dini lower and Dini upper subdifferentials are compact and convex but may be empty.

Theorem 2.20. Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz around x. If x is a local minimum or a local maximum of f, then $\mathbf{0} \in \partial^C f(x)$ and $\mathbf{0} \in \partial^{MP} f(x)$.

Proof. Let x be a local minimum of f. Then for all $d \in \mathbb{R}^n$, we have

$$\begin{aligned} 0 &\leq \liminf_{t\to 0^+} \frac{f(x+td) - f(x)}{t} \\ &\leq \limsup_{t\to 0^+} \frac{f(x+td) - f(x)}{t} \\ &\leq f^C(x; d). \end{aligned}$$

By the definition of the Clarke subdifferential, we obtain $0 \in \partial^C f(x)$. If x is a local maximizer of f, then x is a local minimizer of $-f$, and so, $0 \in \partial^C(-f)(x) = -\partial^C f(x)$.

Similarly, we can prove the result for the Michel-Penot subdifferential. ☐

Remark 2.10. By Corollary 2.7, if a point $x \in \mathbb{R}^n$ is a local minimum of the function $f : \mathbb{R}^n \to \mathbb{R}$, then $f_D(x; d) \geq 0$ and $f^D(x; d) \geq 0$ for any direction d, and hence, the necessary optimality conditions

$$\mathbf{0} \in \partial_D f(x) \subseteq \partial^D f(x) \subseteq \partial^{MP} f(x) \subseteq \partial^C f(x)$$

hold.

Remark 2.11. If $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and locally Lipschitz around x, then from Theorem 2.15, we have

$$\partial f(x) = \partial^D f(x) = \partial^{MP} f(x) = \partial^C f(x).$$

Let us recall the following definitions of upper and lower semicontinuous set-valued maps. These are studied in detail in Appendix A.

Definition 2.9. A set-valued map $\Phi : \mathbb{R}^n \to \mathbb{R}^\ell$ is said to be

(a) *upper semicontinuous* at $x \in \mathrm{Dom}(\Phi)$ if for each open set V in $\mathbb{R}^\ell$ containing $\Phi(x)$, there exists an open neighborhood U of x such that $\Phi(U) \subseteq V$;

(b) *lower semicontinuous* at $x \in \mathrm{Dom}(\Phi)$ if for each open set V in $\mathbb{R}^\ell$ such that $V \cap \Phi(x) \neq \emptyset$, there exists an open neighborhood U of x such that $\Phi(y) \cap V \neq \emptyset$ for all $y \in U$.

Theorem 2.21. Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz around x with Lipschitz constant k. Then the following assertions hold.

(a) Let $\{x_m\}$ and $\{\xi_m\}$ be sequences in $\mathbb{R}^n$ such that $\xi_m \in \partial^C f(x_m)$ for all $m \in \mathbb{N}$. If $x_m \to x \in \mathbb{R}^n$ and $\xi_m \to \xi \in \mathbb{R}^n$, then $\xi \in \partial^C f(x)$, that is, the set-valued map $\partial^C f$ is closed.

(b) The Clarke subdifferential mapping $\partial^C f : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ is upper semicontinuous.

Proof. (a) Let $d \in \mathbb{R}^n$ be given. Then there is a subsequence of $\{\langle \xi_m, d\rangle\}$, again denoted by $\{\langle \xi_m, d\rangle\}$, which converges to $\{\langle \xi, d\rangle\}$ as $m \to \infty$. By the definition of the Clarke subdifferential, we have $\langle \xi_m, d\rangle \leq f^C(x_m; d)$ for all $m \in \mathbb{N}$. Letting $m \to \infty$ and by Proposition 2.2(c), we get $\langle \xi, d\rangle \leq f^C(x; d)$. Since d was arbitrary, we obtain $\xi \in \partial^C f(x)$.

(b) Let $x \in \mathrm{int}(\mathrm{dom}(f))$ and V be an open subset of $\mathbb{R}^n$ containing $\partial^C f(x)$. It is sufficient to show that for any sequence $\{x_m\}$ in $\mathrm{int}(\mathrm{dom}(f))$ with $x_m \to x$ as $m \to \infty$, we have $\partial^C f(x_m) \subseteq V$ for all sufficiently large m. Assume that it is not true. Then for some subsequence $\{x_m\}$, again denoted by $\{x_m\}$, we can find $\xi_m \in \partial^C f(x_m) \setminus V$. By the last part of Theorem 2.18(a), there exists $r > 0$ such that $\partial^C f(x_m) \subseteq \mathbb{B}_r[\mathbf{0}]$ for sufficiently large m. Since $\mathbb{B}_r[\mathbf{0}]$ is compact, the sequence $\{\xi_m\}$ has a limit point ξ. Since $\xi_m \in \partial^C f(x_m) \setminus V$, we conclude that $\xi \in \partial^C f(x) \setminus V$, which contradicts the fact $\partial^C f(x) \subseteq V$. This completes the proof. □

Theorem 2.22 (Sum Rule)**.** Let $g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz around x. Then,

$$\partial^C \left(\sum_{i=1}^m g_i \right)(x) \subseteq \sum_{i=1}^m \partial^C g_i(x) \quad \text{and} \quad \partial^{MP} \left(\sum_{i=1}^m g_i \right)(x) \subseteq \sum_{i=1}^m \partial^{MP} g_i(x). \tag{2.19}$$

Theorem 2.23 (Lebourg's Mean Value Theorem)**.** Let x and y be points in $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz on an open set containing the line segment $[x, y] \subseteq \mathbb{R}^n$. Then there exists a point $z \in\ ]x, y[$ such that

$$f(y) - f(x) \in \langle \partial^C f(z), y - x\rangle = \bigcup_{\xi \in \partial^C f(z)} \{\langle \xi, y - x\rangle\}.$$

Proof. Consider the functions $\psi, \varphi : [0, 1] \to \mathbb{R}$ defined by

$$\psi(\lambda) = f(x + \lambda(y - x)) + \lambda(f(x) - f(y)), \quad \text{for all } \lambda \in [0, 1],$$

and

$$\varphi(\lambda) = f(x + \lambda(y - x)), \quad \text{for all } \lambda \in [0, 1].$$

Clearly, $\psi(0) = \psi(1) = f(x)$ and ψ is continuous, and therefore, ψ attains a local minimum or local maximum at some point $\lambda_0 \in\ ]0,1[$. By Theorem 2.20, $\mathbf{0} \in \partial^C \psi(\lambda_0)$. From Theorem 2.18 and Theorem 2.22, we conclude that

$$\mathbf{0} \in \partial^C \psi(\lambda_0) \subseteq \partial^C \varphi(\lambda_0) + (f(x) - f(y)). \tag{2.20}$$

We show that

$$\partial^C \varphi(\lambda) \subseteq \langle \partial^C f(x + \lambda(y-x)), y - x \rangle, \quad \text{for all } \lambda \in (0,1). \tag{2.21}$$

Since φ is Lipschitz continuous on $]0,1[$, $\partial^C \varphi$ makes sense. For all $\lambda \in\ ]0,1[$, the sets $\partial^C \varphi(\lambda)$ and $\langle \partial^C f(x + \lambda(y-x)), y - x\rangle$ are closed convex subsets of $\mathbb{R}$, and hence, are intervals. Therefore, it is sufficient to prove that

$$\max\left(a\partial^C \varphi(\lambda)\right) \leq \max\left(a\langle \partial^C f(x + \lambda(y-x)), y - x\rangle\right), \quad \text{for } a = \pm 1.$$

Indeed,

$$\begin{aligned}
\max\left(a\partial^C \varphi(\lambda)\right) &= \varphi^C(\lambda; a) = \limsup_{\substack{\hat{\lambda} \to \lambda \\ t \to 0^+}} \frac{\varphi(\hat{\lambda} + ta) - \varphi(\hat{\lambda})}{t} \\
&= \limsup_{\substack{\hat{\lambda} \to \lambda \\ t \to 0^+}} \frac{f(x + (\hat{\lambda} + ta)(y-x)) - f(x + \hat{\lambda}(y-x))}{t} \\
&= \limsup_{\substack{z \to x_\lambda \\ t \to 0^+}} \frac{f(z + ta(y-x)) - f(z)}{t}, \\
&\qquad \text{where } z = x + \hat{\lambda}(y-x) \text{ and } x_\lambda = x + \lambda(y-x) \\
&= f^C(x_\lambda; a(y-x)) = \max\left(\langle \partial^C f(x + \lambda(y-x)), a(y-x)\rangle\right).
\end{aligned}$$

Finally, from (2.20) and (2.21), we get the conclusion on setting $z := x_{\lambda_0}$. □

We conclude this chapter with a mean value theorem in terms of the Dini upper subdifferential.

Theorem 2.24. Let $K \subseteq \mathbb{R}^n$ and $f : K \to \mathbb{R} \cup \{\pm\infty\}$ be finite and upper semicontinuous on an open set containing the segment $[a,b]$ of K. If $\partial^D f(x)$ is an upper semicontinuous set-valued map with nonempty equicontinuous values on $[a,b]$, then there exist $c \in [a,b]$ and $\xi \in \partial^D f(c)$ such that

$$f(b) - f(a) = \langle \xi, b - a\rangle.$$

The assumption that $\partial^D f(x)$ is nonempty and equicontinuous is satisfied when $f^D(x; \cdot)$ is convex and continuous.

For the proof of this theorem, we refer to Penot [174, Corollary 2.4].

Chapter 3

Nonsmooth Convexity

3.1 Introduction

As we have seen in Chapter 1, differentiable convex functions and generalized convex functions can be characterized by means of the gradient function. However, when the function fails to be differentiable, the convexity may be given in terms of a generalized derivative (if it exists). For instance, if $f : \mathbb{R}^n \to \mathbb{R}$ is directionally differentiable, then f is convex if and only if

$$f'(x; y - x) \leq f(y) - f(x), \quad \text{for all } x, y \in \mathbb{R}^n.$$

Various types of generalized derivatives exist in literature and some of them have been studied in Chapter 2. Most of the generalized derivatives, like the directional derivative, the Dini derivative, and the Clarke derivative, share a very important property, namely, positive homogeneity as a function of the direction. Motivated by this fact an attempt was made to unify the generalized derivatives by considering a bifunction $h(x; d)$ with values in $\mathbb{R} \cup \{\pm\infty\}$, where x refers to a point in the domain of the function and d is a direction in $\mathbb{R}^n$. Komlósi [132] defined generalized convexity in terms of this bifunction $h(x; d)$. This definition encompassed most of the existing definitions involving the generalized derivatives.

3.2 Nonsmooth Convexity in Terms of Bifunctions

A natural idea for generalizing convexity of differentiable functions consists of replacing the derivative by a generalized derivative, so that nondifferentiable functions can be considered.

Definition 3.1. Let K be a nonempty convex subset of $\mathbb{R}^n$ and let $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. A function $f : K \to \mathbb{R}$ is said to be

(a) *h-convex* if for all $x, y \in K$, $x \neq y$

$$h(x; y - x) \leq f(y) - f(x); \tag{3.1}$$

(b) *strictly h-convex* if strict inequality holds in (3.1) for all $x \neq y$;

(c) *strongly h-convex with modulus* ρ if there exists a real number $\rho > 0$ such that for all $x, y \in K$, $x \neq y$

$$h(x; y - x) + \rho\|y - x\|^2 \leq f(y) - f(x). \tag{3.2}$$

A function $f : K \to \mathbb{R}$ is called (*strictly, strongly*) *h-concave* if $-f$ is (strictly, strongly) $-h$-convex.

From the above definitions, it is clear that

$$\text{strong } h\text{-convexity} \Rightarrow \text{strict } h\text{-convexity} \Rightarrow h\text{-convexity},$$

and the reverse implications may not necessarily hold.

The next example shows that a strictly h-convex function may not be strongly h-convex.

Example 3.1. Consider the function $f : [0, 1] \to \mathbb{R}$ defined as

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ x^4, & \text{if } 0 < x \leq 1. \end{cases}$$

If we define $h : [0, 1] \times \mathbb{R}$ to $\mathbb{R} \cup \{0, \pm\infty\}$

$$h(x; d) = \begin{cases} -1, & \text{if } x = 0,\ d > 0, \\ 4x^3 d, & \text{otherwise}, \end{cases}$$

then it can easily be verified that f is strictly h-convex but not strongly h-convex because for $x = 0$, the inequality (3.2) holds for all $y \in\ ]0, 1]$ only if $\rho = 0$.

The following example exhibits that the class of strictly h-convex functions is properly contained in the class of h-convex functions.

Example 3.2. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

If we define $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$h(x; d) = \begin{cases} -2, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0, \end{cases}$$

then it can easily be verified that f is h-convex but not strictly h-convex.

The following is an example of a strongly h-convex function, which is not convex.

Example 3.3. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} x^2, & \text{if } x \neq 0, \\ 1, & \text{if } x = 0. \end{cases}$$

Then,

$$f^D(x; d) = \begin{cases} 2xd, & \text{if } x \neq 0, \\ -\infty, & \text{if } x = 0. \end{cases}$$

If we choose $h(x; d) = f^D(x; d)$, we find that f is strongly h-convex with inequality (3.2) holding for $\rho = 1$. Also, we note that f is a discontinuous nonconvex function.

If K is an open convex set, then every differentiable convex function is h-convex, where $h(x; y-x) = \langle \nabla f(x), y-x \rangle$ for all $x, y \in K$. However, we show next that in general a convex function is h-convex provided h is majorized by the Dini upper directional derivative.

Theorem 3.1. Let K be a nonempty convex subset of $\mathbb{R}^n$. If $f : K \to \mathbb{R}$ is (strictly, strongly) convex and h is majorized by the Dini upper directional derivative of f, that is,

$$h(x; d) \leq f^D(x; d), \quad \text{for all } x \in K \text{ and all } d \in \mathbb{R}^n, \tag{3.3}$$

then f is (strictly, strongly) h-convex. Conversely, if f is (strictly, strongly) h-convex and h is both subodd and positively homogeneous in the second argument, then f is (strictly, strongly) convex.

Proof. Since f is convex, we have that for all $x, y \in K$, $x \neq y$, $\lambda \in \,]0, 1]$,

$$f(x + \lambda(y - x)) \leq (1 - \lambda) f(x) + \lambda f(y),$$

thereby implying that

$$f(y) - f(x) \geq f^D(x; y - x),$$

and hence by using (3.3) it follows that f is h-convex.

Conversely, let $x, y \in K$, $x \neq y$, and $w = x + \lambda(y - x)$ for $\lambda \in \,]0, 1[$. By h-convexity of f, we have

$$f(x) - f(w) \geq h(w; x - w) = \lambda h(w; x - y), \tag{3.4}$$

and

$$f(y) - f(w) \geq h(w; y - w) = (1 - \lambda) h(w; y - x). \tag{3.5}$$

Multiplying (3.4) by $1 - \lambda$ and (3.5) by λ and then adding the resultants, we get

$$(1 - \lambda) f(x) + \lambda f(y) - f(w) \geq \lambda(1 - \lambda) \left(h(w; y - x) + h(w; x - y) \right),$$

which by suboddness of h in the second argument implies that f is convex.

The proof for the strongly convex case follows on the same lines and for the strictly convex case the arguments follow as in Theorem 1.10(b) for the differentiable case. ☐

The condition (3.3) in Theorem 3.1 cannot be relaxed, which is illustrated through the following example.

Example 3.4. The function f defined on $\mathbb{R}$ by $f(x) = |x|$ is a convex function. Then,

$$f^D(x;d) = \begin{cases} -d, & \text{if } x < 0, \\ |d|, & \text{if } x = 0, \\ d, & \text{if } x > 0. \end{cases}$$

If h is a bifunction defined on $\mathbb{R} \times \mathbb{R}$ as $h(x;d) = |d|$, then condition (3.3) fails to hold for $x \neq 0$. Clearly, f is not h-convex because for $x = 1$ and $y = 0$,

$$h(x; y - x) > f(y) - f(x).$$

Remark 3.1. The converse implication in Theorem 3.1 fails to hold if h is not subodd in the second argument. The h-convex function f considered in Example 3.2 is not convex and h is not subodd in the second argument.

The following theorem illustrates that (strict, strong) convexity with respect to a bifunction h is preserved with respect to any other bifunction minorizing it.

Theorem 3.2. Let K be a nonempty convex subset of $\mathbb{R}^n$. If $g, h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ are two bifunctions such that

$$g(x;d) \leq h(x;d), \quad \text{for all } x \in K \text{ and all } d \in \mathbb{R}^n,$$

then every (strictly, strongly) h-convex function is also (strictly, strongly) g-convex.

Proof. Follows trivially from the definition. □

We recall that a direction $d \in \mathbb{R}^n$ is said to be a *feasible direction* of K at $x \in K$ if there exists $\sigma > 0$ such that $x + td \in K$ for $0 \leq t < \sigma$. Giorgi and Komlósi [88] referred to a point as an inf-stationary point if $f_D(x;d) \geq 0$ for every feasible direction d.

We extend the notion of inf-stationary point and sup-stationary point using the bifunction h.

Definition 3.2. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. A point $x \in K$ is said to be an *inf-stationary point* of f with respect to h if $h(x;d) \geq 0$ for every feasible direction d of K at x, whereas x is said to be a *sup-stationary point* of f with respect to h if $h(x;d) \leq 0$ for every feasible direction d of K at x.

Analogous to Theorem 1.12 for the differentiable case, we have the corresponding result for nondifferentiable functions in terms of inf-stationary points.

Theorem 3.3. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. If f is (strictly) h-convex function, then every inf-stationary point of f with respect to h is a (unique) global minimum of f over K.

Proof. The proof follows by using the definition of (strict) h-convexity. □

Remark 3.2. The converse of Theorem 3.3 may not necessarily hold. If we again consider the function f given in Example 3.2, then by Theorem 3.2, f is h-convex where

$$h(x; d) = \begin{cases} -2, & \text{if } x = 0, \\ -1, & \text{if } x \neq 0. \end{cases}$$

We observe that none of the global minimum is an inf-stationary point of f.

To establish analogous results for the concave case we assume that h is minorized by the Dini lower directional derivative of f, that is,

$$f_D(x; d) \leq h(x; d), \quad \text{for all } x \in K \text{ and } d \in \mathbb{R}^n. \tag{3.6}$$

Remark 3.3. The conditions (3.3) and (3.6) are satisfied by several known generalized directional derivatives, some of which we have already studied in Chapter 2. For instance, the Dini lower directional derivative, the Dini-Hadamard lower directional derivative, the *upper weak Rockafellar derivative* given by

$$f^{wR}(x; d) = \limsup_{t \to 0^+} \inf_{u \to d} \frac{f(x + tu) - f(x)}{t},$$

the upper derivative or the *adjacent derivative* given as

$$f^A(x; d) = \sup_{U \in N(d)} \limsup_{t \to 0^+} \inf_{u \in U} \frac{f(x + tu) - f(x)}{t},$$

where $N(d)$ denotes the family of neighborhoods of d, are some of the derivatives satisfying inequality (3.3).

Condition (3.6) is satisfied by the Clarke generalized directional derivative, the Dini upper directional derivative, the Dini-Hadamard upper directional derivative, the Michel-Penot directional derivative, and the lower weak Rockafellar derivative defined as

$$f_{wR}(x; d) = \liminf_{t \to 0^+} \sup_{u \to d} \frac{f(x + tu) - f(x)}{t}.$$

We have the following theorems for the concave case.

Theorem 3.4. Let K be a nonempty convex subset of $\mathbb{R}^n$. If $f : K \to \mathbb{R}^n$ is (strictly, strongly) concave and h satisfies (3.6), then f is (strictly, strongly) h-concave. Conversely, if f is (strictly, strongly) h-concave and $-h$ is both subodd and positively homogeneous in the second argument, then f is (strictly, strongly) concave.

Theorem 3.5. Let K be a nonempty convex subset of $\mathbb{R}^n$. If $g, h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ are two bifunctions such that

$$g(x; d) \geq h(x; d), \quad \text{for all } x \in K \text{ and all } d \in \mathbb{R}^n,$$

then every (strictly, strongly) h-concave function is also (strictly, strongly) g-concave.

Theorem 3.6. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. If f is (strictly) h-concave function, then every sup-stationary point of f with respect to h is a (unique) global maximum of f over K.

3.3 Generalized Nonsmooth Convexity in Terms of Bifunctions

Several existing notions of generalized convexity for nondifferentiable functions employ varied kinds of generalized derivatives. In 1981, Diewert [69] initiated the study of pseudoconvexity for arbitrary functions by employing the Dini directional derivatives, while Giorgi and Komlósi [88] studied quasiconvexity using the same directional derivatives. The notions of quasiconvexity and pseudoconvexity for locally Lipschitz functions using the Clarke directional derivative were introduced by Glover [91] and Hiriart-Urruty [103, 104], respectively.

The notion of quasiconvexity via Dini directional derivatives was presented by Giorgi and Komlósi [88] as follows.

A function $f : K \to \mathbb{R}$ defined on a convex set $C \subseteq \mathbb{R}^n$ is *upper-Dini quasiconvex* if for all $x, y \in K$,

$$f(y) \leq f(x) \quad \Rightarrow \quad f^D(x; y - x) \leq 0,$$

and *lower-Dini quasiconvex* if for all $x, y \in K$,

$$f(y) \leq f(x) \quad \Rightarrow \quad f_D(x; y - x) \leq 0.$$

Clearly, every quasiconvex function is upper-Dini quasiconvex and, consequently, lower-Dini quasiconvex. On the other hand, an upper-Dini quasiconvex function may not be quasiconvex. For example, the function considered in Example 3.2 is both upper-Dini and lower-Dini quasiconvex but not quasiconvex. For this function

$$f^D(x; d) = f_D(x; d) = \begin{cases} -\infty, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

Giorgi and Komlósi [88] established that if f is radially continuous on a convex set K, then the notions of quasiconvexity, upper-Dini quasiconvexity, and lower-Dini quasiconvexity are equivalent.

According to Diewert [69] a function $f : K \to \mathbb{R}$ defined on a nonempty convex subset K of $\mathbb{R}^n$ is said to be *upper-Dini pseudoconvex* if for all $x, y \in K$,

$$f(y) < f(x) \quad \Rightarrow \quad f^D(x; y-x) < 0,$$

and *lower-Dini pseudoconvex* if for all $x, y \in K$,

$$f(y) < f(x) \quad \Rightarrow \quad f_D(x; y-x) < 0.$$

From the definitions it is apparent that every upper-Dini pseudoconvex function is lower-Dini pseudoconvex but the reverse implication does not hold. We recall an example given by Diewert [69] to justify this fact.

Example 3.5. Consider a function $f : [0, 1/2] \to \mathbb{R}$ defined as

$$f(x) = \begin{cases} 0, & \text{if } x = 0, \\ -x^2 - \frac{1}{2^{2^n}}, & \text{if } \frac{1}{2^{2^n}} \leq x < \frac{1}{2^{2^{n-1}}}, \quad n = 1, 2, \ldots, \\ -x^2 - \frac{1}{2}, & \text{if } x = \frac{1}{2}. \end{cases}$$

This function zigzags with discontinuities between the functions

$$h_1(x) = -x - x^2 \quad \text{and} \quad h_2(x) = -2x^2.$$

The function f is lower-Dini pseudoconvex but not upper-Dini pseudoconvex because at $x = 0$ and $y = 1/2$,

$$f^D(x; y-x) = 0 \quad \text{and} \quad f(y) < f(x).$$

Glover [91] established that a radially continuous upper-Dini pseudoconvex function is quasiconvex, upper-Dini quasiconvex, and semistrictly quasiconvex (termed as strictly quasiconvex by Glover). Further, he introduced the notion of Clarke-quasiconvexity for locally Lipschitz functions as follows.

A locally Lipschitz function $f : K \to \mathbb{R}$ is said to be *Clarke-quasiconvex* if for all $x, y \in K$,

$$f(y) \leq f(x) \quad \Rightarrow \quad f^C(x; y-x) \leq 0.$$

Hiriart-Urruty [104] termed a locally Lipschitz function $f : K \to \mathbb{R}$ as *Clarke-pseudoconvex* if for all $x, y \in K$,

$$f(y) < f(x) \quad \Rightarrow \quad f^C(x; y-x) < 0.$$

Since

$$f_D(x; d) \leq f^D(x; d) \leq f^C(x; d), \quad \text{for all } x \in K \text{ and all } d \in \mathbb{R}^n,$$

it follows that

$$\boxed{\text{lower-Dini quasiconvexity}} \Leftarrow \boxed{\text{upper-Dini quasiconvexity}} \Leftarrow \boxed{\text{Clarke quasiconvexity}}$$

Similar implications hold in the case of pseudoconvexity as well. Hence, if f is a radially continuous locally Lipschitz Clarke-pseudoconvex function, then it is quasiconvex as well as semistrictly quasiconvex.

We give an example of an upper-Dini pseudoconvex function that is not Clarke-pseudoconvex.

Example 3.6. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2, & \text{if } x > 0, \\ x, & \text{if } x \leq 0, \end{cases}$$

is upper-Dini (lower-Dini) pseudoconvex, where the Dini directional derivatives at x in the direction $d \in \mathbb{R}$ are given by

$$\begin{aligned} f^D(x;d) = f_D(x;d) &= \begin{cases} 2xd, & \text{if } x = 0,\ d > 0 \text{ or } x > 0, \\ d, & \text{if } x = 0,\ d \leq 0 \text{ or } x < 0, \end{cases} \\ &= \begin{cases} 2xd, & \text{if } x > 0, \\ \min\{0, d\}, & \text{if } x = 0, \\ d, & \text{if } x < 0. \end{cases} \end{aligned}$$

As the Clarke generalized directional derivative is

$$f^C(x;d) = \begin{cases} 2xd, & \text{if } x > 0, \\ \max\{0, d\}, & \text{if } x = 0, \\ d, & \text{if } x < 0, \end{cases}$$

it can be seen that f is not Clarke-pseudoconvex because for $x = 0$ and $y = -1$, $f(y) < f(x)$ but $f^C(x; y - x) = 0$.

The following definition by Komlósi [132] subsumes the definitions given earlier by Diewert [69], Hiriart-Urruty [104], and Giorgi and Komlósi [88].

Definition 3.3. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. A function $f : K \to \mathbb{R}$ is said to be

(a) *strictly h-pseudoconvex* if for all $x, y \in K$, $x \neq y$,

$$f(y) \leq f(x) \quad \Rightarrow \quad h(x; y - x) < 0;$$

(b) *h-pseudoconvex* if for all $x, y \in K$, $x \neq y$,

$$f(y) < f(x) \quad \Rightarrow \quad h(x; y - x) < 0;$$

(c) *h-quasiconvex* if for all $x, y \in K$, $x \neq y$,

$$f(y) \leq f(x) \quad \Rightarrow \quad h(x; y - x) \leq 0.$$

If in (a), (b), and (c), the inequalities $<$ and $\leq$ are replaced by $>$ and $\geq$, respectively, then the function f is called *strictly h-pseudoconcave*, *h-pseudoconcave*, and *h-quasiconcave*, respectively.

If f is a differentiable function and $h(x; y - x) = \langle \nabla f(x), y - x \rangle$, then the notion of strict h-pseudoconvexity (respectively, h-pseudoconvexity, h-quasiconvexity) of f is the same as that of strict pseudoconvexity (respectively, pseudoconvexity, quasiconvexity) of f. Observe that if f is strictly h-pseudoconvex (respectively, h-pseudoconvex, h-quasiconvex) then $-f$ is strictly $-h$-pseudoconcave (respectively, $-h$-pseudoconcave, $-h$-quasiconcave).

Remark 3.4. Let K be a nonempty convex subset of $\mathbb{R}^n$. If $g, h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ are two bifunctions such that $g(x; d) \leq h(x; d)$ for all $x \in K$ and all $d \in \mathbb{R}^n$, then every strictly h-pseudoconvex (respectively, h-pseudoconvex, h-quasiconvex) is also strictly g-pseudoconvex (respectively, g-pseudoconvex, g-quasiconvex).

The interrelations among the classes defined so far have been summed up in Figure 3.3.

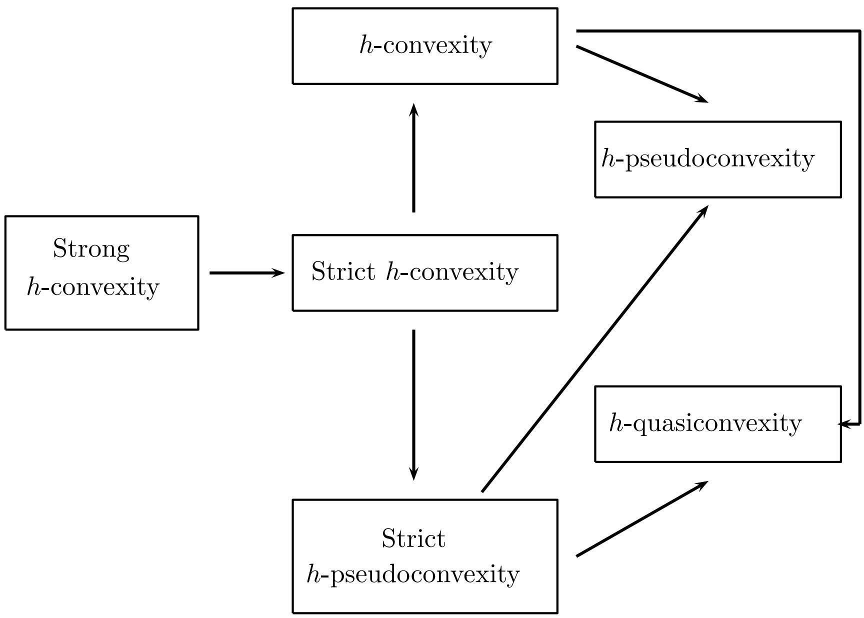


FIGURE 3.1: Relations among different kinds of h-convexities

Unlike the differentiable case, there does not exist any relationship between h-pseudoconvexity and h-quasiconvexity.

Although the class of h-pseudoconvex functions subsumes the classes of h-convex functions and strictly h-pseudoconvex functions, these inclusions are strict.

Example 3.7. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} x^2, & \text{if } x \leq 0, \\ 0, & \text{if } 0 < x \leq 1, \\ x+1, & \text{if } x > 1. \end{cases}$$

For each $(x; d) \in \mathbb{R} \times \mathbb{R}$, let

$$h(x;d) = f^D(x;d) = \begin{cases} 2xd, & \text{if } x \leq 0, \\ 0, & \text{if } x = 1,\ d \leq 0 \text{ or } 0 < x < 1, \\ +\infty, & \text{if } x = 1,\ d > 0, \\ d, & \text{if } x > 1. \end{cases}$$

Then, clearly f is h-pseudoconvex but not h-convex as h-convexity fails at the points $x = 1$ and $y = 2$. Also, by considering the points $x = 1$ and $y = 1/2$, it can be seen that f is not strictly h-pseudoconvex.

However, if h is majorized by the Dini upper directional derivative of f, then every quasiconvex function is h-quasiconvex.

Theorem 3.7. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a quasiconvex function and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction satisfying (3.3), then f is h-quasiconvex.

Proof. Since f is quasiconvex, it follows that for any $x, y \in K$, $x \neq y$,

$$f(y) \leq f(x) \quad \Rightarrow \quad f(x + \lambda(y - x)) \leq f(x), \quad \text{for all } \lambda \in [0, 1].$$

The above inequality implies that $f^D(x; y - x) \leq 0$ and using (3.3), it follows that f is h-quasiconvex. $\square$

If f is radially lower semicontinuous and $h \equiv f^D$, then the converse of Theorem 3.7 holds.

Theorem 3.8. [69] Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be radially lower semicontinuous. Then, f is quasiconvex if and only if it is f^D-quasiconvex.

A more general form of Theorem 3.8 is established in the form of Theorem 4.21. Therefore, we omit the proof of this theorem.

In general, a quasiconvex function may not be h-quasiconvex, as is clear from the following example.

Example 3.8. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} 0, & \text{if } x > 0, \\ x^2, & \text{if } x \leq 0. \end{cases}$$

Then, for each $(x; d) \in \mathbb{R} \times \mathbb{R}$

$$f^D(x; d) = \begin{cases} 0, & \text{if } x \geq 0, \\ 2xd, & \text{if } x < 0. \end{cases}$$

If we choose

$$h(x; d) = \begin{cases} 1, & \text{if } x \geq 0, \\ 2xd, & \text{if } x < 0, \end{cases}$$

then it can be easily verified that f is quasiconvex, but it is not h-quasiconvex at $x = 1$ and $y = 2$.

Theorem 3.9. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a h-pseudoconvex function where $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfies condition (3.3) and is both subodd and positively homogenous in the second argument. Then, f is quasiconvex, and hence h-quasiconvex.

Proof. On the contrary, suppose that there exist points $x, y, z \in K$, $x \neq y$ such that $z \in\ ]x, y[$ and $f(z) > \max\{f(x), f(y)\}$. Then by h-pseudoconvexity of f, it follows that $h(z; x-z) < 0$ and $h(z; y-z) < 0$. Positive homogeneity of h in the second argument implies that $h(z; x-y) < 0$ and $h(z; y-x) < 0$, and hence $h(z; x-y) + h(z; y-x) < 0$, which is a contradiction to the suboddness of h. Thus, f is quasiconvex, and hence by Theorem 3.7 it follows that f is h-quasiconvex. □

The next example demonstrates that the suboddness condition in Theorem 3.9 is imperative. We thus have an example of an h-pseudoconvex function that is not h-quasiconvex.

Example 3.9. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} 1, & \text{if } x \text{ is irrational}, \\ -1, & \text{if } x \text{ is rational}. \end{cases}$$

Define $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ as follows

$$h(x; d) = \begin{cases} -|d|, & \text{if } x \text{ is irrational}, \\ |d|, & \text{if } x \text{ is rational}. \end{cases}$$

Clearly, f is an h-pseudoconvex function but not h-quasiconvex because for $x = 1$ and $y = 1$, $f(y) = f(x)$ and $h(x; y-x) > 0$. Here, h is positively homogeneous in the second argument and satisfies condition (3.3) but h fails to satisfy the suboddness assumption.

The following is an example of an h-quasiconvex function that is not h-pseudoconvex.

Example 3.10. Consider a function $f : \mathbb{R} \to \mathbb{R}$ defined as

$$f(x) = \begin{cases} x^3, & \text{if } x \neq 0, \\ -1, & \text{if } x = 0. \end{cases}$$

Then,

$$f^D(x; d) = \begin{cases} 3x^2 d, & \text{if } x \neq 0, \\ +\infty, & \text{if } x = 0. \end{cases}$$

If we define

$$h(x; d) = -3x^2|d|, \quad \text{for all } (x, d) \in \mathbb{R} \times \mathbb{R},$$

then clearly h satisfies condition (3.3). It can be seen that f is h-quasiconvex but not h-pseudoconvex because for $x = 0$ and $y = -2$, $f(y) < f(x)$ and $h(x; y - x) = 0$. We also observe that f is not upper-Dini quasiconvex because for x and y considered above $f^D(x; y - x) > 0$.

The next theorem generalizes the corresponding result given by Giorgi and Komlósi [88, Theorem 3.4].

Theorem 3.10. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. If f is (strictly) h-pseudoconvex and $x \in K$ is an inf-stationary point of f over K, then x is a (unique) minimum of f on K.

Proof. The proof follows on using the definition of (strict) h-pseudoconvexity. □

Theorem 3.11. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a function. Let $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy (3.6) and $-h$ be both subodd and positively homogeneous in the second argument. Then f is quasiconcave, and hence h-quasiconcave.

Proof. Suppose that f is h-pseudoconcave, then $-f$ is $-h$-pseudoconvex. Also from condition (3.6), it follows that

$$-f_D(x; d) \geq -h(x; d), \quad \text{for all } x \in K \text{ and all } d \in \mathbb{R}^n,$$

that is,

$$(-f)^D(x; d) \geq -h(x; d), \quad \text{for all } x \in K \text{ and all } d \in \mathbb{R}^n. \tag{3.7}$$

Thus using inequality (3.7) and suboddness of $-h$, it follows from Theorem 3.9 that $-f$ is quasiconvex and $-h$-quasiconvex. This implies that f is quasiconcave and h-quasiconcave. □

The following two theorems follow trivially from the definitions of h-quasiconcavity and h-pseudoconcavity.

Theorem 3.12. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a quasiconcave function and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy (3.6). Then, f is h-quasiconcave.

Theorem 3.13. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. If f is (strictly) h-pseudoconcave and $x \in K$ is a sup-stationary point of f over K, then x is a (unique) maximum of f on K.

3.4 Nonsmooth Pseudolinearity

In this section, we study pseudolinearity for the nondifferentiable case. We introduce the notion of pseudolinearity for a function f in terms of a bifunction h.

Definition 3.4. Let K be a nonempty convex subset of $\mathbb{R}^n$. A function $f : K \to \mathbb{R}$ is said to be *h-pseudolinear* if f is both h-pseudoconvex and h-pseudoconcave.

If f is a differentiable h-pseudolinear function with $h(x; d) = \langle \nabla f(x), d \rangle$, then Definition 3.4 reduces to Definition 1.36 given in the first chapter.

Definition 3.5. Let K be a nonempty convex subset of $\mathbb{R}^n$. A function $f : K \to \mathbb{R}$ is said to be

(a) *upper-Dini pseudoconcave* if for all $x, y \in K$,

$$f(y) > f(x) \quad \Rightarrow \quad f^D(x; y - x) > 0;$$

(b) *upper-Dini pseudolinear* if it is both upper-Dini pseudoconvex as well as upper-Dini pseudoconcave.

If $h(x; d) = f^D(x; d)$, then an h-pseudolinear function is upper-Dini pseudolinear. Similarly, we can define a lower-Dini pseudolinear function.

We give two examples of h-pseudolinear functions. It may be noted that both the functions are not pseudolinear. In the first example the function f is both upper-Dini and lower-Dini pseudolinear.

Example 3.11. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} x^2 + 3x + 1, & \text{if } x \geq 0, \\ x, & \text{if } x < 0. \end{cases}$$

Then,

$$f^D(x; d) = f_D(x; d) = \begin{cases} (2x + 3)d, & \text{if } x = 0,\ d \geq 0 \text{ or } x > 0, \\ -\infty, & \text{if } x = 0,\ d < 0, \\ d, & \text{if } x < 0. \end{cases}$$

If we take $h(x;d) = f^D(x;d)$ for all $x, d \in \mathbb{R}$, then it can be seen that f is h-pseudolinear.

We now give an example of an h-pseudolinear function which is neither lower-Dini nor upper-Dini pseudolinear.

Example 3.12. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} \frac{5}{2}x, & \text{if } x \leq 0, \\ x + \frac{1}{2^{2n}}, & \text{if } \frac{1}{2^{(2n+1)}} < x \leq \frac{1}{2^{(2n)}},\ n = 0, 1, 2, \ldots, \\ 4x - \frac{1}{2^{(2n+1)}}, & \text{if } \frac{1}{2^{(2n+2)}} < x \leq \frac{1}{2^{(2n+1)}},\ n = 0, 1, 2, \ldots, \\ 4x - 2, & \text{if } x > 1, \end{cases}$$

then f is a h-pseudolinear function, where $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is

$$h(x;d) = \begin{cases} 5d/2, & \text{if } x \leq 0, \\ d, & \text{if } 1/2^{(2n+1)} < x < 1/2^{2n},\ n = 0, 1, 2, \ldots, \\ 4d, & \text{if } 1/2^{(2n+2)} < x < 1/2^{(2n+1)},\ n = 0, 1, 2, \ldots, \\ \max\{d, 4d\}, & \text{if } x = 1/2^{2n},\ n = 0, 1, 2, \ldots \\ \min\{d, 4d\}, & \text{if } x = 1/2^{(2n+1)},\ n = 0, 1, 2, \ldots, \\ 4d, & \text{if } x > 1. \end{cases}$$

It is worth noting that $f_D(0;d) < h(0;d) < f^D(0;d)$ for all $d > 0$.

The following theorem gives a necessary condition for the function f to be h-pseudolinear.

Theorem 3.14. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy (3.3) and (3.6), and h is positively homogeneous in the second argument. Further, suppose that one of the following conditions hold:

(i) f is a radially continuous function;

(ii) h is odd in the second argument.

If the function f is h-pseudolinear, then for all $x, y \in K$,

$$h(x; y - x) = 0 \quad \text{if and only if} \quad f(x) = f(y). \tag{3.8}$$

Proof. Let $x, y \in K$, $x \neq y$, be such that $h(x; y - x) = 0$. Then by h-pseudoconvexity and h-pseudoconcavity of f, it follows that $f(x) = f(y)$.

Conversely, let $f(x) = f(y)$ for some $x, y \in K$. If $x = y$, then by positive homogeneity of h, it follows that $h(x; \mathbf{0}) = 0$, and hence, $h(x; y - x) = 0$. Let $x \neq y$. Suppose that (i) holds. We assert that $f(x) = f(z)$ for all $z \in\ ]x, y[$. Suppose on the contrary $f(x) > f(z)$ for some $z \in\ ]x, y[$. Then by h-pseudoconcavity of f, we have $h(z; x - z) > 0$. From condition (3.3), it follows that $f^D(z; x-z) > 0$, that is, there exists a sequence $t_n \to 0^+$, $t_n \in\ ]0,1[$ such that $f(z + t_n(x-z)) > f(z)$. Since f is radially continuous, there exists some $t_n \in\ ]0,1[$ such that $f(z) < f(z_n) < f(x)$, where $z_n = z + t_n(x-z)$. As $f(x) = f(y)$, we have $f(z_n) < f(y)$, and by h-pseudoconcavity of f, it follows that $h(z_n; y - z_n) > 0$. Now since h is positively homogeneous in the second argument, it follows that $h(z_n; z - z_n) > 0$ which by h-pseudoconvexity of f implies that $f(z_n) \leq f(z)$, which is clearly a contradiction.

Now let us assume that $f(x) < f(z)$. Then by h-pseudoconcavity of f, we get $h(x; z-x) > 0$. From condition (3.3), it follows that $f^D(x; z-x) > 0$, that is, there exists a sequence $\lambda_n \to 0^+$, $\lambda_n \in\ ]0,1[$ such that

$$f(x + \lambda_n(z-x)) > f(x).$$

Then by radial continuity of f, there exists some $\lambda_n \in\ ]0,1[$ such that $f(x) < f(x_n) < f(z)$, where $x_n = x + \lambda_n(z-x)$. Since $f(y) = f(x) < f(x_n)$, by h-pseudoconvexity of f, it follows that $h(x_n; y - x_n) < 0$. Again employing positive homogeneity of h, we get $h(x_n; z - x_n) < 0$. Using h-pseudoconcavity of f, it follows that $f(x_n) \geq f(z)$, which is a contradiction. Therefore,

$$f(x) = f(z), \quad \text{for all } z \in\ ]x, y[,$$

and hence,

$$f^D(x; y-x) = f_D(x; y-x) = 0.$$

Then the relations (3.3) and (3.6) yield that $h(x; y-x) = 0$.

Now, if (ii) holds then as f is h-pseudolinear it follows from Theorems 3.9 and 3.11 that f is both h-quasiconvex as well as h-quasiconcave, and hence, we have $h(x; y-x) = 0$. $\square$

Remark 3.5. It may be noted that in Example 3.11 and Example 3.12, the conditions (3.3) and (3.6) both hold.

Remark 3.6. The conclusion of Theorem 3.14 may not hold if neither of the assumptions (i) and (ii) are satisfied. The function f considered in Example 3.9 is h-pseudolinear and h satisfies (3.3) and (3.6). Here, h is positively homogeneous but fails to be odd in the second argument, f is not radially continuous and we observe that $f(x) = f(y)$ but $h(x; y-x) \neq 0$ at $x = 1$ and $y = -1$.

We prove the converse of Theorem 3.14 assuming f to be a continuous function.

Theorem 3.15. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a continuous function. Let $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy (3.3) and (3.6), and be positively homogeneous in the second argument. If the condition (3.8) holds for all $x, y \in K$, then f is h-pseudolinear.

Proof. Let $f(y) < f(x)$ for some $x, y \in K$. Then by (3.8), it follows that $h(x; y-x) \neq 0$. Suppose that $h(x; y-x) > 0$, then from (3.3), we have

$$f^D(x; y-x) > 0,$$

which implies that there exists some $\lambda \in \ ]0,1[$ such that

$$f(x + \lambda(y-x)) > f(x).$$

Since f is continuous and $f(y) < f(x) < f(x+\lambda(y-x))$, it follows that there exists some $\bar{\lambda} \in \ ]0,1[$ such that $f(z) = f(x)$, where $z = \bar{\lambda} y + (1-\bar{\lambda})x$. Then by hypothesis, we have $h(x; z-x) = 0$, and using positive homogeneity of h we get $h(x; y-x) = 0$, which is a contradiction. Hence, $h(x; y-x) < 0$. This proves that f is h-pseudoconvex.

Similarly, it can be proved that f is h-pseudoconcave. □

Remark 3.7. The continuity assumption on f in Theorem 3.15 cannot be relaxed. This can be seen from the following example.

Example 3.13. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} e^x, & \text{if } x < 0, \\ -x-1, & \text{if } x \geq 0. \end{cases}$$

Then,

$$f^D(x; d) = f_D(x; d) = \begin{cases} e^x d, & \text{if } x < 0, \\ +\infty, & \text{if } x = 0,\ d < 0, \\ -d, & \text{if } x = 0, d \geq 0 \text{ or } x > 0. \end{cases}$$

If we take $h(x; d) = f^D(x; d)$ for all $x, d \in \mathbb{R}$, then it can be seen that condition (3.8) is trivially satisfied but f is not h-pseudolinear because for $y = 0$ and $x = -1$, $f(y) < f(x)$ whereas $h(x; y-x) > 0$.

The following theorem gives a characterization for continuous h-pseudolinear functions. This extends the result for the differentiable case as given in Theorem 1.38.

Theorem 3.16. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy (3.3) and (3.6), and h is positively homogeneous in the second argument. Further, suppose that one of the following conditions hold:

(i) f is a radially continuous function;

(ii) h is odd in the second argument.

Then, f is h-pseudolinear if and only if there exists a real-valued function p defined on $K \times K$ such that for all $x, y \in K$, $p(x, y) > 0$ and

$$f(y) = f(x) + p(x, y) h(x; y-x).$$

Proof. Suppose f is h-pseudolinear. If $h(x; y-x) = 0$ for some x, $y \in K$, by Theorem 3.14, it follows that $f(y) = f(x)$. In this case we define $p(x, y) = 1$. If $h(x; y-x) \neq 0$, then again by Theorem 3.14, $f(y) \neq f(x)$ and here we define

$$p(x, y) = \frac{f(y) - f(x)}{h(x; y-x)}.$$

We next show that $p(x, y) > 0$. If $f(y) > f(x)$, then by h-pseudoconcavity of f, it follows that $h(x; y-x) > 0$, hence the result holds.

Similarly, if $f(y) < f(x)$ we get the result by h-pseudoconvexity of f.

The converse implication is obvious. □

Theorem 3.17. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy (3.3) and (3.6), and h is positively homogeneous in the second argument. Further, suppose that one of the following conditions hold:

(i) f is a radially continuous function;

(ii) h is odd in the second argument.

If f is h-pseudolinear, then for all $x, y \in K$,

$$h(x; y-x) = 0 \quad \Leftrightarrow \quad f(x) = f(\lambda x + (1-\lambda)y), \quad \text{for all } \lambda \in [0, 1]. \tag{3.9}$$

Proof. For $x, y \in K$, let $h(x; y-x) = 0$. Then for all $\lambda \in \ ]0, 1[$, it follows using the positive homogeneity assumption that

$$h(x; \lambda x + (1-\lambda)y - x) = (1-\lambda)h(x; y-x) = 0.$$

The result then follows by Theorem 3.14. □

As an illustration to Theorem 3.17 consider the following example.

Example 3.14. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} x_1 + x_2, & \text{if } x_1 + x_2 > 0, \\ \frac{x_1+x_2}{2}, & \text{if } x_1 + x_2 \leq 0, \end{cases}$$

then,

$$f^D(x; d) = f_D(x; d) = \begin{cases} d_1 + d_2, & \text{if } x_1 + x_2 > 0, \\ \max\{d_1 + d_2, (d_1 + d_2)/2\}, & \text{if } x_1 + x_2 = 0, \\ (d_1 + d_2)/2, & \text{if } x_1 + x_2 < 0, \end{cases}$$

where $x = (x_1, x_2)$ and $d = (d_1, d_2)$.

Here, f is a continuous h-pseudolinear function with $h(x; d) = f^D(x; d)$, for all $x, d \in \mathbb{R}^2$. We may note that for $x = (1, 2)$, $y = (2, 1)$, $h(x; y-x) = 0$, and clearly $f(x) = f(\lambda x + (1-\lambda)y)$, for all $\lambda \in [0, 1]$.

Remark 3.8. The converse Theorem 3.17 is not necessarily true. The function f considered in Example 3.13 satisfies condition (3.9) but f is not h-pseudolinear. However, from Theorem 3.15 it is clear that the converse holds if f is a continuous function.

3.5 Generalized Nonsmooth Convexity in Terms of Subdifferentials

This section deals with different kinds of convexities defined by means of the Clarke subdifferential. Some characterizations of such kinds of convexities are given.

Definition 3.6. Let K be a nonempty convex subset of $\mathbb{R}^n$. A locally Lipschitz function $f : K \to \mathbb{R}$ is said to be

(a) *generalized convex* if for all $x, y \in K$ and all $\xi \in \partial^C f(x)$, we have

$$\langle \xi, y - x \rangle \leq f(y) - f(x); \tag{3.10}$$

(b) *generalized strongly convex with modulus* ρ if there exists a real number $\rho > 0$ such that for all $x, y \in K$ and all $\xi \in \partial^C f(x)$, we have

$$\langle \xi, y - x \rangle + \rho \|y - x\|^2 \leq f(y) - f(x). \tag{3.11}$$

If the strict inequality holds in (3.10) for all $x \neq y$, then f is said to be *generalized strictly convex.*

Analogously, we can define the generalized convexities in terms of other subdifferentials.

The following result provides characterizations for generalized convex functions and generalized strongly convex functions.

Theorem 3.18. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipshcitz function. Then,

(a) f is (strictly) convex if and only if it is generalized (strictly) convex;

(b) f is strongly convex if and only if it is generalized strongly convex.

Proof. We only prove assertion (b). The "if only" part of assertion (a) is trivial, however, if part of assertion (a) can be obtained by taking $\rho = 0$ in the (b) part.

Suppose that f is strongly convex with modulus $\rho > 0$. Then for all $x, y \in K$ and all $\lambda \in (0, 1)$, we have

$$\lambda[f(\lambda y+(1-\lambda)x)-f(y)]+(1-\lambda)[f(\lambda y+(1-\lambda)x)-f(x)] \leq -\rho\lambda(1-\lambda)\|y-x\|^2.$$

Dividing by $\lambda(1 - \lambda)$ and then taking the limit as λ approaches to 0 from the right, we obtain

$$f(x) - f(y) + f^C(x; y - x) \leq -\rho\|y - x\|^2.$$

By Theorem 2.18(c), we deduce that the inequality (3.11) holds for all $\xi \in \partial^C f(x)$. Hence, f is generalized strongly convex.

Conversely, assume that the inequality (3.11) holds for all $x, y \in K$ and all $\xi \in \partial^C f(x)$. Then for all $\lambda \in [0,1]$ and all $\zeta \in \partial^C f(\lambda y + (1-\lambda)x)$, we have

$$f(y) - f(\lambda y + (1-\lambda)x) \geq (1-\lambda)\langle \zeta, y - x\rangle + \rho(1-\lambda)^2\|y-x\|^2 \tag{3.12}$$

and

$$f(x) - f(\lambda y + (1-\lambda)x) \geq -\lambda\langle \zeta, y - x\rangle + \rho\lambda^2\|y-x\|^2. \tag{3.13}$$

Multiplying (3.12) by λ and (3.13) by $1-\lambda$, and then adding the resultants, we obtain

$$\lambda f(y) + (1-\lambda)f(x) - f(\lambda y + (1-\lambda)x) \geq \rho\lambda(1-\lambda)\|y-x\|^2,$$

which shows that f is strongly convex with constant $\rho > 0$. □

Definition 3.7. Let K be a nonempty convex subset of $\mathbb{R}^n$. A locally Lipschitz function $f : K \to \mathbb{R}$ is said to be

(a) *generalized pseudoconvex* if for all $x, y \in K$ and any $\xi \in \partial^C f(x)$, we have

$$\langle \xi, y - x\rangle \geq 0 \quad \text{implies} \quad f(y) \geq f(x);$$

(b) *generalized strictly pseudoconvex* if for all $x, y \in K$ with $x \neq y$ and any $\xi \in \partial^C f(x)$, we have

$$\langle \xi, y - x\rangle \geq 0 \quad \text{implies} \quad f(y) > f(x);$$

(c) *generalized strongly pseudoconvex with modulus ρ* if there exists a real number $\rho > 0$ such that for all $x, y \in K$ and any $\xi \in \partial^C f(x)$, we have

$$\langle \xi, y - x\rangle \geq 0 \quad \Rightarrow \quad f(y) \geq f(x) + \rho\|y-x\|^2;$$

(d) *generalized quasiconvex* if for all $x, y \in K$ and all $\xi \in \partial^C f(x)$, we have

$$f(y) \leq f(x) \quad \text{implies} \quad \langle \xi, y - x\rangle \leq 0.$$

The class of generalized strictly pseudoconvex functions is a subset of the class of generalized pseudoconvex functions and is a superset of the class of generalized strongly pseudoconvex functions.

Theorem 3.19. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipschitz function. Then f is quasiconvex if and only if it is generalized quasiconvex.

Proof. Let f be a quasiconvex function and let $x, y \in K$ be such that $f(y) \leq f(x)$. By quasiconvexity of f, $f(x + \lambda(y - x)) \leq f(x)$ for all $\lambda \in\ ]0, 1[$. Since f is locally Lipschitz, for $\varepsilon > 0$ small enough, we have

$$\begin{aligned} f^C(x; y - x) &= \limsup_{\substack{x_0 \to x \\ t \to 0^+}} \frac{f(x_0 + t(y - x)) - f(x_0)}{t} \\ &\leq \limsup_{t \to 0^+} \frac{f(x + t(y - x)) - f(x)}{t} + \varepsilon \leq \varepsilon. \end{aligned}$$

The arbitrariness of ε shows that $f^C(x; y - x) \leq 0$. From Theorem 2.18(c), it follows that $\langle \xi, y - x \rangle \leq 0$ for all $\xi \in \partial^C f(x)$, and hence, f is generalized quasiconvex.

Conversely, let f be generalized quasiconvex. Let $x, y \in K$ be such that $f(y) \leq f(x)$. Then by the definition of generalized quasiconvexity of f and Theorem 2.18(c), we obtain $f^C(x; y - x) \leq 0$. Consequently,

$$\begin{aligned} f^D(x; y - x) &= \limsup_{t \to 0^+} \frac{f(x + t(y - x)) - f(x)}{t} \\ &\leq \limsup_{\substack{x_0 \to x \\ t \to 0^+}} \frac{f(x_0 + t(y - x)) - f(x_0)}{t} \\ &= f^C(x; y - x) \leq 0. \end{aligned}$$

The local Lipschitzian property of f implies that for any given $a, b \in K$ and $t_0 \in [0, 1]$, we have

$$\liminf_{t \to t_0} S(t) := \liminf_{t \to t_0} f(a + t(b - a)) = S(t_0).$$

This implies that f is radially lower semicontinuous. Therefore, by Theorem 3.8 f is quasiconvex. □

Lemma 3.1. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be locally Lipschitz. If there exist $x, y \in K$ such that $\langle \zeta, y - x \rangle < 0$ for all $\zeta \in \partial^C f(y)$, then there exists $\mu \in\ ]0, 1[$ such that $f(\mu y + (1 - \mu)x) > f(y)$.

Proof. Since f is locally Lipschitz, we can consider a $\lambda \in\ ]0, 1[$ such that f is locally Lipschitz on a neighborhood of the line segment $[x + \lambda(y - x), y]$. Then by Lebourg's mean value theorem (Theorem 2.23), there exist a point $z \in\ ]x + \lambda(y - x), y[$ and a vector $\xi \in \partial^C f(z)$ such that

$$f(y) - f(x + \lambda(y - x)) = (1 - \lambda)\langle \xi, y - x \rangle.$$

Therefore,

$$\lim_{\lambda \to 0^+} \frac{f(y) - f(x + \lambda(y - x))}{1 - \lambda} = \langle \xi, y - x \rangle < 0,$$

where the last strict inequality is obtained from the hypothesis. Thus, there exists a $\mu \in\ ]0, 1[$ such that $f(y) - f(x + \mu(y - x)) < 0$. □

Theorem 3.20. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be locally Lipschitz. If f is generalized pseudoconvex, then it is semistrictly quasiconvex.

Proof. Assume that f is not semistrictly quasiconvex. Then there exist $x, y \in K$ and $\lambda \in \ ]0,1[$ such that $f(x) < f(y)$ and $f(\lambda x + (1-\lambda)y) \geq f(y)$. Setting $z = \lambda x + (1-\lambda)y$, we get $f(z) \geq f(y) > f(x)$. By generalized pseudoconvexity of f, we obtain

$$\langle \xi, x - z \rangle < 0, \quad \text{for all } \xi \in \partial^C f(z),$$

that is,

$$\langle \xi, y - z \rangle > 0, \quad \text{for all } \xi \in \partial^C f(z).$$

Therefore, again by generalized pseudoconvexity of f, we have $f(z) \leq f(y)$, and hence, $f(z) = f(y)$.

On the other hand, by Lemma 3.1, there exist a vector $\hat{x}$ and $\mu \in \ ]0,1[$ such that $\hat{x} = \mu y + (1-\mu)z \in K$ and $f(\hat{x}) > f(z) = f(y)$ since $\langle \xi, y - z \rangle > 0$, for all $\xi \in \partial^C f(z)$. Thus, using generalized pseudoconvexity of f, we get

$$\langle \hat{\xi}, y - \hat{x} \rangle < 0 \quad \text{and} \quad \langle \hat{\xi}, z - \hat{x} \rangle < 0, \quad \text{for all } \hat{\xi} \in \partial^C f(\hat{x}).$$

However,

$$\langle \hat{\xi}, z - \hat{x} \rangle = \frac{-\mu \langle \hat{\xi}, y - \hat{x} \rangle}{1 - \mu} > 0.$$

This is an obvious contradiction. □

The above result was obtained by Soleimani-damaneh [195] in a more general setting.

Theorem 3.21. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be locally Lipschitz. If f is generalized pseudoconvex, then it is quasiconvex.

Proof. Assume that f is not quasiconvex. Then, there exist $x, y \in K$ and $\hat{\lambda} \in \ ]0,1[$ such that $f(x_{\hat{\lambda}}) > \max\{f(x), f(y)\}$, where $x_{\hat{\lambda}} = x + \hat{\lambda}(y - x)$. Without loss of generality, we may assume that $f(y) \leq f(x)$. Then, we have

$$f(y) \leq f(x) < f(x_{\hat{\lambda}}). \tag{3.14}$$

Let $\varphi(\lambda) = f(x + \lambda(y - x))$. Since f is a continuous function, it follows that $\varphi(\lambda)$ is a continuous function defined on $[0,1]$, and hence, $\varphi(\lambda)$ attains its maximum. From (3.14), we have $\varphi(0) = f(x) < f(x_{\hat{\lambda}})$. So $\lambda = 0$ is not a maximum point. Again from (3.14), we have $f(y) = \varphi(1) \leq f(x) < f(x_{\hat{\lambda}})$ which implies that $\lambda = 1$ is not a maximum point. Hence, there exists $\lambda^* \in \ ]0,1[$ such that

$$f(x^*) = \max_{\lambda \in [0,1]} f(x + \lambda(y - x)),$$

where $x^* = x + \lambda^*(y - x)$. Then, by Theorem 2.20, $\mathbf{0} \in \partial^C f(x^*)$. For $\xi = \mathbf{0}$, we have $\langle \xi, y - x^* \rangle = 0$. Since f is generalized pseudoconvex, we obtain $f(y) \geq f(x^*)$ which contradicts inequality (3.14). In the same way, we get the conclusion when we consider $f(y) > f(x)$. □

The following example shows that a quasiconvex function may not be generalized pseudoconvex.

Example 3.15. Let $K = \,]-1,1[$ and $f : K \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x, & \text{if } x \leq 0, \\ 0, & \text{if } x > 0. \end{cases}$$

The Clarke subdifferential of f is

$$\partial^C f(x) = \begin{cases} \{1\}, & \text{if } x < 0, \\ \{0,1\}, & \text{if } x = 0, \\ \{0\}, & \text{if } x > 0. \end{cases}$$

Then, f is quasiconvex, but it is not generalized pseudoconvex.

Theorem 3.22. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be Lipschitz continuous and quasiconvex. Then, for all $x, y \in K$, $x \neq y$,

$$\langle \zeta, y - x \rangle > 0 \text{ for some } \zeta \in \partial^C f(x) \quad \Rightarrow \quad f(y) \geq f(x).$$

Proof. Let $\zeta \in \partial^C f(x)$ and $\langle \zeta, y-x \rangle > 0$. Then, there exist sequences $\{x_m\} \subseteq K$ and $\zeta_m \in \partial^C f(x_m)$ such that $f(x_m) \to f(x)$ and $\zeta_m \to \zeta$. As, $\langle \zeta, y-x \rangle > 0$, there exists $m_0 \in \mathbb{N}$ such that for any $m \geq m_0$, $\langle \zeta_m, y - x_m \rangle > 0$. Since $\xi_m \in \partial^C f(x_m)$, from Lemma 3.1, we can find a sequence $\lambda_m \to 0^+$ such that $f(x_m + \lambda_m(y - x_m)) > f(x_m)$. The quasiconvexity of f together with the previous inequality implies that for any $\lambda \in [0,1]$,

$$f(x_m + \lambda(y - x_m)) \leq \max\{f(y), f(x_m)\} = f(y).$$

By continuity of f, we deduce that $f(x + \lambda(y - x)) \leq f(y)$, for all $\lambda \in [0,1]$. In particular, for $\lambda = 0$, we have $f(x) \leq f(y)$. □

We remark that the conclusion of Theorem 3.22 can also be written as

$$f(y) < f(x) \quad \Rightarrow \quad \langle \zeta, y - x \rangle \leq 0 \text{ for all } \zeta \in \partial^C f(x).$$

The above three results have been investigated by Ansari and Rezaei [13].

3.6 Generalized Nonsmooth Pseudolinearity in Terms of Clarke Subdifferentials

In this section we study a class of pseudolinear functions defined in terms of the Clarke subdifferentials and present certain characterizations for these functions.

Definition 3.8. Let K be a nonempty subset of $\mathbb{R}^n$. A locally Lipschitz function $f : K \to \mathbb{R}$ is said to be

(a) *generalized pseudoconcave* if for all $x, y \in K$ and for any $\xi \in \partial^C f(x)$,

$$\langle \xi, y - x \rangle \leq 0 \quad \Rightarrow \quad f(y) \leq f(x);$$

(b) *generalized pseudolinear* if it is both generalized pseudoconvex as well as generalized pseudoconcave.

Theorem 3.23. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be locally Lipschitz and generalized pseudolinear. Then, for all $x, y \in K$, $f(x) = f(y)$ if and only if $\langle \xi, y - x \rangle = 0$ for some $\xi \in \partial^C f(x)$.

Proof. Suppose that $\langle \xi, y - x \rangle = 0$ for some $\xi \in \partial^C f(x)$. By the generalized pseudoconvexity of f, we have

$$f(y) \geq f(x). \tag{3.15}$$

Again from the generalized pseudoconcavity of f, we get

$$f(y) \leq f(x). \tag{3.16}$$

Combining (3.15) and (3.16), we obtain

$$f(x) = f(y).$$

Conversely, suppose that $f(x) = f(y)$ for some $x, y \in K$. We need to prove that $\langle \xi, y - x \rangle = 0$ for some $\xi \in \partial^C f(x)$. We first show that for any $\lambda \in \ ]0, 1[$,

$$f(x + \lambda(y - x)) = f(x). \tag{3.17}$$

If $f(x + \lambda(y - x)) > f(x)$, then by generalized pseudoconvexity of f, we have for any $\zeta \in \partial^C f(x + \lambda(y - x))$,

$$\langle \zeta, x - x - \lambda(y - x) \rangle < 0,$$

that is,

$$\langle \zeta, y - x \rangle > 0,$$

which further implies that for $\lambda \in \ ]0, 1[$,

$$\langle \zeta, y - x - \lambda(y - x) \rangle > 0.$$

From generalized pseudoconvexity of f, we have $f(y) \geq f(x + \lambda(y - x))$, which contradicts our assumption that $f(x + \lambda(y - x)) > f(x) = f(y)$. Similarly, using generalized pseudoconcavity of f, we can show that $f(x + \lambda(y - x)) < f(x)$ leads to a contradiction. Hence $f(x + \lambda(y - x)) = f(x)$ for all $\lambda \in [0, 1]$.

Let $0 < \hat{\lambda} < 1$. By Lebourg's mean value theorem (Theorem 2.23), there exist $\lambda \in \ ]0, \hat{\lambda}[$, $\zeta_\lambda \in \partial^C f(z_\lambda)$, where $z_\lambda = x + \lambda(y - x)$, such that

$$0 = f(z_\lambda) - f(x) = \lambda\langle \zeta_\lambda, y - x\rangle,$$

that is,

$$\langle \zeta_\lambda, y - x\rangle = 0.$$

Since $\zeta_\lambda \in \partial^C f(z_\lambda)$ and f is locally Lipschitz, by Theorem 2.18(a), $\|\zeta_\lambda\| \leq k$, where k is the Lipschitz constant of f. Without loss of generality, we may assume that ζ_λ converges to ζ as $\lambda \to 0$. Since $z_\lambda \to x$ as $\lambda \to 0$, by Theorem 2.21(a), $\zeta \in \partial^C f(x)$. Thus, there exists $\zeta \in \partial^C f(x)$ such that $\langle \zeta, y - x\rangle = 0$. This completes the proof. $\square$

Theorem 3.24. Let K be a nonempty convex subset of $\mathbb{R}^n$. If a function $f : K \to \mathbb{R}$ is generalized pseudolinear, then there exists a function $p : K \times K \to \mathbb{R}_+$ such that for all $x, y \in K$ and for some $\xi \in \partial^C f(x)$,

$$f(y) = f(x) + p(x, y)\langle \xi, y - x\rangle. \tag{3.18}$$

Proof. Let f be generalized pseudolinear. We have to construct a function $p : K \times K \to \mathbb{R}_+$ such that for all $x, y \in K$ and for some $\xi \in \partial^C f(x)$,

$$f(y) = f(x) + p(x, y)\langle \xi, y - x\rangle.$$

Let $x, y \in K$. If $\langle \xi, y - x\rangle = 0$ for some $\xi \in \partial^C f(x)$, then we define $p(x, y) = 1$. By Theorem 3.23, we have $f(y) = f(x)$, and thus (3.18) holds.

If $\langle \xi, y - x\rangle \neq 0$ for all $\xi \in \partial^C f(x)$, we define

$$p(x, y) = \frac{f(y) - f(x)}{\langle \xi, y - x\rangle}, \quad \text{for a fixed } \xi \in \partial^C f(x). \tag{3.19}$$

Then we have to show that $p(x, y) > 0$.

If $f(y) > f(x)$, then by the generalized pseudoconcavity of f, we have $\langle \xi, y - x\rangle > 0$ for all $\xi \in \partial^C f(x)$. From (3.19), we get $p(x, y) > 0$.

Similarly, if $f(y) < f(x)$, we can get $p(x, y) > 0$ by using generalized pseudoconvexity of f. $\square$

Chapter 4

Monotonicity and Generalized Monotonicity

4.1 Introduction

A well-known aspect of a differentiable convex function is that it is closely related to monotonicity. It is known that a differentiable function is convex if and only if its gradient is a monotone map. Generalized monotone maps also provide a characterization for generalized convex functions and they are not new in the literature. For instance, a differentiable function is quasiconvex if and only if its gradient has a property that, quite naturally, was called quasimonotonicity. It is very interesting to observe that the first appearance of generalized monotonicity (well before the birth of this terminology) dates back to 1936. Even more remarkable, it occurred independently and almost at the same time in two seminal articles: the one, by Georgescu-Roegen (1936), dealt with the concept of a local preference in consumer theory of economics, and the other one, by Wald (1936), contained the first rigorous proof of the existence of a competitive general equilibrium. It is also remarkable that various axioms on revealed preferences in consumer theory are in fact generalized monotone conditions. The recognition of close links between the already well-established field of generalized convexity and the relatively undeveloped field of generalized monotonicity gave a boost to both and led to a major increase in research activities. Today generalized monotonicity is frequently used in complementarity problems, variational inequalities, fixed point theory, optimization, equilibrium problems, and so on. This can be extended to nondifferentiable functions as well, through the use of generalized directional derivatives, subdifferentials, or multivalued maps. Similar connections were discovered between generalized convex functions and certain classes of maps, generically called generalized monotone.

4.2 Monotonicity and Its Relation with Convexity

As a generalization of an (strictly) increasing function of one variable, the notion of (strict) monotonicity arises quite naturally and has been considered by many authors. Most of the generalized monotone maps that exist in literature have been defined in such a way that in case of a gradient map they characterize some type of generalized convexity of the underlying function.

Definition 4.1. Let K be a nonempty subset of $\mathbb{R}^n$. A map $F : K \to \mathbb{R}^n$ is said to be

(a) *monotone* if for all $x, y \in K$, $x \neq y$, we have

$$\langle F(y) - F(x), y - x \rangle \geq 0;$$

(b) *strictly monotone* if for all $x, y \in K$, $x \neq y$, we have

$$\langle F(y) - F(x), y - x \rangle > 0;$$

(c) *strongly monotone* with modulus σ if there exists a real number $\sigma > 0$ such that for all $x, y \in K$, $x \neq y$, we have

$$\langle F(y) - F(x), y - x \rangle \geq \sigma \|y - x\|^2.$$

Clearly, a strictly monotone map is monotone but the converse is not true. For example, the map $F : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $F(x_1, x_2) = (2x_1, 0)$ is monotone but not strictly monotone, as the definition fails at $x = (0, 1)$, $y = (0, 2)$.

Also, every strongly monotone map is strictly monotone but the converse is not true. For instance, the map $F : \mathbb{R} \to \mathbb{R}$ defined by

$$F(x) = \begin{cases} 1 + x^2, & \text{if } x \geq 0, \\ 1 - x^2, & \text{if } x < 0, \end{cases}$$

is strictly monotone but it is not strongly monotone. We observe that if we restrict the domain of the function F defined above to $[1, \infty)$, then it is strongly monotone with modulus $\sigma = 2$.

In Theorem 1.17, it was observed that a differentiable function f is convex on an open convex set K if and only if for every $x, y \in K$,

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0,$$

and is strictly convex on an open convex set K if for every $x, y \in K$, $x \neq y$,

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle > 0.$$

Hence we can restate Theorem 1.17 as follows:

Theorem 4.1. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A differentiable function $f : K \to \mathbb{R}$ is

(a) convex if and only if its gradient ∇f is monotone;

(b) strictly convex if and only if its gradient ∇f is strictly monotone.

Analogous to Theorem 4.1, we have the following theorem.

Theorem 4.2. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A differentiable function $f : K \to \mathbb{R}$ is strongly convex with modulus $\rho > 0$ if and only if its gradient ∇f is strongly monotone with modulus $\sigma = 2\rho$.

Proof. If f is strongly convex, then for all $x, y \in K$, we have

$$\langle \nabla f(x), y - x\rangle + \rho\|y - x\|^2 \le f(y) - f(x).$$

Interchanging the roles of x and y, we have

$$\langle \nabla f(y), x - y\rangle + \rho\|x - y\|^2 \le f(x) - f(y).$$

By adding the above inequalities, we obtain the conclusion with $\sigma = 2\rho$.

Conversely, assume that ∇f is strongly monotone with modulus σ, then for all $x, y \in K$, $x \neq y$, we have

$$\langle \nabla f(y) - \nabla f(x), y - x\rangle \ge \sigma\|y - x\|^2.$$

Define a function $h : [0, 1] \to \mathbb{R}$ as

$$h(\lambda) = f((1 - \lambda)x + \lambda y), \quad \text{for all } \lambda \in [0, 1].$$

Then, we have

$$f(y) - f(x) = h(1) - h(0) \text{ and } h'(\lambda) = \langle \nabla f((1 - \lambda)x + \lambda y), y - x\rangle.$$

Since ∇f is strongly monotone with modulus σ, we have for all $\lambda \in \,]0, 1]$,

$$\langle \nabla f((1 - \lambda)x + \lambda y) - \nabla f(x), y - x\rangle \ge \sigma\lambda\|x - y\|^2.$$

Thus,

$$h'(\lambda) \ge \langle \nabla f(x), y - x\rangle + \sigma\lambda\|x - y\|^2.$$

On integrating with respect to λ over $[0, 1]$, we have

$$h(1) - h(0) \ge \langle \nabla f(x), y - x\rangle + (\sigma/2)\|x - y\|^2,$$

which implies that f is strongly convex with $\rho = \sigma/2$. □

Next we define generalized monotone maps and relate generalized convexity with generalized monotonicity of its gradient function. Karamardian [124] introduced the concept of pseudomonotone maps whereas the notions of strict pseudomonotonicity and quasimonotonicity were introduced by Hassouni [101] and independently by Karamardian and Schaible [125].

Definition 4.2. Let K be a nonempty subset of $\mathbb{R}^n$. A map $F : K \to \mathbb{R}^n$ is said to be

(a) *quasimonotone* if for all $x, y \in K$, $x \neq y$, we have

$$\langle F(x), y - x \rangle > 0 \quad \Rightarrow \quad \langle F(y), y - x \rangle \geq 0;$$

(b) *pseudomonotone* if for all $x, y \in K$, $x \neq y$, we have

$$\langle F(x), y - x \rangle \geq 0 \quad \Rightarrow \quad \langle F(y), y - x \rangle \geq 0;$$

(c) *strictly pseudomonotone* if for all $x, y \in K$, $x \neq y$, we have

$$\langle F(x), y - x \rangle \geq 0 \quad \Rightarrow \quad \langle F(y), y - x \rangle > 0.$$

Clearly, a (strictly) monotone map is (strictly) pseudomonotone but the converse is not true. For example, the map $F : \mathbb{R} \to \mathbb{R}$ defined by $F(x) = xe^{-x^2}$ is (strictly) pseudomonotone but the definition of (strict) monotonicity fails at $x = 1$, $y = 2$.

A strictly pseudomonotone map is pseudomonotone but the converse is not true. For instance, the map $F : \mathbb{R} \to \mathbb{R}$ defined as $F(x) = \max\{x, 0\}$ is pseudomonotone but it is not strictly pseudomonotone.

We also observe that every pseudomonotone map is quasimonotone but the converse is not true as the map $F(x) = x^2$ is quasimonotone on $\mathbb{R}$ but it is not pseudomonotone on $\mathbb{R}$.

The following result gives a characterization of quasiconvex functions.

Theorem 4.3. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A differentiable function $f : K \to \mathbb{R}^n$ is quasiconvex if and only if its gradient ∇f is quasimonotone.

Proof. Assume that f is quasiconvex but ∇f is not quasimonotone. Then, there exist $x, y \in K$, $x \neq y$, such that

$$\langle \nabla f(x), y - x \rangle > 0 \quad \text{and} \quad \langle \nabla f(y), x - y \rangle > 0.$$

Since f is quasiconvex, the above inequalities respectively lead to $f(y) > f(x)$ and $f(x) > f(y)$, which are contradictory to each other.

Conversely, assume on the contrary that ∇f is quasimonotone but f is not quasiconvex. Then, there exist $x, y \in K$ such that $f(y) \leq f(x)$ and $f(z) > f(x)$, where $z = (1-\lambda)x + \lambda y$ for some $\lambda \in \,]0, 1[$. By the mean value theorem, there exist $u = (1 - \alpha)z + \alpha x$, $v = (1 - \beta)z + \beta y$ for some $\alpha, \beta \in \,]0, 1[$ such that

$$f(z) - f(x) = \langle \nabla f(u), z - x \rangle,$$

and

$$f(z) - f(y) = \langle \nabla f(v), z - y \rangle.$$

Since $f(z) > f(x) \geq f(y)$, we have

$$\langle \nabla f(u), u - x \rangle > 0, \qquad \langle \nabla f(v), v - y \rangle > 0,$$

which leads to

$$\langle \nabla f(u), v - u \rangle > 0, \qquad \langle \nabla f(v), u - v \rangle > 0.$$

This contradicts the quasimonotonicity of ∇f. □

As expected we have a similar characterization for (strict) pseudoconvexity of a function in terms of the (strict) pseudomonotonicity of the gradient map.

Theorem 4.4. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A differentiable function $f : K \to \mathbb{R}^n$ is

(a) pseudoconvex if and only if its gradient ∇f is pseudomonotone;

(b) strictly pseudoconvex if and only if its gradient ∇f is strictly pseudomonotone.

Proof. (a) Assume that f is pseudoconvex but ∇f is not pseudomonotone. Then, there exist $x, y \in K$, $x \neq y$, such that

$$\langle \nabla f(x), y - x \rangle \geq 0 \quad \text{and} \quad \langle \nabla f(y), x - y \rangle > 0.$$

Since f is pseudoconvex, the first inequality leads to $f(y) \geq f(x)$, and the second one leads to $f(x) > f(y)$ as every pseudoconvex function is quasiconvex. We thus arrive at a contradiction as the two conclusions are contradictory to each other.

Conversely, assume on the contrary that ∇f is pseudomonotone but f is not pseudoconvex. Then, there exist $x, y \in K$ such that

$$\langle \nabla f(x), y - x \rangle \geq 0 \quad \text{and} \quad f(y) < f(x).$$

By the mean value theorem, there exists $z = (1 - \lambda)x + \lambda y$ for some $\lambda \in\ ]0, 1[$ such that

$$f(y) - f(x) = \langle \nabla f(z), y - x \rangle = (1/\lambda)\langle \nabla f(z), z - x \rangle.$$

Since $f(y) < f(x)$, it follows that $\langle \nabla f(z), z - x \rangle < 0$. Now by pseudomonotonicity of ∇f, we have

$$\langle \nabla f(x), z - x \rangle < 0, \text{ that is, } \langle \nabla f(x), y - x \rangle < 0,$$

which leads to a contradiction.

(b) Proof follows on similar lines. □

As observed earlier every pseudomonotone map is quasimonotone. We now give sufficient conditions for a quasimonotone map to be pseudomonotone.

Theorem 4.5. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A quasimonotone map $F : K \to \mathbb{R}^n$ is pseudomonotone if F is continuous and for all $x, y \in K$ with $F(x) \neq \mathbf{0}$ implies that there exists $\lambda \in \ ]0, 1[$ such that

$$\langle F((1-\lambda)x + \lambda y), y - x\rangle \geq 0.$$

Proof. Assume on the contrary that F is not pseudomonotone. Then, there exist $x, y \in K$, $x \neq y$, such that

$$\langle F(x), y - x\rangle \geq 0 \quad \text{and} \quad \langle F(y), y - x\rangle < 0.$$

Clearly, $\hat{\lambda} = \sup\{\lambda \in [0, 1] : \langle F((1-\lambda)x + \lambda y), y - x\rangle \geq 0\}$ is attained in $[0, 1[$ as F is continuous. If we let $u = x + \hat{\lambda}(y - x)$ and $z = u + (1 - \hat{\lambda})(y - x)$, then

$$\langle F(u), z - u\rangle = 0, \tag{4.1}$$

and

$$\langle F(u + \lambda(z - u)), z - u\rangle < 0, \quad \text{for all } \lambda \in \ ]0, 1]. \tag{4.2}$$

Hence, by the hypothesis, we have $F(u) \neq \mathbf{0}$. Choose v such that

$$\langle F(u), v - u\rangle > 0.$$

By taking $\lambda = 1$ in inequality (4.2), we have $\langle F(z), z - u\rangle < 0$. By continuity of F, there exists $t > 0$ such that

$$\langle F(z + t(v - u)), z - u + t(v - u)\rangle < 0,$$

which by quasimonotonicity of F implies that

$$\langle F(u), z - u + t(v - u)\rangle \leq 0.$$

Using equation (4.1), we obtain

$$\langle F(u), v - u\rangle \leq 0,$$

which contradicts our choice of v. □

Corollary 4.1. Let K be a nonempty open convex subset of $\mathbb{R}^n$. If $F : K \to \mathbb{R}^n$ is a continuous map such that $F(x) \neq \mathbf{0}$ for all $x \in K$, then F is pseudomonotone if and only if F is quasimonotone.

The following concepts of strict and semistrict quasimonotonicity were introduced by Blum and Oettli [35].

Definition 4.3. Let K be a nonempty convex subset of $\mathbb{R}^n$. A map $F : K \to \mathbb{R}^n$ is said to be

(a) *strictly quasimonotone* if F is quasimonotone and for all $x, y \in K$, $x \neq y$, there exists $z \in \ ]x, y[$ such that $\langle F(z), y - x\rangle \neq 0$;

(b) *semistrictly quasimonotone* if F is quasimonotone and for $x, y \in K$, $x \neq y$,

$$\langle F(x), y - x\rangle > 0 \Rightarrow \text{ there exists } z \in \,](x+y)/2, y[$$
$$\text{such that } \langle F(z), y - x\rangle > 0.$$

Obviously, a pseudomonotone map is semistrictly quasimonotone, and a strictly pseudomonotone map is strictly quasimonotone.

The following diagram gives the relationship among different classes of monotone maps defined above.

$$\begin{array}{ccccc}
\text{Strong monotonicity} & & & & \\
\Downarrow & & & & \\
\text{Strict monotonicity} & \Rightarrow & \text{Monotonicity} & & \\
\Downarrow & & \Downarrow & & \\
\text{Strict pseudomonotonicity} & \Rightarrow & \text{Pseudomonotonicity} & \Rightarrow & \text{Semistrict} \\
\Downarrow & & \Downarrow & & \text{Quasimonotonicity} \\
\text{Strict quasimonotonicity} & \Rightarrow & \text{Quasimonotonicity} & & \\
\Downarrow & & & & \\
\text{Semistrict quasimonotonicity} & & & &
\end{array}$$

Theorem 4.6. Let K be a nonempty convex subset of $\mathbb{R}^n$. If $F : K \to \mathbb{R}^n$ is strictly quasimonotone, then it is also semistrictly quasimonotone.

Proof. If $\langle F(x), y - x\rangle > 0$ for all $x, y \in K$, $x \neq y$, then

$$\langle F(x), z - x\rangle > 0, \quad \text{for all } z \in \,]x, y[.$$

Since F is quasimonotone, we have $\langle F(z), z - x\rangle \geq 0$ which implies that

$$\langle F(z), y - x\rangle \geq 0, \quad \text{for all } z \in \,]x, y[.$$

Since F is strictly quasimonotone, there exists $\hat{z} \in \,](x+y)/2, y[$ such that $\langle F(\hat{z}), y - x\rangle \neq 0$. Thus, we have $\langle F(\hat{z}), y - x\rangle > 0$, that is, F is semistrictly quasimonotone. □

We now link strict quasiconvexity of a function with strict quasimonotonicity of its gradient.

Theorem 4.7. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A differentiable function $f : K \to \mathbb{R}$ is strictly quasiconvex if and only if its gradient ∇f is strictly quasimonotone.

Proof. Assume that f is strictly quasiconvex, then it is quasiconvex. By Theorem 4.3, it follows that ∇f is quasimonotone. For all $x, y \in K$, $x \neq y$, consider the function $h(t) = f((1-t)x + ty)$ defined on $[0,1]$. Since f is strictly quasiconvex, it follows that h is not constant on $[0,1]$ which implies that $h'(t) \neq 0$ for at least one $t \in \ ]0,1[$, where $h'(t) = \langle \nabla f((1-t)x+ty), y-x \rangle$. This implies that there exists $z \in \ ]x,y[$ such that $\langle \nabla f(z), y-x \rangle \neq 0$, that is, ∇f is strictly quasimonotone.

Conversely, let ∇f be strictly quasimonotone, then ∇f is quasimonotone which by Theorem 4.3 leads to the fact that f is quasiconvex. Thus, for every $x, y \in K$, $x \neq y$, we have

$$f((1-t)x + ty) \leq \max\{f(x), f(y)\}, \quad \text{for all } t \in \ ]0,1[.$$

Let us assume that $\max\{f(x), f(y)\} = f(y)$, that is, $f(x) \leq f(y)$. It is enough to show that $f(z) < f(y)$ for every $z \in \ ]x,y[$. Suppose on the contrary that $f(z) = f(y)$ for some $z \in \ ]x,y[$. By quasiconvexity of f, it follows that $f((1-t)z + ty) \leq f(z)$ for all $t \in [0,1]$.

We now deal with two cases:

Case 1. Let $f(x) < f(y)$. In this case, we will show that f is constant on $[z,y]$. If it is not so, then there exists $v \in \ ]z,y[$ such that $f(v) < f(z)$. As, $f(x) < f(y) = f(z)$, we have $z \in \ ]x,v[$ and $f(z) > \max\{f(x), f(v)\}$ which contradicts the quasiconvexity of f.

Case 2. Let $f(x) = f(y)$. In this case, we will show that f is constant on $[z,y]$ or $[x,z]$. If it is not so, there exists $u \in \ ]x,z[$ and $v \in \ ]z,y[$ such that $f(u) < f(z)$ and $f(v) < f(z)$. Hence, we have

$$z \in \ ]u,v[\quad \text{and} \quad f(z) > \max\{f(u), f(v)\},$$

which contradicts the quasiconvexity of f.

If f is constant on $[z,y]$, then the function $h(t) = f((1-t)z+ty)$ is constant on $[0,1]$, and hence, $\langle \nabla f(w), z-y \rangle = 0$ for all $w \in \ ]z,y[$, which contradicts the strict quasimonotonicity of f. Similarly, we arrive at a contradiction to strict quasimonotonicity of f if it is constant on $]x,z[$. □

The next theorem relates semistrict quasiconvexity of a function with semistrict quasimonotonicity of its gradient.

Theorem 4.8. Let K be a nonempty open convex subset of $\mathbb{R}^n$. A differentiable function $f : K \to \mathbb{R}$ is semistrictly quasiconvex if and only if its gradient ∇f is semistrictly quasimonotone.

Proof. Assume that f is semistrictly quasiconvex, then it is quasiconvex. By Theorem 4.3, it follows that ∇f is quasimonotone. If for all $x, y \in K$, $x \neq y$, we have $\langle \nabla f(x), y-x \rangle > 0$, then

$$\langle \nabla f(x), z-x \rangle \geq 0, \quad \text{for all } z \in [x,y].$$

By quasimonotonicity of ∇f, we have $\langle \nabla f(z), z-x \rangle \geq 0$, which implies

$$\langle \nabla f(z), y-x \rangle \geq 0, \quad \text{for every } z \in [x,y].$$

Consider the function $h(t) = f((1-t)x + ty)$ defined on $[0,1]$. Since f is semistrictly quasiconvex, it follows that $h'(t) \geq 0$, that is, h is increasing on $[0,1]$. We need to show that there exists $z \in \,](x+y)/2, y[$ such that $\langle \nabla f(z), y - x\rangle > 0$, that is, h is not constant on $[1/2, 1]$. Suppose, on the contrary, h is constant on $[1/2,1]$. Since $h'(0) > 0$, it follows that $h(1) > h(0)$. Thus, we have $f(x) < f(y)$ and $f(u) = f(y)$ for every $u \in \,](x+y)/2, y[$, which contradicts semistrict quasiconvexity of f.

Conversely, let ∇f be semistrictly quasimonotone, hence ∇f is quasimonotone, which by Theorem 4.3 leads to the fact that f is quasiconvex. Suppose that f is not semistrictly quasiconvex, then there exist $x, y \in K$ and $z = (1-\alpha)x + \alpha y$ for some $\alpha \in \,]0,1[$ such that $f(x) < f(z) = f(y)$. In this case, we will show that f is constant on $[z, y]$. If it is not so, then there exists $v \in \,]z, y[$ such that $f(v) < f(z)$. As, $f(x) < f(y) = f(z)$, we have $z \in \,]x, v[$ and $f(z) > \max\{f(x), f(v)\}$, which contradicts quasiconvexity of f. Let $h(t) = f((1-t)x + ty)$, then $h(0) < h(1)$. By the mean value theorem there exists $t_1 \in \,]0,1[$ such that $h'(t_1) > 0$. By semistrict quasimonotonicity of ∇f, there exists $t_2 \in \,](1+t_1)/2, 1[$ such that $h'(t_2) > 0$. Proceeding in this manner, we get a sequence $\{t_k\}$ such that $h'(t_k) > 0$ and $t_{k+1} \in \,](1+t_k)/2, 1[$. Clearly, $t_k \to 1$, and hence, $t_k > \alpha$ for some k. Since h is constant on $[\alpha, 1]$, we have $h'(t_k) = 0$, which is a contradiction. □

4.3 Nonsmooth Monotonicity and Generalized Monotonicity in Terms of a Bifunction

In the last section convex and generalized convex functions were characterized by the monotonicity and the generalized monotonicity of the gradient maps, respectively. For nonsmooth functions, a natural extension of these results would be to replace the gradient maps by generalized directional derivatives.

The following definition uses a unified approach to define monotonicity and its generalizations in terms of a bifunction $h(x; d)$.

Definition 4.4. Let K be a nonempty convex subset of $\mathbb{R}^n$. A bifunction $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be

(a) *monotone* if for all $x, y \in K$, $x \neq y$,

$$h(x; y - x) + h(y; x - y) \leq 0; \tag{4.3}$$

(b) *strictly monotone* if the strict inequality holds in inequality (4.3) for all $x, y \in K$, $x \neq y$;

(c) *strongly monotone* if for all $x, y \in K$, $x \neq y$, there exists $\sigma > 0$ such that

$$h(x; y - x) + h(y; x - y) \leq -\sigma \|y - x\|^2;$$

(d) *pseudomonotone* if for all $x, y \in K$, $x \neq y$,

$$h(x; y - x) \geq 0 \quad \Rightarrow \quad h(y; x - y) \leq 0;$$

(e) *strictly pseudomonotone* if for all $x, y \in K$, $x \neq y$,

$$h(x; y - x) \geq 0 \quad \Rightarrow \quad h(y; x - y) < 0;$$

(f) *quasimonotone* if for all $x, y \in K$, $x \neq y$,

$$h(x; y - x) > 0 \quad \Rightarrow \quad h(y; x - y) \leq 0;$$

(g) *strictly quasimonotone* if it is quasimonotone and for all $x, y \in K$, $x \neq y$, there exists $z \in \ ]x, y[$ such that either $h(z; y - x) \neq 0$ or $h(z; x - y) \neq 0$;

(h) *semistrictly quasimonotone* if it is quasimonotone and for all $x, y \in K$, $x \neq y$,

$$h(x; y - x) > 0 \quad \Rightarrow \quad \text{there exists } z \in \ ](x + y)/2, y[\quad \text{such that } h(z, y - x) > 0.$$

Clearly, every strongly monotone bifunction is strictly monotone, and every strictly monotone bifunction is monotone.

The bifunction $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined as $h(x; d) = \min\{0, xd^2\}$ is monotone but not strictly monotone, as the definition fails at $x = 1$ and $y = 2$.

The bifunction $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given by $h(x; d) = \min\{0, d\}$ is strictly monotone but not strongly monotone. For $x = 0$ and $y > 0$, the condition $h(x; y - x) + h(y; x - y) \leq -\sigma \|y - x\|^2$ reduces to $y \geq \sigma \|y\|^2$ from which it follows that $\sigma \to 0$ as $y \to \infty$.

However, the bifunction $h : (-\infty, -1] \times \mathbb{R} \to \mathbb{R}$ given by $h(x; d) = xd^2$ is strongly monotone with $\sigma = 2$.

It is obvious that every monotone bifunction is pseudomonotone, every strictly pseudomonotone bifunction is pseudomonotone, and that every pseudomonotone bifunction is quasimonotone. The inclusions between these classes of bifunctions are strict. For instance, the bifunction $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by

$$h(x; d) = \begin{cases} 2xd, & \text{if } x \geq 1, \\ 0, & \text{if } 0 \leq x < 1, \\ d, & \text{if } x < 0, \end{cases}$$

is pseudomonotone but is not strictly pseudomonotone at the points $x = 0$ and $y = 1/2$. Moreover, this bifunction is not monotone as the definition of monotonicity fails at $x = -1$ and $y = 1/2$.

Also, the bifunction $h(x; d) = x^2 d^3$ defined on $\mathbb{R} \times \mathbb{R}$ is quasimonotone but pseudomonotonicity is violated at $x = 0$ and $y = -1$.

Remark 4.1. Every pseudomonotone bifunction is semistrictly quasimonotone provided it is positively homogeneous in the second argument, for if $h(x; y-x) > 0$ for some $x, y \in K$, then by positive homogeneity of h, it follows that $h(x; z-x) > 0$ for all $z \in \,]x, y[$ and from pseudomonotonicity of h, it follows that $h(z; x-z) < 0$ or $h(z; x-y) < 0$.

Now onward, throughout this chapter, we assume K to be a nonempty convex subset of $\mathbb{R}^n$.

Remark 4.2. If $g : K \times \mathbb{R}^n \to \mathbb{R}$ is a bifunction such that $g(x; d) \leq h(x; d)$ for all $x \in K$, $d \in \mathbb{R}^n$ and if h is monotone (respectively, strictly monotone, strongly monotone, pseudomonotone, strictly pseudomonotone, quasimonotone), then g is also monotone (respectively, strictly monotone, strongly monotone, pseudomonotone, strictly pseudomonotone, quasimonotone). Since $f_{DH}(x; d) \leq f^R(x; d)$, it follows that if the Rockafellar directional derivative $f^R(x; d)$ is monotone, then the Dini-Hadamard lower directional derivative $f_{DH}(x; d)$ is also monotone. Similarly if the Clarke derivative $f^C(x; d)$ is monotone, then the generalized derivatives $f_{DH}(x; d)$, $f_D(x; d)$, $f^D(x; d)$ are also monotone as

$$f_{DH}(x; d) \leq f_D(x; d) \leq f^D(x; d) \leq f^C(x; d).$$

Theorem 4.9. If $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is strictly quasimonotone and $g : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is a bifunction such that $g(x; d) \leq h(x; d)$ for all $(x, d) \in K \times \mathbb{R}^n$, then g is strictly quasimonotone provided h is odd in the second argument.

Proof. Since h is strictly quasimonotone, it is quasimonotone, and hence, it follows that g is quasimonotone. By the definition of strict quasimonotonicity and the oddness of h in the second argument, it follows that for any $x, y \in K$, $x \neq y$, there exists $z \in \,]x, y[$ such that $h(z; y-x) \neq 0$. If $h(z; y-x) < 0$, then $g(z; y-x) < 0$ and the proof is completed. If $h(z; y-x) > 0$, then from oddness of h, it follows that $h(z; x-y) < 0$, and then $g(z; x-y) < 0$. □

The oddness assumption in Theorem 4.9 cannot be relaxed.

Example 4.1. [33] Let $h : [-1, 0] \times \mathbb{R} \to \mathbb{R}$ be defined as

$$h(x; d) = \begin{cases} -d, & \text{if } x = 0,\ d < 0, \\ 0, & \text{if } x \leq 0,\ d \geq 0, \\ d, & \text{if } x < 0,\ d < 0. \end{cases}$$

Then, h is strictly quasimonotone but fails to be semistrictly quasimonotone because for $x = 0$, $y = -1$, $h(x; y-x) > 0$ but for all $z \in \,]-1, -1/2[$, we observe that $h(z; y-x) < 0$.

Theorem 4.10. If $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is strictly pseudomonotone, then it is strictly quasimonotone.

Proof. From strict pseudomonotonicity of h, it follows that h is pseudomonotone, and hence, quasimonotone. Suppose, on the contrary, there exist $x, y \in K$, $x \neq y$, such that

$$h(z; y-x) = 0 \text{ and } h(z; x-y) = 0, \quad \text{for all } z \in \;]x, y[.$$

Let $u, v \in \;]x, y[$, then we have,

$$h(u; v-u) = 0, \quad h(u; u-v) = 0, \quad h(v; v-u) = 0 \quad \text{and } h(v; u-v) = 0.$$

Since $h(u; v-u) = 0$ and h is strictly pseudomonotone, we have

$$h(v; u-v) < 0,$$

which is a contradiction. □

Next example shows that the converse of the above theorem does not hold.

Example 4.2. Let $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined as

$$h(x; d) = \begin{cases} xd, & \text{if } x < 0, \\ 0, & \text{if } x = 0,\ d \geq 0, \\ -d, & \text{if } x = 0,\ d < 0 \text{ or } x > 0. \end{cases}$$

Then h is strictly quasimonotone but it is not strictly pseudomonotone because for $x = 2$ and $y = 0$, $h(x; y-x) > 0$ and $h(y; x-y) = 0$.

We now define the concept of proper quasimonotonicity for a bifunction h. This notion plays a crucial role in establishing the existence of a solution to Minty variational inequality problem defined in terms of a bifunction (see Chapter 6).

Definition 4.5. A bifunction $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be *properly quasimonotone* if for every $y_1, y_2, \ldots, y_m \in K$ and $x \in \text{co}(\{y_1, y_2, \ldots, y_m\})$ there exists $i \in \{1, 2, \ldots, m\}$ such that $h(y_i; x - y_i) \leq 0$.

Example 4.3. Consider the bifunction $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by

$$h(x; d) = d - 1.$$

Then, we assert that h is properly quasimonotone. Suppose, on the contrary, that there exist $y_1, y_2, \ldots, y_m \in K$ and $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$ with $\sum_{i=1}^m \lambda_i = 1$ and $x = \sum_{i=1}^m \lambda_i y_i$ such that

$$h(y_i; x - y_i) > 0, \quad \text{for all } i = 1, 2, \ldots, m.$$

This implies that $x - y_i - 1 > 0$ for all $i = 1, 2, \ldots, m$, which on multiplying by $\lambda_i \geq 0$ and summing over i yields a contradiction. Thus, h is properly quasimonotone.

To show that under certain assumptions a pseudomonotone bifunction is properly quasimonotone, we define the notion of proper suboddness.

Definition 4.6. A function $g : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be proper subodd if

$$g(d_1) + g(d_2) + \cdots + g(d_m) \geq 0,$$

where $d_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, m$ with $\sum_{i=1}^{m} d_i = 0$ for any $m \geq 2$.

When $m = 2$, Definition 4.6 reduces to the definition of suboddness of a function (see Remark 1.2(a)).

Theorem 4.11. If $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is a pseudomonotone bifunction that is both positively homogeneous and proper subodd in the second argument, then h is properly quasimonotone.

Proof. Suppose, on the contrary, there exist $y_1, y_2, \ldots, y_m \in K$ and $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$ with $\sum_{i=1}^{m} \lambda_i = 1$ and $x = \sum_{i=1}^{m} \lambda_i y_i$ such that

$$h(y_i; x - y_i) > 0, \quad \text{for all } i = 1, 2, \ldots, m.$$

By using the pseudomonotonicity of h, we obtain

$$h(x; y_i - x) < 0, \quad \text{for all } i = 1, 2, \ldots, m,$$

which implies

$$\sum_{i=1}^{m} h(x; \lambda_i(y_i - x)) < 0, \qquad \text{where } \sum_{i=1}^{m} \lambda_i(y_i - x) = 0,$$

which is a contradiction to the proper suboddness of h. Thus, h is properly quasimonotone. □

The next theorem establishes that a properly quasimonotone bifunction h is quasimonotone under positive homogeneity assumption.

Theorem 4.12. If $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is positively homogeneous in the second argument and is properly quasimonotone, then h is quasimonotone.

Proof. Let $x, y \in K$ be such that $h(x; y - x) > 0$. Let $y_\lambda = \lambda x + (1 - \lambda)y$ for all $\lambda \in \ ]0, 1[$. Then,

$$h(x; y_\lambda - x) = h(x; (1 - \lambda)(y - x)) = (1 - \lambda)h(x; y - x) > 0.$$

Since h is properly quasimonotone and $y_\lambda \in \mathrm{co}(\{x, y\})$, from the above relation, it follows that

$$h(y; \lambda(x - y)) = h(y; y_\lambda - y) \leq 0.$$

Thus, by positive homogeneity of h, we obtain $h(y; x - y) \leq 0$. □

Remark 4.3. Theorem 4.12 may not hold in the absence of positive homogeneity of h. For example, consider the bifunction $h : K \times \mathbb{R} \to \mathbb{R}$ where $K = [0, 1]$, defined as

$$h(x; d) = \begin{cases} 1, & \text{if } x(d + x) = 0. \\ 0, & \text{otherwise.} \end{cases}$$

Then, h is properly quasimonotone but it is not quasimonotone because for $x = 1$ and $y = 0$, $h(x; y - x) > 0$ and $h(y; x - y) > 0$.

Remark 4.4. The converse of Theorem 4.12 may not hold. Consider the bifunction $h : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ defined by

$$h(x; d) = \begin{cases} -d_1 - d_2, & \text{if } x = (0, 1), \\ d_1, & \text{if } x = (0, 0), \\ d_2, & \text{if } x = (1, 0), \\ 0, & \text{otherwise.} \end{cases}$$

Then, it can be verified that h is quasimonotone but not properly quasimonotone because for $y_1 = (0, 1)$, $y_2 = (1, 0)$, $y_3 = (0, 0)$, and $\lambda_i = 1/3$ for $i = 1, 2, 3$, we have $x = (1/3, 1/3) \in \text{co}(\{y_1, y_2, y_3\})$ and $h(y_i; x - y_i) > 0$ for all $i = 1, 2, 3$.

Theorem 4.13. If $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is a strictly quasimonotone bifunction that is positively homogeneous and subodd in the second argument, then h is semistrictly quasimonotone.

Proof. Let $x, y \in K$ be such that $h(x; y - x) > 0$. Then by positive homogeneity of h in the second argument, it follows that $h(x; z - x) > 0$ for all $z \in \;]x, y[$ and quasimonotonicity of h then yields $h(z; x - z) \leq 0$. From the strict quasimonotonicity of h, we conclude the existence of $\bar{z} \in ((x + y)/2, y)$ such that $h(\bar{z}; x - y) \neq 0$, and therefore, $h(\bar{z}; x - \bar{z}) \neq 0$. This leads to $h(\bar{z}; x - \bar{z}) < 0$, that is, $h(\bar{z}; x - y) < 0$, and the suboddness of h thereby implies that h is semistrictly quasimonotone. $\square$

The suboddness assumption in Theorem 4.13 cannot be relaxed.

Example 4.4. Let $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined as

$$h(x; d) = \begin{cases} -d, & \text{if } x = 0, \; d < 0, \\ 0, & \text{if } x \leq 0, \; d \geq 0, \\ d, & \text{if } x < 0, \; d < 0. \end{cases}$$

Then h is strictly quasimonotone but fails to be semistrictly quasimonotone because for $x = 0, y = -1$, $h(x; y - x) > 0$ but for all $z \in \;]-1, -1/2[$, we observe that $h(z; y - x) < 0$.

4.4 Relation between Nonsmooth Monotonicity and Nonsmooth Convexity

In this section, we study the relations of nonsmooth convex and nonsmooth generalized convex functions defined in terms of bifunctions with the corresponding notions of nonsmooth monotonicity and nonsmooth generalized monotonicity, respectively. Sach and Penot [186] established the following result.

Theorem 4.14. If $f : K \to \mathbb{R}$ is radially upper semicontinuous, $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is a bifunction satisfying (3.3) and is both subodd and positively homogeneous in the second argument, then the following statements are equivalent:

(a) f is convex;

(b) f is h-convex;

(c) h is monotone.

Proof. (a) $\Leftrightarrow$ (b) Follows from Theorem 3.1.

(b) $\Rightarrow$ (c) Since f is h-convex, we have for all $x, y \in K$, $x \neq y$,

$$h(x; y-x) \leq f(y) - f(x). \tag{4.4}$$

On interchanging the roles of x and y in inequality (4.4) and adding the resultant inequality with inequality (4.4), we have h is monotone.

(c) $\Rightarrow$ (b) By Diewert's mean value theorem (Theorem 2.11(a)), there exists $\theta \in [0, 1[$ such that for $w = x + \theta(y - x)$, we have

$$f(y) - f(x) \geq f^D(w; y-x).$$

If $\theta = 0$, then using (3.3), we have

$$f(y) - f(x) \geq f^D(x; y-x) \geq h(x; y-x).$$

If $\theta \neq 0$, then using (3.3) and positive homogeneity of h, we have

$$f(y) - f(x) \geq f^D(w; y-x) \geq h(w; y-x) = \theta^{-1} h(w; w-x). \tag{4.5}$$

Since h is subodd in the second argument, we have

$$h(w; w-x) \geq -h(w; x-w), \tag{4.6}$$

and using monotonicity of h, we further get

$$-h(w; x-w) \geq h(x; w-x). \tag{4.7}$$

Using inequalities (4.6) and (4.7) along with inequality (4.5), we have

$$f(y) - f(x) \geq \theta^{-1} h(x; w - x) = h(x; y - x),$$

where the last equality follows from the positive homogeneity of h in the second argument. □

From Theorem 4.14 we can immediately deduce the following result.

Corollary 4.2. If $f : K \to \mathbb{R}$ is radially upper semicontinuous, then f is convex if and only if f_D or f^D is monotone.

Corollary 4.3. If the assumptions of Theorem 4.14 hold, then the following statements hold:

(a) If f is strictly h-convex, then h is strictly monotone.

(b) If f is strongly h-convex with modulus ρ, then h is strongly monotone with $\sigma = 2\rho$.

The following example shows that the suboddness assumption in Theorem 4.14 is imperative.

Example 4.5. Let $K = \mathbb{R}$ and $f : K \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} -1, & \text{if } x \neq 0, \\ 1, & \text{if } x = 0. \end{cases}$$

Then, f is upper semicontinuous and

$$f^D(x; d) = \begin{cases} -\infty, & \text{if } x = 0,\ d \neq 0, \\ 0, & \text{if } x \neq 0 \text{ or } x = 0,\ d = 0. \end{cases}$$

We observe that $h \equiv f^D$ is monotone and f is h-convex but f is not convex.

We have the following variant of Theorem 4.14 for radially lower semicontinuous functions.

Theorem 4.15. Let $f : K \to \mathbb{R}$ be radially lower semicontinuous and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction that satisfies (3.6). If h is positively homogeneous in the second argument and is (strictly) monotone, then f is (strictly) convex.

Proof. Suppose, on the contrary, f is not convex. Then there exist $x, y \in K$, $x \neq y$, and $\lambda \in\ ]0,1[$ such that

$$f(z) > \lambda f(y) + (1 - \lambda) f(x), \tag{4.8}$$

where $z = x + \lambda(y - x)$. By Diewert's mean value theorem (Theorem 2.11(b)), there exist $u \in [x, z[$ and $v \in\ ]z, y]$ such that

$$f_D(u; z - x) \geq f(z) - f(x) \quad \text{and} \quad f_D(v; z - y) \geq f(z) - f(y).$$

Using (4.8), we have

$$f_D(u; z-x) > \lambda(f(y)-f(x)) \quad \text{and} \quad f_D(v; z-y) > (1-\lambda)(f(x)-f(y)).$$

Adding the above two inequalities and employing (3.6), we get

$$h(u; y-x) + h(v; x-y) > 0.$$

Clearly, $v-u = \mu(y-x)$ for some $\mu \in \ ]0,1[$, and hence by positive homogeneity assumption, we get

$$h(u; v-u) + h(v; u-v) > 0,$$

which contradicts the monotonicity of h.

The result for the strict case follows on the same lines. □

The converse of Theorem 4.15 may not be true in general. For example, the function $f(x) = |x|$ is convex and the bifunction $h(x; d) = |d|$, satisfies (3.6) and is positively homogeneous in the second argument but h is not monotone.

The following theorem is a variant of Theorem 4.14 for radially lower semicontinuous functions. Similar characterization for radially lower semicontinuous convex functions in terms of the monotonicity of the Clarke-Rockafellar directional derivative has also been provided by Luc and Swaminathan [151].

Theorem 4.16. Let $f : K \to \mathbb{R}$ be radially lower semicontinuous and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction that satisfies (3.3) and (3.6). If h is both subodd and positively homogeneous in the second argument, then the assertions (a) to (c) of Theorem 4.14 are equivalent.

Proof. The proof of the parts (a) $\Leftrightarrow$ (b) and (b) $\Rightarrow$ (c) is similar to Theorem 4.14. The proof of the part (c) $\Rightarrow$ (a) follows from Theorem 4.15. □

From Theorem 4.16 we can immediately deduce the following result given by Komlósi [134].

Corollary 4.4. If $f : K \to \mathbb{R}$ is radially lower semicontinuous, then it is (strictly) convex if and only if f^D or f_D is (strictly) monotone.

Theorem 4.17. If $f : K \to \mathbb{R}$ is a quasiconvex function and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is a bifunction satisfying (3.3), then h is quasimonotone.

Proof. Assume on the contrary that h is not quasimonotone. Then there exist $x, y \in K$, $x \neq y$ such that

$$h(x; y-x) > 0 \quad \text{and} \quad h(y; x-y) > 0,$$

which on using (3.3) respectively yields that

$$f^D(x; y-x) > 0 \quad \text{and} \quad f^D(y; x-y) > 0.$$

These inequalities imply that there exist $u, v \in \ ]x, y[$ such that

$$f(u) > f(x) \quad \text{and} \quad f(v) > f(y).$$

This implies that

$$\max\{f(u), f(v)\} > \max\{f(x), f(y)\},$$

which contradicts the quasiconvexity of f. □

Corollary 4.5. If $f : K \to \mathbb{R}$ is a quasiconvex function, then both f^D and f_D are quasimonotone.

Theorem 4.18. Let $f : K \to \mathbb{R}$ be a radially upper semicontinuous function, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be both subodd and positively homogeneous in the second argument and satisfy (3.3). If h is quasimonotone, then f is quasiconvex.

Proof. Suppose on the contrary that f is not quasiconvex. Then there exist points $x, y \in K$, $x \neq y$, and $z \in \ ]x, y[$ such that $f(z) > \max\{f(x), f(y)\}$. By Diewert's mean value theorem (Theorem 2.11(a)), there exist $\lambda, \bar{\lambda} \in [0, 1[$ such that for $w = z + \lambda(x - z)$, $\bar{w} = z + \bar{\lambda}(y - z)$ we have

$$0 > f(x) - f(z) \geq f^D(w; x - z) \geq h(w; x - z), \tag{4.9}$$

$$0 > f(y) - f(z) \geq f^D(\bar{w}; y - z) \geq h(\bar{w}; y - z). \tag{4.10}$$

If λ, $\bar{\lambda}$ both are zero then adding inequalities (4.9) and (4.10), we get a contradiction to our assumption that h is subodd in the second argument. So, without loss of generality, we may assume that $\lambda > 0$. Then,

$$\bar{w} - w = \alpha(z - x) = \bar{\alpha}(y - z),$$

for some $\alpha, \bar{\alpha} > 0$. By positive homogeneity of h in the second argument and suboddness of h, from inequalities (4.9) and (4.10), we get

$$\begin{aligned} h(w; \bar{w} - w) = \alpha h(w; z - x) \geq -\alpha h(w; x - z) > 0, \\ h(\bar{w}; w - \bar{w}) = \bar{\alpha} h(\bar{w}; z - y) \geq -\bar{\alpha} h(\bar{w}; y - z) > 0, \end{aligned}$$

which contradicts the quasimonotonicity of h. □

Corollary 4.6. If $f : K \to \mathbb{R}$ is a radially upper semicontinuous function such that f^D or f_D is quasimonotone, then f is quasiconvex.

The following equivalence follows directly from Theorem 4.18 and the definition of quasimonotonicity of the bifunction h.

Theorem 4.19. Let $f : K \to \mathbb{R}$ be a radially upper semicontinuous function and h be a bifunction satisfying (3.3). If h is both subodd and positively homogeneous in the second argument, then the following statements are equivalent:

(a) f is quasiconvex;

(b) f is h-quasiconvex;

(c) h is quasimonotone.

Proof. (a) $\Rightarrow$ (b) follows from Theorem 3.7, (a) $\Rightarrow$ (c) follows from Theorem 4.17, and (c) $\Rightarrow$ (a) follows from Theorem 4.18.

(b) $\Rightarrow$ (c). Suppose h is not quasimonotone. Then there exist $x, y \in K$, $x \neq y$ such that

$$h(x; y-x) > 0 \quad \text{and} \quad h(y; x-y) > 0.$$

Since f is h-quasiconvex, we have $f(y) > f(x)$ and $f(x) > f(y)$, which contradict each other. ☐

We have the following connection between quasimonotonicity of h and quasiconvexity of f when the function f is radially lower semicontinuous and satisfies condition (3.6) instead of condition (3.3).

Theorem 4.20. Let $f : K \to \mathbb{R}$ be a radially lower semicontinuous function and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction satisfying (3.6). If h is quasimonotone, then f is quasiconvex.

Proof. Suppose f is not quasiconvex, then there exist $x, y \in K$, $x \neq y$ such that

$$f(z) > f(x) \quad \text{and} \quad f(z) > f(y), \quad \text{for some } z \in \ ]x, y[.$$

Then by Theorem 2.11(b), there exist $u \in [x, z[$ and $v \in \ ]z, y]$ such that

$$f_D(u; z-x) > 0 \quad \text{and} \quad f_D(v; z-y) > 0.$$

Let $v - u = \lambda(z - x)$, $u - v = \mu(z - y)$ for some $\lambda, \mu > 0$. By positive homogeneity of h in the second argument, we get

$$h(u; v-u) = \lambda h(u; z-x) \geq \lambda f_D(u; z-x) > 0,$$

$$h(v; u-v) = \mu h(v; z-y) \geq \mu f_D(v; z-y) > 0,$$

which contradict the quasimonotonicity of h. ☐

Theorem 4.21. Let $f : K \to \mathbb{R}$ be a radially lower semicontinuous function and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction that satisfies (3.3) and (3.6). If h is positively homogeneous in the second argument, then assertions (a) to (c) of Theorem 4.19 are equivalent.

Proof. (a) $\Rightarrow$ (b) follows from Theorem 3.7. (b) $\Rightarrow$ (c) follows as in the proof Theorem 4.19. (c) $\Rightarrow$ (a) follows from Theorem 4.20. ☐

From Theorem 4.21 we can immediately deduce the following result of Komlósi [134].

Corollary 4.7. If $f : K \to \mathbb{R}$ is radially lower semicontinuous, then f is quasiconvex if and only if f_D or f^D is quasimonotone.

We have a complete characterization of proper quasimonotonicity in terms of quasimonotonicity of the bifunction h if h happens to be a bifunction majorizing the lower Dini derivative of a radially lower semicontinuous function.

Theorem 4.22. If $f : K \to \mathbb{R}$ is radially lower semicontinuous, $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfies (3.3) and is positively homogeneous in the second argument, then h is quasimonotone if and only if it is properly quasimonotone.

Proof. Suppose that h is quasimonotone and let, if possible, $y_1, y_2, \ldots, y_m \in K$ and $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$ with $\sum_{i=1}^m \lambda_i = 1$ and $x = \sum_{i=1}^m \lambda_i y_i$ such that

$$h(y_i; x - y_i) > 0, \quad \text{for all } i = 1, 2, \ldots, m.$$

Since f is radially lower semicontinuous and h is quasimonotone, so by Theorem 4.21, we have that f is quasiconvex. This yields $f(y_i) < f(x)$ for all $i = 1, 2, \ldots, m$. Thereby implying that none of the points $y_1, y_2, \ldots, y_m$ can be a maximum point for f on the convex set $\mathrm{co}(\{y_1, y_2, \ldots, y_m\})$. But this contradicts the well known property of vertex maximum of quasiconvex functions over a polyhedron as given in Theorem 1.28. Therefore, h is properly quasimonotone.

The converse part follows from Theorem 4.12. $\square$

We next obtain a characterization for a radially upper semicontinuous h-pseudoconvex function via the pseudomonotonicity of the bifunction h.

Theorem 4.23. Let $f : K \to \mathbb{R}$ be radially upper semicontinuous and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be subodd and positively homogenous in the second argument that satisfies (3.3). Then, f is h-pseudoconvex if and only if h is pseudomonotone.

Proof. Let f be h-pseudoconvex and $x, y \in K$ be such that $h(x; y - x) > 0$. Then by condition (3.3), it follows that $f^D(x; y - x) > 0$ and this yields the existence of a point $z \in\ ]x, y[$ such that $f(z) > f(x)$. By Theorem 3.9, f is quasiconvex, hence we get, $f(y) \geq f(z) > f(x)$. By the definition of h-pseudoconvexity of f, we get $h(y; x - y) < 0$.

Conversely, suppose that h is pseudomonotone. Let $x, y \in K$ be such that $f(x) > f(y)$. Then by Diewert's mean value theorem (Theorem 2.11), there exists $\lambda \in [0, 1[$ and $w = x + \lambda(y - x)$ such that

$$0 > f(y) - f(x) \geq f^D(w; y - x) \geq h(w; y - x).$$

If $\lambda = 0$, then we have $h(x; y - x) < 0$. If $\lambda > 0$, by positive homogeneity of h, we obtain $h(w; w - x) < 0$, and further employing suboddness of h, we get $h(w; x - w) > 0$. Using the pseudomonotonicity and positive homogeneity of h in the second argument we get $h(x; y - x) < 0$. $\square$

The suboddness assumption in Theorem 4.23 cannot be relaxed as is evident from the following example.

Example 4.6. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(x) = x^3$ and let

$$h(x; d) = \min\{0, 3x^2 d\}.$$

Then, h is pseudomonotone but f is not h-pseudoconvex because for $x = 0$, $y = -2$, $f(y) < f(x)$ and $h(x; y - x) = 0$. Note that h is not subodd as $h(2; 1) + h(2; -1) < 0$.

If in Theorem 4.23 instead of radial upper semicontinuity we assume the function to be radially lower semicontinuous, then the conclusion of the theorem may not hold.

Example 4.7. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} 0, & \text{if } x \neq 0, \\ -1, & \text{if } x = 0. \end{cases}$$

Then for $h(x; d) = 0$, condition (3.3) is trivially satisfied, and h is subodd and pseudomonotone. However, f is not h-pseudoconvex as for $x = 1$, $y = 0$, $f(y) < f(x)$ but $h(x; y - x) = 0$.

Corollary 4.8. Let $f : K \to \mathbb{R}$ be a radially lower semicontinuous function, $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction such that $-h$ is both subodd and positively homogeneous in the second argument and h satisfies condition (3.6). Then, f is h-pseudoconcave if and only if $-h$ is pseudomonotone.

Proof. Assume that f is h-pseudoconcave, then $-f$ is $-h$-pseudoconvex. Since h satisfies condition (3.6), we have

$$-h(x; d) \leq -f_D(x; d) = (-f)^D(x; d), \quad \text{for all } (x; d) \in C \times \mathbb{R}^n.$$

Since f is radially lower semicontinuous, it follows that $-f$ is radially upper semicontinuous. With these observations, the proof can easily be deduced from Theorem 4.23. □

We now obtain an analog of Theorem 4.23 for radially lower semicontinuous functions where h is required to satisfy the condition (3.6).

Theorem 4.24. Let $f : K \to \mathbb{R}$ be radially lower semicontinuous and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction that satisfies both (3.3) and (3.6). If h is both positively homogeneous and subodd in the second argument, then h is pseudomonotone if and only if f is h-pseudoconvex.

Proof. If f is h-pseudoconvex, then by Theorem 3.9, f is h-quasiconvex. Suppose that h is not pseudomonotone, then there exist $x, y \in K$, $x \neq y$ such that

$$h(x; y - x) \geq 0 \quad \text{and} \quad h(y; x - y) > 0.$$

Then by h-pseudoconvexity and h-quasiconvexity of f, it follows that $f(y) \geq f(x)$ and $f(x) > f(y)$, respectively, which are contradictory to each other.

Conversely, suppose that h is pseudomonotone but f is not h-pseudoconvex. Then there exist $x, y \in K$, $x \neq y$ such that

$$h(x; y - x) \geq 0 \quad \text{and} \quad f(y) < f(x).$$

By Diewert's mean value theorem (Theorem 2.11(b)), there exists $z \in [x, y[$ such that

$$f_D(z; x - y) \geq f(x) - f(y) > 0.$$

Using condition (3.6), we then have $h(z; x - y) > 0$. Since $h(x; y - x) \geq 0$, by positive homogeneity of h in the second argument, it follows that $h(x; z - x) \geq 0$, which further implies that $h(z; x - z) \leq 0$ as h is pseudomonotone. Again positive homogeneity assumption yields $h(z; x - y) \leq 0$, which contradicts the inequality $h(z; x - y) > 0$. □

In Theorem 4.24, it is worth noting that condition (3.6) is needed only in the direct part of the proof. A characterization for lower semicontinuous (h-quasiconvex) h-pseudoconvex functions in terms of (quasimonotonicity) pseudomonotonicity of the bifunction h can also be found in Komlósi [132] wherein instead of the suboddness property of h, f is assumed to satisfy nonconstancy property, that is, there does not exist any line segment in K on which f is constant.

We give a characterization for strict pseudoconvexity in terms of strict pseudomonotonicity of the associated bifunction.

Theorem 4.25. Let $f : K \to \mathbb{R}$ be radially lower semicontinuous and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction associated to f such that it satisfies (3.6). If h is positively homogeneous in the second argument, then h is strictly pseudomonotone if and only if f is strictly h-pseudoconvex.

Proof. Suppose that f is strictly h-pseudoconvex and h is not strictly pseudomonotone. Then there exist $x, y \in K$, $x \neq y$ such that

$$h(x; y - x) \geq 0 \quad \text{and} \quad h(y; x - y) \geq 0.$$

Then by strict h-pseudoconvexity of f, it follows that

$$f(y) > f(x) \quad \text{and} \quad f(x) > f(y),$$

which clearly is not possible.

Conversely, suppose that h is strictly pseudomonotone and f is not h-pseudoconvex. Then there exist $x, y \in K$, $x \neq y$ such that

$$h(x; y - x) \geq 0 \quad \text{and} \quad f(y) \leq f(x).$$

By Diewert's mean value theorem (Theorem 2.11(b)), there exists $z \in \,]x, y]$ such that

$$f^D(z; x - y) \geq f(x) - f(y) \geq 0.$$

By using condition (3.6), we have

$$h(z; x - y) \geq 0.$$

Since $h(x; y - x) \geq 0$, by positive homogeneity of h in the second argument, it follows that $h(x; z - x) \geq 0$, which further implies that $h(z; x - z) < 0$ as h is strictly pseudomonotone. Again, by positive homogeneity assumption, we have $h(z; x - y) < 0$, which is a contradiction. □

4.5 Nonsmooth Pseudoaffine Bifunctions and Nonsmooth Pseudolinearity

In this section, we study pseudomonotone bifunction h for which $-h$ is also pseudomonotone. Such bifunctions are used to characterize h-pseudolinear functions.

Definition 4.7. A bifunction $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be *pseudoaffine* if h and $-h$ are both pseudomonotone.

Similar notions are known in literature, for instance, a bifunction $T : K \times K \to \mathbb{R}$ satisfying Definition 4.7 was referred to as *bipseudomonotone map* by Ansari et al. [6] and was called *PPM-map* by Bianchi and Schaible [34] when $h(x; d) = \langle F(x), d \rangle$.

The following is an example of a pseudoaffine bifunction, which is positively homogeneous as well as odd in the second argument.

Example 4.8. The bifunction $h : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} \cup \{\pm\infty\}$ defined by

$$h(x; d) = \begin{cases} \frac{e^{x_1}\left(d_1^2 + d_2^2\right)}{d_1 + d_2} & \text{if } d_1 \neq -d_2, \\ 0, & \text{if } d_1 = -d_2, \end{cases}$$

where $d = (d_1, d_2)$ and $x = (x_1, x_2)$ is pseudoaffine.

We derive necessary and sufficient conditions for a bifunction h to be pseudoaffine. These characterizations will be instrumental in studying the solution sets of variational inequality problems involving such maps.

Theorem 4.26. Let $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be radially continuous in the first argument and be odd as well as positively homogeneous in the second argument. Then, h is pseudoaffine if and only if for all $x, y \in K$,

$$h(x; y - x) = 0 \quad \Rightarrow \quad h(y; x - y) = 0. \tag{4.11}$$

Proof. Suppose that h is pseudoaffine and $x, y \in K$. In view of the positive homogeneity assumption the result is obvious if $x = y$. Let us now suppose $h(x; y - x) = 0$ for $x \neq y$. Pseudomonotonicity of h and $-h$ imply that $h(y; x - y) \leq 0$ and $h(y; x - y) \geq 0$, and hence we have the result.

Conversely, let us assume that h is not pseudomonotone. Then there exist $x, y \in K$, $x \neq y$ such that

$$h(x; y - x) > 0 \quad \text{and} \quad h(y; x - y) \geq 0.$$

Define $g : [0, 1] \to \mathbb{R} \cup \{\pm\infty\}$ by

$$g(\alpha) = h(x + \alpha(y - x); y - x).$$

We have $g(0) > 0$ and

$$g(1) = h(y; y - x) = -h(y; x - y) \leq 0.$$

Since h is radially continuous in the first argument there exists $\bar{\alpha} \in \,]0, 1]$ such that $g(\bar{\alpha}) = 0$, that is,

$$h(z; y - x) = 0, \quad \text{where } z = x + \bar{\alpha}(y - x).$$

By positive homogeneity of h, we have $h(z; z - x) = 0$, which in turn implies using the oddness of h that $h(z; x - z) = 0$. Using (4.11), we obtain $h(x; z - x) = 0$ and positive homogeneity of h leads to $h(x; y - x) = 0$, which is a contradiction.

Similarly, it can be established that $-h$ is pseudomonotone. $\square$

Corollary 4.9. Let $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be radially continuous in the first argument and be odd as well as positively homogeneous in the second argument. Then, h is pseudoaffine if and only if for all $x, y \in K$,

$$h(x; y - x) = 0 \quad \Rightarrow \quad h(\lambda x + (1 - \lambda)y; x - y) = 0, \quad \text{for all } \lambda \in [0, 1]. \tag{4.12}$$

Proof. Let us suppose that h is pseudoaffine and $x, y \in K$. By positive homogeneity of h the result holds trivially if $x = y$. Let $h(x; y - x) = 0$ for $x \neq y$ and let $z = \lambda x + (1 - \lambda)y$ for $\lambda \in [0, 1]$. Since the result is obvious for $\lambda = 1$, we assume that $\lambda \in [0, 1[$. Now using positive homogeneity of h, we have

$$\begin{aligned} 0 &= h(x; y - x) \\ &= h(x; (1 - \lambda)^{-1}(z - x)) \\ &= (1 - \lambda)^{-1} h(x; z - x), \end{aligned}$$

which implies that $h(x; z - x) = 0$. Since h is pseudoaffine, so from Theorem 4.26, we have $h(z; x - z) = 0$, and from positive homogeneity of h it follows that $h(z; x - y) = 0$.

Conversely, by taking $\lambda = 0$ in relation (4.12), condition (4.11) is satisfied, and hence the result follows from Theorem 4.26. $\square$

The following characterization due to Bianchi and Schaible [34] can be directly deduced from Theorem 4.26.

Corollary 4.10. Let $F : K \to \mathbb{R}^n$ be a continuous map. Then, F is pseudoaffine (or a PPM map) if and only if for all $x, y \in K$,

$$\langle F(x), y - x\rangle = 0 \quad \Leftrightarrow \quad \langle F(y), y - x\rangle = 0.$$

We may note that in Theorem 4.26, the oddness assumption is needed only in the sufficiency part. We now give an example to show that this assumption cannot be relaxed.

Example 4.9. Let $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ be defined as

$$h(x; d) = \begin{cases} d, & \text{if } d \geq 0, \\ -2d, & \text{if } d < 0. \end{cases}$$

Then, h is a continuous function of x satisfying the positive homogeneity assumption but h is not odd in the second argument because $h(1; 3)+h(1; -3) \neq 0$. The condition (4.11) is trivially satisfied but h is not pseudoaffine because for $x = 1$ and $y = 4$, $h(x; y - x) > 0$ but $h(y; x - y) \not< 0$.

Using Theorem 4.23 and Corollary 4.8 we have the following characterization for h-pseudolinear function.

Theorem 4.27. Let $f : K \to \mathbb{R}$ be a radially continuous function, $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction that is both odd and positively homogeneous in the second argument and satisfies conditions (3.3) and (3.6). Then, f is h-pseudolinear if and only if h is pseudoaffine.

The next result is a consequence of Theorem 4.26 and Corollary 4.9.

Theorem 4.28. Let $f : K \to \mathbb{R}$ be a radially continuous function, $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction that is both odd and positively homogeneous in the second argument and satisfies conditions (3.3) and (3.6). Moreover, if h is radially continuous in the first argument, then f is h-pseudolinear if and only if one of the following holds:

(i) $h(x; y - x) = 0 \;\Rightarrow\; h(y; x - y) = 0$;

(ii) $h(x; y - x) = 0 \;\Rightarrow\; h(\lambda x + (1 - \lambda)y; x - y) = 0$ for all $\lambda \in [0, 1]$.

4.6 Generalized Monotonicity for Set-Valued Maps

This section deals with the generalized monotonicity for set-valued maps. Characterizations for nondifferentiable functions have been derived via the generalized monotonicity of the subdifferential map.

Definition 4.8. Let K be a nonempty subset of $\mathbb{R}^n$. A set-valued map $F : K \to 2^{\mathbb{R}^n}$ is said to be

(a) *generalized monotone* if for every pair of distinct points $x, y \in K$, and for all $u \in F(x)$, $v \in F(y)$, we have

$$\langle v - u, y - x \rangle \geq 0;$$

(b) *generalized strictly monotone* if for every pair of distinct points $x, y \in K$, and for all $u \in F(x)$, $v \in F(y)$, we have

$$\langle v - u, y - x \rangle > 0;$$

(c) *generalized strongly monotone* if there exists a constant $\sigma > 0$ such that for every pair of distinct points $x, y \in K$, and for all $u \in F(x)$, $v \in F(y)$, we have

$$\langle v - u, y - x \rangle \geq \sigma \|y - x\|^2;$$

(d) *generalized pseudomonotone* if for every pair of distinct points $x, y \in K$, and for all $u \in F(x)$, $v \in F(y)$, we have

$$\langle u, y - x \rangle \geq 0 \quad \Rightarrow \quad \langle v, y - x \rangle \geq 0;$$

(e) *generalized weakly pseudomonotone* [139] if for every pair of distinct points $x, y \in K$, and for all $u \in F(x)$, we have

$$\langle u, y - x \rangle \geq 0 \quad \Rightarrow \quad \langle v, y - x \rangle \geq 0, \quad \text{for some } v \in F(y);$$

(f) *generalized strictly pseudomonotone* if for every pair of distinct points $x, y \in K$, and for all $u \in F(x)$, $v \in F(y)$, we have

$$\langle u, y - x \rangle \geq 0 \quad \Rightarrow \quad \langle v, y - x \rangle > 0;$$

(g) *generalized strongly pseudomonotone* if there exists a constant $\sigma > 0$ such that for every pair of distinct points $x, y \in K$, and for all $u \in F(x)$, $v \in F(y)$, we have

$$\langle u, y - x \rangle \geq 0 \quad \Rightarrow \quad \langle v, y - x \rangle \geq \sigma \|y - x\|^2;$$

(h) *generalized partially relaxed strongly pseudomonotone* if there exists a constant $\beta > 0$ such that for every pair of distinct points $x, y \in K$, and for all $u \in F(x)$, $v \in F(y)$, we have

$$\langle v - u, y - x \rangle \geq -\beta \|y - x\|^2;$$

(i) *generalized quasimonotone* if for every pair of distinct points $x, y \in K$, and for all $u \in F(x)$, $v \in F(y)$, we have

$$\langle u, y - x \rangle > 0 \quad \Rightarrow \quad \langle v, y - x \rangle \geq 0;$$

The following implications follow directly from the definitions.

Generalized strong monotonicity	$\Rightarrow$	Generalized strict monotonicity	$\Rightarrow$	Generalized monotonicity
$\Downarrow$		$\Downarrow$		$\Downarrow$
Generalized strong pseudomonotonicity	$\Rightarrow$	Generalized strict pseudomonotonicity	$\Rightarrow$	Generalized pseudomonotonicity
$\Downarrow$			$\swArrow$	$\Downarrow$
Generalized partially relaxed strong pseudomonotonicity		Generalized weak pseudomonotonicity		Generalized quasimonotonicity

FIGURE 4.1: Relations among different kinds of monotonicities

The following example shows that generalized monotonicity does not imply generalized strict monotonicity.

Example 4.10. Let $K = [0, +\infty[$ and $F : K \to 2^K$ be defined as

$$F(x) = \begin{cases} \{0\}, & \text{if } x \neq 0, \\ [-1, 0], & \text{if } x = 0. \end{cases}$$

Then, F is generalized monotone but not generalized strictly monotone which can easily be shown by taking the points $x = 0$ and $y = 3$ in the definition.

The next example illustrates that the class of generalized strongly monotone maps is strictly contained in the class of generalized strictly monotone maps.

Example 4.11. Let $K = \mathbb{R}$ and $F : K \to 2^K$ be defined as

$$F(x) = \begin{cases} \{x^3 - 1\}, & \text{if } x < 0, \\ [-1, 1], & \text{if } x = 0, \\ \{x + 1\}, & \text{if } x > 0. \end{cases}$$

Then, F is generalized strictly monotone map but not generalized strongly monotone.

The following is an example of a set-valued map that is generalized strictly pseudomonotone but not generalized strictly monotone.

Example 4.12. Let $K = \mathbb{R}$ and $F : K \to 2^K$ be defined as

$$F(x) = \begin{cases} \left\{t : x^2 \leq t \leq 3x^2\right\}, & \text{if } x \geq 0, \\ \left\{t : -2x^2 \leq t \leq -x^2\right\}, & \text{if } x < 0. \end{cases}$$

Then, F is generalized strictly pseudomonotone and by considering the points $y = 1/2$ and $x = 1/3$ we can check that F is not generalized strictly monotone.

The map in the subsequent example shows the implication that generalized pseudomonotonicity implies generalized weak pseudomonotonicity is one way only.

Example 4.13. Let $K = \mathbb{R}$ and $F : K \to 2^K$ be defined by

$$F(x) = \begin{cases} \left\{t : -x^2 \leq t \leq x^2\right\}, & \text{if } x < 0, \\ \{0\}, & \text{if } x = 0, \\ \left\{t : x^2 \leq t \leq 2x^2\right\}, & \text{if } x > 0. \end{cases}$$

Then, F is generalized weakly pseudomonotone. It can be verified by taking the points $x = -1$ and $y = 0$ that generalized pseudomonotonicity gets violated.

We next present an example to show that generalized quasimonotonicity does not imply generalized pseudomonotonicity. This example also demonstrates that generalized pseudomonotonicity does not imply generalized monotonicity.

Example 4.14. Let $K = \mathbb{R}$ and $F : K \to 2^K$ be defined by

$$F(x) = \begin{cases} \left\{t : x^2 \leq t \leq 2x^2\right\}, & \text{if } x \neq 0, \\ \{0\}, & \text{if } x = 0. \end{cases}$$

Then, F is generalized quasimonotone but not generalized pseudomonotone at the points $x = 0$ and $y = 1$.

Finally, we provide an example of a set-valued map that is generalized pseudomonotone but not generalized strictly pseudomonotone.

Example 4.15. Let $K = \mathbb{R}$ and $F : K \to 2^K$ be defined by

$$F(x) = \begin{cases} \{0\}, & \text{if } x \neq 0, \\ \{1, 2\}, & \text{if } x = 0. \end{cases}$$

Then, F is generalized pseudomonotone. By taking $x = 0$ and $y = 1$, it can be seen that generalized strict pseudomonotonicity is violated. Also, it must be noted that the set-valued map F is not generalized monotone.

The following results are due to Fan et al. [74].

Theorem 4.29. Let K be nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipschitz function. Then, the following assertions hold.

(a) f is convex (respectively, strictly convex) if and only if the Clarke subdifferential $\partial^C f$ is a generalized monotone (respectively, generalized strictly monotone) set-valued map;

(b) f is strongly convex with modulus $\rho > 0$ if and only if the Clarke subdifferential $\partial^C f$ is a generalized strongly monotone set-valued map with constant $\sigma = 2\rho$.

Proof. (a) Suppose that f is convex. Then by Theorem 3.18(a), for all $x, y \in K$ and all $\xi \in \partial^C f(x)$ and $\zeta \in \partial^C f(y)$, we have

$$\langle \zeta - \xi, y - x \rangle = \langle \zeta, y - x \rangle - \langle \xi, y - x \rangle \geq f(y) - f(x) - f(y) + f(x) = 0.$$

Hence, $\partial^C f$ is generalized monotone.

Conversely, assume that $\partial^C f$ is a generalized monotone set-valued map. Then by Lebourg's mean value theorem (Theorem 2.23), for $x, y \in K$ with $x \neq y$ and any $\lambda \in \ ]0,1[$, there exist $\lambda_1, \lambda_2 \in \ ]0,1[$ with $0 < \lambda_1 < \lambda < \lambda_2 < 1$, $p \in \partial^C f(u)$, and $q \in \partial^C f(v)$ such that

$$\begin{aligned} f(y) - f(\lambda y + (1-\lambda)x) &= (1-\lambda)\langle p, y - x \rangle, \\ f(x) - f(\lambda y + (1-\lambda)x) &= -\lambda \langle q, y - x \rangle, \end{aligned}$$

where $u = x + \lambda_2(y - x) \in \ ]y, \lambda y + (1-\lambda)x[$, and $v = x + \lambda_1(y - x) \in \ ]\lambda y + (1-\lambda)x, x[$. Using generalized monotonicity of $\partial^C f$, we obtain

$$\langle q - p, v - u \rangle = (\lambda_1 - \lambda_2)\langle q - p, y - x \rangle \geq 0.$$

Therefore, for all $x, y \in K$ with $x \neq y$ and all $\lambda \in \ ]0,1[$, we have

$$f(\lambda y + (1-\lambda)x) \leq \lambda f(y) + (1-\lambda) f(x).$$

This inequality also holds for $x = y$ or $\lambda = 0$ and 1.

(b) Suppose that f is strongly convex with modulus $\rho > 0$. We want to show that $\sigma = 2\rho$. For any given $x, y \in K$, $\xi \in \partial^C f(x)$, and $\zeta \in \partial^C f(y)$, it follows from Theorem 3.18(b) that

$$\begin{aligned} f(y) - f(x) &\geq \langle \xi, y - x \rangle + \rho \|y - x\|^2, \\ f(x) - f(y) &\geq \langle \zeta, x - y \rangle + \rho \|y - x\|^2. \end{aligned}$$

Proceeding to the next step, we have

$$\langle \zeta - \xi, y - x \rangle \geq 2\rho \|y - x\|^2.$$

Therefore, $\sigma = 2\rho$.

Conversely, let $\partial^C f$ be generalized strongly monotone with constant $\sigma > 0$. Assume to the contrary that f is not strongly convex. Then for all $\rho > 0$, there exist $x, y \in K$ with $x \neq y$ and $\xi \in \partial^C f(x)$ such that

$$f(y) - f(x) \leq \langle \xi, y - x \rangle + \rho \|y - x\|^2.$$

Then by Lebourg's mean value theorem (Theorem 2.23), there exist $\lambda \in \,]0, 1[$ and $\gamma \in \partial^C f(u)$ such that

$$\langle \gamma - \xi, y - x \rangle \leq \rho \|y - x\|^2,$$

where $u = \lambda y + (1 - \lambda)x$. Consequently, using generalized strong monotonicity of $\partial^C f$, we have

$$\sigma \lambda^2 \|y - x\|^2 = \sigma \|x - u\|^2 \leq \langle \gamma - \xi, u - x \rangle \leq \rho \lambda \|y - x\|^2,$$

that is, $\rho \geq \lambda \sigma$, which contradicts the arbitrariness of ρ. □

Definition 4.9. Let K be a nonempty convex set contained in an open subset of $\mathbb{R}^n$. A locally Lipschitz function $f : K \to \mathbb{R}$ is said be *sharply pseudoconvex* if there exists a constant $\theta > 0$ such that for all $x, y \in K$, $\lambda \in [0, 1]$ and $\xi \in \partial^C f(x)$,

$$\langle \xi, y - x \rangle \geq 0 \;\Rightarrow\; f(y) \geq f(y + \lambda(y - x)) + \lambda(1 - \lambda)\theta \|y - x\|^2.$$

Theorem 4.30. Let K be a nonempty convex set contained in an open subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipschitz function. Then,

(a) f is strictly pseudoconvex if and only if the Clarke subdifferential $\partial^C f$ is a generalized strictly pseudomonotone set-valued map;

(b) f is sharply pseudoconvex with modulus $\theta > 0$ if and only if the Clarke subdifferential $\partial^C f$ is a generalized strongly pseudomonotone set-valued map with constant θ.

Proof. (a) Assume that f is strictly pseudoconvex. For any given $x, y \in K$ with $x \neq y$ and $\xi \in \partial^C f(x)$, $\zeta \in \partial^C f(y)$, let

$$\langle \xi, y - x \rangle \geq 0. \tag{4.13}$$

We want to show that $\langle \zeta, y - x \rangle > 0$. Assume to the contrary that $\langle \zeta, y - x \rangle \leq 0$. Then the strict pseudoconvexity of f implies that

$$f(x) > f(y). \tag{4.14}$$

On the other hand, it follows from the inequality (4.13) that $f(y) > f(x)$, which contradicts the inequality (4.14). Hence, $\partial^C f$ is generalized strictly pseudomonotone.

Conversely, suppose that $\partial^C f$ is generalized strictly pseudomonotone. For any given $x, y \in K$ with $x \neq y$ and $\xi \in \partial^C f(x)$, let the inequality (4.13) hold. We have to show that $f(y) > f(x)$. Assume to the contrary that $f(y) \leq f(x)$. By Lebourg's mean value theorem (Theorem 2.23), there exist $\lambda \in \ ]0,1[$ and $\gamma \in \partial^C f(\lambda y + (1-\lambda)x)$ such that

$$f(y) - f(x) = \langle \gamma, y - x \rangle \leq 0.$$

Consequently,

$$\langle \gamma, x - (\lambda y + (1-\lambda)x) \rangle = -\lambda \langle \gamma, y - x \rangle \geq 0.$$

By generalized strict pseudomonotonicity of $\partial^C f$, we have

$$\langle \xi, x - (\lambda y + (1-\lambda)x) \rangle = -\lambda \langle \xi, y - x \rangle > 0,$$

that is, $\langle \xi, y - x \rangle < 0$, a contradiction to the inequality (4.13). Hence, f is strictly pseudoconvex.

(b) Assume that f is sharply pseudoconvex with constant $\theta > 0$. For any given $x, y \in K$, $\xi \in \partial^C f(x)$, $\zeta \in \partial^C f(y)$, and $t \in [0,1]$, let the inequality (4.13) hold. By sharp pseudoconvexity of f, we have

$$\limsup_{t \to 0^+} \frac{f(y + t(x-y)) - f(y)}{t} + \theta \|y - x\|^2 \leq 0.$$

Since f is locally Lipschitz and K is a convex set contained in an open subset of $\mathbb{R}^n$, for any any given $\varepsilon > 0$ and $t \in \ ]0,1[$ small enough, there exists a constant $\delta > 0$ such that for all z with $\|z - y\| < \delta$, we have $z, z + t(x-y) \in K$ and

$$\frac{f(z + t(x-y)) - f(z)}{t} \leq \frac{f(y + t(x-y)) - f(y)}{t} + \varepsilon.$$

Hence,

$$\begin{aligned} f^C(y; x-y) + \theta \|y - x\|^2 &= \limsup_{\substack{z \to y \\ t \to 0^+}} \frac{f(z + t(x-y)) - f(z)}{t} + \theta \|y - x\|^2 \\ &\leq \limsup_{t \to 0^+} \frac{f(y + t(x-y)) - f(y)}{t} + \varepsilon + \theta \|y - x\|^2 \\ &\leq \varepsilon. \end{aligned}$$

Since ε was arbitrary, we conclude that $f^C(y; x-y) + \theta \|y - x\|^2 \leq 0$. By Theorem 2.18(c), $f^C(x; d) = \max\{\langle \xi, d \rangle : \xi \in \partial^C f(x)\}$, and therefore we have

$$\langle \xi, y - x \rangle \geq \theta \|y - x\|^2.$$

Thus, $\partial^C f$ is generalized strongly pseudomonotone with constant θ.

Conversely, for any given $x, y \in K$, $\xi \in \partial^C f(x)$, and $\lambda \in \ ]0,1[$, let $x_\lambda := y + \lambda(x - y)$, and the inequality (4.13) holds. By Lebourg's mean value theorem (Theorem 2.23), there exist $t \in \ ]0, \lambda[$ and $\gamma \in \partial^C f(x_t)$ such that

$$f(y) - f(x_\lambda) = \lambda \langle \gamma, y - x \rangle, \quad \text{where } x_t = y + t(x - y).$$

Since $\langle \xi, y-x\rangle \geq 0$ implies $\langle \xi, x_t - x\rangle \geq 0$, by generalized strong pseudomonotonicity of $\partial^C f$, we have

$$(1-t)\langle \gamma, y-x\rangle = \langle \gamma, x_t - x\rangle \geq \theta\|x - x_t\|^2 = \theta(1-t)^2\|y-x\|^2.$$

Therefore,

$$f(y) - f(y+\lambda(x-y)) \geq \theta\lambda(1-t)\|y-x\|^2 \geq \theta\lambda(1-\lambda)\|y-x\|^2,$$

and hence f is sharply pseudoconvex. □

Theorem 4.31. Let K be nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipschitz function. If the set-valued map $\partial^C f$ is generalized strongly pseudomonotone with constant $\sigma > 0$, then f is strongly pseudoconvex with constant $\rho = \sigma/4$.

Proof. For any given $x, y \in K$ and $\xi \in \partial^C f(x)$, let $z = \frac{x+y}{2}$, and the inequality (4.13) holds. By Lebourg's mean value theorem (Theorem 2.23), there exist $\lambda_1, \lambda_2 \in\ ]0,1[$ with $0 < \lambda_1 < 1/2 < \lambda_2 < 1$, $\gamma \in \partial^C f(u)$, and $\tau \in \partial^C f(v)$ such that

$$f(y) - f(z) = (1/2)\langle \gamma, y-x\rangle,$$
$$f(z) - f(x) = (1/2)\langle \tau, y-x\rangle,$$

where $u = y + \lambda_1(x-y)$ and $v = y + \lambda_2(x-y)$. Obviously, inequality (4.13) implies that

$$\langle \xi, y-u\rangle \geq 0 \quad \text{and} \quad \langle \xi, y-v\rangle \geq 0.$$

By generalized strong pseudomonotonicity of $\partial^C f$, it follows that

$$\begin{aligned} f(y) - f(z) &= (1/2)(1-\lambda_1)\sigma\langle \gamma, u-y\rangle \\ &\geq (1/2)\sigma(1-\lambda_1)\|u-y\|^2 \\ &= (1/2)\sigma(1-\lambda_1)\|y-x\|^2. \end{aligned}$$

Similarly, we can deduce that

$$f(z) - f(x) \geq (1/2)\sigma(1-\lambda_2)\|y-x\|^2.$$

Hence,

$$\begin{aligned} f(y) - f(x) &\geq (1/2)\sigma\left((1-\lambda_1) + (1-\lambda_2)\right)\|y-x\|^2 \\ &> (1/2)\sigma(1-\lambda_1)\|y-x\|^2 > (\sigma/4)\|y-x\|^2. \end{aligned}$$

This proves the assertion. □

From Theorems 4.30(b) and 4.31, we obtain the following corollary.

Corollary 4.11. Let K be nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipschitz function. If f is sharply pseudoconvex with constant θ, then it is strongly pseudoconvex with constant $\theta/4$.

Theorem 4.32. Let K be nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipschitz function. If the Clarke subdifferential $\partial^C f$ is a generalized pseudomonotone set-valued map, then f is generalized pseudoconvex, and hence quasiconvex.

Proof. Consider $x, y \in K$ such that $f(y) < f(x)$. Since f is continuous, there exists $\lambda \in \ ]0,1[$ such that $f(x+\lambda(y-x)) > f(y)$. Since f is locally Lipschitz on an open set containing $[y, x+\lambda(y-x)]$, by Lebourg's mean value theorem (Theorem 2.23), there exist $z \in \ ]y, x+\lambda(y-x)[$, and $\xi_\lambda \in \partial^C f(z)$ such that $z = x + \mu(y-x)$ for some $\mu \in \ ]\lambda, 1[$ and

$$\begin{aligned} 0 < f(x+\lambda(y-x)) - f(y) &= \langle \xi_\lambda, x+\lambda(y-x) - y\rangle \\ &= (\lambda - 1)\langle \xi_\lambda, y - x\rangle. \end{aligned}$$

Therefore,

$$\langle \xi_\lambda, y - x\rangle < 0.$$

By generalized pseudomonotonicity of $\partial^C f$, there exists $\xi \in \partial^C f(x)$ such that $\langle \xi, y - x\rangle < 0$, which leads to the generalized pseudoconvexity of f.

By Theorem 3.21, f is quasiconvex. □

Theorem 4.33. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipschitz function. If f is generalized pseudoconvex, then the Clarke subdifferential $\partial^C f$ is a generalized pseudomonotone set-valued map.

Proof. Suppose that f is generalized pseudoconvex. Consider $x, y \in K$ such that

$$\langle \xi, y - x\rangle > 0, \quad \text{for all } \xi \in \partial^C f(x). \tag{4.15}$$

We first show that $\mathbf{0} \notin \partial^C f(y)$. Indeed, relation (4.15) and Theorem 2.18(c) together imply that

$$f^C(x; y-x) = \limsup_{\substack{z \to x \\ t \to 0^+}} \frac{f(z + t(y-x)) - f(z)}{t} > 0,$$

that is, there exist $\{x_m\} \subseteq \mathbb{R}^n$ and $\{t_m\} \subseteq \mathbb{R}$ such that $x_m \to x$, $t_m \to 0^+$ and

$$\frac{f(x_m + t_m(y - x_m)) - f(x_m)}{t_m} > 0,$$

and therefore,

$$f(x_m + t_m(y - x_m)) - f(x_m) > 0.$$

Since f is locally Lipschitz and generalized pseudoconvex, it follows from Theorem 3.21 that f is quasiconvex, and hence

$$f(x_m) < f(x_m + t_m(y - x_m)) \leq f(y).$$

By generalized pseudoconvexity, $\mathbf{0} \notin \partial^C f(y)$.

Now, from relation (4.15) it follows that for some $\xi \in \partial^C f(x)$, there exists $\varepsilon > 0$ such that

$$\langle \xi, u - x \rangle > 0, \quad \text{for all } u \in \mathbb{B}_\varepsilon(y), \tag{4.16}$$

which together with generalized pseudoconvexity imply that

$$f(u) \geq f(x), \quad \text{for all } u \in \mathbb{B}_\varepsilon(y). \tag{4.17}$$

From inequality (4.16), it follows that

$$\langle \xi, y - x \rangle > 0.$$

If $\partial^C f$ is not generalized strongly pseudomonotone, then $\langle \zeta, y-x \rangle > 0$ for some $\zeta \in \partial^C f(y)$, which on using generalized pseudoconvexity of f yields $f(x) \geq f(y)$. Combining the last inequality with the inequality (4.17), it follows that y is a local minimizer of f, and hence $\mathbf{0} \in \partial^C f(y)$, a contradiction. $\square$

Theorem 4.34. Let K be nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipschitz function. Then, f is quasiconvex if and only if the Clarke subdifferential $\partial^C f$ is a generalized quasimonotone set-valued map.

Proof. Assume that f is quasiconvex. For any given $x, y \in K$, $\xi \in \partial^C f(x)$, and $\zeta \in \partial^C f(y)$, let

$$\langle \xi, y - x \rangle > 0. \tag{4.18}$$

We want to show that $\langle \zeta, y-x \rangle \geq 0$. Suppose to the contrary that $\langle \zeta, y-x \rangle < 0$. Then by Theorem 3.19, we have $f(x) > f(y)$, which implies that $\langle \xi, y-x \rangle \leq 0$. This contradicts (4.18).

Conversely, assume that f is not quasiconvex. Then there exist $x, y \in K$ with $x \neq y$ and $\lambda \in \ ]0,1[$ such that $f(x) \leq f(y)$ and $f(z) > f(y)$, where $z = x + \lambda(y - x)$. By Lebourg's mean value theorem (Theorem 2.23), there exist $\lambda_1, \lambda_2 \in \ ]0,1[$ with $0 < \lambda_1 < \lambda < \lambda_2 < 1$, $\gamma \in \partial^C f(u)$, and $\tau \in \partial^C f(v)$ such that

$$f(z) - f(x) = \langle \gamma, z - x \rangle > 0,$$

$$f(z) - f(y) = \langle \tau, z - y \rangle > 0,$$

where $u = x + \lambda_1(y - x) \in \ ]x, z[$ and $v = x + \lambda_2(y - x) \in \ ]z, y[$. Proceeding to the next step, we have

$$\langle \gamma, v - u \rangle = (\lambda_2 - \lambda_1)\langle \gamma, y - x \rangle = \frac{\lambda_2 - \lambda_1}{\lambda}\langle \gamma, z - x \rangle > 0,$$

$$\langle \tau, v - u \rangle = -(\lambda_2 - \lambda_1)\langle \tau, x - y \rangle = -\frac{\lambda_2 - \lambda_1}{1 - \lambda}\langle \tau, z - y \rangle < 0.$$

This contradicts the generalized quasimonotonicity of $\partial^C f$, and hence the result is proved. $\square$

Penot and Quang [175] studied the generalized monotonicities and generalized convexities involving the Clarke-Rockafellar subdifferential [184] defined as

$$\partial^{CR} f(x) = \left\{ \xi \in \mathbb{R}^n : \langle \xi, d \rangle \leq f^{CR}(x; d), \text{ for all } d \in \mathbb{R}^n \right\},$$

where $f^{CR}(x; d)$ is defined as

$$f^{CR}(x; d) = \sup_{\varepsilon > 0} \limsup_{(y,\alpha) \downarrow_f x;\ t \downarrow 0} \inf_{u \in \mathbb{B}_\varepsilon(d)} \frac{f(y + tu) - \alpha}{t},$$

where $(y, \alpha) \downarrow_f x$ means that $y \to x$, $\alpha \to f(x)$, $\alpha \geq f(y)$.

Part II

Nonsmooth Variational Inequalities and Nonsmooth Optimization

Chapter 5

Elements of Variational Inequalities

5.1 Introduction

The subject of variational inequalities has its origin in the calculus of variation associated with the minimization of infinite dimensional functionals. On the basis of a contact problem studied by Signorini, Fichera [78] carried on a research and proposed the terminology of a variational inequality for the formulation of the Signorini's contact problem. Stampacchia [197] simultaneously began the study of variational inequalities. In this research, he implicated some scholars, like, Hartman, Lions, etc. [29, 129]. Auslender [22], Brézis [38], Browder [40], Hartman and Stampacchia [100], Lions and Stampacchia [148], Minty [161, 162], Rockafellar [181], and Stampacchia [198] obtained several existence results for a solution of variational inequalities with or without using monotonicity condition.

Variational inequality has shown to be an important mathematical model in the study of many real problems, in particular equilibrium problems. It provides us with a tool for formulating and qualitatively analyzing the equilibrium problems in terms of existence and uniqueness of solutions, stability, and sensitivity analysis, and provides us with algorithms for computational purposes.

The theory of finite dimensional variational inequality was initiated by Smith [194] and Dafermos [59]. Smith [194] formulated the traffic assignment problem in the form of a variational inequality. Dafermos [59] recognized that the traffic network equilibrium conditions has a structure of a variational inequality. Details on variational inequalities related to traffic equilibrium problems and equilibrium models are provided in other studies [70, 99, 136, 167, 173] and the references therein.

In fact, the last three decades have witnessed an exceptional interest for variational inequalities, and an enormous number of papers and books have been devoted to this topic. More and more problems arising from the economic world, as the spatial price equilibrium problem, the oligopolistic market equilibrium problem, the migration problem, and many others, were formulated in terms of finite dimensional variational inequalities and solved using this theory [59, 70, 136, 167].

Several well-known problems from mathematical programming, such as system of nonlinear equations, optimization problems, complementarity problems, and fixed point problems, can be written in the form of a variational inequality problem. In the recent decades, a variational inequality problem and its various generalizations and extensions have been studied and analyzed quite extensively. Variational inequalities have not only stimulated new and deep results in many different branches of mathematics and engineering sciences but have also provided a unified and general framework for studying many unilateral free and moving boundary value problems arising in elastostatics fluid flow through porous media [24, 29, 41, 109, 129, 167] and so forth.

5.2 Variational Inequalities and Related Problems

Let K be a nonempty subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be a mapping. The *variational inequality problem* (VIP) is to determine a vector $\bar{x} \in K$ such that

$$\langle F(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K. \tag{5.1}$$

The inequality (5.1) is called a *variational inequality* (VI).

Roughly speaking, the variational inequality (5.1) states that the vector $F(\bar{x})$ must be at a nonobtuse angle with all the feasible vectors emanating from $\bar{x}$. In other words, the vector $\bar{x}$ is a solution of VIP if and only if $F(\bar{x})$ forms a nonobtuse angle with every vector of the form $y - \bar{x}$, for all $y \in K$.

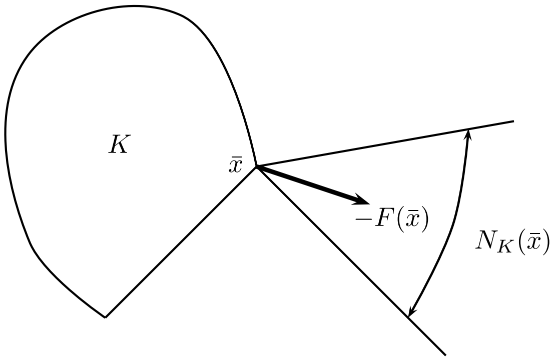


FIGURE 5.1: The solution of VIP and the normal cone

Geometrically, a vector $\bar{x}$ is a solution of VIP if and only if $-F(\bar{x}) \in N_K(\bar{x})$, where $N_K(\bar{x})$ is the normal cone to K at the point $\bar{x}$ and is defined by

$$N_K(x) = \begin{cases} \{d \in \mathbb{R}^n : \langle d, y - x\rangle \leq 0 \text{ for all } y \in K\}, & \text{if } x \in K \\ \emptyset, & \text{otherwise.} \end{cases}$$

Clearly, $\bar{x} \in K$ is a solution of VIP if and only if

$$\mathbf{0} \in F(\bar{x}) + N_K(\bar{x}). \tag{5.2}$$

The inclusion (5.2) is called a *generalized equation.*

The simplest example of a variational inequality problem is the problem of solving a system of nonlinear equations, which is given in the form of the following result.

Proposition 5.1. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a mapping. A vector $\bar{x} \in \mathbb{R}^n$ is a solution of VIP if and only if $F(\bar{x}) = \mathbf{0}$.

Proof. Let $F(\bar{x}) = \mathbf{0}$. Then, obviously, the inequality (5.1) holds with equality.

Conversely, suppose that $\bar{x}$ satisfies the inequality (5.1). Then, by taking $y = \bar{x} - F(\bar{x})$ in (5.1), we get

$$\langle F(\bar{x}), \bar{x} - F(\bar{x}) - \bar{x}\rangle = \langle F(\bar{x}), -F(\bar{x})\rangle \geq 0,$$

that is, $-\|F(\bar{x})\|^2 \geq 0$, which implies that $F(\bar{x}) = \mathbf{0}$. □

Both unconstrained and constrained optimization problems can be formulated as a variational inequality problem. If $F(x)$ is the gradient of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, then the following result provides a relationship between an optimization problem and a variational inequality problem.

Proposition 5.2. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a differentiable function. If $\bar{x}$ is a solution of the following optimization problem

$$\min f(x), \quad \text{subject to } x \in K, \tag{5.3}$$

then $\bar{x}$ is a solution of VIP with $F = \nabla f$.

Proof. For any $y \in K$, define a function $\varphi : [0, 1] \to \mathbb{R}$ by

$$\varphi(\lambda) = f(\bar{x} + \lambda(y - \bar{x})), \quad \text{for all } \lambda \in [0, 1].$$

Since $\varphi(\lambda)$ attains its minimum at $\lambda = 0$, therefore $\varphi'(0) \geq 0$, that is,

$$\langle \nabla f(\bar{x}), y - \bar{x}\rangle \geq 0, \quad \text{for all } y \in K. \tag{5.4}$$

Hence, $\bar{x}$ is a solution of VIP with $F \equiv \nabla f$. □

We observe that the inequality (5.4) is nothing but the first-order necessary optimality condition for the optimization problem (5.3). This condition becomes sufficient when the function f satisfies convexity or generalized convexity assumption, namely, pseudoconvexity.

Proposition 5.3. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a pseudoconvex function. If $\bar{x}$ is a solution of VIP with $F(\bar{x}) = \nabla f(\bar{x})$, then it is a solution of the optimization problem (5.3).

Proof. Suppose that $\bar{x}$ is a solution of VIP but not an optimal solution of the optimization problem (5.3). Then, there exists a vector $y \in K$ such that $f(y) < f(\bar{x})$. By pseudoconvexity of f, we have $\langle \nabla f(\bar{x}), y - \bar{x} \rangle < 0$, which is a contradiction to the fact that $\bar{x}$ is a solution of VIP. $\square$

Remark 5.1. It is always possible to associate the constrained optimization problem (5.3) with the variational inequality (5.4), whereas given the VIP, it is not always possible to define a constrained optimization problem for which VIP represents the first-order optimality condition. However, this can be achieved by proving the existence of a differentiable function f such that $F = \nabla f$.

Theorem 5.1. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable function. Then, there exists a differentiable function f such that $F = \nabla f$ if and only if the Jacobian matrix $J(F)(y)$ is symmetric, for all $y \in \mathbb{R}^n$.

We conclude that although VIP encompasses an optimization problem, a variational inequality problem can be reformulated as an optimization problem only when the Jacobian $J(F)(y)$ is symmetric. The variational inequality problem, therefore, is a more general problem as it can also handle a function F with an asymmetric Jacobian.

Let $K = \mathbb{R}^n_+$ and $F : K \to \mathbb{R}^n$ be a mapping. The *complementarity problem* (CP) is to find $\bar{x} \in \mathbb{R}^n_+$ such that

$$F(\bar{x}) \in \mathbb{R}^n_+ \quad \text{and} \quad \langle F(\bar{x}), \bar{x} \rangle = 0. \tag{5.5}$$

Let K be a closed convex cone in $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be a mapping. The *nonlinear complementarity problem* (NCP) is to find a vector $\bar{x} \in K$ such that

$$F(\bar{x}) \in K^* \quad \text{and} \quad \langle F(\bar{x}), \bar{x} \rangle = 0, \tag{5.6}$$

where K^* is the dual cone of K.

When the mapping F is affine, that is, $F(x) = Mx + b$, where M is an $n \times n$ matrix and $b \in \mathbb{R}^n$, then the above problem is known as the *linear complementarity problem.*

Geometrically, the nonlinear complementarity problem is to find a nonnegative vector $\bar{x}$ such that its image $F(\bar{x})$ is also nonnegative and is orthogonal to $\bar{x}$.

See other studies for further details and applications of complementarity problems [8, 18, 66, 70, 108, 109, 120, 121, 122, 123, 124, 125, 165, 166, 187, 188, 199, 211].

The next result provides a relationship between the nonlinear complementarity problem and the variational inequality problem.

Proposition 5.4. If K is a closed convex pointed cone in $\mathbb{R}^n$, then VIP and NCP have precisely the same solution sets.

Proof. Suppose that $\bar{x} \in K$ is a solution of VIP. Then,

$$\langle F(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K. \tag{5.7}$$

In particular, taking $y = x + \bar{x}$ in the above inequality, we get

$$\langle F(\bar{x}), x \rangle \geq 0, \quad \text{for all } x \in K,$$

which implies that $F(\bar{x}) \in K^*$.

By substituting $y = 2\bar{x}$ in inequality (5.7), we obtain

$$\langle F(\bar{x}), \bar{x} \rangle \geq 0, \tag{5.8}$$

and again taking $y = \mathbf{0}$ in inequality (5.7), we get

$$\langle F(\bar{x}), -\bar{x} \rangle \geq 0. \tag{5.9}$$

Inequalities (5.8) and (5.9) together imply that $\langle F(\bar{x}), \bar{x} \rangle = 0$. Hence, $\bar{x}$ is a solution of NCP.

Conversely, suppose that $\bar{x} \in K$ is a solution of NCP, then we have

$$\langle F(\bar{x}), \bar{x} \rangle = 0 \text{ and } \langle F(\bar{x}), y \rangle \geq 0, \quad \text{for all } y \in K.$$

Thus,

$$\langle F(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

Hence, $\bar{x}$ is a solution of VIP. $\square$

Let K be a nonempty subset of $\mathbb{R}^n$ and $T : K \to K$ be a mapping. The *fixed point problem* (FPP) is to find $\bar{x} \in K$ such that

$$T(\bar{x}) = \bar{x}. \tag{5.10}$$

The relationships between a variational inequality and a fixed point problem are given in the following results.

Proposition 5.5. Let K be a nonempty subset of $\mathbb{R}^n$ and $T : K \to K$ be a mapping. If the mapping $F : K \to K$ is defined by

$$F(x) = x - T(x), \tag{5.11}$$

then VIP (5.1) coincides with FPP (5.10).

Proof. Let $\bar{x} \in K$ be a fixed point of the problem (5.10). Then, $F(\bar{x}) = \mathbf{0}$, and thus $\bar{x}$ solves (5.1).

Conversely, suppose that $\bar{x}$ solves (5.1) with $F(\bar{x}) = \bar{x} - T(\bar{x})$. Then, $T(\bar{x}) \in K$ and letting $y = T(\bar{x})$ in (5.1) gives $-\|\bar{x} - T(\bar{x})\|^2 \geq 0$, that is, $\bar{x} = T(\bar{x})$. $\square$

Proposition 5.6. Let K be a nonempty closed convex subset of $\mathbb{R}^n$. An element $\bar{x} \in K$ is a solution of VIP if and only if for any $\gamma > 0$, $\bar{x}$ is a fixed point of the mapping $P_K(I - \gamma F) : K \to K$, that is, $\bar{x} = P_K(\bar{x} - \gamma F(\bar{x}))$, where $P_K(\bar{x} - \gamma F(\bar{x}))$ denotes the projection of $\bar{x} - \gamma F(\bar{x})$ onto K.

Proof. Suppose that $\bar{x}$ is a solution of VIP. Then,

$$\langle F(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

Multiplying the above inequality by $-\gamma < 0$, and adding $\langle \bar{x}, y - \bar{x} \rangle$ to both sides, we get

$$\langle \bar{x}, y - \bar{x} \rangle \geq \langle \bar{x} - \gamma F(\bar{x}), y - \bar{x} \rangle, \quad \text{for all } y \in K.$$

From Theorem 1.5, it follows that $\bar{x} = P_K(\bar{x} - \gamma F(\bar{x}))$.

Conversely, if $\bar{x} = P_K(\bar{x} - \gamma F(\bar{x}))$ for any $\gamma > 0$, then again invoking Theorem 1.5, we have

$$\langle \bar{x}, y - \bar{x} \rangle \geq \langle \bar{x} - \gamma F(\bar{x}), y - \bar{x} \rangle, \quad \text{for all } y \in K,$$

which yields,

$$\langle F(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

Thus, $\bar{x}$ is a solution of VIP. □

This equivalent formulation is very useful in order to introduce some computational methods based on the projection operator.

5.3 Basic Existence and Uniqueness Results

This section deals with the existence and uniqueness results for a solution of VIP. We observe that VIP does not always admit a solution. For example, if $K = \mathbb{R}$ and $F(x) = e^x$, then VIP has no solution.

We first mention the well-known Brouwer's fixed point theorem for a continuous function defined on a nonempty compact convex subset of $\mathbb{R}^n$.

Theorem 5.2 (Brouwer's Fixed Point Theorem)**.** Let K be a nonempty compact convex subset of $\mathbb{R}^n$ and $T : K \to K$ be a continuous mapping. Then, T has at least one fixed point, that is, there exists $\bar{x} \in K$ such that $T(\bar{x}) = \bar{x}$.

Proof. Let $\mathbb{B}$ be a closed ball in $\mathbb{R}^n$ such that $K \subseteq \mathbb{B}$. From Corollary 1.3, P_K is nonexpansive, and hence, continuous. Therefore, the mapping $T \circ P_K : \mathbb{B} \to K \subseteq \mathbb{B}$ is also continuous on $\mathbb{B}$ into itself. Thus, by Theorem B.2, $T \circ P_K$ has at least one fixed point $\bar{x} \in K$, that is, $(T \circ P_K)(\bar{x}) = \bar{x} \in K$. Since $P_K(x) = x$ for all $x \in K$, we have $T(\bar{x}) = \bar{x}$. □

Theorem 5.3. Let K be a nonempty compact convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be a continuous mapping. Then, VIP admits at least one solution.

Proof. Since P_K and $I - \gamma F$ are continuous mappings for any $\gamma > 0$. $P_K \circ (I - \gamma F)$ is a continuous function defined on a nonempty compact convex set K. By Brouwer's fixed point theorem (Theorem 5.2), it follows that $P_K \circ (I - \gamma F)$ admits a fixed point. The conclusion then follows from Proposition 5.6. □

Let $K_r = K \cap \mathbb{B}_r[\mathbf{0}]$, where $r > 0$ is any real number. The variational inequality problem defined over K_r, denoted by $(\text{VIP})_r$, is defined as follows: Find $\bar{x}_r \in K_r$ such that

$$\langle F(\bar{x}_r), y - \bar{x}_r \rangle \geq 0, \quad \text{for all } y \in K_r.$$

If $K_r \neq \emptyset$ and K is a closed convex subset of $\mathbb{R}^n$, then K_r is nonempty compact and convex. Further, if $F : K \to \mathbb{R}^n$ is continuous, then by Theorem 5.3, there exists at least one solution $\bar{x}_r$ of $(\text{VIP})_r$.

The next theorem gives a necessary and sufficient condition for the existence of a solution of VIP.

Theorem 5.4. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be a continuous mapping. The VIP admits a solution if and only if there exist $r > 0$ and a solution $\bar{x}_r$ of $(\text{VIP})_r$ such that $\|\bar{x}_r\| < r$.

Proof. Clearly, if there exists a solution $\bar{x}$ of VIP, then $\bar{x}$ is a solution of $(\text{VIP})_r$ whenever $\|\bar{x}\| < r$.

Conversely, suppose that there exist $r > 0$ and a solution $\bar{x}_r$ of $(\text{VIP})_r$ such that $\|\bar{x}_r\| < r$. Let $y \in K$ be arbitrary. Since $\|\bar{x}_r\| < r$, we can always find $\lambda > 0$ sufficiently small such that $w = \bar{x}_r + \lambda(y - \bar{x}_r) \in K_r$. Consequently,

$$0 \leq \langle F(\bar{x}_r), w - \bar{x}_r \rangle = \lambda \langle F(\bar{x}_r), y - \bar{x}_r \rangle.$$

Since y was arbitrary, it follows that $\bar{x}_r$ is a solution of VIP. □

The study of variational inequalities over unbounded domains is usually based on a coercivity condition, which is set in order to guarantee the existence of solutions. The subsequent corollary ensures the existence of a solution of VIP using a coercivity condition.

Corollary 5.1. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be a continuous mapping such that

$$\frac{\langle F(x) - F(x_0), x - x_0 \rangle}{\|x - x_0\|} \to \infty \quad \text{as } \|x\| \to \infty, \quad x \in K, \tag{5.12}$$

for some $x_0 \in K$. Then, VIP always has a solution.

Proof. Choose $s > \|F(x_0)\|$ and $r > \|x_0\|$ such that

$$\langle F(x) - F(x_0), x - x_0 \rangle \geq s\|x - x_0\|, \quad \text{for } \|x\| \geq r,\ x \in K.$$

Then,

$$\begin{aligned} \langle F(x), x - x_0 \rangle &\geq s\|x - x_0\| + \langle F(x_0), x - x_0 \rangle \\ &\geq s\|x - x_0\| - \|F(x_0)\|\ \|x - x_0\| \\ &\geq (s - \|F(x_0)\|)(\|x\| - \|x_0\|) > 0, \quad \text{for } \|x\| = r. \end{aligned} \tag{5.13}$$

Let $x_r \in K_r$ be a solution of $(\text{VIP})_r$. Then,

$$\langle F(x_r), x_r - x_0 \rangle = -\langle F(x_r), x_0 - x_r \rangle \leq 0,$$

and thus it follows from (5.13) that $\|x_r\| \neq r$. Hence, $\|x_r\| < r$ and the result follows from Theorem 5.4. □

Monotone maps and their generalizations play an important role in the theory of variational inequalities [25, 57, 61, 125, 161, 163, 205]. A detailed study of monotone mappings has already been developed in Section 4.2. If the mapping F in the formulation of VIP possesses some additional properties of monotonicity, then the existence and uniqueness of a solution of VIP becomes more obtainable.

The following result shows that strict monotonicity provides the uniqueness of a solution of VIP, if it exists.

Theorem 5.5. Let K be a nonempty subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be strictly monotone. Then, the solution of VIP is unique, if it exists.

Proof. Let, if possible, x_1 and x_2, $x_1 \neq x_2$ be two solutions of VIP. Then,

$$\langle F(x_1), y - x_1 \rangle \geq 0, \quad \text{for all } y \in K, \tag{5.14}$$

and

$$\langle F(x_2), y - x_2 \rangle \geq 0, \quad \text{for all } y \in K. \tag{5.15}$$

Taking $y = x_2$ in inequality (5.14) and $y = x_1$ in inequality (5.15) and then adding the resultant inequalities, we get

$$\langle F(x_1) - F(x_2), x_1 - x_2 \rangle \leq 0,$$

a contradiction to the fact that F is strictly monotone. Hence, $x_1 = x_2$. □

Remark 5.2. We note that the strict monotonicity does not ensure the existence of a solution of VIP. For example, consider $K = \{x \in \mathbb{R} : x \geq 0\}$ and $F(x) = -e^{-x} - 1$. Then F is strictly monotone, but there is no $x \in K$ such that $F(x) \geq 0$. Hence, there is no $x \in K$ such that the variational inequality (5.1) holds.

The following theorem guarantees the existence of a unique solution of VIP.

Theorem 5.6. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be a continuous mapping. If F is strongly monotone, then there exists a unique solution of VIP.

Proof. If F is strongly monotone with modulus σ, then for a given $x_0 \in K$, we have

$$\frac{\langle F(x) - F(x_0), x - x_0 \rangle}{\|x - x_0\|} \geq \sigma \|x - x_0\|, \quad \text{for all } x \in K.$$

It follows that F is coercive, that is, condition (5.12) holds. Corollary 5.1 guarantees the existence of a solution of VIP. The uniqueness of the solution follows from the fact that strong monotonicity implies strict monotonicity. □

In 1962, Minty [161] gave a complete characterization of the solutions of VIP in terms of the solutions of the following problem, known as the *Minty variational inequality problem* (MVIP): Find $\bar{x} \in K$ such that

$$\langle F(y), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K, \tag{5.16}$$

where $F : K \to \mathbb{R}^n$ and K is a nonempty subset of $\mathbb{R}^n$. The inequality (5.16) is known as the *Minty variational inequality* (MVI). It is important to observe that while in MVIP the inequality must be satisfied at $F(y)$ for all $y \in K$, whereas in VIP the inequality needs to hold only for the point $F(\bar{x})$ where $\bar{x} \in K$ is the solution of the problem. Recently, this problem and its generalizations have been studied [4, 53, 54, 85, 129, 133, 155, 199].

To distinguish between a variational inequality and Minty variational inequality, we sometime write Stampacchia variational inequality (SVI) instead of a variational inequality.

Contrary to the Stampacchia variational inequality problem (SVIP), the Minty variational inequality problem (MVIP) is a sufficient optimality condition for the optimization problem (5.3), which becomes necessary if the objective function f is pseudoconvex and differentiable.

Theorem 5.7. [85] Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a differentiable function. The following statements hold:

(a) If $\bar{x} \in K$ is a solution of MVIP with $F \equiv \nabla f$, then $\bar{x}$ is a solution of optimization problem (5.3).

(b) If f is pseudoconvex and $\bar{x} \in K$ is a solution of the optimization problem (5.3), then it is a solution of MVIP with $F \equiv \nabla f$.

Proof. (a) Let $y \in K$ be arbitrary. Consider the function $\varphi(\lambda) = f(\bar{x}+\lambda(y-\bar{x}))$ for all $\lambda \in [0,1]$. Since $\varphi'(\lambda) = \langle \nabla f(\bar{x}+\lambda(y-\bar{x})), y-\bar{x}\rangle$ and $\bar{x}$ is a solution of MVIP with $F \equiv \nabla f$, it follows that

$$\varphi'(\lambda) = \langle \nabla f(\bar{x}+\lambda(y-\bar{x})), y-\bar{x}\rangle \geq 0, \quad \text{for all } \lambda \in [0,1].$$

This implies that φ is a nondecreasing function on $[0,1]$, and therefore,

$$f(y) = \varphi(1) \geq \varphi(0) = f(\bar{x}).$$

Thus, $\bar{x}$ is a solution of the optimization problem (5.3).

(b) Let $\bar{x}$ be an optimal solution of the optimization problem (5.3). Then, for all $y \in K$, $f(\bar{x}) \leq f(y)$. Since f is a pseudoconvex differentiable function, by Theorem 1.35, f is quasiconvex. Then by Theorem 1.26, we have

$$\langle \nabla f(y), y-\bar{x}\rangle \geq 0, \quad \text{for all } y \in K.$$

Thus, $\bar{x}$ is a solution of MVIP. □

Definition 5.1. Let K be a nonempty convex subset of $\mathbb{R}^n$. A mapping $F : K \to \mathbb{R}^n$ is said to be

(a) *lower hemicontinuous* if for any fixed $x, y \in K$, the function $\lambda \mapsto F(x+\lambda(y-x))$ defined on $[0,1]$ is lower semicontinuous;

(b) *upper hemicontinuous* if for any fixed $x, y \in K$, the function $\lambda \mapsto F(x+\lambda(y-x))$ defined on $[0,1]$ is upper semicontinuous;

(c) *hemicontinuous* if for any fixed $x, y \in K$, the mapping $\lambda \mapsto F(x+\lambda(y-x))$ defined on $[0,1]$ is continuous, that is, if F is continuous along the line segments in K.

The following Minty lemma is an important tool in the theory of variational inequalities when the mapping is monotone and the domain is convex.

Lemma 5.1 (Minty Lemma)**.** Let K be a nonempty subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be a mapping. The following assertions hold.

(a) If K is convex and F is hemicontinuous, then every solution of MVIP is a solution of VIP.

(b) If F is pseudomonotone, then every solution of VIP is a solution of MVIP.

Proof. (a) Let $\bar{x} \in K$ be a solution of MVIP. Then, for any $y \in K$ and $\lambda \in \,]0,1]$, $z_\lambda = \bar{x}+\lambda(y-\bar{x}) \in K$, and hence

$$\langle F(z_\lambda), z_\lambda - \bar{x}\rangle \geq 0, \quad \text{for all } \lambda \in \,]0,1],$$

which implies that

$$\langle F(\bar{x}+\lambda(y-\bar{x})), y-\bar{x}\rangle \geq 0, \quad \text{for all } \lambda \in \,]0,1].$$

By the hemicontinuity of F, we have

$$\langle F(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

Hence, $\bar{x}$ is a solution of VIP.

(b) It is obvious that every solution of VIP is a solution of MVIP by pseudomonotonicity of F. □

Lemma 5.2. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be hemicontinuous and pseudomonotone. Then, the solution set of VIP is closed and convex.

Proof. In view of Lemma 5.2, the solution sets of VIP and MVIP are the same. Therefore, it is sufficient to show that the solution set of MVIP is closed and convex.

Let $\bar{x}$ and $\hat{x}$ be any two solutions of MVIP. Then, for any $y \in K$,

$$\langle F(y), y - \bar{x} \rangle \geq 0 \quad \text{and} \quad \langle F(y), y - \hat{x} \rangle \geq 0.$$

Multiplying the first inequality by $\lambda \in [0, 1]$ and the second inequality by $1-\lambda$, and then by adding the resultants, we get

$$\langle F(y), \lambda y - \lambda\bar{x} + (1-\lambda)y - (1-\lambda)\hat{x} \rangle = \langle F(y), y - (\lambda\bar{x} + (1-\lambda)\hat{x}) \rangle \geq 0.$$

Hence, $\lambda\bar{x} + (1-\lambda)\hat{x}$ is a solution of MVIP. Thus, the solution set of MVIP is convex.

Let $\{x_m\}$ be a sequence of solutions of MVIP such that $x_m \to \bar{x}$ as $m \to \infty$. Then, for all $y \in K$,

$$\langle F(y), y - x_m \rangle \geq 0, \quad \text{for all } m.$$

Since $\langle F(y), y - x_m \rangle \to \langle F(y), y - \bar{x} \rangle$ as $m \to \infty$, we have $\langle F(y), y - \bar{x} \rangle \geq 0$, and thus $\bar{x}$ is a solution of MVIP. Therefore, the solution set of MVIP is closed. □

Theorem 5.8. Let K be a nonempty compact convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be hemicontinuous and pseudomonotone. Then, VIP has a solution.

Proof. Define two set-valued maps $P, Q : K \to 2^K$ by

$$P(y) = \{x \in K : \langle F(x), y - x \rangle \geq 0\}, \quad \text{for all } y \in K,$$

and

$$Q(y) = \{x \in K : \langle F(y), y - x \rangle \geq 0\}, \quad \text{for all } y \in K.$$

Then, P is a KKM map. Indeed, let $\{y_1, y_2, \ldots, y_m\}$ be a finite subset of K and $\tilde{x} \in \text{co}(\{y_1, y_2, \ldots, y_m\})$. Then, $\tilde{x} = \sum_{i=1}^m \lambda_i y_i$ for some $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$ with $\sum_i^m \lambda_i = 1$. If $\tilde{x} \notin \bigcup_{i=1}^m P(y_i)$, then,

$$\langle F(\tilde{x}), y_i - \tilde{x} \rangle < 0, \quad \text{for all } i = 1, 2, \ldots, m,$$

and so,

$$\sum_{i=1}^{m} \lambda_i \langle F(\tilde{x}), y_i - \tilde{x} \rangle < 0.$$

Therefore, we have

$$\begin{aligned} 0 = \langle F(\tilde{x}), \tilde{x} - \tilde{x} \rangle &= \left\langle F(\tilde{x}), \sum_{i=1}^{m} \lambda_i y_i - \sum_{i=1}^{m} \lambda_i \tilde{x} \right\rangle \\ &= \left\langle F(\tilde{x}), \sum_{i=1}^{m} \lambda_i \left(y_i - \tilde{x} \right) \right\rangle \\ &= \sum_{i=1}^{m} \lambda_i \left\langle F(\tilde{x}), y_i - \tilde{x} \right\rangle < 0, \end{aligned}$$

a contradiction. Thus, we must have $\mathrm{co}(\{y_1, y_2, \ldots, y_m\}) \subseteq \bigcup_{i=1}^{m} P(y_i)$, and hence P is a KKM-map.

Since F is a pseudomonotone map, $P(y) \subseteq Q(y)$ for all $y \in K$, and thus, Q is also a KKM-map.

For each $y \in K$, $Q(y)$ is closed. Indeed, let $\{x_m\}$ be a sequence in $Q(y)$ for any fixed $y \in K$, such that x_m converges to $\hat{x} \in K$. Then,

$$\langle F(y), y - x_m \rangle \geq 0, \quad \text{for all } m.$$

Since $\langle F(y), y - x_m \rangle \to \langle F(y), y - \hat{x} \rangle$ as $m \to \infty$, we have $\langle F(y), y - \hat{x} \rangle \geq 0$, and thus, $\hat{x} \in Q(y)$. Therefore, $Q(y)$ is a closed subset of the compact set K, and so it is compact. By the Fan-KKM theorem (Theorem B.3), $\bigcap_{y \in K} Q(y) \neq \emptyset$. Hence, there exists $\bar{x} \in K$ such that

$$\langle F(y), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

By Lemma 5.1, $\bar{x}$ is a solution of VIP. □

Example 5.1. Let $K = [1, +\infty[$ and $F : K \to \mathbb{R}$ be defined as $F(x) = x$ for all $x \in K$. Since K is not compact, Theorem 5.8 cannot be applied. However, it is easy to see that $\bar{x} = 1 \in K$ is such that $\langle F(\bar{x}), y - \bar{x} \rangle = \bar{x}(y - \bar{x}) = y - 1 \geq 0$ for all $y \in K$, that is, $\bar{x} = 1$ is a solution of VIP.

When K is not necessarily compact, we have the following results.

Theorem 5.9. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be hemicontinuous and pseudomonotone. Assume that there exist a compact subset C of $\mathbb{R}^n$ and $\tilde{y} \in K \cap C$ such that

$$\langle F(x), \tilde{y} - x \rangle < 0, \quad \text{for all } x \in K \setminus C. \tag{5.17}$$

Then, VIP has a solution.

Proof. Let the set-valued maps $P, Q : K \to 2^K$ be the same as in the proof of Theorem 5.8. Let $\tilde{y} \in K$ and the set C be the same as in the hypothesis. Then, we want to show that $P(\tilde{y})$ is compact. If $P(\tilde{y}) \nsubseteq C$, then there exists $x \in P(\tilde{y})$ such that $x \in K \setminus C$. It follows that

$$\langle F(x), \tilde{y} - x \rangle \geq 0,$$

which contradicts (5.17). Therefore, we have $P(\tilde{y}) \subseteq C$. Then, the closure $\mathrm{cl}(P(\tilde{y}))$ of $P(\tilde{y})$ is a closed subset of the compact set C, and hence compact. As we have seen in Theorem 5.8, P is a KKM-map. Therefore, by Theorem B.3, $\bigcap_{y \in K} \mathrm{cl}(P(y)) \neq \emptyset$. Since for each $y \in K$, $Q(y)$ is closed and $P(y) \subseteq Q(y)$, we have

$$\mathrm{cl}(P(y)) \subseteq \mathrm{cl}(Q(y)) = Q(y).$$

Therefore,

$$\emptyset \neq \bigcap_{y \in K} \mathrm{cl}(P(y)) \subseteq \bigcap_{y \in K} Q(y).$$

Thus, there exists $\bar{x} \in K$ such that

$$\langle F(y), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

By Lemma 5.1, $\bar{x}$ is a solution of VIP. □

Theorem 5.10. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be hemicontinuous and pseudomonotone. Assume that there exist a compact subset D of $\mathbb{R}^n$ and $\tilde{y} \in K \cap D$ such that

$$\langle F(\tilde{y}), \tilde{y} - x \rangle < 0, \quad \text{for all } x \in K \setminus D. \tag{5.18}$$

Then, VIP has a solution.

Proof. Let the set-valued maps $P, Q : K \to 2^K$ be the same as in the proof of Theorem 5.8. Let $\tilde{y} \in K$ and the set D be the same as in the hypothesis. By the same argument as in the proof of Theorem 5.8, we derive that Q is a KKM map and, for each $y \in K$, $Q(y)$ is closed.

We show that $Q(\tilde{y}) \subseteq K \cap D$. If $Q(\tilde{y}) \nsubseteq D$, then there exists $x \in Q(\tilde{y})$ such that $x \in K \setminus D$. It follows that

$$\langle F(\tilde{y}), \tilde{y} - x \rangle \geq 0,$$

which contradicts (5.18). Hence, $Q(\tilde{y}) \subseteq D$, and thus, $Q(\tilde{y}) \subseteq K \cap D$. Since K is closed and D is compact, $K \cap D$ is compact. Therefore, $Q(\tilde{y})$ is a closed subset of the compact set $K \cap D$, and hence, $Q(\tilde{y})$ is compact. Then, by Theorem B.3, $\bigcap_{y \in K} Q(y) \neq \emptyset$. The rest of the proof follows on the lines of the proof of Theorem 5.8. □

We establish the following existence result for a solution of VIP by using a Browder-type fixed point theorem.

Theorem 5.11. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be hemicontinuous and pseudomonotone. Assume that there exist a nonempty compact convex subset B of K and a nonempty compact subset D of K such that for each $x \in K \setminus D$, there exists $\tilde{y} \in B$ such that $\langle F(x), \tilde{y} - x\rangle < 0$. Then, VIP has a solution.

Proof. For each $x \in K$, define set-valued maps $P, Q : K \to 2^K$ by

$$P(x) = \{y \in K : \langle F(y), y - x\rangle < 0\}$$

and

$$Q(x) = \{y \in K : \langle F(x), y - x\rangle < 0\}.$$

It is clear that for each $x \in K$, $Q(x)$ is convex. By pseudomonotonicity of F, we have $P(x) \subseteq Q(x)$, and hence $\mathrm{co}(P(x)) \subseteq \mathrm{co}(Q(x)) = Q(x)$ for all $x \in K$.

For each $y \in K$, the complement of $P^{-1}(y)$ in K is

$$[P^{-1}(y)]^c = \{x \in K : \langle F(y), y - x\rangle \geq 0\}$$

is closed in K, and thus $P^{-1}(y)$ is open in K.

Assume that for all $x \in K$, $P(x)$ is nonempty. Then, all the conditions of Theorem B.4 are satisfied, and therefore there exists $\hat{x} \in K$ such that $\hat{x} \in Q(\hat{x})$. It follows that

$$0 = \langle F(\tilde{x}), \tilde{x} - \tilde{x}\rangle < 0,$$

a contradiction. Hence, there exists $\bar{x} \in K$ such that $P(\bar{x}) = \emptyset$. This implies that for all $y \in K$,

$$\langle F(y), y - \bar{x}\rangle \geq 0,$$

that is, $\bar{x} \in K$ is a solution of MVIP. By Lemma 5.1, $\bar{x} \in K$ is a solution of VIP. $\square$

Unlike the Stampacchia variational inequality, the continuity of F and the boundedness and convexity of K do not ensure the existence of a solution of the Minty variational inequality and that some generalized monotonicity conditions on F are needed.

Example 5.2. Let $K = [0, 4]$ and $F(x) = \cos(\pi x/2)$. Then, K is convex and compact, F is continuous on K, but the solution set of MVIP is empty. However, the solution set of VIP is the set $\{0, 1, 3\}$.

Daniilidis and Hadjisavvas [61] showed that the notion of proper quasimonotonicity, which they introduced earlier [60], is a sufficient condition for the existence of a solution of MVIP. An earlier version of proper quasimonotonicity was introduced by Zhou and Chen [208] under the name of 0-diagonally quasiconcavity.

Definition 5.2. Let K be a nonempty subset of $\mathbb{R}^n$. A mapping $F : K \to \mathbb{R}^n$ is said to be *properly quasimonotone* if for every $x_1, x_2, \ldots, x_m \in K$ and every $y \in \mathrm{co}(\{x_1, x_2, \ldots, x_m\})$, there exists $i \in \{1, 2, \ldots, m\}$ such that

$$\langle F(x_i), y - x_i \rangle \leq 0.$$

Theorem 5.12. Let K be a nonempty compact convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be properly quasimonotone. Then, MVIP has a solution.

Proof. Define the set-valued mapping $Q : K \to 2^K$ by

$$Q(y) = \{x \in K : \langle F(y), y - x \rangle \geq 0\}, \quad \text{for all } y \in K.$$

For any $y_1, y_2, \ldots, y_m \in K$ and $\tilde{y} \in \mathrm{co}(\{y_1, y_2, \ldots, y_m\})$, proper quasimonotonicity of F implies that $\tilde{y} \in \bigcup_{i=1}^m Q(y_i)$. Also, for each $y \in K$, $Q(y)$ is a closed subset of the compact set K, and hence compact. Therefore, by Theorem B.3, it follows that $\bigcap_{y \in K} Q(y) \neq \emptyset$. Thus, any $\bar{x} \in \bigcap_{y \in K} Q(y)$ is a solution of MVIP. □

Further, John [118] showed that the condition of proper quasimonotonicity in Theorem 5.12 cannot be relaxed any further. The existence of a solution of VIP under quasimonotonicity and dense pseudomonotonicity has been studied by Hadjisavvas and Schaible [97] and Luc [149], respectively.

Now, we present the existence result for VIP under the assumption of pseudomonotonicity in the sense of Brézis [37].

Definition 5.3. Let K be a nonempty subset of $\mathbb{R}^n$. A function $F : K \to \mathbb{R}^n$ is said to be *B-pseudomonotone* if for each $x \in K$ and every sequence $\{x_m\}$ in K converging to x with

$$\liminf_{m \to \infty} \langle F(x_m), x - x_m \rangle \geq 0,$$

we have

$$\langle F(x), y - x \rangle \geq \limsup_{m \to \infty} \langle F(x_m), y - x_m \rangle, \quad \text{for all } y \in K.$$

Theorem 5.13. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be B-pseudomonotone such that for each finite subset A of K, $x \mapsto \langle F(x), y - x \rangle$ is upper semicontinuous on $\mathrm{co}(A)$. Assume that there exist a nonempty compact subset D of K and an element $\tilde{y} \in D$ such that for all $x \in K \setminus D$, $\langle F(x), \tilde{y} - x \rangle < 0$. Then, VIP has a solution.

Proof. For each $x \in K$, define a set-valued map $P : K \to 2^K$ by

$$P(x) = \{y \in K : \langle F(x), y - x \rangle < 0\}.$$

Then for all $x \in K$, $P(x)$ is convex. Let A be a finite subset of K. Then for all $y \in \mathrm{co}(A)$,

$$[P^{-1}(y)]^c \cap \mathrm{co}(A) = \{x \in \mathrm{co}(A) : \langle F(x), y - x \rangle \geq 0\}$$

is closed in $\mathrm{co}(A)$, by upper semicontinuity of the map $x \mapsto \langle F(x), y - x\rangle$ on $\mathrm{co}(A)$. Hence, $P^{-1}(y) \cap \mathrm{co}(A)$ is open in $\mathrm{co}(A)$.

Suppose that $x, y \in \mathrm{co}(A)$ and $\{x_m\}$ is a sequence in K converging to x such that

$$\langle F(x_m), \lambda y + (1-\lambda)x - x_m\rangle \geq 0, \quad \text{for all } m \in \mathbb{N} \text{ and all } \lambda \in [0,1].$$

For $\lambda = 0$, we have

$$\langle F(x_m), x - x_m\rangle \geq 0, \quad \text{for all } m \in \mathbb{N},$$

and therefore,

$$\liminf_{m\to\infty} \langle F(x_m), x - x_m\rangle \geq 0.$$

By B-pseudomonotonicity of F, we have

$$\langle F(x), y - x\rangle \geq \limsup_{m\to\infty} \langle F(x_m), y - x_m\rangle. \tag{5.19}$$

For $\lambda = 1$, we have

$$\langle F(x_m), y - x_m\rangle \geq 0, \quad \text{for all } m \in \mathbb{N},$$

and therefore,

$$\liminf_{m\to\infty} \langle F(x_m), y - x_m\rangle \geq 0. \tag{5.20}$$

From inequalities (5.19) and (5.20), we obtain

$$\langle F(x), y - x\rangle \geq 0,$$

and thus, $y \notin P(x)$.

Assume that for all $x \in D$, $P(x)$ is nonempty. Then, all the conditions of Theorem B.5 are satisfied. Hence, there exists $\hat{x} \in K$ such that $\hat{x} \in P(\hat{x})$, that is,

$$0 = \langle F(\hat{x}), \hat{x} - \hat{x}\rangle < 0,$$

a contradiction. Thus, there exists $\bar{x} \in D \subseteq K$ such that $P(\bar{x}) = \emptyset$, that is,

$$\langle F(\bar{x}), y - \bar{x}\rangle \geq 0, \quad \text{for all } y \in K.$$

Hence, $\bar{x}$ is a solution of VIP. □

Corollary 5.2. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be B-pseudomonotone such that for each finite subset A of K, $x \mapsto \langle F(x), y - x\rangle$ is upper semicontinuous on $\mathrm{co}(A)$. Assume that there exists $\tilde{y} \in K$ such that

$$\lim_{\|x\|\to\infty,\ x \in K} \langle F(x), \tilde{y} - x\rangle < 0. \tag{5.21}$$

Then, the VIP has a solution.

Proof. Let

$$\alpha = \lim_{\|x\| \to \infty,\ x \in K} \langle F(x), \tilde{y} - x \rangle.$$

Then by inequality (5.21), $\alpha < 0$. Let $r > 0$ be such that $\|\tilde{y}\| \leq r$ and

$$\langle F(x), \tilde{y} - x \rangle < \frac{\alpha}{2}, \quad \text{for all } x \in K \text{ with } \|x\| > r.$$

Let $\mathbb{B}_r = \{x \in K : \|x\| \leq r\}$ be the closed unit ball. Then, $\mathbb{B}_r$ is a nonempty and compact subset of K. Note that for any $x \in K \setminus \mathbb{B}_r$, $\langle F(x), \tilde{y} - x \rangle < \frac{\alpha}{2} < 0$, and the conclusion follows from Theorem 5.13 by taking $D = \mathbb{B}_r$. □

As an application of Corollary 5.2, we derive the following result on the existence of a fixed point of a nonlinear operator.

Corollary 5.3. Let K be a nonempty convex subset of $\mathbb{R}^n$. Let $F : K \to \mathbb{R}^n$ be B-pseudomonotone such that for each finite subset A of K, $x \mapsto \langle F(x), y - x \rangle$ is lower semicontinuous on $\mathrm{co}(A)$. Assume that there exists $\tilde{y} \in K$ such that

$$\lim_{\|x\| \to \infty,\ x \in K} \langle x - F(x), \tilde{y} - x \rangle < 0.$$

Then, there exists a fixed point $\bar{x} \in K$ of F, that is, $F(\bar{x}) = \bar{x}$.

Proof. Define a nonlinear mapping $S : K \to \mathbb{R}^n$ by $S(x) = x - F(x)$ for all $x \in K$. Then, obviously, S satisfies all the conditions of Corollary 5.2. Hence, there exists $\bar{x} \in K$ such that

$$\langle S(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

Let $y = F(\bar{x})$, we have $\|\bar{x} - F(\bar{x})\|^2 \leq 0$. Therefore, $F(\bar{x}) = \bar{x}$. □

5.4 Gap Functions

As observed earlier, a system of equations is a simplest example of a variational inequality problem. Therefore, it is natural that several algorithms developed for solving variational inequalities are generalizations of the classical methods for systems of equations, such as Newton's method, projection method, and nonlinear Jacobi (diagonalization) method. Yet another approach of solving a variational inequality is to transform it into an equivalent minimization problem. An obvious advantage of this approach lies in the fact that a minimization problem may be solved by descent algorithms that possess a global convergence property. A typical situation where the problem can be reformulated as an optimization problem is when $F(x)$ is a gradient map, that is, if there exists a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ such that

$$F(x) = \nabla f(x), \quad \text{for all } x \in \mathbb{R}^n. \tag{5.22}$$

As mentioned in Theorem 5.1, if the mapping F is continuously differentiable, a necessary and sufficient condition for F to satisfy the condition (5.22) is that the Jacobian matrix $J(F)(x)$ is symmetric for all x. Hence, it follows that if the Jacobian of the map F in VIP is symmetric and positive semidefinite, then there exists a convex function such that (5.22) holds, and solving VIP is equivalent to solving the optimization problem (5.3). Unfortunately, we cannot expect this symmetry condition to hold in many practical equilibrium models. Hence, for a general asymmetric variational inequality, a different optimization approach has been proposed by Auslender [22] in the form of a gap function. The term "gap function" was first used by Hearn [102] in connection with a convex optimization problem. The gap functions are useful in studying error bounds and in developing algorithms to find the approximate solutions of VIP or optimization problem. One of the most important feature of a gap function for a variational inequality is that the variational inequality can be transformed into an optimization problem, and thus the powerful optimization solution methods and algorithms can be applied for finding the approximate solutions of variational inequalities. Different kinds of gap functions, also known as merit functions, for variational inequalities have been studied in the literature during the last three decades. See other studies for details on gap functions [16, 43, 70, 143, 145, 155, 158, 173, 178, 196, 202, 203, 212].

Definition 5.4. A function $\varphi : \mathbb{R}^n \to \mathbb{R}$ is said to be a *gap function* for VIP if it satisfies the following properties:

(i) $\varphi(x) \geq 0$ for all $x \in \mathbb{R}^n$,

(ii) $\varphi(\bar{x}) = 0$ if and only if $\bar{x}$ solves VIP.

From the definition it follows that gap functions provide a certificate of optimality to the variational inequality problems. The following gap function for VIP was proposed by Auslender [22], and therefore it is known as *Auslender's gap function.* Some properties of Auslender's gap function can be found in Auslender [23] and Li et al. [145].

Theorem 5.14. The function $\varphi_A : \mathbb{R}^n \to \mathbb{R}$ defined as

$$\varphi_A(x) = \sup_{y \in K} \langle F(x), x - y \rangle, \quad \text{for all } x \in \mathbb{R}^n, \tag{5.23}$$

is a gap function for VIP.

Proof. Let $y = x$, then $\varphi_A(x) \geq \langle F(x), x - x \rangle = 0$.

To prove the second condition of a gap function, we first let $\varphi_A(\bar{x}) = 0$. Then,

$$\sup_{y \in K} \langle F(\bar{x}), \bar{x} - y \rangle = 0,$$

that is,

$$\langle F(\bar{x}), \bar{x} - y \rangle \leq 0, \quad \text{for all } y \in K,$$

and hence $\bar{x}$ is a solution of VIP.

Conversely, let $\bar{x}$ be a solution of VIP. Then,

$$\langle F(\bar{x}), \bar{x} - y\rangle \leq 0, \quad \text{for all } y \in K,$$

and therefore

$$\sup_{y \in K} \langle F(\bar{x}), \bar{x} - y\rangle \leq 0,$$

and the equality holds when $y = \bar{x}$. So, $\varphi_A(\bar{x}) = 0$. □

Gap functions play a vital role in formulating a variational inequality problem as an optimization problem. By the virtue of the defining properties of the gap function φ, we can transform the VIP into the following optimization problem with optimal value zero:

$$\min \varphi(x), \quad \text{subject to } x \in K. \tag{5.24}$$

Consequently, methods for solving an optimization problem can also be utilized for finding the solutions of variational inequality problems. In fact, the gap function approach is an important research method in the variational inequality theory. These functions have been known to monitor convergence of sequences to a solution of a variational inequality. In economic equilibrium theory, the gap function corresponds to the excess demand function and is a better measure of proximity to an equilibrium problem than the quasi-welfare function $\int_0^x F(t)dt$ (when the latter is unambiguously defined), since it is directly related to the perception of the market structure by the economic agents. In a traffic equilibrium framework, the gap function measures the difference between the actual (perceived) travel costs and minimal (shortest path) costs.

Since the Auslender's gap function is not differentiable, it is not easy to find a suitable algorithm for the optimization problem (5.24). To overcome the nondifferentiability of the Auslender's gap function φ_A, Auchmuty [21] and Fukushima [81] independently proposed the following regularized gap function:

$$\varphi_F(x) = -\langle F(x), H(x) - x\rangle - \frac{1}{2}\langle H(x) - x, H(x) - x\rangle, \tag{5.25}$$

where $H(x) = P_K(x - F(x))$. From Proposition 5.6, we know that $\bar{x} \in K$ is a solution of VIP if and only if $\bar{x} = H(\bar{x})$. The function φ_F defined by (5.25) is known as the *Fukushima's gap function*.

Theorem 5.15. The function $\varphi_F : \mathbb{R}^n \to \mathbb{R}$ defined by (5.25) is a gap function for VIP.

Proof. By completing square, equation (5.25) can be rewritten as

$$\varphi_F(x) = \frac{1}{2}\left\{\|F(x)\|^2 - \|H(x) - (x - F(x))\|^2\right\}.$$

Clearly, $\|F(x)\|$ is the distance between the point $x - F(x)$ (which may not be in K) and $x \in K$; also $\|H(x) - (x - F(x))\|$ is the distance between the point $x - F(x)$ and its projection $H(x)$ on K. Obviously, the first distance is always greater than or equal to the second, and thus $\varphi_F(x) \geq 0$ for all $x \in K$. Furthermore, $\varphi_F(x) = 0$, that is, the two distances are equal if and only if $H(x) = x$. Then the conclusion follows from Proposition 5.6. □

Corollary 5.4. $\bar{x}$ is a solution of VIP if and only if it is a solution of the following minimization problem:

$$\min_{x \in \mathbb{R}^n} \varphi_F(x). \tag{5.26}$$

Unlike the Auslender's gap function, the Fukushima's gap function enjoys some nice properties that make it an attractive alternative to solve VIP.

Theorem 5.16. If $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuous, then the Fukushima's gap function φ_F defined by (5.25) is also continuous. Furthermore, if F is differentiable, then φ_F is also differentiable with

$$\nabla \varphi_F(x) = F(x)^\top - \langle \nabla F(x) - I, H(x) - x \rangle, \tag{5.27}$$

where I is the identity mapping.

Proof. Since the projection operator H is continuous, the continuity of φ_F follows from its definition.

For the differentiability part, let

$$h(x, y) = \langle F(x), y - x \rangle + \frac{1}{2}\|y - x\|^2.$$

Then, the Fukushima's gap function φ_F defined by (5.25) can be rewritten as

$$\varphi_F(x) = -\min\{h(x, y) : y \in K\},$$

where the minimum is uniquely attained at $y = H(x)$. Also, by Theorem 1.7, Chapter 4 in Auslender [23], φ_F is differentiable. Finally, taking the derivative of h with respect to the first argument, we obtain

$$\frac{\partial h}{\partial x} = \langle \nabla F(x), y - x \rangle - F(x)^\top + (x - y)^\top.$$

Since

$$\nabla \varphi_F(x) = \left[\frac{\partial h}{\partial x}\right]\Bigg|_{(x, H(x))},$$

we get the conclusion. □

Several other properties of Fukushima's gap function can be found in Fukushima [81] and Goh and Yang [93]. Different kinds of gap functions can be found in the literature [16, 43, 70, 93, 84, 136, 143, 145, 155, 158, 173, 178, 196, 202, 203, 212].

As we have seen, the Minty variational inequality and Stampacchia variational inequality are closely related to each other. In fact, if the map F, involved in the formulations of these inequalities, is hemicontinuous and pseudomonotone, then both problems have the same solution set. Therefore, it is also interesting to study the gap functions for a Minty variational inequality problem. Marcotte and Zhu [155] proposed the following gap function for Minty variational inequality problem:

$$\Phi(x) = \sup_{y \in K} \langle F(y), x - y \rangle, \quad \text{for all } x \in \mathbb{R}^n. \tag{5.28}$$

It is called the *dual gap function*, while the Auslender's gap function φ_A defined by (5.23) is also known as the *primal gap function*. Since the function Φ is the pointwise supremum of affine functions, it is closed and convex. Moreover, $\Phi(x) \geq 0$ for all $x \in K$; and when F is continuous and pseudomonotone, $\bar{x} \in K$ is a solution of VIP if and only if $\Phi(\bar{x}) = 0$. Hence, any solution of VIP is a global minimizer for the following convex optimization problem:

$$\min_{x \in K} \Phi(x), \tag{5.29}$$

with zero optimal value. The dual gap function Φ can be rewritten as

$$\Phi(x) = \langle F(\tilde{y}), x - \tilde{y} \rangle, \tag{5.30}$$

where $\tilde{y}$ is any point in the set $\Lambda(x) = \arg\max_{z \in K} \langle F(z), x - z \rangle$.

Definition 5.5. Let K be a nonempty subset of $\mathbb{R}^n$. A vector-valued map $F : K \to \mathbb{R}^n$ is said to be *pseudomonotone*$^+$ if F is pseudomonotone and, for all $x, y \in K$,

$$\langle F(x), y - x \rangle \geq 0 \text{ and } \langle F(y), y - x \rangle = 0 \quad \Rightarrow \quad F(y) = F(x).$$

If F is pseudomonotone$^+$, then the dual gap function Φ enjoys the following nice properties, which were studied by Marcotte and Zhu [155].

Theorem 5.17. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be continuous and pseudomonotone$^+$. Then,

(a) F is constant over the set of solutions of VIP;

(b) for any solution $\bar{x}$ of VIP, F is constant and equal to $F(\bar{x})$ over $\Lambda(\bar{x})$;

(c) for any solution $\bar{x}$ of VIP, $\Lambda(\bar{x})$ is equal to the set of solutions of VIP;

(d) if K is compact, then Φ is continuously differentiable over the set of solutions of VIP, and $\nabla\Phi(\bar{x}) = F(\bar{x})$ for every solution $\bar{x}$ of VIP.

Proof. (a) Let $\bar{x}$ and $\hat{x}$ be any two solutions of VIP. Then from the Minty lemma (Lemma 5.1), $\bar{x}$ is a solution of MVIP, and therefore

$$\langle F(y), y-\bar{x}\rangle \geq 0, \quad \text{for all } y \in K.$$

In particular,

$$\langle F(\hat{x}), \hat{x}-\bar{x}\rangle \geq 0.$$

Since $\hat{x}$ is a solution of VIP, we have

$$\langle F(\hat{x}), \hat{x}-\bar{x}\rangle \leq 0.$$

From the above two inequalities, we obtain

$$\langle F(\hat{x}), \hat{x}-\bar{x}\rangle = 0.$$

Since $\bar{x}$ is a solution of VIP, we have

$$\langle F(\bar{x}), \hat{x}-\bar{x}\rangle \geq 0.$$

This together with the pseudomonotonicity$^+$ yields $F(\bar{x}) = F(\hat{x})$, and hence F is constant.

(b) For any solution $\bar{x}$ of VIP and $\bar{y} \in \Lambda(\bar{x})$, we have

$$\Phi(\bar{x}) = \langle F(\bar{y}), \bar{x}-\bar{y}\rangle = 0. \tag{5.31}$$

Since $\bar{x}$ is a solution of VIP, we have

$$\langle F(\bar{x}), \bar{y}-\bar{x}\rangle \geq 0.$$

This together with $\langle F(\bar{y}), \bar{y}-\bar{x}\rangle = 0$ imply, by the pseudomonotonicity$^+$ of F, that $F(\bar{x}) = F(\bar{y})$.

(c) Let $\tilde{x} \in \Lambda(\bar{x})$. From (b), we have

$$F(\bar{x}) = F(\tilde{x}) \quad \text{and} \quad \langle F(\tilde{x}), \bar{x}-\tilde{x}\rangle = 0.$$

This implies that for all $y \in K$,

$$\begin{aligned}\langle F(\tilde{x}), \tilde{x}-y\rangle &= \langle F(\tilde{x}), \tilde{x}-\bar{x}\rangle + \langle F(\tilde{x}), \bar{x}-y\rangle \\ &= \langle F(\bar{x}), \bar{x}-y\rangle \leq 0,\end{aligned}$$

and thus $\tilde{x}$ is a solution of VIP.

Conversely,

$$\langle F(\tilde{x}), \bar{x}-\tilde{x}\rangle = 0,$$

for any solution $\tilde{x}$ of VIP, and hence $\tilde{x} \in \Lambda(\bar{x})$.

(d) From a result of Danskin [63], the directional derivative of Φ at $\bar{x}$ in the direction d is given by

$$\Phi'(\bar{x}; d) = \max_{y \in \Lambda(\bar{x})} \langle F(y), d\rangle = \langle F(\bar{x}), d\rangle,$$

since, by (b), F is constant and equal to $F(\bar{x})$ over $\Lambda(\bar{x})$. Thus, Φ is continuously differentiable at every solution of VIP, with gradient $\nabla\Phi(\bar{x}) = F(\bar{x})$. □

By using the dual gap function, Marcotte and Zhu [155] studied the relationship among a global error bound, weak sharpness of the solution set, minimum principle sufficiency property, and finite termination of descent algorithms for the solution of VIP. Most of the results from Marcotte and Zhu [155] require the assumptions that F is pseudomonotone$^+$ and K is compact. Zhang et al. [212] further studied the dual gap function by relaxing the pseudomonotonicity$^+$ of F and compactness of K. They also studied some other properties of the dual gap function and established some applications of these properties. The dual gap function and its properties were further studied by Wu and Wu [202]. They basically characterized the Gâteaux differentiability of the dual gap function and presented several sufficient conditions for its directional derivative expression.

Definition 5.6. Let K be a nonempty subset of $\mathbb{R}^n$. A pair $(\bar{x}, \bar{y}) \in K \times K$ is said to be a *saddle point* of the function $\Psi : K \times K \to \mathbb{R}$ if

$$\Psi(x, \bar{y}) \le \Psi(\bar{x}, \bar{y}) \le \Psi(\bar{x}, y), \quad \text{for all } x, y \in K.$$

The solutions of SVIP and MVIP gives a saddle point of Ψ as can be seen from the following theorem whose proof is a straightforward application of the definitions.

Define $\Psi(x, y) := \langle F(x), x - y\rangle$ for all $(x, y) \in K \times K$.

Theorem 5.18. (a) $(\bar{x}, \bar{y}) \in K \times K$ is a saddle point of Ψ if and only if $\bar{x}$ solves MVIP and $\bar{y}$ solves SVIP.

(b) If $(\bar{x}, \bar{y}) \in K \times K$ is a saddle point of Ψ, then $\Psi(\bar{x}, \bar{y}) = 0$.

Thus, it follows that a method to find the solutions of SVIP and MVIP can be based on the search of the saddle point of Ψ. Further, if we know even one solution of SVIP, then it can be used to search the solution of MVIP and vice versa. In fact, suppose if $\bar{y}$ is a solution of SVIP with $F(\bar{y}) \neq \mathbf{0}$, then the set $\{x \in K : \Psi(x, \bar{y}) = 0\}$ contains the solutions of MVIP. Similar conditions can be obtained for the solutions of SVIP if we start with a solution $\bar{x}$ of MVIP.

5.5 Solution Methods

In Section 5.3, we have seen several conditions that guarantee the existence of a solution of VIP. A large number of solution methods have been proposed in the literature in the last three decades [70, 92, 99, 136, 137, 173]. In this section, we consider only two very well-known solution methods, namely, the projection method and auxiliary principle method, to compute the approximate solutions of VIP. These two methods are very popular due to their simplicity and clarity. The convergence properties of these methods serve as a basis for understanding more complicated methods.

Throughout the section, unless otherwise specified, we assume that K is a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ is a mapping.

Projection Method. For a given $x_0 \in \mathbb{R}^n$, compute x_{m+1} by the rule:

$$x_{m+1} = P_K(x_m - \gamma F(x_m)), \quad m = 0, 1, 2, \ldots, \tag{5.32}$$

where P_K denotes the projection operator and $\gamma > 0$ is a constant.

In view of Proposition 5.6, the iteration (5.32) is well-defined. Moreover, x_{m+1} is also the unique solution of the following auxiliary variational inequality: Find $x_{m+1} \in K$ such that

$$\langle F(x_m) + \gamma^{-1}(x_{m+1} - x_m), y - x_{m+1}\rangle \geq 0, \quad \text{for all } y \in K, \tag{5.33}$$

where $\gamma > 0$ is a constant.

The following convergence result for the sequence generated by iteration (5.32) is one of the most strongest convergence results.

Theorem 5.19. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be strongly monotone with constant $\sigma > 0$ and Lipschitz continuous with constant $k > 0$. If $\{x_m\}$ is a sequence generated by the iteration (5.32) with $\gamma \in \ ]0, 2\sigma/k^2[$, then it converges to a unique solution $\bar{x}$ of VIP.

Proof. By Theorem 5.6, VIP has a unique solution $\bar{x} \in K$. Then, by Proposition 5.6, we have

$$\bar{x} = P_K(\bar{x} - \gamma F(\bar{x})). \tag{5.34}$$

By nonexpansiveness of P_K, and strong monotonicity and Lipschitz continuity of F, we have

$$\begin{aligned} \|x_{m+1} - \bar{x}\|^2 &= \|P_K(x_m - \gamma F(x_m)) - P_K(\bar{x} - \gamma F(\bar{x}))\|^2 \\ &= \|(x_m - \bar{x}) - \gamma(F(x_m) - F(\bar{x}))\|^2 \\ &= \|x_m - \bar{x}\|^2 - 2\gamma\langle F(x_m) - F(\bar{x}), x_m - \bar{x}\rangle \\ &\quad + \gamma^2\|F(x_m) - F(\bar{x})\|^2 \\ &\leq (1 - 2\gamma\sigma + \gamma^2 k^2)\|x_m - \bar{x}\|^2. \end{aligned} \tag{5.35}$$

Then,

$$\|x_{m+1} - \bar{x}\| \leq \theta\|x_m - \bar{x}\|,$$

where $\theta = \sqrt{1 - 2\gamma\sigma + \gamma^2 k^2}$. For $\gamma \in \ ]0, 2\sigma/k^2[$, $\theta \in \ ]0, 1[$, and hence it follows that x_m converges to $\bar{x}$. ☐

Definition 5.7. Let K be a nonempty subset of $\mathbb{R}^n$. A mapping $F : K \to \mathbb{R}^n$ is said to be *co-coercive* (or *inverse strongly monotone*) with constant $\mu > 0$ if for every pair of points $x, y \in K$,

$$\langle F(x) - F(y), x - y\rangle \geq \mu\|F(x) - F(y)\|^2.$$

Example 5.3. If $F(x) = Ax + b$, where A is symmetric positive semidefinite $n \times n$ matrix, $b \in \mathbb{R}^n$, then F is co-coercive.

Remark 5.3. It is clear that co-coercive mappings are monotone but not necessarily strongly monotone. However, a strongly monotone and Lipschitz continuous map is co-coercive. Also, co-coercivity of a mapping F (with constant μ) is equivalent to strong monotonicity of the inverse mapping F^{-1} and implies the Lipschitz continuity of F with constant $1/\mu$. Moreover, if a mapping F is strongly monotone with constant σ and Lipschitz continuous with constant k, then it is co-coercive with constant $\mu = \sigma/k^2$. A monotone map is not necessarily co-coercive. For example, consider the map $F : [1, \infty[\ \to \mathbb{R}$ defined by $F(x) = x^2$. Then, F is monotone but not co-coercive. However, if $K = [0, \infty[$, then the map $F(x) = x^2$ is co-coercive but not strongly monotone.

Lemma 5.3. Let K be a nonempty closed convex subset of $\mathbb{R}^n$. If $F : K \to \mathbb{R}^n$ is a co-coercive mapping with constant μ, then the mapping

$$T(x) = x - P_K(x - \gamma F(x)), \tag{5.36}$$

is co-coercive with constant $\mu' = 1 - \frac{\gamma}{4\mu}$ where $\gamma \in \ ]0, 4\mu[$.

Proof. Let $x, y \in K$ be arbitrary and $u = T(x)$, $v = T(y)$. Then, using inequality (1.6), we have

$$\begin{aligned}
& \langle (u - \gamma F(x)) - (v - \gamma F(y)), (x - u) - (y - v)\rangle \\
= \quad & \langle (x - \gamma F(x)) - (y - \gamma F(y)), P_K(x - \gamma F(x)) - P_K(y - \gamma F(y))\rangle \\
& - \|P_K(x - \gamma F(x)) - P_K(y - \gamma F(y))\|^2 \geq 0.
\end{aligned}$$

It follows that

$$\begin{aligned}
& \langle u - v, x - y\rangle \\
\geq \quad & \|u - v\|^2 + \gamma\langle F(x) - F(y), x - y\rangle - \gamma\langle F(x) - F(y), u - v\rangle \\
\geq \quad & \|u - v\|^2 + \gamma\mu\|F(x) - F(y)\|^2 - \gamma\langle F(x) - F(y), u - v\rangle \\
= \quad & \left(1 - \frac{\gamma}{4\mu}\right)\|u - v\|^2 + \left\|\sqrt{\frac{\gamma}{4\mu}}(u - v) - \sqrt{\gamma\mu}(F(x) - F(y))\right\|^2 \\
\geq \quad & \mu'\|u - v\|^2,
\end{aligned}$$

as desired. □

The following result is due to Zhu and Marcotte [214].

Proposition 5.7. Let $F : K \to \mathbb{R}^n$ be a mapping and I be the identity mapping on K. If the mapping $F - \mu I$ is Lipschitz continuous with constant β with $\beta \leq \mu$, then F is co-coercive. Conversely, if F is co-coercive with constant β with $\beta > 1/2\mu$, then $F - \mu I$ is Lipschitz continuous with constant μ.

Proof. Set $G = F - \mu I$. For all $x, y \in K$, we have

$$\begin{aligned}
\|F(y) - F(x)\|^2 &= \|G(y) - G(x) + \mu y - \mu x\|^2 \\
&= \|G(y) - G(x)\|^2 + 2\mu\langle G(y) - G(x), y - x\rangle + \mu^2\|y - x\|^2 \\
&\leq \left(\beta^2 + \mu^2\right)\|y - x\|^2 + 2\mu\langle G(y) - G(x), y - x\rangle \\
&\leq 2\mu\langle (G(y) + \mu y) - (G(x) + \mu x), y - x\rangle.
\end{aligned}$$

This implies that

$$\langle F(y) - F(x), y - x\rangle \geq \frac{1}{2\mu}\|F(y) - F(x)\|^2,$$

that is, F is co-coercive with constant $1/2\mu$.

For the converse part, we observe that

$$\begin{aligned}
& \|(F - \mu I)(y) - (F - \mu I)(x)\|^2 \\
= \; & \|F(y) - F(x)\|^2 - 2\mu\langle F(y) - F(x), y - x\rangle + \mu^2\|y - x\|^2 \\
\leq \; & \mu^2\|y - x\|^2,
\end{aligned}$$

that is, $F - \mu I$ is Lipschitz continuous. $\square$

It is clear from Proposition 5.6 that every zero of the mapping T defined by (5.36) is a solution of VIP.

The following theorem provides the convergence of the sequence generated by iteration (5.32) under co-coercivity assumption.

Theorem 5.20. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be a co-coercive mapping with constant $\mu > 0$. Assume that VIP is solvable. If $\{x_m\}$ is a sequence generated by the iteration (5.32) with $\gamma \in \;]0, 2\mu[$, then it converges to a solution $\bar{x}$ of VIP.

Proof. Let $\bar{x} \in K$ be a solution of VIP. Then from Lemma 5.3, the map T as defined in (5.35) is co-coercive with constant $\mu' = 1 - \gamma/4\mu$, and therefore, we have

$$\begin{aligned}
\|x_{m+1} - \bar{x}\|^2 &= \|(x_m - T(x_m)) - (\bar{x} - T(\bar{x}))\|^2 \\
&= \|x_m - \bar{x}\|^2 - 2\langle T(x_m) - T(\bar{x}), x_m - \bar{x}\rangle + \|T(x_m) - T(\bar{x})\|^2 \\
&\leq \|x_m - \bar{x}\|^2 - (2\mu' - 1)\|T(x_m) - T(\bar{x})\|^2 \\
&= \|x_m - \bar{x}\|^2 - (2\mu' - 1)\|T(x_m)\|^2 \\
&\leq \|x_m - \bar{x}\|^2
\end{aligned}$$

because $2\mu' - 1 = 1 + (1 - \gamma/2\mu) - 1 > 0$. It follows that the sequence $\{x_m\}$ is bounded, and hence it has limit points. Moreover,

$$\lim_{m\to\infty} T(x_m) = 0.$$

Taking an arbitrary limit point $\tilde{x}$ of $\{x_m\}$, we have $T(\tilde{x}) = \mathbf{0}$ by continuity of F and the projection mapping. By Proposition 5.6, $\tilde{x}$ is a solution of VIP, that is, we can replace $\bar{x}$ by $\tilde{x}$ in the above inequalities and the monotone decrease of the distance $\|x_m - \tilde{x}\|$ yields $\lim_{m\to\infty} x_m = \tilde{x}$. □

For classical optimization problem, Cohen and Zhu [49, 50, 52] introduced the so-called auxiliary problem principle as a general framework to describe and analyze computational algorithms ranging from gradient or subgradient to decomposition/coordination algorithms. Cohen [51] further extended this approach to the computation of solutions to variational inequalities, which is close to the approach found in Glowinski et al. [92]. Zhu and Marcotte [214] considered the following iterative schemes based upon the auxiliary problem framework developed by Cohen [51]. From now onward, we assume that the solution set of VIP is nonempty.

Auxiliary Variational Principle. Let θ be a positive parameter and consider, for a given iterate x_m, the auxiliary variational inequality problem (AVIP) that consists of finding x_{m+1} such that

$$\langle \theta F(x_m) + \nabla f(x_{m+1}) - \nabla f(x_m), y - x_{m+1} \rangle \geq 0, \quad \text{for all } y \in K, \tag{5.37}$$

where $f : K \to \mathbb{R}$ is a mapping.

Theorem 5.21. Let K be a nonempty convex bounded subset of $\mathbb{R}^n$ and $F : K \to \mathbb{R}^n$ be co-coercive with constant $\mu > 0$. Let $f : K \to \mathbb{R}$ be strongly convex with modulus $\rho > 0$ such that its gradient ∇f is Lipschitz continuous with constant k. Then, there exists a unique solution $x_{m+1} \in K$ to (5.37). If

$$0 < \theta < 4\mu\rho, \tag{5.38}$$

then the sequence $\{x_m\}$ generated by (5.37) is bounded and converges to a solution of VIP.

Proof. First we prove the existence of a solution of problem (5.37). For the sake of simplicity, we rewrite inequality (5.37) as follows: Find $\bar{x} \in K$ such that

$$\langle \theta F(x_m) + \nabla f(\bar{x}) - \nabla f(x_m), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

For each fixed m and every $y \in K$, define

$$P(y) = \{x \in K : \langle \theta F(x_m) + \nabla f(x) - \nabla f(x_m), y - x \rangle \geq 0\}.$$

Since for each $y \in K$, $y \in P(y)$, we have $P(y)$ is nonempty.

We assert that P is a KKM map. Suppose that there is a finite subset $\{y_1, y_2, \ldots, y_q\}$ of K and $\lambda_i \geq 0$, for all $i = 1, 2, \ldots, q$ with $\sum_{i=1}^q \lambda_i = 1$ such that $\hat{x} = \sum_{i=1}^q \lambda_i y_i \notin P(y_i)$ for all i. Then, we have

$$\langle \theta F(x_m) + \nabla f(\hat{x}) - \nabla f(x_m), y_i - \hat{x} \rangle < 0, \quad \text{for all } i = 1, 2, \ldots, q.$$

Therefore,

$$\sum_{i=1}^{q} \lambda_i \langle \theta F(x_m) + \nabla f(\hat{x}) - \nabla f(x_m), y_i - \hat{x} \rangle < 0,$$

that is,

$$0 = \langle \theta F(x_m) + \nabla f(\hat{x}) - \nabla f(x_m), \hat{x} - \hat{x} \rangle < 0,$$

a contradiction. Hence, P is a KKM-map.

Since $\mathrm{cl}(P(y))$ is a closed subset of the bounded set K, it is compact. Hence, by Theorem B.3,

$$\bigcap_{y \in K} \mathrm{cl}(P(y)) \neq \emptyset.$$

Let $\bar{x} \in \bigcap_{y \in K} \mathrm{cl}(P(y))$. Then, there exists a sequence $\{x_p\}$ in $P(y)$ such that $x_p \to \bar{x}$. Therefore,

$$\langle \theta F(x_m) + \nabla f(x_p) - \nabla f(x_m), y - x_p \rangle \geq 0,$$

and hence

$$\lim_{p \to \infty} \langle \theta F(x_m) + \nabla f(x_p) - \nabla f(x_m), y - x_p \rangle \geq 0,$$

that is,

$$\langle \theta F(x_m) + \nabla f(\bar{x}) - \nabla f(x_m), y - \bar{x} \rangle \geq 0.$$

Therefore, $\bar{x} \in K$ is a solution of the inequality (5.37).

We next prove uniqueness of solution of problem (5.37). Let $\bar{x}_1$ and $\bar{x}_2$ be two solutions of problem (5.37). Then, for all $y \in K$,

$$\langle \theta F(x_m) + \nabla f(\bar{x}_1) - \nabla f(x_m), y - \bar{x}_1 \rangle \geq 0, \tag{5.39}$$

and

$$\langle \theta F(x_m) + \nabla f(\bar{x}_2) - \nabla f(x_m), y - \bar{x}_2 \rangle \geq 0. \tag{5.40}$$

Taking $y = \bar{x}_2$ in inequality (5.39) and $y = \bar{x}_1$ in inequality (5.40) and then adding the resultant inequalities, we get

$$\theta \langle F(x_m), \bar{x}_2 - \bar{x}_1 \rangle + \langle \nabla f(\bar{x}_1) - \nabla f(x_m), \bar{x}_2 - \bar{x}_1 \rangle +$$

$$\theta \langle F(x_m), \bar{x}_1 - \bar{x}_2 \rangle + \langle \nabla f(\bar{x}_2) - \nabla f(x_m), \bar{x}_1 - \bar{x}_2 \rangle \geq 0.$$

Thus,

$$\langle \nabla f(\bar{x}_1), \bar{x}_2 - \bar{x}_1 \rangle \geq -\langle \nabla f(\bar{x}_2), \bar{x}_1 - \bar{x}_2 \rangle.$$

By strong convexity of f, we obtain

$$f(\bar{x}_2) - f(\bar{x}_1) - \rho \|\bar{x}_1 - \bar{x}_2\|^2 \geq -f(\bar{x}_1) + f(\bar{x}_2) + \rho \|\bar{x}_2 - \bar{x}_1\|^2,$$

and therefore,

$$\rho \|\bar{x}_1 - \bar{x}_2\|^2 \leq 0.$$

Since $\rho > 0$, we get $\bar{x}_1 = \bar{x}_2$, and hence, the solution of inequality (5.37) is unique.

Let $\bar{x}$ be any fixed solution of VIP. For each $y \in K$, define a functional

$$\Lambda(y) = f(\bar{x}) - f(y) - \langle \nabla f(y), \bar{x} - y \rangle.$$

By strong convexity of f, we have

$$\Lambda(y) = f(\bar{x}) - f(y) - \langle \nabla f(y), \bar{x} - y \rangle \geq \rho \|y - \bar{x}\|^2. \tag{5.41}$$

By strong convexity of f and inequality (5.37) with $y = \bar{x}$, we get

$$\begin{aligned}
& \Lambda(x_m) - \Lambda(x_{m+1}) \\
= \;& f(x_{m+1}) - f(x_m) - \langle \nabla f(x_m), \bar{x} - x_m \rangle + \langle \nabla f(x_{m+1}), \bar{x} - x_{m+1} \rangle \\
= \;& f(x_{m+1}) - f(x_m) - \langle \nabla f(x_m), \bar{x} - x_{m+1} \rangle - \langle \nabla f(x_m), x_{m+1} - x_m \rangle \\
& + \langle \nabla f(x_{m+1}), \bar{x} - x_{m+1} \rangle \\
= \;& f(x_{m+1}) - f(x_m) - \langle \nabla f(x_m), x_{m+1} - x_m \rangle \\
& + \langle \nabla f(x_{m+1}) - \nabla f(x_m), \bar{x} - x_{m+1} \rangle \\
\geq \;& \rho \|x_m - x_{m+1}\|^2 + \langle \nabla f(x_{m+1}) - \nabla f(x_m), \bar{x} - x_{m+1} \rangle \\
\geq \;& \rho \|x_m - x_{m+1}\|^2 - \theta \langle F(x_m), \bar{x} - x_{m+1} \rangle \\
= \;& \rho \|x_m - x_{m+1}\|^2 + \theta \langle F(x_m), x_{m+1} - \bar{x} \rangle.
\end{aligned} \tag{5.42}$$

We set $y = x_{m+1}$ in VIP and combine it with (5.42) to get

$$\begin{aligned}
\Lambda(x_m) - \Lambda(x_{m+1}) \;&\geq\; \rho \|x_m - x_{m+1}\|^2 + \theta \langle F(x_m) - F(\bar{x}), x_{m+1} - \bar{x} \rangle \\
&=\; \rho \|x_m - x_{m+1}\|^2 + Q,
\end{aligned}$$

where,

$$\begin{aligned}
Q \;&=\; \theta \langle F(x_m) - F(\bar{x}), x_{m+1} - \bar{x} \rangle \\
&=\; \theta \langle F(x_m) - F(\bar{x}), x_m - \bar{x} \rangle + \theta \langle F(x_m) - F(\bar{x}), x_{m+1} - x_m \rangle \\
&\geq\; \theta \left[\mu \|F(x_m) - F(\bar{x})\|^2 + \langle F(x_m) - F(\bar{x}), x_{m+1} - x_m \rangle \right] \\
&\geq\; \theta \left[-\frac{1}{4\mu} \|x_{m+1} - x_m\|^2 \right].
\end{aligned}$$

Therefore,

$$\Lambda(x_m) - \Lambda(x_{m+1}) \geq \left(\rho - \frac{\theta}{4\mu} \right) \|x_{m+1} - x_m\|^2. \tag{5.43}$$

If $x_{m+1} = x_m$ for some m, then x_m is a solution of VIP. Otherwise, it follows from (5.38), $\Lambda(x_m) - \Lambda(x_{m+1})$ is nonnegative and from which we have

$$\lim_{m \to \infty} \|x_{m+1} - x_m\| = 0.$$

Further, from (5.41),

$$\|x_m - \bar{x}\|^2 \leq \left(\frac{1}{\rho} \right) \Lambda(x_m),$$

and the sequence $\{\Lambda(x_m)\}$ is decreasing, we affirm that the sequence $\{x_m\}$ is bounded.

Let $\bar{x}$ be any cluster point of the sequence $\{x_m\}$. Taking the limit in (5.37), we conclude that $\bar{x}$ is a solution of VIP. Let

$$\bar{\Lambda}(x_m) = f(\bar{x}) - f(x_m) - \langle \nabla f(x_m), \bar{x} - x_m \rangle \geq \rho \|\bar{x} - x_m\|^2. \tag{5.44}$$

By the above argument, we know that $\{\bar{\Lambda}(x_m)\}$ is also a decreasing sequence, and from Theorem 1.20, we have

$$\bar{\Lambda}(x_m) \leq \frac{k}{2} \|\bar{x} - x_m\|^2,$$

from which it follows that

$$\lim_{m \to \infty} \bar{\Lambda}(x_m) = 0. \tag{5.45}$$

Combining (5.44) and (5.45), we conclude that the sequence $\{x_m\}$ converges to $\bar{x}$. $\square$

The above method has been extended in Ansari and Yao [17] and Ding [64] for a more general variational inequality.

Chapter 6

Nonsmooth Variational Inequalities

6.1 Introduction

It is established in Chapter 5 that for any differentiable optimization problem over a convex feasible region, the first-order necessary optimality conditions describe a variational inequality problem involving the derivative of the objective function. However, the nonsmooth phenomena occur frequently in optimization, and also the assumption of differentiability is not too realistic since in many practical applications the functions involved need not be differentiable. Moreover, there are some functions that are theoretically differentiable, but the structure of the calculation of their derivatives is so complex that nonsmooth behavior can be detected in the numerical sense. For this reason, there have been substantial efforts to provide appropriate tools for the study of more general problems in mathematical optimization. In this connection, one of the approaches presented in Chapter 3 of the book was to consider the directional derivatives in an abstract way as a bifunction that is a positively homogenous function of the direction. In this chapter, we endeavor to study those variational inequality problems that pertain to the optimization problems involving a nonsmooth objective. It is easy to observe that in case the objective function fails to be differentiable, then a necessary optimality condition may be derived in terms of a generalized directional derivative $h(x; d)$.

6.2 Nonsmooth Variational Inequalities in Terms of a Bifunction

Motivated by the optimality conditions, for a nondifferentiable optimization problem, in terms of generalized directional derivatives we define a variational inequality problem defined by means of a bifunction h.

Let K be a nonempty subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. The variational inequality problem in terms of a bifunction h is defined as follows:

$(\text{VIP})_h$ Find $\bar{x} \in K$ such that $h(\bar{x}; y - \bar{x}) \geq 0, \quad$ for all $y \in K$.

If we take $h(x; y - x) = \langle F(x), y - x \rangle$, where $F : \mathbb{R}^n \to \mathbb{R}^n$, then $(\text{VIP})_h$ reduces to the VIP studied in Chapter 5.

As we have seen in Chapter 5, the Minty variational inequality problem is closely related to VIP, and also provides a necessary and sufficient optimality condition for a differentiable optimization problem under convexity or pseudoconvexity assumption. Therefore, the study of the Minty variational inequality defined by means of a bifunction h is also very important in the theory of nonsmooth variational inequalities. The Minty variational inequality problem (MVIP) in terms of a bifunction h is defined as follows:

$(\text{MVIP})_h$ Find $\bar{x} \in K$ such that $h(y; \bar{x} - y) \leq 0, \quad$ for all $y \in K$.

To prove the equivalence between $(\text{VIP})_h$ and $(\text{MVIP})_h$, we introduce the following concept of upper sign continuity.

Definition 6.1. Let K be nonempty convex subset of $\mathbb{R}^n$. A bifunction $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be *upper sign continuous* if for all $x, y \in K$ and $\lambda \in \,]0, 1[$,

$$h(x + \lambda(y - x); x - y) \leq 0 \quad \text{implies} \quad h(x; y - x) \geq 0.$$

This notion of upper sign continuity for a bifunction extends the concept of upper sign continuity introduced by Hadjisavvas [95].

Clearly, every subodd radially upper semicontinuous bifunction is upper sign continuous.

The following lemma is a generalization of the Minty lemma (Lemma 5.1).

Lemma 6.1. Let K be nonempty convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a pseudomonotone and upper sign continuous bifunction such that h is positively homogeneous in the second argument. Then, $\bar{x} \in K$ is a solution of $(\text{VIP})_h$ if and only if it is a solution of $(\text{MVIP})_h$.

Proof. The pseudomonotonicity of h implies that every solution of $(\text{VIP})_h$ is a solution of $(\text{MVIP})_h$.

Conversely, let $\bar{x} \in K$ be a solution of $(\text{MVIP})_h$. Then,

$$h(y; \bar{x} - y) \leq 0, \quad \text{for all } y \in K. \tag{6.1}$$

Since K is convex, we have $y_\lambda = \bar{x} + \lambda(y - \bar{x}) \in K$ for all $\lambda \in \,]0, 1[$. Therefore, inequality (6.1) becomes

$$h(y_\lambda; \bar{x} - y_\lambda) \leq 0.$$

As $\bar{x} - y_\lambda = \lambda(\bar{x} - y)$ and h is positively homogeneous in the second argument, we have

$$h(y_\lambda; \bar{x} - y) \leq 0.$$

Thus, the upper sign continuity of h implies that $\bar{x} \in K$ is a solution of $(\text{VIP})_h$. ☐

6.3 Relation between an Optimization Problem and Nonsmooth Variational Inequalities

Let us recall the optimization problem:

$$\text{(P)} \qquad \min f(x), \quad \text{subject to } x \in K,$$

where K is a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ is a function.

In the subsequent theorems, we relate the solutions of the problems (P) and $(\text{VIP})_h$.

Theorem 6.1. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. If f is h-convex and $\bar{x} \in K$ is a solution of $(\text{VIP})_h$, then $\bar{x}$ solves the problem (P).

Proof. By h-convexity of f, we have

$$f(y) - f(\bar{x}) \geq h(\bar{x}; y - \bar{x}), \quad \text{for all } y \in K.$$

Since $\bar{x}$ is a solution of $(\text{VIP})_h$, we have

$$h(\bar{x}; y - \bar{x}) \geq 0, \quad \text{for all } y \in K.$$

The last two inequalities together imply that

$$f(y) - f(\bar{x}) \geq 0, \quad \text{for all } y \in K,$$

that is, $\bar{x}$ is a solution of problem (P). □

The h-convexity assumption in Theorem 6.1 can be weakened to h-pseudoconvexity.

Theorem 6.2. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction. If f is h-pseudoconvex and $\bar{x} \in K$ is a solution of $(\text{VIP})_h$, then $\bar{x}$ solves problem (P).

Proof. Follows directly using the definition of h-pseudoconvexity. □

The following example illustrates Theorem 6.2.

Example 6.1. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} x^2, & \text{if } x \leq 0, \\ 0, & \text{if } 0 < x \leq 1, \\ x + 1, & \text{if } x > 1. \end{cases}$$

Let $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined as

$$h(x;d) = f^D(x;d) = \begin{cases} 2xd, & \text{if } x \leq 0, \\ 0, & \text{if } x = 1,\ d \leq 0 \text{ or } 0 < x < 1, \\ +\infty, & \text{if } x = 1,\ d > 0, \\ d, & \text{if } x > 1. \end{cases}$$

Then, f is h-pseudoconvex and the solution sets of the problems (P) and $(\mathrm{VIP})_h$ coincide and are given by $\{x : x \in [0,1]\}$.

The following example shows that Theorem 6.2 may not hold if the function f is not h-pseudoconvex.

Example 6.2. Let $f : \mathbb{R} \to \mathbb{R}$ be a function defined as

$$f(x) = \begin{cases} 1, & \text{if } x > 0, \\ -1, & \text{if } x \leq 0. \end{cases}$$

Let $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a bifunction defined as

$$h(x;d) = f^D(x;d) = \begin{cases} 0, & \text{if } x = 0,\ d \leq 0 \text{ or } x \neq 0, \\ +\infty, & \text{if } x = 0,\ d > 0. \end{cases}$$

Here, f is not h-pseudoconvex because the definition fails at $x = 2$ and $y = -1$. We may note that each $x > 0$ is a solution of $(\mathrm{VIP})_h$, but it is not a solution of the problem (P).

For the converse of the Theorem 6.1 to hold, we do not require the function f to be h-convex. However, we assume that the function f and the bifunction h satisfy the condition (3.6).

Theorem 6.3. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a function, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy the condition (3.6). If $\bar{x}$ is an optimal solution of the problem (P), then $\bar{x} \in K$ is a solution of $(\mathrm{VIP})_h$.

Proof. Since K is convex and $\bar{x}$ is an optimal solution of problem (P), for any $y \in K$ we have

$$f(\bar{x}) \leq f(\bar{x} + \lambda(y - \bar{x})), \quad \text{for all } \lambda \in \ ]0,1].$$

This implies that

$$\frac{f(\bar{x} + \lambda(y - \bar{x})) - f(\bar{x})}{\lambda} \geq 0, \quad \text{for all } \lambda \in \ ]0,1].$$

Taking $\liminf$ as $\lambda \to 0^+$, we obtain

$$f_D(\bar{x}; y - \bar{x}) \geq 0, \quad \text{for all } y \in K,$$

which on using (3.6) implies that

$$h(\bar{x}; y - \bar{x}) \geq 0, \quad \text{for all } y \in K.$$

Hence, $\bar{x}$ is a solution of $(\mathrm{VIP})_h$. $\square$

Thus, it is possible to identify the solutions of the optimization problem (P) with those of the $(\mathrm{VIP})_h$ provided the objective function is h-convex or h-pseudoconvex.

Next example shows that the condition (3.6) in Theorem 6.3 cannot be relaxed.

Example 6.3. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} \frac{-x}{2}, & \text{if } x < 0, \\ x^2, & \text{if } x \geq 0. \end{cases}$$

Then, for $(x, d) \in \mathbb{R} \times \mathbb{R}$,

$$f_D(x; d) = \begin{cases} \frac{-d}{2}, & \text{if } x = 0,\ d < 0 \text{ or } x < 0, \\ 2xd, & \text{if } x = 0, d \geq 0 \text{ or } x > 0. \end{cases}$$

Choose

$$h(x; d) = \begin{cases} \frac{-d}{2}, & \text{if } x \leq 0, \\ 2xd, & \text{if } x > 0, \end{cases}$$

then condition (3.6) is not satisfied at $x = 0$ and $d > 0$. Here we may note that $x = 0$ is a solution of the problem (P), but it is not a solution of $(\mathrm{VIP})_h$.

Theorem 6.4. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a function such that

$$h(x; y - x) > f(y) - f(x), \quad \text{for all } x, y \in K \text{ and } x \neq y. \tag{6.2}$$

Then, every solution of the problem (P) is a solution of $(\mathrm{VIP})_h$.

Proof. Assume that $\bar{x}$ is a solution of the problem (P) but not a solution of $(\mathrm{VIP})_h$. Then, there exists $y \in K$ such that

$$h(\bar{x}; y - \bar{x}) < 0. \tag{6.3}$$

From (6.2), we reach a contradiction to our assumption that $\bar{x}$ is a solution of the problem (P). Hence, $\bar{x}$ is a solution of $(\mathrm{VIP})_h$. □

Next we establish that a solution of the Minty variational inequality problem $(\mathrm{MVIP})_h$ is an optimal solution of the problem (P) under specific assumptions.

Theorem 6.5. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f : K \to \mathbb{R}$ be a radially lower semicontinuous function, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy condition (3.6) and be positively homogeneous in the second argument. If $\bar{x} \in K$ is a solution of $(\mathrm{MVIP})_h$, then it is a solution of the problem (P).

Proof. Let $\bar{x} \in K$ be a solution of $(\text{MVIP})_h$. Then,

$$h(y; \bar{x} - y) \leq 0, \quad \text{for all } y \in K. \tag{6.4}$$

Let $y \in K$, $y \neq \bar{x}$ be arbitrary. Since f is radially lower semicontinuous, by Theorem 2.11(b), there exists $\theta \in [0, 1[$ such that for $w = y + \theta(\bar{x} - y)$, we have

$$f_D(w; \bar{x} - y) \geq f(\bar{x}) - f(y). \tag{6.5}$$

As $\theta < 1$, by the positive homogeneity of h in the second argument, we have from relation (6.5) and the condition (3.6) that

$$(1 - \theta)^{-1} h(w; \bar{x} - w) \geq f(\bar{x}) - f(y).$$

From (6.4), we have

$$0 \geq h(w; \bar{x} - w) \geq (1 - \theta)(f(\bar{x}) - f(y)),$$

and as $\theta < 1$, it follows that $f(\bar{x}) - f(y) \leq 0$. Since $y \in K$ was arbitrary, it follows that $\bar{x}$ is a solution of problem (P). □

As in the differentiable case, the problem $(\text{MVIP})_h$ is a necessary optimality condition under the assumption of the convexity (or strict pseudoconvexity) of f.

Theorem 6.6. Let K be a nonempty convex subset of $\mathbb{R}^n$, $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction, and $f : K \to \mathbb{R}$ be a h-convex function. If $\bar{x} \in K$ is solution of problem (P), then it solves $(\text{MVIP})_h$.

Proof. Since f is h-convex, we have

$$f(\bar{x}) - f(y) - h(y; \bar{x} - y) \geq 0, \quad \text{for all } y \in K.$$

Since $\bar{x}$ is a solution of problem (P), we obtain

$$0 \geq f(\bar{x}) - f(y) \geq h(y; \bar{x} - y), \quad \text{for all } y \in K,$$

thus $\bar{x}$ solves $(\text{MVIP})_h$. □

In the following theorem, we relax the h-convexity assumption, but we add some other assumptions.

Theorem 6.7. Let K be a nonempty convex subset of $\mathbb{R}^n$, $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy condition (3.3) and be positively homogeneous and subodd in the second argument, and $f : K \to \mathbb{R}$ be a h-pseudoconvex function. If $\bar{x} \in K$ is solution of problem (P), then it solves $(\text{MVIP})_h$.

Proof. Since $\bar{x}$ is a solution of problem (P), we have

$$f(\bar{x}) \leq f(y), \quad \text{for all } y \in K.$$

By Theorem 3.9, f is h-quasiconvex, and hence

$$h(y; \bar{x} - y) \leq 0, \quad \text{for all } y \in K,$$

thus $\bar{x}$ solves $(\text{MVIP})_h$. □

Remark 6.1. It can be seen that in Example 6.1 the solution set of the problem $(\text{MVIP})_h$ also turns out to be the set $\{x : x \in [0, 1]\}$.

Remark 6.2. If f is a differentiable convex function then $h(x; y - x) = \langle \nabla f(x), y - x \rangle$, so all the conditions of Theorem 6.6 and Theorem 6.7 are automatically satisfied, and thus we can deduce Theorem 5.7 from these theorems.

6.4 Existence Criteria

An essential aspect of the study of variational inequalities is to provide sufficient conditions for the existence of their solutions. This research area has played an important role in theory and practical applications, and has facilitated many significant contributions. In this section, we present some existence results for the solutions of $(\text{VIP})_h$ and $(\text{MVIP})_h$ under different conditions on the underlying set K and the bifunction h.

Theorem 6.8. Let K be a nonempty compact convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a pseudomonotone bifunction such that h is proper subodd in the second argument and the function $x \mapsto h(y; x - y)$ is lower semicontinuous. Then, $(\text{MVIP})_h$ has a solution $\bar{x} \in K$. Furthermore, if h is upper sign continuous and positively homogeneous in the second argument, then $\bar{x} \in K$ is a solution of $(\text{VIP})_h$.

Proof. For all $y \in K$, we define two set-valued maps $P, Q : K \to 2^K$ by

$$P(y) = \{x \in K : h(x; y - x) \geq 0\}$$

and

$$Q(y) = \{x \in K : h(y; x - y) \leq 0\}.$$

We first prove that the set-valued map P is a KKM-map. Let $\{y_1, y_2, \ldots, y_m\}$ be a finite subset of K and $\hat{x} \in \text{co}(\{y_1, y_2, \ldots, y_m\})$, then $\hat{x} = \sum_{i=1}^m \lambda_i y_i$ with $\lambda_i \geq 0$ and $\sum_{i=1}^m \lambda_i = 1$. If $\hat{x} \notin \bigcup_{i=1}^m P(y_i)$, then

$$h(\hat{x}; y_i - \hat{x}) < 0, \quad \text{for each } i = 1, 2, \ldots, m.$$

Since $\lambda_i \geq 0$ with $\sum_{i=1}^m \lambda_i = 1$, we have

$$\sum_{i=1}^{m} \lambda_i h(\hat{x}; y_i - \hat{x}) < 0. \tag{6.6}$$

As

$$\sum_{i=1}^{m} \lambda_i (y_i - \hat{x}) = \sum_{i=1}^{m} \lambda_i y_i - \sum_{i=1}^{m} \lambda_i \hat{x} = \hat{x} - \hat{x} = \mathbf{0},$$

by proper suboddness of h in the second argument, we have

$$\sum_{i=1}^{m} h(\hat{x}; \lambda_i (y_i - \hat{x})) = \sum_{i=1}^{m} \lambda_i h(\hat{x}; y_i - \hat{x}) \geq 0,$$

a contradiction to the inequality (6.6). Therefore, $\mathrm{co}(\{y_1, y_2, \ldots, y_m\}) \subseteq \bigcup_{i=1}^m P(y_i)$, which imples that P is a KKM-map. The pseudomonotonicity of h implies that $P(y) \subseteq Q(y)$ for all $y \in K$, and hence Q is a KKM-map.

We claim that $Q(y)$, for all $y \in K$, is a closed set in K. Indeed, let $\{x_m\}$ be a sequence in $Q(y)$ that converges to $x \in K$. Then, $h(y; x_m - y) \leq 0$. Since h is lower semicontinuous in the second argument, we have

$$h(y; x - y) \leq \liminf_{m \to \infty} h(y; x_m - y) \leq 0.$$

Therefore, $h(y; x - y) \leq 0$, and hence $x \in Q(y)$. Thus, $Q(y)$ is closed in K.

Further, since K is compact, it follows that $Q(y)$ is compact for all $y \in K$. Then, by the Fan-KKM the theorem (Theorem B.3),

$$\bigcap_{y \in K} Q(y) \neq \emptyset,$$

that is, there exists $\bar{x} \in K$ such that

$$h(y; \bar{x} - y) \leq 0, \quad \text{for all } y \in K.$$

Thus, $\bar{x} \in K$ is a solution of $(\mathrm{MVIP})_h$.

By Lemma 6.1, $\bar{x} \in K$ is a solution of $(\mathrm{VIP})_h$. □

When K is not necessarily compact, we have the following result.

Theorem 6.9. Let K be a nonempty convex subset of $\mathbb{R}^n$, $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a pseudomonotone bifunction such that h is proper subodd in the second argument, and the function $x \mapsto h(y; x - y)$ is lower semicontinuous. Assume that there exists a nonempty compact convex subset D of K such that for all $x \in K \setminus D$, there exists $\tilde{y} \in D$ satisfying $h(\tilde{y}; x - \tilde{y}) < 0$. Then, $(\mathrm{MVIP})_h$ has a solution $\bar{x} \in D$. Furthermore, if h is upper sign continuous and positively homogeneous in the second argument, then $\bar{x} \in D$ is a solution of $(\mathrm{VIP})_h$.

Proof. Let $\{y_1, y_2, \ldots, y_m\}$ be a finite subset of K and let $Q = \text{co}(D \cup \{y_1, y_2, \ldots y_m\})$. Then, Q is compact and convex. By Theorem 6.8, there exists $\bar{x} \in Q$ such that

$$h(y; \bar{x} - y) \leq 0, \quad \text{for all } y \in Q.$$

From the hypothesis, $\bar{x} \in D$. In particular, we have $\bar{x} \in D$ such that

$$h(y_i; \bar{x} - y_i) \leq 0, \quad \text{for all } i = 1, 2, \ldots, m.$$

Since D is compact, by lower semicontinuity of the function $x \mapsto h(y; x - y)$, we have

$$G(y) = \{x \in D : h(y; x - y) \leq 0\}$$

is closed in D, and hence compact. Therefore, $\{G(y)\}_{y \in K}$ has a nonempty intersection property, and hence

$$\bigcap_{y \in K} G(y) \neq \emptyset.$$

Thus, there exists $\bar{x} \in D$ such that $h(y; \bar{x} - y) \leq 0$ for all $y \in K$.

By Lemma 6.1, $\bar{x} \in D$ is a solution of $(\text{VIP})_h$. $\square$

The next theorem shows that under the assumption of strict pseudomonotonicity of the bifunction h, the problem $(\text{VIP})_h$ possesses at most one solution.

Theorem 6.10. If K is a nonempty convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is strictly pseudomonotone, then $(\text{VIP})_h$ has a unique solution, if one exists.

Proof. Let $\bar{x}, \hat{x} \in K$, $\bar{x} \neq \hat{x}$ be two solutions of $(\text{VIP})_h$. Then, we have

$$h(\hat{x}; \bar{x} - \hat{x}) \geq 0, \tag{6.7}$$

and

$$h(\bar{x}; \hat{x} - \bar{x}) \geq 0. \tag{6.8}$$

Since h is strictly pseudomonotone, it follows from the inequality (6.8) that

$$h(\hat{x}; \bar{x} - \hat{x}) < 0,$$

a contradiction to (6.7). Hence, $\bar{x} = \hat{x}$. $\square$

The following theorem guarantees the existence of a solution of the problem $(\text{MVIP})_h$ under a proper quasimonotonicity assumption.

Theorem 6.11. Let K is a nonempty compact convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction such that the following conditions hold:

(i) h is continuous in the second argument;

(ii) h is properly quasimonotone.

Then, $(\mathrm{MVIP})_h$ is solvable.

Proof. For each $y \in K$, define a set-valued mapping $Q : K \to 2^K$ by

$$Q(y) = \{x \in K : h(y; x - y) \leq 0\}.$$

Then by assumption (ii), it follows that for each $y \in K$, $h(y; \mathbf{0}) \leq 0$ implying that $y \in Q(y)$, and hence $Q(y) \neq \emptyset$.

Next we prove that Q is a KKM-map. On the contrary, suppose that there exist $y_1, y_2, \ldots, y_m \in K$ and $x \in \mathrm{co}(\{y_1, y_2, \ldots, y_m\})$ such that $x \notin \bigcup_{i=1}^m Q(y_i)$. This implies that $h(y_i; x - y_i) > 0$ for all $i = 1, 2, \ldots, m$ which contradicts the assumption (ii). Thus, Q is a KKM-map.

Assumption (i) ensures that $Q(y)$ is a closed subset of the compact set K, and hence $Q(y)$ is compact for each $y \in K$. Then, by Theorem B.3, we have $\bigcap_{y \in K} Q(y) \neq \emptyset$, that is, $(\mathrm{MVIP})_h$ is solvable. □

The following definition of pseudomonotonicity in the sense of Brézis [37] generalizes the one considered by Chowdhury and Tan [45].

Definition 6.2. Let K be a nonempty subset of $\mathbb{R}^n$. A function $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is said to be *B-pseudomonotone* if for each $x \in K$ and every sequence $\{x_m\}$ in K converging to x with

$$\liminf_{m \to \infty} h(x_m; x - x_m) \geq 0,$$

we have

$$h(x; y - x) \geq \limsup_{m \to \infty} h(x_m; y - x_m), \quad \text{for all } y \in K.$$

We present the following existence result for a solution of $(\mathrm{VIP})_h$ under the B-pseudomonotonicity assumption.

Theorem 6.12. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be B-pseudomonotone such that for each finite subset A of K, the map $x \mapsto h(x; y - x)$ is upper semicontinuous on $\mathrm{co}(A)$, the map $y \mapsto h(x; y - x)$ is convex, and $h(x; \mathbf{0}) = 0$ for all $x \in K$. Assume that there exist a nonempty compact subset D of K and an element $\tilde{y} \in D$ such that for all $x \in K \setminus D$, $h(x; \tilde{y} - x) < 0$. Then, $(\mathrm{VIP})_h$ has a solution.

Proof. For each $x \in K$, define a set-valued map $P : K \to 2^K$ by

$$P(x) = \{y \in K : h(x; y - x) < 0\}.$$

Then by convexity of the map $y \mapsto h(x; y - x)$, $P(x)$ is convex for all $x \in K$. Let A be a finite subset of K. Then for all $y \in \mathrm{co}(A)$,

$$[P^{-1}(y)]^c \cap \mathrm{co}(A) = \{x \in \mathrm{co}(A) : h(x; y - x) \geq 0\}$$

is closed in $\operatorname{co}(A)$ by upper semicontinuity of the map $x \mapsto h(x; y-x)$ on $\operatorname{co}(A)$. Hence, $P^{-1}(y) \cap \operatorname{co}(A)$ is open in $\operatorname{co}(A)$.

Suppose that $x, y \in \operatorname{co}(A)$ and $\{x_m\}$ is a sequence in K converging to x such that

$$h(x_m; \lambda y + (1-\lambda)x - x_m) \geq 0, \quad \text{for all } m \in \mathbb{N} \text{ and all } \lambda \in [0,1].$$

For $\lambda = 0$, we have

$$h(x_m; x - x_m) \geq 0, \quad \text{for all } m \in \mathbb{N},$$

and therefore,

$$\liminf_{m\to\infty} h(x_m; x - x_m) \geq 0.$$

By B-pseudomonotonicity of h, we have

$$h(x; y - x\rangle \geq \limsup_{m\to\infty} h(x_m; y - x_m). \tag{6.9}$$

For $\lambda = 1$, we have

$$h(x_m; y - x_m) \geq 0, \quad \text{for all } m \in \mathbb{N},$$

and therefore,

$$\liminf_{m\to\infty} h(x_m; y - x_m) \geq 0. \tag{6.10}$$

From inequalities (6.9) and (6.10), we obtain

$$h(x; y - x) \geq 0,$$

and thus $y \notin P(x)$.

Assume that for all $x \in D$, $P(x)$ is nonempty. Then, all the conditions of Theorem B.5 are satisfied. Hence, there exists $\hat{x} \in K$ such that $\hat{x} \in P(\hat{x})$, that is,

$$0 = h(\hat{x}; \hat{x} - \hat{x}\rangle < 0,$$

a contradiction. Thus, there exists $\bar{x} \in D \subseteq K$ such that $P(\bar{x}) = \emptyset$, that is,

$$h(\bar{x}; y - \bar{x}) \geq 0, \quad \text{for all } y \in K.$$

Hence, $\bar{x}$ is a solution of $(\text{VIP})_h$. □

Corollary 6.1. Let K be a nonempty closed and convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be B-pseudomonotone such that for each finite subset A of K, $x \mapsto h(x; y-x)$ is upper semicontinuous on $\operatorname{co}(A)$, the map $y \mapsto h(x; y-x)$ is convex, and $h(x; \mathbf{0}) = 0$ for all $x \in K$. Assume that there exists $\tilde{y} \in K$ such that

$$\lim_{\|x\|\to\infty,\ x\in K} h(x; \tilde{y} - x) < 0. \tag{6.11}$$

Then, $(\text{VIP})_h$ has a solution.

Proof. Let

$$\alpha = \lim_{\|x\|\to\infty,\ x\in K} h(x;\tilde{y}-x).$$

By inequality (6.11), $\alpha < 0$. Let $r > 0$ be such that $\|\tilde{y}\| \leq r$ and

$$h(x;\tilde{y}-x) < \frac{\alpha}{2}, \quad \text{for all } x \in K \text{ with } \|x\| > r.$$

Let $\mathbb{B}_r = \{x \in K : \|x\| \leq r\}$. Then $\mathbb{B}_r$ is a nonempty and compact subset of K. Note that for any $x \in K \setminus \mathbb{B}_r$, $h(x;\tilde{y}-x) < \frac{\alpha}{2} < 0$, and the conclusion follows from Theorem 6.12. □

Now we present some existence results for a solution of $(\text{VIP})_h$ without any kind of monotonicity.

Theorem 6.13. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction such that the following conditions hold:

(i) for all $x \in K$, $h(x;\mathbf{0}) \geq 0$;

(ii) for all $x \in K$, the function $y \mapsto h(x;y-x)$ is convex;

(iii) $\liminf_{x\to x^*} h(x;y-x) \leq h(x^*,y-x^*)$ for all $y \in K$ whenever $x \to x^*$ in K;

(iv) there exist a nonempty compact convex subset C of K and a nonempty compact subset D of K such that for each $x \in K \setminus D$, there exists $\tilde{y} \in C$ such that $h(x;\tilde{y}-x) < 0$.

Then, $(\text{VIP})_h$ has a solution in K.

Proof. For all $y \in K$, define

$$S(y) = \{x \in K : h(x,y-x) \geq 0\}.$$

Then, the solution set of $(\text{VIP})_h$ is $S := \bigcap_{y\in K} S(y)$. We note that for each $y \in K$, $S(y)$ is closed. Indeed, let $\{x_m\}$ be a sequence in $S(y)$ such that $x_m \to x \in K$. Then by condition (iii), we have

$$h(x;y-x) \geq \liminf_{x_m\to x} h(x_m;y-x_m) \geq 0, \quad \text{for all } y \in K.$$

Hence $x \in S(y)$. So for all $y \in K$, $S(y)$ is closed.

Now we prove that the solution set S is nonempty. Assume, to the contrary, $S = \emptyset$. Then for each $x \in K$, the set

$$Q(x) = \{y \in K : x \notin S(y)\} = \{y \in K : h(x;y-x) < 0\} \neq \emptyset.$$

Since h is convex in the second variable, we have for each $x \in K$, $Q(x)$ is convex. Thus, $Q : K \to 2^K$ defines a set-valued map such that for each $x \in K$, $Q(x)$ is nonempty and convex. Note that for each $y \in K$, the set

$$\begin{aligned} Q^{-1}(y) &= \{x \in K : y \in Q(x)\} = \{x \in K : h(x; y - x) < 0\} \\ &= \{x \in K : h(x; y - x) \geq 0\}^c = [S(y)]^c \end{aligned}$$

is open in K.

Thus, the set-valued map $Q : K \to 2^K$ satisfies all the conditions of Theorem B.4, and therefore there exists a point $\hat{x} \in K$ such that $\hat{x} \in Q(\hat{x})$, that is, $0 \leq h(\hat{x}, \hat{x} - \hat{x}) < 0$, a contradiction. Hence, the solution set S is nonempty. $\square$

Remark 6.3. The condition (iv) of Theorem 6.13 can be repalced by the following condition:

(iv)$'$ There exists $\tilde{y} \in K$ such that $\liminf\limits_{\|x\| \to +\infty} h(x; \tilde{y} - x) < 0$.

Proof. By condition (iv)$'$, there exists $r > 0$ such that $\|\tilde{y}\| < r$ and if $x \in K$ with $\|x\| \geq r$, we have $h(x; \tilde{y} - x) < 0$. Let $\mathbb{B}_r[\mathbf{0}]$ be a closed ball of radius r. Then, $\mathbb{B}_r[\mathbf{0}]$ is nonempty compact convex subset of $\mathbb{R}^n$. By taking $C = D = \mathbb{B}_r[\mathbf{0}] \cap K$ in assumption (iv) of Theorem 6.13, we get the conclusion. $\square$

If K is a compact subset of $\mathbb{R}^n$, then condition (iv) of Theorem 6.13 is automatically satisfied. Hence, we have the following result.

Corollary 6.2. Let K be a nonempty compact convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction such that the following conditions hold:

(i) for all $x \in K$, $h(x; \mathbf{0}) \geq 0$;

(ii) for all $x \in K$, the function $y \mapsto h(x; y - x)$ is convex;

(iii) $\liminf\limits_{x \to x^*} h(x; y - x) \leq h(x^*, y - x^*)$ for all $y \in K$ whenever $x \to x^*$ in K.

Then, $(\text{VIP})_h$ has a solution in K.

Remark 6.4. Condition (iii) in Theorem 6.13 and Corollary 6.2 are satisfied if h is continuous in both the arguments.

We now give an example in support of Corollary 6.2.

Example 6.4. Let $K = [-2, 3]$ and $h : K \times \mathbb{R} \to \mathbb{R}$ be defined as

$$h(x; d) = x^2 + d.$$

Then, h is continuous in both the arguments, and conditions (i) and (ii) of Corollary 6.2 are also satisfied. It can be seen that the solution set of $(\text{VIP})_h$ is given by $[-2, -1] \cup [2, 3]$.

6.5 Gap Functions and Saddle Point Characterization

This section deals with the study of an Auslender type [22] gap function for $(\mathrm{VIP})_h$.

Definition 6.3. An extended real-valued function $\varphi : K \to \mathbb{R} \cup \{-\infty\}$ is said to be a *gap function* for $(\mathrm{VIP})_h$ if it satisfies the following properties:

(i) $\varphi(x) \leq 0$ for all $x \in K$;

(ii) $\varphi(\bar{x}) = 0$ if and only if $\bar{x}$ is a solution of $(\mathrm{VIP})_h$.

Theorem 6.14. The function $\varphi(x)$ defined by

$$\varphi(x) = \inf_{y \in K} h(x; y - x),$$

is a gap function for $(\mathrm{VIP})_h$, provided that $h(x; \mathbf{0}) = 0$ for all $x \in K$.

Proof. For $x \in K$,

$$\varphi(x) = \inf_{y \in K} h(x; y - x) \leq h(x; \mathbf{0}) = 0.$$

Thus, $\varphi(x) \leq 0$ for all $x \in K$.

If $\varphi(\bar{x}) = 0$, then

$$\inf_{y \in K} h(\bar{x}; y - \bar{x}) = 0,$$

which implies that

$$h(\bar{x}; y - \bar{x}) \geq 0, \quad \text{for all } y \in K.$$

Thus, $\bar{x}$ solves $(\mathrm{VIP})_h$.

Conversely, suppose $\bar{x}$ solves $(\mathrm{VIP})_h$, then

$$h(\bar{x}; y - \bar{x}) \geq 0, \quad \text{for all } y \in K$$

which implies that

$$\inf_{y \in K} h(\bar{x}; y - \bar{x}) \geq 0.$$

But as $h(\bar{x}; \bar{x} - \bar{x}) = 0$ we conclude that

$$\inf_{y \in K} h(\bar{x}; y - \bar{x}) = 0,$$

that is, $\varphi(\bar{x}) = 0$. Thus, $\varphi(x)$ is a gap function for $(\mathrm{VIP})_h$. □

Remark 6.5. An element $\bar{x} \in K$ is a solution to $(\mathrm{VIP})_h$ if and only if $\bar{x}$ solves

$$(\mathrm{MP}) \qquad \max \varphi(x) \quad \text{subject to } x \in K$$

and $\varphi(\bar{x}) = 0$.

Example 6.5. Let $K = [0,1] \times [0,1]$ and $h : K \times \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$h(x; d) = x_1 d_1 + d_2^2,$$

where $x = (x_1, x_2)$ and $d = (d_1, d_2)$. Then, the gap function for $(\text{VIP})_h$ is given by

$$\begin{aligned} \varphi(x) &= -\sup_{y \in K} \left\{ x_1(x_1 - y_1) + (x_2 - y_2)^2 \right\} \\ &= -\max_{y \in K} \left\{ -x_1 y_1 + (x_2 - y_2)^2 \right\} - x_1^2 \\ &= -x_1^2. \end{aligned}$$

The problem (MP) has an optimal at the points $\{(0, x_2) : x_2 \in [0,1]\}$ and thus by Remark 6.5 this set is the set of solutions of $(\text{VIP})_h$.

One of the many useful applications of gap functions is in deriving the so-called error bounds, that is, upper estimates on the distance to the solution set of $(\text{VIP})_h$. We now obtain an upper bound for the gap function φ using the notion of strong monotonicity of h.

Theorem 6.15. Let $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ satisfy the following conditions:

(i) h is positively homogeneous in the second argument;

(ii) h is strongly monotone with modulus $\sigma > 0$;

(iii) $h(x; \mathbf{0}) = 0$ for all $x \in K$.

If $\bar{x} \in K$ is a solution of $(\text{VIP})_h$, then there exists a positive constant μ such that $\varphi(x) \le -\mu \|x - \bar{x}\|^2$ for all $x \in K$.

Proof. Let $\bar{x} \in K$ be a solution of $(\text{VIP})_h$. Then

$$h(\bar{x}; y - \bar{x}) \ge 0, \quad \text{for all } y \in K. \tag{6.12}$$

Since h is strongly monotone with modulus $\sigma > 0$, we have

$$h(\bar{x}; y - \bar{x}) + h(y; \bar{x} - y) \le -\sigma \|y - \bar{x}\|^2, \quad \text{for all } y \in K. \tag{6.13}$$

From inequalities (6.12) and (6.13), we obtain

$$h(y; \bar{x} - y) \le -\sigma \|y - \bar{x}\|^2, \quad \text{for all } y \in K. \tag{6.14}$$

For any $x \in K$, let $y = \bar{x} + \lambda(x - \bar{x})$ for $\lambda \in \ ]0,1[$, then

$$\varphi(x) \le h(x; \bar{x} + \lambda(x - \bar{x}) - x) = h(x; (1 - \lambda)(\bar{x} - x)).$$

By positive homogeneity of h in the second argument, we obtain $\varphi(x) \le (1 - \lambda) h(x; \bar{x} - x)$, and taking $y = x$ in (6.14), it follows that

$$\varphi(x) \le -\sigma(1 - \lambda)\|x - \bar{x}\|^2 = -\mu \|x - \bar{x}\|^2, \quad \text{where } \mu = \sigma(1 - \lambda) > 0.$$

This completes the proof. □

Theorem 6.16. If $h(x;\mathbf{0}) = 0$ for all $x \in K$, then the function $\varphi' : K \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\varphi'(y) = \sup_{x \in K} h(x; y - x)$$

is a gap function for $(\text{MVIP})_h$, that is, it satisfies the conditions

(i) $\varphi'(y) \geq 0$ for all $y \in K$;

(ii) $\varphi'(\bar{y}) = 0$ if and only if $\bar{y}$ solves $(\text{MVIP})_h$.

Define the Lagrangian function $\psi : K \times K \to \mathbb{R} \cup \{\pm\infty\}$ as

$$\psi(x, y) = h(x; y - x). \tag{6.15}$$

Then $\bar{x} \in K$ is a solution of $(\text{VIP})_h$ if and only if

$$\psi(\bar{x}, y) \geq 0, \quad \text{for all } y \in K,$$

and $\bar{y} \in K$ is a solution of $(\text{MVIP})_h$ if and only if

$$\psi(x, \bar{y}) \leq 0, \quad \text{for all } x \in K.$$

Definition 6.4. A pair $(\bar{x}, \bar{y}) \in K \times K$ is said to be a *saddle point* of the Lagrangian ψ if for every $x, y \in K$ we have

$$\psi(x, \bar{y}) \leq \psi(\bar{x}, \bar{y}) \leq \psi(\bar{x}, y).$$

It can be easily seen that $(\bar{x}, \bar{y}) \in K \times K$ is a saddle point of the Lagrangian ψ if and only if

$$\sup_{x \in K} \inf_{y \in K} \psi(x, y) = \inf_{y \in K} \sup_{x \in K} \psi(x, y) = \psi(\bar{x}, \bar{y}).$$

Theorem 6.17. Let $h(x;\mathbf{0}) = 0$ for all $x \in K$ and ψ be defined as in (6.15). Then,

(a) $(\text{VIP})_h$ admits a solution $\bar{x}$ if and only if

$$\sup_{x \in K} \inf_{y \in K} \psi(x, y) = 0,$$

and the supremum is attained at $\bar{x}$;

(b) $(\text{MVIP})_h$ admits a solution $\bar{y}$ if and only if

$$\inf_{y \in K} \sup_{x \in K} \psi(x, y) = 0,$$

and the infimum is attained at $\bar{y}$;

(c) $\bar{x}$ and $\bar{y}$ are solutions of $(\text{VIP})_h$ and $(\text{MVIP})_h$, respectively, if and only if $(\bar{x}, \bar{y})$ is a saddle point of ψ on $K \times K$.

Proof. (a) Since $\varphi(x) = \inf_{y \in K} \psi(x, y)$ is a gap function for $(\text{VIP})_h$, we have $\bar{x} \in K$ is a solution of $(\text{VIP})_h$ if and only if $\varphi(\bar{x}) = 0$, that is,

$$0 = \varphi(\bar{x}) = \sup_{x \in K} \inf_{y \in K} \psi(x, y).$$

(b) Since $\varphi'(y) = \sup_{x \in K} \psi(x, y)$ is a gap function for $(\text{MVIP})_h$, therefore, it follows that $\bar{y} \in K$ is a solution of $(\text{MVIP})_h$ if and only if

$$0 = \varphi'(\bar{y}) = \inf_{y \in K} \sup_{x \in K} \psi(x, y).$$

(c) Let $\bar{x}$ and $\bar{y}$ be the solutions of $(\text{VIP})_h$ and $(\text{MVIP})_h$, respectively. Then from parts (a) and (b), it follows that $(\bar{x}, \bar{y})$ is a saddle point of ψ on $K \times K$.

Conversely, suppose $(\bar{x}, \bar{y}) \in K \times K$ is a saddle point of ψ on $K \times K$, that is,

$$\begin{aligned} h(x; \bar{y} - x) &\leq h(\bar{x}; \bar{y} - \bar{x}) \\ &\leq h(\bar{x}; y - \bar{x}), \quad \text{for all } (x, y) \in K \times K. \end{aligned}$$

Taking $x = \bar{y}$ and $y = \bar{x}$ in the above inequalities, we get

$$h(\bar{x}; \bar{y} - \bar{x}) = 0,$$

which implies that $\bar{x}$ solves $(\text{VIP})_h$ and $\bar{y}$ solves $(\text{MVIP})_h$. □

Chapter 7

Characterizations of Solution Sets of Optimization Problems and Nonsmooth Variational Inequalities

7.1 Introduction

During the last few decades, optimization theory has been evolving in all possible directions at an astonishing rate. New algorithmic and theoretical techniques have been developed, diffusion into other disciplines has proceeded at a rapid pace, and our knowledge of all aspects of the field has grown even more profound. Due to the recent developments, the theory of optimization is becoming increasingly pivotal in mathematics as well as interdisciplinary areas, especially in the interplay between mathematics and many other sciences like computer science, physics, engineering, and operations research. The study of variational inequalities is also a part of this development because solutions of optimization problems can often be related to the solutions of variational inequalities.

An important study in mathematical programming is the characterization of the solution set of an optimization problem having multiple solutions and it is fundamental for understanding the behavior and development of solution methods for such problems. The objective of this chapter is to study the characterizations of the solution sets of optimization problems having h-convex or h-pseudolinear objective function and pseudoaffine variational inequality problems when one of the solution is known.

7.2 Characterizations of the Solution Set of an Optimization Problem with a Pseudolinear Objective Function

In this section, we characterize the solution set of the following optimization problem:

$$\text{(P)} \qquad \min f(x), \quad \text{subject to } x \in K,$$

where K is a nonempty subset of $\mathbb{R}^n$, and $f : K \to \mathbb{R}$ is a function. Let $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction associated to f such that h is positively homogeneous in the second argument and satisfies conditions (3.3) and (3.6).

Throughout this section, we denote by S the solution set of the problem (P) and assume that S is nonempty. By combining conditions (3.3) and (3.6), we mention the following condition:

$$f_D(x; d) \leq h(x; d) \leq f^D(x; d), \quad \text{for all } x \in K \text{ and } d \in \mathbb{R}^n. \tag{7.1}$$

Theorem 7.1. Let K be a nonempty convex subset of $\mathbb{R}^n$, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction satisfying condition (7.1) and be positively homogeneous in the second argument. Let the objective function $f : K \to \mathbb{R}$ be h-pseudolinear. If one of the following conditions hold:

(i) f is a radially continuous function;

(ii) h is odd in the second argument;

then the solution set S of problem (P) is convex.

Proof. Let $x_1, x_2 \in S$. Then, $f(x_1) = f(x_2)$, and by Theorem 3.14, we have

$$h(x_1; x_2 - x_1) = 0 \quad \text{and} \quad h(x_2; x_1 - x_2) = 0.$$

Also, for $\lambda \in [0, 1]$, we have

$$\begin{aligned} h(x_2; \lambda x_1 + (1-\lambda)x_2 - x_2) &= h(x_2; \lambda(x_1 - x_2)) \\ &= \lambda h(x_2; x_1 - x_2) = 0. \end{aligned}$$

Again invoking Theorem 3.14, it follows that

$$f(\lambda x_1 + (1-\lambda)x_2) = f(x_2),$$

and thus we have

$$\lambda x_1 + (1-\lambda)x_2 \in S, \quad \text{for all } \lambda \in [0, 1].$$

Hence, S is a convex set. □

We now characterize the solution set of the problem (P) in terms of any of its solution points when the objective function is h-pseudolinear.

Theorem 7.2. Let K be a nonempty convex subset of $\mathbb{R}^n$, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction satisfying condition (7.1) and be positively homogeneous in the second argument. Let the objective function $f : K \to \mathbb{R}$ be h-pseudolinear. Further, suppose that one of the following conditions hold:

(i) f is a radially continuous function;

(ii) h is odd in the second argument.

If $\bar{x} \in S$, then $S = S_1 = S_2$, where

$$\begin{aligned} S_1 &= \{x \in K : h(x; \bar{x} - x) = 0\}, \\ S_2 &= \{x \in K : h(\bar{x}; x - \bar{x}) = 0\}. \end{aligned}$$

Proof. A point $x \in S$ if and only if $f(x) = f(\bar{x})$. Then from Theorem 3.14, we have $f(x) = f(\bar{x})$ if and only if $h(x; \bar{x} - x) = 0$. Thus, $S = S_1$.

Similarly, we can prove that $S = S_2$. $\square$

Example 7.1. To exemplify Theorem 7.2, consider the problem (P) where the function $f : \mathbb{R}^2 \to \mathbb{R}$ is given by

$$f(x_1, x_2) = \begin{cases} x_1 + x_2, & \text{if } x_1 + x_2 \geq 0, \\ (x_1 + x_2)/2, & \text{if } x_1 + x_2 < 0. \end{cases}$$

Let K be given by

$$K = \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 \geq -1,\ x_1 \leq 0,\ x_2 \leq 0\right\}.$$

Clearly, $\bar{x} = (0, -1)$ is a solution of the problem (P). If $h(x; d) = f^D(x; d)$ for $(x; d) \in K \times \mathbb{R}^2$, then by using the above characterization the solution set S is given by

$$\begin{aligned} S &= \{x \in K : h(\bar{x}; x - \bar{x}) = 0\} \\ &= \{(x_1, x_2) \in K : (x_1 + x_2 + 1)/2 = 0\} \\ &= \{(x_1, x_2) \in K : x_1 + x_2 + 1 = 0\} \\ &= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 = 1,\ x_1 \leq 0,\ x_2 \leq 0\}. \end{aligned}$$

Theorem 7.3. Let K be a nonempty convex subset of $\mathbb{R}^n$, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction satisfying condition (7.1) and be both odd as well as positively homogeneous in the second argument. Let the objective function $f : K \to \mathbb{R}$ be h-pseudolinear. If $\bar{x} \in S$, then the solution set S can also be characterized as

$$S = S_3 = \{x \in K : h(\lambda x + (1 - \lambda)\bar{x}; \bar{x} - x) = 0, \text{ for all } \lambda \in [0, 1]\}.$$

Proof. Let $x \in S$. By Theorem 7.1, we have $\lambda x+(1-\lambda)\bar{x} \in S$ for all $\lambda \in\ ]0,1]$, and therefore from Theorem 7.3 it follows that

$$\begin{aligned} 0 &= h(\lambda x + (1-\lambda)\bar{x}; \bar{x} - \lambda x - (1-\lambda)\bar{x}) \\ &= \lambda h(\lambda x + (1-\lambda)\bar{x}; \bar{x} - x). \end{aligned}$$

As $\lambda > 0$, we have

$$h(\lambda x + (1-\lambda)\bar{x}; \bar{x} - x) = 0.$$

Also since $x \in S$, from Theorem 7.2 and the oddness of h, we have

$$h(\bar{x}; \bar{x} - x) = 0.$$

Thus, from the previous two statements we conclude that $S \subseteq S_3$.

Conversely, let $x \in S_3$. Then by taking $\lambda = 1$ in particular, we have

$$h(x; \bar{x} - x) = 0,$$

and hence from Theorem 7.2 it follows that $x \in S$. Thus, we have $S_3 \subseteq S$. Hence, $S = S_3$. $\square$

Combining Theorem 7.2 and Theorem 7.3, we immediately deduce the following result of Jeyakumar and Yang [115].

Corollary 7.1. Let the function f in the problem (P) be continuously differentiable pseudolinear and $\bar{x}$ be any point of the solution set S of problem (P). Then, $S = \hat{S}_1 = \hat{S}_2 = \hat{S}_3$, where

$$\begin{aligned} \hat{S}_1 &= \{x \in K : \langle \nabla f(\bar{x}), \bar{x} - x\rangle = 0\}, \\ \hat{S}_2 &= \{x \in K : \langle \nabla f(x), \bar{x} - x\rangle = 0\}, \\ \hat{S}_3 &= \{x \in K : \langle \nabla f(\lambda\bar{x} + (1-\lambda)x), \bar{x} - x\rangle = 0 \text{ for all } \lambda \in [0,1]\}. \end{aligned}$$

Theorem 7.4. Let K be a nonempty convex subset of $\mathbb{R}^n$, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction satisfying condition (7.1) and be positively homogeneous in the second argument. Let the objective function $f : K \to \mathbb{R}$ be h-pseudolinear. Further, suppose that one of the following conditions hold:

(i) f is a radially continuous function;

(ii) h is odd in the second argument.

If $\bar{x} \in S$, then $S = S_4 = S_5$, where

$$\begin{aligned} S_4 &= \{x \in K : h(x; \bar{x} - x) \geq 0\}, \\ S_5 &= \{x \in K : h(\bar{x}; x - \bar{x}) \leq 0\}. \end{aligned}$$

Proof. The inclusion $S \subseteq S_4$ follows immediately from Theorem 7.2. For the converse inclusion, assume that $x \in K$ satisfies

$$h(x; \bar{x} - x) \geq 0.$$

As f is h-pseudoconvex, it follows that

$$f(x) \leq f(\bar{x}),$$

but $\bar{x}$ is a solution of problem (P), and hence $f(\bar{x}) \leq f(x)$. Therefore, $f(x) = f(\bar{x})$, that is, $x \in S$. Thus, $S = S_4$.

Similarly, we can prove $S = S_5$ by using the h-pseudoconcavity of f. $\square$

Theorem 7.5. Let K be a nonempty convex subset of $\mathbb{R}^n$, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction satisfying (7.1) and be both odd as well as positively homogeneous in the second argument. Let the objective function $f : K \to \mathbb{R}$ be h-pseudolinear. If $\bar{x} \in S$, then the solution set $S = S_6$, where

$$S_6 = \{x \in K : h(\lambda x + (1 - \lambda)\bar{x}; \bar{x} - x) \geq 0, \text{ for all } \lambda \in [0, 1]\}.$$

Proof. We note that $S_6 \subseteq S_4 = S$. Let $x \in S$, then from Theorem 7.3 we have

$$h(\lambda x + (1 - \lambda)\bar{x}; \bar{x} - x) = 0, \quad \text{for all } \lambda \in [0, 1].$$

This implies that $x \in S_6$, and hence $S \subseteq S_6$. Thus, $S = S_6$. $\square$

The following result due to Jeyakumar and Yang [115] is a direct consequence of Theorem 7.5.

Corollary 7.2. Let the function f in the problem (P) be continuously differentiable pseudolinear and $\bar{x}$ be any point of the solution set S of problem (P). Then, $S = \hat{S}_4 = \hat{S}_5 = \hat{S}_6$, where

$$\begin{aligned} \hat{S}_4 &= \{x \in K : \langle \nabla f(x), \bar{x} - x \rangle \geq 0\} \\ \hat{S}_5 &= \{x \in K : \langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq 0\} \\ \hat{S}_6 &= \{x \in K : \langle \nabla f(\lambda \bar{x} + (1 - \lambda)x), \bar{x} - x \rangle \geq 0 \text{ for all } \lambda \in [0, 1]\}. \end{aligned}$$

Theorem 7.6. Let K be a nonempty convex subset of $\mathbb{R}^n$, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction satisfying (7.1) and be positively homogeneous in the second argument. Let the objective function $f : K \to \mathbb{R}$ be h-pseudolinear. Further, suppose that one of the following conditions hold:

(i) f is a radially continuous function;

(ii) h is odd in the second argument.

If $\bar{x} \in S$, then $S = S^* = S^{**}$, where

$$\begin{aligned} S^* &= \{x \in K : h(\bar{x}; x - \bar{x}) = h(x; \bar{x} - x)\}, \\ S^{**} &= \{x \in K : h(\bar{x}; x - \bar{x}) \leq h(x; \bar{x} - x)\}. \end{aligned}$$

Proof. It is clear from Theorem 7.2 that $S \subseteq S^*$. Let $x \in K$ be such that

$$h(\bar{x}; x - \bar{x}) \leq h(x; \bar{x} - x).$$

Suppose that $x \notin S$. Then $f(x) > f(\bar{x})$. Using h-pseudoconcavity of f, we have $h(\bar{x}; x - \bar{x}) > 0$, and thus $h(x; \bar{x} - x) > 0$. By applying h-pseudoconvexity of f, we get $f(x) \leq f(\bar{x})$, which is a contradiction. Thus, $x \in S$, and hence $S^{**} \subseteq S$. It follows that $S \subseteq S^* \subseteq S^{**} \subseteq S$, and therefore $S = S^* = S^{**}$. □

An easy deduction from Theorem 7.6 is the following result proved by Jeyakumar and Yang [115].

Corollary 7.3. For the problem (P), assume that f is continuously differentiable and pseudolinear. If $\bar{x} \in S$, then $S = \hat{S}^* = \hat{S}^{**}$, where

$$\begin{aligned} \hat{S}^* &= \{x \in K : \langle \nabla f(\bar{x}), \bar{x} - x \rangle = \langle \nabla f(x), x - \bar{x} \rangle\}, \\ \hat{S}^{**} &= \{x \in K : \langle \nabla f(\bar{x}), \bar{x} - x \rangle \geq \langle \nabla f(x), x - \bar{x} \rangle\}. \end{aligned}$$

7.3 Characterizations of Solution Sets of Variational Inequalities Involving Pseudoaffine Bifunctions

In this section, we focus our study on variational inequalities involving pseudoaffine bifunctions.

Let $\bar{S} = \{x \in K : h(x; y - x) \geq 0,$ for all $y \in K\}$ denote the solution set of $(\text{VIP})_h$, and assume that $\bar{S}$ is nonempty.

The following theorem is a direct consequence of pseudoaffineness of the bifunction h.

Theorem 7.7. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a pseudoaffine bifunction. Then, $\bar{x} \in K$ is a solution of $(\text{VIP})_h$ if and only if $h(x; \bar{x} - x) \leq 0$, for all $x \in K$.

Theorem 7.8. Let K be a nonempty convex subset of $\mathbb{R}^n$, and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a pseudoaffine bifunction such that it is radially continuous in the first argument and both odd as well as positively homogeneous in the second argument. If $\bar{x} \in \bar{S}$, then

$$\begin{aligned} \bar{S} &\subseteq \{x \in K : h(x; \bar{x} - x) = 0\} \\ &= \{x \in K : h(\bar{x}; x - \bar{x}) = 0\} \\ &= \{x \in K : h(\lambda x + (1 - \lambda)\bar{x}; \bar{x} - x) = 0 \text{ for all } \lambda \in [0, 1]\}. \end{aligned}$$

Proof. Let $x \in \bar{S}$, then

$$h(x; y - x) \geq 0, \quad \text{for all } y \in K.$$

In particular, for $y = \bar{x}$ we have $h(x; \bar{x} - x) \geq 0$. As $\bar{x}$ is a solution of $(\mathrm{VIP})_h$, from Theorem 7.7, it follows that

$$h(x; \bar{x} - x) \leq 0.$$

Thus, we have

$$h(x; \bar{x} - x) = 0,$$

which implies that

$$\bar{S} \subseteq \{x \in K : h(x; \bar{x} - x) = 0\}.$$

The rest of the result follows from Theorem 4.26 and Corollary 4.9. $\square$

Remark 7.1. Theorem 7.8 gives a partial characterization of the solution set $\bar{S}$, that is, the implication in Theorem 7.8 is, in general, one way. If $\bar{x} \in \bar{S}$ and $h(\bar{x}; x - \bar{x}) = 0$ for some $x \in K$, it does not always imply that $x \in \bar{S}$, as can be seen from the following example.

Example 7.2. Let $K = [-1, 1] \times [-1, 1]$ and let $h : K \times \mathbb{R}^2 \to \mathbb{R}$ be defined as

$$h(x; d) = \begin{cases} (1 + x_1^2 + x_2^2) d_1^3 / d_2^2, & \text{if } d_2 \neq 0, \\ 0, & \text{if } d_2 = 0, \end{cases}$$

where $d = (d_1, d_2)$ and $x = (x_1, x_2)$. We note that h is a continuous function of x and is both odd as well as positively homogeneous as a function of d. By using condition (4.11), it can be easily verified that h is a pseudoaffine bifunction. Clearly, $\bar{x} = (-1, 1)$ is a solution of $(\mathrm{VIP})_h$ and $x = (1, 1)$ satisfies $h(\bar{x}; x - \bar{x}) = 0$ but $x \notin \bar{S}$ because for $y = (0, 1/2)$, $h(x; y - x) = -12 < 0$.

Definition 7.1. Let K be a nonempty convex subset of $\mathbb{R}^n$. A mapping $F : K \to \mathbb{R}^n$ is said to be a *G-map* if there exists a positive function $k(x, y)$ on $K \times K$ such that for all $x, y \in K$,

$$\langle F(x), y - x \rangle = 0 \quad \Rightarrow \quad F(x) = k(x, y) F(y).$$

Bianchi and Schaible [34] proved that if $F : K \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a continuous G-map, then the solution set $\bar{S}$ coincides with the set $\{x \in K : h(\bar{x}; x - \bar{x}) = 0\}$, where $h(\bar{x}; x - \bar{x}) = \langle F(\bar{x}), x - \bar{x} \rangle$. Jeyakumar and Yang [114] proved that if F is the gradient of a pseudolinear function, then $\bar{S} = \{x \in K : \langle F(x), \bar{x} - x \rangle = 0\}$.

We prove that the solution set $\bar{S}$ of $(\mathrm{VIP})_h$ coincides with the set $\{x \in K : h(\bar{x}; x - \bar{x}) = 0\}$ if h is a bifunction associated to some h-pseudolinear function.

Theorem 7.9. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction, and let $f : K \to \mathbb{R}$ be an h-pseudolinear function such that h satisfies condition (7.1) and is positively homogeneous in the second argument. Suppose that one of the following conditions hold:

(i) f is a radially continuous function;

(ii) h is odd in the second argument.

If $\bar{x} \in \bar{S}$, then

$$\begin{aligned} \bar{S} &= \{x \in K : h(x; \bar{x} - x) = 0\} \\ &= \{x \in K : h(\bar{x}; x - \bar{x}) = 0\}. \end{aligned}$$

Proof. The result follows from Theorems 6.2, 6.3, and 7.2. □

Theorem 7.10. Let K be a nonempty convex subset of $\mathbb{R}^n$, $h : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a bifunction, and $f : K \to \mathbb{R}$ be an h-pseudolinear function such that h satisfies condition (7.1) and is both odd as well as positively homogeneous in the second argument. If $\bar{x} \in \bar{S}$, then

$$\bar{S} = \{x \in K : h(\lambda x + (1 - \lambda)\bar{x}; \bar{x} - x) = 0 \text{ for all } \lambda \in [0, 1]\}.$$

Proof. The result follows from Theorems 6.2, 6.3, and 7.3. □

Remark 7.2. Under the assumptions of Theorem 7.8, it can be shown using Theorems 6.2, 6.3, and 7.4 that

$$\begin{aligned} \bar{S} &= \{x \in K : h(x; \bar{x} - x) \geq 0\} \\ &= \{x \in K : h(\bar{x}; x - \bar{x}) \leq 0\}. \end{aligned}$$

7.4 Lagrange Multiplier Characterizations of the Solution Set of an Optimization Problem

The subject of devising necessary and sufficient conditions for the existence of an optimum point lies at the heart of optimization. The necessary and sufficient optimality conditions allow one to devise efficient numerical methods for practical solutions of a given optimization problem.

Lagrange multipliers have long been used in optimality conditions for problems with constraints, but recently their role has come to be understood from many different angles. The theory of Lagrange multipliers has been one of the major research areas in nonlinear optimization and there have been a variety of different approaches. Lagrange multipliers were originally introduced for problems with equality constraints. Inequality constraint problems were addressed considerably later. The best known optimality criteria for nonlinear programming problems have been given by John [117] and Kuhn and Tucker [140].

Theorems of the alternative play a very important role in establishing the existence of the Lagrangian multipliers in necessary optimality conditions in constrained optimization. Theorems of alternative are direct applications of separation theorems for convex sets. Loosely speaking, theorems of alternative state that if there are two systems of equalities and inequalities in terms of linear or convex functions, both the systems cannot be solved simultaneously. Some of the important theorems of alternative for systems involving convex functions are the Gordan-type theorems of alternative, and Motzkin-type theorems of alternative. For more theorems one may refer to Mangasarian [153].

The following theorem of alternative, which is a particular case of Corollary 4.2.2 in Mangasarian [153], plays a crucial role in the development of necessary optimality conditions.

Theorem 7.11. [153] Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f_1, f_2 : K \to \mathbb{R}$ be convex functions. If $f_1(x) < 0$, $f_2(x) \leq 0$ has no solution in K, then there exist $s_1, s_2 \in \mathbb{R}$ such that $s_1, s_2 \geq 0$, $(s_1, s_2) \neq (0, 0)$ and $s_1 f_1(x) + s_2 f_2(x) \geq 0$, for all $x \in K$.

Consider the following optimization problem:

$$\text{(MP)} \qquad \min f(x) \quad \text{subject to } g_j(x) \leq 0, \quad j = 1, 2, \ldots, m,$$

where $f, g_j : K \to \mathbb{R}$ are functions defined on a nonempty convex subset of $\mathbb{R}^n$.

We first obtain necessary optimality conditions for this problem in terms of associated bifunctions assuming that the constraints, which are h_j-pseudolinear, satisfy Slater's constraint qualification. Using these optimality conditions, we then derive characterizations for the solution set of (MP) when the objective function is h-pseudolinear.

Let $\tilde{S}$ denote the solution set of the problem (MP). Let $I = \{1, 2, \ldots, m\}$ and for a feasible point $x \in K$, let $A(x)$ denote the set of active indices at x.

Theorem 7.12. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f, g_j : K \to \mathbb{R}$, $j \in I$, be functions, and $h, h_j : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$, $j \in I$, be sublinear in the second argument such that h satisfies condition (7.1) and

$$(g_j)_D(x; d) \leq h_j(x; d) \leq (g_j)^D(x; d), \quad \text{for all } j \in I. \tag{7.2}$$

Further, suppose that g_j is h_j-pseudolinear as well as radially continuous and Slater's constraint qualification holds, that is, there exists $\hat{x} \in K$ such that $g_j(\hat{x}) < 0$ for all $j = 1, 2, \ldots, m$. If $\bar{x} \in K$, then there exists $\bar{u} \in \mathbb{R}^m_+$ such that

$$\langle \bar{u}, g(\bar{x}) \rangle = 0 \tag{7.3}$$

and

$$h(\bar{x}; x - \bar{x}) + \sum_{j=1}^{m} \bar{u}_j h_j(\bar{x}; x - \bar{x}) \geq 0, \quad \text{for all } x \in K. \tag{7.4}$$

Proof. We assert that the system

$$\begin{aligned} h(\bar{x}; x - \bar{x}) &< 0, \\ h_j(\bar{x}; x - \bar{x}) &\leq 0, \quad j \in A(\bar{x}), \end{aligned}$$

has no solution $x \in K$. Suppose on the contrary $x \in K$ satisfies the above inequalities. Then,

$$f_D(\bar{x}; x - \bar{x}) < 0,$$

and

$$h_j(\bar{x}; x - \bar{x}) \leq 0, \quad j \in A(\bar{x}).$$

By positive homogeneity of h_j, there exists a sequence $t_m \to 0^+$ and a positive integer k such that for all $m \geq k$

$$f(\bar{x} + t_m(x - \bar{x})) < f(\bar{x}),$$

$$h_j(\bar{x}; \bar{x} + t_m(x - \bar{x}) - \bar{x}) \leq 0, \quad \text{for all } j \in A(\bar{x}),$$

where the last inequality follows from the positive homogeneity of h_j in the second argument. To arrive at a contradiction it is enough to prove that $\bar{x} + t_p(x - \bar{x})$ is feasible for some integer $p \geq k$. Since g_j is h_j-pseudoconcave, it follows that

$$g_j(\bar{x} + t_m(x - \bar{x})) \leq g_j(\bar{x}) = 0, \quad \text{for all } m \geq k \text{ and } j \in A(\bar{x}).$$

Since $g_j(\bar{x}) < 0$ for $j \notin A(\bar{x})$ and g_j is radially continuous, there exists a positive integer $p \geq k$ such that

$$g_j(\bar{x} + t_p(x - \bar{x})) < 0, \quad \text{for all } j \notin A(\bar{x}).$$

Hence $\bar{x}+t_p(x-\bar{x})$ is a feasible solution of (MP). Since $h(\bar{x}; x-\bar{x})$, $h_j(\bar{x}; x-\bar{x})$, $j \in I$ are sublinear functions of x and hence convex, therefore by Theorem 7.11, there exist nonnegative scalars $\bar{v}$ and $\bar{u}_j$, $j \in A(\bar{x})$ are not all zero such that

$$\bar{v}h(\bar{x}; x - \bar{x}) + \sum_{j \in A(\bar{x})} \bar{u}_j h_j(\bar{x}; x - \bar{x}) \geq 0, \quad \text{for all } x \in K.$$

If $\bar{v} = 0$, then we get

$$\sum_{j \in A(\bar{x})} \bar{u}_j h_j(\bar{x}; x - \bar{x}) \geq 0, \quad \text{for all } x \in K.$$

Since Slater's constraint qualification holds there exists some $\hat{x} \in K$ such that $g_j(\hat{x}) < 0$ for all $j \in A(\bar{x})$, which by h_j-pseudoconvexity of g_j gives

$$h_j(\bar{x}; \hat{x} - \bar{x}) < 0, \quad \text{for all } j \in A(\bar{x}).$$

As $\bar{u}_j$, $j \in A(\bar{x})$ are not all zero, we have

$$\sum_{j \in A(\bar{x})} \bar{u}_j h_j(\bar{x}; \hat{x} - \bar{x}) < 0,$$

which is a contradiction. Therefore, $\bar{v} \neq 0$, and hence the result follows by defining $\bar{u}_j = 0$ for $j \notin A(\bar{x})$. □

Remark 7.3. If $K = \mathbb{R}^n$, f is a differentiable function, g_js are differentiable pseudolinear functions, and the Slater's constraint qualification holds, then the optimality conditions in Theorem 7.12 reduce to the Kuhn-Tucker optimality conditions (1.39) to (1.41) given in Chapter 1.

The next theorem provides a sufficient optimality condition for the optimization problem (MP).

Theorem 7.13. Let K be a nonempty convex subset of $\mathbb{R}^n$, $f, g_j : K \to \mathbb{R}$, $j \in I$ be functions, and $h, h_j : K \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$, $j \in I$ be positively homogeneous in the second argument satisfying (7.1) and (7.2), respectively. Further, suppose that f is h-pseudoconvex, g_j is h_j-pseudolinear, and h_j, $j \in I$ is subodd in the second argument. If $\bar{x} \in K$ and there exists $\bar{u} \in \mathbb{R}^m_+$ such that (7.3) and (7.4) hold, then $\bar{x}$ is an optimal solution of (MP).

Proof. Suppose $\bar{x}$ is not an optimal solution of (MP). Then, there exists a feasible point $x \in K$ such that $f(x) < f(\bar{x})$. Since f is h-pseudoconvex, it follows that

$$h(\bar{x}; x - \bar{x}) < 0. \tag{7.5}$$

Since x is feasible, for each $j \in A(\bar{x})$, we have

$$g_j(x) \le 0 = g_j(\bar{x}).$$

By Theorem 3.9, it follows that g_j is h_j-quasiconvex, and hence from the above relation we have

$$h_j(\bar{x}; x - \bar{x}) \le 0, \quad \text{for all } j \in A(\bar{x}).$$

Since $\bar{u}_j \ge 0$ for all $j \in A(\bar{x})$, we have

$$\sum_{j \in A(\bar{x})} \bar{u}_j h_j(\bar{x}; x - \bar{x}) \le 0.$$

Since $\langle \bar{u}, g(\bar{x}) \rangle = 0$ and $g_j(\bar{x}) < 0$ for all $j \notin A(\bar{x})$, therefore it follows that $\bar{u}_j = 0$ for all $j \notin A(\bar{x})$, and thus the above relation can also be written as

$$\sum_{j=1}^{m} \bar{u}_j h_j(\bar{x}; x - \bar{x}) \le 0,$$

which along with optimality condition (7.4) yields that

$$h(\bar{x}; x - \bar{x}) \ge 0,$$

a contradiction to (7.5). Thus, we conclude that $\bar{x}$ is an optimal solution of (MP). □

For $\bar{x} \in \tilde{S}$, let

$$M(\bar{x}) = \Big\{ u \in \mathbb{R}^m_+ : h(\bar{x}; x - \bar{x}) + \sum_{j=1}^{m} u_j h_j(\bar{x}; x - \bar{x}) \ge 0,$$

$$\text{for all } x \in K,\ \langle u, g(\bar{x})\rangle = 0\Big\}.$$

For any $u \in M(\bar{x})$, let

$$B_u(\bar{x}) = \{j \in A(\bar{x}) : \bar{u}_j \neq 0\},$$

and $C_u(\bar{x}) = I \backslash B_u(\bar{x})$.

Theorem 7.14. Let $\bar{x} \in \tilde{S}$. Suppose that all the conditions of Theorem 7.12 are satisfied, f is h-pseudolinear, and $u \in M(\bar{x})$, then

$$\tilde{S} = \Big\{x \in K : h(\bar{x}; x - \bar{x}) \leq 0,\ h_j(\bar{x}; x - \bar{x}) = 0 \text{ for all } j \in B_u(\bar{x}),$$
$$g_j(x) \leq 0 \text{ for all } j \in C_u(\bar{x})\Big\}.$$

Proof. Let

$$\tilde{S}^* = \Big\{x \in K : h(\bar{x}; x - \bar{x}) \leq 0,\ h_j(\bar{x}; x - \bar{x}) = 0 \text{ for all } j \in B_u(\bar{x}),$$
$$g_j(x) \leq 0 \text{ for all } j \in C_u(\bar{x})\Big\}.$$

Assume that $x \in \tilde{S}$. By Theorem 3.9, f is h-quasiconvex, which together with $f(x) = f(\bar{x})$ implies that

$$h(\bar{x}; x - \bar{x}) \leq 0.$$

Since $u \in M(\bar{x})$, we have

$$h(\bar{x}; x - \bar{x}) + \sum_{j=1}^{m} u_j h_j(\bar{x}; x - \bar{x}) \geq 0,$$

and hence we get

$$\sum_{j=1}^{m} u_j h_j(\bar{x}; x - \bar{x}) \geq 0.$$

As $\langle u, g(\bar{x})\rangle = 0$ and $g_j(\bar{x}) < 0$ for all $j \notin A(\bar{x})$, therefore $u_j = 0$ for all $j \notin A(\bar{x})$, and thus the above inequality becomes

$$\sum_{j \in A(\bar{x})} u_j h_j(\bar{x}; x - \bar{x}) \geq 0,$$

which is equivalent to

$$\sum_{j \in B_u(\bar{x})} u_j h_j(\bar{x}; x - \bar{x}) \geq 0. \tag{7.6}$$

Since x is feasible, we have

$$g_j(x) \leq 0 = g_j(\bar{x}), \quad \text{for all } j \in B_u(\bar{x}).$$

Employing Theorem 3.9 it follows that g_j is h_j-quasiconvex, and thus we have

$$h_j(\bar{x}; x - \bar{x}) \leq 0, \quad \text{for all } j \in B_u(\bar{x}). \tag{7.7}$$

This implies that

$$\sum_{j\in B_u(\bar{x})} u_j h_j(\bar{x}; x-\bar{x}) \leq 0,$$

which in conjunction with (7.6) gives

$$\sum_{j\in B_u(\bar{x})} u_j h_j(\bar{x}; x-\bar{x}) = 0.$$

From (7.7) and the above relation we gather that

$$h_j(\bar{x}; x-\bar{x}) = 0, \quad \text{for all } j \in B_u(\bar{x}).$$

Hence, $x \in \tilde{S}^*$.

Conversely, let $x \in \tilde{S}^*$. Then

$$h_j(\bar{x}; x-\bar{x}) = 0, \quad \text{for all } j \in B_u(\bar{x}), \tag{7.8}$$

and by h_j-pseudoconcavity of g_j, it follows that

$$g_j(x) \leq g_j(\bar{x}) = 0, \quad \text{for all } j \in B_u(\bar{x}),$$

which along with the hypothesis $g_j(x) \leq 0$ for all $j \in C_u(\bar{x})$ proves that x is feasible. Since $\bar{x}$ is an optimal solution, by (7.8) and the optimality condition (7.4), we have

$$h(\bar{x}; x-\bar{x}) \geq 0.$$

As $x \in \tilde{S}^*$, we have $h(\bar{x}; x-\bar{x}) \leq 0$, and hence $h(\bar{x}; x-\bar{x}) = 0$. Using h-pseudoconvexity and h-pseudoconcavity of f, it follows that $f(x) = f(\bar{x})$. Thus, $x \in \tilde{S}$. □

Remark 7.4. In the problem (MP), let $K = \mathbb{R}^n$ and C denote the feasible set, that is,

$$C = \{x \in \mathbb{R}^n : g_j(x) \leq 0, \text{ for all } j \in I\}.$$

Since g_j is h_j-pseudolinear, the conditions $h_j(\bar{x}; x-\bar{x}) = 0$ for all $j \in B_u(\bar{x})$ and $g_j(x) \leq 0$ for all $j \in C_u(\bar{x})$ together imply that $g_j(x) \leq 0$ for all $j \in I$. Therefore, the characterization given in Theorem 7.14 reduces to

$$\tilde{S} = \{x \in C : h(\bar{x}; x-\bar{x}) \leq 0\}.$$

The above characterization was also obtained in Theorem 7.4, though under a different set of assumptions.

Remark 7.5. If the function f in the problem (MP) is a differentiable pseudolinear function and there are no explicit constraints g_j in the problem (MP), then for $\bar{x} \in \tilde{S}$ the solution set of the problem is given as

$$\tilde{S} = \{x \in K : \langle \nabla f(\bar{x}), x-\bar{x}\rangle \leq 0\}.$$

In the problem (MP) when $K = \mathbb{R}^n$, f is a differentiable pseudolinear function, and the constraints are affine, then we obtain the following characterization as given by Dinh et al. [67].

Corollary 7.4. For the problem (MP), let $K = \mathbb{R}^n$, f be a differentiable pseudolinear function, and the constraints $g_j(x) = \langle a_j, x\rangle - b_j$ for all $j \in I$, where $a_j \in \mathbb{R}^n$ and $b_j \in \mathbb{R}$. Further, suppose that $\bar{x} \in \tilde{S}$, $u \in M(\bar{x})$ and Slater's constraint qualification holds. Then

$$\tilde{S} = \big\{x \in \mathbb{R}^n : \langle a_j, x\rangle - b_j = 0, \text{ for all } j \in B_u(\bar{x}), \ \langle a_j, x\rangle - b_j \leq 0, \\ \text{for all } j \in C_u(\bar{x})\big\}.$$

Proof. Observe that when f is differentiable and $g_j(x) = \langle a_j, x\rangle - b_j$ for all $j \in I$, we have

$$h(\bar{x}; x - \bar{x}) = \langle \nabla f(\bar{x}), x - \bar{x}\rangle$$

and

$$h_j(\bar{x}; x - \bar{x}) = \langle a_j, x\rangle - b_j, \quad \text{for all } j \in B_u(\bar{x}).$$

Using Remark 7.3, it is clear that if $\bar{x} \in \tilde{S}$ and $u \in M(\bar{x})$, then $\langle a_j, x\rangle - b_j = 0$ for all $j \in B_u(\bar{x})$ implies that

$$\langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0.$$

Hence the result follows from Theorem 7.14. □

We illustrate with an example the computation of the solution set $\tilde{S}$ using the characterization given in Theorem 7.14.

Example 7.3. Consider the problem

$$\text{(P1)} \qquad \min f(x) \quad \text{subject to } g_j(x) \leq 0, \quad j = 1, 2,$$

where $f, g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$ are functions given by

$$f(x) = \begin{cases} x_1 + x_2, & \text{if } x_1 + x_2 > 0, \\ x_1 + x_2 - 1, & \text{if } x_1 + x_2 \leq 0, \end{cases}$$

$g_1(x) = -x_1 - x_2$, $g_2(x) = -x_1$, where $x = (x_1, x_2)$.

For $d = (d_1, d_2)$, define

$$h(x; d) = \begin{cases} d_1 + d_2, & \text{if } x_1 + x_2 \neq 0, \text{ for all } d \text{ or } x_1 + x_2 = 0, \ d_1 + d_2 \leq 0, \\ +\infty, & \text{if } x_1 + x_2 = 0, \ d_1 + d_2 > 0, \end{cases}$$

$$h_1(x; d) = -d_1 - d_2,$$

and

$$h_2(x; d) = -d_1.$$

We note that the functions h, h_1, and h_2 are sublinear functions of d, for each fixed $x \in \mathbb{R}^2$. The set of feasible solutions is

$$\left\{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 \geq 0,\ x_1 \geq 0\right\}.$$

Now $f(x_1, x_2) \geq -1$, for every feasible point (x_1, x_2), and thus $\bar{x} = (0,\, 0)$ is an optimal solution of the problem (P1). Also, $A(\bar{x}) = \{1, 2\}$ and there exists $\hat{x} = (1, 1) \in \tilde{S}$ such that $g_j(\hat{x}) < 0$ for all $j \in A(\bar{x})$. Moreover, $M(\bar{x}) = \{(1, 0)\}$ and $B_u(\bar{x}) = \{1\}$ for $u = (1, 0)$. Using the above characterization the set of optimal solutions for the problem (P1) is given by

$$\begin{aligned}
\tilde{S} &= \left\{(x_1, x_2) \in \mathbb{R}^2 : h(\bar{x}; x - \bar{x}) \leq 0,\ h_1(\bar{x}; x - \bar{x}) = 0,\ g_2(x) \leq 0\right\} \\
&= \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 \leq 0,\ \ - x_1 - x_2 = 0,\ \ - x_1 \leq 0\right\} \\
&= \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 = 0,\ x_1 \geq 0\right\}.
\end{aligned}$$

Mangasarian [154] proved that for a differentiable convex program the solution set can be characterized completely by a constant gradient of the objective function when the objective function is twice differentiable on some open convex set and by the subdifferential when the objective function is continuous and the relative interior of the solution set is nonempty. We now derive a characterization for the solution set $\tilde{S}$ of (MP) when f is a h-convex function and each of the constraint g_j is an h_j-pseudolinear function.

For $\bar{x} \in \tilde{S}$ and $u \in M(\bar{x})$, we denote the cardinality of the set $B_u(\bar{x})$ by k_u. However, if $B_u(\bar{x})$ is an empty set then k_u takes the value zero.

We first establish a property of the solution set that the active constraints with positive Lagrange multipliers at a solution remain active at all the solutions of (MP) and that the Lagrangian function

$$L(x, u) = f(x) + \langle u, g(x) \rangle$$

as a function of x remains constant on the solution set $\tilde{S}$, for a fixed Lagrange multiplier u.

Theorem 7.15. For the problem (MP), assume that $\bar{x} \in \tilde{S}$ and all the conditions of Theorem 7.12 are satisfied. Further if f is h-convex, then for a fixed $\bar{u} \in M(\bar{x})$ and each $x \in \tilde{S}$, $\langle \bar{u}, g(x) \rangle = 0$ and the Lagrangian function $L(\cdot, \bar{u})$ is constant on $\tilde{S}$.

Proof. Since $f(x) = f(\bar{x})$ for each $x \in \tilde{S}$ and f is h-convex, we get

$$h(\bar{x}; x - \bar{x}) \leq 0,$$

which together with optimality conditions $\langle \bar{u}, g(\bar{x}) \rangle = 0$ and (7.4) implies that

$$\sum_{j \in A(\bar{x})} \bar{u}_j h_j(\bar{x}; x - \bar{x}) \geq 0. \tag{7.9}$$

For each $x \in \tilde{S}$ we have

$$g_j(x) \leq 0 = g_j(\bar{x}), \quad \text{for all } j \in A(\bar{x}).$$

Since g_j is h_j-pseudoconvex, it follows from Theorem 3.9 that g_j is h_j-quasiconvex. Hence,

$$h_j(\bar{x}; x - \bar{x}) \leq 0, \quad \text{for all } j \in A(\bar{x}).$$

This in conjunction with (7.9) implies

$$\bar{u}_j h_j(\bar{x}; x - \bar{x}) = 0, \quad \text{for all } j \in A(\bar{x}).$$

Since $\bar{u}_j \neq 0$ for $j \in B_{\bar{u}}(\bar{x})$, we obtain $h_j(\bar{x}; x - \bar{x}) = 0$ for each $j \in B_{\bar{u}}(\bar{x})$.

By h_j-pseudoconvexity and h_j-pseudoconcavity of g_j, it follows that

$$g_j(x) = g_j(\bar{x}) = 0, \quad \text{for all } j \in B_{\bar{u}}(\bar{x}).$$

Moreover, as $\bar{u}_j = 0$ for $j \in C_u(\bar{x})$, we obtain $\langle \bar{u}, g(x) \rangle = 0$ for each $x \in \tilde{S}$. Hence, for each $x \in \tilde{S}$,

$$f(x) + \langle \bar{u}, g(x) \rangle = f(\bar{x}) + \langle \bar{u}, g(\bar{x}) \rangle,$$

that is, $L(\cdot, \bar{u})$ is constant on $\tilde{S}$. $\square$

Theorem 7.16. Let $\bar{x} \in \tilde{S}$ and suppose that all the conditions of Theorem 7.12 are satisfied. Further, suppose that f is h-convex and $u \in M(\bar{x})$, then $\tilde{S} = \tilde{S}_1 = \tilde{S}_2$, where

$$\begin{aligned}
\tilde{S}_1 &= \Big\{ x \in K : h(\bar{x}; x - \bar{x}) = 0,\ |h(x; \bar{x} - x)| + k_u u_j h_j(\bar{x}; x - \bar{x}) \leq 0 \\
&\qquad \text{for all } j \in B_u(\bar{x}), g_j(x) \leq 0,\ \text{for all } j \in C_u(\bar{x}) \Big\} \\
\tilde{S}_2 &= \Big\{ x \in K : h(\bar{x}; x - \bar{x}) \leq 0,\ |h(x; \bar{x} - x)| + k_u u_j h_j(\bar{x}; x - \bar{x}) \leq 0, \\
&\qquad \text{for all } j \in B_u(\bar{x}), g_j(x) \leq 0,\ \text{for all } j \in C_u(\bar{x}) \Big\}.
\end{aligned}$$

Proof. First we show the inclusion $\tilde{S} \subseteq \tilde{S}_1$. Let $x \in \tilde{S}$, then $f(x) = f(\bar{x})$ and h-convexity of f imply

$$h(\bar{x}; x - \bar{x}) \leq 0. \tag{7.10}$$

Since x is feasible we have $g_j(x) \leq 0 = g_j(\bar{x})$ for all $j \in B_u(\bar{x})$. By Theorem 3.9, g_j is h_j-quasiconvex, and thus we have $h_j(\bar{x}; x - \bar{x}) \leq 0$ for all $j \in B_u(\bar{x})$, which further implies

$$k_u u_j h_j(\bar{x}; x - \bar{x}) \leq 0, \quad \text{for all } j \in B_u(\bar{x}). \tag{7.11}$$

As $\bar{x}$ is an optimal solution of (MP) and $u \in M(\bar{x})$, therefore the optimality condition (7.4) reduces to

$$h(\bar{x}; x - \bar{x}) + \sum_{j \in A(\bar{x})} u_j h_j(\bar{x}; x - \bar{x}) \geq 0,$$

which together with (7.11) implies

$$h(\bar{x}; x - \bar{x}) \geq 0.$$

This again in conjunction with (7.10) gives

$$h(\bar{x}; x - \bar{x}) = 0. \tag{7.12}$$

By interchanging the roles of x and $\bar{x}$ and using optimality conditions at x, it can similarly be established that

$$h(x; \bar{x} - x) = 0. \tag{7.13}$$

Hence from (7.11) to (7.13) and feasibility of x, it follows that $x \in \tilde{S}_1$.

Clearly, $\tilde{S}_1 \subseteq \tilde{S}_2$. Next, we show that $\tilde{S}_2 \subseteq \tilde{S}$. Assume that $x \in \tilde{S}_2$. Then,

$$k_u u_j h_j(\bar{x}; x - \bar{x}) \leq 0, \quad \text{for all } j \in B_u(\bar{x}),$$

that is,

$$h_j(\bar{x}; x - \bar{x}) \leq 0, \quad \text{for all } j \in B_u(\bar{x}).$$

By h_j-pseudoconcavity of g_j, it follows that

$$g_j(x) \leq g_j(\bar{x}) = 0, \quad \text{for all } j \in B_u(\bar{x}),$$

which along with the hypothesis $g_j(x) \leq 0$ for all $j \in C_u(\bar{x})$ proves that x is feasible. Since $u \in M(\bar{x})$, we have

$$h(\bar{x}; x - \bar{x}) + \sum_{j \in A(\bar{x})} u_j h_j(\bar{x}; x - \bar{x}) \geq 0.$$

Using the hypothesis $h(\bar{x}; x - \bar{x}) \leq 0$, it is obvious that

$$\sum_{j \in A(\bar{x})} u_j h_j(\bar{x}; x - \bar{x}) \geq 0,$$

or equivalently

$$\sum_{j \in B_u(\bar{x})} u_j h_j(\bar{x}; x - \bar{x}) \geq 0.$$

But as

$$|h(x; \bar{x} - x)| + k_u u_j h_j(\bar{x}; x - \bar{x}) \leq 0, \quad \text{for all } j \in B_u(\bar{x}),$$

we have

$$h(x; \bar{x} - x) \geq \sum_{j \in B_u(\bar{x})} u_j h_j(\bar{x}; x - \bar{x}),$$

which implies that $h(x; \bar{x} - x) \geq 0$. By h-convexity of f, it follows that $f(x) \leq f(\bar{x})$. Since $\bar{x} \in \tilde{S}$, we obtain $f(x) = f(\bar{x})$, that is, $x \in \tilde{S}$. $\square$

The following corollary is an immediate consequence of Theorem 7.16.

Corollary 7.5. For the problem (MP), suppose that f is a differentiable convex function, the constraints are linear, that is, $g_j(x) = \langle a_j, x\rangle - b_j$ for all j and Slater's constraint qualification holds. Let $\bar{x} \in \tilde{S}$ and $u \in M(\bar{x})$, then

$$\tilde{S} = \{x \in K : \langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0,\ \ |\langle \nabla f(x), \bar{x} - x\rangle| + k_u u_j(\langle a_j, x\rangle - b_j) \le 0,$$
$$\text{for all } j \in B_u(\bar{x}),\ \langle a_j, x\rangle \le b_j,\ \text{for all } j \in C_u(\bar{x})\}.$$

We now consider the case $K = \mathbb{R}^n$ and establish that the gradient of the objective function is constant on the solution set.

Corollary 7.6. For the problem (MP), assume that $K = \mathbb{R}^n$, f is a differentiable convex function, the constraints are linear, that is, $g_j(x) = \langle a_j, x\rangle - b_j$ for all j, and Slater's constraint qualification holds. If $\bar{x} \in \tilde{S}$ and $u \in M(\bar{x})$, then

$$\tilde{S} = \{x \in \mathbb{R}^n : \langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0,\ \langle a_j, x\rangle - b_j = 0,\ \text{for all } j \in B_u(\bar{x}),$$
$$\langle a_j, x\rangle \le b_j,\ \text{for all } j \in C_u(\bar{x}),\ \nabla f(\bar{x}) = \nabla f(x)\}.$$

Proof. Let $x \in \tilde{S}$, then from Corollary 7.5 we have that x satisfies

$$\langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0. \tag{7.14}$$

Since f is convex, we have for each $y \in \mathbb{R}^n$

$$\begin{aligned} f(y) - f(x) &= f(y) - f(\bar{x}) \ge \langle \nabla f(\bar{x}), y - \bar{x}\rangle \\ &= \langle \nabla f(\bar{x}), y - x\rangle + \langle \nabla f(\bar{x}), x - \bar{x}\rangle \\ &= \langle \nabla f(\bar{x}), y - x\rangle, \end{aligned}$$

which implies that $\nabla f(\bar{x})$ is a subgradient of f at x. Since$f$ is differentiable the only subgradient of f at x is $\nabla f(x)$, hence it follows that

$$\nabla f(\bar{x}) = \nabla f(x). \tag{7.15}$$

The relations (7.14) and (7.15) together imply that $\langle \nabla f(x), \bar{x} - x\rangle = 0$. From Remark 7.3, it follows that

$$\begin{aligned} \langle \nabla f(x), x - \bar{x}\rangle &= -\sum_{j \in A(\bar{x})} u_j \langle a_j, x - \bar{x}\rangle \\ &= -\sum_{j \in A(\bar{x})} u_j(\langle a_j, x\rangle - b_j) \\ &= -\sum_{j \in B_u(\bar{x})} u_j(\langle a_j, x\rangle - b_j). \end{aligned} \tag{7.16}$$

Hence, we have

$$\sum_{j \in B_u(\bar{x})} u_j(\langle a_j, x\rangle - b_j) = 0.$$

Since $u_j > 0$ and $\langle a_j, x\rangle - b_j \leq 0$ for all $j \in B_u(\bar{x})$, it follows that $\langle a_j, x\rangle - b_j = 0$ for all $j \in B_u(\bar{x})$. Thus, we have

$$\begin{aligned}\tilde{S} &\subseteq \Big\{x \in \mathbb{R}^n : \langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0,\ \langle a_j, x\rangle - b_j = 0 \text{ for all } j \in B_u(\bar{x}),\\ &\qquad \langle a_j, x\rangle \leq b_j,\ \text{for all } j \in C_u(\bar{x}),\ \nabla f(\bar{x}) = \nabla f(x)\Big\}\\ &\subseteq \Big\{x \in \mathbb{R}^n : \langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0,\ |\langle \nabla f(x), \bar{x} - x\rangle| + k_u u_j(\langle a_j, x\rangle - b_j) \leq 0\\ &\qquad \text{for all } j \in B_u(\bar{x}),\ \langle a_j, x\rangle \leq b_j,\ \text{for all } j \in C_u(\bar{x})\Big\}\\ &= \tilde{S}.\end{aligned}$$

This completes the proof. □

Remark 7.6. In the solution set given in Corollary 7.6, the conditions $\langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0$ and $\langle a_j, x\rangle - b_j = 0$ for all $j \in B_u(\bar{x})$ are equivalent. This observation is in view of condition (7.16). Therefore, the solution set $\tilde{S}$ is actually given by

$$\tilde{S} = \Big\{x \in \mathbb{R}^n : \langle a_j, x\rangle - b_j = 0 \text{ for all } j \in B_u(\bar{x}),\ \langle a_j, x\rangle \leq b_j,\\ \text{for all } j \in C_u(\bar{x}),\ \nabla f(\bar{x}) = \nabla f(x)\Big\}.$$

Consider the quadratic program

$$\text{(QP)} \qquad \min f(x) = \frac{1}{2}\langle x, Qx\rangle + \langle d, x\rangle, \quad \text{subject to } \langle a_j, x\rangle \leq b_j,\ j \in I,$$

where Q is an $n \times n$ symmetric positive semidefinite matrix, d and a_j, $j \in I$ are vectors in $\mathbb{R}^n$, and $b_j \in \mathbb{R}$.

Here we assume that the quadratic function f is defined on $\mathbb{R}^n$. Then in view of Remark 7.6, we can immediately deduce Corollary 2.3 in Jeyakumar et al. [112], which is as follows.

Corollary 7.7. Let $\tilde{K}$ be the solution set of (QP), $\bar{x} \in \tilde{K}$, and $u \in M(\bar{x})$. Further, suppose that Slater's constraint qualification holds. Then,

$$\tilde{K} = \Big\{x \in \mathbb{R}^n : \langle a_j, x\rangle - b_j = 0,\ \text{for all } j \in B_u(\bar{x}),\ \langle a_j, x\rangle \leq b_j,\\ \text{for all } j \in C_u(\bar{x}),\ Q\bar{x} = Qx\Big\}.$$

We now deduce a characterization for the problem (MP) in the absence of the constraints, that is, for the problem of minimizing f over the convex domain K. These corollaries follow immediately from Theorem 7.16.

Corollary 7.8. Consider the problem (MP) with $I = \phi$. If f is h-convex, $h(x; d)$ is a sublinear function of d, $\bar{x} \in \tilde{S}$, then

$$\begin{aligned}\tilde{S} &= \{x \in K : h(\bar{x}; x - \bar{x}) = 0,\ h(x; \bar{x} - x) = 0\}\\ &= \{x \in K : h(\bar{x}; x - \bar{x}) \leq 0,\ h(x; \bar{x} - x) = 0\}.\end{aligned}$$

Corollary 7.9. Consider the problem (MP) with $I = \phi$. If f is a differentiable convex function, $\bar{x} \in \tilde{S}$, then

$$\begin{aligned}\tilde{S} &= \{x \in K : \langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0,\ \langle \nabla f(x), \bar{x} - x\rangle = 0\} \\ &= \{x \in K : \langle \nabla f(\bar{x}), x - \bar{x}\rangle \leq 0,\ \langle \nabla f(x), \bar{x} - x\rangle = 0\}.\end{aligned}$$

Remark 7.7. If f is a twice continuously differentiable convex function and K is an open convex set then it can be seen as in the proof of Lemma 1 of Mangasarian [154] that $\nabla f(\bar{x}) = \nabla f(x)$. Thus under these conditions the solution set $\tilde{S}$ is given by

$$\begin{aligned}\tilde{S} &= \{x \in K : \langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0,\ \nabla f(\bar{x}) = \nabla f(x)\} \\ &= \{x \in K : \langle \nabla f(\bar{x}), x - \bar{x}\rangle \leq 0,\ \nabla f(\bar{x}) = \nabla f(x)\}.\end{aligned}$$

This is precisely Theorem 1 of Mangasarian [154].

We give examples to elucidate the characterization given in Theorem 7.16.

Example 7.4. Consider the problem

$$\text{(P2)} \qquad \min f(x), \quad \text{subject to } g_j(x) \leq 0,\ j = 1, 2,$$

where $f, g_1, g_2 : \mathbb{R} \to \mathbb{R}$ are functions given by

$$\begin{aligned} f(x) &= \begin{cases} -x, & \text{if } x < 0, \\ 0, & \text{if } 0 \leq x \leq 2, \\ x - 2, & \text{if } x > 2, \end{cases} \\ g_1(x) &= -x^3 - x, \\ g_2(x) &= \begin{cases} x - 3, & \text{if } x \geq 3, \\ \frac{x-3}{2}, & \text{if } x < 3. \end{cases}\end{aligned}$$

Define,

$$h(x; d) = \begin{cases} -d, & \text{if } x = 0,\ d \leq 0 \text{ or } x < 0, \\ 0, & \text{if } x = 0,\ d > 0 \text{ or } x = 2,\ d \leq 0 \text{ or } 0 < x < 2, \\ d, & \text{if } x = 2,\ d > 0 \text{ or } x > 2, \end{cases}$$

$$h_1(x; d) = (-3x^2 - 1)d,$$

and

$$h_2(x; d) = \begin{cases} d, & \text{if } x = 3,\ d > 0 \text{ or } x > 3, \\ \frac{d}{2} & \text{if } x = 3,\ d \leq 0 \text{ or } x < 3. \end{cases}$$

The set of feasible solutions of the problem (P2) is $\{x \in \mathbb{R} : 0 \leq x \leq 3\}$ and $\bar{x} = 0$ is an optimal solution of the problem (P2).

For $\bar{x} = 0$, the set $A(\bar{x}) = \{1\}$ and $M(\bar{x}) = \{(0, 0)\}$. Clearly $k_u = 0$ for $u = (0, 0) \in M(\bar{x})$ as the set $B_u(\bar{x})$ is an empty set. Also, Slater's constraint qualification is satisfied as

$$g_j(\hat{x}) < 0, \quad \text{for all } j \in A(\bar{x}),\ \text{where } \hat{x} = 2.$$

The solution set of (P2) is given as

$$\begin{aligned}\tilde{S} &= \{x \in \mathbb{R} : h(0;x) \leq 0,\ h(x;-x) = 0,\ g_j(x) \leq 0,\ j = 1,2\} \\ &= \{x \in \mathbb{R} : 0 \leq x \leq 2\}.\end{aligned}$$

We consider the problem given by Jeyakumar et al. [112] and obtain its solution set using the characterization given in Theorem 7.16.

Example 7.5. Consider the following problem

$$\text{(P3)} \qquad \min f(x), \quad \text{subject to } g_j(x) \leq 0,\ j = 1,2,$$

where $f, g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$ are functions given by

$$\begin{aligned} f(x_1,x_2) &= (x_1^2 + x_2^2)^{1/2} + x_1 + x_2, \\ g_1(x_1,x_2) &= x_1 + x_2, \\ g_2(x_1,x_2) &= -x_1. \end{aligned}$$

For $d = (d_1, d_2)$ and $x = (x_1, x_2)$, define

$$\begin{aligned} h(x;d) &= \begin{cases} \frac{x_1 d_1 + x_2 d_2}{(x_1^2+x_2^2)^{1/2}} + d_1 + d_2, & \text{if } x \neq 0, \\ (d_1^2 + d_2^2)^{1/2} + d_1 + d_2, & \text{if } x = 0, \end{cases} \\ h_1(x;d) &= d_1 + d_2, \\ h_2(x;d) &= -d_1. \end{aligned}$$

Then, $f(x_1,x_2) \geq 0$ for all (x_1,x_2) in the feasible set. Clearly, $\bar{x} = (0,0) \in \tilde{S}$ and $A(\bar{x}) = \{1,2\}$. For $u = (0,1) \in M(\bar{x})$ we have $k_u = 1$ and the solution set is given by

$$\begin{aligned} \tilde{S} &= \big\{x = (x_1,x_2) \in \mathbb{R}^2 : h(\bar{x}; x - \bar{x}) \leq 0,\ |h(x; \bar{x} - x)| \\ &\qquad + h_2(\bar{x}; x - \bar{x}) \leq 0,\ g_1(x) \leq 0\big\} \\ &= \big\{(x_1,x_2) \in \mathbb{R}^2 : (x_1^2 + x_2^2)^{1/2} + x_1 + x_2 \leq 0, \\ &\qquad (x_1^2 + x_2^2)^{1/2} + x_1 + x_2 - x_1 \leq 0,\ x_1 + x_2 \leq 0\big\} \\ &= \big\{(x_1,x_2) \in \mathbb{R}^2 : (x_1^2 + x_2^2)^{1/2} + x_2 \leq 0,\ x_1 + x_2 \leq 0\big\}. \end{aligned}$$

The condition $(x_1^2 + x_2^2)^{1/2} + x_2 \leq 0$ is equivalent to $x_2 \leq 0$ and $x_1^2 = 0$. Thus,

$$\tilde{S} = \{(x_1,x_2) \in \mathbb{R}^2 : x_1 = 0,\ x_2 \leq 0\}.$$

We conclude the chapter with an example that gives a better understanding of Corollary 7.6.

Example 7.6. If in the problem (P3) the functions f, g_1, g_2 are given by

$$\begin{aligned} f(x_1,x_2) &= x_1^2, \\ g_1(x_1,x_2) &= 1-x_2, \\ g_2(x_1,x_2) &= 1-x_1, \end{aligned}$$

then for $d=(d_1,d_2)$ and $x=(x_1,x_2)$,

$$\langle \nabla f(x), d\rangle = 2x_1d_1, \quad \langle \nabla g_1(x), d\rangle = -d_2, \text{ and } \langle \nabla g_2(x), d\rangle = -d_1.$$

Clearly, $f(x_1,x_2) \geq 1$ for all (x_1,x_2) in the feasible set. We note that for $\hat{x}=(2,2)\in\mathbb{R}^2$ the Slater's constraint qualification holds. For $\bar{x}=(1,1)\in\tilde{K}$ the set $A(\bar{x})=\{1,2\}$ and $M(\bar{x})=\{(0,2)\}$. Thus, for $u=(0,2)$ we note that $B_u(\bar{x})=\{2\}$, and so $k_u=1$. Using the characterization given in Corollary 7.6 the solution set can be described as

$$\begin{aligned} \tilde{S} &= \big\{x\in\mathbb{R}^2 : \langle \nabla f(\bar{x}), x-\bar{x}\rangle = 0,\ \langle a_2, x\rangle - b_2 = 0,\ \langle a_1, x\rangle \leq b_1, \\ &\quad \nabla f(\bar{x}) = \nabla f(x)\big\} \\ &= \big\{(x_1,x_2)\in\mathbb{R}^2 : 2(x_1-1)=0,\ 1-x_2\leq 0\big\} \\ &= \big\{(x_1,x_2)\in\mathbb{R}^2 : x_1=1,\ x_2\geq 1\big\}. \end{aligned}$$

Chapter 8

Nonsmooth Generalized Variational Inequalities and Optimization Problems

8.1 Introduction

As we have seen in Section 5.2, the variational inequality problem and complementarity problem are equivalent if the the underlying set is a convex cone. Karamardian [122, 123] and Saigal [187] considered the complementarity problem for set-valued maps, known as the *generalized complementarity problem.* If the underlying map in the formulation of the variational inequality problem (VIP) is a set-valued map, then the VIP is called a *generalized variational inequality problem* (GVIP). This problem was considered by Karamardian [122, 123] and Saigal [187] to prove the existence of a solution of a generalized complementarity problem. Basically, they extended Proposition 5.4 for set-valued maps and established an equivalence between a generalized complementarity problem and a generalized variational inequality problem under the condition that the underlying set is a convex cone. We have also seen in Section 5.2 that the solution set of a variational inequality problem coincides with the solution set of a minimization problem if the underlying function is convex and differentiable. However, if the function in the optimization problem is not necessarily differentiable, then the generalized variational inequality problem provides a necessary and sufficient condition for a solution of a minimization problem. A large number of papers on the existence of solutions of generalized variational inequality problems have appeared in the literature [8, 9, 42, 62, 65, 66, 75, 108, 109, 122, 123, 128, 135, 139, 146, 172, 187, 191, 192, 199, 205, 207, 210].

8.2 Generalized Variational Inequalities and Related Topics

Let K be a nonempty subset of $\mathbb{R}^n$, and $F : K \to 2^{\mathbb{R}^n}$ be a set-valued map with nonempty values. The *generalized variational inequality problem* (GVIP) is to find $\bar{x} \in K$ and $\bar{u} \in F(\bar{x})$ such that

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K. \tag{8.1}$$

An element $\bar{x} \in K$ is said to be a *strong solution* of GVIP if there exists $\bar{u} \in F(\bar{x})$ such that the inequality (8.1) holds, and in that case $(\bar{x}, \bar{u})$ is called a solution of GVIP.

The weak form of the GVIP is the problem of finding $\bar{x} \in K$ such that for each $y \in K$ there exists $\bar{u} \in F(\bar{x})$ that satisfies

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0. \tag{8.2}$$

It is called a *weak generalized variational inequality problem* (WGVIP). An element $\bar{x} \in K$ is said to be a *weak solution* of GVIP if for each $y \in K$, there exists $\bar{u} \in F(\bar{x})$ such that the inequality (8.2) holds. It should be noted that $\bar{u}$ in WGVIP depends on y. Of course, if F is a single-valued map, then both the problems mentioned above reduce to the variational inequality problem (5.1).

Clearly, every strong solution of GVIP is a weak solution. However, the converse is not true in general.

Example 8.1. Let $K = \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1 \leq 0, \ -\sqrt{-x_1} \leq x_2 \leq 0\right\}$, $f(x_1, x_2) = \sqrt{x_1^2 + x_2^2} + x_2$, for all $(x_1, x_2) \in K$. If $(x_1, x_2) = (0, 0) = \mathbf{0}$, then

$$\begin{aligned} \partial f\left(x_1, x_2\right) &= \left\{(\zeta_1, \zeta_2) \in \mathbb{R}^2 : \zeta_1^2 + \zeta_2^2 \leq 1\right\} + \{(0, 1)\} \\ &= \left\{(\zeta_1, \zeta_2) \in \mathbb{R}^2 : \zeta_1^2 + (\zeta_2 - 1)^2 \leq 1\right\}. \end{aligned}$$

If $(x_1, x_2) \neq (0, 0)$, then

$$\partial f(x_1, x_2) = \left\{\left(\frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \frac{x_2}{\sqrt{x_1^2 + x_2^2}} + 1\right)\right\}.$$

It can be easily checked that for all $\bar{u} = (\bar{u}_1, \bar{u}_2) \in \partial f(0, 0) = \partial f(\mathbf{0})$, there exists $y = (y_1, y_2) \in K$ such that

$$\langle \bar{u}, y - \mathbf{0} \rangle = \bar{u}_1 y_1 + \bar{u}_2 y_2 < 0.$$

Thus, $\mathbf{0}$ is not a strong solution of GVIP with $F = \partial f$. However, it is a weak solution of GVIP because for all $(y_1, y_2) \in K$ there exists $(\bar{u}_1, \bar{u}_2) \in \partial f(0,0)$ such that

$$\langle \bar{u}, y - \mathbf{0} \rangle = \bar{u}_1 y_1 + \bar{u}_2 y_2 \geq 0.$$

Moreover, the set of strong solutions of GVIP is empty and the set of weak solutions of GVIP is $\{(0,0)\}$.

The following lemma says that every weak solution of GVIP is a strong solution if the set-valued map F is nonempty compact and convex valued.

Lemma 8.1. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $F : K \to 2^{\mathbb{R}^n}$ be a set-valued map such that for each $x \in K$, $F(x)$ is nonempty, compact, and convex. Then, every weak solution of GVIP is a strong solution.

Proof. Let $\bar{x} \in K$ be a weak solution of GVIP. Then, for each $y \in K$ there exists $\bar{u} \in F(\bar{x})$ such that

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0,$$

that is,

$$\sup_{u \in F(\bar{x})} \langle u, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

Since $F(\bar{x})$ is compact and convex, by Theorem B.7, we have

$$\sup_{u \in F(\bar{x})} \inf_{y \in K} \langle u, y - \bar{x} \rangle = \inf_{y \in K} \sup_{u \in F(\bar{x})} \langle u, y - \bar{x} \rangle \geq 0. \tag{8.3}$$

Note that the real-valued function $u \mapsto \inf_{y \in K} \langle u, y - \bar{x} \rangle$ is convex and lower semicontinuous. Since $F(\bar{x})$ is compact, it follows from (8.3) that there exists $\bar{u} \in F(\bar{x})$ such that

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

Consequently, $\bar{x}$ is a strong solution of GVIP. □

If $K = \mathbb{R}^n$, then clearly WGVIP reduces to the following *set-valued inclusion problem*. Find $\bar{x} \in \mathbb{R}^n$ such that

$$\mathbf{0} \in F(\bar{x}). \tag{8.4}$$

We consider the generalized complementarity problem, which is one of the most important problems from operations research. For details on complementarity problems and their generalizations, we refer to other studies and the references therein [18, 70, 108, 109, 166, 211].

Let K be a convex cone in $\mathbb{R}^n$ with its dual cone $K^* = \{u \in \mathbb{R}^n : \langle u, x \rangle \geq 0 \text{ for all } x \in K\}$. The *generalized complementarity problem* (GCP) is to find $\bar{x} \in K$ and $\bar{u} \in F(\bar{x})$ such that

$$\bar{u} \in K^* \quad \text{and} \quad \langle \bar{u}, \bar{x} \rangle = 0. \tag{8.5}$$

Proposition 8.1. $(\bar{x}, \bar{u})$ is a solution of GVIP if and only if it is a solution of GCP.

Proof. Let $(\bar{x}, \bar{u})$ be a solution of GVIP. Then, $\bar{u} \in F(\bar{x})$ and

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K,$$

that is, for all $y \in K$, $\langle \bar{u}, \bar{x} \rangle \leq \langle \bar{u}, y \rangle$. Since $\mathbf{0} \in K$, we have $\langle \bar{u}, \bar{x} \rangle \leq 0$. Also, since $\bar{x} \in K$ and K is a cone, we have $\lambda \bar{x} \in K$ for all $\lambda \geq 0$. Therefore, for all $\lambda \geq 0$ we have $\langle \bar{u}, \bar{x} \rangle \leq \langle \bar{u}, \lambda \bar{x} \rangle$. Hence, we deduce that $(\lambda - 1)\langle \bar{u}, \bar{x} \rangle \geq 0$ and considering a $\lambda > 1$, we obtain $\langle \bar{u}, \bar{x} \rangle \geq 0$. So, finally we have $\langle \bar{u}, \bar{x} \rangle = 0$. This implies that $\langle \bar{u}, y \rangle \geq 0$ for all $y \in K$, that is, $\bar{u} \in K^*$. Hence, $(\bar{x}, \bar{u})$ is a solution of GCP.

Conversely, let $(\bar{x}, \bar{u})$ be a solution of GCP. Then, $\bar{x} \in K$, $\bar{u} \in F(\bar{x}) \cap K^*$ and $\langle \bar{u}, \bar{x} \rangle = 0$. Since $\bar{u} \in K^*$, we have $\langle \bar{u}, y \rangle \geq 0$ for all $y \in K$. This together with the previous equality implies that $\langle \bar{u}, y - \bar{x} \rangle \geq 0$ for all $y \in K$, and hence $(\bar{x}, \bar{u})$ is a solution of GVIP. □

Let K be a nonempty subset of $\mathbb{R}^n$ and $T : K \to 2^K$ be a set-valued map with nonempty values. The *set-valued fixed point problem* (in short, SVFPP) associated with T is to find $\bar{x} \in K$ such that

$$\bar{x} \in T(\bar{x}). \tag{8.6}$$

The point $\bar{x} \in K$ is called a fixed point of T if the relation (8.6) holds. This problem can be converted into a generalized variational inequality formulation as shown next in the set-valued version of Proposition 5.5.

Proposition 8.2. Let K be a nonempty subset of $\mathbb{R}^n$ and $T : K \to 2^K$ be a set-valued map with nonempty values. If the set-valued map $F : K \to 2^{\mathbb{R}^n}$ is defined by

$$F(x) = x - T(x), \tag{8.7}$$

then an element $\bar{x} \in K$ is a strong solution of GVIP (8.1) if and only if it is a fixed point of T.

Proof. Let $\bar{x} \in K$ be a fixed point of T. Then, $\bar{x} \in T(\bar{x})$, and therefore $\mathbf{0} \in F(\bar{x})$. Thus, $\bar{x}$ satisfies the inequality (8.1) with $\bar{u} = \mathbf{0}$, and hence $\bar{x}$ is a strong solution of GVIP.

Conversely, suppose that $\bar{x}$ is a strong solution of GVIP and satisfies the relation (8.7). Then, there exists $\bar{u} \in F(\bar{x})$ such that

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K, \tag{8.8}$$

and $F(\bar{x}) = \bar{x} - T(\bar{x})$. Therefore, there exists $\bar{v} \in T(\bar{x})$ such that $\bar{u} = \bar{x} - \bar{v}$ and

$$\langle \bar{x} - \bar{v}, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

In particular, we can take $y = \bar{v}$, and thus we obtain $\|\bar{x} - \bar{v}\|^2 \leq 0$, that is, $\bar{x} = \bar{v} \in T(\bar{x})$. Hence, $\bar{x}$ is a fixed point of T. □

Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a function. Consider the following minimization problem:

$$\min f(x), \quad \text{subject to } x \in K. \tag{8.9}$$

The following result shows that the GVIP with $F(x) = \partial f(x)$, the subdifferential of a convex function f, is a necessary and sufficient optimality condition for the optimization problem (8.9).

Proposition 8.3. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a convex function. If $\bar{x} \in K$ is a solution of the minimization problem (8.9), then it is a strong solution of GVIP with $F(x) = \partial f(x)$ for all $x \in K$. Conversely, if $(\bar{x}, \bar{u})$ is a solution of GVIP with $\bar{u} \in \partial f(\bar{x})$, then $\bar{x}$ solves the minimization problem (8.9).

Proof. Let $\bar{x} \in K$ be a solution of the minimization problem (8.9). Then by Theorem 1.47 it follows that $\mathbf{0} \in \partial f(\bar{x})$. Hence, $(\bar{x}, \mathbf{0})$ is a solution of GVIP, that is, $\bar{x}$ is a strong solution of GVIP with $F(x) = \partial f(x)$ for all $x \in K$.

Conversely, assume that $(\bar{x}, \bar{u})$ is a solution of GVIP with $F(x) = \partial f(x)$ for all $x \in K$. Then, $\bar{u} \in \partial f(\bar{x})$ and

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K. \tag{8.10}$$

Since $\bar{u} \in \partial f(\bar{x})$, we have

$$\langle \bar{u}, y - \bar{x} \rangle \leq f(y) - f(\bar{x}), \quad \text{for all } y \in \mathbb{R}^n. \tag{8.11}$$

By combining inequalities (8.10) and (8.11), we obtain

$$f(\bar{x}) \leq f(y), \quad \text{for all } y \in K.$$

Hence, $\bar{x}$ is a solution of the minimization problem (8.9). $\square$

It can be easily seen that if $\bar{x}$ is a weak solution of GVIP, even then it is a solution of the minimization problem (8.9).

8.3 Basic Existence and Uniqueness Results

In this section, we present some basic existence and uniqueness results for various types of solutions of GVIP.

Theorem 8.1. [187] Let K be a nonempty compact convex subset of $\mathbb{R}^n$ and $F : K \to 2^{\mathbb{R}^n}$ be an upper semicontinuous set-valued map such that for each $x \in K$, $F(x)$ is nonempty, compact, and convex. Then, there exists a solution $(\bar{x}, \bar{u})$ of GVIP.

Proof. For each $x \in K$ and $u \in \mathbb{R}^n$, define a set-valued map $T : K \times \mathbb{R}^n \to 2^K$ by

$$T(x,u) = \left\{ y \in K : \langle u, y - x\rangle = \min_{z\in K}\langle u, z - x\rangle \right\}.$$

Then, for each $(x,u) \in K \times \mathbb{R}^n$, $T(x,u)$ is convex. Indeed, let $y_1, y_2 \in T(x,u)$ and $\lambda \in [0,1]$. Then,

$$\langle u, y_1 - x\rangle \le \langle u, z - x\rangle, \quad \text{for all } z \in K, \tag{8.12}$$

and

$$\langle u, y_2 - x\rangle \le \langle u, z - x\rangle, \quad \text{for all } z \in K. \tag{8.13}$$

Multiplying inequality (8.12) by λ and inequality (8.13) by $1-\lambda$ and then on adding the resultant inequalities, we obtain

$$\langle u, \lambda y_1 + (1-\lambda)y_2 - x\rangle \le \langle u, z - x\rangle, \quad \text{for all } z \in K,$$

that is, $\lambda y_1 + (1-\lambda)y_2 \in T(x,u)$. Hence, $T(x,u)$ is convex.

It can be easily seen that $T(x,u)$ is a closed subset of the compact set K, and hence compact for each $(x,u) \in K \times \mathbb{R}^n$. It can also be easily seen that T is a closed set-valued map. Since K is compact, by Theorem A.10, T is upper semicontinuous set-valued map with compact and convex values. By Theorem A.7, $F(K) = \bigcup_{x\in K} F(x)$ is a compact subset of $\mathbb{R}^n$. Hence, $D = \mathrm{co}(F(K))$ the convex hull of $F(K)$, is a compact convex set. Consequently, the set-valued map $P : K \times D \to 2^{K\times D}$, defined by $P(x,u) = (T(x,u), F(x))$, is upper semicontinuous with nonempty, compact, and convex values. By the Kakutani fixed point theorem (Theorem B.6), there exists $(\bar{x},\bar{u}) \in K \times D$ such that $(\bar{x},\bar{u}) \in P(\bar{x},\bar{u})$. Hence, $\bar{u} \in F(\bar{x})$ and $\bar{x} \in T(\bar{x},\bar{u})$, that is, for all $y \in K$, $\langle \bar{u}, y - \bar{x}\rangle \ge \langle \bar{u}, \bar{x} - \bar{x}\rangle = 0$. Thus, $(\bar{x},\bar{u})$ is a solution of GVIP. □

If K is not necessarily bounded, then we have the following result.

Theorem 8.2. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to 2^{\mathbb{R}^n}$ be an upper semicontinuous set-valued map such that for each $x \in K$, $F(x)$ is nonempty, compact, and convex. If there exist an element $\tilde{y} \in K$ and a constant $r > \|\tilde{y}\|$ such that

$$\max_{u\in F(x)}\langle u, \tilde{y} - x\rangle \le 0, \tag{8.14}$$

for all $x \in K$ with $\|x\| = r$, then there exists a solution $(\bar{x},\bar{u})$ of GVIP.

Proof. Let $\mathbb{B}_r := \{x \in K : \|x\| \le r\}$. Since $\mathbb{B}_r$ is compact and convex, by Theorem 8.1, there exist $\bar{x} \in \mathbb{B}_r$ and $\bar{u} \in F(\bar{x})$ such that

$$\langle \bar{u}, y - \bar{x}\rangle \ge 0, \quad \text{for all } y \in \mathbb{B}_r. \tag{8.15}$$

We distinguish two cases.

Case I: $\|\bar{x}\| = r$. Since $\tilde{y} \in \mathbb{B}_r$, it follows from the inequalities (8.14) and

(8.15) that $\langle \bar{u}, \tilde{y} - \bar{x} \rangle = 0$. Now, let $x \in K$ and choose $\lambda \in \ ]0,1[$ small enough so that $z = \lambda x + (1-\lambda)\tilde{y} \in \mathbb{B}_r$. Then, we have

$$\begin{aligned} 0 &\leq \langle \bar{u}, z - \bar{x} \rangle \\ &= \lambda \langle \bar{u}, x - \bar{x} \rangle + (1-\lambda)\langle \bar{u}, \tilde{y} - \bar{x} \rangle \\ &= \lambda \langle \bar{u}, x - \bar{x} \rangle. \end{aligned}$$

Consequently, $(\bar{x}, \bar{u})$ is a solution of GVIP.

CASE II: $\|\bar{x}\| < r$. For any given $x \in K$, we can choose $\lambda \in \ ]0,1[$ small enough so that $w = \lambda x + (1-\lambda)\bar{x} \in \mathbb{B}_r$. Then, as in Case I, it can be shown that $(\bar{x}, \bar{u})$ is a solution of GVIP. □

Some existence results for a solution of GVIP under the assumption that the underlying set K is convex but neither bounded nor closed are derived by Fang and Peterson [75].

The following problem is the set-valued version of the Minty variational inequality problem, known as the *generalized Minty variational inequality problem* (in short, GMVIP). Find $\bar{x} \in K$ such that for all $y \in K$ and all $v \in F(y)$, we have

$$\langle v, y - \bar{x} \rangle \geq 0. \tag{8.16}$$

A weak form of the generalized Minty variational inequality problem is the following problem, which is called the *weak generalized Minty variational inequality problem* (WGMVIP). Find $\bar{x} \in K$ such that for all $y \in K$, there exists $v \in F(y)$ satisfying the inequality (8.16).

A solution of WGMVIP is called a *weak solution* of GMVIP. It is clear that every solution of GMVIP is a weak solution of GMVIP where $F(x) = \partial f(x)$ for all $x \in K$.

The following result provides a necessary and sufficient condition for a solution of the minimization problem (8.9).

Proposition 8.4. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a convex function. Then, $\bar{x} \in K$ is a solution of the minimization problem (8.9) if and only if it is a solution of GMVIP (8.16) with $F(x) = \partial f(x)$ for all $x \in K$.

Proof. Let $\bar{x} \in K$ be a solution of GMVIP (8.16) but not a solution of minimization problem (8.9). Then, there exists $z \in K$ such that

$$f(z) < f(\bar{x}). \tag{8.17}$$

By Theorem 1.50, there exist $\lambda \in \ ]0,1[$ and $v \in \partial f(z(\lambda))$, where $z(\lambda) = \lambda z + (1-\lambda)\bar{x}$, such that

$$\langle v, z - \bar{x} \rangle = f(z) - f(\bar{x}). \tag{8.18}$$

By combining (8.17) and (8.18), we obtain

$$\langle v, z - \bar{x} \rangle < 0.$$

Since $\lambda(z - \bar{x}) = z(\lambda) - \bar{x}$, we have

$$\langle v, z(\lambda) - \bar{x} \rangle < 0,$$

a contradiction to our supposition that $\bar{x}$ is a solution of GMVIP (8.16).

Conversely, suppose that $\bar{x} \in K$ is a solution of the minimization problem (8.9). Then, we have

$$f(y) - f(\bar{x}) \geq 0, \quad \text{for all } y \in K. \tag{8.19}$$

Since f is convex, we deduce that

$$\langle v, y - \bar{x} \rangle \geq f(y) - f(\bar{x}), \quad \text{for all } y \in K \text{ and all } v \in \partial f(y). \tag{8.20}$$

From inequalities (8.19) and (8.20), it follows that $\bar{x}$ is a solution of GMVIP (8.16). □

Definition 8.1. Let K a nonempty convex subset of $\mathbb{R}^n$. A set-valued map $F : K \to 2^{\mathbb{R}^n}$ is said to be *generalized hemicontinuous* if for any $x, y \in K$ and for all $\lambda \in [0, 1]$, the set-valued mapping

$$\lambda \mapsto \langle F(x + \lambda(y - x)), y - x \rangle = \bigcup_{w \in F(x+\lambda(y-x))} \langle w, y - x \rangle$$

is upper semicontinuous at $\mathbf{0}$.

Now we present some existence results for solutions of GVIP under different kinds of generalized monotonicities, which have already been discussed in Section 4.6.

The following result, which was established by Konnov and Yao [139], is a set-valued version of the Minty lemma.

Lemma 8.2 (Generalized Linearization Lemma)**.** Let K be a nonempty convex subset of $\mathbb{R}^n$ and $F : K \to 2^{\mathbb{R}^n}$ be a set-valued map with nonempty values. The following assertions hold.

(a) If F is generalized hemicontinuous, then every solution of WGMVIP is a solution of WGVIP.

(b) If F is generalized pseudomonotone, then every solution of WGVIP is a solution of GMVIP.

(c) If F is generalized weakly pseudomonotone, then every solution of WGVIP is a solution of WGMVIP.

Proof. (a) Let $\bar{x} \in K$ be a solution of WGMVIP but not a solution of WGVIP. Then, for some $y \in K$ and all $\bar{u} \in F(\bar{x})$, we have $\langle \bar{u}, y - \bar{x} \rangle < 0$, that is, $\langle \bar{u}, y - \bar{x} \rangle \in -\text{int}(\mathbb{R}_+)$. Set $x_\lambda = \lambda y + (1 - \lambda)\bar{x}$. Then, by generalized hemicontinuity of F, there exists $\delta > 0$ such that $\langle w, y - \bar{x} \rangle \in -\text{int}(\mathbb{R}_+)$ for all $w \in F(x_\lambda)$, $\lambda \in\ ]0, \delta[$. Since $\lambda(y - \bar{x}) = x_\lambda - \bar{x}$, we have $\langle w, x_\lambda - \bar{x} \rangle < 0$ for all $w \in F(x_\lambda)$, $\lambda \in\ ]0, \delta[$. This contradicts our hypothesis.

(b) and (c) follow directly from the definitions of pseudomonotonicity and weak pseudomonotonicity, respectively. □

Theorem 8.3. Let K be a nonempty compact convex subset of $\mathbb{R}^n$ and $F : K \to 2^{\mathbb{R}^n}$ be a generalized pseudomonotone and generalized hemicontinuous set-valued map such that for each $x \in K$, $F(x)$ is nonempty. Then, there exists a solution $\bar{x} \in K$ of WGVIP. If, in addition, the set $F(\bar{x})$ is also compact and convex, then $\bar{x} \in K$ is a strong solution of GVIP.

Proof. For each $y \in K$, define two set-valued maps $P, Q : K \to 2^K$ by

$$P(y) = \{x \in K : \exists u \in F(x),\ \langle u, y - x\rangle \geq 0\},$$

and

$$Q(y) = \{x \in K : \forall v \in F(y),\ \langle v, y - x\rangle \geq 0\},$$

respectively. We divide the proof into five steps.

(i) We claim that P is a KKM-map, that is, the convex hull $\mathrm{co}(\{y_1, y_2, \ldots, y_m\})$ of every finite subset $\{y_1, y_2, \ldots, y_m\}$ of K is contained in the corresponding union $\bigcup_{i=1}^m P(y_i)$.

Let $\hat{x} \in \mathrm{co}(\{y_1, y_2, \ldots, y_m\})$. Then,

$$\hat{x} = \sum_{i=1}^m \lambda_i y_i, \quad \text{for some } \lambda_i \geq 0 \text{ with } \sum_{i=1}^m \lambda_i = 1.$$

If $\hat{x} \notin \bigcup_{i=1}^m P(y_i)$, then for all $w \in F(\hat{x})$,

$$\langle w, y_i - \hat{x}\rangle < 0, \quad \text{for all } i = 1, 2, \ldots, m.$$

For all $w \in F(\hat{x})$, it follows that

$$\begin{aligned} 0 &= \langle w, \hat{x} - \hat{x}\rangle \\ &= \left\langle w, \sum_{i=1}^m \lambda_i y_i - \sum_{i=1}^m \lambda_i \hat{x} \right\rangle \\ &= \left\langle w, \sum_{i=1}^m \lambda_i (y_i - \hat{x}) \right\rangle \\ &= \sum_{i=1}^m \lambda_i \langle w, y_i - \hat{x}\rangle < 0, \end{aligned}$$

which is a contradiction. Therefore, we must have

$$\mathrm{co}(\{y_1, y_2, \ldots, y_m\}) \subseteq \bigcup_{i=1}^m P(y_i),$$

and hence P is a KKM-map.

(ii) We show that $P(y) \subseteq Q(y)$ for all $y \in K$, and hence Q is a KKM-map.

By generalized pseudomonotonicity of F, we have that $P(y) \subseteq Q(y)$ for all $y \in K$. Since P is a KKM-map, so is Q.

(iii) We assert that $\bigcap_{y\in K} P(y) = \bigcap_{y\in K} Q(y)$.

From step (ii), we have

$$\bigcap_{y\in K} P(y) \subseteq \bigcap_{y\in K} Q(y),$$

and from Lemma 8.2, we have

$$\bigcap_{y\in K} P(y) \supseteq \bigcap_{y\in K} Q(y).$$

Therefore, the conclusion follows.

(iv) We prove that for each $y \in K$, $Q(y)$ is a closed subset of K.

For any fixed $y \in K$, let $\{x_m\}$ be a sequence in $Q(y)$ such that $x_m \to \tilde{x} \in K$. Since $x_m \in Q(y)$, for all $v \in F(y)$, we have $\langle v, y - x_m\rangle \geq 0$ for all m. As $\langle v, y - x_m\rangle$ converges to $\langle v, y - \tilde{x}\rangle$, therefore $\langle v, y - \tilde{x}\rangle \geq 0$, and hence $\tilde{x} \in Q(y)$. Consequently, $Q(y)$ is closed.

(v) Finally, we show that the WGVIP is solvable.

From step (iv), $Q(y)$ is a closed subset of the compact set K, and hence it is compact. By step (ii) and Theorem B.3, we have $\bigcap_{y\in K} Q(y) \neq \emptyset$. Consequently, by step (iii), we also have $\bigcap_{y\in K} P(y) \neq \emptyset$. Hence, there exists $\bar{x} \in K$ such that

$$\forall y \in K,\ \exists\ \bar{u} \in F(\bar{x}) \quad : \quad \langle \bar{u}, y - \bar{x}\rangle \geq 0. \tag{8.21}$$

Thus, $\bar{x}$ is a solution of WGVIP.

If, in addition, the set $F(\bar{x})$ is also compact and convex, then by Lemma 8.1, $\bar{x} \in K$ is a strong solution of GVIP. □

Definition 8.2. Let K be a nonempty subset of $\mathbb{R}^n$. A set-valued map $F : K \to 2^{\mathbb{R}^n}$ is said to be

(a) *generalized coercive* [139] if there exist a bounded subset D of $\mathbb{R}^n$ and $\tilde{y} \in D \cap K$ such that for all $u \in F(x)$,

$$\langle u, \tilde{y} - x\rangle < 0, \quad \text{for all } x \in K \setminus D; \tag{8.22}$$

(b) *generalized α-coercive* [139] if there exist a point $\tilde{y} \in K$ and a number $\alpha > 0$ such that for all $u \in F(x)$,

$$\langle u, \tilde{y} - x\rangle < 0, \quad \text{if } x \in K \text{ and } \|\tilde{y} - x\| > \alpha. \tag{8.23}$$

When K is not necessarily bounded we assume some coercivity condition so as to impose some sort of boundedness on the solution in order to establish the existence of at least one solution. For instance, the generalized coercivity condition defined above ensures that no point outside the bounded set D is a candidate for solution.

Theorem 8.4. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to 2^{\mathbb{R}^n}$ be a generalized pseudomonotone, generalized coercive, and generalized hemicontinuous set-valued map such that for each $x \in K$, $F(x)$ is nonempty. Then, there exists a solution $\bar{x} \in K$ of WGVIP. If, in addition, the set $F(\bar{x})$ is also compact and convex, then $\bar{x}$ is a strong solution of GVIP.

Proof. Let P and Q be the same as in the proof of Theorem 8.3. Choose a bounded subset D of $\mathbb{R}^n$ and $\tilde{y} \in D \cap K$ such that for all $u \in F(x)$, relation (8.22) holds.

The proof goes through steps (i)–(iv) in the proof of Theorem 8.3 and the following two additional steps, steps (v) and (vi).

(v) We claim that the closure $\mathrm{cl}(P(\tilde{y}))$ of $P(\tilde{y})$ is compact.

If $P(\tilde{y}) \not\subseteq D$, then there exists $\tilde{x} \in P(\tilde{y})$ such that $\tilde{x} \in K \setminus D$. It follows that, for some $\tilde{u} \in F(\tilde{x})$, $\langle \tilde{u}, \tilde{y} - \tilde{x} \rangle \geq 0$, a contradiction to the relation (8.22). Hence, we must have $P(\tilde{y}) \subseteq D$. Therefore, $\mathrm{cl}(P(\tilde{y}))$ is a compact subset of D.

(vi) Finally, we exhibit that WGVIP is solvable.

From step (i), (v), and Theorem B.3, we have $\bigcap_{y \in K} \mathrm{cl}(P(y)) \neq \emptyset$. By step (iv) and the generalized pseudomonotonicity of F, $Q(y)$ is closed and $P(y) \subseteq Q(y)$ for all $y \in K$, respectively. Thus, $\mathrm{cl}(P(y)) \subseteq Q(y)$ for all $y \in K$. Consequently,

$$\bigcap_{y \in K} Q(y) \neq \emptyset.$$

From step (iii),

$$\bigcap_{y \in K} P(y) = \bigcap_{y \in K} Q(y) \neq \emptyset.$$

Hence, there exists $\bar{x} \in K$ such that

$$\forall y \in K, \ \exists \bar{u} \in F(\bar{x}) \quad : \quad \langle \bar{u}, y - \bar{x} \rangle \geq 0. \tag{8.24}$$

Thus, $\bar{x}$ is a solution of WGVIP.

If, in addition, the set $F(\bar{x})$ is also compact and convex, then by Lemma 8.1, $\bar{x}$ is a strong solution of GVIP. □

Theorem 8.5. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $F : K \to 2^{\mathbb{R}^n}$ be a generalized weakly pseudomonotone and generalized hemicontinuous set-valued map such that for each $x \in K$, $F(x)$ is nonempty compact. Suppose that at least one of the following assumptions holds:

(i) K is compact.

(ii) K is closed and F is generalized α-coercive.

Then, there exists a solution $\bar{x} \in K$ of WGVIP. If, in addition, the set $F(\bar{x})$ is also convex, then $\bar{x}$ is a strong solution of GVIP.

Proof. For each $y \in K$, define two set-valued maps $P, Q : K \to 2^K$ by

$$P(y) = \{x \in K : \exists u \in F(x) \text{ such that } \langle u, y - x\rangle \geq 0\},$$

and

$$Q(y) = \{x \in K : \exists v \in F(y) \text{ such that } \langle v, y - x\rangle \geq 0\},$$

respectively. In order to prove the present theorem under assumption (i), it suffices to follow the proof of Theorem 8.3 and replace steps (ii) and (iv) in the proof of Theorem 8.3 by the following steps.

(ii′) We prove that $P(y) \subseteq Q(y)$ for all $y \in K$, and Q is a KKM-map.

By generalized weak pseudomonotonicity of F, we have $P(y) \subseteq Q(y)$ for all $y \in K$. Since P is a KKM-map, so is Q.

(iv′) We assert that for each $y \in K$, $Q(y)$ is a closed subset of K.

For any fixed $y \in K$, let $\{x_m\}$ be a sequence in $Q(y)$ such that $x_m \to \tilde{x} \in K$. Since $x_m \in Q(y)$, for some $v_m \in F(y)$, we have $\langle v_m, y - x_m\rangle \geq 0$ for all m. Since $F(y)$ is compact, we may assume, by taking a subsequence if necessary, that $\{v_m\}$ converges to some $v \in F(y)$. As K is compact, $\{x_m\}$ is bounded. Therefore, $\langle v_m - v, y - x_m\rangle$ converges to 0, but $\langle v, y - x_m\rangle$ converges to $\langle v, y - \tilde{x}\rangle$. Hence $\langle v_m, y - x_m\rangle$ converges to $\langle v, y - \tilde{x}\rangle$. Now we get $(x_m, \langle v_m, y - x_m\rangle)$ converges to $(\tilde{x}, \langle v, y - \tilde{x}\rangle)$. Therefore, for $v \in F(y)$,

$$\langle v, y - \tilde{x}\rangle \geq 0,$$

and thus, $\tilde{x} \in Q(y)$. Consequently, $Q(y)$ is a closed subset of the compact set K.

Let us consider the case under the assumption (ii). Consider the closed ball $\mathbb{B}_r[\mathbf{0}]$ with center at origin and radius $r > 0$. If $K \cap \mathbb{B}_r[\mathbf{0}] \neq \emptyset$, part (i) guarantees the existence of a solution $\bar{x}_r$ of the following problem $(\text{WGVIP})_r$: Find $\bar{x}_r \in K \cap \mathbb{B}_r[\mathbf{0}]$ such that

$$\forall y \in K \cap \mathbb{B}_r[\mathbf{0}], \quad \exists \bar{u} \in F(\bar{x}_r) \quad : \quad \langle \bar{u}, y - \bar{x}_r\rangle \geq 0.$$

We observe that $\{\bar{x}_r : r > 0\}$ must be bounded. Otherwise, we can choose r large enough so that $r \geq \|\tilde{y}\|$ and $\alpha < \|\tilde{y} - \bar{x}_r\|$, where $\tilde{y}$ satisfies the generalized α-coercivity condition of F. It follows that, for all $\bar{u} \in F(\bar{x}_r)$, $\langle \bar{u}, \tilde{y} - \bar{x}_r\rangle < 0$, that is, $\bar{x}_r$ is not a solution of $(\text{WGVIP})_r$, a contradiction. Therefore, there exists $r > 0$ such that $\|\bar{x}_r\| < r$. Choose any $x \in K$. Then, we can choose $\varepsilon > 0$ small enough such that $\bar{x}_r + \varepsilon(x - \bar{x}_r) \in K \cap \mathbb{B}_r[\mathbf{0}]$. If we suppose that, for every $\bar{u} \in F(\bar{x}_r)$, $\langle \bar{u}, x - \bar{x}_r\rangle < 0$, then

$$\langle \bar{u}, \bar{x}_r + \varepsilon(x - \bar{x}_r) - \bar{x}_r\rangle = \varepsilon\langle \bar{u}, x - \bar{x}_r\rangle < 0,$$

that is, $\bar{x}_r$ is not a solution of $(\text{WGVIP})_r$, a contradiction. Thus, $\bar{x}_r$ is a solution of WGVIP.

If, in addition, the set $F(\bar{x})$ is also compact and convex, then by Lemma 8.1, $\bar{x}$ is a strong solution of GVIP. □

Under different kinds of coercivity conditions, the existence results for a weak solution of GVIP were derived by Daniilidis and Hadjisavvas [62].

The following result shows that the solution of GVIP is unique provided that F is generalized strictly pseudomonotone.

Proposition 8.5. Let K be a nonempty closed convex subset of $\mathbb{R}^n$. If $F : K \to 2^{\mathbb{R}^n}$ is a generalized strictly pseudomonotone set-valued map, then the solution $(\bar{x}, \bar{u})$ of GVIP is unique, if it exists.

Proof. Suppose that there exist two distinct solutions $(\bar{x}, \bar{u})$ and $(\hat{x}, \hat{u})$ of GVIP. Then, $\bar{u} \in F(\bar{x})$, $\hat{u} \in F(\hat{x})$,

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K,$$

and

$$\langle \hat{u}, y - \hat{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

It follows that

$$\langle \bar{u}, \hat{x} - \bar{x} \rangle \geq 0, \tag{8.25}$$

and

$$\langle \hat{u}, \bar{x} - \hat{x} \rangle \geq 0. \tag{8.26}$$

Since F is generalized strictly pseudomonotone, it follows from the inequality (8.25) that

$$\langle \hat{u}, \bar{x} - \hat{x} \rangle < 0,$$

a contradiction to the inequality (8.26). Therefore, the solution $(\bar{x}, \bar{u})$ of GVIP is unique. □

Definition 8.3. [207] Let K be a nonempty subset of $\mathbb{R}^n$. A set-valued map $F : K \to 2^{\mathbb{R}^n}$ is said to be *generalized α-pseudomonotone* if there exist $x_0 \in K$, $u_0 \in F(x_0)$, and $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$, with

$$\alpha(0) = 0, \quad \alpha(r) > 0 \text{ for } r > 0, \text{ and } \liminf_{r \to +\infty} \alpha(r) > \|u_0\|,$$

such that for every pair of distinct points $x, y \in K$ and for all $u \in F(x)$, $v \in F(y)$, we have

$$\langle u, y - x \rangle \geq 0 \quad \Rightarrow \quad \langle v, y - x \rangle \geq \|y - x\| \alpha(\|y - x\|).$$

The following existence and uniqueness result under generalized α-pseudomonotonicity is due to Yao [205].

Theorem 8.6. Let K be a nonempty closed convex subset of $\mathbb{R}^n$ and $F : K \to 2^{\mathbb{R}^n}$ be a generalized α-pseudomonotone and generalized hemicontinuous set-valued map such that for each $x \in K$, $F(x)$ is nonempty, compact, and convex. Then, there exists a unique solution $(\bar{x}, \bar{u})$ of GVIP.

Proof. By the generalized α-pseudomonotonicity of F, there exist $x_0 \in K$, $u_0 \in F(x_0)$, and $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$, with

$$\alpha(0) = 0, \quad \alpha(r) > 0 \text{ for } r > 0, \text{ and } \liminf_{r \to +\infty} \alpha(r) > \|u_0\|,$$

such that for every pair of distinct points $x, y \in K$ and for any $u \in F(x)$, $v \in F(y)$, we have

$$\langle u, y - x \rangle \geq 0 \quad \Rightarrow \quad \langle v, y - x \rangle \geq \|y - x\| \alpha(\|y - x\|).$$

For each $y \in K$, let

$$P(y) = \left\{ x \in K : \inf_{u \in F(x)} \langle u, x - y \rangle \leq 0 \right\},$$

and

$$Q(y) = \left\{ x \in K : \sup_{v \in F(y)} \langle v, x - y \rangle \leq 0 \right\}.$$

We claim that $P(x_0)$ is bounded. If $P(x_0)$ is not bounded, then there exists a sequence $\{x_m\}$ in $P(x_0)$ such that $\|x_m\| \to +\infty$. For each m, since $F(x_m)$ is compact, there exists $u_m \in F(x_m)$ such that

$$\langle u_m, x_m - x_0 \rangle = \inf_{u \in F(x_m)} \langle u, x_m - x_0 \rangle \leq 0.$$

Thus, $\langle u_m, x_0 - x_m \rangle \geq 0$ for all m. Since F is generalized α-pseudomonotone, we have

$$\langle u_0, x_0 - x_m \rangle \geq \|x_0 - x_m\| \alpha(\|x_0 - x_m\|), \quad \text{for all } m.$$

Consequently, we have $\alpha(\|x_0 - x_m\|) \leq \|u_0\|$ for all m, from which it follows that

$$\|u_0\| < \liminf_{m \to +\infty} \alpha(\|x_0 - x_m\|) \leq \|u_0\|,$$

which is a contradiction. Therefore, $P(x_0)$ is bounded as claimed.

By employing the same argument as in the proof of Theorem 8.4, we have $\bigcap_{y \in K} P(y) \neq \emptyset$. Hence, there exists $\bar{x} \in K$ such that

$$\inf_{u \in F(\bar{x})} \langle u, \bar{x} - y \rangle \leq 0, \quad \text{for all } y \in K,$$

that is,

$$\sup_{y \in K} \inf_{u \in F(\bar{x})} \langle u, \bar{x} - y \rangle \leq 0.$$

By Theorem B.7, we have

$$\inf_{u \in F(\bar{x})} \sup_{y \in K} \langle u, \bar{x} - y \rangle = \sup_{y \in K} \inf_{u \in F(\bar{x})} \langle u, \bar{x} - y \rangle \leq 0.$$

Since $F(\bar{x})$ is compact, there exists $\bar{u} \in F(\bar{x})$ such that

$$\langle \bar{u}, \bar{x} - y \rangle \leq 0, \quad \text{for all } y \in K.$$

Finally, by Proposition 8.5, $(\bar{x}, \bar{u})$ is unique. $\square$

Since every generalized strongly monotone set-valued map is generalized α-pseudomonotone, the above result also holds under generalized strong monotonicity in place of generalized α-pseudomonotonicity.

Similar results under different assumptions have been proven [66, 191, 205, 207]. Konnov [135] established some existence results for a weak solution of GVIP.

Let K be nonempty subset of $\mathbb{R}^n$ and $F : K \to 2^{\mathbb{R}^n}$ be a set-valued map. A function $f : K \to \mathbb{R}^n$ is said to be a selection of F on K if $f(x) \in F(x)$ for all $x \in K$.

Furthermore, the function f is called a continuous selection of F on K if it is a selection of F and also continuous.

For results on the existence of a continuous selection we refer to Repovs and Semenov [179] and the references therein.

The variational inequality problem (VIP) defined by f and K, which we have already studied in Chapter 5, is to find $\bar{x} \in K$ such that

$$\langle f(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

The following lemma provides a relation between VIP and GVIP, which is the key to establish the existence of a solution of GVIP by using VIP.

Lemma 8.3. If f is a selection of F on K, then every solution of VIP is a strong solution of GVIP.

Proof. Let $\bar{x} \in K$ be a solution of VIP. Then,

$$\langle f(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

Let $\bar{u} = f(\bar{x})$. Then, $\bar{u} \in F(\bar{x})$ and

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K.$$

Hence, $(\bar{x}, \bar{u})$ is a solution of GVIP. $\square$

The following lemma provides a connection between the generalized pseudomonotonicity of a set-valued map and its selection. Its proof follows directly from the definitions, and therefore is omitted.

Lemma 8.4. Let $f : K \to \mathbb{R}^n$ be a selection of the set-valued map $F : K \to 2^{\mathbb{R}^n}$. If F is generalized pseudomonotone, then f is pseudomonotone.

The following definition of generalized B-pseudomonotonicity was introduced by Ansari and Khan [5], which is a generalization of the definition of B-pseudomonotonicity given in Definition 5.3.

Definition 8.4. Let K be a nonempty subset of $\mathbb{R}^n$. A set-valued map F :

$K \to 2^{\mathbb{R}^n}$ is said to be *generalized B-pseudomonotone* if for each $x \in K$ and every sequence $\{x_m\}$ in K converging to x with

$$\liminf_{m\to\infty} \langle u_m, x - x_m\rangle \geq 0, \quad \text{for all } u_m \in F(x_m)$$

we have, for all $u \in F(x)$,

$$\langle u, y - x\rangle \geq \limsup_{m\to\infty} \langle u_m, y - x_m\rangle, \quad \text{for all } u_m \in F(x_m) \text{ and } y \in K.$$

Lemma 8.5. Let $f : K \to \mathbb{R}^n$ be a selection of the set-valued map $F : K \to 2^{\mathbb{R}^n}$. If F is generalized B-pseudomonotone, then f is B-pseudomonotone.

By using Lemmas 8.4 and 8.5, and the results presented in Chapter 5, we can derive several existence results for solutions of GVIP.

8.4 Gap Functions for Generalized Variational Inequalities

We have seen in Section 5.4 that by using a gap function, one can convert a variational inequality problem into an optimization problem. This section deals with the study of the gap functions for GVIP, WGVIP, and GMVIP. Yang and Yao [204] considered the following gap functions for GVIP and WGVIP.

Definition 8.5. Let K be a nonempty subset of $\mathbb{R}^n$. A function $\Phi : K \to \mathbb{R}$ is said to be a *gap function* for GVIP if and only if it satisfies the following properties:

(i) $\Phi(x) \geq 0$ for all $x \in K$,

(ii) $\Phi(\bar{x}) = 0$ if and only if $\bar{x}$ is a strong solution of GVIP.

In a similar way, we can define gap functions for WGVIP and GMVIP.

We define two functional $\Phi_1 : K \times \mathbb{R}^n \to \mathbb{R}$ and $\Phi : K \to \mathbb{R}$ by

$$\Phi_1(x, u) = \max_{y\in K} \langle u, x - y\rangle, \tag{8.27}$$

and

$$\Phi(x) = \min_{u\in F(x)} \Phi_1(x, u). \tag{8.28}$$

If K is compact, then Φ_1 is well defined and upper semicontinuous function in x. Also, if for all $x \in K$, $F(x)$ is compact, then Φ is well-defined. For $x \in K$ and $u \in F(x)$, it is easy to see that

$$\Phi_1(x, u) = \max_{y\in K} \langle u, x - y\rangle \geq 0.$$

Theorem 8.7. Assume that K is a compact subset of $\mathbb{R}^n$ and $F(x)$ is compact for all $x \in K$. The function Φ defined by (8.28) is a gap function for GVIP.

Proof. Clearly, $\Phi_1(x,u) \geq 0$ for all $x \in K$ and $u \in F(x)$. Thus, $\Phi(x) \geq 0$ for all $x \in K$.

If $\Phi(\bar{x}) = 0$, then there exists $\bar{u} \in F(\bar{x})$ such that $\Phi_1(\bar{x},\bar{u}) = 0$. Consequently, we have

$$\max_{y \in K} \langle \bar{u}, \bar{x} - y \rangle = 0,$$

if and only if,

$$\langle \bar{u}, \bar{x} - y \rangle \leq 0, \quad \text{for all } y \in K,$$

that is, $\bar{x}$ is a strong solution of GVIP. □

By Theorem 8.7, finding a solution of GVIP is equivalent to finding a global solution to the following optimization problem:

$$\min \Phi(x), \quad \text{subject to } x \in K, \tag{8.29}$$

with $\Phi(\bar{x}) = 0$.

Next, we define a gap function for WGVIP. To this end, for $x \in K$, let

$$S_x = \{u : \text{ where } u : K \to F(x)\},$$

that is, S_x is the set of all mappings u from K to $F(x)$. Let $x \in K$ and $u \in S_x$. Then, $u_y \in F(x)$ for all $y \in K$.

Define two mappings $\Phi_1^* : K \times S_x \to \mathbb{R}$ and $\Phi^* : K \to \mathbb{R}$ by

$$\Phi_1^*(x,u) = \max_{y \in K} \langle u_y, x - y \rangle, \tag{8.30}$$

and

$$\Phi^*(x) = \min_{u \in S_x} \Phi_1^*(x,u). \tag{8.31}$$

Theorem 8.8. The function Φ^* defined by (8.31) is a gap function for WGVIP.

Proof. It is clear that $\Phi_1^*(x,u) \geq 0$ for all $x \in K$ and $u \in S_x$, and hence, $\Phi^*(x) \geq 0$ for all $x \in K$.

If $\Phi^*(\bar{x}) = 0$, then there exists $\bar{u} \in S_{\bar{x}}$ such that $\Phi_1^*(\bar{x},\bar{u}) = 0$. Consequently, we have

$$\max_{y \in K} \langle \bar{u}_y, \bar{x} - y \rangle = 0,$$

if and only if,

$$\langle \bar{u}_y, \bar{x} - y \rangle \leq 0, \quad \text{for all } y \in K,$$

that is, $\bar{x}$ is a solution of WGVIP.

Assume that $\bar{x} \in K$ is a solution of WGVIP. Then, for each $y \in K$, there exists $\bar{u}_y \in F(\bar{x})$ such that

$$\langle \bar{u}_y, \bar{x} - y \rangle \leq 0.$$

Thus, a mapping $\bar{u} : K \to F(\bar{x})$ has been defined. Then, $\bar{u} \in S_{\bar{x}}$ and

$$\langle \bar{u}_y, \bar{x} - y \rangle \leq 0, \quad \text{for all } y \in K.$$

Hence,

$$\Phi_1^*(\bar{x}, \bar{u}) = \max_{y \in K} \langle \bar{u}_y, \bar{x} - y \rangle \leq 0.$$

So, $\Phi_1^*(\bar{x}, \bar{u}) = 0$. Also, for any $u \in S_{\bar{x}}$, $\langle u_y, \bar{x} - \bar{x} \rangle = 0$, from which it follows that $\Phi_1^*(\bar{x}, u) = 0$, and consequently, $\Phi^*(\bar{x}) = 0$. □

Now, we study the gap function for GMVIP (8.16), which was investigated by Ansari and Yao [16] and Crouzeix [57].

We define a functional $\Psi : K \to \mathbb{R}$ as

$$\Psi(x) = \sup \left\{ \langle v, x - y \rangle : y \in K, \ v \in F(y) \right\}. \tag{8.32}$$

This function Ψ is nonnegative on K, closed, and convex on $\mathbb{R}^n$ as a supremum of affine functions.

We also set

$$m = \inf_{x \in K} \Psi(x) \quad \text{and} \quad M = \{x \in K : \Psi(x) = m\}.$$

Theorem 8.9. The functional Ψ defined by (8.32) is a gap function for GMVIP.

Proof. (i) Since $\langle u, x - x \rangle = 0$ for all $x \in K$ and $u \in F(x)$, we have

$$\Psi(x) \geq 0, \quad \text{for all } x \in K. \tag{8.33}$$

(ii) Suppose that $\bar{x} \in K$ is a solution of GMVIP, then

$$\langle v, \bar{x} - y \rangle \leq 0, \quad \text{for all } y \in K \text{ and } v \in F(y),$$

and hence,

$$\sup\{\langle v, \bar{x} - y \rangle : y \in K, \ v \in F(y)\} \leq 0. \tag{8.34}$$

This implies that $\Psi(\bar{x}) \leq 0$. Combining (8.33) and (8.34), we obtain

$$\Psi(\bar{x}) = 0. \tag{8.35}$$

Conversely, let $\Psi(\bar{x}) = 0$. From (8.32), we have

$$\Psi(\bar{x}) \geq \langle v, \bar{x} - y \rangle, \quad \text{for all } y \in K \text{ and } v \in F(y),$$

and hence

$$\langle v, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in K \text{ and } v \in F(y).$$

Therefore, $\bar{x} \in K$ is a solution of GMVIP. □

Theorem 8.10. If the solution set of GMVIP is nonempty, then $m = 0$, and the set M and the solution set of GMVIP are precisely the same.

Proof. Assume that the solution set of GMVIP is nonempty. Then from (8.35), $m = 0$.

Let $\bar{x} \in K$ be a solution of GMVIP. Then, $\Psi(\bar{x}) = 0$. But from (8.33), we have $\Psi(x) \geq 0$ for all $x \in K$, and hence $\Psi(\bar{x}) \leq \Psi(x)$ for all $x \in K$. Therefore, $\bar{x} \in M$.

Conversely, assume that $\bar{x} \in M$. Then, $\Psi(\bar{x}) = 0$, and thus $\bar{x}$ is a solution of GMVIP. □

8.5 Generalized Variational Inequalities in Terms of the Clarke Subdifferential and Optimization Problems

This section deals with generalized variational inequalities defined by means of the Clarke subdifferential. Like the classical variational inequality problem, these variational inequalities have strong links with minimization problems.

Let $f : K \to \mathbb{R}$ be a locally Lipschitz function. The *weak generalized variational inequality problem* (WGVIP) and *generalized Minty variational inequality problem* (GMVIP) in terms of the Clarke subdifferential are defined as follows:

$$\text{WGVIP} \begin{cases} \text{Find } \bar{x} \in K \text{ such that for all } y \in K, \text{ there exists } \bar{u} \in \partial^C f(\bar{x}) \\ \text{satisfying } \langle \bar{u}, y - \bar{x} \rangle \geq 0. \end{cases} \tag{8.36}$$

$$\text{GMVIP} \begin{cases} \text{Find } \bar{x} \in K \text{ such that for all } y \in K \text{ and all } v \in \partial^C f(y), \\ \text{we have } \langle v, y - \bar{x} \rangle \geq 0. \end{cases} \tag{8.37}$$

Here we present a result that provides the equivalence between the solutions of GMVIP and WGVIP.

Theorem 8.11. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be locally Lipschitz and generalized pseudoconvex. Then, $\bar{x} \in K$ is a solution of WGVIP (8.36) if and only if it is a solution of GMVIP (8.37).

Proof. By Theorem 4.33, $\partial^C f$ is generalized pseudomonotone. Therefore, by Lemma 8.2, every solution of WGVIP (8.36) is a solution of GMVIP (8.37).

Conversely, let $\bar{x} \in K$ be a solution of GMVIP (8.37). Consider any $y \in K$ and any sequence $\{\lambda_m\} \to 0^+$ with $\lambda_m \in \,]0, 1]$. Since K is convex,

$$y_m := \bar{x} + \lambda_m (y - \bar{x}) \in K.$$

Since $\bar{x} \in K$ is a solution of GMVIP (8.37), there exist $v_m \in \partial^C f(y_m)$ such that

$$\langle v_m, y_m - \bar{x} \rangle \geq 0,$$

which implies that

$$\langle v_m, y - \bar{x} \rangle \geq 0.$$

By Theorem 2.18, there exists $k > 0$ such that $\|v_m\| \leq k$ for sufficiently large m. So, we can assume that the sequence $\{v_m\}$ converges to v. Since the set-valued map $y \mapsto \partial f^C(y)$ is closed, $v_m \in \partial^C f(y_m)$ and $y_m \to \bar{x}$ as $m \to \infty$, we have $v \in \partial^C f(\bar{x})$. Thus, for any $y \in K$, there exists $\bar{u} \in \partial^C f(\bar{x})$ such that

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0.$$

Hence, $\bar{x} \in K$ is a solution of WGVIP (8.36). □

The following result provides a necessary and sufficient condition for a solution of optimization problem (8.9).

Theorem 8.12. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be locally Lipschitz and generalized pseudoconvex. Then, $\bar{x} \in K$ is a solution of the optimization problem (8.9) if and only if it is a solution of GMVIP (8.37).

Proof. Let $\bar{x} \in K$ be a solution of GMVIP (8.37) but not a solution of the optimization problem (8.9). Then, there exists $z \in K$ such that

$$f(z) < f(\bar{x}). \tag{8.38}$$

Set $z(\lambda) := \bar{x} + \lambda(z - \bar{x})$ for all $\lambda \in [0, 1]$. Since K is convex, $z(\lambda) \in K$ for all $\lambda \in [0, 1]$. By Lebourg's mean value theorem (Theorem 2.23), there exist $\lambda \in \ ]0, 1[$ and $v \in \partial^C f(z(\lambda))$ such that

$$\langle v, z - \bar{x} \rangle = f(z) - f(\bar{x}). \tag{8.39}$$

By combining (8.38)–(8.39), we obtain

$$\langle v, z - \bar{x} \rangle < 0.$$

Since $\lambda(z - \bar{x}) = z(\lambda) - \bar{x}$, we have

$$\langle v, z(\lambda) - \bar{x} \rangle < 0,$$

which contradicts our supposition that $\bar{x}$ is a solution of GMVIP (8.37).

Conversely, suppose that $\bar{x} \in K$ is a solution of the optimization problem (8.9). Then, we have

$$f(\bar{x}) \leq f(y), \quad \text{for all } y \in K. \tag{8.40}$$

Since f is generalized pseudoconvex, by Theorem 3.21, f is quasiconvex. Further, by Theorem 3.19, f is generalized quasiconvex, and hence

$$\langle v, \bar{x} - y \rangle \leq 0, \quad \text{for all } \ y \in K, \ v \in \partial^C f(y).$$

Thus, $\bar{x}$ is a solution of GMVIP (8.37). □

Theorem 8.13. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a generalized convex function. If $\bar{x} \in K$ is a solution WGVIP (8.36), then it is a solution of the optimization problem (8.9).

Proof. Since $\bar{x} \in K$ is a solution of WGVIP (8.36), for any $y \in K$, there exists $\bar{u} \in \partial^C f(\bar{x})$ such that

$$\langle \bar{u}, y - \bar{x} \rangle \geq 0. \tag{8.41}$$

Since f is generalized convex, we have

$$\langle \bar{u}, y - \bar{x} \rangle \leq f(y) - f(\bar{x}) \quad \text{for any } y \in K. \tag{8.42}$$

By combining (8.41) and (8.42), we obtain

$$f(y) \geq f(\bar{x}), \quad \text{for all } y \in K.$$

Thus, $\bar{x} \in K$ is a solution of the optimization problem (8.9). □

Theorem 8.14. Let K be a nonempty subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be locally Lipschitz such that $-f$ is generalized strictly convex, that is, for all $x, y \in K$,

$$\langle u, y - x \rangle > f(y) - f(x), \quad \text{for all } u \in \partial^C f(x). \tag{8.43}$$

If $\bar{x} \in K$ is a solution of the optimization problem (8.9), then it is a solution of WGVIP (8.36).

Proof. Suppose that $\bar{x}$ is not a solution of WGVIP (8.36). Then, there exists $y \in K$ such that

$$\langle u, y - \bar{x} \rangle < 0, \quad \text{for all } u \in \partial^C f(\bar{x}). \tag{8.44}$$

Combining inequalities (8.43) and (8.44), we obtain

$$f(y) < f(\bar{x})$$

which contradicts our supposition that $\bar{x}$ is a solution of the optimization problem (8.9). □

The problems WGVIP and GMVIP defined by means of Dini upper subdifferentials were considered and studied by Al-Homidan and Ansari [2]. Almost all the results mentioned above have been extended for these problems [2].

8.6 Characterizations of Solution Sets of an Optimization Problem with Generalized Pseudolinear Objective Function

In this section, we present some characterizations of solution sets of an optimization problem with a generalized pseudolinear objective function.

Consider the following optimization problem:

$$\text{(P)} \qquad \min f(x) \quad \text{subject to } x \in K,$$

where K is a nonempty subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ is a function. We denote by S the solution set of the problem (P) and assume that S is nonempty. We know that the solution set S is convex if the function f is convex. The next result shows that this is also true for the case when the objective function is generalized pseudolinear.

Proposition 8.6. Let K be a nonempty convex subset of $\mathbb{R}^n$. The solution set S of problem (P) is convex if $f : K \to \mathbb{R}$ is locally Lipschitz and generalized pseudolinear.

Proof. Suppose that $x_1, x_2 \in S$. Then $f(x_1) \leq f(y)$ and $f(x_2) \leq f(y)$ for all $y \in K$. Since $x_1, x_2 \in S$, we have $f(x_1) = f(x_2)$. By Theorem 3.23, for some $\xi_1 \in \partial^C f(x_1)$, $\langle \xi_1, x_2 - x_1 \rangle = 0$, that is, $\langle \xi_1, x_1 - x_2 \rangle = 0$, and so $-\lambda\langle \xi_1, x_1 - x_2 \rangle = 0$ for all $\lambda \in [0, 1]$. Therefore,

$$\langle \xi_1, x_1 - x_1 + \lambda(x_2 - x_1) \rangle = -\lambda\langle \xi_1, x_1 - x_2 \rangle = 0.$$

Since f is generalized pseudolinear, $f(x_1 + \lambda(x_2 - x_1)) = f(x_1)$, and therefore $x_1 + \lambda(x_2 - x_1)$ is also a solution of the optimization problem, and thus the solution set of problem (P) is convex. $\square$

We give some characterizations of the solution set of a generalized pseudolinear program in terms of any one of its solutions.

Theorem 8.15. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $f : K \to \mathbb{R}$ be a locally Lipschitz generalized pseudolinear function. Let $\bar{x} \in S$, then $S = S_1 = S_2$, where

$$\begin{aligned} S_1 &= \left\{ x \in K : \langle \xi, \bar{x} - x \rangle = 0 \text{ for some } \xi \in \partial^C f(x) \right\}, \\ S_2 &= \left\{ x \in K : \langle \zeta, \bar{x} - x \rangle = 0 \text{ for some } \zeta \in \partial^C f(\bar{x}) \right\}. \end{aligned}$$

Proof. A point $x \in S$ if and only if $f(x) = f(\bar{x})$. Then from Theorem 3.23, we have $f(x) = f(\bar{x})$ if and only if $\langle \xi, \bar{x} - x \rangle = 0$ for some $\xi \in \partial^C f(x)$. Also, $f(\bar{x}) = f(x)$ if and only if $\langle \zeta, x - \bar{x} \rangle = 0$ for some $\zeta \in \partial^C f(\bar{x})$, equivalently $\langle \zeta, \bar{x} - x \rangle = 0$ for some $\zeta \in \partial^C f(\bar{x})$. $\square$

Corollary 8.1. Let K and f be the same as in Theorem 8.15 and let $\bar{x} \in S$. Then $S = S_3 = S_4$, where

$$\begin{aligned} S_3 &= \left\{ x \in K : \langle \xi, \bar{x} - x \rangle \geq 0 \text{ for some } \xi \in \partial^C f(x) \right\}, \\ S_4 &= \left\{ x \in K : \langle \zeta, \bar{x} - x \rangle \geq 0 \text{ for some } \zeta \in \partial^C f(\bar{x}) \right\}. \end{aligned}$$

Proof. It is clear from Theorem 8.15 that $S \subseteq S_3$. We prove that $S_3 \subseteq S$. Assume that $x \in S_3$, that is, $x \in K$ is such that $\langle \xi, \bar{x} - x \rangle \geq 0$ for some $\xi \in \partial^C f(x)$. Then by generalized pseudoconvexity of f, it follows that $f(\bar{x}) \geq f(x)$. This implies that $x \in S$, and hence $S_3 \subseteq S$. Similarly, we can easily prove that $S = S_4$ by making use of generalized pseudoconcavity of f. □

Theorem 8.16. Let K be a nonempty convex subset of $\mathbb{R}^n$. If $f : K \to \mathbb{R}$ is locally Lipschitz and generalized pseudolinear and $\bar{x} \in S$, then $S = S_5 = S_6$, where

$$\begin{aligned} S_5 &= \left\{ x \in K : \langle \zeta, \bar{x} - x \rangle = \langle \xi, x - \bar{x} \rangle \text{ for some } \xi \in \partial^C f(x),\ \zeta \in \partial^c f(\bar{x}) \right\}, \\ S_6 &= \left\{ x \in K : \langle \zeta, \bar{x} - x \rangle \geq \langle \xi, x - \bar{x} \rangle \text{ for some } \xi \in \partial^C f(x),\ \zeta \in \partial^c f(\bar{x}) \right\}. \end{aligned}$$

Proof. Let $x \in S$. Then by Theorem 8.15, for some $\xi \in \partial^C f(x)$ and some $\zeta \in \partial^c f(\bar{x})$,

$$\langle \xi, \bar{x} - x \rangle = 0 = \langle \zeta, \bar{x} - x \rangle, \tag{8.45}$$

that is,

$$\langle \xi, x - \bar{x} \rangle = 0 = \langle \zeta, \bar{x} - x \rangle. \tag{8.46}$$

Thus $x \in S_5$, and hence $S \subseteq S_5$. $S_5 \subseteq S_6$ is obvious.

We now prove that $S_6 \subseteq S$. Assume that $x \in S_6$. Then for some $\xi \in \partial^C f(x)$ and some $\zeta \in \partial^c f(\bar{x})$,

$$\langle \zeta, \bar{x} - x \rangle \geq \langle \xi, x - \bar{x} \rangle. \tag{8.47}$$

Suppose that $x \notin S$. Then $f(\bar{x}) < f(x)$. By the generalized pseudoconcavity of f, we have

$$\langle \zeta, x - \bar{x} \rangle > 0,$$

that is,

$$\langle \zeta, \bar{x} - x \rangle < 0.$$

By using (8.47), we have

$$\langle \xi, x - \bar{x} \rangle < 0 \quad \text{or} \quad \langle \xi, \bar{x} - x \rangle > 0.$$

Now using generalized pseudoconvexity of f, it follows that $f(\bar{x}) \geq f(x)$, which is a contradiction to the fact that $f(\bar{x}) < f(x)$. Hence, $x \in S$. □

Appendix A

Set-Valued Maps

Definition A.1. Let X and Y be two nonempty sets. A *set-valued map* or *multivalued map* or *point-to-set map* or *multifunction* $F : X \to 2^Y$ from X to Y is a map that associates with any $x \in X$ a subset $F(x)$ of Y; the set $F(x)$ is called the *image of* x under F. F is called *proper* if there exists at least an element $x \in X$ such that $F(x) \neq \emptyset$. In this case the set $\mathrm{Dom}(F) = \{x \in X : F(x) \neq \emptyset\}$ is called the *domain* of F. Actually, a set-valued map F is characterized by its *graph*, the subset of $X \times Y$ defined by

$$\mathrm{Graph}(F) = \{(x, y) : y \in F(x)\}.$$

Indeed, if A is a nonempty subset of the product space $X \times Y$, then the graph of a set-valued map F is defined by

$$y \in F(x) \quad \text{if and only if} \quad (x, y) \in A.$$

The domain of F is the projection of $\mathrm{Graph}(F)$ on X and the *image* of F, the subset of Y defined by

$$\mathrm{Im}(F) = \bigcup_{x \in X} F(x) = \bigcup_{x \in \mathrm{Dom}(F)} F(x),$$

is the projection of $\mathrm{Graph}(F)$ on Y. A set-valued map F from X to Y is called *strict* if $\mathrm{Dom}(F) = X$, that is, if the image $F(x)$ is nonempty for all $x \in X$. Let K be a nonempty subset of X and F be a strict set-valued map from K to Y. It may be useful to extend it to the set-valued map F_K from X to Y defined by

$$F_K(x) = \begin{cases} F(x), & \text{when } x \in K, \\ \emptyset, & \text{when } x \notin K, \end{cases}$$

whose domain $\mathrm{Dom}(F_K)$ is K.

When F is a set-valued map from X to Y and $K \subseteq X$, we denote by $F_{|_K}$ its restriction to K.

Example A.1 (Inverse Function)**.** If $f : X \to Y$ is a single-valued map, then its inverse f^{-1} can be considered as a set-valued map $F : Y \to 2^X$ defined by

$$F(y) = f^{-1}(y), \quad \text{for all } y \in \mathrm{Im}(f).$$

This set-valued map is strict when f is surjective, and single-valued when f

is injective. This map plays an important role when we study the equation $f(x) = y$ and the behavior of the set of solutions of $f^{-1}(y)$ as y ranges over Y.

Definition A.2. Let $F : X \to 2^Y$ be a set-valued map. For a nonempty subset A of X, we write

$$F(A) = \bigcup_{x \in A} F(x).$$

If $A = \emptyset$, we write $F(\emptyset) = \emptyset$. The set $F(A)$ is called the *image* of A under the set-valued map F.

Theorem A.1. Let $\{A_\alpha\}_{\alpha\in\Lambda}$ be a family of nonempty subsets of X and $F : X \to 2^Y$ be a set-valued map.

(a) If $A_1 \subseteq A_2$, then $F(A_1) \subseteq F(A_2)$;

(b) $F\left(\bigcup_{\alpha\in\Lambda} A_\alpha\right) = \bigcup_{\alpha\in\Lambda} F(A_\alpha)$;

(c) $F\left(\bigcap_{\alpha\in\Lambda} A_\alpha\right) \subset \bigcap_{\alpha\in\Lambda} F(A_\alpha)$.

Definition A.3. Let F_1 and F_2 be two set-valued maps from X to Y.

(a) The *union* of F_1 and F_2 is a set-valued map $(F_1 \cup F_2)$ from X to Y defined by

$$(F_1 \cup F_2)(x) = F_1(x) \cup F_2(x) \quad \text{for all } x \in X.$$

(b) The *intersection* of F_1 and F_2 is a set-valued map $(F_1 \cap F_2)$ from X to Y defined by

$$(F_1 \cap F_2)(x) = F_1(x) \cap F_2(x) \quad \text{for all } x \in X.$$

(c) The *Cartesian product* of F_1 and F_2 is a set-valued map $(F_1 \times F_2)$ from X to $Y \times Y$ defined by

$$(F_1 \times F_2)(x) = F_1(x) \times F_2(x) \quad \text{for all } x \in X.$$

(d) If F_1 is a set-valued map from X to Y and F_2 is another set-valued map from Y to Z, then the *composition product* of F_2 by F_1 is a set-valued map $(F_2 \circ F_1)$ from X to Z defined by

$$(F_2 \circ F_1)(x) = F_2(F_1(x)).$$

Theorem A.2. If F_1 and F_2 are two set-valued maps from X to Y and A is a nonempty subset of X, then

(a) $(F_1 \cup F_2)(A) = F_1(A) \cup F_2(A)$;

(b) $(F_1 \cap F_2)(A) \subseteq F_1(A) \cap F_2(A)$;

(c) $(F_1 \times F_2)(A) \subseteq F_1(A) \times F_2(A)$;

(d) $(F_2 \circ F_1)(A) = F_2(F_1(A))$.

Definition A.4. If F is a set-valued map from X to Y, then the *inverse* F^{-1} of F is defined by

$$F^{-1}(y) = \{x \in X : y \in F(x)\}, \quad \text{for all } y \in Y.$$

Further, let B be a subset of Y. The *upper inverse image* $F^{-1}(B)$ and *lower inverse image* $F_+^{-1}(B)$ of B under F are defined by

$$F^{-1}(B) = \{x \in X : F(x) \cap B \neq \emptyset\}$$

and

$$F_+^{-1}(B) = \{x \in X : F(x) \subseteq B\}.$$

We also write $F^{-1}(\emptyset) = \emptyset$ and $F_+^{-1}(\emptyset) = \emptyset$. It is clear that the definition of inverse of F^{-1} that $(F^{-1})^{-1} = F$ and that $y \in F(x)$ if and only if $x \in F^{-1}(y)$.

We have the following relations between domain, graphs, and images of F and F^{-1}:

$$\mathrm{Dom}(F^{-1}) = \mathrm{Im}(F), \qquad Im(F^{-1}) = \mathrm{Dom}(F), \quad \text{and}$$

$$\mathrm{Graph}(F^{-1}) = \{(y, x) \in Y \times X : (x, y) \in \mathrm{Graph}(F)\}.$$

Theorem A.3. Let $\{B_\alpha\}_{\alpha \in \Lambda}$ be a family of nonempty subsets of Y, $A \subseteq X$, and $B \subseteq Y$. Let $F : X \to 2^Y$ be a set-valued map.

(a) If $B_1 \subseteq B_2$, then $F^{-1}(B_1) \subseteq F^{-1}(B_2)$;

(b) $A \subset F_+^{-1}(F(A))$;

(c) $B \subset F(F_+^{-1}(B))$;

(d) $X \setminus F_+^{-1}(B) = F_+^{-1}(Y \setminus B)$;

(e) $F_+^{-1}\left(\bigcup_{\alpha \in \Lambda} B_\alpha\right) \subset \bigcup_{\alpha \in \Lambda} F_+^{-1}(B_\alpha)$;

(f) $F_+^{-1}\left(\bigcap_{\alpha \in \Lambda} B_\alpha\right) = \bigcap_{\alpha \in \Lambda} F_+^{-1}(B_\alpha)$;

(g) $F^{-1}(F(A)) \subset A$;

(h) $F(F^{-1}(B)) \subset B \cap F(X)$;

(i) $X \setminus F^{-1}(B) = F_+^{-1}(Y \setminus B)$;

(j) $F^{-1}\left(\bigcap_{\alpha\in\Lambda} B_\alpha\right) \subset \bigcap_{\alpha\in\Lambda} F^{-1}(B_\alpha)$;

(k) $F^{-1}\left(\bigcup_{\alpha\in\Lambda} B_\alpha\right) = \bigcup_{\alpha\in\Lambda} F^{-1}(B_\alpha)$.

Theorem A.4. Let $F_1, F_2 : X \to 2^Y$ be set-valued maps such that $(F_1 \cap F_2)(x) \neq \emptyset$ for all $x \in X$ and let $B \subseteq Y$. Then

(a) $(F_1 \cup F_2)^{-1}(B) = F_1^{-1}(B) \cup F_2^{-1}(B)$;

(b) $(F_1 \cap F_2)^{-1}(B) \subset F_1^{-1}(B) \cap F_2^{-1}(B)$;

(c) $(F_1 \cup F_2)_+^{-1}(B) = F_{1+}^{-1}(B) \cap F_{2+}^{-1}(B)$;

(d) $(F_1 \cap F_2)_+^{-1}(B) \subset F_{1+}^{-1}(B) \cup F_{2+}^{-1}(B)$.

Let K be a subset of $\mathbb{R}^n$. Then the distance from a point $x \in \mathbb{R}^n$ to the set K is defined by

$$d(x, K) = \inf_{y\in K} \|x - y\|,$$

where $\|x\| = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$ for all $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. The *ball of radius* $r > 0$ *around* K is denoted by

$$\mathbb{B}_r[K] = \{x \in X : d(x, K) \leq r\}.$$

The sphere $\mathbb{B}_r[K]$ is a neighborhood of K. If K is compact, then $\mathbb{B}_r[K]$ is closed.

Indeed, if K is compact then it is totally bounded. Therefore, the set $\mathbb{B}_r[K]$ can be expressed as the union of a finite number of closed sets. Thus, it is closed.

When K is compact, each neighborhood of K contains such a sphere around K.

If the images of a set-valued map F are closed, bounded, compact, and so on, we say that F is *closed valued, bounded valued, compact valued*, and so on.

Definition A.5. A set-valued map $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is called *upper semicontinuous* at $x_0 \in \mathrm{Dom}(F)$ if for any neighborhood U of $F(x_0)$, there exists $\delta > 0$ such that $F(x) \subseteq U$ for all $x \in \mathbb{B}_\delta(x_0)$.

It is said to be *upper semicontinuous* on $\mathbb{R}^n$ if it is upper semicontinuous at every point of $\mathrm{Dom}(F)$.

When $F(x)$ is compact for all $x \in \mathbb{R}^n$, F is upper semicontinuous at x_0 if and only if for any $\epsilon > 0$, there exists $\delta > 0$ such that $F(x) \subseteq \mathbb{B}_\varepsilon(F(x))$ for all $x \in \mathbb{B}_\delta(x_0)$.

We observe that this definition is a natural adaptation of the definition of a continuous single-valued map. Why then do we use the adjective upper semicontinuous instead of continuous? One of the reason is that the celebrated characterization of continuous maps—"a single-valued map f is continuous at x if and only if it maps sequences converging to x to sequences converging to $f(x)$"—does not hold true any longer in the set-valued case. Indeed, the set-valued version of this characterization leads to the following definition.

Definition A.6. A set-valued map $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is called *lower semicontinuous* at $x_0 \in \mathrm{Dom}(F)$ if for any $y \in F(x_0)$ and for any sequence of elements $x_m \in \mathrm{Dom}(F)$ converging to x_0 there exists a sequence of elements $y_m \in F(x_m)$ converging to y.

It is said to be *lower semicontinuous* on $\mathbb{R}^n$ if it is lower semicontinuous at every point $x \in \mathrm{Dom}(F)$.

The above definition of lower semicontinuity could be interpreted as follows:

F is *lower semicontinuous* at x_0 if for any $y_0 \in F(x_0)$ and any neighborhood $V(y_0)$ of y_0 there exists $\delta > 0$ such that $F(x) \cap V(y_0) \neq \emptyset$ for all $x \in \mathbb{B}_\delta(x_0)$.

Actually, as in the single-valued case, this definition is equivalent to the following definition.

For any open subset $U \subseteq Y$ such that $U \cap F(x_0) \neq \emptyset$, there exists $\delta > 0$ such that $F(x) \cap U \neq \emptyset$ for all $x \in \mathbb{B}_\delta(x_0)$.

When $F(x)$ is compact, F is lower semicontinuous at x_0 if and only if for any $\varepsilon > 0$, there exists $\delta > 0$ such that $F(x) \cap S_\varepsilon(F(x_0)) \neq \emptyset$ for all $x \in \mathbb{B}_\delta(x_0)$.

Definition A.7. A set-valued map $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is said to be *continuous* at $x_0 \in \mathbb{R}^n$ if it is both lower semicontinuous as well as upper upper semicontinuous at x_0. It is said to be *continuous* on $\mathbb{R}^n$ if it is continuous at every point $x \in \mathbb{R}^n$.

It is not easy for the readers to prove the set-valued maps are upper semicontinuous or lower semicontinuous by using the definitions of these continuities. So, we have the following characterizations for these definitions.

Theorem A.5. A set-valued map $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is lower semicontinuous if and only if $F^{-1}(G)$ is open for every open subset G of $\mathbb{R}^\ell$.

Corollary A.1. A set-valued map $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is lower semicontinuous if and only if $F_+^{-1}(H)$ is closed for every closed subset H of $\mathbb{R}^\ell$.

Theorem A.6. Let $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ be a set-valued map such that $F(x)$ is compact for each $x \in \mathbb{R}^n$. Then, F is upper semicontinuous if and only if for each open subset G of $\mathbb{R}^\ell$ the set

$$F_+^{-1}(G) = \{x \in \mathbb{R}^n : F(x) \subseteq G\}$$

is open.

Corollary A.2. Let $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ be a set-valued map such that $F(x)$ is compact for each $x \in \mathbb{R}^n$. Then, F is upper semicontinuous if and only if for each closed subset H of $\mathbb{R}^\ell$ the set $F^{-1}(H)$ is closed.

Theorem A.7. Let $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ be an upper semicontinuous set-valued map such that for all $x \in \mathbb{R}^n$, $F(x)$ is compact. Then the image $F(K)$ of a compact subset K of $\mathbb{R}^n$ is compact.

Definition A.8. A set-valued map $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is said to be *closed* if whenever $x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}^\ell$, $y_0 \notin F(x_0)$, there exist two neighborhoods $U(x_0)$ of x_0 and $V(y_0)$ of y_0 such that $F(x) \cap V(y_0) = \emptyset$ for all $x \in U(x_0)$.

We state some useful characterizations of closed set-valued maps.

(a) F is closed if and only if its graph Graph(F) is a closed set.

(b) F is closed if and only if for any sequences $\{x_m\}$ and $\{y_m\}$ such that $x_m \to x_0$, $y_m \to y_0$ and $y_m \in F(x_m)$ for all m, we have $y_0 \in F(x_0)$.

Theorem A.8. Let K be a nonempty compact subset of $\mathbb{R}^n$. If the set-valued map $F : K \to 2^{\mathbb{R}^\ell}$ is closed, then the set $F(K)$ is closed.

Theorem A.9. If $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is an upper semicontinuous set-valued map with compact values, then it is closed.

Remark A.1. In general if $f : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is a single-valued continuous map from $\mathbb{R}^n$ onto $\mathbb{R}^\ell$, then the inverse map $f^{-1} : \mathbb{R}^\ell \to 2^{\mathbb{R}^n}$ is a set-valued map and it has a closed graph. But it is not necessarily upper semicontinuous.

Theorem A.10. Let D be nonempty compact subset of $\mathbb{R}^\ell$. If $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is a closed set-valued map, then F is upper semicontinuous.

Theorem A.11. [31, Theorem 1, pp. 115] If $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is a lower semicontinuous real-valued function and $F : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ is a lower semicontinuous set-valued map with nonempty values, then the real-valued functions M and N defined by

$$M(x) = \sup\{f(x, y) : y \in F(x)\} \quad \text{and} \quad N(x) = \inf\{f(x, y) : y \in F(x)\},$$

respectively, are lower semicontinuous.

Theorem A.12. [31, Theorem 2, pp. 116] If $f : \mathbb{R}^n \times \mathbb{R}^\ell \to \mathbb{R}$ is a upper semicontinuous real-valued function and $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is a upper semicontinuous set-valued map with nonempty values, then the real-valued functions M and N defined by

$$M(x) = \max\{f(x, y) : y \in F(x)\} \quad \text{and} \quad N(x) = \min\{f(x, y) : y \in F(x)\},$$

respectively, are upper semicontinuous.

Theorem A.13. [31, Theorem 3, pp. 116] If $f : \mathbb{R}^\ell \to \mathbb{R}$ is a upper semicontinuous real-valued function and $F : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ is a continuous set-valued map with nonempty values, then the real-valued functions M and N defined by

$$M(x) = \max\{f(y) : y \in F(x)\} \quad \text{and} \quad N(x) = \min\{f(y) : y \in F(x)\},$$

respectively, are continuous. Furthermore, the set-valued maps $\Phi, \Psi : \mathbb{R}^n \to 2^{\mathbb{R}^\ell}$ defined by

$$\Phi(x) = \{y \in \mathbb{R}^\ell : y \in F(x),\ f(y) = M(x)\}$$

and

$$\Psi(x) = \{y \in \mathbb{R}^\ell : y \in F(x),\ f(y) = N(x)\},$$

respectively, are upper semicontinuous set-valued maps.

Appendix B

Elements of Nonlinear Analysis

Definition B.1. A family $\mathcal{F}$ of functions from $\mathbb{R}^n$ to $\mathbb{R}^\ell$ is said to be *equicontinuous* at a point $x \in \mathbb{R}^n$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|f(x) - f(y)\| < \varepsilon$ for all $f \in \mathcal{F}$ and all y such that $\|x - y\| < \delta$. The family is *equicontinuous* if it is equicontinuous at each point of $\mathbb{R}^n$.

Definition B.2. A collection $\mathcal{C} = \{C_1, C_2, \ldots\}$ of subsets of $K \subseteq \mathbb{R}^n$ is said to have the *finite intersection property* if every finite subcollection of $\mathcal{C}$ has nonempty intersection, that is, for every finite collection $\{C_1, C_2, \ldots, C_m\}$ of $\mathcal{C}$ we have $\bigcap_{i=1}^m C_i \neq \emptyset$.

Theorem B.1. A subset K of $\mathbb{R}^n$ is compact if and only if every collection of closed sets in K having the finite intersection property has a nonempty intersection.

Theorem B.2 (Brouwer's Fixed Point Theorem)**.** Let $\mathbb{B}$ be a closed ball in $\mathbb{R}^n$ and $T : \mathbb{B} \to \mathbb{B}$ be continuous. Then T admits at least one fixed point.

Definition B.3. Let K be a nonempty convex subset of $\mathbb{R}^n$. A set-valued map $P : K \to 2^K$ is said to be a *KKM-map* if for every finite subset $\{x_1, x_2, \ldots, x_m\}$ of K,

$$\text{co}\{x_1, x_2, \ldots, x_m\} \subseteq \bigcup_{i=1}^m P(x_i),$$

where $\text{co}\{x_1, x_2, \ldots, x_m\}$ denotes the convex hull of $\{x_1, x_2, \ldots, x_m\}$.

The following Fan-KKM theorem and the Browder-type fixed point theorem for set-valued maps will be the key tools to establish existence results for solutions of nonsmooth vector variational-like inequalities.

Theorem B.3 (Fan-KKM Theorem)**.** [71] Let K be a nonempty convex subset of $\mathbb{R}^n$ and $P : K \to 2^K$ be a KKM-map such that $P(x)$ is closed for all $x \in K$, and $P(x)$ is compact for at least one $x \in K$. Then, $\bigcap_{x \in K} P(x) \neq \emptyset$.

Theorem B.4. [15] Let K be a nonempty convex subset of $\mathbb{R}^n$ and $P, Q : K \to 2^K$ be two set-valued maps. Assume that the following conditions hold:

(i) For each $x \in K$, $\text{co}P(x) \subseteq Q(x)$ and $P(x)$ is nonempty;

(ii) For each $y \in K$, $P^{-1}(y) = \{x \in K : y \in P(x)\}$ is open in K;

(iii) If K is not compact, assume that there exist a nonempty compact convex subset B of K and a nonempty compact subset D of K such that for each $x \in K \setminus D$ there exists $\tilde{y} \in B$ such that $\tilde{y} \in P(x)$.

Then, there exists $\bar{x} \in K$ such that $\bar{x} \in Q(\bar{x})$.

Theorem B.5. Let K be a nonempty convex subset of $\mathbb{R}^n$ and $P : K \to 2^K$ a set-valued map. Assume that the following conditions hold:

(i) for all $x \in K$, $P(x)$ is convex;

(ii) for each finite subset A of K and for all $y \in \mathrm{co}(A)$, $P^{-1}(y) \cap \mathrm{co}(A)$ is open in $\mathrm{co}(A)$;

(iii) for each finite subset A of K and all $x, y \in \mathrm{co}(A)$ and every sequence $\{x_m\}$ in K converging to x such that $\lambda y + (1-\lambda)x \notin P(x_m)$ for all $m \in \mathbb{N}$ and all $\lambda \in [0,1]$, we have $y \notin P(x)$;

(iv) there exist a nonempty compact subset D of K and an element $\tilde{y} \in D$ such that $\tilde{y} \in P(x)$ for all $x \in K \setminus D$;

(v) for all $x \in D$, $P(x)$ is nonempty.

Then, there exists $\hat{x} \in K$ such that $\hat{x} \in P(\hat{x})$.

Theorem B.6 (Kakutani)**.** [119] Let K be a nonempty compact convex subset of $\mathbb{R}^n$ and $P : K \to 2^K$ be a set-valued map such that for each $x \in K$, $P(x)$ is nonempty, compact, and convex. Then P has a fixed point, that is, there exists $\bar{x} \in K$ such that $\bar{x} \in P(\bar{x})$.

Theorem B.7 (Kneser)**.** [131] Let K be a nonempty convex subset of $\mathbb{R}^n$ and D be a nonempty compact convex subset of $\mathbb{R}^m$. Suppose that $f : K \times D \to \mathbb{R}$ is lower semicontinuous in the second argument and concave in the first argument. Then

$$\min_{y \in D} \sup_{x \in K} f(x,y) = \sup_{x \in K} \min_{y \in Y} f(x,y).$$

Bibliography

[1] Al-Homidan, S., Ansari, Q.H.: Generalized Minty vector variational-like inequalities and vector optimization problems. J. Optim. Theory Appl. **144**, 1–11 (2010)

[2] Al-Homidan, S., Ansari, Q.H.: Relations between generalized vector variational-like inequalities and vector optimization. Taiwanese J. Math **16**, 987–998 (2012)

[3] Ansari, Q.H.: Metric Spaces, Including Fixed Point Theory and Set-Valued Maps. Narosa Publishing House, New Delhi (2010)

[4] Ansari, Q.H. (Editor): Topics in Nonlinear Analysis and Optimization. World Education, Delhi (2012)

[5] Ansari, Q.H., Khan, Z.: Relatively B-pseudomonotone variational inequalities over product of sets. J. Inequal. Pure Appl. Math. **4**, Article 6 (2003)

[6] Ansari, Q.H., Konnov, I.V., Yao, J.C.: Characterizations of solutions for vector equilibrium problems. J. Optim. Theory Appl. **113**, 435–447 (2002)

[7] Ansari, Q.H., Lee, G.M.: Nonsmooth vector optimization problems and Minty vector variational inequalities. J. Optim. Theory Appl. **145**, 1–16 (2010)

[8] Ansari, Q.H., Lai, T.C., Yao, J.C.: On the equivalence of extended generalized complementarity and generalized least-element problems. J. Optim. Theory Appl. **102**, 277–288 (1999)

[9] Ansari, Q.H., Lin, Y.C., Yao, J.C.: General KKM theorem with applications to minimax and variational inequalities. J. Optim. Theory Appl. **104**, 41–57 (2000)

[10] Ansari, Q.H., Rezaei, M.: Existence results for Stampacchia and Minty type vector variational inequalities. Optimization **59**, 1053–1065 (2010)

[11] Ansari, Q.H., Rezaei, M.: Invariant pseudolinearity with applications. J. Optim. Theory Appl. **153**, 587–601 (2012)

[12] Ansari, Q.H., Rezaei, M.: Generalized pseudolinearity. Optim. Lett. **6**, 241–251 (2012)

[13] Ansari, Q.H., Rezaei, M.: Generalized vector variational-like inequalities and vector optimization in Asplund spaces. Optimization (2012)

[14] Ansari, Q.H., Rezaei, M., Zafarani, J.: Generalized vector variational-like inequalities and vector optimization. J. Global Optim. **53**, 271–284 (2012)

[15] Ansari, Q.H., Yao, J.C.: A fixed point theorem and its applications to a system of variational inequalities. Bull. Austral. Math. Soc. **59**, 433–442 (1999)

[16] Ansari, Q.H., Yao, J.C.: Generalized variational-like inequalities and a gap function. Bull. Austral. Math. Soc. **59**, 33–44 (1999)

[17] Ansari, Q.H., Yao, J.C.: Iterative scheme for solving mixed variational-like inequalities. J. Optim. Theory Appl. **108**, 527–541 (2001)

[18] Ansari, Q.H., Yao, J.-C.: Some equivalences among nonlinear complementarity problems, least-element problems and variational inequality problems in ordered spaces. In: Topics in Nonconvex Optimization, Theory and Applications, S.K. Mishra (ed.), Springer-Verlag, New York, Berlin, pp. 1–25 (2011)

[19] Arrow, K.J., Enthoven, A.C.: Quasi-concave programming. Econometrica **29**, 779–800 (1961)

[20] Aubin, J.-P.: Optima and Equilibria. Springer-Verlag, Berlin, New York (1993)

[21] Auchmuty, G.: Variational principles for variational inequalities. Numer. Funct. Anal. Optim. **10**, 863–874 (1989)

[22] Auslender, A.: Problémes de Minimax via l'Analyse Convexe et les Inégalitiés Variationnelles: Théorie et Algorithmes. Springer-Verlag, Berlin (1972)

[23] Auslender, A.: Optimisation: Méthodes Numériques. Masson, Paris (1976)

[24] Auslender, A., Teboulle, M.: Asymptotic Cones and Functions in Optimization and Variational Inequalities. Springer-Verlag, New York (2003)

[25] Aussel, D., Luc, D.T.: Existence conditions in general quasimonotone variational inequalities. Bull. Austral. Math. Soc. **71**, 285–303 (2005)

[26] Avriel, M.: Nonlinear Programming: Analysis and Methods. Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1976)

[27] Avriel, M., Diewert, W.E., Schaible, S., Ziemba W.T: Introduction to concave and generalized concave functions. In: Generalized Concavity in Optimization and Economics, S. Schaible, W.T. Ziemba (Eds.), pp. 21–50 (1981)

[28] Avriel, M., Diewert, W.E., Schaible, S., Zang, I.: Generalized Concavity. Plenum Press, New York, London (1988)

[29] Baiocchi, C., Capelo, A.: Variational and Quasivariational Inequalities: Applications to Free Boundary Problems. John Wiley & Sons, Chichester, New York, Brisbane, Tokyo (1984)

[30] Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming, Theory and Algorithms (Third Edition). John Wiley & Sons, Inc., Hoboken, New Jersey (2006)

[31] Berge, C.: Topological Spaces, Including a Treatment of Multi-Valued Functions, Vector Spaces and Convexity. Oliver & Boyd, Edinburgh and London (1963)

[32] Bertsekas, D.P., Nedić, A., Ozdaglar, E.: Convex Analysis and Optimization. Athena Scientific, Belmont, Massachusetts (2003)

[33] Bianchi. M.: Generalized quasimonotonicity and strong pseudomonotonicity of bifunctions. Optimization **1**, 1–10 (1996)

[34] Bianchi, M., Schaible, S.: An extension of pseudolinear functions and variational inequality problems. J. Optim. Theory Appl. **104**, 59–71 (2000)

[35] Blum, E., Oettli, W.: Variational principles for equilibrium problems. In: Parametric Optimization and Related Topics III, J. Guddat, H. Th. Jongen, B. Kummer, F. Nožička (eds.), Peter Lang Verlag, Frankfurt am Main, pp. 79–88 (1993)

[36] Borwein, J.M., Lewis, A.S.: Convex Analysis and Nonlinear Optimization: Theory and Examples. Canadian Mathematical Society, Springer, New York (2000)

[37] Brézis, H.: Équations et inéquations non linéaires dans les espaces vectoriels en dualité. Ann. Inst. Fourier (Grenoble) **18**, 115–175 (1968)

[38] Brézis, H.: Operateurs Maximaux Monotones. North-Holland Publication, Amsterdam (1973)

[39] Browder, F.E.: Nonlinear monotone operators and convex sets in Banach spaces. Bull. Amer. Math. Soc. **71**, 780–785 (1965)

[40] Browder, F.E.: Existence and approximation of solutions of nonlinear variational inequalities. Proc. Natl. Acad. Sci. USA **56**, 1080–1086 (1966)

[41] Carl, S., Le, V.K., Montreanu, D.: Nonsmooth Variational Problems and Their Inequalities: Comparison Principles and Applications. Springer-Verlag, New York (2007)

[42] Chen, G.Y., Craven, B.D.: A vector variational inequality and optimization over an efficient set. ZOR-Math. Meth. Oper. Res. **34**, 1–12 (1990)

[43] Chen, G.Y., Goh, C.J., Yang, X.Q.: On gap functions and duality of variational inequality problems. J. Math. Anal. Appl. **214**, 658–673 (1997)

[44] Chew, K.L., Choo, E.U.: Pseudolinearity and efficiency. Math. Prog. **28**, 226–239 (1984)

[45] Chowdhury, M.S.R., Tan, K.-K.: Generalization of Ky Fan's minimax inequality with applications to generalized variational inequalities for pseudo-monotone operators and fixed point theorems. J. Math. Anal. Appl. **204**, 910–929 (1996)

[46] Chowdhury, M.S.R., Tan, K.-K.: Generalized variational inequalities for quasi-monotone operators and applications. Bull. Polish Acad. Sci. Math. **45**, 25–54 (1997)

[47] Clarke, F.H.: Generalized gradients and applications. Trans. Amer. Math. Soc. **205**, 247–262 (1975)

[48] Clarke, F.H.: Optimization and Nonsmooth Analysis. John Wiley & Sons, New York, Chichester, Brisbane (1983), and SIAM, Philadelphia (1990)

[49] Cohen, G.: Optimization by decomposition and coordination: A unified approach. IEEE Trans. Automat. Control **AC-23**, 222–232 (1978)

[50] Cohen, G.: Auxiliary problem principle and decomposition of optimization problems. J. Optim. Theory Appl., **32**, 277–305 (1980)

[51] Cohen, G.: Auxiliary problem principle extended to variational inequalities. J. Optim. Theory Appl. **59**, 325–333 (1988)

[52] Cohen, G., Zhu, D. L.: Decomposition coordination methods in large-scale optimization problems: The nondiffferentiable case and the use of augmented Lagrangians. In: Advances in Large-Scale Systems, Theory and Applications, J.B. Cruz (ed.), JAI Press, Greenwich, Connecticut, Vol. 1, pp. 203–266 (1984)

[53] Crespi, G.P., Ginchev, I., Rocca, M.: Minty variational inequalities, increase-along-rays property and optimization. J. Optim. Theory Appl. **123**, 479–496 (2004)

[54] Crespi, G.P., Ginchev, I., Rocca, M.: Existence of solutions and star-shapedness in Minty variational inequalities. J. Global Optim. **32**, 485–494 (2005)

[55] Crespi, G.P., Ginchev, I., Rocca, M.: Some remarks on the Minty vector variational principle. J. Math. Anal. Appl. **345**, 165–175 (2008)

[56] Crouzeix J.-P.: Conditions for convexity of quasiconvex functions. Math. Oper. Res. **5**, 120–125 (1980)

[57] Crouzeix J.-P.: Pseudomonotone variational inequality problems: existence of solutions. Math. Prog. **78**, 305–314 (1997)

[58] Crouzeix, J.-P., Martinez-Legaz, J.E., Volle, M.: Generalized Convexity, Generalized Monotonicity: Recent Results. Kluwer Academic Publishers, Dordrecht, Boston, London (1998)

[59] Dafermos, S.: Traffic equilibrium and variational inequalities. Transp. Sci. **14**, 42–54 (1980)

[60] Daniilidis, A., Hadjisavvas, N.: On the subdifferentials of quasiconvex and pseudoconvex functions and cyclic monotonicity. J. Math. Anal. Appl. **237**, 30–42 (1999)

[61] Daniilidis, A., Hadjisavvas, N.: Characterization of nonsmooth semistrictly quasiconvex and strictly quasiconvex functions. J. Optim. Theory Appl. **102**, 525–536 (1999)

[62] Daniilidis, A., Hadjisavvas, N.: Coercivity conditions and variational inequalities. Math. Prog. **86**, 433–438 (1999)

[63] Danskin, J.M.: The theory of max-min with applications. SIAM J. Appl. Math. **14**, 641–664 (1966)

[64] Ding, X.P.: General algorithm of solutions for nonlinear variational inequalities in Banach spaces. Comput. Math. Appl. **34**, 131–137 (1997)

[65] Ding, X.P., Tan, K.-K.: Generalized variational inequalities and generalized quasi-variational inequalities. J. Math. Anal. Appl. **148**, 497–508 (1990)

[66] Ding, X.P., Tarafdar, E.: Monotone generalized variational inequalities and generalized complementarity problems. J. Optim. Theory Appl. **88**, 107–122 (1996)

[67] Dinh, N., Jeyakumar, V., Lee, G.M.: Lagrange multiplier characterizations of solution sets of constrained pseudolinear optimization problems. Applied Mathematics Research Report AMR05/24, University of New South Wales, Sydney, Australia (2005)

[68] de Finetti, B.: Sulle stratificazioni convesse. Ann. Mat. Pura Appl. **30**, 173–183 (1949)

[69] Diewert, W.E.: Alternative characterizations of six kinds of quasiconcavity in the nondifferentiable case with applications to nonsmooth programming. In: Generalized Concavity in Optimization and Economics, S., Schaible, W.T., Ziemba, (eds.), Academic Press, New York, pp. 51–93 (1981)

[70] Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I and II. Springer-Verlag, New York, Berlin, Heidelberg (2003)

[71] Fan, K.: A generalization of Tychonoffs fixed point theorem. Math. Ann. **142**, 305-310 (1961)

[72] Fan, K.: Application of a theorem concerning sets with convex sections. Math. Ann. **163**, 189-203 (1966)

[73] Fan, K.: A minimax inequality and applications. In: Inequalities III, O. Shisha (ed.), Academic Press, New York, pp. 103-113 (1972)

[74] Fan, L., Liu, S., Gao, S.: Generalized monotonicity and convexity of non-differentiable functions. J. Math. Anal. Appl. **279**, 276–289 (2003)

[75] Fang, S.C., Peterson, E.L.: Generalized variational inequalities. J. Optim. Theory Appl. **38**, 363–383 (1982)

[76] Fenchel, W.: On conjugate convex functions. Canad. J. Math. 1, 73–77 (1949)

[77] Fenchel, W.: Convex Cones, Sets and Functions. Mimeographed Lecture Notes, Princeton University, Princeton, New Jersey (1953)

[78] Fichera, G.: Problemi elettrostatici con vincoli unilaterali; il problema di Signorini con ambigue condizioni al contorno. Atti Acad. Naz. Lincei. Mem. Cl. Sci. Fis. Mat. Nat. Sez. I **7**, 91–140 (1964)

[79] Fiedler, M., Pták, V.: Some generalizations of positive definiteness and monotonicity. Numer. Math. **9**, 163–172 (1966)

[80] Fitzpatrick, P.M.: Advanced Calculus: A Course in Mathematical Analysis. PWS Publishing Company, Boston (1996)

[81] Fukushima, M.: Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problem. Math. Prog. **53**, 99–110 (1992)

[82] Giannessi, F.: Theorems of alternatives for multifunctions with applications to optimization—General results. Publication No. **127**, Group of Optimization and Operations Research, University of Pisa, Italy (1985)

[83] Giannessi, F.: Semidifferenatiable functions and necessary optimality conditions. J. Optim. Theory Appl. **60**, 191–241 (1989)

[84] Giannessi, F.: On some connections among variational inequalities, combinatorial and continuous optimization. Ann. Oper. Res. **58**, 181–200 (1995)

[85] Giannessi, F.: On Minty variational principle. In: New Trends in Mathematical Programming, F., Giannessi, S., Komlòsi, T. Tapcsáck, (eds.), Kluwer Academic Publisher, Dordrecht, Holland, pp. 93–99 (1998)

[86] Giorgi, G., Guerraggio, A., Thierfelder, J.: Mathematics of Optimization: Smooth and Nonsmooth Case. Elsevier B.V., Amsterdam (2004)

[87] Giorgi, G., Komlósi, S.: Dini derivatives in optimization—Part I. Riv. Mat. Sci. Econom. Social **15**(1), 3–30 (1992)

[88] Giorgi, G., Komlósi, S.: Dini derivatives in optimization—Part II. Riv. Mat. Sci. Econom. Social **15**(2), 3–24 (1992)

[89] Giorgi, G., Komlósi, S.: Dini derivatives in optimization—Part III, Riv. Mat. Sci. Econom. Social **18**(1), 47–63 (1995)

[90] Giles, J.R.: Convex Analysis with Application in the Differentiation of Convex Functions. Pitman Advanced Publishing Program, Boston, London, Melbourne (1982)

[91] Glover, B.M.: Generalized convexity in nondifferentiable programming. Bull. Austral. Math. Soc. **30**, 193–218 (1984)

[92] Glowinski, R., Lions, J.L., Trémolières, R.: Numerical Analysis of Variational Inequalities. North-Holland, Amsterdam (1981)

[93] Goh, C.J., Yang, X.Q.: Duality in Optimization and Variational Inequalities. Taylor & Francis, London and New York (2002)

[94] Greenberg, H.J., Pierskalla, W.P.: A review of quasi-convex functions. Oper. Res. **19**, 1553–1570 (1971)

[95] Hadjisavvas, N.: Continuity and maximality properties of pseudomonotone operators. J. Convex Anal. **10**, 465–475 (2003)

[96] Hadjisavvas, N., Komlósi, S., Schaible, S.: Handbook of Generalized Convexity and Generalized Monotonicity. Springer, New York (2005)

[97] Hadjisavvas, N., Schaible, S.: Quasimonotone variational inequalities in Banach spaces. J. Optim. Theory Appl. **90**, 95–111 (1996)

[98] Hadjisavvas, N., Schaible, S., Wong, N.-C.: Pseudomonotone operators: A survey of the theory and its applications. J. Optim. Theory Appl. **152**, 1–20 (2012)

[99] Harker, P.T., Pang, J.-S.: Finite dimensional variational inequalities and nonlinear complementarity problems: A survey of theory, algorithms and applications. Math. Prog. **48**, 161–220 (1990)

[100] Hartmann, P., Stampacchia, G.: On some nonlinear elliptic differential functional equations. Acta Math. **115**, 271–310 (1966)

[101] Hassouni, A.: Sous-differentiels des fonctions quasiconvexes, Thése de troisiéme cycle, Université Paul Sabatier, Toulouse (1983)

[102] Hearn, D.W.: The gap function of a convex program. Oper. Res. Lett., **1**, 67–71 (1982)

[103] Hiriart-Urruty, J.-B.: On optimality conditions in non-differentiable programming. Math. Prog. **14**, 73–86 (1978)

[104] Hiriart-Urruty, J.-B.: New concepts in nondifferentiable programming. Bull. Soc. Math. France, Mém. **60**, 57–85 (1979)

[105] Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms I. Springer-Verlag, Berlin, Heidelberg (1993)

[106] Hiriart-Urruty, J.-B., Lemaréchal, C.: Fundamentals of Convex Analysis. Springer-Verlag, Berlin, Heidelberg, New York (2001)

[107] Ioffe, A.D.: Approximate subdifferentials and applications, I: The finite dimensional theory. Trans. Amer. Math. Soc. **281**, 389–416 (1984)

[108] Isac, G.: Complementarity Problems. Lecture Notes in Mathematics, Vol. 1528, Springer-Verlag, Berlin, Heidelberg (1992)

[109] Isac, G.: Topological Methods in Complementarity Theory. Kluwer Academic Publishers, Dordrecht, Boston, London (2000)

[110] Jahn, J.: Vector Optimization: Theory, Applications and Extensions. Springer-Verlag, Heidelberg, Dordrecht, London, New York (2011)

[111] Jensen, J.L.W.V.: Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Math. **30**, 175–193 (1906)

[112] Jeyakumar, V., Lee, G.M., Dinh, N.: Lagrange multiplier conditions characterizing the optimal solutions sets of cone-constrained convex programs. J. Optim. Theory Appl. **123**, 83–103 (2004)

[113] Jeyakumar, V., Luc, D.T.: Nonsmooth Vector Functions and Continuous Optimization. Springer-Verlag, New York (2008)

[114] Jeyakumar, V., Yang, X.Q.: Convex composite multi-objective nonsmooth programming. Math. Prog. Ser. A **59**, 325–343 (1993)

[115] Jeyakumar, V., Yang, X.Q.: On characterizing the solution sets of pseudolinear programs. J. Optim. Theory Appl. **87**, 747–755 (1995)

[116] Jofré, A., Penot, J.-P.: Comparing new notions of tangent cones. J. Lond. Math. Soc. **40**, 280–290 (1989)

[117] John, F.: Extremum problems with inequalities as subsidiary conditions. Studies and Essays Presented to R. Courant on His 60th Birthday, January 8, 1948. Interscience Publishers, Inc., New York, pp. 187–204 (1948)

[118] John, R.: A note on Minty variational inequalities and generalized monotonicity, In: Generalized Convexity and Generalized Monotonicity (Karlovarsi, 1999), Lecture Notes in Economics and Mathematical Systems, **502**, Springer, Berlin, pp. 240–246 (2001)

[119] Kakutani, S.: A generalization of Brouwer's fixed point theorem. Duke Math. J. **8**, 457–459 (1941)

[120] Karamardian, S.: The nonlinear complementarity problems with applications, Part 1. J. Optim. Theory Appl. **4**, 87–98 (1969)

[121] Karamardian, S.: The nonlinear complementarity problems with applications, Part 2. J. Optim. Theory Appl. **4**, 167–181 (1969)

[122] Karamardian, S.: Generalized complementarity problem. J. Optim. Theory Appl. **8**, 161–168 (1971)

[123] Karamardian, S.: The complementarity problem. Math. Prog. **2**, 107–129 (1972)

[124] Karamardian, S.: Complementarity over cones with monotone and pseudomonotone maps. J. Optim. Theory Appl. **18**, 445–454 (1976)

[125] Karamardian, S., Schaible, S.: Seven kinds of monotone maps. J. Optim. Theory Appl. **66**, 37–46 (1990)

[126] Karamardian, S., Schaible, S., Crouzeix, J.P.: Characterizations of generalized monotone maps. J. Optim. Theory Appl. **76**, 399–413 (1993)

[127] Karush, W.: Minima of Functions of Several Variables with Inequalities as Side Conditions. M.S. Thesis, Department of Mathematics, University of Chicago (1939)

[128] Kim, W.K., Tan, K.-K.: A variational inequality in non-compact sets and its applications. Bull. Austral. Math. Soc. **46**, 139–148 (1992)

[129] Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Academic Press, New York (1980)

[130] Knaster, B., Kuratowski, C., Mazurkiewicz, S.: Ein beweis des fixpunktsatzes für n dimensionale simplexe. Fund. Math. **14**, 132–137 (1929)

[131] Kneser, H.: Sur un théorème fondamental de la théorie des jeux. C.R. Math. Acad. Sci. Paris **234**, 2418–2429 (1952)

[132] Komlósi, S.: Generalized monotonicity and generalized convexity. J. Optim. Theory Appl. **84**(2), 361–376 (1995)

[133] Komlósi, S.: On the Stampacchia and Minty variational inequalities. In: Generalized Convexity and Optimization for Economic and Financial Decisions, G., Giorgi, F. Rossi, (eds.), Pitagora Editrice, Bologna, Italy, pp. 231–260 (1999)

[134] Komlósi, S.: Generalized convexity and generalized derivatives. In: Handbook of Generalized Convexity and Generalized Monotonicity, N., Hadjisavvas, S., Komlósi, S. Schaible, (eds.), Springer-Verlag, Berlin, Heidelberg, New York, pp. 421–463 (2005)

[135] Konnov, I.V.: On quasimonotone variational inequalities J. Optim. Theory Appl. **99**, 165–181 (1998)

[136] Konnov, I.V.: Equilibrium Models and Variational Inequalities. Elsevier, Amsterdam, Boston, Heidelberg, London, New York (2007)

[137] Konnov, I.V.: Combined Relaxation Methods for Variational Inequalities. Springer-Verlag, Berlin (2001)

[138] Konnov, I.V., Luc, D.T., Rubinov, A. (eds.): Generalized Convexity and Related Topics. Lecture Notes in Economics and Mathematical Systems, Vol. **583**, Springer-Verlag, Berlin, Heidelberg, New York (2007)

[139] Konnov, I.V., Yao, J.C.: On the generalized vector variational inequality problem. J. Math. Anal. Appl. **206**, 42–58 (1997)

[140] Kuhn, H.W., Tucker, A.W.: Nonlinear programming. Proceedings of the second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, pp. 481–492 (1951)

[141] Lalitha, C.S., Mehta, M.: Vector variational inequalities with cone-pseudomonotone bifunctions. Optimization **54**, 327–338 (2005)

[142] Lalitha, C.S., Mehta, M.: Characterization of the solution sets of pseudolinear programs and pseudoaffine variational inequality problems. J. Nonlinear Convex Anal. **8**(1), 87–98 (2007)

[143] Larsson, T., Patriksson, M.: A class of gap functions for variational inequalities. SIAM J. Control Optim. **64**, 53–79 (1994)

[144] Lee, G.M.: On relations between vector variational inequality and vector optimization problem. In: Progress in Optimization, II: Contributions from Australasia, X.Q. Yang, A.I. Mees, M.E. Fisher, and L.S. Jennings (eds.), Kluwer Academic Publisher, Dordrecht, Holland, pp. 167–179 (2000).

[145] Li, S.J., Hou, S.H., Chen, G.Y.: Generalized differential properties of the Auslender gap function for variational inequalities. J. Optim. Theory Appl. **124**, 739–749 (2005)

[146] Lin, L.J., Yang, M.F., Ansari, Q.H., Kassay, G.: Existence results for Stampacchia and Minty type implicit variational inequalities with multivalued maps. Nonlinear Anal. **61**, 1–19 (2005)

[147] Lions, J.-L., Stampacchia, G.: Inéquations variationnelles non coercives. C.R. Math. Acad. Sci. Paris **261**, 25–27 (1965)

[148] Lions, J.-L., Stampacchia, G.: Variational inequalities. Comm. Pure Appl. Math. **20**, 493–519 (1967)

[149] Luc, D.T.: Existence results for densely pseudomonotone variational inequalities. J. Math. Anal. Appl. **254**, 291–308 (2001)

[150] Lucchetti, R.: Convexity and Well-Posed Problems. Canadian Mathematical Society, Springer, New York (2006)

[151] Luc, D.T., Swaminathan, S.: A characterization of convex functions. Nonlinear Anal. **20**, 697–701 (1993)

[152] Mangasarian, O.L.: Pseudo-convex functions. J. Soc. Indust. Appl. Math. Ser. A Control **3**, 281–290 (1965)

[153] Mangasarian, O.L.: Nonlinear Programming. McGraw-Hill Publishing Company, New York (1969)

[154] Mangasarian, O.L.: A simple characterization of solution sets of convex programs. Oper. Res. Lett. **7**, 21–26 (1988)

[155] Marcotte, P., Zhu, D.: Weak sharp solutions of variational inequalities. SIAM J. Optim. **9**, 179–189 (1999)

[156] Martos, B.: Quasi-convexity and quasi-monotonicity in nonlinear programming. Studia Sci. Math. Hungar. **2**, 265–273 (1967)

[157] Martos, B.: Nonlinear Programming: Theory and Methods. Akadmia Kiad, Budapest (1975)

[158] Mastroeni, G.: Gap functions for equilibrium problems. J. Global Optim. **27**, 411–426 (2003)

[159] McShane, E.J.: Integration. Princeton University Press, Princeton, New Jersey (1944)

[160] Michel, P., Penot, J.-P.: Calcul sous-différentiel pour des fonctions Lipschitziennes et non Lipschitziennes. C. R. Math. Acad. Sci. Paris. **298**, 269–272 (1984)

[161] Minty, G.J.: Monotone (nonlinear) operators in Hilbert space. Duke Math. J. **29**, 341–346 (1962)

[162] Minty, G.J.: On the generalization of a direct method of the calculus of variations. Bull. Amer. Math. Soc. **73**, 315–321 (1967)

[163] Minty, G.J.: On some aspects of the theory of monotone operators. In: Theory and Applications of Monotone Operators, A. Ghizzetti, (ed.), Proc. NATO Adv. Study Institute, Venice, Oderisi, Gubbio, pp. 67–82 (1968)

[164] Mishra, S.K., Wang, S.Y.: Vector variational-like inequalities and nonsmooth vector optimization problems. Nonlinear Anal. **64**, 1939–1945 (2006)

[165] Moré, J.J.: Coercivity conditions in nonlinear complementarity problems. SIAM Rev. **16**, 1–16 (1974)

[166] Murty, K.G.: Linear Complementarity, Linear and Nonlinear Programming. Heldermann Verlag, Berlin (1988)

[167] Nagurney, A.: Network Economics: A Variational Inequality Approach. Academic Publishers, Dordrecht, Netherlands (1993)

[168] von Neumann, J.: Zur Theorie der Gesellschaftsspiele. Math. Ann. **100**, 295–320 (1928)

[169] Newman, P.: Some properties of concave functions. J. Econom. Theory **1**, 291–314 (1969)

[170] Niculescu, C., Persson, L.-E.: Convex Functions and Their Applications: A Contemporary Approach. Canadian Mathematical Society, Springer, New York (2006)

[171] Nikaido, N.: Convex Structures and Economic Theory. Academic Press, New York (1968)

[172] Parida, J., Sen, A.: A variational-like inequality for multifunctions with applications. J. Math. Anal. Appl. **124**, 73–81 (1987)

[173] Patriksson, M.: Nonlinear Programming and Variational Inequality Problems: A Unified Approach. Kluwer Academic Publishers, Dordrecht, Boston, London (1999)

[174] Penot, J.-P.: On the mean value theorem. Optimization **19**, 147–156 (1988)

[175] Penot, J.-P., Quang, P.H.: Generalized convexity of functions and generalized monotonicity of set-valued maps. J. Optim. Theory Appl. **92**, 343–356 (1997)

[176] Ponstein, J.: Seven kinds of convexity. SIAM Rev. **9**, 115–119 (1967)

[177] Popoviciu, T.: Deux remarques sur les fonctions convexes. Bull. Math. Soc. Sci. Roumaine. **20**, 45–49 (1939)

[178] Qu, B., Wang, C.Y., Zhang, J.Z.: Convergence and error bound of a method for solving variational inequality problems via generalized D-gap function. J. Optim. Theory Appl. **119**, 535–552 (2003)

[179] Repovs, D., Semenov, P.V.: Continuous Selections of Multivalued Mappings. Kluwer Academic Publishers, Dordrecht, Netherlands (1998)

[180] Roberts, A.W., Varberg, D.E.: Convex Functions. Academic Press, New York, London (1973)

[181] Rockafellar, R.T.: Convex functions, monotone operators, and variational inequalities. In: Theory and Applications of Monotone Operators, A. Ghizzetti, (ed.), Proc. NATO Adv. Study Institute, Venice, Oderisi, Gubbio (1968)

[182] Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)

[183] Rockafellar, R.T.: Directionally Lipschitzian functions and subdifferential calculus. Proc. Lond. Math. Soc. **39**, 331–355 (1979)

[184] Rockafellar, R.T.: Generalized directional derivatives and subgradients of nonconvex functions. Canad. J. Math. **32**, 257–280 (1980)

[185] Rockafellar, R.T.: The Theory of Subgradients and Its Applications to Problems of Optimization: Convex and Nonconvex Functions. Heldermann-Verlag, Berlin (1981)

[186] Sach, P.H., Penot, J.-P.: Characterizations of generalized convexities via generalized directional derivatives. Numer. Funct. Anal. Optim. **19**, 615–634 (1998)

[187] Saigal, R.: Extension of generalized complementarity problem. Math. Oper. Res. **1**, 260–266 (1976)

[188] Schaible, S., Yao, J.-C.: On the equivalence of nonlinear complementarity problems. Math. Prog. **70**, 191–2000 (1995)

[189] Schaible, S., Ziemba, W.T. (eds.): Generalized Concavity in Optimization and Economics. Academic Press, New York (1981)

[190] Schirotzek, W.: Nonsmooth Analysis. Springer, Berlin, Heidelberg (2007)

[191] Shih, M.-H., Tan, K.-K.: Browder-Hartman-Stampacchia variational inequalities for multi-valued monotone operators. J. Math. Anal. Appl. **134**, 431–440 (1988)

[192] Siddiqi, A.H., Ansari, Q.H.: An iterative method for generalized variational inequalities. Math. Japon. **34**, 475–482 (1989)

[193] Slater, M.: Lagrange multipliers revisited: A contribution to nonlinear programming. Cowles Commission Discussion Paper, Mathematics, No. 403 (1950)

[194] Smith, M.J.: The existence, uniqueness and stability of traffic equilibria. Transp. Res. **13** B, 295–304 (1979)

[195] Soleimani-damaneh, M.: Characterization of nonsmooth quasiconvex and pseudoconvex functions, J. Math. Anal. Appl. **330**, 1387–1392 (2007)

[196] Solodov, M., Tseng, P.: Some methods based on the D-gap function for solving monotone variational inequalities. Comput. Optim. Appl. **17**, 255–277 (2000)

[197] Stampacchia, G.: Formes bilineaires coercivitives sur les ensembles convexes. C. R. Math. Acad. Sci. Paris **258**, 4413–4416 (1964)

[198] Stampacchia, G.: Variational inequalities. In: Theory and Applications of Monotone Operators, A. Ghizzetti, (ed.), Proc. NATO Adv. Study Institute, Venice, Oderisi, Gubbio (1968)

[199] Tarafdar, E.U., Chowdhury, M.S.R.: Topological Methods for Set-Valued Nonlinear Analysis. World Scientific, New Jersey, London, Singapore, Hong Kong (2008)

[200] Thomson, B.S., Bruckner, J.B., Bruckner, A.M.: Elementary Real Analysis. Prentice-Hall, Inc., New York (2001)

[201] Weir, T., Jeyakumar, V.: A class of nonconvex functions and mathematical programming. Bull. Austral. Math. Soc. **38**, 177–189 (1988)

[202] Wu, Z., Wu, S.-Y.: Gâteaux differentiability of the dual gap function of a variational inequality. European J. Oper. Res. **190**, 328–344 (2008)

[203] Yang, X.Q.: On the gap functions of prevariational inequalities. J. Optim. Theory Appl. **116**, 437–452 (2003)

[204] Yang, X.Q., Yao, J.C.: Gap functions and existence of solutions to set-valued vector variational inequalities. J. Optim. Theory Appl. **115**, 407–417 (2002)

[205] Yao, J.C.: Variational inequalities with generalized monotone operators. Math. Oper. Res. **19**, 691–705 (1994)

[206] Yao, J.C.: Multi-valued variational inequalities with K-pseudomonotone operators. J. Optim. Theory Appl. **83**, 391–403 (1994)

[207] Yao, J.C., Guo, J.-S.: Variational and generalized variational inequalities with discontinuous mappings. J. Math. Anal. Appl. **182**, 371–392 (1994)

[208] Zhou, J.X., Chen, G.: Diagonal convexity conditions for problems in convex analysis and quasi-variational inequalities. J. Math. Anal. Appl. **132**, 213–225 (1988)

[209] Yu, S.Y., Yao, J.C.: On vector variatonal inequalities. J. Optim. Theory Appl. **89**, 749–769 (1996)

[210] Yuan, G. X.-Z.: KKM Theory and Applications in Nonlinear Analysis. Marcel Dekker, Inc., New York (1999)

[211] Zeng, L.-C., Ansari, Q.H., Yao, Y.-C.: Equivalence of generalized mixed complementarity and generalized mixed least element problems in ordered spaces. Optimization, **58**, 63–76 (2009)

[212] Zhang, J., Wan, C., Xiu, N.: The dual gap function for variational inequalities. Appl. Math. Optim. **48**, 129–148 (2003)

[213] Zhu, D., Marcotte, P.: New classes of generalized monotonicity. J. Optim. Theory Appl. **87**, 457–471 (1995)

[214] Zhu, D., Marcotte, P.: Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities. SIAM J. Optim. **6**, 714–726 (1996)

Operations Research/Optimization

Until now, no book addressed convexity, monotonicity, and variational inequalities together. **Generalized Convexity, Nonsmooth Variational Inequalities, and Nonsmooth Optimization** covers all three topics, including new variational inequality problems defined by a bifunction.

The first part of the book focuses on generalized convexity and generalized monotonicity. The authors investigate convexity and generalized convexity for both the differentiable and nondifferentiable case. For the nondifferentiable case, they introduce the concepts in terms of a bifunction and the Clarke subdifferential.

The second part offers insight into variational inequalities and optimization problems in smooth as well as nonsmooth settings. The book discusses existence and uniqueness criteria for a variational inequality, the gap function associated with it, and numerical methods to solve it. It also examines characterizations of a solution set of an optimization problem and explores variational inequalities defined by a bifunction and set-valued version given in terms of the Clarke subdifferential.

Integrating results on convexity, monotonicity, and variational inequalities into one unified source, this book deepens your understanding of various classes of problems, such as systems of nonlinear equations, optimization problems, complementarity problems, and fixed-point problems. The book shows how variational inequality theory not only serves as a tool for formulating a variety of equilibrium problems, but also provides algorithms for computational purposes.

CRC Press
Taylor & Francis Group
an **informa** business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487

711 Third Avenue
New York, NY 10017

2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

K13101

ISBN: 978-1-4398-6820-1
90000
9 781439 868201