Springer Optimization and Its Applications 50
Nonconvex Optimization and Its Applications

Shashi Kant Mishra *Editor*

Topics in Nonconvex Optimization

Theory and Applications

Springer

Springer Optimization and Its Applications 50

Series Editor:
Panos M. Pardalos

Subseries:
Nonconvex Optimization and Its Applications

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

Shashi Kant Mishra
Editor

Topics in Nonconvex Optimization

Theory and Applications

Springer

Editor
Shashi Kant Mishra
Banaras Hindu University
Faculty of Science
Dept. of Mathematics
Varanasi
India
bhu.skmishra@gmail.com

ISSN 1931-6828
ISBN 978-1-4419-9639-8 e-ISBN 978-1-4419-9640-4
DOI 10.1007/978-1-4419-9640-4
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011929034

Printed on acid-free paper

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

I would like to dedicate this volume to my teacher Prof. R. N. Mukherjee, who introduced this wonderful field of mathematics to me. I would also like to dedicate this volume to Prof. B. D. Craven who showed me the path in this research area.

Foreword

It is a great pleasure to learn that the Centre for Interdisciplinary Mathematical Sciences and the Department of Mathematics, Banaras Hindu University organized an Advanced Training Programme on Nonconvex Optimization and Its Applications. This programme was organized to introduce the subject to young researchers and college teachers working in the area of nonconvex optimization.

During the five-day period several eminent professors from all over the country working in the area of optimization gave expository to advanced level lectures covering the following topics.

(i) Quasi-convex optimization
(ii) Vector optimization
(iii) Penalty function methods in nonlinear programming
(iv) Support vector machines and their applications
(v) Portfolio optimization
(vi) Nonsmooth analysis
(vii) Generalized convex optimization

Participants were given copies of the lectures. I understand from Dr. S. K. Mishra, the main organizer of the programme, that the participants thoroughly enjoyed the lectures related to nonconvex programming. I am sure the students will benefit greatly from this kind of training programme and I am confident that Dr. Mishra will conduct a more advanced programme of this kind soon. I also appreciate the efforts taken by him to get these lectures published by Springer. I am sure this volume will serve as excellent lecture notes in optimization for students and researchers working in this area.

Chennai, April 2010

Thiruvenkatachari Parthasarathy
INSA Senior Scientist
Indian Statistical Institute, Chennai, India

Preface

Optimization is a multidisciplinary research field that deals with the characterization and computation of minima and/or maxima (local/global) of nonlinear, nonconvex, nonsmooth, discrete, and continuous functions. Optimization problems are frequently encountered in modelling of complex real-world systems for a very broad range of applications including industrial and systems engineering, management science, operational research, mathematical economics, seismic optimization, production planning and scheduling, transportation and logistics, and many other applied areas of science and engineering. In recent years there has been growing interest in optimization theory.

The present volume contains 16 full-length papers that reflect current theoretical studies of generalized convexity and its applications in optimization theory, set-valued optimization, variational inequalities, complementarity problems, cooperative games, and the like. All these papers were refereed and carefully selected from those delivered at the Advanced Training Programme on Nonconvex Optimization and Its Applications held at the DST-Centre for Interdisciplinary Mathematical Sciences, Department of Mathematics, Banaras Hindu University, Varanasi, India, March 22–26, 2010.

I would like to take this opportunity, to thank all the authors whose contributions make up this volume, all the referees whose cooperation helped in ensuring the scientific quality of the papers, and all the people from the DST-CIMS and Department of Mathematics, Banaras Hindu University, whose assistance was indispensable in running the training programme. I would also like to thank to all the participants of the advanced training programme, especially those who travelled a long distance within India in order to participate. Finally, we express our appreciation to Springer for including this volume in their series. We hope that the volume will be useful for students, researchers, and those who are interested in this emerging field of applied mathematics.

Varanasi, April, 2010 *Shashi Kant Mishra*

Contents

List of Contributors

Qamrul Hasan Ansari
Department of Mathematics, Aligarh Muslim University, Aligarh, UP, India,
e-mail: qhansari@gmail.com

Pooja Arora
Department of Operational Research, University of Delhi, Delhi 110 007, India,
e-mail: poojaarora82@rediffmail.com

R. Arora
Department of Mathematics, University of Delhi, Delhi 110 007, India,
e-mail: ruchiiq@yahoo.com

Guneet Bhatia
Department of Mathematics, University of Delhi, Delhi 110 007, India,
e-mail: guneet172@yahoo.co.in

Surajit Borkotokey
Department of Mathematics, Dibrugarh University, Asom, India,
e-mail: surajitbor@yahoo.com

L. Coladas Uria
Department of Statistics and Operations Research, Faculty of Mathematics,
University of Santiago de Compostela, 15782 Santiago de Compostela, Spain,
e-mail: luis.coladas@usc.es

A. K. Das
Indian Statistical Institute, 203, B. T. Road, Kolkata 700 108, India,
e-mail: akdas@isical.ac.in

Anulekha Dhara
Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas,
New Delhi 110 016, India, e-mail: anulekha_iitd@yahoo.com

Anjana Gupta
Department of Mathematics, MAIT, GGSIPU, Delhi, India,
e-mail: guptaanjana2003@yahoo.co.in

Pankaj Gupta
Department of Operational Research, University of Delhi, Delhi, India,
e-mail: pgupta@or.du.ac.in

Sy-Ming Guu
College of Management, Yuan-Ze University, 135 Yuan-Tung Road, Chung-Li, Taoyuan, 32003, Taiwan, e-mail: iesmguu@gmail.com

Abdelouahed Hamdi
Department of Mathematics and Computer Science, Kuwait University, Kuwait,
e-mail: whamdi99@yahoo.fr

Phan Quoc Khanh
Department of Mathematics, International University of Hochiminh City, Linh Trung, Thu Duc, Hochiminh City, Vietnam, e-mail: pqkhanh@hcmiu.edu.vn

Bhawna Kohli
Department of Mathematics, University of Delhi, Delhi 110 007, India,
e-mail: bhawna.kohli@rediffmail.com

Kin Keung Lai
Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, e-mail: mskklai@cityu.edu.hk

C. S. Lalitha
Department of Mathematics, University of Delhi South Campus, Benito Jaurez Road, New Delhi 110 021, India, e-mail: cslalitha@rediffmail.com

Lai-Jiu Lin
Department of Mathematics, National Changhua University for Education, Changhua 50058, Taiwan, ROC, e-mail: maljlin@math.ncue.edu.tw

Mukesh Kumar Mehlawat
Department of Operational Research, University of Delhi, Delhi, India,
e-mail: mukesh0980@yahoo.com

Aparna Mehra
Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110 016, India, e-mail: apmehra@maths.iitd.ac.in

S. K. Mishra
Department of Mathematics, Faculty of Science, Banaras Hindu University, Varanasi 221 005, India, e-mail: bhu.skmishra@gmail.com

C. Nahak
Department of Mathematics, Indian Institute of Technology, Kharagpur, India,
e-mail: cnahak@maths.iitkgp.ernet.in

Sudarsan Nanda
Department of Mathematics, Indian Institute of Technology, Kharagpur, India,
e-mail: snanda@maths.iitkgp.ernet.in

S. K. Neogy
Indian Statistical Institute, 7, S. J. S. Sansanwal Marg, New Delhi 110 016, India,
e-mail: skn@isid.ac.in

S. K. Padhan
Department of Mathematics, Indian Institute of Technology, Kharagpur, India,
e-mail: sarojpadhan@gmail.com

J. S. Rautela
Department of Mathematics, Faculty of Applied Sciences and Humanities, Echelon Institute of Technology, Faridabad 121 002, India,
e-mail: sky_dreamz@rediffmail.com

Kalpana Shukla
Department of Mathematics, Faculty of Science, Banaras Hindu University, Varanasi 221 005, India, e-mail: bhukalpanabhu@gmail.com

Vinay Singh
Department of Mathematics, Faculty of Science, Banaras Hindu University, Varanasi 221 005, India, e-mail: vinaybhu1981@gmail.com

Nguyen Dinh Tuan
Department of Mathematics, University of Economics of Hochiminh City, Nguyen Dinh Chieu Street, D. 3, Hochiminh City, Vietnam, e-mail: ndtuan73@yahoo.com

Shouyang Wang
Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, e-mail: sywang@amss.ac.cn

Jen-Chih Yao
Center for General Education, Kaohsiung Medical University, Kaohsiung 80708, Taiwan, e-mail: yaojc@kmu.edu.tw

Y. X. Zhao
Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, e-mail: zyx@amss.ac.cn

Chapter 1
Some Equivalences Among Nonlinear Complementarity Problems, Least-Element Problems, and Variational Inequality Problems in Ordered Spaces

Qamrul Hasan Ansari and Jen-Chih Yao

Abstract In this survey chapter we introduce several Z-type single-valued maps as well as set-valued maps. We present several equivalences among different types of nonlinear programs, different types of least-element problems, and different types of variational inequality problems under certain regularity and growth conditions.

1.1 Introduction

It is well known that the theory of complementarity problems has been become a very effective and powerful tool in the study of a wide class of linear and nonlinear problems in optimization, economics, game theory, mechanics, engineering, and so on, see, for example [9, 15–17], and the references therein. For a long time, a great deal of effort has gone into the study of the equivalence of complementarity problems and other problems. In 1980, Cryer and Dempster [10] studied the equivalence of linear complementarity problems, linear programs, least-element problems, variational inequality problems, and minimization problems in vector lattice Hilbert spaces. In 1981, Riddle [28] established the equivalence of complementarity and least-element problems as well as several related problems. In 1995, Schaible and Yao [30] proved the equivalence of these problems by introducing strictly pseudomonotone Z-maps operating on Banach lattices. In 1999, Ansari et al. [1] extended the results of Schaible and Yao [30] for point-to-set maps and established equivalence among generalized complementarity problems, generalized least-element problems, generalized variational inequality problems,

Qamrul Hasan Ansari
Department of Mathematics, Aligarh Muslim University, Aligarh 202 002, India,
e-mail: qhansari@gmail.com

Jen-Chih Yao
Center for General Education, Kaohsiung Medical University, Kaohsiung 80708, Taiwan,
e-mail: yaojc@kmu.edu.tw

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_1,

and minimization problems. In [34] Yin, Xu, and Zhan established the equivalence of F-complementarity, variational inequality, and least-element problems in the Banach space setting. Very recently, Huang and Fang [14] introduced several classes of strong vector F-complementarity problems and gave their relationships with the least element problems of feasible sets. Furthermore, in [36], Zeng and Yao first gave an equivalence result for variational-like inequality problems and least element problems.

In this survey chapter we introduce several Z-type single-valued maps as well as multivalued maps. We present several equivalences among different types of nonlinear programs, least-element problems, complementarity problems, and variational inequality problems under certain regularity and growth conditions.

1.2 Preliminaries

In this section, we introduce some notations and definitions that are used in the sequel.

Let B be a real Banach space with its dual B^*, and let $K \subseteq B$ be a closed convex cone. Let K^* be the dual cone of K; that is,

$$K^* = \{u \in B^* : \langle u, x\rangle \geq 0 \text{ for all } x \in K\},$$

where $\langle u, x\rangle$ denotes the pairing between $u \in B^*$ and $x \in B$.

The vector ordering induced by K on B and induced by K^* on B^* is denoted by $\leq$:

$$x \leq y \quad \text{if and only if } y - x \in K, \quad \text{for all } x, y \in B,$$
$$u \leq v \quad \text{if and only if } v - u \in K^*, \quad \text{for all } u, v \in B^*.$$

Nonzero elements of K^* are said to be *positive*, and $u \in K^*$ is said to be *strictly positive* if

$$\langle u, x\rangle > 0, \quad \text{for all } x \in K,\ x \neq 0.$$

The space B is a vector lattice with respect to $\leq$ if each pair $x, y \in B$ has a unique infimum $x \wedge y$ characterized by the properties

$$x \wedge y \leq x,\ x \wedge y \leq y,\ z \leq x, z \leq y \quad \text{if and only if } z \leq x \wedge y.$$

If B is a vector lattice, so is B^* with respect to the ordering $\leq$ induced by K^*; see, for example, [22].

Proposition 1.1 ([2, pp. 533]). *Let K be a nonempty convex subset of B and let $f : K \to \mathbb{R}$ be a lower semicontinuous and convex functional. Then, f is weakly lower semicontinuous.*

Remark 1.1. From Proposition 1.1, we can see that, if $f : K \to \mathbb{R}$ is upper semicontinuous and concave, then f is weakly upper semicontinuous.

Definition 1.1. Let Ω be an open subset of a real Banach space B. A function $f : \Omega \to \mathbb{R}$ is said to be *Gâteaux differentiable at* $x \in \Omega$ if there exists $\nabla f(x) \in B^*$ such that

$$\lim_{t \to 0^+} \frac{f(x+th) - f(x)}{t} = \langle \nabla f(x), h \rangle, \quad \forall h \in B.$$

$\nabla f(x)$ is called the *Gâteaux derivative of* f at the point x. The function f is Gâteaux differentiable in Ω if it is Gâteaux differentiable at each point of Ω.

Let K be a closed subset of B and $f : K \to \mathbb{R}$. By saying f is Gâteaux differentiable in K we mean that f is Gâteaux differentiable in an open set neighborhood of K.

Definition 1.2 ([3]). Let Ω be an open subset of a real Banach space B and $f : \Omega \to \mathbb{R}$ be Gâteaux differentiable. The function f is said to be

(i) *Pseudoconvex* on Ω if for every pair of points $x, y \in \Omega$, we have

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

(ii) *Strictly pseudoconvex* on Ω if for every pair of distinct points $x, y \in \Omega$, we have

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

The relation of (strict) pseudoconvexity and (strict) pseudomonotonicity is the following.

Theorem 1.1 ([19, 20]). *Let Ω be an open convex subset of a real Banach space B and $f : \Omega \to \mathbb{R}$ be Gâteaux differentiable. Then f is (strictly) pseudoconvex on Ω if and only if $\nabla f : \Omega \to B^*$ is (strictly) pseudomonotone.*

We note that if $f : \Omega \to \mathbb{R}$ is strictly pseudoconvex, then the solution of $\min_{x \in \Omega} f(x)$ is unique provided a solution exists [3].

Definition 1.3. Let $f : B \to \mathbb{R}$ be a functional. Then an element $u \in B^*$ is called a *subgradient* of f at the point $x \in B$ if $f(x)$ is finite and

$$\langle u, y - x \rangle \leq f(y) - f(x), \qquad \forall y \in B.$$

The set of all subgradients of f at x is called the *subdifferential* of f at x and is denoted by $\partial f(x)$. That is,

$$\partial f(x) = \{u \in B^* : \langle u, y - x \rangle \leq f(y) - f(x)\}, \qquad \forall y \in B,$$

and therefore the subdifferential of f is the point-to-set map $\partial f : x \mapsto \partial f(x)$ from B to B^*.

Lemma 1.1 ([52]). *Let $(X,\|\cdot\|)$ be a normed vector space and $\mathscr{H}$ be a Hausdorff metric on the collection $CB(X)$ of all nonempty, closed, and bounded subsets of X, which is defined as*

$$\mathscr{H}(U,V) = \max\left\{\sup_{u\in U}\inf_{v\in V}\|u-v\|, \sup_{v\in V}\inf_{u\in U}\|u-v\|\right\},$$

for U and V in $CB(X)$, where the metric d is induced by $d(u,v)=\|u-v\|$. If U and V are compact sets in X, then for each $u\in U$, there exists $v\in V$ such that

$$\|u-v\| \leq \mathscr{H}(U,V).$$

Let D be a nonempty subset of a topological vector space X. A point-to-set map $G : D \to 2^X$ is called a *KKM map* if for each finite subset $\{x_1,\ldots,x_n\} \subseteq D$,

$$\text{co}\{x_1,\ldots,x_n\} \subseteq \bigcup_{i=1}^{n} G(x_i),$$

where $\text{co}\{x_1,\ldots,x_2\}$ denotes the convex hull of $\{x_1,\ldots,x_n\}$.

Lemma 1.2 ([11]). *Let D be an arbitrary nonempty subset of a Hausdorff topological vector space X. Let the point-to-set map $G : D \to 2^X$ be a* KKM *map such that $G(x)$ is closed for all $x\in D$ and is compact for at least one $x\in D$. Then $\cap_{x\in D} G(x) \neq \emptyset$.*

1.3 Equivalence of Nonlinear Complementarity Problems and Least-Element Problems

Given are a closed convex cone $K \subseteq B$, $T : K \to B^*$ and $f : K \to \mathbb{R}$ whose special properties do not concern us for the moment. We denote by $\mathscr{F}$ the *feasible set* of T with respect to K; that is,

$$\mathscr{F} = \{x \in B : x \in K \text{ and } T(x) \in K^*\}.$$

In this section, we consider the following problems.

(I) *Nonlinear program* : For a given $u \in B^*$, find $x \in \mathscr{F}$ such that

$$\langle u,x\rangle = \min_{y\in\mathscr{F}} \langle u,y\rangle.$$

(II) *Least-element problem* : Find $x \in \mathscr{F}$ such that

$$x \leq y, \qquad \forall y \in \mathscr{F}.$$

(III) *Complementarity problem* : Find $x \in \mathscr{F}$ such that

$$\langle u, x \rangle = 0.$$

(IV) *Variational inequality problem* : Find $x \in K$ such that

$$\langle T(x), y - x \rangle \geq 0, \qquad \forall y \in K.$$

(V) *Unilateral minimization problem* : Find $x \in K$ such that

$$f(x) = \min_{y \in K} f(y).$$

The equivalence of (I) and (II) on the one hand, and among (III), (IV), and (V) is well known; see, for example [18, 28]. The purpose of this section is to investigate suitable conditions under which these five problems are equivalent.

Definition 1.4 ([28]). Let B be a Banach space that is also a vector lattice with positive cone K; let $T : K \to B^*$ be a mapping. Then T is called a *Z-map relative to* K if for any $x, y, z \in K$,

$$\langle T(x) - T(y), z \rangle \leq 0, \quad \text{whenever } (x - y) \wedge z = 0.$$

In the case where T is linear, Definition 1.4 reduces to the definition of *condition* Z in [10]. In the case where $B = \mathbb{R}^n$ and K is the nonnegative orthant, T is a Z-map relative to K if and only if it is *off-diagonally antitone* in the sense of [27].

Definition 1.5 ([18, 20, 28]). Let B be a Banach space, K a nonempty convex subset of B, and $T : K \to B^*$ a mapping. Then T is called

(i) *Pseudomonotone* if for any $x, y \in K$,

$$\langle T(y), x - y \rangle \geq 0 \quad \text{implies} \quad \langle T(x), x - y \rangle \geq 0$$

(ii) *Strictly pseudomonotone* if for any distinct points $x, y \in K$

$$\langle T(y), x - y \rangle \geq 0 \quad \text{implies} \quad \langle T(x), x - y \rangle > 0$$

(iii) *Hemicontinuous* if it is continuous on the line segments in K with respect to weak* topology in B^*; that is, for any fixed $x, y, z \in K$, the function

$$t \mapsto \langle T(x + ty), z \rangle, \qquad 0 \leq t \leq 1$$

is continuous

(iv) *Positive at infinity* if for any $x \in K$, there exists a positive real number $\rho(x)$ such that $\langle T(y), y - x \rangle > 0$ for every $y \in K$ such that $\|y\| \geq \rho(x)$.

Lemma 1.3 ([30]). *Let K be a convex cone in a Banach space B and $T : K \to B^*$ be (strictly) pseudomonotone. Then for each fixed $z \in K$, the operator $T_z : K \to 2^{B^*}$ defined by*

$$T_z(x) = T(x+z), \qquad \forall x \in K$$

is also (strictly) pseudomonotone.

Proof. For any $x, y \in K$, suppose that $\langle T_z(y), x-y \rangle \geq 0$. Then $\langle T(y+z), x-y \rangle \geq 0$, from which it follows that $\langle T(y+z), (x+z)-(y+z) \rangle \geq 0$. Because T is pseudomonotone, we have

$$\langle T(x+z), (x+z)-(y+z) \rangle \geq 0$$

and hence

$$\langle T_z(x), x-y \rangle \geq 0.$$

Therefore, T_z is also pseudomonotone. The case where T is strictly pseudomonotone can be dealt with by a similar argument.

We need the following result to derive the equivalence of problems (I)–(V) under suitable conditions.

Theorem 1.2. *Let K be a nonempty, closed, bounded convex subset of a reflexive Banach space B and let $T : K \to B^*$ be weakly pseudomonotone and hemicontinuous. Then there exist $x \in K$ such that*

$$\langle T(x), y-x \rangle \geq 0, \qquad \forall y \in K.$$

Furthermore, if in addition T is strictly pseudomonotone, the solution is unique.

Theorem 1.2 is an extension of classical existence results for variational inequalities due to [4, 13]. By employing Theorem 1.2, we obtain the following result for perturbed variational inequalities.

Proposition 1.2 ([30]). *Let K be a nonempty, closed, convex cone in a reflexive Banach space B, and $T : K \to B^*$ be pseudomonotone, hemicontinuous, and positive at infinity. Then for each fixed $z \in K$, there exist $x \in K$ such that*

$$\langle T(x+z), y-x \rangle \geq 0, \qquad \forall y \in K. \tag{1.1}$$

If, in addition, T is strictly pseudomonotone, then for each $z \in K$, (1.1) *has a unique solution.*

Proof. For each $z \in K$, we define $T_z : K \to B^*$ by

$$T_z(x) = T(x+z), \qquad \forall x \in K.$$

Then obviously, T_z is hemicontinuous.

By Lemma 1.2, T_z is also pseudomonotone. Let $\rho = \|z\| + \rho(z)$, where $\rho(z)$ is defined as in the definition of positive at infinity. Let

$$D = \{y + z : y \in K, \|y\| \leq \rho\},$$

which is a closed, bounded, convex subset of a reflexive Banach space B. Then by Theorem 1.2, there exist $x \in K$ with $\|x\| \leq \rho$ such that

$$\langle T_z(x), y - x\rangle \geq 0, \qquad \forall y \in K \qquad \text{with } \|y\| \leq \rho. \tag{1.2}$$

We note that $\|x\| < \rho$. Suppose that $\|x\| = \rho$; then

$$\|x + z\| \geq \|x\| - \|z\| = \rho(z).$$

T is positive at infinity, thus we have

$$\langle T(x + z), x\rangle > 0,$$

or

$$\langle T_z(x), x\rangle > 0. \tag{1.3}$$

On the other hand, letting $y = 0$ in (1.2), we have

$$\langle T_z(x), x\rangle \leq 0,$$

which is a contradiction of (1.3). Therefore, $\|x\| < \rho$ and by standard technique it can be shown that x is indeed a solution of (1.1).

If, in addition, T is strictly pseudomonotone, then by Lemma 1.3, T_z is also strictly pseudomonotone. Consequently, the solution is unique.

In the remaining part of this section, we assume that B is a real Banach space and K is a closed convex cone of B, and, whenever the ordering induced by K is mentioned, $(B, \leq)$ is a vector lattice.

Now we establish the equivalence of problems (I)–(V) under suitable conditions.

Proposition 1.3 ([30]). *Let $T : K \longrightarrow B^*$ be the Gâteaux derivative of $f : K \longrightarrow \mathbb{R}$. Then any solution of* (V) *is also a solution of* (IV). *If in addition, T is pseudomonotone, then, conversely, any solution of* (IV) *is also a solution of* (V).

Proposition 1.4 ([18, Lemma 3.1]). *Let $T : K \longrightarrow B^*$. Then x is a solution of* (III) *if and only if it is a solution of* (IV).

Proposition 1.5 ([30]). *Suppose that $T : K \longrightarrow B^*$ is strictly pseudomonotone and a Z-map relative to Z. Then any solution of* (IV) *is also a solution of* (II).

Proposition 1.6 ([18, Lemma 3.1]). *Let $T : K \longrightarrow B^*$ and $u \in K^*$. Then any solution of* (II) *is a solution of* (I).

Proposition 1.7 ([30]). *Let B be a reflexive Banach space. Assume that $T : K \longrightarrow B^*$ is a Z-map relative to K, strictly pseudomonotone, hemicontinuous, and positive at* *infinity. Then the feasible set* $\mathscr{F} = \{x \in B : x \in K$ *and* $T(x) \in K^*\}$ *is a* $\wedge$*-sublattice; that is,* $x \in \mathscr{F}$ *and* $y \in \mathscr{F}$ *imply* $x \wedge y \in \mathscr{F}$.

Proposition 1.8 ([30]). *Let B be a reflexive Banach space. Assume that* $T : K \to B^*$ *is a Z-map relative to K, strictly pseudomonotone, hemicontinuous, and positive at infinity. Let* $u \in K^*$ *be strictly positive. Then Problem* (I) *corresponding to u has at most one solution, and any solution of* (I) *is also a solution of* (II).

By combining Propositions 1.3 and 1.5–1.7, we have the following main result of this section.

Theorem 1.3. *Let K be a closed convex cone in a reflexive Banach space B such that B is a vector lattice with respect to the order* $\leq$ *induced by K. Let* $T : K \to B^*$ *be a Z-map relative to K, strictly pseudomonotone, hemicontinuous, and positive at infinity. If* $u \in K^*$ *is a strictly positive element, then there exists* $x \in \mathscr{F}$ *which is a solution of problems* (I)–(IV). *Moreover, the solution x is unique. If T is the Gâteaux derivative of* $f : K \to \mathbb{R}$*, then x is also a unique solution of problem* (V).

Corollary 1.1. *Let K be a closed convex cone in a reflexive Banach space B such that B is a vector lattice with respect to the order* $\leq$ *induced by K. Let* $T : K \to B^*$ *be a Z-map relative to K, strongly pseudomonotone and hemicontinuous. If* $u \in K^*$ *is a strictly positive element, then there exists* $x \in \mathscr{F}$ *which is a solution of problems* (I)–(IV). *Moreover, the solution x is unique. If T is the Gâteaux derivative of* $f : K \to \mathbb{R}$*, then x is also a unique solution of problem* (V).

The following example illustrates that the extension of Riddell's result is not empty.

Example 1.1. Let $B = \mathbb{R}^n$ with the Euclidean norm. Then $B^* = \mathbb{R}^n$. The pairing between $x = (x_1, \ldots, x_n) \in B$ and $u = (u_1, \ldots, u_n) \in B^*$ is given by

$$\langle u, x \rangle = \sum_{i=1}^{n} u_i x_i.$$

Let K be the nonnegative orthant. Then $K^* = K$ and the reduced ordering makes B a vector lattice with

$$x \wedge y = (z_1, \ldots, z_n), \qquad z_i = \min\langle y_i, x_i \rangle.$$

Let $T : [0, \infty) \to \mathbb{R}$ be defined as $T(x) = 2 + (1/10)x + \sin x$ for $x \geq 0$. Then it can be checked that T is strictly pseudomonotone and a Z-map relative to $[0.\infty)$. T is also positive at infinity. Note that T is not monotone because $\langle T(x) - T(y), x - y \rangle < 0$ for $x = (3/2)\pi$ and $y = 0$.

1.4 Equivalence Between Variational-Like Inequality Problem and Least-Element Problem

Let B be a real Banach space with norm $\|\cdot\|$ and dual B^*. Let $K \subset B$ be a nonempty convex subset, $f : K \to B^*$ be a single-valued mapping, and $\varphi : K \to \mathbb{R}$ be a convex functional. For a given mapping $\eta : K \times K \to B$, we consider the following *variational-like inequality problem* of finding $x^* \in K$ such that

$$\langle f(x^*), \eta(x,x^*)\rangle \geq F(x^*) - F(x), \quad \text{for all } x \in K. \tag{1.4}$$

If $B = H$ is a real Hilbert space, $K = H$, $\eta(x,y) = x - y$ for all $x, y \in H$, $f : H \to H$ is a single-valued mapping, and $F : H \to \mathbb{R}$ is a linear continuous functional, then the problem (1.4) reduces to the following *variational inequality problem*. Find $x^* \in K$ such that

$$\langle f(x^*), x - x^*\rangle \geq F(x^*) - F(x), \quad \text{for all } x \in K. \tag{1.5}$$

If $F \equiv 0$, then the problem (1.4) reduces to the following *variation-like inequality problem*: Find $x^* \in K$ such that

$$\langle f(x^*), \eta(x,x^*)\rangle \geq 0, \quad \text{for all } x \in K. \tag{1.6}$$

The problem (1.6) is studied in the setting of finite-dimensional Eucludian space in [26] and infinite-dimensional spaces in [31].

If $K \subset B$ is a closed convex cone, and $\eta(x,y) = x - y$ for all $x, y \in K$, then the problem (1.4) reduces to the *variational inequality problem*: find $x^* \in K$ such that

$$\langle f(x^*), x - x^*\rangle \geq F(x^*) - F(x), \quad \text{for all } x \in K. \tag{1.7}$$

In order to study the F-complementarity problem, Yin, Xu, and Zhang [34] introduced and considered the problem (1.7), and established the equivalence between problem (1.7) and the F-complementarity problem in the case when $F : K \to \mathbb{R}$ is positively homogeneous. More precisely, let B be a real Banach space and B^* the dual space. Let K be a closed convex cone in B, $f : K \to B^*$ and $F : K \to \mathbb{R}$. The *F-complementarity problem* is to find $x^* \in K$ such that

$$\langle x^*, f(x^*)\rangle + F(x^*) = 0 \quad \text{and} \quad \langle x, f(x^*)\rangle + F(x) \geq 0, \quad \text{for all } x \in K.$$

Furthermore, by virtue of the existence of solutions of problem (1.7), they studied the equivalence between the F-complementarity problem and the least element problem.

In this section, we establish the existence results for solutions of variational-like inequality problems in the case when $K \subset B$ is a nonempty closed convex subset containing zero. Furthermore, we prove that the feasible sets of problem (1.4) are $\wedge$-sublattices in the vector lattice. Moreover, we investigate the equivalence between problem (1.4) and the least element problems. The results of this section improve and generalize the results of Yin et al. [34] by extending the variational inequality problem (1.7) in [34] to the variational-like inequality problem (1.4). In addition, these results also generalize and extend the corresponding results in [26, 28, 30].

We give some notations and definitions that are used in the rest of this section.

Definition 1.6. Let $f : K \to B^*$ and $\eta : K \times K \to B$. f is said to be η-*hemicontinuous* on K if for every fixed $x, y \in K$, the function

$$t \mapsto \langle f(x + t(y - x)), \eta(y, x)\rangle$$

is continuous at 0^+. In particular, if $\eta(x, y) = x - y$ for all $x, y \in K$, then f is said to be hemicontinuous on K.

Definition 1.7. Let $f : K \to B^*$ and $\eta : K \times K \to B$. Let $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ be a nonnegative function and $F : K \to \mathbb{R}$ be a convex functional.

(i) f is said to be *strictly* $\eta - \alpha$*-monotone* on K if for each $x, y \in K$ and $x \neq y$,

$$\langle f(x) - f(y), \eta(x, y)\rangle > \alpha(\|x - y\|).$$

In the case where $\alpha(t) = 0$, f is said to be *strictly* η*-monotone* on K. In particular, if $\eta(x, y) = x - y$ for all $x, y \in K$, then f is said to be *strictly* α*-monotone* on K.

(ii) f is said to be $\eta - F$*-pseudomonotone* on K if for each $x, y \in K$ and $x \neq y$,

$$\langle f(y), \eta(x, y)\rangle \geq F(y) - F(x) \Longrightarrow \langle f(x), \eta(x, y)\rangle \geq F(y) - F(x).$$

In particular, if $\eta(x, y) = x - y$ for all $x, y \in K$, then f is said to be F*-pseudomonotone* on K.

(iii) f is said to be *strictly* $\eta - F$*-pseudomonotone* on K if for each $x, y \in K$,

$$\langle f(y), \eta(x, y)\rangle \geq F(y) - F(x) \Longrightarrow \langle f(x), \eta(x, y)\rangle > F(y) - F(x).$$

In particular, if $\eta(x, y) = x - y$ for all $x, y \in K$, then f is said to be *strictly* F*-pseudomonotone* on K.

(iv) f is said to satisfy the η*-coercive condition* with respect to F if for any given $y \in K$, there exists a positive number $\rho(y)$ such that

$$\langle f(x + y), \eta(x, 0)\rangle + F(x) > F(0)$$

for all $x \in K$ with $\|x\| = \rho(y)$. In particular, if $\eta(x, y) = x - y$ for all $x, y \in K$, then f is said to satisfy the *coercive condition* with respect to F.

It is clear that strictly $\eta - \alpha$-monotone $\Rightarrow$ strictly η-monotone $\Rightarrow$ strictly $\eta - F$-pseudomonotone $\Rightarrow$ $\eta - F$-pseudomonotone.

Remark 1.2. If $\eta(x, y) = x - y$ for all $x, y \in K$, then Definitions 1.6 and 1.7 reduce to Definitions 2.1 and 2.2 in Yin, Xu, and Zhang [34], respectively. Definition 1.6 with $\eta(x, y) = x - y$ was previously introduced by Riddell [28].

Definition 1.8 ([34]). Let $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ be a nonnegative function and $F : K \to \mathbb{R}$ a functional, where $K + K \subset K$. F is said to be *α-bounded* on K if for each $x, y \in K$,

$$F(x) + F(y) - F(x+y) \leq \min\{\alpha(\|x\|), \alpha(\|y\|)\}.$$

Throughout this section, unless otherwise specified, we assume that B is a real Banach space and that $K \subset B$ is a nonempty, closed, convex subset containing zero.

Theorem 1.4 ([36]). *Let B be a reflexive Banach space, and $F : K \to \mathbb{R}$ a lower semicontinuous and convex functional. Let $f : K \to B^*$ be an η-hemicontinuous and $\eta - F$-pseudomonotone mapping, where $\eta : K \times K \to B$ has the properties:*

(i) $\eta(x,y) + \eta(y,x) = 0$ for all $x, y \in K$.
(ii) $\eta(\cdot,\cdot)$ is affine in the first variable.
(iii) For each fixed $y \in K$, $x \mapsto \eta(y,x)$ is sequentially continuous from the strong topology to the weak topology.

Assume that there exists a positive number $r > 0$ such that

$$\langle f(x), \eta(x,0)\rangle + F(x) > F(0), \quad \textit{for all } x \in K \textit{ with } \|x\| = r. \tag{1.8}$$

Then the variational-like inequality problem (1.4) *has a solution in K. In particular, if f is strictly $\eta - F$-pseudomonotone, then the solution is unique.*

As consequences of Theorem 1.4, we immediately obtain the following corollaries.

Corollary 1.2 ([34, Theorem 3.1]). *Let B be a reflexive Banach space, and $F : K \to \mathbb{R}$ a lower semicontinuous and convex functional. Let $f : K \to B^*$ be a hemicontinuous and F-pseudomonotone mapping. If there exists a positive number $r > 0$ such that*

$$\langle f(x), x\rangle + F(x) > F(0), \quad \textit{for all } x \in K \textit{ with } \|x\| = r,$$

then the variational inequality problem (1.7) *has a solution in K. In particular, if f is strictly F-pseudomonotone on K, then the solution is unique.*

Corollary 1.3 ([34, Corollary 3.2]). *Let B be a reflexive Banach space, and $F : K \to \mathbb{R}$ a lower semicontinuous and convex functional. Let $f : K \to B^*$ be a hemicontinuous and strictly monotone mapping. If f satisfies the coercive condition with respect to F, then for any given $z \in K$, there exists a unique element $x^* \in K$ such that*

$$\langle x - x^*, f(x^* + z)\rangle \geq F(x^*) - F(x), \quad \textit{for all } x \in K.$$

Following the idea of Yin, Xu, and Zhang [34], we define the feasible set of the variational-like inequality problem (1.4) as follows,

$$\mathscr{D} = \{w \in K : \langle f(w), \eta(u, u \wedge w)\rangle + F(u - u \wedge w) \geq 0 \text{ for all } u \in K\}.$$

In particular, if $\eta(x,y) = x - y$ for all $x, y \in K$, then the feasible set of the variational-like inequality problem (1.4) reduces to that of the variational inequality problem (1.7); that is,

$$\mathscr{D} = \{x \in K : \langle f(x), y - y \wedge x\rangle + F(y - y \wedge x) \geq 0 \text{ for all } u \in K\}.$$

Definition 1.9. Let $(B, \leq)$ be a vector lattice. A function $f : K \to B^*$ is said to be an $\eta - Z$*-mapping* on K if for each $u, v, w \in K$,

$$v \wedge (w - u) = 0 \Rightarrow \langle f(w) - f(u), \eta(u + v, u)\rangle \leq 0.$$

In particular, if $\eta(x,y) = x - y$ for all $x, y \in K$, then f is said to be a *Z-mapping* on K.

Theorem 1.5 ([36]). *Let B be a reflexive Banach space, and $(B, \leq)$ a vector lattice. Let $F : K \to \mathbb{R}$ be a functional and $f : K \to B^*$ an η – Z-mapping, where $\eta : K \times K \to B$ is a mapping such that $\eta(x,y) + \eta(y,x) = 0$ for all $x, y \in K$. Assume that the following conditions are satisfied.*

(i) There exists a nonnegative function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ such that

(a) f is strictly η – α-monotone on K.
(b) F is α-bounded on K.

(ii) For any given $z \in K$, there exists $x^* \in K$ such that

$$\langle f(x^* + z), \eta(u, u \wedge z + x^*)\rangle \geq F(x^*) - F(u - u \wedge z) \quad \text{for all } u \in K.$$

If the feasible set $\mathscr{D}$ of the variational-like inequality problem (1.4) *is nonempty, then $\mathscr{D}$ is a $\wedge$-sublattice of B.*

Corollary 1.4 ([36]). *Let B be a reflexive Banach space, and $(B, \leq)$ a vector lattice. Let $F : K \to \mathbb{R}$ be a lower semicontinuous and convex functional, $f : K \to B^*$ a hemicontinuous Z-mapping, and f satisfies the coercive condition with respect to F. Assume that there exists a nonnegative function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ such that*

(i) f is strictly α-monotone on K.
(ii) F is α-bounded on K.

If the feasible set $\mathscr{D}$ of the variational inequality problem (1.7) *is nonempty, then $\mathscr{D}$ is a $\wedge$-sublattice of B.*

Theorem 1.6 ([36]). *Let B be a reflexive Banach space and $(B, \leq)$ be a vector lattice. Let $F : K \to \mathbb{R}$ be a functional and $f : K \to B^*$ an η – Z-mapping, where $\eta : K \times K \to B$ is a mapping such that $\eta(x,y) + \eta(y,x) = 0$ for all $x, y \in K$. Assume that there exists a nonnegative function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ such that the condition* (i) *in Theorem* 10.20 *is satisfied. If the variational-like inequality problem* (1.4) *has a solution x^* in the feasible set $\mathscr{D}$, then x^* is the least element of $\mathscr{D}$.*

Corollary 1.5 ([36]). *Let $(B, \leq)$ be a vector lattice. Let $F : K \to \mathbb{R}$ be a positively homogeneous and convex functional, and $f : K \to B^*$ a Z-mapping. Assume that there exists a nonnegative function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ such that the conditions* (i) *and* (ii) *in Corollary* 10.5 *are satisfied. If the variational inequality problem* (1.7) *has a solution x^* in K, then x^* is the least element of $\mathscr{D}$.*

Now, from Theorems 1.4–1.6 we immediately obtain the following result.

Theorem 1.7 ([36]). *Let B be a reflexive Banach space, and $(B, \leq)$ a vector lattice. Assume that the following conditions are satisfied.*

(i) $F : K \to \mathbb{R}$ is a lower semicontinuous and convex functional.
(ii) $f : K \to B^*$ is an η-semicontinuous η – Z-mapping, where $\eta : K \times K \to B$ has the following properties.

(a) $\eta(x,y) + \eta(y,x) = 0$ for all $x, y \in K$.
(b) $\eta(\cdot,\cdot)$ is affine in the first variable.
(c) For each fixed $y \in K$, $x \mapsto \eta(y,x)$ is sequentially continuous from the strong topology to the weak topology.

(iii) There exists a positive number $r > 0$ such that

$$\langle \eta(x,0), f(x)\rangle + F(x) > F(0), \quad \textit{for all } x \in K \textit{ with } \|x\| = r.$$

(iv) There exists a nonnegative function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ such that

(a) f is strictly η – α-monotone on K.
(b) F is α-bounded on K.

(v) For any given $z \in K$, there exists $x^* \in K$ satisfying the following inequality.

$$\langle f(x^* + z), \eta(u, u \wedge z + x^*)\rangle \geq F(x^*) - F(u - u \wedge z), \quad \textit{for all } u \in K.$$

Then the variational-like inequality problem (1.4) *has a unique solution x^* in K. In particular, if this solution x^* lies in $\mathscr{D}$, then $\mathscr{D}$ is a $\wedge$-sublattice of B, and x^* is the least element of $\mathscr{D}$.*

Finally, from Corollaries 1.3, 1.4, and 1.5 we immediately have the following corollary.

Corollary 1.6 ([36]). *Let B be a reflexive Banach space, and $(B, \leq)$ a vector lattice. Assume that the following conditions are satisfied.*

(i) $F : K \to \mathbb{R}$ is a lower semicontinuous, positively homogeneous and convex functional.
(ii) $f : K \to B^*$ is a semicontinuous Z-mapping.
(iii) f satisfies the coercive condition with respect to F.
(iv) There exists a nonnegative function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ such that

(a) f is strictly α-monotone on K.
(b) F is α-bounded on K.

Then the variational inequality problem (1.7) *has a unique solution x^* in the feasible set $\mathscr{D}$ of itself, $\mathscr{D}$ is a $\wedge$-sublattice of B, and x^* is the least element of $\mathscr{D}$.*

1.5 Equivalence Between Extended Generalized Complementarity Problems and Generalized Least-Element Problem

In this section, we extend the formulations and results of Section 1.3 for set-valued maps.

Given is a closed convex cone $K \subseteq B$ and $T : K \to 2^{B^*}$, where 2^B is the family of all nonempty subsets of B. We denote by $\mathscr{F}$, the *feasible set* of T with respect to K; that is,

$$\mathscr{F} = \{x \in B : x \in K \text{ and } T(x) \cap K^* \neq \emptyset\}.$$

We consider the following problems.

(I) *Generalized nonlinear program*: For a given $u \in B^*$, find $x \in \mathscr{F}$ such that

$$\langle u, x \rangle = \min_{y \in \mathscr{F}} \langle u, y \rangle.$$

(II) *Generalized least-element problem*: Find $x \in \mathscr{F}$ such that

$$x \leq y, \quad \forall y \in \mathscr{F}.$$

(III) *Extended generalized complementarity problem*: Find $x \in K$ and $u \in T(x) \cap K^*$ such that

$$\langle u, x \rangle = 0.$$

(IV) *Generalized variational inequality problem*: Find $x \in K$ and $u \in T(x)$ such that

$$\langle u, y - x \rangle \geq 0, \qquad \forall y \in K.$$

The equivalence of (III) and (IV) has been studied by Saigal [29]. The main object of this section is to investigate suitable conditions under which these four problems are equivalent.

Definition 1.10. Let B be a Banach space that is also a vector lattice with positive cone K; let $T : K \to 2^{B^*}$ be a point-to-set map. Then T is called

(i) *Z-map relative to K* if for any $x, y, z \in K$,

$$\langle u - v, z \rangle \leq 0, \qquad \forall u \in T(x) \quad \text{and} \quad v \in T(y), \quad \text{whenever } (x - y) \wedge z = 0$$

(ii) *Monotone* if for any $x, y \in K$,

$$\langle u - v, x - y \rangle \geq 0, \qquad \forall u \in T(x) \quad \text{and} \quad v \in T(y)$$

(iii) *Weakly pseudomonotone* if for any $x,y \in K$ and for every $v \in T(y)$, we have $\langle v, x-y \rangle \geq 0$ implies $\langle u, x-y \rangle \geq 0$, for some $u \in T(x)$
(iv) *Weakly strictly pseudomonotone* if for any $x,y \in K$, and for every $v \in T(y)$ we have $\langle v, x-y \rangle \geq 0$ implies $\langle u, x-y \rangle > 0$, for some $u \in T(x)$
(v) *v-hemicontinuous* if for any $x,y \in K$, and $\alpha \in [0,1]$, the point-to-set map $\alpha \mapsto (T(x+\alpha(y-x)), y-x)$, is upper semicontinuous at 0^+, where $(T(x+\alpha(y-x)), y-x) = \{\langle u, y-x \rangle : u \in T(x+\alpha(y-x))\}$
(vi) *Positive at infinity* if for any $x \in K$, there exists a positive real number $\rho(x)$ such that for all $v \in T(y)$, $\langle v, y-x \rangle > 0$, $\forall y \in K$ such that $\|y\| \geq \rho(x)$

If T is a point-to-point map then the above definition (i) reduces to the definition of the Z-map relative to K in [23].

Lemma 1.4 ([1]). *Let K be a convex cone in a Banach space B and let $T : K \longrightarrow 2^{B^*}$ be a weakly (strictly) pseudomonotone point-to-set map. Then for each fixed $z \in K$, the point-to-set map $T_z : K \longrightarrow 2^{B^*}$ defined by*

$$T_z(x) = T(x+z), \qquad \forall x \in K$$

is also weakly (strictly) pseudomonotone.

Now we consider the problem (IV) for a particular set-valued map and some other problems.

Given $K \subseteq B$, $f : K \longrightarrow \mathbb{R}$ and $\phi : B \longrightarrow \mathbb{R}$.

(IV)$'$ *Generalized variational inequality problem:* Find $x \in K$ and $u \in \partial f(x)$ such that

$$\langle u, y-x \rangle \geq 0, \quad \forall y \in K.$$

(V) *Unilateral minimization problem:* Find $x \in K$ such that

$$f(x) = \min_{y \in K} f(y).$$

The set of solutions for (V) is denoted by E.

(VI) *Optimization over an efficient set:* Find $x \in K$ such that

$$\phi(x) = \min_{y \in E} \phi(y).$$

The equivalences of these problems have been investigated by Chen and Graven [5].

We need the following result to derive the equivalence of problems (I)–(IV) under suitable conditions. The proof of the following theorem follows from [23, Corollary 4.1] and [33, Lemma 3.1].

Theorem 1.8 ([1]). *Let K be a nonempty, closed, bounded convex subset of a reflexive Banach space B and let $T : K \to 2^{B^*}$ be weakly pseudomonotone, compact-valued, and v-hemicontinuous. Then there exist $x \in K$ and $u \in T(x)$ such that*

$$\langle u, y - x \rangle \geq 0, \quad \forall y \in K.$$

Furthermore, if in addition T is weakly strictly pseudomonotone, the solution is unique.

We use Theorem 1.8 to prove the existence result for perturbed generalized variational inequalities.

Proposition 1.9 ([1]). *Let K be a nonempty, closed convex cone in a reflexive Banach space B, and $T : K \to 2^{B^*}$ be weakly pseudomonotone, compact-valued, v-hemicontinuous, and positive at infinity. Then for each fixed $z \in K$, there exist $x \in K$ and $u \in T(x+z)$ such that*

$$\langle u, y - x \rangle \geq 0, \qquad \forall y \in K. \tag{1.9}$$

In the remaining part of this section, B is a real Banach space and K is a closed convex cone of B, and, whenever the ordering induced by K is mentioned, $(B, \leq)$ is a vector lattice.

We establish equivalence of problems (I)–(IV) under suitable conditions.

Proposition 1.10 ([1]). *Let $T : K \to 2^{B^*}$ be a set-valued map. Then any solution of* (II) *is a solution of* (I).

Proof. Let x be a solution of (II). Then $x \in \mathscr{F}$ such that $x \leq y$, $\forall y \in \mathscr{F}$. Because $x \in \mathscr{F}$, we have $x \in K$ and $T(x) \cap K^* \neq \emptyset$. Let $u \in T(x) \cap K^*$. Then $y - x \in K$ and $u \in K^*$ imply that

$$\langle u, y - x \rangle \geq 0 \quad \text{or} \quad \langle u, y \rangle \geq \langle u, x \rangle;$$

that is, $\langle u, x \rangle = \min_{y \in \mathscr{F}} \langle u, y \rangle$.

Proposition 1.11 ([1]). *Let B be a reflexive Banach space. Assume that $T : K \to 2^{B^*}$ is a Z-map relative to K, weakly strictly pseudomonotone, compact valued, v-hemicontinuous, and positive at infinity. Then the feasible set $\mathscr{F} = \{x \in B : x \in K \text{ and } T(x) \cap K^* \neq \emptyset\}$ is a $\wedge$-sublattice; that is, $x \in \mathscr{F}$ and $y \in \mathscr{F}$ imply $x \wedge y \in \mathscr{F}$.*

Proof. Suppose that $x, y \in \mathscr{F}$ and let $z = x \wedge y$. Because $x, y \in \mathscr{F}$, we have $x, y \in K$, $T(x) \cap K^* \neq \emptyset$, and $T(y) \cap K^* \neq \emptyset$. Let $u_1 \in T(x) \cap K^*$ and $u_2 \in T(y) \cap K^*$. Because $x, y \in K$, we have $x \geq 0$, $y \geq 0$ imply $z \geq 0$ and hence $z \in K$.

It remains to show that $T(z) \cap K^* \neq \emptyset$.

By Proposition 1.9, there exist $x^* \in K$ and $u^* \in T(x^* + z)$ such that

$$\langle u^*, w - x^* \rangle \geq 0, \qquad \forall w \in K. \tag{1.10}$$

For any $y \in K$, $y + x^* \in K$, Thus

$$\langle u^*, y + x^* - x^* \rangle \geq 0, \qquad \forall y \in K;$$

that is, $\langle u^*, y \rangle \geq 0, \forall y \in K$ and hence $u^* \in K^*$.

Now, we prove that $z = x^* + z$. For that, let $z_0 = x \wedge (x^* + z)$. Inasmuch as $x^* + z \geq z$ and $x \geq z$, we have $z_0 \geq z$ and so $z_0 - z \geq 0$ and thus $z_0 - z \in K$. Now from (1.10), we get

$$x^* \in K \quad \text{and} \quad u^* \in T(x^* + z) \quad \text{such that } \langle u^*, (z_0 - z) - x^* \rangle \geq 0. \tag{1.11}$$

Suppose that $x^* + z \neq z_0$. Then by the weakly strict pseudomonotonicity of T and (1.11), we have

$$\langle v^*, z_0 - (x^* + z) \rangle > 0 \quad \text{for some } v^* \in T(z_0). \tag{1.12}$$

$(x - z_0) \wedge (x^* + z - z_0) = (x \wedge (x^* + z)) - z_0 = 0$, and T is a Z-map relative to K, therefore we have

$$\langle u_1 - v^*, (x^* + z) - z_0 \rangle \leq 0, \qquad \forall u_1 \in T(x) \quad \text{and} \quad v^* \in T(z_0). \tag{1.13}$$

Because $u_1 \in T(x) \cap K^*$, we have $u_1 \in K^*$. Also because $(x^* + z) - z_0 \in K$, we have

$$\langle u_1, (x^* + z) - z_0 \rangle \geq 0,$$

or

$$\langle -u_1, -z_0 + (x^* + z) \rangle \leq 0. \tag{1.14}$$

Adding (1.13) and (1.14), we have

$$\langle v^*, z_0 - (x^* + z) \rangle \leq 0,$$

which is a contradiction of (1.12). Hence $z_0 = x^* + z$ and by the definition of z_0, we conclude that $x^* + z \leq x$.

Replacing the above argument with y in place of x, we can show $x^* + y \leq y$. Thus $x^* + y \leq x \wedge y = z$. But on the other hand, $x^* + z \geq z$ so $x^* + z = z$ and the proof is completed.

Proposition 1.12 ([1]). *Let B be a reflexive Banach space. Assume that $T : K \to 2^{B^*}$ is a Z-map relative to K, weakly strictly pseudomonotone, compact-valued, v-hemicontinuous, and positive at infinity. Let $u \in K^*$ be strictly positive. Then Problem* (I) *corresponding to u has at most one solution, and any solution of* (I) *is also a solution of* (II).

Proof. Suppose that $x, y \in \mathscr{F}$ are solutions of problem (I). Then by Proposition 1.11, $x \wedge y \in \mathscr{F}$ and hence

$$\langle u, x \wedge y \rangle \geq \langle u, x \rangle$$

with strict inequality of $x \wedge y \neq x$, because u is strictly positive. But x is an optimal solution, therefore the strict inequality is impossible. Hence $x \wedge y = x$. Similarly $x \wedge y = y$ and thus $x = y$ by the uniqueness of $x \wedge y$.

Suppose that x is a solution of (I) corresponding to u and let $z \in \mathscr{F}$. Then by Proposition 1.11, $z \wedge x \in \mathscr{F}$. By the optimality of x and the positivity of u, we have

$$\langle u, x \rangle \leq \langle u, z \wedge x \rangle \leq \langle u, x \rangle = \min_{y \in \mathscr{F}} \langle u, y \rangle$$

Consequently, $z \wedge x$ solves (I). By the uniqueness, $z \wedge x = x$ and so $x \leq z$, $\forall z \in \mathscr{F}$. Hence x is a solution of (II).

Proposition 1.13 ([33]). (x, u) *is a solution of* (III) *if and only if it is a solution of* (IV).

Proposition 1.14 ([1]). *Let $T : K \to 2^{B^*}$ be weakly strictly pseudomonotone and a Z-map relative to K. Then any solution of* (IV) *is a solution of* (II).

By combining Propositions 1.9 and 1.11–1.13, we have the main result of this section as follows.

Theorem 1.9 ([1]). *Let K be closed convex cone of a reflexive Banach space B such that B is a vector lattice with respect to the order $\leq$ induced by K. Let $T : K \to 2^{B^*}$ be a Z-map relative to K, weakly strictly pseudomonotone, compact-valued, v-hemicontinuous, and positive at infinity. If $u \in K^*$ is a strictly positive element, then there exists $x \in \mathscr{F}$ which is a solution of problems* (I)–(IV). *Moreover, the solution x is unique.*

We observe that Theorem 1.9 is an extension of Theorem 1.3.

Definition 1.11. A point-to-set map $T : K \to 2^{B^*}$ is called *weakly strongly pseudomonotone* if there exists $\beta > 0$ and for any $x, y \in K$, $x \neq y$, $v \in T(y)$ such that $\langle v, x - y \rangle \geq 0$ imply that

$$\langle u, x - y \rangle \geq \beta \|x - y\|^2, \quad \text{for some } u \in T(x).$$

We also see that, by the same arguments as in [23], a weakly strongly pseudomonotone operator with nonempty compact values is positive at infinity and hence the following result is a consequence of Theorem 1.9.

Corollary 1.7 ([1]). *Let K be a closed cone of a reflexive Banach space B such that B is a vector lattice with respect to the order $\leq$ induced by K. Let $T : K \to 2^{B^*}$ be a Z-map relative to K, weakly strongly pseudomonotone, compact-valued, v-hemicontinuous, and positive at infinity. If $u \in K^*$ is a strictly positive element, then there exists $x \in \mathscr{F}$ which is a solution of problems* (I)–(IV). *Moreover, the solution x is unique.*

1.6 Equivalence Between Generalized Mixed Complementarity Problems and Generalized Mixed Least-Element Problem

Given are a closed convex cone $K \subseteq B$, $A : B^* \to B^*$, $f : K \to \mathbb{R}$, and $T : K \to 2^{B^*}$. We denote by S the feasible set of T with respect to K, A, and f; that is,

$$S := \{x \in K : \langle Au, y\rangle + f(y) - f(0) \geq 0 \text{ forsome } u \in T(x) \text{ andall } y \in K\}.$$

We consider the following problems.

(I) *Generalized mixed nonlinear program:* For given $u \in B^*$, find $x \in S$ such that

$$\langle Au, x\rangle = \min_{y \in S} \langle Au, y\rangle.$$

(II) *Generalized mixed least-element problem:* Find $x \in S$ such that

$$x \leq y, \quad \text{forall } y \in S.$$

(III) *Generalized mixed complementarity problem:* Find $x \in K$ and $u \in T(x)$ such that

$$\langle Au, x\rangle + f(x) = 0 \quad \text{and} \quad \langle Au, y\rangle + f(y) \geq 0, \quad \text{for all } y \in K.$$

(IV) *Generalized mixed variational inequality problem:* Find $x \in K$ and $u \in T(x)$ such that

$$\langle Au, y - x\rangle + f(y) - f(x) \geq 0, \quad \text{for all } y \in K.$$

We remark that if $A \equiv I$ the identity operator on B^*, and $f(x) = 0$, then S reduces to the feasible set considered in the previous section. In this case, the above problems (I)–(IV) reduce, respectively, to the problems considered in Section 1.5. We further remark that, for $A \equiv I$ the identity operator on B^*, and $f(x) = 0$, the equivalence of (III) and (IV) has been studied by Saigal [29]. Moreover, whenever K is a nonempty, closed, bounded, convex subset of a reflexive Banach space B, problem (IV) has been considered in [23, 33]. In this case, the equivalence of these problems was established in the previous section.

The main objective of this section is to investigate suitable conditions under which the above problems (I)–(IV) are equivalent.

Definition 1.12. Let B be a Banach space that is also a vector lattice with positive cone K and let $T : K \to 2^{B^*}$ be a point-to-set map. Let $A : B^* \to B^*$ be a mapping, and $f : K \to \mathbb{R}$ be a functional. Then, T is called

(i) A *Z-type map relative to* K *and* A if, for any $x, y, z \in K$,

$$\langle Au - Av, z\rangle \leq 0, \quad \text{for all } u \in T(x) \text{ and } v \in T(y),$$

whenever $(x - y) \wedge z = 0$.

(ii) *Strictly a-monotone with respect to A* if there exists a function $a : \mathbb{R}_+ \to \mathbb{R}_+$ where $\mathbb{R}_+ = [0,\infty)$, such that for any $x \neq y \in K$,

$$\langle Au - Av, x-y\rangle > a(\|x-y\|), \quad \text{for all } u \in T(x) \text{and } v \in T(y);$$

in particular, if $a(t) = 0$, then T is said to be *strictly monotone with respect to A*.

(iii) *Pseudomonotone-type with respect to A and f* if for any $x, y \in K$, the existence of $u \in T(x)$ such that

$$\langle Au, y-x\rangle + f(y) - f(x) \geq 0$$

implies

$$\langle Av, y-x\rangle + f(y) - f(x) \geq 0, \quad \text{for all } v \in T(y).$$

(iv) *Strictly pseudomonotone-type with respect to A and f* if for any $x, y \in K$, the existence of $u \in T(x)$ such that

$$\langle Au, y-x\rangle + f(y) - f(x) \geq 0$$

implies

$$\langle Av, y-x\rangle + f(y) - f(x) > 0, \quad \text{for all } v \in T(y).$$

(v) υ*-hemicontinuous* if for any $x, y \in K$ and $\alpha \in [0,1]$, the point-to-set map $\alpha \mapsto (T(x+\alpha(y-x)), y-x)$ is upper semicontinuous at 0^+, where

$$(T(x+\alpha(y-x)), y-x) = \{\langle u, y-x\rangle : u \in T(x+\alpha(y-x))\}.$$

(vi) $\mathscr{H}$*-hemicontinuous* if, for any $x, y \in K$ and $\alpha \in (0,1)$, there holds

$$\mathscr{H}(T(x+\alpha(y-x)), T(x)) \to 0, \quad \text{as } \alpha \to 0^+,$$

where $\mathscr{H}$ is the Hausdorff metric defined on $CB(B^*)$.

(vii) *Quasi-positive at infinity with respect to A and f* if for any $x \in K$, there exists a positive real number $\rho(x)$ such that

$$\langle Aw, y\rangle + f(y) - f(0) > 0$$

for all $y \in K$ and $w \in T(x+y)$ with $\|y\| = \rho(x)$.

Remark 1.3.

(a) If $A \equiv I$ the identity operator on B^*, then Definition 1.12 (i) reduces to the concept of a Z-map relative to K; See Section 1.3.

(b) If $A \equiv I$ the identity operator on B^*, and T is a single-valued map of K into B^*, then Definition 1.12(ii)–(iv) and (viii) reduce to the concepts of strict a-monotonicity, pseudomonotonicity, strict pseudomonotonicity, and coercivity in [34], respectively.

Definition 1.13 ([34]). Let $a : \mathbb{R}_+ \to \mathbb{R}_+$ be a function and $f : K \to \mathbb{R}$ be a functional. Then f is said to be *a-bounded* if

$$f(x) + f(y) - f(x+y) \leq \min\{a(\|x\|), a(\|y\|)\}, \quad \text{for all } x, y \in K.$$

We need the following results to derive the equivalence of problems (I)–(IV) under suitable conditions.

Theorem 1.10 ([35, Theorem 3.1]). *If* $(\hat{x}, \hat{u})$ *is a solution of* (III) *then it is a solution of* (IV). *Conversely, if* $(\hat{x}, \hat{u})$ *is a solution of* (IV) *and* $f : K \to \mathbb{R}$ *is a functional such that* $f(\lambda x) = \lambda f(x)$ *for all* $x \in K$ *and* $\lambda > 0$, *then it is a solution of* (III).

In the rest of the section, unless otherwise specified, we assume that B is reflexive Banach space.

Theorem 1.11 ([35, Theorem 3.2]). *Let* $A : B^* \to B^*$ *be a continuous map,* $f : K \to \mathbb{R}$ *be a lower semicontinuous and convex functional, and* $T : K \to 2^{B^*}$ *be a nonempty compact-valued multifunction that is* $\mathscr{H}$*-hemicontinuous and pseudomonotone-type with respect to* A *and* f. *If there exists* $r > 0$ *such that*

$$\langle Av, y \rangle + f(y) - f(0) > 0, \quad \text{for all } y \in K \text{ and } v \in T(y) \text{ with } \|y\| = r, \tag{1.15}$$

then there exists a solution $(\hat{x}, \hat{u})$ *of* (IV). *Suppose additionally that* T *is strictly pseudomonotone-type with respect to* A *and* f*; then* $\hat{x}$ *is unique.*

Corollary 1.8 ([35, Corollary 3.1]). *Let* $A : B^* \to B^*$ *be a continuous map,* $f : K \to \mathbb{R}$ *be a lower semicontinuous and convex functional such that* $f(\lambda x) = \lambda f(x)$ *for all* $x \in K$ *and* $\lambda > 0$, *and* $T : K \to 2^{B^*}$ *be a nonempty compact-valued multifunction that is* $\mathscr{H}$*-hemicontinuous and pseudomonotone-type with respect to* A *and* f. *If there exists* $r > 0$ *such that*

$$\langle Av, y \rangle + f(y) - f(0) > 0, \quad \text{for all } y \in K \text{ and } v \in T(y) \text{ with } \|y\| = r,$$

then there exists a solution $(\hat{x}, \hat{u})$ *of* (III). *Suppose additionally that* T *is strictly pseudomonotone-type with respect to* A *and* f*; then* $\hat{x}$ *is unique.*

Corollary 1.9 ([35, Corollary 3.2]). *Let* $A : B^* \to B^*$ *be a continuous map,* $f : K \to \mathbb{R}$ *be a lower semicontinuous and convex functional, and* $T : K \to 2^{B^*}$ *be a nonempty compact-valued multifunction that is* $\mathscr{H}$*-hemicontinuous and strictly monotone with respect to* A. *If* T *is quasi-positive at infinity with respect to* A *and* f, *then for any given* $z \in K$, *there exist* $\hat{x}_z \in K$ *and* $\hat{u}_z \in T(\hat{x}_z + z)$ *such that*

$$\langle A\hat{u}_z, y - \hat{x}_z \rangle + f(y) - f(\hat{x}_z) \geq 0, \quad \text{for all } y \in K. \tag{1.16}$$

Moreover, $\hat{x}_z$ *is unique for each given* $z \in K$.

We provide the sufficient condition under which the feasible set S is a $\wedge$-sublattice of B and consider the existence of a least element of S.

Theorem 1.12 ([35, Theorem 4.1]). *Let $(B,\leq)$ be a vector lattice. Let $A: B^* \to B^*$ be a continuous map, $f: K \to \mathbb{R}$ be a lower semicontinuous and convex functional such that $f(0)=0$, and $T: K \to 2^{B^*}$ be a nonempty compact-valued multifunction that is a $\mathscr{H}$-hemicontinuous and Z-type map relative to K and A, and quasi-positive at infinity with respect to A and f. Assume that there exists a function $a: \mathbb{R}_+ \to \mathbb{R}_+$ such that*

(i) T is strictly a-monotone with respect to A.
(ii) f is a-bounded.

If the feasible set S is nonempty, then S is a $\wedge$-sublattice of B; that is, $x \in S$ and $y \in S$ imply $x \wedge y \in S$.

Definition 1.14. Let $(B,\leq)$ be a vector lattice. If there exists an element $\hat{x} \in S$ such that $\hat{x} \leq x$ for all $x \in S$, where S is the feasible set of T with respect to K, A, and f, then $\hat{x}$ is called a *least element of S*.

As pointed out in the following theorem, under suitable conditions the solvability of (III) can imply the existence of a least element of S.

Theorem 1.13 ([35, Theorem 4.2]). *Suppose that all conditions in Theorem* 1.12 *are satisfied. If $f(0)=0$, then the solvability of* (III) *implies the existence of a least element of S.*

The following main result in this section follows immediately from Theorems 1.12 and 1.13 and Corollary 1.8.

Theorem 1.14 ([35, Theorem 4.3]). *Let $(B,\leq)$ be a vector lattice. Assume that the following conditions are satisfied.*

(i) $A: B^* \to B^*$ is a continuous map.
(ii) $f: K \to \mathbb{R}$ is a lower semicontinuous and convex functional such that $f(\lambda x) = \lambda f(x)$ for all $x \in K$ and $\lambda \geq 0$.
(iii) $T: K \to 2^{B^*}$ is a nonempty compact-valued multifunction that is an $\mathscr{H}$-hemicontinuous and Z-type map relative to K and A.
(iv) T is quasi-positive at infinity with respect to A and f.
(v) There exists a function $a: \mathbb{R}_+ \to \mathbb{R}_+$ such that T is strictly a-monotone with respect to A, and such that f is a-bounded.

Then the following statements hold.

(a) There exists a solution $(\hat{x},\hat{u})$ of (III)*, and $\hat{x}$ is unique.*
(b) S is a $\wedge$-sublattice of B and there exists a least element of S, where S is the feasible set of T with respect to K, A, and f.

In the remaining part of this section, B is a real Banach space and K is a closed convex cone of B; whenever the ordering induced by K is mentioned, $(B,\leq)$ is a vector lattice.

We derive the equivalence of problems (I)–(IV) under suitable conditions.

Proposition 1.15 ([35, Proposition 5.1]). *Let $A : K^* \to B^*$ be an operator. Then any solution of* (II) *is a solution of* (I) *for every $u \in B^*$.*

Proposition 1.16 ([35, Proposition 5.2]). *Let $A : K^* \to B^*$ be an operator. Assume that all conditions in Theorem* 1.14 *are satisfied. If for $u \in B^*$, $Au \in K^*$ is strictly positive, then problem* (I) *corresponding to u has at most one solution and any solution of* (I) *is also a solution of* (II).

By combining Propositions 1.15 and 1.16 and Theorem 1.14, we have the following result.

Theorem 1.15 ([35, Theorem 5.1]). *Let K be a closed convex cone of a reflexive Banach space B with its dual B^* such that B is a vector lattice with respect to the order $\leq$ induced by K. Assume that the following conditions are satisfied.*

(i) $A : B^* \to K^*$ is a continuous map.
(ii) $f : K \to \mathbb{R}$ is a lower semicontinuous and convex functional such that $f(\lambda x) = \lambda f(x)$ for all $x \in K$ and $\lambda \geq 0$.
(iii) $T : K \to 2^{B^*}$ is a nonempty compact-valued multifunction that is an $\mathscr{H}$-hemicontinuous and Z-type map relative to K and A.
(iv) T is quasi-positive at infinity with respect to A and f.
(v) There exists a function $a : \mathbb{R}_+ \to \mathbb{R}_+$ such that T is strictly a-monotone with respect to A, and such that f is a-bounded.

If $Au \in K^*$ is strictly positive for some $u \in B^*$, then there exist $\hat{x} \in S$ and $\hat{u} \in T(\hat{x})$ such that $\hat{x}$ is a solution of problems (I) *and* (II)*, and such that $(\hat{x}, \hat{u})$ is a solution of problems* (III) *and* (IV). *Moreover, $\hat{x}$ is unique.*

Remark 1.4. We observe that Theorem 1.15 is an improvement and extension of Theorem 1.3 and Theorem 1.9.

References

1. 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).
2. Ansari, Q.H., Yao, J.C.: Iterative schemes for solving mixed variational-like inequalities. *J. Optim. Theory Appl.* **108**, 527–541 (2001).
3. Avriel, M., Diewert, W.E., Schaible, S., Zang, I.: *Generalized Concavity*. Plenum Press, New York (1988).
4. Browder, F.E.: Nonlinear monotone operators and convex sets in Banach spaces. *Bull. Amer. Math. Soc.* **71**, 780–785 (1965).
5. Chen, G.Y., Graven, B.D.: A vector variational inequality and optimization over an efficient set. *ZOR-Meth. Model. Oper. Res.* **34**, 1–12 (1990).
6. Chen, G.Y., Yang, X.Q.: The vector complementarity problem and its equivalence with the weak minimal element in ordered spaces. *J. Math. Anal. Appl.* **153**, 136–158 (1990).
7. Cottle, R.W., Pang, J.S.: On solving linear complementarity problems as linear programs. *Math. Prog. Study* **7**, 88–107 (1978).

8. Cottle, R.W., Pang, J.S.: A least element theory of solving linear complementarity problems as linear programs. *Math. Oper. Res.* **3**, 155–170 (1978).
9. Cottle, R.W., Pang, J.S., Stone, R.E.: *The Linear Complementarity Problem*. Academic Press, New York (1992).
10. Cryer, C.W., Dempster, M.A.H.: Equivalence of linear complementarity problems and linear programs in vector lattice Hilbert spaces. *SIAM J. Control Optim.* **18**, 76–90 (1980).
11. Fan, K.: A generalization of Tychonoff's fixed-point theorem. *Math. Ann.* **142**, 305–310 (1961).
12. Fang Y.-P., Huang N.-J.: On least element problems for feasible sets in vector complementarity problems. *Math. Meth. Oper. Res.* **60**, 369–377 (2004).
13. Hartman, P., Stampacchia, G.: On some non-linear elliptic differential-functional equations. *Acta Math.* **115**, 271–310 (1966).
14. Huang, N.-J., Fang, Y.-P.: Strong vector F-complementarity problem and least element problem of feasible set. *Nonlinear Anal.* **61**, 901–918 (2005).
15. Hyers, D.H., Isac, G., Rassias, T.M.: *Topics in Nonlinear Analysis and Applications*. World Scientific Press, Singapore (1997).
16. Isac, G.: *Complementarity Problems*. Lecture Notes in Mathematics, Vol. 1528, Springer-Verlag, Berlin (1992).
17. Kalashnikov, V.V.: *Complementarity Problems and Generalized Oligopoly Models*. Habilitation Thesis, Central Economics and Mathematics Institute, Moscow (in Russian) (1995).
18. Karamardian, S.: Generalized complementarity problem. *J. Optim. Theory Appl.* **8**, 161–168 (1971).
19. Karamardian, S.: Complementarity over cones with monotone and pseudomonotone maps. *J. Optim. Theory Appl.* **18**, 445–454 (1976).
20. Karamardian, S. Schaible, S.: Seven kinds of monotone maps. *J. Optim. Theory Appl.* **66**, 37–46 (1990).
21. Karamardian, S., Schaible, S., Crouzeix, J.P.: Characterizations of generalized monotone maps. *J. Optim. Theory Appl.* **76**, 399–413(1993).
22. Kelly, J.L., Namioka, I.: *Linear Topological spaces*. Springer-Verlag, New York (1976).
23. Konnov, I.V., Yao, J.C.: On the generalized vector variational inequality problem. *J. Math. Anal. Appl.* **206**, 42–58 (1997).
24. Lions, J.-L., Stampacchia, G.: Variational inequalities. *Comm. Pure Appl. Math.* **20**, 493–519 (1967).
25. Nadler Jr. S.B.: Multi-valued contraction mappings. *Pacific J. Math.* **30**, 475–488 (1969).
26. Parida, J., Sahoo, M., Kumar, A.: A variational-like inequality problem. *Bull. Austral. Math. Soc.* **39**, 225–231 (1989).
27. Rheinboldt, W.C.: On M-functions and their application to nonlinear Gauss-Seidel inetrations and network flows. *J. Math. Anal. Appl.* **32**, 274–307 (1970).
28. Riddell, R.C.: Equivalence of nonlinear complementarity problems and least-element problems in Banach lattices. *Math. Oper. Res.* **6**, 462–474 (1981).
29. Saigal, R.: Extension of the generalized complementarity problems. *Math. Oper. Res.* **1**, 260–266 (1976).
30. Schaible, S., Yao, J.C.: On the equivalence of nonlinear complementarity problems and least-element problems. *Math. Prog.* **70**, 191–200 (1995).
31. Siddiqi, A.H., Khaliq, A., Ansari, Q.H.: On variational-like inequalities. *Ann. Sci. Math. Québec* **18**, 95–104 (1994).
32. Yang, X.Q.: Vector complementarity and minimal element problems. *J. Optim. Theory Appl.* **77**, 483–495 (1993).
33. Yao, J.C.: Multi-valued variational inequalities with K-pseudomonotone operators. *J. Optim. Theory Appl.* **82**(2), 391–403 (1994).
34. Yin, H.Y., Xu, C.X., Zhang, Z.X.: The F-complementarity problem and its equivalence with the least element problem. *Acta Math. Sinica* **44**, 679–676 (2001) (in Chinese).

35. Zeng, L.C., Ansari, Q.H., Yao, J.C.: Equivalence of generalized mixed complementarity and generalized mixed least element problems in ordered spaces. *Optimization* **58**(1), 63–76 (2009).
36. Zeng, L.C., Yao, J.C.: A class of variational-like inequality problems and its equivalence with the least element problems. *J. Nonlinear Convex Anal.* **6**(2), 259–270 (2005).

Chapter 2
Generalized Monotone Maps and Complementarity Problems

S. K. Neogy and A. K. Das

Abstract In this chapter, we present some classes of generalized monotone maps and their relationship with the corresponding concepts of generalized convexity. We present results of generalized monotone maps that are used in the analysis and solution of variational inequality and complementarity problems. In addition, we obtain various characterizations and establish a connection between affine pseudomonotone mapping, affine quasimonotone mapping, positive-subdefinite matrices, generalized positive-subdefinite matrices, and the linear complementarity problem. These characterizations are useful for extending the applicability of Lemke's algorithm for solving the linear complementarity problem.

2.1 Introduction

Generalized monotonicity plays an important role in solving mathematical programming, complementarity problems, and variational inequalities. Generalized monotone maps are of fundamental importance and arise in economic applications. Different types of generalized monotonicity are related to various kinds of generalized convexity of the underlying function. It is well known [25] that a differentiable function is convex if and only if its gradient is a monotone map; see also [16]. In [27], the notion of a monotone map is generalized to that of a pseudomonotone map and a differentiable pseudoconvex function is characterized by the pseudomonotonicity of the gradient. A similar relationship exists between strictly pseudoconvex and quasiconvex functions as well as strongly convex and strongly pseudoconvex and the corresponding monotonicity property [29] of their gradient. Generalized monotone maps provide first-order characterizations of

S. K. Neogy
Indian Statistical Institute, 7, S. J. S. Sansanwal Marg, New Delhi 110 016, India,
e-mail: skn@isid.ac.in

A. K. Das
Indian Statistical Institute, 203, B. T. Road, Kolkata 700 108, India, e-mail: akdas@isical.ac.in

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_2,

generalized convex functions. During the last few decades extensive research has been devoted to generalized convexity in view of finding solutions of nonconvex optimization problems. A large number of articles have appeared on this subject where several existence results and algorithmic implications are studied. The importance of various kinds of generalized monotonicity concepts both for theory and for solution methods of variational inequalities and complementarity problems are well known. John [23, 24] presents uses of generalized concavity and generalized monotonicity in consumer theory and general equilibrium theory. In [3], application of pseudomonotone maps to economics is discussed and it is shown that the concept of pseudomonotonicity is strongly related to a notion of rationality of consumer behaviour. In this chapter, we discuss various characterizations and the role of generalized monotone maps that are used in the analysis and solution of complementarity problems.

Given a nonempty subset K of R^n and a mapping $\mathscr{F} : R^n \to R^n$, the variational inequality problem $\mathrm{VI}(K, \mathscr{F})$ is to find a vector $x^* \in K$ such that

$$(x - x^*)^t \mathscr{F}(x^*) \geq 0 \qquad \forall\, x \in K. \tag{2.1}$$

One typically assumes that the set K is closed and convex. The set K is polyhedral in many applications. When $K = R^n_+$, the nonnegative orthant of R^n, the above problem reduces to the nonlinear complementarity problem NCP$(\mathscr{F})$ which is stated as follows.

Find a vector x^* such that

$$x^* \in R^n_+, \qquad \mathscr{F}(x^*) \in R^n_+, \qquad x^{*t}\mathscr{F}(x^*) = 0. \tag{2.2}$$

For a given matrix $A \in R^{n\times n}$ and a vector $q \in R^n$ when $\mathscr{F}(x)$ is an affine function (i.e., $\mathscr{F}(x) = Ax + q$) then the problem NCP$(\mathscr{F})$ reduces to the linear complementarity problem LCP(q,A). Complementarity problems are treated as a part of mathematical programming and equilibrium problems. The complementarity problem has gained importance because it is a unified study of several optimization problems and game problems. This subject has a wide range of applications encompassing fields such as economics, control theory, engineering, game theory, and optimization. For a comprehensive survey of theory, algorithms, and applications on finite-dimensional variational inequalities and nonlinear complementarity problems, we refer the reader to the article by Harker and Pang [22].

Given a convex cone K in R^n and a mapping $\mathscr{F} : R^n \to R^n$, the generalized complementarity problem GCP$(K, \mathscr{F})$ is to find a vector $x^* \in K$ such that

$$\mathscr{F}(x^*) \in K^* \quad \text{and} \quad x^{*t}\mathscr{F}(x^*) = 0, \tag{2.3}$$

where K^* is the dual cone of K; that is $K^* = \{y \in R^n \;:\; y^t x \geq 0, \;\forall\, x \in K\}$. The feasible set of GCP$(K, \mathscr{F})$ is defined as

$$\mathrm{FEA}(K, \mathscr{F}) = \{x \in K \;:\; \mathscr{F}(x) \in K^*\}.$$

The problem GCP$(K, \mathscr{F})$ is said to be feasible if FEA$(K, \mathscr{F})$ is nonempty.

Geometrically, the problem GCP$(K,\mathscr{F})$ finds a vector $x^* \in K$ with the property that its image under the mapping $\mathscr{F}$ lies in the dual cone of K which is orthogonal to x^*. The nonnegative orthant is self-dual (i.e., $(R^n_+)^* = R^n_+$), therefore it is easy to see that GCP$(R^n_+,\mathscr{F})$ reduces to NCP$(\mathscr{F})$ as given by (2.2). Karamardian [26] obtained a relationship of the solution set between the generalized complementarity problem and variational inequality and proved that GCP$(K,\mathscr{F})$ and VI$(K,\mathscr{F})$ have the same solution set. See also [7, p. 31] and [22].

Proposition 2.1 ([26]). *Let K be a convex cone. Then $x^* \in K$ solves the problem* VI$(K,\mathscr{F})$ *if and only if x^* solves GCP$(K,\mathscr{F})$.*

Even though every generalized complementarity problem is a variational inequality problem, the converse is not true in general. The most basic result on the existence of a solution to the variational inequality problem VI$(K,\mathscr{F})$ requires the set K to be compact and convex and the mapping $\mathscr{F}$ to be continuous. See [22] and the references cited therein. The basic existence result is presented below.

Theorem 2.1. *Let K be a nonempty, compact, and convex subset of R^n and let the map $\mathscr{F} : K \to R^n$ be continuous. Then there exists a solution to the problem VI$(K,\mathscr{F})$.*

2.2 Preliminaries

We consider matrices and vectors with real entries. Let R^n_+ denote the nonnegative orthant in R^n and $R^{n\times n}$ denote the set of all $n \times n$ real matrices. For any matrix $A \in R^{m\times n}$, a_{ij} denotes its ith row and jth column entry. For any matrix $A \in R^{m\times n}$, let $A_{i\cdot}$ denote its ith row and $A_{\cdot j}$ denote its jth column. For any set $\alpha \subseteq \{1,2,\ldots,n\}$, $\bar{\alpha}$ denotes its complement in $\{1,2,\ldots,n\}$. If A is a matrix of order n, $\alpha \subseteq \{1,2,\ldots,n\}$ and $\beta \subseteq \{1,2,\ldots,n\}$, then $A_{\alpha\beta}$ denotes the submatrix of A consisting of only the rows and columns of A whose indices are in α and β, respectively. Any vector $x \in R^n$ is a column vector unless otherwise specified and x^t denotes the row transpose of x. For any matrix $A \in R^{n\times n}$, A^t denotes its transpose. We say that a vector $y \in R^n$ is *unisigned* if either $y \in R^n_+$ or $-y \in R^n_+$. Given a symmetric matrix $S \in R^{n\times n}$, its *inertia* is the triple $(\nu_+(S), \nu_-(S), \nu_0(S))$ where $\nu_+(S), \nu_-(S)$, and $\nu_0(S)$ denote the number of positive, negative, and zero eigenvalues of S, respectively. Given $x \in R^n$, x^+ and x^- are the vectors of R^n defined by $x_i^+ := \max\{x_i, 0\}$ and $x_i^- := \max\{-x_i, 0\}$ $\forall\, i$. Clearly, $x = x^+ - x^-$. A cone is said to be pointed if $K \cap (-K) = \{0\}$. A cone is said to be solid if its interior is nonempty. Given $\Omega \subseteq R^n$, we denote the interior of Ω by int(Ω).

Let $\Omega \subseteq R^n$ be a convex set and $f : \Omega \to R$. Different kinds of generalized convexity were established in the literature by retaining some of the properties of convex functions and a large number of articles have appeared on this subject. Each type of generalized monotone map is related to a generalized convex function. We recall the definitions of generalized convex functions and review some of the characterizations from the literature [2, 33] which are needed for further discussions.

Definition 2.1. f is said to be

(i) *Convex* on Ω if for all $x,y \in \Omega$, and $0 \leq \lambda \leq 1$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$
(ii) *Strictly convex* on Ω if for all $x,y \in \Omega$, $x \neq y$ and $0 < \lambda < 1$, $f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$
(iii) *Quasiconvex* on Ω if for all $x,y \in \Omega$ and $0 \leq \lambda \leq 1$, $f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}$
(iv) *Strictly quasiconvex* on Ω if for all $x,y \in \Omega$ and $0 < \lambda < 1$, $f(x) < f(y) \Rightarrow f(\lambda x + (1-\lambda)y) < f(y)$

We assume differentiability of the function f on the open convex set $\Omega \subseteq R^n$ for providing the definition of following generalized convex functions.

Definition 2.2. f is said to be

(i) *A pseudoconvex* function if for all $x,y \in \Omega$, $(y-x)^t \nabla f(x) \geq 0 \Rightarrow f(y) \geq f(x)$.
(ii) *Strictly pseudoconvex* if for all $x,y \in \Omega$, $x \neq y$, $(y-x)^t \nabla f(x) \geq 0 \Rightarrow f(y) > f(x)$.

Theorem 2.2 ([2, 3.5.11 Theorem, p. 143]). *Let $\Omega \subseteq R^n$ be a nonempty open convex set and $f : \Omega \to R$ be a differentiable pseudoconvex function. Then f is both strictly quasiconvex and quasiconvex.*

Theorem 2.3 ([33, p. 134, 146], [2, p. 137]). *Let $\Omega \subseteq R^n$ be a nonempty open convex set and f be a differentiable function defined on Ω. Then f is quasiconvex if and only if for all $x,y \in \Omega$, either one of the following statements holds true.*

$$(y-x)^t \nabla f(x) > 0 \Rightarrow f(y) > f(x). \tag{2.4}$$

$$f(y) \leq f(x) \Rightarrow (y-x)^t \nabla f(x) \leq 0. \tag{2.5}$$

2.3 Different Types of Generalized Monotone Maps

Various kinds of generalized monotonicity concepts have been introduced in the literature. A large number of publications have appeared which deal with the concepts and characterizations of generalized monotonicity for different subclasses of maps. In this chapter, we present a brief review of basic generalized monotonicity concepts which are needed for presentation of the results that deal with variational inequalities and in particular complementarity problems. We now recall the following definitions from [27, 29].

Definition 2.3. Let $\Omega \subset R^n$ and $\mathscr{F} : \Omega \to R^n$. $\mathscr{F}$ is said to be

(i) *Monotone* on Ω if $x,y \in \Omega \Rightarrow (y-x)^t(\mathscr{F}(y) - \mathscr{F}(x)) \geq 0$
(ii) *Strictly monotone* on Ω if $x,y \in \Omega$, $x \neq y \Rightarrow (y-x)^t(\mathscr{F}(y) - \mathscr{F}(x)) > 0$
(iii) *Pseudomonotone* on Ω if $x,y \in \Omega$, $(y-x)^t \mathscr{F}(x) \geq 0 \Rightarrow (y-x)^t \mathscr{F}(y) \geq 0$

(iv) *Strictly pseudomonotone* on Ω if $x,y \in \Omega,\ x \neq y,\ \ (y-x)^t\mathscr{F}(x) \geq 0 \Rightarrow (y-x)^t\mathscr{F}(y) > 0$

(v) *Quasimonotone* on Ω if $x,y \in \Omega$, $(y-x)^t\mathscr{F}(x) > 0 \Rightarrow (y-x)^t\mathscr{F}(y) \geq 0$

The following statement follows from the above definitions.

$$\text{Monotonicity} \Rightarrow \text{pseudomonotonicity} \Rightarrow \text{quasimonotonicity.}$$

For further details on different kinds of generalized monotonicity and their relationship, see [21, 29]. The following lemma is useful.

Lemma 2.1 ([27, Lemma 3.1. p. 449]). *Let $\Omega \subset R^n$ and $\mathscr{F} : \Omega \to R^n$ be pseudomonotone on Ω. Then for every $x,y \in \Omega$, we have*

$$(y-x)^t\mathscr{F}(x) > 0 \Rightarrow (y-x)^t\mathscr{F}(y) > 0.$$

The following theorem establishes the equivalence of convexity of a function and monotonicity of its gradient.

Theorem 2.4 ([29]). *Let $\Omega \subset R^n$ and $f : \Omega \to R$ be a differentiable function on an open convex set Ω of R^n. Then f is convex (strictly convex) if and only if $\mathscr{F} = \nabla f$ is monotone (strictly monotone) on Ω.*

The following theorem generalizes a well-known result of a convex mathematical program (the solution set of a convex mathematical program is convex) to a variational inequality problem with a pseudomonotone type of map. See [22] and references cited therein.

Theorem 2.5. *Let K be a nonempty, closed, and convex subset of R^n and the map $\mathscr{F} : K \to R^n$ be continuous and pseudomonotone from $K \to R^n$. Then x^* solves the problem VI$(K,\mathscr{F})$ if and only if $x^* \in K$ and*

$$(y-x)^t\mathscr{F}(y) \geq 0\ \forall\, y \in K.$$

In particular, the solution set of VI$(K,\mathscr{F})$ *is convex if it is nonempty.*

Even though a variational inequality problem can have more than one solution, if we assume $\mathscr{F}$ to be strictly monotone on K then VI$(K,\mathscr{F})$ can have at most one solution. We now state the following existence theorem for the generalized complementarity problem GCP$(K,\mathscr{F})$.

Theorem 2.6. *Let K be a solid, pointed, closed, convex cone in R^n. If $\mathscr{F}$ is continuous and strictly monotone with respect to K and if the GCP$(K,\mathscr{F})$ is feasible, then the GCP$(K,\mathscr{F})$ has a unique solution.*

Karamardian and Schaible [29] studied different generalizations of monotone maps where different generalizations of monotonicity correspond to some kind of generalized convexity of the function f. For the gradient map $\mathscr{F} = \nabla f$, the following result is observed in [27, 29]. We present the proof for pseudomonotone maps along the same lines of Karamardian [27]. For quasimonotone maps, the proof technique is similar and we refer the reader to the article of Karamardian and Schaible [29].

Theorem 2.7. *Let $\Omega \subset R^n$ be an open convex set and $f : \Omega \to R$ be differentiable on Ω. Then f is pseudoconvex if and only if $\mathscr{F} = \nabla f$ is pseudomonotone on Ω.*

Proof. Suppose that f is pseudoconvex. Let $x, y \in \Omega$ such that $(y-x)^t \nabla f(x) \geq 0$. From the definition of pseudoconvexity, it follows that $f(y) \geq f(x)$. By Theorem 2.2, pseudoconvexity implies quasiconvexity. Consequently, $f(y) \geq f(x) \Rightarrow (y-x)^t \nabla f(y) \geq 0$. Therefore ∇f is pseudomonotone on Ω.

To prove the converse suppose that ∇f is pseudomonotone on Ω. Let $x, y \in \Omega$, $x \neq y$ such that

$$(y-x)^t \nabla f(x) \geq 0. \tag{2.6}$$

To show f is pseudoconvex, we need to show that $f(y) \geq f(x)$.

Assume to the contrary that

$$f(y) < f(x). \tag{2.7}$$

From the mean value theorem, we have

$$f(y) - f(x) = (y-x)^t \nabla f(\bar{x}), \tag{2.8}$$

where

$$\bar{x} = \bar{\lambda} x + (1-\bar{\lambda}) y \tag{2.9}$$

for some $0 < \bar{\lambda} < 1$. From (2.7), (2.8), and (2.9), we get

$$(x-\bar{x})^t \nabla f(\bar{x}) > 0. \tag{2.10}$$

From (2.10) and Lemma 2.1, we get

$$(x-\bar{x})^t \nabla f(x) > 0.$$

From (2.9), this implies

$$(x-y)^t \nabla f(x) > 0.$$

However, this contradicts (2.6). This completes the proof. ∎

Theorem 2.8. *Let $\Omega \subset R^n$ be an open convex set and $f : \Omega \to R$ be differentiable on Ω. Then f is quasiconvex if and only if $\mathscr{F} = \nabla f$ is quasimonotone on Ω.*

Proof. Suppose that f is quasiconvex. Let $x, y \in \Omega$ such that $(y-x)^t \nabla f(x) > 0$. From (2.4), it follows that $f(y) > f(x)$. Now by (2.5), $f(x) < f(y) \Rightarrow (x-y)^t \nabla f(y) \leq 0 \Rightarrow (y-x)^t \nabla f(y) \geq 0$. Therefore ∇f is quasimonotone on Ω.

For the converse part of theorem, the argument is similar to the earlier one. ∎

Karamardian, Schaible, and Crouzeix [30] obtained first-order necessary and sufficient conditions for a map to be pseudomonotone or quasimonotone. Let $\Omega \subset R^n$ be an open convex set and $\mathscr{F} : \Omega \to R^n$ be differentiable with Jacobian matrix $J_{\mathscr{F}}(x)$ evaluated at $x \in \Omega$. Let the projection of $\mathscr{F}$ on v be defined by $\psi : I_{x,v} \to R$, $v \in R^n$ where

$$\psi_{x,v}(t) = v^t \mathscr{F}(x+tv), \qquad I_{x,v} = \{t \mid x+tv \in \Omega\}. \tag{2.11}$$

Theorem 2.9 ([30, Theorem 4.1, p. 404]). *Let $\mathscr{F} : \Omega \to R^n$ be differentiable on the open convex set $\Omega \subset R^n$. Then*

(i) $\mathscr{F}$ is quasimonotone on Ω if and only if

$$v^t \mathscr{F}(x) = 0 \Rightarrow v^t J_{\mathscr{F}}(x)v \geq 0; \tag{2.12}$$

$$v^t \mathscr{F}(x) = v^t J_{\mathscr{F}}(x)v = 0, \qquad \hat{t} < 0, v^t \mathscr{F}(x+\hat{t}v) > 0 \Rightarrow \exists\, \tilde{t} > 0, \qquad \tilde{t} \in I_{x,v} \tag{2.13}$$

$$\textit{such that } v^t \mathscr{F}(x+tv) \geq 0\ \forall\, 0 \leq t \leq \tilde{t}.$$

(ii) $\mathscr{F}$ is pseudomonotone on Ω if and only if

$$\begin{gathered} v^t \mathscr{F}(x) = 0 \Rightarrow v^t J_{\mathscr{F}}(x)v \geq 0; \\ v^t \mathscr{F}(x) = v^t J_{\mathscr{F}}(x)v = 0 \ \Rightarrow\ \exists\, \tilde{t} > 0, \qquad \tilde{t} \in I_{x,v} \\ \textit{such that } v^t \mathscr{F}(x+tv) \geq 0\ \forall\, 0 \leq t \leq \tilde{t}. \end{gathered} \tag{2.14}$$

Karamardian, Schaible, and Crouzeix [30] also obtained somewhat different sufficient conditions for a map to be pseudomonotone and strictly pseudomonotone. For the proof we refer the reader to [30].

Theorem 2.10 ([30, Theorem 4.2, p. 406]). *Let $\mathscr{F} : \Omega \to R^n$ be differentiable on the open convex set $\Omega \subset R^n$. Then*

(i) $\mathscr{F}$ is pseudomonotone on Ω if for every $x \in \Omega$ and $v \in R^n$, we have

$$\begin{gathered} v^t \mathscr{F}(x) = 0 \Rightarrow v^t J_{\mathscr{F}}(x)v \geq 0; \\ v^t \mathscr{F}(x) = v^t J_{\mathscr{F}}(x)v = 0 \Rightarrow\ \exists\, \varepsilon > 0 \quad \textit{such that } v^t J_{\mathscr{F}}(x+tv)v \geq 0 \\ \forall\, t \in I_{x,v}, \qquad |t| \leq \varepsilon. \end{gathered} \tag{2.15}$$

(ii) $\mathscr{F}$ is strictly pseudomonotone on Ω if for every $x \in \Omega$ and $v \in R^n$

$$v^t \mathscr{F}(x) = 0 \Rightarrow v^t J_{\mathscr{F}}(x)v \geq 0.$$

For the proof, we refer the reader to [30]. Karamardian, Schaible, and Crouzeix [30] also presented an example from [1] to show that the sufficient condition (2.15) in Theorem 2.10 is not a necessary condition.

2.4 Generalized Monotonicity of Affine Maps

Generalized monotone affine maps have been considered in [9–11, 18, 19, 35, 37, 39]. The special case of the map $\mathscr{F} : \Omega \to R^n$ is an affine map

$$\mathscr{F}(x) = Ax + q,$$

where $A \in R^{n\times n}$ and $q \in R^n$. Special cases of Theorem 2.9 for an affine map $\mathscr{F} : \Omega \to R^n$ are given below.

Theorem 2.11 ([30, Theorem 5.1, p. 408]). *Let $\Omega \subset R^n$ be open and convex. Then for an affine map $\mathscr{F} : \Omega \to R^n$ of the form $\mathscr{F}(x) = Ax + q$ is quasimonotone if and only if it is pseudomonotone on Ω.*

Proof. Note that pseudomonotonicity $\Rightarrow$ quasimonotonicity. Therefore we need to show that quasimonotonicity imples pseudomonotonicity. The Jacobian matrix $J_{\mathscr{F}}(x) = A$ is independent of x. Note that by part (i) of Theorem 2.9 the condition (2.12)

$$v^t\mathscr{F}(x) = 0 \Rightarrow v^t J_{\mathscr{F}}(x)v \geq 0$$

holds. This reduces to $v^t(Ax+q) = 0 \Rightarrow v^tAv \geq 0$.

The function $\psi : I_{x,v} \to R,\ v \in R^n$, which is a linear function of t, is given by

$$\psi(t) = (v^tAv)t + v^t(Ax+q).$$

It is easy to see that the condition (2.14) is always satisfied because

$$v^t(Ax+q) = v^tAv = 0 \Rightarrow v^t[A(x+tv)+q] = 0,\ \forall\, t.$$

Now by part (ii) of Theorem 2.9, it follows that $\mathscr{F}$ is pseudomonotone. ■

Theorem 2.11 is not true if Ω is not open. Karamardian, Schaible, and Crouzeix [30] provide an example. For a continuous map $\mathscr{F}$ (not necessarily affine), Crouzeix and Schaible [11] observe the following result which uses the nonemptiness of $int(\Omega)$.

Theorem 2.12. *Assume that $\mathscr{F} : \Omega \to R^n$ is continuous on Ω and quasimonotone on $int(\Omega)$. Then it is also quasimonotone on Ω.*

Theorem 2.13 ([30, Theorem 5.2, p. 409]). *Let $\Omega \subset R^n$ be open and convex. Then for an affine map $\mathscr{F} : \Omega \to R^n$ where $\mathscr{F}(x) = Ax + q$ is pseudomonotone if and only if for every $x \in \Omega$ and $v \in R^n$ we have*

$$v^t(Ax+q) = 0 \Rightarrow v^tAv \geq 0.$$

Proof. It is easy to see that the condition (2.14) is always satisfied inasmuch as

$$v^t(Ax+q) = v^tAv = 0 \Rightarrow v^t[A(x+tv)+q] = 0,\ \forall\, t.$$

From Theorem 2.9, the result follows. ■

Karamardian, Schaible, and Crouzeix [30] observe that the sufficiency part of the above theorem remains valid for an arbitrary convex set Ω (not necessarily open).

2.5 Generalized Monotone Affine Maps on R^n_+ and Positive-Subdefinite Matrices

Generalized monotone affine maps arise in linear complementarity problems. In [9], Crouzeix et al. obtained new characterizations of generalized monotone affine maps on R^n_+ using positive subdefinite matrices. A is called a *positive-subdefinite matrix* if for all $x \in R^n$, $x^tAx < 0$ implies A^tx is unisigned. The class of positive-subdefinite matrices (PSBD) is a generalization of the class of positive-semidefinite (PSD) matrices. The study of pseudoconvex and quasiconvex quadratic forms leads to this new class of matrices, and it is useful in the study of quadratic programming problem. The class of symmetric positive-subdefinite matrices was introduced by Martos [34] in connection with a characterization of a pseudoconvex function. Martos did an interesting study of these matrices. Cottle and Ferland [5] followed the path set by Martos in [34] and among other things, obtained converses for some of Martos's results. Rao [40] obtained a characterization of merely positive-subdefinite matrices which enabled the easy recognition of quasiconvex and pseudoconvex quadratic forms. He also studied this class with respect to generalized inverse (g-inverse). Martos was considering the Hessians of quadratic functions, therefore he was concerned only about symmetric matrices. Later Crouzeix et al. [9] and Mohan, Neogy, and Das [35] studied nonsymmetric PSBD matrices in the context of generalized monotonicity and the linear complementarity problem. In this section characterizations of generalized monotone affine maps on R^n_+ using positive-subdefinite matrices, the properties of PSBD matrices, and their applications to linear complementarity problem are presented. It is not surprising that many properties of PSD matrices are lost through the generalization. It is useful to review some matrix classes and their properties which form the basis for further discussions.

The *linear complementarity problem* is a fundamental problem that arises in optimization, game theory, economics, and engineering. It can be stated as follows.

Given a matrix $A \in R^{n\times n}$ and a vector $q \in R^n$, the linear complementarity problem is to find a vector $x \in R^n$ such that

$$x \geq 0, \qquad Ax + q \geq 0, \tag{2.16}$$

$$x^t(Ax+q) = 0. \tag{2.17}$$

This problem is denoted as LCP(q,A). The LCP is normally identified as a problem of mathematical programming and provides a unifying framework for several optimization problems such as linear programming, linear fractional programming, convex quadratic programming, and bimatrix game problems. More specifically, the LCP models the optimality conditions of these problems. The early motivation for studying the linear complementarity problem was that the KKT optimality conditions for linear and quadratic programs reduce to an LCP. The algorithm presented by Lemke and Howson [32] to compute an equilibrium pair of strategies to a bimatrix game, later extended by Lemke [5] (known as Lemke's algorithm) to solve an LCP(q,A), contributed significantly to the development of linear complementarity theory. In fact, the study of the LCP really came into prominence only when Lemke and Howson [32] and Lemke [5] showed that the problem of computing a Nash equilibrium point of a bimatrix game can be posed as an LCP following the publication by Cottle [4]. However, Lemke's algorithm does not solve every instance of the linear complementarity problem, and in some instances of the problem may terminate inconclusively without either computing a solution to it or showing that no solution to it exists. Extending the applicability of Lemke's algorithm to more matrix classes has been considered by many researchers including Eaves [12, 13], Garcia [17], Karamardian [28], and Todd [41]. For recent books on the linear complementarity problem and its applications, see Cottle, Pang, and Stone [7], Murty [36], and Facchinei and Pang [15]. Matrix classes play an important role in studying the theory and algorithms of LCP. The study of special properties of the data matrix A has historically been an important part of LCP research. A variety of classes of matrices is introduced in the context of the linear complementarity problem. Many of the matrix classes encountered in the context of the LCP are commonly found in several applications. Some of these matrix classes are of interest because they characterize certain properties of the LCP and they offer certain nice features from the viewpoint of algorithms. Several algorithms have been designed for the solution of the linear complementarity problem. Many of these methods are matrix class dependent. They work only for LCPs with some special classes of matrices and can give no information otherwise.

It is well known that the positive-semidefiniteness of a matrix A is equivalent to the monotonicity of the affine mapping $\mathscr{F}(x) = Ax + q$, where $A \in R^{n\times n}$ and $q \in R^n$. The class of PSD matrices is a subclass of positive-subdefinite matrices. Let A be a given $n \times n$ matrix, not necessarily symmetric.

Definition 2.4. We say that a real square matrix A of order n is *positive subdefinite* (*PSBD*) if for all $x \in R^n$

$$x^tAx < 0 \text{ implies either } A^t x \leq 0 \text{ or } A^t x \geq 0.$$

A is said to be *merely positive-subdefinite* (*MPSBD*) if A is a PSBD matrix but not positive-semidefinite (PSD).

Definition 2.5. A is said to be a $P\,(P_0)$-*matrix* if all its principal minors are positive (nonnegative).

A subclass of P_0 occurs in Markov chain analysis and in the study of global univalence in economic theory [38].

Definition 2.6. A is called *copositive* (C_0) (*strictly copositive* (C)) if $x^tAx \geq 0\ \forall\, x \geq 0$ $(x^tAx > 0\ \forall\, 0 \neq x \geq 0)$. $A \in R^{n\times n}$ is said to be *conegative* if $x^tAx \leq 0\ \forall\, x \geq 0$.

A is said to be *copositive-plus* (C_0^+) if $A \in C_0$ and the following implication holds.

$$[x^t Ax = 0, x \geq 0] \Rightarrow (A + A^t)x = 0.$$

We say that $A \in R^{n\times n}$ is *copositive-star* (C_0^*) if $A \in C_0$ and the following implication holds.

$$[x^t Ax = 0, Ax \geq 0, x \geq 0] \Rightarrow A^t x \leq 0.$$

A is called *copositive (strictly copositive, copositive-plus, PSD, PD)* of order k, $0 \leq k \leq n$, if every principal submatrix of order k is copositive (strictly copositive, copositive-plus, PSD, PD).

Definition 2.7. A is said to be *column sufficient* if for all $x \in R^n$ the following implication holds.

$$x_i(Ax)_i \leq 0\ \forall\ i\ \Rightarrow\ x_i(Ax)_i = 0\ \forall\ i.$$

A is said to be *row sufficient* if A^t is column sufficient.

A is *sufficient* if A and A^t are both column sufficient.

A matrix A is *sufficient of order k* if all its $k \times k$ principal submatrices are sufficient.

For details on sufficient matrices, see [6, 8, 43].

Definition 2.8. $A \in R^{n\times n}$ is called a *Q-matrix* (or a *matrix satisfying the Q-property*) if for every $q \in R^n$, LCP(q,A) has a solution.

Given a matrix $A \in R^{n\times n}$ and a vector $q \in R^n$ we define the feasible set $F(q,A) = \{x \geq 0 \mid Ax + q \geq 0\}$ and the solution set of LCP(q,A) by $S(q,A) = \{x \in F(q,A) \mid x^t(Ax+q) = 0\}$. We say that A is a *Q_0-matrix* if $F(q,A) \neq \emptyset$ implies $S(q,A) \neq \emptyset$.

A is said to be a *completely Q (Q_0)-matrix* if all its principal submatrices are Q (Q_0)-matrices.

We recall that given a matrix $A \in R^{n\times n}$ and a vector $q \in R^n$, an affine map $\mathscr{F}(x) = Ax + q$ is said to be *pseudomonotone* on R^n_+ if

$$(y-x)^t(Ax+q) \geq 0, \qquad y \geq 0, \qquad x \geq 0 \Rightarrow (y-x)^t(Ay+q) \geq 0.$$

Given $A \in R^{n\times n}$ and $q \in R^n$, Crouzeix et al. [9] prove the following necessary and sufficient condition for an affine map $\mathscr{F}(x) = Ax + q$ to be pseudomonotone on R^n_+.

Proposition 2.2 ([9]). *An affine map $\mathscr{F}$ is pseudomonotone on R^n_+ if and only if*

$$x \in R^n,\ x^t Ax < 0 \Rightarrow \begin{cases} A^t x \geq 0 \text{ and } x^t q \geq 0 & \text{or} \\ A^t x \leq 0,\ x^t q \leq 0 & \text{and } x^t(Ax^- + q) < 0. \end{cases}$$

A necessary and sufficient condition for an affine map $\mathscr{F}(x) = Ax + q$ to be quasimonotone on R^n_+ is given below.

Proposition 2.3 ([9]). *An affine map $\mathscr{F}$ is quasimonotone on R^n_+ if and only if*

$$x \in R^n,\ x^tAx < 0 \Rightarrow \begin{cases} A^tx \geq 0 & \text{and } x^tq \geq 0 \text{ or} \\ A^tx \leq 0 & \text{and } x^tq \leq 0. \end{cases}$$

The above proposition shows that A is PSBD when $\mathscr{F}$ is quasimonotone (a fortiori, pseudomonotone) on R^n_+.

Theorem 2.14 ([9]). *Assume that $\mathscr{F}$ is quasimonotone on R^n_+ and $q \neq 0$. Then $\mathscr{F}$ is pseudomonotone on R^n_+.*

Definition 2.9. We say that a matrix $A \in R^{n\times n}$ is *pseudomonotone* if $\mathscr{F}(x) = Ax$ is pseudomonotone on the nonnegative orthant.

Theorem 2.15 ([18, Corollary 4]). *If A is pseudomonotone, then A is a row sufficient matrix.*

A row sufficient matrix belongs to Q_0 [7, p. 159], therefore LCP(q,A) is solvable by Lemke's algorithm where A is a pseudomonotone matrix. Gowda [19] showed with an example that the transpose of a pseudomonotone matrix need not be in Q_0 and hence need not be pseudomonotone. However, if A is pseudomonotone then under certain conditions only A^t is a Q_0-matrix. These conditions are stated below in the following theorem.

Theorem 2.16 ([19]). *Suppose that $A \in R^{n\times n}$ is pseudomonotone. Then under each of the following conditions A^t satisfies the copositive star property and hence belongs to Q_0.*

(i) The diagonals of A consist only of zeros.
(ii) The system $0 \neq d \geq 0, A^td = 0$ has no solution.
(iii) A is invertible.
(iv) $A \in R_0$.
(v) A is normal (i.e., $AA^t = A^tA$).

Gowda [18] observes the following results. For the proofs of these results we refer the reader to the article [18] by Gowda.

Theorem 2.17. *Suppose that LCP(q,A) is feasible and the map $\mathscr{F}(x) = Ax + q$ is pseudomonotone. Then $A \in C_0 \cap P_0$.*

Gowda [18] provides the following example to show that the above result may not hold if the feasibility condition is dropped. Let

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad q = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Then the map $\mathscr{F}(x) = Ax + q$ is pseudomonotone but $A \notin C_0 \cap P_0$. For details see [18].

The following result is a corollary of the above theorem.

Corollary 2.1 ([18]). *Suppose that A is symmetric. Then the pseudomonotone mapping $\mathscr{F}(x) = Ax + q$ is monotone (i.e., A is positive semidefinite (PSD)) if and only if A is copositive.*

The following theorem in [18] shows that the pseudomonotonicity can be described in terms of a single variable in R^n.

Theorem 2.18 ([18, p. 375]). *For the pair (q,A), let*

(i) $\mathscr{A} := \{x : (A^t x)_i > 0$ *for some* i *and* $x^t(Ax^- + q) \leq 0\}$.
(ii) $\mathscr{B} := \{x : (A^t x)_i < 0$ *for some* i *and* $x^t(Ax^- + q) \geq 0\}$.
(iii) $\mathscr{C} := \{x : x^t(Ax^- + q) \geq 0\}$, $\mathscr{D} := \{x : x^t(Ax^+ + q) \geq 0\}$.

The mapping $\mathscr{F}(x) = Ax + q$ is pseudomonotone if and only if

(a) $x^t Ax \geq 0 \; \forall \, x \in \mathscr{A} \cup \mathscr{B}$.
(b) $\mathscr{C} \subseteq \mathscr{D}$.

It is easy to see that if A is PSD then for any q, mapping $\mathscr{F}(x) = Ax + q$ is pseudomonotone. Gowda [18] proves that even the converse is true.

Corollary 2.2 ([18, p. 376]). *$A \in R^{n \times n}$ is PSD if and only if for every q, the mapping $\mathscr{F}(x) = Ax + q$ is pseudomonotone.*

Theorem 2.19 ([18]). *Suppose that $A \in R^{n \times n}$ has no zero column. If the map $\mathscr{F}(x) = Ax + q$ is pseudomonotone and LCP(q,A) is feasible, then A is pseudomonotone.*

Theorem 2.20 ([18]). *Suppose that A is pseudomonotone. Then $A \in P_0 \cap Q_0$ and every feasible LCP(q,A) is solvable (by Lemke's algorithm).*

The following result follows immediately from the above two theorems.

Theorem 2.21 ([18]). *Suppose that $A \in R^{n \times n}$ has no zero column. If LCP(q,A) is feasible and the map $\mathscr{F}(x) = Ax + q$ is pseudomonotone, then $A \in P_0 \cap Q_0$ and every feasible LCP(q',A) is solvable (by Lemke's algorithm).*

Theorem 2.22 ([18]). *Suppose that the map $\mathscr{F}(x) = Ax + q$ is pseudomonotone and LCP(q,A) is feasible. Then LCP(q,A) is solvable (by Lemke's algorithm).*

By presenting the following example Gowda [18] shows that stronger conclusions in the above theorem are not possible. For other values q', feasibility of LCP(q',A) does not necessarily imply solvability. Let

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad q = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad q' = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Then the map $\mathscr{F}(x) = Ax + q$ is pseudomonotone. LCP(q,A) and LCP(q',A) are feasible but LCP(q',A) is not solvable. Therefore, $A \notin Q_0$.

The following theorem relates the concept of pseudomonotonicity of a matrix A to the class of PSBD matrices.

Theorem 2.23 ([9, Theorem 3.3]). *$A \in R^{n\times n}$ is pseudomonotone if and only if A is PSBD and copositive with the additional condition in the case where $A = ab^t$ that $b_i = 0 \Rightarrow a_i = 0$.*

In fact, the class of pseudomonotone matrices coincides with the class of matrices which are both PSBD and C_0^*. For more details on pseudomonotone and C_0^* matrices see [20].

In general PSBD matrices need not be P_0 or Q_0. We provide the following example.

Example 2.1. Suppose

$$A = \begin{bmatrix} 0 & -3 \\ -5 & 0 \end{bmatrix} \quad \text{and} \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

We show that A is PSBD by showing that for all $x \in R^n$,

$$x^t Ax < 0 \quad \text{implies either} \quad A^t x \leq 0 \text{ or } A^t x \geq 0.$$

Then $x^t Ax = -8x_1x_2 < 0$ implies x_1 and x_2 are of same sign. Clearly $A \in$ PSBD because

$$A^t x = \begin{bmatrix} -5x_2 \\ -3x_1 \end{bmatrix}$$

implies either $A^t x \leq 0$ or $A^t x \geq 0$ but $A \notin P_0$.

The following example shows that PSBD matrices need not be Q_0 in general.

Example 2.2. Let

$$A = \begin{bmatrix} 1 & 0 \\ 4 & 0 \end{bmatrix} \quad \text{and} \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

But $x^t Ax = x_1^2 + 4x_1x_2 < 0$ implies x_1 and x_2 are of different sign. Clearly $A \in$ PSBD because

$$A^t x = \begin{bmatrix} x_1 + 4x_2 \\ 0 \end{bmatrix}$$

implies either $A^t x \leq 0$ or $A^t x \geq 0$. Taking

$$q = \begin{bmatrix} -1 \\ -8 \end{bmatrix}$$

we note that LCP(q,A) is feasible but the problem has no complementary solution. Therefore A is not a Q_0 matrix.

Theorem 2.24 ([9, Proposition 2.1]). *Let $A = ab^t$ with $a,b \in R^n$, $a,b \neq 0$.*
A is PSBD if and only if one of the following conditions holds.

(i) $\exists$ a $t > 0$ such that $b = ta$.
(ii) For all $t > 0$, $b \neq ta$ and either $b \geq 0$ or $b \leq 0$.

Further suppose that $A \in MPSBD$. Then $A \in C_0$ if and only if either ($a \geq 0$ and $b \geq 0$) or ($a \leq 0$ and $b \leq 0$) and $A \in C_0^*$ if and only if A is copositive and $a_i = 0$ whenever $b_i = 0$.

Gowda [18] conjectured that pseudomonotonicity of a matrix A implies pseudomonotonicity of a matrix A^t. Obviously, Gowda's conjecture is true when A is PSD. Crouzeix et al. [9] show that the conjecture is also true when rank$(A) \geq 2$ but it is not true for matrices of rank1.

Combining [9, Theorem 2.1] and [9, Proposition 2.5], we get the following theorem on PSBD matrices.

Theorem 2.25 ([35]). *Suppose $A \in R^{n\times n}$ is PSBD and rank$(A) \geq 2$. Then A^t is PSBD and at least one of the following conditions holds.*

(i) A is PSD.
(ii) $(A + A^t) \leq 0$.
(iii) A is C_0^*.

Theorem 2.26 ([9, Proposition 2.2]). *Assume that $A \in R^{n\times n}$ is MPSBD and rank$(A) \geq 2$. Then*

(a) $\nu_-(A + A^t) = 1$.
(b) $(A + A^t)x = 0 \ \Leftrightarrow \ Ax = A^t x = 0$.

Because a PSBD matrix is a natural generalization of a PSD matrix, it is of interest to determine which of the properties of a PSD matrix also holds for a PSBD matrix. In particular, we may ask whether

(i) A is PSBD if and only if $(A + A^t)$ is PSBD.
(ii) Any PPT (Principal Pivot Transform)[42] of a PSBD matrix is a PSBD matrix.

Mohan, Neogy, and Das [35] observe that these properties are not carried over to PSBD matrices. However PSBD is a complete class in the sense of [7, 3.9.5].

Theorem 2.27 ([35]). *Suppose $A \in R^{n\times n}$ is a PSBD matrix. Then $A_{\alpha\alpha} \in PSBD$ where $\alpha \subseteq \{1, \ldots, n\}$.*

Theorem 2.28 ([35]). *Suppose $A \in R^{n\times n}$ is a PSBD matrix. Let $D \in R^{n\times n}$ be a positive diagonal matrix. Then $A \in PSBD$ if and only if $DAD^t \in PSBD$.*

Theorem 2.29 ([35]). *PSBD matrices are invariant under principal rearrangement; that is if $A \in R^{n\times n}$ is a PSBD matrix and $P \in R^{n\times n}$ is any permutation matrix, then $PAP^t \in PSBD$.*

Lemma 2.2. *Suppose $A \in R^{n\times n}$ is a PSBD matrix with rank$(A) \geq 2$ and $A + A^t \leq 0$. Then at least one of the following conditions holds.*

(i) A is PSD.
(ii) $A \leq 0$.

Theorem 2.30 ([35]). *Suppose $A \in R^{n\times n}$ is a PSBD matrix with rank$(A) \geq 2$. Then A is a Q_0 matrix.*

Proof. By Theorem 2.25 and Lemma 2.2, it follows that either $A \in$ PSD or $A \leq 0$ or $A \in C_0^*$. Therefore $A \in Q_0$ (see [7]). ■

Theorem 2.31 ([35]). *Suppose A is a PSBD $\cap$ C_0 matrix with rank$(A) \geq 2$. Then $A \in R^{n\times n}$ is a sufficient matrix.*

Proof. Note that by Theorem 2.25, A^t is a PSBD $\cap$ C_0 matrix with rank$(A^t) \geq 2$. Now by Theorem 2.23, A and A^t are pseudomonotone. Hence A and A^t are row sufficient by Theorem 2.15. Therefore, A is sufficient. ■

The following theorem provides a new sufficient condition to solve LCP(q,A) by Lemke's algorithm.

Theorem 2.32. *Suppose $A \in R^{n\times n}$ can be written as $M+N$ where $M \in$ MPSBD $\cap$ C_0^+, rank$(M) \geq 2$, and $N \in C_0$. If the system $q+Mx-N^ty \geq 0$, $y \geq 0$ is feasible, then Lemke's algorithm for LCP(q,A) with covering vector $d > 0$ terminates with a solution.*

For the proof of the above result we refer the reader to the article by Mohan, Neogy, and Das [35]. The proof follows along similar lines to the proof of Evers [14].

2.6 Generalized Positive-Subdefinite Matrices

The class of generalized positive-subdefinite (GPSBD) matrices is an interesting matrix class introduced by Crouzeix and Komlósi [10]. This class is a generalization of the class of symmetric positive-subdefinite (PSBD) matrices introduced by Martos [34] and nonsymmetric PSBD matrices studied by Crouzeix et al. [9]. The solution set of a linear complementarity problem ($S(q,A)$) can be linked with the set of KKT-stationary points ($S''(q,A)$) of the corresponding quadratic programming problem. The row-sufficient matrices have been characterized by Cottle, Pang, and Venkateswarn [8] as the class for which the solution set of LCP(q,A) is the same as the solution set of KKT points of the corresponding quadratic program. In [37], Neogy and Das showed that the property ($S''(q,A) \subseteq S(q,A)$) holds for generalized positive-subdefinite matrices under some additional assumptions and identified a large subclass of GPSBD matrices as row-sufficient matrices. This has practical relevance to the study of quadratic programming and interior point algorithms.

Definition 2.10. A matrix $A \in R^{n\times n}$ is called a *generalized positive-subdefinite matrix* (GPSBD) [10] if there exist nonnegative multipliers s_i, t_i with $s_i + t_i = 1$, $i = 1, 2, \ldots, n$ such that

$$\forall\, x \in R^n, \qquad x^t Ax < 0 \Rightarrow \begin{cases} \text{either} & -s_i x_i + t_i (A^t x)_i \geq 0 \quad \text{for all } i, \\ \text{or} & -s_i x_i + t_i (A^t x)_i \leq 0 \quad \text{for all } i. \end{cases} \tag{2.18}$$

Let S and T be two nonnegative diagonal matrices with diagonal elements s_i, t_i, where $s_i + t_i = 1$ for $i = 1, \ldots, n$. Note that S and T are independent of x. A matrix $A \in R^{n \times n}$ is said to be GPSBD if there exist two nonnegative diagonal matrices S and T with $S + T = I$ such that

$$\forall\, x \in R^n, \qquad x^t Ax < 0 \Rightarrow \begin{cases} \text{either} & -Sx + TA^t x \geq 0, \\ \text{or} & -Sx + TA^t x \leq 0. \end{cases} \tag{2.19}$$

Note that GPSBD reduces to PSBD if $S = 0$. A is called *nondegenerate* GPSBD if for all $x \in R^n$, $x^t Ax < 0$ implies $-Sx + TA^t x \neq 0$ and unisigned; that is, at least one of the inequalities in (2.18) should hold as a strict inequality. A is said to be a *merely generalized positive-subdefinite* (MGPSBD) *matrix* if A is a GPSBD matrix but not a PSBD matrix. The following example is a nontrivial example of a GPSBD matrix.

Example 2.3. Let

$$A = \begin{bmatrix} 0 & 5 & 0 \\ -4 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Note that $\nu_-(A + A^t) = 1$.

Then for any $x = [x_1 \;\; x_2 \;\; x_3]^t$, $x^t Ax = x_1 x_2 < 0$ implies x_1 and x_2 are of opposite sign. Clearly, $A^t x = [-4x_2 \;\; 5x_1 + x_3 \;\; -x_2]^t$ and for $x = [-1 \;\; 1 \;\; 7]^t$, $x^t Ax < 0$ does not imply $A^t x$ is unisigned. Therefore, A is not a PSBD matrix. However, with the choice $s_1 = 0, s_2 = 1$ and $s_3 = 0$, it is easy to check that A is a GPSBD matrix.

Theorem 2.33. *Suppose $A \in MGPSBD \cap C_0$ with $0 < t_i < 1$ for all i. Then A is a row-sufficient matrix.*

Proof. Suppose $x_i (A^t x)_i \leq 0$ for $i = 1, \ldots, n$. Let $I_1 = \{i \,:\, x_i > 0\}$ and $I_2 = \{i \,:\, x_i < 0\}$. We need to consider three cases.

Case I. $I_2 = \emptyset$. Then $x^t Ax = x^t A^t x = \sum_i x_i (A^t x)_i \leq 0$. Because $A \in C_0$, $[x_i (A^t x)_i] = 0$, $\forall\, i$.

Case II. $I_1 = \emptyset$. Then $(-x)^t A^t (-x) = x^t A^t x = \sum_i x_i (A^t x)_i \leq 0$. Because $A \in C_0$, $[x_i (A^t x)_i] = 0$, $\forall\, i$.

Case III. Suppose there exists a vector x such that $x_i (A^t x)_i \leq 0$ for $i = 1, 2, \ldots, n$ and $x_k (A^t x)_k < 0$ for at least one $k \in \{1, 2, \ldots, n\}$. Let $I_1 \neq \emptyset$ and $I_2 \neq \emptyset$. $x^t Ax = x^t A^t x = \sum_i [x_i (A^t x)_i] < 0$. This implies $-s_i x_i + t_i (A^t x)_i \geq 0$, $\forall\, i$ or $-s_i x_i + t_i (A^t x)_i \leq 0$, $\forall\, i$.

Without loss of generality, assume $-s_i x_i + t_i (A^t x)_i \geq 0$, $\forall\, i$. Then for all $i \in I_1$, $-s_i x_i^2 + t_i x_i (A^t x)_i \geq 0$. This implies $[x_i (A^t x)_i] \geq (s_i / t_i) x_i^2 > 0$, $\forall\, i \in I_1$. Therefore, $\sum_{i \in I_1} [x_i (A^t x)_i] > 0$. Because $x_i (A^t x)_i \leq 0$ for $i = 1, \ldots, n$, this leads to a contradiction.

Therefore, $[x_i (A^t x)_i] = 0$, $\forall\, i$. So A is row sufficient. ■

Neogy and Das [37] provide an example to show that the assumption in the above theorem $0 < t_i < 1\ \forall\, i$ cannot be relaxed.

The following theorem extends the result of Evers [14] and the result obtained in Theorem 2.32 in an earlier section for solving LCP(q,A) by Lemke's algorithm when A satisfies certain conditions stated in the following theorem.

Theorem 2.34. *Suppose* $A \in R^{n\times n}$ *can be written as* $M+N$ *where* $M \in MGPSBD \cap C_0^+$*, is nondegenerate with* $0 < t_i < 1,\ \forall\, i$*, and* $N \in C_0$*. If the system* $q+Mx-N^t y \geq 0,\ y \geq 0$ *is feasible, then Lemke's algorithm for LCP*(q,A) *with covering vector* $d > 0$ *terminates with a solution.*

The following example demonstrates that the class MGPSBD $\cap\, C_0^+$ is nonempty.

Example 2.4 ([37]). Consider the copositive-plus matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 8 & 0 & 1 \end{bmatrix}.$$

Take $x = [-1\ \ -1\ \ 1]^t$. It is easy to check that A is not MPSBD. However, with choice $s_i = \frac{1}{2}\ \forall\, i$, A is a MGPSBD matrix.

The following result is a consequence of the characterization of row-sufficient matrices observed by Cottle, Pang, and Venkateswarn [8].

Lemma 2.3. *Suppose* $A \in MGPSBD \cap C_0$ *with* $0 < t_i < 1$ *for all* i*. For each vector* $q \in R^n$*, if* (x^*, u^*) *is a Karush–Kuhn–Tucker pair of the quadratic program* $QP(q,A):\ [\min x^t(Ax+q);\ x \geq 0,\ Ax+q \geq 0]$*, then* x^* *solves* $LCP(q,A):\ [x \geq 0,\ Ax+q \geq 0,\ x^t(Ax+q) = 0]$.

Proof. From Theorem 2.33 and [8, Theorem 4, p. 238], the result follows. ■

Remark 2.1. From Lemma 2.3, it follows that the solution set of a linear complementarity problem $(S(q,A))$ is related to the set of KKT-stationary points $(S''(q,A))$ of the corresponding quadratic programming problem and the statement $S''(q,A) \subseteq S(q,A)$ holds for MGPSBD matrices with some additional assumptions as stated in Theorem 2.33. For details see Neogy and Das [37].

Acknowledgements The authors wish to thank the unknown referees for their valuable suggestions.

References

1. Avriel, M., Schaible, S.: Second-order characterizations of pseudoconvex functions. *Math. Program.* **14**, 170–185 (1978)

2. Bazara, M. S., Sherali, H. D., Shetty, C. M.: *Nonlinear Programming, Theory and Algorithms.* John Wiley and Sons, Inc., Hoboken, NJ (2006)
3. Brighi, L., John, R.: Characterizations of pseudomonotone maps and economic equilibrium. *J. Statist. Manage. Sys.* **5**, 253–273 (2002)
4. Cottle, R. W.: Nonlinear programs with positively bounded Jacobians. Ph.D. Thesis, Department of Mathematics, University of California, Berkeley (1964)
5. Cottle, R. W., Ferland, J. A.: Matrix-theoretic criteria for the quasiconvexity and pseudoconvexity of quadratic functions. *Linear Algebra Appl.* **5**, 123–136 (1972)
6. Cottle, R. W., Guu, S.-M.: Two characterizations of sufficient matrices. *Linear Algebra Appl.* **170**, 56–74 (1992)
7. Cottle, R. W., Pang, J. S., Stone, R. E.: *The Linear Complementarity Problem*. Academic Press, New York (1992)
8. Cottle, R. W., Pang, J. S., Venkateswarn, V. Sufficient matrices and the linear complementarity problem. *Linear Algebra and Its Appl.* **114/115**, 231–249 (1989)
9. Crouzeix, J.-P., Hassouni, A., Lahlou, A., Schaible, S.: Positive subdefinite matrices, generalized monotonicity and linear complementarity problems. *SIAM J. Matrix Anal. Appl.* **22**, 66–85 (2000)
10. Crouzeix, J.-P., Komlósi, S.: The linear complementarity problem and the class of generalized positive subdefinite matrices, In: Giannessi, F., Pardalos, P., Rapcsak, T. (eds.) *Optimization Theory: Recent Developments from Matrahaza*, pp. 45–63. Kluwer Academic Dordrecht (2002)
11. Crouzeix, J.-P., Schaible, S.: Generalized monotone affine maps. *SIAM J. Matrix Anal. Appl.* **17**, 992–997 (1996)
12. Eaves, B. C.: On the basic theorem of complementarity. *Math. Program.* **1**, 68–75 (1971)
13. Eaves, B. C.: The linear complementarity problem. *Manage. Sci.* **17** 612–634 (1971)
14. Evers, J. J. M.: More with the Lemke complementarity algorithm. *Math. Program.* **15**, 214–219 (1978)
15. Facchinei, F., Pang, J.-S.: *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer-Verlag, New York (2003)
16. Fiedler, M., Pták, V.: Some generalizations of positive definiteness and monotonicity. *Numerische Mathematik* **9**, 163–172 (1966)
17. Garcia, C. B.: Some classes of matrices in linear complementarity theory. *Math. Program.* **5**, 299–310 (1973)
18. Gowda, M. S.: Affine pseudomonotone mappings and the linear complementarity problem. *SIAM J. Matrix Anal. Appl.* **11**, 373–380 (1990)
19. Gowda, M. S.: On the transpose of a pseudomonotone matrix and the linear complementarity problem. *Linear Algebra Appl.* **140**, 129–137 (1990)
20. Gowda, M. S.: Pseudomonotone and copositive star matrices. *Linear Algebra Appl.* **113**, 107–118 (1990)
21. Hadjisavvas N., Komlosi S., Schaible S. (Eds.): *Handbook of Generalized Convexity and Generalized Monotonicity, Nonconvex Optimization and its Applications* **76**, Springer, New York (2005)
22. Harker, P. T., Pang, J. S.: Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. *Math. Program.* **48**, 161–220 (1990)
23. John, R.: Quasimonotone individual demand. *Optimization* **47** 201–209 (2000)
24. John, R.: Uses of generalized convexity and generalized monotonicity in economics, in *Handbook of Generalized Convexity and Generalized Monotonicity*, Hadjisavvas N., Komlosi S., Schaible S. (Eds.), *Nonconvex Optimization and Its Applications*, **76**, pp. 619–666. Springer (2005)
25. Karamardian, S.: The nonlinear complementarity problem with applications, Part 2. *J. Optim. Theory Appl.* **4**, 167–181 (1969)
26. Karamardian, S.: Generalized complementarity problems. *J. Optim. Theory Appl.* **8**, 161–167 (1971)

27. Karamardian, S.: Complementarity problems over cones with monotone and pseudomonotone maps. *J. Optim. Theory Appl.* **18**, 445–454 (1976)
28. Karamardian, S.: An existence theorem for complementarity problems. *J. Optim. Theory Appl.* **19**, 227–232 (1976)
29. Karamardian, S., Schaible, S.: Seven kinds of monotone maps. *J. Optim. Theory Appl.* **66**, 37–46 (1990)
30. Karamardian, S., Sehaible, S., Crouzeix, J.-P.: Characterizations of generalized monotone maps. *J. Optim. Theory Appl.* **76**, 399–413 (1993)
31. Lemke, C. E.: Bimatrix equilibrium points and mathematical programming. *Manage. Sci.* **11**, 681–689 (1965)
32. Lemke, C. E., Howson, J. T. Jr.: Equilibrium points of bimatrix games. *SIAM J. Appl. Math.* **12**, 413–423 (1964)
33. Mangasarian, O. L.: *Nonlinear Programming*. McGraw-Hill, New York (1969)
34. Martos, B.: Subdefinite matrices and quadratic forms. *SIAM J. Appl. Math.* **17**, 1215–1223 (1969)
35. Mohan, S. R., Neogy, S. K., Das, A. K.: More on positive subdefinite matrices and the linear complementarity problem. *Linear Algebra Appl.* **338** 275-285 (2001)
36. Murty, K. G.: *Linear Complementarity, Linear and Nonlinear Programming*. Heldermann Verlag, West Berlin (1988)
37. Neogy, S. K., Das, A. K.: Some properties of generalized positive subdefinite matrices. *SIAM J. Matrix Anal. Appl.* **27**, 988-995 (2006)
38. Nikaido, H.: *Convex Structures and Economic Theory*, Academic Press, New York (1968)
39. Pini, R., Schaible, S.: Invariance properties of generalized monotonicity. *Optimization.* **28**, 211–222 (1994)
40. Rao, P. S. S. N. V. P.: A characterisation of merely positive subdefinite matrices and related results. Sankhya series A: *Indian J. Statist.* **39**, 387–395 (1977)
41. Todd, M. J.: Orientation in complementary pivot algorithms. *Math. Ope. Res.* **1**, 54-66 (1976)
42. Tucker, A. W.: Principal pivotal transforms of square matrices, *SIAM Rev.* **5**, 305–305 (1963)
43. Väliaho, H.: Criteria for sufficient matrices. *Linear Algebra Appl.* **233**, 109–129 (1996)

Chapter 3
Optimality Conditions Without Continuity in Multivalued Optimization Using Approximations as Generalized Derivatives

Phan Quoc Khanh and Nguyen Dinh Tuan

Abstract We propose a notion of approximations as generalized derivatives for multivalued mappings and establish both necessary and sufficient conditions of orders 1 and 2 for various kinds of efficiency in multivalued vector optimization without convexity and even continuity. Compactness assumptions are also relaxed. Our theorems include several recent existing results in the literature as special cases.

3.1 Introduction and Preliminaries

Differentiability assumptions are often crucial for a classical problem in all areas of continuous mathematics, because derivatives are local linear approximations for the involved nonlinear mappings and then supply a much simpler approximated linear problem, replacing the original nonlinear problem. However, such differentiability assumptions are too severe and not satisfied in many practical situations. Relaxing these assumptions has been one of the main ideas in optimization for more than three decades now and constituted an important field of research called nonsmooth optimization. Most of contributions in this field are based on using generalized derivatives, which are local approximations bearing not the whole linearity but still parts of linearity. Many notions of generalized derivatives have been proposed. Each of them is suitable for a class of problems. The Clarke derivative [5] was introduced for locally Lipschitz mappings; the quasidifferentiability of Demyanov and Rubinov [6] requires directional differentiability to be defined; the approximate Jacobian proposed in [8] (later renamed the pseudo-Jacobian) exists only for

Phan Quoc Khanh
Department of Mathematics, International University of Hochiminh City, Linh Trung, Thu Duc, Hochiminh City, Vietnam, e-mail: pqkhanh@hcmiu.edu.vn

Nguyen Dinh Tuan
Department of Mathematics and Statistics, University of Economics of Hochiminh City, Nguyen Dinh Chieu Street, D. 3, Hochiminh City, Vietnam, e-mail: ndtuan73@yahoo.com

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_3,

continuous mappings, and so on. The approximations, introduced in [11] for order 1 and in [1] for order 2, are defined for general mappings that are even discontinuous. In this note we extend these definitions to the case of multifunctions.

The major goal proposed for generalized derivatives is establishing optimality conditions in nonsmooth optimization problems. We see from the very beginning of classical optimization that derivatives play a fundamental role in the Fermat theorem, the first necessary optimality condition. We would say that all generalized derivatives are used in similar ways as the classical derivative in the Fermat theorem. In the literature we observe only [11, 1, 2, 12–14, 16, 17] which deal with this kind of approximations as generalized derivatives. This notion was used in [11] to study metric regularity and in [1] for establishing second-order necessary optimality conditions in the compactness case. Second-order approximations of scalar functions are used for support functions in [2] to scalarize vector problems so that second-order optimality conditions can be established, but under strict (first-order) differentiability and compactness assumptions. In [12–14] we used first- and second-order approximations of single-valued mappings to derive first- and second-order necessary and sufficient conditions for various kinds of efficiency in nonsmooth vector optimization problems of several types.

In this chapter we develop the results of our talk presented (but unpublished) at an international conference [16]. Namely, after extending the notion of first- and second-order approximations of a mapping to the case of a multivalued mapping, we use this notion to establish both necessary and sufficient conditions of both orders 1 and 2 for weak and firm efficiencies in multivalued vector optimization with set constraints, without continuity and convexity assumptions. In [17] we develop such optimality conditions also for proper efficiency and in problems with functional constraints. The problem under our consideration here is as follows. Throughout this chapter, unless otherwise specified, let X and Y be normed spaces, Y being partially ordered by a convex cone C with nonempty interior, $S \subseteq X$ be a nonempty subset, and $F : X \to 2^Y$ be a multifunction (i.e., a multivalued mapping). We are concerned with the problem

$$\min F(x), \text{ subject to } x \in S. \tag{P}$$

Here "min" means minimizing: finding efficient solutions in the sense defined by the end of this section. The layout of the chapter is as follows. In the rest of this section we recall definitions and preliminaries needed for our later investigation. Section 3.2 is devoted to defining first- and second-order approximations of a multivalued mapping. In Section 3.3 we establish necessary conditions of order 1 for weak efficiency and sufficient conditions of order 1 for firm efficiency of problem (P). We develop such conditions for these kinds of efficiency, but of order 2, in the final Section 3.4.

Our notations are rather standard. $\mathbb{N} = \{1, 2, \ldots n, \ldots\}$ and $\|.\|$ stands for the norm in any normed space (the context makes it clear what space is concerned). B_X denotes the open unit ball in X and $B_X(x,r) = \{z \in X \mid \|x - z\| < r\}$; X^* is the topological dual of X with $\langle .,. \rangle$ being the canonical pairing. $L(X,Y)$ denotes the space of all bounded linear mappings from X into Y and $B(X,X,Y)$ that of all bounded bilinear mappings from $X \times X$ into Y. For a cone $C \subseteq X$, $C^* = \{x^* \in X^* \mid \langle x^*, c\rangle \geq 0, \forall c \in C\}$ is the positive polar cone of C. For $A \subseteq X$, int$A$, cl$A$, and bd$A$ denote the interior, closure, and boundary of A, respectively. For $t > 0$ and $k \in \mathbb{N}$, $o(t^k)$ stands for a moving point (in a normed space) such that $o(t^k)/t^k \to 0$ as $t \to 0^+$. We use the following tangent sets of $A \subseteq X$ at $x_0 \in A$.

(a) The contingent (or Bouligand) cone of A at x_0 (see [3]) is

$$T(A,x_0) = \{v \in X \mid \exists t_n \to 0^+, \exists v_n \to v, \forall n \in \mathbb{N}, x_0 + t_n v_n \in A\}.$$

(b) The second-order contingent set of A at (x_0, v) (see [3]) is

$$T^2(A,x_0,v) = \left\{w \in X \mid \exists t_n \to 0^+, \exists w_n \to w, \forall n \in \mathbb{N}, x_0 + t_n v + \frac{1}{2}t_n^2 w_n \in A\right\}.$$

(c) The asymptotic second-order tangent cone of A at (x_0, v) (see [4, 18]) is

$$T''(A,x_0,v) = \left\{w \in X \mid \exists (t_n, r_n) \to (0^+, 0^+) : \frac{t_n}{r_n} \to 0, \exists w_n \to w.\right.$$
$$\left.\forall n \in \mathbb{N}, x_0 + t_n v + \frac{1}{2}t_n r_n w_n \in A\right\}.$$

Lemma 3.1 ([10]). *Assume that X is a finite-dimensional space $\mathbb{R}^m$ and $x_0 \in A \subseteq X$. If $x_n \in A \setminus \{x_0\}$ tends to x_0, then there exists $v \in T(A,x_0) \setminus \{0\}$ and a subsequence, denoted again by x_n, such that, for $t_n = \|x_n - x_0\|$,*

(i) $(1/t_n)(x_n - x_0) \to v$.
(ii) Either $z \in T^2(A,x_0,v) \cap v^{\perp}$ exists such that $(x_n - x_0 - t_n v)/\frac{1}{2}t_n^2 \to z$ or $z \in T''(A,x_0,v) \cap v^{\perp} \setminus \{0\}$ and $r_n \to 0^+$ with $(t_n/r_n) \to 0^+$ exist such that $(x_n - x_0 - t_n v)/\frac{1}{2}t_n r_n \to z$, where $v^{\perp} = \{y \in \mathbb{R}^m \mid \langle y, v\rangle = 0\}$.

Recall now notions of efficiency in vector optimization. Consider a subset V of the objective space Y. A point $y_0 \in V$ is called an efficient point (weak efficient point, strict efficient point, respectively) of V if

$$(V - y_0) \cap -C \subseteq (-C) \cap C$$

$$((V - y_0) \cap -\text{int}C = \emptyset,$$

$$(V - y_0) \cap (-C \setminus \{0\}) = \emptyset, \quad \text{respectively}).$$

The set of efficient, weak efficient and strict efficient points are denoted by $\text{Min}_C V$, $\text{WMin}_C V$, and $\text{StrMin}_C V$, respectively. Now apply these notions to problem (P). A point (x_0, y_0) with $x_0 \in S$ and $y_0 \in F(x_0)$ is said to be a local weak efficient solution of (P) if there is a neighborhood U of x_0 such that, $\forall x \in S \cap U$,

$$(F(x) - y_0) \cap -\text{int}C = \emptyset \tag{3.1}$$

and (x_0,y_0) is called a local efficient solution if (3.1) is replaced by

$$(F(x)-y_0)\cap -C \subseteq (-C)\cap C.$$

We extend the firm efficiency notion (see [9, 15]) to the case of multivalued optimization as follows.

Definition 3.1. Let $x_0 \in S$, $y_0 \in F(x_0)$ and $m \in \mathbb{N}$. Then (x_0,y_0) is said to be a *local firm efficient solution of order m* if there are a neighborhood U of x_0 and $\gamma > 0$ such that $y_0 \in \mathrm{StrMin}_C F(x_0)$ and, for all $x \in S\cap U \setminus \{x_0\}$,

$$(F(x)-y_0)\cap (B_Y(0,\gamma\|x-x_0\|^m)-C)=\emptyset.$$

In the sequel let LWE(P), LE(P), and LFE(m, P) stand for the sets of the local weakly efficient solutions, of the local efficient solutions, and of the local firm efficient solutions of order m, respectively, of problem (P). Then it is clear that, for $p,m \in \mathbb{N}$ with $p \geq m$,

$$\mathrm{LFE}(m,\mathrm{P}) \subseteq \mathrm{LFE}(p,\mathrm{P}) \subseteq \mathrm{LE(P)} \subseteq \mathrm{LWE(P)}.$$

Hence, necessary conditions for the rightmost term are valid also for the others and sufficient conditions for the leftmost term hold true for the others as well.

For a multifunction $H : X \to 2^Y$ the domain of H is

$$\mathrm{dom}H = \{x \in X \mid H(x) \neq \emptyset\}.$$

H is said to be upper semicontinuous (usc) at $x_0 \in \mathrm{dom}H$ if for all open set $V \supseteq H(x_0)$, there is a neighborhood U of x_0 such that $V \supseteq H(U)$. H is termed lower semicontinuous (lsc) at $x_0 \in \mathrm{dom}H$ if for all open set $V \cap H(x_0) \neq \emptyset$, there is a neighborhood U of x_0 such that for all $x \in U$, $V \cap H(x) \neq \emptyset$.

3.2 First- and Second-Order Approximations of Multifunctions

Consider a multifunction $H : X \to 2^Y$, $x_0 \in \mathrm{dom}H$ and $y_0 \in F(x_0)$.

Definition 3.2.

(i) A subset $A_H(x_0,y_0)$ of $L(X,Y)$ is said to be a *first-order approximation* of H at (x_0,y_0) if there exists a neighborhood U of x_0 such that, for all $x \in U\cap \mathrm{dom}H$, there are positive r_x with $r_x\|x-x_0\|^{-1} \to 0^+$ and $y \in H(x)$ satisfying

$$y-y_0 \in A_H(x_0,y_0)(x-x_0)+r_x B_Y.$$

(ii) A subset $A_H^S(x_0,y_0)$ of $L(X,Y)$ is called a *first-order strong approximation* of H at (x_0,y_0) if there exists a neighborhood U of x_0 such that, for all $x \in U \cap \mathrm{dom}H$, there is positive r_x with $r_x\|x-x_0\|^{-1} \to 0^+$ such that, for all $y \in H(x)$,

$$y-y_0 \in A_H^S(x_0,y_0)(x-x_0)+r_x B_Y.$$

(iii) A pair $(A_H(x_0,y_0), B_H(x_0,y_0))$, where $A_H(x_0,y_0) \subseteq L(X,Y)$ and $B_H(x_0,y_0) \subseteq B(X,X,Y)$, is called a *second-order approximation* of H at (x_0,y_0) if $A_H(x_0,y_0)$ is a first-order approximation of H at (x_0,y_0) and there is a neighborhood U of x_0 such that, for all $x \in U \cap \mathrm{dom}H$, there are positive r_x^2 with $r_x^2\|x-x_0\|^{-2} \to 0^+$ and $y \in H(x)$ satisfying

$$y - y_0 \in A_H(x_0,y_0)(x-x_0) + B_H(x_0,y_0)(x-x_0,x-x_0) + r_x^2 B_Y.$$

(iv) A pair $(A_H^S(x_0,y_0), B_H^S(x_0,y_0))$, where $A_H^S(x_0,y_0) \subseteq L(X,Y)$ and $B_H^S(x_0,y_0) \subseteq B(X,X,Y)$, is termed a *second-order strong approximation* of H at (x_0,y_0) if $A_H^S(x_0,y_0)$ is a first-order strong approximation of H at (x_0,y_0) and there is a neighborhood U of x_0 such that, for all $x \in U \cap \mathrm{dom}H$, there exists positive r_x^2 with $r_x^2\|x-x_0\|^{-2} \to 0^+$ such that, for all $y \in H(x)$,

$$y - y_0 \in A_H^S(x_0,y_0)(x-x_0) + B_H^S(x_0,y_0)(x-x_0,x-x_0) + r_x^2 B_Y.$$

In this chapter we impose on these approximations the following relaxed compactness.

Definition 3.3.

(i) Let M_n and M be in $L(X,Y)$. The sequence M_n is said to *pointwise converge* to M and is written as $M_n \xrightarrow{p} M$ or $M = \text{p-lim} M_n$ if $\lim M_n(x) = M(x)$ for all $x \in X$. A similar definition is adopted for $N_n, N \in B(X,X,Y)$.

(ii) A subset $A \subseteq L(X,Y)$ ($B \subseteq B(X,X,Y)$, respectively) is called (sequentially) *asymptotically pointwise compact*, or (sequentially) *asymptotically p-compact* if

(a) Each norm bounded sequence $\{M_n\} \subseteq A$ ($\subseteq B$, respectively) has a subsequence $\{M_{n_k}\}$ and $M \in L(X,Y)$ ($M \in B(X,X,Y)$, respectively) such that $M = \text{p-}\lim M_{n_k}$.

(b) For each sequence $\{M_n\} \subseteq A$ ($\subseteq B$, respectively) with $\lim \|M_n\| = \infty$, the sequence $\{M_n/\|M_n\|\}$ has a subsequence which pointwise converges to some $M \in L(X,Y) \setminus \{0\}$ ($M \in B(X,X,Y) \setminus \{0\}$, respectively).

(iii) If in (ii), pointwise convergence, (i.e., p-lim) is replaced by convergence (i.e., lim), a subset $A \subseteq L(X,Y)$ (or $B \subseteq B(X,X,Y)$) is called (sequentially) *asymptotically compact*.

Because only sequential convergence is considered in this chapter, we omit the word "sequentially". For $A \subseteq L(X,Y)$ and $B \subseteq B(X,X,Y)$ we adopt the notations:

$$p-clA = \{M \in L(X,Y) : \exists (M_n) \subseteq A, M = p - \lim M_n\}, \tag{3.2}$$

$$p-clB = \{N \in B(X,X,Y) : \exists (N_n) \subseteq B, N = p - \lim N_n\}, \tag{3.3}$$

$$A_\infty = \{M \in L(X,Y) : \exists (M_n) \subseteq A, \exists t_n \to 0^+, M = \lim t_n M_n\}, \tag{3.4}$$

$$p-A_\infty = \{M \in L(X,Y) : \exists (M_n) \subseteq A, \exists t_n \to 0^+, M = p - \lim t_n M_n\}, \tag{3.5}$$

$$p-B_\infty = \{N \in B(X,X,Y) : \exists (N_n) \subseteq B, \exists t_n \to 0^+, N = p - \lim t_n N_n\}. \tag{3.6}$$

The sets (3.2), (3.3) are pointwise closures; (4) is just the definition of the recession cone of A. So (3.5), (3.6) are pointwise recession cones.

Remark 3.1.

(i) If X is finite-dimensional, a convergence occurs if and only if the corresponding pointwise convergence does, but in general the "if" does not hold; see [[12], Example 3.1].

(ii) If X and Y are finite-dimensional, every subset is asymptotically p-compact and asymptotically compact but in general the asymptotical compactness is stronger, as shown by [[12], Example 3.2].

(iii) Assume that $\{M_n\} \subseteq L(X,Y)$ is norm bounded. If $x_n \to x$ in X and $M_n \xrightarrow{p} M$ in $L(X,Y)$, then $M_n x_n \to Mx$ in Y. Similarly, if $x_n \to x$, $y_n \to y$ in X, $N_n \xrightarrow{p} N$ in $B(X,X,Y)$ and $\{N_n\}$ is norm bounded then $N_n(x_n,y_n) \to N(x,y)$ in Y.

Indeed, the conclusion is derived from the following evaluations.

$$\|M_n x_n - Mx\| \leq \|M_n x_n - M_n x\| + \|M_n x - Mx\| \leq \|M_n\|\|x_n - x\| + \|M_n x - Mx\|;$$

$$\begin{aligned}\|N_n(x_n,y_n) - N(x,y)\| &\leq \|N_n(x_n,y_n) - N_n(x_n,y)\| + \|N_n(x_n,y) - N_n(x,y)\| \\ &\quad + \|N_n(x,y) - N(x,y)\| \leq \|N_n\|\|x_n\|\|y_n - y\| \\ &\quad + \|N_n\|\|x_n - x\|\|y\| + \|N_n(x,y) - N(x,y)\|.\end{aligned}$$

The following example gives a multivalued map F, which is neither usc nor lsc at x_0, but has even second-order strong approximations.

Example 3.1. Let $F : \mathbb{R} \to 2^{\mathbb{R}}$ be defined by

$$F(x) = \begin{cases} \{y \in \mathbb{R} \mid y \geq \sqrt{x}\} & \text{if } x > 0, \\ \{y \in \mathbb{R} \mid y \leq \frac{1}{x}\} & \text{if } x < 0, \\ \{0\} & \text{if } x = 0. \end{cases}$$

Let $(x_0,y_0) = (0,0)$. Then F is neither usc nor lsc at x_0 but F has the following approximations, for fixed positive α and $\beta > 0$.

$$\begin{aligned} &A_F(x_0,y_0) = (\alpha,+\infty), \qquad A_F^S(x_0,y_0) = (\beta,+\infty), \\ &B_F(x_0,y_0) = B_F^S(x_0,y_0) = \{0\}. \end{aligned}$$

In the next example F is not usc at x_0 but $A_F(x_0,y_0)$ is even a singleton.

Example 3.2. Let $F : \mathbb{R}^2 \to 2^{\mathbb{R}}$ be defined by

$$F(x_1,x_2) = \begin{cases} \left\{y \in \mathbb{R} \mid \frac{2}{3}|x_1|^{\frac{3}{2}} + x_2^2 \leq y \leq \frac{1}{|x_1|+|x_2|}\right\} & \text{if } (x_1,x_2) \neq (0,0), \\ \{0\} & \text{if } (x_1,x_2) = (0,0). \end{cases}$$

Then F is not usc at $x^0 = (0,0)$. But for $y_0 = 0$ we have

$$\begin{aligned} A_F(x^0, y_0) &= \{0\}, \\ A_F^S(x^0, y_0) &= (\mathbb{R} \setminus \{0\}) \times \{0\} \cup \{0\} \times (\mathbb{R} \setminus \{0\}), \\ B_F(x^0, y_0) &= \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix} \mid \alpha > 1 \right\}, \\ B_F^S(x^0, y_0) &= \{0\}. \end{aligned}$$

Note that a similar example for a single-valued mapping does not exist, because a single-valued mapping has a first-order approximation at x_0 being a singleton if and only if it is Fréchet differentiable at x_0 and hence continuous at this point.

3.3 First-Order Optimality Conditions

Theorem 3.1 (Necessary condition). *Consider problem (P). Assume that $A_F(x_0, y_0)$ is an asymptotically p-compact first-order approximation of F at (x_0, y_0). If $(x_0, y_0) \in LWE(P)$ then, for each $v \in T(S, x_0)$ there is $M \in p - clA_F(x_0, y_0) \cup (p\text{-}A_F(x_0, y_0)_\infty \setminus \{0\})$ such that*

$$Mv \notin -int\, C.$$

Proof. Let $v \in T(S, x_0)$ be arbitrary and fixed. By the definition of a contingent cone, there is $(t_n, v_n) \to (0^+, v)$ such that $x_0 + t_n v_n \in S$ for all $n \in \mathbb{N}$. By the weak efficiency of (x_0, y_0) one has, for large n and all $y \in F(x_0 + t_n v_n)$,

$$y - y_0 \notin -\text{int}\, C.$$

On the other hand, as $A_F(x_0, y_0)$ is a first-order approximation, there are positive r_n with $r_n t_n^{-1} \to 0^+$ and $y_n \in F(x_0 + t_n v_n)$ such that

$$y_n - y_0 \in A_F(x_0, y_0)(t_n v_n) + r_n B_Y.$$

Therefore, $M_n \in A_F(x_0, y_0)$ and $\bar{y}_n \in r_n B_Y$ exist such that

$$M_n(t_n v_n) + \bar{y}_n \notin -\text{int}\, C. \tag{3.7}$$

If $\{M_n\}$ is norm bounded, one can assume that $M_n \xrightarrow{\text{p}} M \in \text{p} - \text{cl}A_F(x_0, y_0)$. Dividing (3.7) by t_n and passing to the limit one gets $Mv \notin -\text{int}\, C$. If $\{M_n\}$ is unbounded, one can assume that $\|M_n\| \to \infty$ and

$$\frac{M_n}{\|M_n\|} \xrightarrow{\text{p}} M \in \text{p} - A_F(x_0, y_0)_\infty \setminus \{0\}.$$

Dividing (3.7) by $\|M_n\| t_n$ one obtains in the limit $Mv \notin -\text{int}\, C$. ∎

If F is single-valued, Theorem 3.1 collapses to Theorem 3.3 of [13], which was shown there to improve or include many existing results. The following example shows that, for F being multivalued, Theorem 3.1 is easily applied.

Example 3.3. Let $X = Y = \mathbb{R}, S = [0, +\infty), C = \mathbb{R}_+, x_0 = y_0 = 0$, and

$$F(x) = \begin{cases} \{y \in \mathbb{R} \mid y \leq \frac{1}{\sqrt[3]{x}}\} & \text{if } x > 0, \\ \{y \in \mathbb{R} \mid y \geq \sqrt[3]{-x}\} & \text{if } x < 0, \\ \{0\} & \text{if } x = 0. \end{cases}$$

Then $T(S,x_0) = S$ and for a fixed $\alpha < 0$ we have $A_F(x_0,y_0) = (-\infty,\alpha)$, $\mathrm{cl}A_F(x_0,y_0) = (-\infty,\alpha]$, $A_F(x_0,y_0)_\infty = (-\infty,0]$. Taking $v = 1 \in T(S,x_0)$ one sees that, for all $M \in \mathrm{cl}A_F(x_0,y_0) \cup (A_F(x_0,y_0)_\infty \setminus \{0\}) = (-\infty,0)$,

$$Mv = M \in -\mathrm{int}\, C.$$

Due to Theorem 3.1, (x_0,y_0) is not a local (weakly efficient) solution of problem (P).

Theorem 3.2 (Sufficient condition). *Consider problem (P) with X being finite-dimensional. Assume that* $A_F^S(x_0,y_0)$ *is an asymptotically p-compact first-order strong approximation of F at* (x_0,y_0)*,* $x_0 \in S$*, and* $y_0 \in StrMin_C F(x_0)$*. Impose further that, for all* $v \in T(S,x_0) \setminus \{0\}$ *and all* $M \in p - clA_F^S(x_0,y_0) \cup (p - A_F^S(x_0,y_0)_\infty \setminus \{0\})$*, one has*

$$Mv \notin -\mathrm{cl}C.$$

Then $(x_0,y_0) \in LFE(1,P)$.

Proof. Reasoning ad absurdum, suppose the existence of $x_n \in S \cap B_X(x_0,(1/n)) \setminus \{x_0\}$ such that, for each $n \in \mathbb{N}$, there is $y_n \in F(x_n)$ such that

$$y_n - y_0 \in B_Y\left(0, \frac{1}{n}\|x_n - x_0\|\right) - C.$$

As X is finite-dimensional, we can assume that $(x_n - x_0)/\|x_n - x_0\|$ tends to a point v in $T(S,x_0) \setminus \{0\}$. On the other hand, for large n there is positive r_n with $r_n\|x_n - x_0\|^{-1} \to 0^+$ such that

$$y_n - y_0 \in A_F^S(x_0,y_0)(x_n - x_0) + r_n B_Y.$$

Hence, there are $M_n \in A_F^S(x_0,y_0)$ and $\bar{y}_n \in r_n B_Y$ such that

$$M_n(x_n - x_0) + \bar{y}_n \in B_Y\left(0, \frac{1}{n}\|x_n - x_0\|\right) - C.$$

Arguing similarly as in the final part of the proof of Theorem 3.1, we obtain $M \in \mathrm{p} - \mathrm{cl}A_F^S(x_0,y_0) \cup (\mathrm{p} - A_F^S(x_0,y_0)_\infty \setminus \{0\})$ such that $Mv \in -\mathrm{cl}\, C$, a contradiction. ■

Theorem 3.2 includes Theorem 3.4 of [13] as a special case where F is single-valued. The following example explains how to employ Theorem 3.2.

Example 3.4. Let $X = \mathbb{R}$, $Y = \mathbb{R}^2$, $S = [0,+\infty)$, $C = \mathbb{R}^2_+$, $x_0 = 0$, $(y_0,z_0) = (0,0) \in Y$, and

$$F(x) = \begin{cases} \{(y,z) \in \mathbb{R}^2 \mid y \geq \sqrt[3]{x}, z = x\} & \text{if } x > 0, \\ \{(0,0)\} & \text{if } x = 0, \\ \emptyset & \text{if } x < 0. \end{cases}$$

Then $(y_0,z_0) \in \mathrm{StrMin}_C F(x_0)$ and for any fixed $\alpha > 0$ we can take a strong approximation as follows.

$$A^S_F(x_0,(y_0,z_0)) = \{(y,z) \in \mathbb{R}^2 \mid y > \alpha, z = 1\},$$

$$\mathrm{cl}A^S_F(x_0,(y_0,z_0)) = \{(y,z) \in \mathbb{R}^2 \mid y \geq \alpha, z = 1\},$$

$$A^S_F(x_0,(y_0,z_0))_\infty = \{(y,z) \in \mathbb{R}^2 \mid y \geq 0, z = 0\}.$$

It is clear that $(\forall v \in T(S,x_0) \setminus \{0\} = (0,+\infty))$ one has $\forall M \in \mathrm{cl}A^S_F(x_0,(y_0,z_0))$, $Mv = (yv,v) \notin -C$. Furthermore, $Mv = (yv,0) \notin -C$ for all $M \in A^S_F(x_0,(y_0,z_0))_\infty \setminus \{0\}$. By Theorem 3.2, $(x_0,(y_0,z_0)) \in \mathrm{LFE}(1,\mathrm{P})$.

3.4 Second-Order Optimality Conditions

Theorem 3.3 (Necessary condition). *For problem (P) assume that $(A_F(x_0,y_0), B_F(x_0,y_0))$ is an asymptotically p-compact second-order approximation of F at (x_0,y_0) with $A_F(x_0,y_0)$ being norm bounded. Assume further that $(x_0,y_0) \in LWE(P)$. Then*

(i) For all $v \in T(S,x_0)$, there exists $M \in p-clA_F(x_0,y_0)$ such that $Mv \notin -int\,C$.
(ii) For all $v \in T(S,x_0)$ with $A_F(x_0,y_0)v \subseteq -bdC$ one has

(a) For each $w \in T^2(S,x_0,v)$, either $\overline{M} \in p-clA_F(x_0,y_0)$ and $\overline{N} \in p-clB_F(x_0,y_0)$ exist such that

$$\overline{M}w + 2\overline{N}(v,v) \notin -int\,C,$$

or there is $\overline{N} \in p-B_F(x_0,y_0)_\infty \setminus \{0\}$ satisfying

$$\overline{N}(v,v) \notin -int\,C.$$

(b) For each $w \in T''(S,x_0,v)$, either $M' \in p-clA_F(x_0,y_0)$ and $N' \in p-B_F(x_0,y_0)_\infty$ exist such that

$$M'w + N'(v,v) \notin -int\,C,$$

or one has $N' \in p - B_F(x_0,y_0)_\infty \setminus \{0\}$ *with*

$$N'(v,v) \notin -int\, C.$$

Proof.

(i) This assertion follows from Theorem 3.1.

(ii) (a) Let $v \in T(S,x_0)$ with $A_F(x_0,y_0)(v) \subseteq -bdC$ and $w \in T^2(S,x_0,v)$. Then, there are $x_n \in S$ and $t_n \to 0^+$ such that

$$w_n := (x_n - x_0 - t_n v)/\frac{1}{2}t_n^2 \to w.$$

By the definition of the second-order approximation, there are $M_n \in A_F(x_0,y_0)$, $N_n \in B_F(x_0,y_0)$ and $o(\|x_n - x_0\|^2)$ such that, for large n,

$$M_n(x_n - x_0) + N_n(x_n - x_0, x_n - x_0) + o(\|x_n - x_0\|^2) \in F(x_n) - y_0.$$

The weak efficiency of (x_0,y_0) implies then, for some $o(t_n^2) \in Y$,

$$M_n w_n + 2N_n\left(v + \frac{1}{2}t_n w_n, v + \frac{1}{2}t_n w_n\right) + o(t_n^2)/\frac{1}{2}t_n^2 \notin -\text{int}\, C. \qquad (3.8)$$

We can assume that $M_n \xrightarrow{p} \overline{M}$ for some $\overline{M} \in \text{p} - \text{cl}A_F(x_0,y_0)$. If $\{N_n\}$ is norm bounded then $N_n \xrightarrow{p} \overline{N}$ for some $\overline{N} \in \text{p} - \text{cl}B_F(x_0,y_0)$. From (3.8) we get in the limit

$$\overline{M}w + 2\overline{N}(v,v) \notin -\text{int}\, C.$$

If $\{N_n\}$ is unbounded, we can assume $\|N_n\| \to \infty$ and $(N_n/\|N_n\|) \xrightarrow{p} \overline{N}$ for some $\overline{N} \in \text{p} - B_f(x_0)_\infty \setminus \{0\}$. Dividing (3.8) by $\|N_n\|$ and passing to the limit gives $\overline{N}(v,v) \notin -\text{int}\, C$.

(b) For any $w \in T''(S,x_0,v)$, there are $x_n \in S$ and $(t_n,r_n) \to (0^+,0^+)$ with $(t_n/r_n) \to 0^+$ such that

$$w_n := (x_n - x_0 - t_n v)/\frac{1}{2}t_n r_n \to w.$$

Similarly as in (a) we have M_n' and N_n' satisfying the following relation, corresponding to (3.8),

$$M_n' w_n + \left(\frac{2t_n}{r_n}\right) N_n' \left(v + \frac{1}{2}t_n w_n, v + \frac{1}{2}t_n w_n\right) + o(t_n^2)/\frac{1}{2}t_n r_n \notin -\text{int}\, C. \quad (3.9)$$

We can assume that $M_n' \xrightarrow{p} M' \in \text{p} - \text{cl}A_F(x_0,y_0)$. There are three possibilities.

(α) $(2t_n/r_n)N_n' \to 0$. From (3.9) we get in the limit

$$M'w \notin -\text{int}\, C.$$

(β) If $(2t_n/r_n)\|N_n^{'}\| \to a > 0$, then $\|N_n\| \to \infty$ and we can assume that $(N_n^{'}/\|N_n^{'}\|) \xrightarrow{p} N^{'} \in p - B_F(x_0, y_0)_\infty \setminus \{0\}$. Passing (3.9) to the limit yields

$$M^{'}w + aN^{'}(v,v) \notin -\text{int}\, C.$$

(γ) If $(2t_n/r_n)\|N_n^{'}\| \to \infty$, then dividing (3.9) by $(2t_n/r_n)\|N_n^{'}\|$ and passing to the limit gives

$$N^{'}(v,v) \notin -\text{int}\, C$$

If F is single-valued, Theorem 3.3 collapses to Theorem 4.10 of [13]. The example below gives an application of Theorem 3.3 to a multivalued case. ■

Example 3.5. Let $X = \mathbb{R}^2$, $Y = \mathbb{R}$, $S = \{(x,z) \in \mathbb{R}^2 \mid z = |x|^{3/2}\}$, $C = \mathbb{R}_+$, $(x_0, z_0) = (0,0) \in X$, $y_0 = 0 \in Y$, and

$$F(x,z) = \begin{cases} \left\{y \in \mathbb{R} \mid -\frac{2}{3}|x|^{\frac{3}{2}} + z^2 - z \le y \le \frac{1}{x^2+z^2}\right\} & \text{if } (x,z) \ne (0,0) \\ \{0\} & \text{if } (x,z) = (0,0). \end{cases}$$

Then, for a fixed $\alpha < 0$,

$$T(S,(x_0,z_0)) = \{(x,z) \in \mathbb{R}^2 \mid z = 0\},$$

$$A_F((x_0,z_0),y_0) = \{(0,-1)\},$$

$$B_F((x_0,z_0),y_0) = \left\{ \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \mid t < \alpha \right\},$$

$$clB_F((x_0,z_0),y_0) = \left\{ \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \mid t \le \alpha \right\},$$

$$B_F((x_0,z_0),y_0)_\infty = \left\{ \begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix} \mid t \le 0 \right\}.$$

Taking $v = (1,0) \in T(S,(x_0,z_0))$ one has

$$A_F((x_0,z_0),y_0)v = \{0\} \subseteq -bdC,$$

$$T^2(S,(x_0,z_0),v) = \emptyset,$$

$$T^{''}(S,(x_0,z_0),v) = \mathbb{R} \times \mathbb{R}_+.$$

Hence, for $w = (0,1) \in T^{''}(S,(x_0,z_0),v)$ one obtains

$$(0,-1)w + N(v,v) = -1 + t < 0$$

for all $N \in B_F((x_0,z_0),y_0)_\infty$ and

$$N(v,v) = t < 0$$

for all $N \in B_F((x_0,z_0),y_0)_\infty \setminus \{0\}$. Taking into account Theorem 3.3, one sees that $((x_0,z_0),y_0)$ is not a local weakly efficient solution of problem (P) in this case.

Theorem 3.4 (Sufficient condition). *Consider problem (P) with X being finite-dimensional. Assume that* $x_0 \in S$ *and* $y_0 \in \mathrm{StrMin}_C F(x_0)$. *Assume further that* $(A_F^S(x_0,y_0), B_F^S(x_0,y_0))$ *is an asymptotically p-compact second-order strong approximation of F at* (x_0,y_0) *with* $A_F^S(x_0,y_0)$ *being norm bounded. Then* $(x_0,y_0) \in LFE(2,P)$ *if*

(i) For all $v \in T(S,x_0) \setminus \{0\}$, $A_F^S(x_0,y_0)v \subseteq clC$.

(ii) For each $v \in T(S,x_0) \setminus \{0\}$ *with* $\overline{M}v \in -clC$ *for some* $\overline{M} \in p-clA_F^S(x_0,y_0)$, *for each* $N \in p-B_F^S(x_0,y_0)_\infty \setminus \{0\}$, *one has* $N(v,v) \notin -clC$ *and*

(a) $\forall w \in T^2(S,x_0,v) \cap v^\perp$, $\forall M \in p-clA_F^S(x_0,y_0)$, $\forall N \in p-clB_F^S(x_0,y_0)$, $Mw + 2N(v,v) \notin -clC$.

(b) $\forall w \in T''(S,x_0,v) \cap v^\perp \setminus \{0\}$, $\forall M \in p-clA_F^S(x_0,y_0)$, $\forall N \in p-B_F^S(x_0,y_0)_\infty$, $Mw + N(v,v) \notin -clC$.

Proof. Suppose to the contrary that $x_n \in S \cap B_X(x_0,(1/n)) \setminus \{x_0\}$ exists such that

$$(F(x_n) - y_0) \cap \left(B_Y\left(0, \frac{1}{n}t_n^2\right) - C \right) \neq \emptyset, \tag{3.10}$$

where $t_n = \|x_n - x_0\|$. We can assume that $(1/t_n)(x_n - x_0) \to v \in T(S,x_0) \setminus \{0\}$. By (10) and by the definition of first-order strong approximations, for large n, there exist $\overline{M}_n \in A_F^S(x_0,y_0)$ and $o(t_n)$ such that

$$\overline{M}_n(x_n - x_0) + o(t_n) \in B_Y\left(0, \frac{1}{n}t_n^2\right) - C. \tag{3.11}$$

The norm boundedness of $A_F^S(x_0,y_0)$ allows us to assume that $\overline{M}_n \xrightarrow{\text{p}} \overline{M} \in \text{p}-\text{cl}A_F^S(x_0,y_0)$. Dividing (3.11) by t_n we get, in the limit, $\overline{M}v \in -\text{cl}\ C$. According to Lemma 3.1, there are only the following two possibilities.

(α) One has $w_n := (x_n - x_0 - t_n v)/\frac{1}{2}t_n^2 \to w \in T^2(S,x_0,v) \cap v^\perp$. By the definition of the second-order strong approximation, (10) implies the existence of $M_n \in A_F^S(x_0,y_0)$, $N_n \in B_F^S(x_0,y_0)$ and $o(\|x_n - x_0\|^2)$ such that, for large n,

$$M_n(x_n - x_0) + N_n(x_n - x_0, x_n - x_0) + o(\|x_n - x_0\|^2) \in B_Y\left(0, \frac{1}{n}t_n^2\right) - C.$$

This can be rewritten as

$$M_n w_n + 2N_n\left(v + \frac{1}{2}t_n w_n, v + \frac{1}{2}t_n w_n\right) + o(t_n^2)/\frac{1}{2}t_n^2 = d_n/\frac{1}{2}t_n^2 - c_n', \tag{3.12}$$

where $d_n \in B_Y(0,(1/n)t_n^2)$ and $c_n' = (c_n + t_n M_n v)/\frac{1}{2}t_n^2 \in cl\ C$, because $c_n \in C$ and $A_F^S(x_0,y_0)v \subseteq \text{cl}\,C$. We can assume that $M_n \xrightarrow{\text{p}} M \in \text{p}-\text{cl}A_F^S(x_0,y_0)$. If $\{N_n\}$

is norm bounded, we can assume further that $N_n \xrightarrow{\text{p}} N \in \text{p} - \text{cl}B_F^S(x_0, y_0)$. In the limit (3.12) gives the contradiction

$$Mw + 2N(v,v) \in -\text{cl}\, C.$$

If $\{N_n\}$ is unbounded, we can assume that $\|N_n\| \to \infty$ and $(N_n/\|N_n\|) \xrightarrow{\text{p}} N \in \text{p} - B_F(x_0,y_0)_\infty \setminus \{0\}$. We divide (3.12) by $\|N_n\|$ and pass it to limit to get $N(v,v) \in -\text{cl}\, C$, also a contradiction.

(β) There is $r_n \to 0^+$ such that $(t_n/r_n) \to 0^+$ and

$$w_n := (x_n - x_0 - t_n v)/\frac{1}{2}t_n r_n \to w \in T''(S, x_0, v) \cap v^\perp \setminus \{0\}.$$

Similarly as for the case (α), there are $M_n \in A_F^S(x_0,y_0)$, $N_n \in B_F^S(x_0,y_0)$, and $o(t_n^2)$ such that, for large n,

$$M_n w_n + \left(\frac{2t_n}{r_n}\right) N_n \left(v + \frac{1}{2}r_n w_n, v + \frac{1}{2}r_n w_n\right) + o(t_n^2)/\frac{1}{2}t_n r_n = d_n/\frac{1}{2}t_n r_n - c_n^{'}, \tag{3.13}$$

where $d_n \in B_Y(0,(1/n)t_n^2)$ and $c_n^{'} = (c_n + t_n M_n v)/\frac{1}{2}t_n r_n \in \text{cl}\, C$. We can assume that $M_n \xrightarrow{\text{p}} M \in \text{p} - \text{cl}A_F^S(x_0,y_0)$. There are three subcases as follows.

- $(2t_n/r_n)N_n \to 0$. Passing (3.13) to limit one gets $Mw \in -\text{cl}\ C$, contradicting assumption (ii) (b) (with $N = 0 \in \text{p} - B_F^S(x_0,y_0)_\infty$).
- $(2t_n/r_n)\|N_n\| \to a > 0$. Then $\|N_n\| \to \infty$ and we can assume that $(N_n/\|N_n\|) \xrightarrow{\text{p}} N \in \text{p} - B_F(x_0,y_0)_\infty \setminus \{0\}$. Dividing (3.13) by $(2t_n/r_n)\|N_n\|$ and passing to limit we obtain the contradiction

$$Mw + aN(v,v) \in -\text{cl}\, C.$$

- $(2t_n/r_n)\|N_n\| \to \infty$. Then $\|N_n\| \to \infty$ and assume that $(N_n/\|N_n\|) \xrightarrow{\text{p}} N \in \text{p} - B_F(x_0,y_0)_\infty \setminus \{0\}$. Dividing (3.13) by $(2t_n/r_n)\|N_n\|$ we get in the limit $N(v,v) \notin -\text{cl}\, C$ which is absurd. ■

Theorem 3.4 strictly contains Theorem 4.12 of [13] as a special case. We interpret the use of Theorem 3.4 by the following example.

Example 3.6. Let $X = \mathbb{R}$, $Y = \mathbb{R}^2$, $S = [0,+\infty)$, $C = \mathbb{R}_+^2$, $x_0 = 0$, $(y_0,z_0) = (0,0) \in Y$, and

$$F(x) = \begin{cases} \{(y,z) \in \mathbb{R}^2 \mid y = x^2, \frac{3}{4}|x|^{4/3} \le z \le |x|^{4/3}\} & \text{if } x \ge 0, \\ \emptyset & \text{if } x < 0. \end{cases}$$

Then $(y_0,z_0) \in \text{StrMin}_C F(x_0)$, $T(S,x_0) = S$ and, for a fixed $\alpha > 0$,

$$A_F^S(x_0,(y_0,z_0)) = \{0\} \times [0,1] = \text{cl}\, A_F^S(x_0,(y_0,z_0)),$$

$$B_F^S(x_0,(y_0,z_0)) = \{(1,z) \mid z > \alpha\},$$

$$\mathrm{cl}B_F^S(x_0,(y_0,z_0)) = \{(1,z) \mid z \geq \alpha\},$$

$$B_F^S(x_0,(y_0,z_0))_\infty = \{(0,z) \mid z \geq 0\}.$$

It is easy to check that, for all $v \in T(S,x_0) \setminus \{0\}$, one has

$$A_F^S(x_0,(y_0,z_0))v = \{(0,\beta v) \mid \beta \in [0,1]\} \subseteq \mathrm{cl}\, C,$$

$$N(v,v) = (0,zv^2) \notin -\mathrm{cl}\, C,$$

$\forall N \in B_F^S(x_0,(y_0,z_0))_\infty \setminus \{0\}$, and

$$Mw + 2N(v,v) = (2v^2, 2zv^2) \notin -\mathrm{cl}\, C,$$

$\forall w \in T^2(S,x_0,v) \cap v^\perp = \{0\}$, $\forall M \in \mathrm{cl}A_F^S(x_0,(y_0,z_0))$, $\forall N \in \mathrm{cl}B_F^S(x_0,(y_0,z_0))$, and $T''(S,x_0,v) \cap v^\perp \setminus \{0\} = \emptyset$. Now that all assumptions of Theorem 3.4 are satisfied, $(x_0,(y_0,z_0)) \in \mathrm{LFE}(2,\mathrm{P})$.

Summarizing, it should be noted that each of the necessary and sufficient conditions presented in this chapter is an extension to the multivalued case of the corresponding result in [13] for the single-valued case. The results of [13] were shown in [13] to be sharper than the corresponding theorems in [15] and better in use than many recent results in the literature, because the assumptions are very relaxed.

References

1. Allali, K., Amahroq, T. Second-order approximations and primal and dual necessary optimality conditions. *Optimization* **40** (1997), 229–246.
2. Amahroq, T., Gadhi, N. Second-order optimality conditions for the extremal problem under inclusion constraints. *J. Math. Anal. Appl.* **285** (2003), 74–85.
3. Aubin, J. P., Frankowska, H. *Set-Valued Analysis*, Birkhäuser, Boston, 1990.
4. Cambini, A., Martein, L., Vlach, M. Second-order tangent sets and optimality conditions. *Math. Japon.* **49** (1999), 451–461.
5. Clarke, F. H. *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York, 1983.
6. Demyanov, V. F., Rubinov, A. M. On quasidifferentiable functionals. *Soviet Math. Doklady* **21** (1980), 14–17.
7. Ginchev, I., Guerraggio, A., Rocca, M. Second-order conditions for $C^{1,1}$ constrained vector optimization. *Math. Program.* **104** (B) (2005), 389–405.
8. Jeyakumar, V., Luc, D. T. Approximate Jacobian matrices for nonsmooth continuous maps and C^1-optimization. *SIAM J. Control Optim.* **36** (1998), 1815–1832.
9. Jiménez, B. Strict efficiency in vector optimization. *J. Math. Anal. Appl.* **265** (2002), 264–284.
10. Jiménez, B., Novo, V. Optimality conditions in differentiable vector optimization via second-order tangent sets. *Appl. Math. Optim.* **49** (2004), 123–144.
11. Jourani, A., Thibault, L. Approximations and metric regularity in mathematical programming in Banach spaces. *Math. Oper. Res.* **18** (1992), 390–400.
12. Khanh, P. Q., Tuan, N. D. First and second-order optimality conditions using approximations for nonsmooth vector optimization in Banach spaces. *J. Optim. Theory Appl.* **130** (2006), 289–308.

13. Khanh, P. Q., Tuan, N. D. First and second-order approximations as derivatives of mappings in optimality conditions for nonsmooth vector optimization. *Appl. Math. Optim.* **58** (2008), 147–166.
14. Khanh, P. Q., Tuan, N. D. Optimality conditions using approximations for nonsmooth vector optimization problems under general inequality constraints. *J. Convex Anal.* **16** (2009), 169–186.
15. Khanh, P. Q., Tuan, N. D. Optimality conditions for nonsmooth multiobjective optimization using Hadamard directional derivatives. *J. Optim. Theory Appl.* **133** (2007), 341–357.
16. Khanh, P. Q., Tuan, N. D. Approximations as generalized differentiability and optimality conditions without continuity assumptions in multivalued optimization. *Inter. Confer. Nonlinnear Analysis in Engineering Mechanics Today*, Hochiminh City, December 12–15, 2006.
17. Khanh, P. Q., Tung, N. M. Approximations as generalized derivatives: Calculus and optimality conditions in multivalued optimization, submitted for publication.
18. Penot, J. P. Recent advances on second-order optimality conditions. *Optimization* (Nguyen, V. H., Strodiot, J. J., Tossings, P., Eds), Springer, Berlin, 357–380, 2000.

Chapter 4
Variational Inequality and Complementarity Problem

Sudarsan Nanda

4.1 Introduction

Variational inequality and complementarity have much in common, but there has been little direct contact between the researchers of these two related fields of mathematical sciences. Several problems arising from fluid mechanics, solid mechanics, structural engineering, mathematical physics, geometry, mathematical programming, and so on have the formulation of a variational inequality or complementarity problem. People working in applied mathematics mostly deal with the infinite-dimensional case and they deal with variational inequality whereas people working in operations research mostly deal with the finite-dimensional problem and they use the complementarity problem. Variational inequality is a formulation for solving the problem where we have to optimize a functional. The theory is derived by using the techniques of nonlinear functional analysis such as fixed point theory and the theory of monotone operators, among others.

In this chapter we give a brief review of the subject. The chapter is divided into four sections. Section 4.2 deals with nonlinear operators, which are required to describe the results. Sections 4.3 and 4.4, respectively, deal with variational inequality and the complementarity problem. Section 4.5 describes semi-inner-product spaces and variational inequality in semi-inner-product spaces.

4.2 Nonlinear Operators

In this section we discuss certain nonlinear operators, which are useful in the study of variational inequalities and the complementarity problem.

Let X be a real normed linear space and let X^* be the dual space of X. Let the pairing between $x \in X$ and $x^* \in X^*$ be denoted by (x^*, X). Let T be a mapping of the subset $D(T)$ of X into X^*. T is said to be monotone if

Sudarsan Nanda
Department of Mathematics, IIT Kharagpur, India, e-mail: snanda@maths.iitkgp.ernet.in

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_4,

$$(T_x - T_y, x-y) \geq 0, \qquad \forall x,y \in D(T),$$

and strictly monotone if T is monotone and the strict inequality holds whenever $x \neq y$. T is called α-monotone if there is a continuous strictly increasing function $\alpha : [0,1) \rightarrow [0,1]$ with $\alpha(0) = 0$ and $\alpha(r) \rightarrow \infty$ such that

$$(T_x - T_y, x-y) \geq \|x-y\|\alpha(\|x-y\|)$$

for all x,y in $D(T)$. T is strongly monotone if $\alpha(r) = cr$ for some constant $c > 0$. T is coercive on subset K of $D(T)$ if there exists a function $c : (0,\infty) \rightarrow [0,\infty]$ where $c(r) \rightarrow \infty$ as $r \rightarrow \infty$ such that

$$(Tx,x) \geq \|x\|c(\|x\|), \qquad \forall x \in K.$$

Thus T is coercive on K if K is bound, and T is coercive on an unbounded K if and only if

$$\frac{(Tx,x)}{\|x\|} \rightarrow \infty \quad \text{as} \quad \|x\| \rightarrow \infty, \qquad x \in K.$$

T is hemicontinuous if $D(T)$ is convex and for any x,y in $D(T)$, the map $t \mapsto T$ $(tx+(1-t)y)$ of $[0,1]$ to X^* is continuous for the natural topology of $[0,1]$ and the weak topology of X^*.

Example 4.1.

(a) Let $f : R \rightarrow R$ be a monotonically increasing function. Then f is a monotone operator.
(b) Let H be a Hilbert space and $T : H \rightarrow H$ be a compact self-adjoint linear operator. Then T is a monotone operator if all the eigenvalues of T are nonnegative.
(c) Let H be a Hilbert space. An operator $T : H \rightarrow H$ is said to be nonexpansive if $\|Tx - Ty\| \leq \|x-y\|$ for all $x,y \in H$.
If T is nonexpansive, then $I - T$ is monotone operator.
(d) Let H be a Hilbert space and C a closed convex subset of H. Let Px denote the point of minimum distance of C from x; that is,

$$Px = \left\{ z \in C : \|z-x\| = \inf_{y \in C} \|y-x\| \right\}.$$

Then P is monotone operator on H,
(e) Let H be a Hilbert space. Then an operator $T : H \rightarrow H$ is said to be accretive if $\|x-y\| \leq \|Tx - Ty\|$ for all $x,y \in H$.
Then $T : H \rightarrow H$ is monotone iff $(I + \lambda T)$ is accretive for every $\lambda > 0$.

Theorem 4.1. *If $T : D(T) \subset X \rightarrow X^*$ is α-monotone, then it is strictly monotone (hence monotone) and coercive. In particular every strongly monotone operator is strictly monotone and coercive.*

Definition 4.1. Let X be nls and let X^* be its dual. A map $T : X \rightarrow X^*$ is said to be a *duality map* if for any $x \in X$.
(i) $(T_x, x) = \|T_x\|\|x\|$ and (ii) $\|T_x\| = \|x\|$.

A duality map can be constructed in any nls in the following way. By the Hahn–Banach theorem, for any $x \in X$, there exists at least one bounded linear functional $Y_x \in X^*$ such that $\|y_x\| = 1$ and $(y_x, x) = \|x\|$. Taking one such functional y_x and setting $T_x = \|x\| y_x$ and $T(-x) = \|x\| y_x$, we get $\|T_x\| = \|x\|$ and $(T_x, x) = \|T_x\| \|x\|$.

Theorem 4.2. *In general a duality map $T : X \to X^*$ is multivalued. It is single-valued if X^* is strictly convex.*

Theorem 4.3. *If $T : X \to X^*$ is a duality map, then it is monotone and coercive; if furthermore X is strictly convex, then T is strictly monotone.*

Theorem 4.4. *Let X be a real Banach space and $F : X \to X^*$ be a nonlinear operator. If Gateaux derivative $F'(x)$ exists for every $x \in X$ and is positive-semi-definite, then F is monotone.*

Theorem 4.5. *Let f be a proper convex function defined on X. If f is differentiable, the ∇f is monotone.*

Theorem 4.6. *Let f be a proper differentiable function defined on X. If ∇f is monotone, then f is convex.*

4.3 Variational Inequalities

In this section we discuss some basic properties of variational inequalities. Before we state the definition we first discuss some examples where variational inequalities arise.

Example 4.2. Let $I = [a, b] \subset R$. Let f be a real-valued differentiable function defined on I. Suppose we seek the points $x \in I$ for which

$$f(x) = \min_{y \in I} f(y).$$

Then three cases arise in this case:

(i) $a < x < y \Rightarrow f'(x) = 0$.
(ii) $a = x \Rightarrow f'(x) \geq 0$.
(iii) $x = b \Rightarrow f'(x) \leq 0$.

All three cases can be put together as a single inequality as follows.

$$f'(y)(y - x) \geq 0, \qquad \forall y \in I.$$

This is an example of variational inequality.

Example 4.3. Let K be a closed convex set in R^n and let $f : K \to R$ be differentiable. We characterize the points $x \in K$ for which

$$f(x) = \min_{y \in K} f(y).$$

If there exists $x \in K$ that satisfies the above equation and if $F(x) = grad f(x)$, then x is a solution of the following inequality

$$x \in K : (Fx, y - x) \geq 0, \qquad \forall y \in K.$$

Conversely, if f is differentiable and convex and if the above inequality is satisfied by x, then

$$f(x) = \min_{y \in K} f(y).$$

Example 4.4. Let Ω be a bounded open domain in R^n with the boundary T. In some problems of mechanics we seek a real-valued function $x \to u(X)$ which in Ω, satisfies the classical equation

$$-\Delta u - u = f, \qquad f \in \Omega, \qquad u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i{}^2} \tag{4.1}$$

with boundary conditions

$$u \geq 0, \qquad \frac{\partial u}{\partial v} \geq 0, \qquad u \frac{\partial u}{\partial v} = 0 \quad \text{on} \quad \Gamma, \tag{4.2}$$

where $\partial / \partial v$ denotes differentiation along the outward normal to Γ. If we write

$$J(v) = \frac{1}{2} a(v, v) - (f, v),$$

where

$$a(u, v) = \sum_{i=1}^{n} \int_{\Omega} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} dx + \int_{\Omega} uv, dx$$

$$(f, v) = \int_{\Omega} f v dx$$

and if we introduce the closed convex set K defined by

$$K = \{v : v \geq 0 \text{ on } \Gamma\}$$

then the problem given by (4.1) and (4.2) is equivalent to finding $u \in K$ such that

$$J(u) = \inf_{v \in K} J(v).$$

This admits a unique solution u characterized by

$$u \in K, \qquad a(u, v - u) \geq (f, v - u) \qquad \forall v \in K.$$

This is called a variational inequality problem.

We now state the problem in the most general setting.

Let X be a reflexive real Banach space and let X^* be its dual. Let T be a monotone hemicontinuous mapping from X into X^* and let K be a nonempty, closed, convex subset of the domain $D(T)$ of T. Then a variational inequality is stated as follows.

$$x \in K : (Tx, y - x) \geq 0, \qquad \forall y \in K. \tag{4.3}$$

Any $x \in X$ that satisfies (4.3) is called a solution of the variational inequality. We write $S(T,K)$ to denote the set of all solutions of the variational inequality (4.3). In fact we consider a more general inequality which is stated as follows.

For each given element $w_0 \in X^*$,

$$x \in K : (Tx - w_0, y - x) \geq 0, \qquad \forall y \in K. \tag{4.4}$$

Inequality (4.3) can also be written by replacing the subset K of X by an extended real-valued function defined on X. For any subset K of X, let δ_k, called the indicator function of K, be the function defined on X by

$$\delta_k(y) = \quad \begin{cases} 0 & \text{if } y \in K \\ \infty & \text{if } y \notin K. \end{cases}$$

Then it is easy to verify that $x \in K$ is a solution of (4.3) if and only if

$$(Tx, y - x) \geq \delta_k(x) - \delta_k(y) \qquad \forall y \in K.$$

Therefore we consider, as a generalization of inequality (4.3), the inequalities of the form:

$$x \in X : (Tx, y - x) \geq f(x) - f(y) \qquad \forall y \in X, \tag{4.5}$$

where f is an arbitrary extended real-valued function defined on X.

Observe that if $f = 0$, then (4.5) reduces to the VI(4.3) and if $T = 0$, then we are in the framework of the calculus of variations where we minimize the extended real-valued f; that is, we have

$$f(x) \leq f(y), \qquad \forall y \in X \quad \text{or} \quad f(x) = \min_{y \in X} f(y).$$

Theorem 4.7. *Let T be a monotone, hemicontinuous mapping of subset $D(T)$ of X into X^* and K a convex subset of $D(T)$. Then for given element $w_0 \in X^*$, any solution of inequality* (4.4) *is also a solution of the inequality*

$$(Ty - w_0, y - x) \geq 0 \qquad \forall y \in K. \tag{4.6}$$

Theorem 4.8. *Let T be a hemicontinuous map of X into X^*. Suppose that for any pair of vectors $x_0 \in K$ and $w_0 \in X^*$,*

$$(Ty - w_0, y - x_0) \geq 0 \qquad \forall y \in X. \tag{4.7}$$

Then $TX_0 = w_0$.

The following result gives uniqueness of solutions when it exists.

Theorem 4.9. *If the mapping T from X into X^* is strictly monotone, then the inequality* (4.4) *can have at most one solution.*

Theorem 4.10. *If either the mapping T is strictly monotone or the function f is strictly convex, then the inequality* (4.5) *can have at most one solution.*

We now state the following fundamental theorem for variational inequality that appears in Hartmann and Stampacchia [20]

Theorem 4.11. *Let T be a monotone hemicontinuous map of a closed convex subset K of a reflexive real Banach space X, with $0 \in K$, into X^* and if K is not bounded, let T be coercive on K. Then for each given element $w_0 \in X^*$ there is an $x \in K$ such that the inequality* (4.4) *holds:*

$$x \in K : (Tx - w_0, y - x) \geq \qquad \forall y \in K.$$

4.4 Complementarity Problem

Several problems arising in various fields such as mathematical programming, game theory, economics, mechanics, and geometry have the mathematical formulation of a complementarity problem.

Definition 4.2. Let X be a reflexive real Banach space and let X^* be its dual. Let K be a closed convex cone in X with $0 \in K$. The *polar* of K is the cone K^* defined by

$$K^* = \{y \in X^* : (y,x) \geq 0 \; \forall x \in K\}.$$

Obviously $K^* \neq \phi$ because $0 \in K^*$. It also easy to see that K^* is a closed convex cone in X^*. Let T be a map from K into X^*. Then the complementarity problem (CP in short) is to find an $x \in X$ such that

$$x \in K, \qquad Tx \in K^*, \qquad (Tx,x) = 0.$$

The following theorem proves the equivalence between the complementarity problem and the variational inequality over a closed convex cone. We write

$$S(T,K) = \{x : x \in K, (Tx, y - x) \geq 0 \quad \forall y \in K\}$$

and $C(T,K) = \{x : x \in K, Tx \in K^*, (Tx,x) = 0\}$. We have the following.

Theorem 4.12. $C(T,K) = S(T,K)$.

Remark 4.1.

(a) It should be noted that the solution of a complementarity problem, if it exists, is unique if the operator T is strictly monotone. $C(T,K) = S(T,K)$ for a closed convex cone K, therefore the proof is same as that of Theorem 4.9.

(b) Regarding the existence it must be noted that the solution may not exist only under the assumption of hemicontinuity and monotonicity (even strict monotonicity) of the operator T. For example, let $X = R, K = \{x \in R : x \geq 0\}$, so that $K = K^*$ and K is a closed convex cone. Let $T : K \to R$ be defined by

$$Tx = -\frac{1}{1+x}.$$

Then T is hemicontinuous and strictly monotone. $(Tx, x) = 0$ implies $x = 0$ but $TO = -1 \notin K^*$.

We now discuss the existence of solutions of the complementarity problem. We have the following.

Theorem 4.13. *Let $T : K \to X^*$ be hemicontinuous, monotone, and coercive. Then the complementarity problem has a solution. In particular if T is hemicontinuous and α-monotone, then the solution exists and is unique.*

Theorem 4.14. *Let $T : K \to X^*$ be hemicontinuous and monotone and let $TO \in K^*$. Then the complementarity problem has a solution.*

Theorem 4.15. *Let $T : C \to B^*$ be hemicontinuous and monotone such that there is an $x \in C$ with $Tx \in int\, C^*$. Then there is an x_0 such that*

$$x_0 \in C, \qquad Tx_0 \in C^*, \quad \textit{and} \quad (Tx_0, x_0) = 0. \tag{4.8}$$

If, furthermore, T is strictly monotone, then there is a unique x_0 satisfying (4.8).

In order to prove the theorem we need the following result, which is due to Browder. See Browder [5] and Mosco [33]. This is a special case of Theorem 4.7.

Let T be a monotone hemicontinuous map of a closed, convex, bounded subset K of B, with $0 \in K$, into B^*. Then there is an $x_0 K$ such that

$$(Tx_0, y - x_0) \geq 0 \qquad \forall y \in K.$$

Now observe that if $e \in C^*$ but $e \notin \text{int}\, C^*$, the sets $D_r(e)$ need not be bounded. In this case we cannot conclude $y = 0$ from the fact that $(e, y) = 0$. Consider the case when $B = R^2$, $C = R^2$, and $e = (1,0)$. Then for each $r > 0$, $D(e)$ contains the positive y-axis and hence is unbounded.

We note that this theorem fails to hold if the requirement that there exists $x \in C$ with $Tx \in \text{int}\, C^*$ is dropped.

Take $B = R^3$, $C = \{(x,y,z) \in R^3 : x, z \geq 0, 2xz \geq y^2\}$. Define T by $T(x,y,z) = (x+1, y+1, 0)$. Then T is monotone, hemicontinuous (even bounded); $(1,-1,1) \in C$ and $T(1,-1,1) = (2,0,0) \in C^*$. If $u = (x,y,z \in C)$ with $Tu \in C^*$, then $y = -1$ and hence $x > 0$. Hence for any such u, $(Tu, u) = x(x+1) > 0$.

Finite-Dimensional Case

Let K be a closed convex cone in R^n and f a map from K into R^n such that

$$x \in K, \qquad f(x) \in K^*, \qquad (f(x),x) = 0.$$

In particular, if $K = R^n_+$

$$\{x = (x_1,x_2,\ldots,x_n) \in R^n : x_i \geq 0, i = 1,2,\ldots,n\},$$

then the problem can be stated as follows.

$$x \geq 0, \qquad f(x) \geq 0, \qquad (f(x),x) = \sum_{i=1}^{n} x_i f_i(x) = 0.$$

If, furthermore, $f(x) = Mx + b$ where M is a given real square matrix of order n and b is a given column vector in R^n, then the above problem is called a linear complementary problem (LCP in short) and it can be stated as follows.

Finf $w_1,w_2,\ldots,w_n$ and $x_1,x_2,\ldots,x_n$,

$$w = Mx + b, \qquad w \geq 0, \qquad x \geq 0, \qquad w_i x_i = 0, \qquad i = 1,2,\ldots,n.$$

Otherwise, in general, the problem is known as a nonlinear complementarity problem (NCP in short).

We now illustrate the LCP by a numerical example.

As a specific example of an LCP in R^n, let

$$n = 2, \qquad \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad \begin{pmatrix} -5 \\ -6 \end{pmatrix}$$

In this case the problem is to solve:

$$w_1 - 2x_1 - x_2 = -5$$
$$w_2 - x_1 - 2x_2 = -6$$
$$w_1,\ w_2,\ x_1,\ x_2 \geq 0,\ w_1 x_1 = w_2 x_2 = 0.$$

This can be expressed in the form of a vector equation as

$$w_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + w_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + x_1 \begin{pmatrix} -2 \\ -1 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} -5 \\ -6 \end{pmatrix} \tag{4.9}$$

$$w_1,\ w_2,\ x_1,\ x_2 \geq 0, \qquad w_1 x_1 = w_2 x_2 = 0. \tag{4.10}$$

The solution for the given LCP is $(w_1,w_2,x_1,x_2) = (0,0,(4/3)(7/3))$

As special cases we have the following results for R^n.

Theorem 4.16. *Let $f : K \to R^n$ be continuous, monotone, and have a solution. In particular if f is continuous and strongly monotone, then the solution exists uniquely.*

Theorem 4.17. *Let $f : K \to R^n$ be continuous, monotone, and such that $f(0) \in K^*$ (or $f(0) = 0$). Then there exists a solution to the NCP.*

Theorem 4.18. *Let $f : K \to R^n$ be continuous, monotone, and such that there exists an $x \in K$ with $f(x) \in \text{int}K^*$. Then there exists a solution to the NCP. Lemke* [27] *and Eaves* [13] *discussed the existence of stationary points and the nature of the set of all stationary points of the pair (f,K) in R^n where $K = R^n_+$. Lemke* [27] *discussed the linear case by considering affine functions. A basic theorem of Lemke* [27] *states the following.*

Theorem 4.19 (Lemke). *Given an affine map $f : R^n_+ \to R^n$ and $d \in R^n_+$ there is a piecewise affine map $x : R_+ \to R^n_+$ such that $x(t)$ is a stationary point of (f, D^n_t) with $d.x(t) = t$ where $D^n_t = \{x \in R^n_+ : d.x \leq t\}$.*

Definition 4.3. Let M be a square matrix of order n. M is said to be *positive-definite* if

$$y^T My = \sum_{i=1}^{n} \sum_{j=1}^{n} y_i m_{ij} y_j > 0 \qquad \forall 0 \neq y \in R^n,$$

positive-semidefinite if

$$y^T My \geq 0, \qquad \forall y \in R^n,$$

and *copositive matrix* if $y^T My \geq 0$ for all $y \geq 0$ and *strictly copositive* if strictly inequality holds for all $y \geq 0$,

Copositive plus if it is a copositive matrix and if $y^T(M + M^T) = 0$, whenever $y \geq 0$ satisfies $y^T My = 0$,

P-matrix if all principal subdeterminents of M are positive, *Q-matrix* if the LCP has a solution for every $q \in R^n$, *nondegenerate matrix* if all its principal subdeterminants are nonzero, *degenerate matrix* if it is not nondegenerate, *z-matrix* if m_{ij} for all $i \neq j$, and *J-matrix* if

$$Mz \geq, \qquad z^T Mz \geq 0, \qquad z \geq 0 \Rightarrow z = 0.$$

4.5 Semi-Inner-Product Space and Variational Inequality

In this section we discuss the concept of semi-inner product (sip. for short), which was introduced by Lumer in the year 1961 and subsequently studied by Giles and several other mathematicians. We then study variational inequality on sip space.

Let V be a complex vector space. A sip on V is a complex function $[,]$ on $V \times V$ with the following properties: for $x, y, z \in V$ and $\lambda \in C$,

(i) $[x + y, z] + [x, z] + [y, z]$,
$[\lambda x, y] = \lambda [x, y]$.

(ii) $[x,x] > 0$ for $x \neq 0$.
(iii) $|[x,y]|^2 \leq [x,x][y,y]$.

V along with a sip defined on it is called a sip space. A sip space V has the homogeneity property when the sip satisfies

(iv) $[x,\lambda y] = \bar{\lambda}[x,y]$.

With the aim of carrying over Hilbert-space type arguments to the theory of Banach spaces Lumer [30] introduced the concept of sip. But the generality of the axiom system defining the sip is a serious limitation of any extensive development of a theory of sip spaces parallel to the theory of inner-product spaces. Let X be a normed linear space and let X^* be its dual.

The unit ball of X is

$$U = \{x \in X : \|x\| \leq 1\} \text{ and its boundary.}$$
$$S = \{x \in X : \|x\| = 1\} \text{ is the unit sphere of } X.$$
$$U^* = \{f \in X^* : \|f\| \leq 1\} \text{ is the unit ball and.}$$
$$S^* = \{f \in X : \|f\| = 1\} \text{ is the unit sphere of } X^*.$$

The conjugate norm is also denoted by $\|.\|$.

Theorem 4.20. *A sip space V is a normed linear space with the norm* $\|x\| = [x,x]^{1/2}$. *Every normed linear space can be made into a sip space (in general, in infinitely many different ways) with the homogeneity property.*

Theorem 4.21. *A Hilbert space H can be made into a sip space in a unique way. A sip space is an ip space if and only if the norm it induces satisfies parallelogram law.*

Continuous and Uniform Sip Spaces

A continuous sip space is a sip space V where the sip has the additional property:

(v) For all $(x,y) \in S \times S$,

$Re\,\lfloor y, x+\lambda y\rfloor \to Re\,\lfloor y,x\rfloor$ for all $\lambda \to 0$.

The space is a uniform continuous sip space when the above limit is approached uniformly for all $(x,y) \in S \times S$.

A uniform sip space is a uniformly continuous sip space where the induced normed linear space is uniformly continuous and complete.

Examples (L_p space for $1 < p < \infty$). The real Banach space $L_p(X,S,\mu)$, where $1 < p < \infty$, can readily be expressed as a uniform sip space with sip defined by

$$[y,x] = \frac{1}{\|x\|}\int_X y|x|^{p-1} sgn\,x\, d\mu$$

For x,y in any sip space V, x is said to be normal to y and y is transversal to x if $[y,x]=0$. A vector $x \in V$ is normal to a subspace N and N is transversal to x if x is normal to each $y \in N$.

A Banach space X is said to be smooth at a point $x \in S$ if and only if there exists a unique hyperplane of support at x, that is, there exists only one continuous linear functional $1_x \in E^*$ with $\|1_x\| = 1$ and $1_x(x) = 1$. X is said to be a smooth Banach space if it is smooth at every $x \in S$.

The norm of X is said to be Gateaux differentiable if for all $x,y \in S$ and real λ,

$$\lim_{\lambda \to 0} \frac{\|x+\lambda y\| - \|x\|}{\lambda} \quad \text{exists.}$$

The norm is said to be uniformly Fréchet differentiable if this limit is approached uniformly for $(x,y) \in S \times S$. Note that X is smooth at $xinS$ if and only if the norm is Gateaux differentiable at x. We have the following.

Theorem 4.22. *In a continuous sip space x is normal to y if and only if $\|x+\lambda y\| > \|x\|$ for all complex numbers λ.*

Theorem 4.23. *A sip space is a continuous (uniformly continuous) sip space iff the norm is Gateaux (uniformly Frechet) differentiable.*

Lemma 4.1. *In a continuous sip space that is uniformly convex and complete in its norm, there exists a nonzero vector normal to every proper closed vector subspace.*

Lemma 4.2. *A sip space is strictly convex if whenever $[x,y] = \|x\|\|y\|, x,y \neq 0$; then $y = \lambda x$ for some real $\lambda > 0$.*

Theorem 4.24 (Generalized Riesz–Fischer Theorem). *In a continuous sip space V that is uniformly convex and complete in its norm, to every continuous linear functional $f \in V^*$, there exists a unique vector $y \in V$ such that*

$$f(x) = [x,y], \qquad x \in V.$$

Theorem 4.25. *For a uniform sip space M, the dual space M^* is a uniform sip space w. r. t. the sip defined by*

$$\lfloor f_x, f_y \rfloor = \lfloor y, x \rfloor.$$

Theorem 4.26. *Every finite-dimensional, strictly convex, continuous sip space is a uniform sip space.*

Theorem 4.27. *Let X be a continuous sip space that is uniformly convex and complete in its norm. If A is a bounded linear operator from X into itself, then there is a unique bounded linear operator A^+ such that*

$$[Ax,y] = \lfloor x, A^+, y \rfloor.$$

A^+ is called the generalized adjoint of A: The proof uses Theorem 4.24 *and is similar to that of the corresponding theorem for Hilbert space operators. Note that if X is a Hilbert space, then the generalized adjoint is the usual Hilbert space adjoint.*

We now discuss variational inequality and the complementarity problem in semi-inner-product space under certain contractive type conditions on the operators.

Let K be a closed convex subset of a sip X. If $T : K \to K$, then a variational inequality (VI for short) is stated as follows.

$$x \in K : \lfloor Tx, y - x \rfloor \geq 0 \qquad \forall y \in K.$$

If K is a closed convex cone, then the polar or dual of K, denoted by K^*, is defined by

$$K^* = \{z \in X^* : [z,x] \geq 0, \quad \forall x \in K\}.$$

If K is a closed convex cone, then the complementarity problem (CP for short) is defined as follows. $x \in X$ such that

$$x \in K, \qquad Tx \in K^* \quad \text{and} \quad [Tx, x] = 0.$$

Observe that if K is a closed convex cone, then (VI) and (CP) are equivalent.

We have the following.

Theorem 4.28. *Let X be uc and ss and K a nonempty, closed, convex subset of X. Let $T : K \to K$ satisfy any one of the following conditions.*

(i) $\|Tx - Ty\| \leq a\|x - y\| + b\|Tx - x\| + c\|Ty - y\|$,
where $-1 < a < a < 0, \quad b \geq 0, \quad c \geq 0, \quad a + b + c = 0$,

(ii) $\|Tx - Ty\| \leq a_1\|x - y\| + a_2\|x - Tx\| + a_3\|y - Ty\| + a_4\|x - Ty\| + a_5\|y - Tx\|$,
where $-1 < a_1 < 0, \quad a_2, a_3, a_4, a_5 \geq 0, \quad \sum_{i=1}^{5} a_i = 0$

Then there is a unique $y_0 \in K$ such that $\lfloor Ty_0, z - y_0 \rfloor \geq 0$ for all $x \in K$.

Proof. Because K is a nonempty, closed, convex subset of X and X is uc, for every $y \in K$ there is a unique $x \in K$ closest to $y - Ty$; that is,

$$\|x - y + Ty\| \leq \|z - y + Ty\|$$

for every $z \in K$ (see Edelstein [14]). Let the correspondence $y \mapsto x$ be denoted by θ. Let $z \in K$ and let $0 \leq \lambda \leq 1$. Because K is convex, $(1 - \lambda)x + \lambda z \in K$. Define a map

$$h : [0,1] \to R^+ \quad by$$

$$h(\lambda) = \|y - Ty - (1 - \lambda)x - \lambda z\|^2$$

X is uc and ss, thus h is a continuously differentiable function of λ and

$$h'(\lambda) = 2\lfloor y - Ty - (1 - \lambda)x - \lambda z, x - z \rfloor.$$

Because x is the unique element closest to $y - Ty$ we must have $h'(0) \geq 0$. Therefore

$$\lfloor y - Ty - x, x - z \rfloor \geq 0, \qquad \forall z \in K. \tag{4.11}$$

Let $y_1, y_2 \in K$ and $y_1 \neq y_2$. Let $\theta(y_1) = x_1, \quad \theta(y_2) = x_2$. It follows from (4.11) that

$$\lfloor y_1 - Ty_1 - \theta y_1, -\theta y_2 + \theta y_1 \rfloor \geq 0 \tag{4.12}$$

and

$$\lfloor y_2 - Ty_2 - \theta y_2, -\theta y_2 - \theta y_1 \rfloor \geq 0. \tag{4.13}$$

From (4.12) and (4.13) we get

$$\lfloor y_1 - Ty_1 - y_2 + Ty_2 - \theta y_1 + \theta y_2, \theta y_1 - \theta y_2 \rfloor \geq 0.$$

Therefore

$$\begin{aligned} \|\theta y_1 - \theta y_2\|^2 &= \lfloor \theta y_1 - \theta y_2, \theta y_1 - \theta y_2 \rfloor \\ &= \lfloor y_1 - Ty_1 - y_2 + Ty_2, \theta y_1 - \theta y_2 \rfloor \\ &\leq \|y_1 - Ty_1 - y_2 + Ty_2\| \|\theta y_1 - \theta y_2\|. \end{aligned}$$

If T satisfies condition (i) Theorem 4.28, then

$$\|\theta y_1 - \theta y_2\| \leq (1+a)\|y_1 - y_2\| + b\|Ty_1 - y_1\| + c\|Ty_2 - y_2\|,$$

where $0 < 1+a < 1$, $b \geq 0$, $c \geq 0$, $a+b+c = 1$.

If T satisfies condition (ii) then

$$\begin{aligned} \|Tx - Ty\| &\leq (1+a_1)\|x-y\| + a_2\|x - Tx\| \\ &\quad + a_3\|y - Ty\| + a_4\|x - Ty\| + a_5\|y - Tx\|, \end{aligned}$$

where $0 < 1+a_1 < 1, a_2, a_3, a_4, a_5 \geq 0, (1+a_1) + \sum_{i=1}^{4} a_i = 1$.

Therefore (see Reich [48], Gregus [18], and Ghosh [16] for case (i), Hardy and Rogers [19] for case (ii), and Joshi and Bose [23] and Rhoades [49] for both cases) θ has a unique fixed point, say y_0; hence $\theta(y_0) = y_0$. It now follows from (4.11) that for all $z \in K$,

$$\lfloor Ty_0, z - y_0 \rfloor \geq 0$$

and this completes the proof. ∎

Theorem 4.29. *Let X be a Hilbert space and K a closed convex cone and let the conditions of the previous theorem be satisfied. Then the CP has a unique solution; that is, there is a unique $y_0 \in X$ such that*

$$y_0 \in K, \qquad Ty_0 \in K^* \quad \textit{and} \quad (Ty_0, y_0) = 0.$$

Proof. We have

$$(Ty_0, y_0 - z) \leq 0 \quad \forall z \in K.$$

K is a cone and $y_0 \in K$, thus it follows that $2y_0 \in K$ and hence $(Ty_0, y_0) \geq 0$. Thus $(Ty_0, y_0) = 0$.

Furthermore for all $z \in K$,

$$(Ty_0, z) \geq (Ty_0, y_0) = 0.$$

Thus $Ty_0 \in K^*$ and this completes the proof. ■

Theorem 4.30. *Let X be uc and ss and K a nonempty, closed, convex subset of X. Let $T : K \to K$ be nonexpansive. Then there exists some $y_0 \in K$ such that*

$$\lfloor Ty_0 + y_0, y_0 \rfloor = 0.$$

Proof. Proceeding as in Theorem 4.28 we obtain

$$\|\theta y_1 - \theta y_2\| \leq \|y_1 - Ty_1 - y_2 + Ty_2\|.$$

If T is nonexpansive, then

$$\begin{aligned} \|\theta y_1 - \theta y_2\| &\leq \|Ty_1 - Ty_2\| \\ &\leq \|y_1 - y_2\|. \end{aligned}$$

Hence $\theta/2$ is nonexpansive and thus Theorem 4.2.3, p. 98 of [23], $\theta/2$ has a fixed point, say y_0, that is, $\theta(y_0)2y_0$. Hence it follows that

$$\lfloor Ty_0 + y_0, z - 2y_0 \rfloor \geq, \qquad \forall z \in K.$$

Because $0 \in K$, $\lfloor Ty_0 + y_0, y_0 \rfloor \leq 0$. Because K is a cone and $y_0 \in K$, $3y_0 \in K$ and thus

$$\lfloor Ty_0 + y_0, y_0 \rfloor \geq 0.$$

Therefore $\lfloor Ty_0 + y_0, y_0 \rfloor = 0$ and this completes the proof. ■

References

1. M. S. Bazaras, J. J. Goode, and M. Z. Nashed, A nonlinear complementarity problem in mathematical programming in Banach space, *Proc. Amer. Math. Soc.* 35 (1972), 165–170.
2. F. E. Browder, Nonlinear variational inequalities and maximal monotone mappings in Banach spaces, *Math. Annalen*, 176 (1968), 89–113.
3. ______, Existence and approximation of solutions of nonlinear variational inequalities, *Proc. Natl. Acad. Sci. U. S.* 56 (1966a), 1080–1086.
4. ______, On the unification of calculus of variation and the theory of monotone nonlinear operators in Banach spaces, ibid (1966b), 419–425.
5. ______, Nonlinear monotone operators and convex sets in Banach spaces, *Bull. Amer. Math. Soc.* 71 (1965a) 780–758.
6. ______, On a theorem of Beurling and Livingston, *Canad. J. Math.* 17 (1965b), 367–372.
7. ______, Variational boundary value problems for quasilinear elliptic equations of arbitrary order, *Proc. Natl. Acad. Sci., U. S.* 50 (1963), 31–37.
8. B. D. Craven and B. Mond. Complementarity over arbitrary cone, *Z. O. R.* (1977).

9. R. Chandrasekharan, A special case of the complementarity pivot problem, *Opsearch*, 7 (1970) 263–268.
10. R. W. Cottle and G. B. Dantzig, Complementarity pivot theory of mathematical programming, *Linear Algebra Appl.* 1 (1968), 108–125.
11. O. D. Donato and G. Maier, Mathematical-programming methods for the inelastic analysis of reinforced concrete frames allowing for limited rotation capacity, *Int. J. Numer. Meth. Eng.* 4 (1972) 307–329.
12. A. T. Dash and S. Nanda, A complementarity problem in mathematical programming in Banach space, *J. Math. Anal. Appl.* 98 (1984).
13. B. C. Eaves, On the basic theorem of completemtarity, *Math. Program.* 1 (1971) 68–75.
14. M. Edelstein, On nearest points of sets in uniformly convex Banach spaces, *J. London Math. Soc.* 43 (1968) 375–377.
15. B. Garcia, Some classes of matrices in linear complementarity problem, *Math. Program.* 5 (1973), 299–310.
16. K. M. Ghosh, Fixed-point theorems, *Pure Math Manuscript* 3 (1984) 51–53.
17. J. R. Giles, Classes of semi-inner-product spaces, *Trans. Amer. Math. Soc.* 129 (1967), 436–46.
18. M. Gregus, A fixed-point theorem in Banach space, *Boll. Un. Mat. Ital.* 5(17)—A (1980) 193–198.
19. G. E. Hardy and T. D. Rogers, A generalization of a fixed-point theorem of Reich, *Canad, Math. Bull.* 16 (1973) 201–206.
20. Hartmann and G. Stampacchia, On some nonlinear elliptic differential functional equations, *Acta Math.* 125 (1956) 271–310.
21. G. J. Habetler and A. L. Price, Existence theory for generalized nonlinear complementarity problem, *J. Optim. Theory Appl.* 7 (1971) 223–239.
22. G. J. Habetler and M. M. Kostreva, On a direct algorithm for nonlinear complementarity problems *SIAM J. Control Optim.* 16(3) (1978) 504–511.
23. M. C. Joshi and R. K. Bose, *Some Topics in Nonlinear Functional Analysis*, Wiley Eastern Ltd. (1985).
24. S. Karamaradian, The nonlinear complementarity problem with applications I, II, *J. Optim. Theory Appl.* 4 (1969) 87–98, 167–181.
25. ______, Generalized complementarity problem, ibid 8 (1971) 161–168.
26. J. L. Lions and G. Stampacchia, Variational inequalilities, *Comm. Pure Appl. Math.* 20 (1967) 493–519.
27. C. E. Lemke, Bimatrix equilibrium points and mathematical programming, *Manage. Sci. Ser. A*, 11 (1965) 168–169.
28. ______, Recent Results on Complementarity Problems, in Nonlinear Programming (J. B. Rosen, O. L. Mangasarian and K. Ritter, Eds.), Academic New York (1970).
29. C. E. Lemke and J. T. Howson Jr., Equilibrium points of bimatrix games, *SIAM J. Appl. Math.* 12 (1964), 413–423.
30. G. Lumer, Semi-inner-product spaces, *Trans. Amer. Math. Soc.* 100 (1961) 29–43.
31. B. Mond. On the complex complementarity problem, *Bull. Austral. Math. Soc.* 9 (1973) 249–257.
32. O. G. Mancino and G. G. Stampacchia, Convex programming and variational inequalities, *J. Optim. Theory Appl.* 9(1) (1972), 3–23.
33. U. Mosco, Convergence of convex sets and solutions of variational inequalities, *Adv. Math.* 3 (1966), 510–585.
34. ______, A remark on a theorem of F. E. Browder, *J. Math. Anal. Appl.* 20 (1967a), 90–93.
35. ______, Approximation of the solution of some variational inequalities, *Ann Scuolo Normals sup. Pisa* 21 (1967b), 373–934; 765.
36. K. G. Murty, On a characterization of P-matrices, *SIAM J. Appl. Math.* 20(3) (1971), 378–384.
37. ______, On the number of solutions of the complementarity problem and spanning properties of complementary cones, *Linear Algebra Appl.* 5 (1972), 65–108.
38. ______, Note on a Bard-type scheme for solving the complementarity problem, *Opsearch* 11, 2–3 (1974), 123–130.

39. Sribatsa Nanda and Sudarsan Nanda, A complex nonlinear complementarity problem, *Bull. Austral. Math. Soc.* 19 (1978) 437–444.
40. ______, On stationary points and the complementarity problem, ibid 20 (1979a) 77–86.
41. ______, A nonlinear complementarity problem in mathematical programming in Hilbert space, ibid 20 (1979b) 233–236.
42. ______, A nonlinear complementarity problem in Banach space, ibid 21 (1980) 351–356.
43. Sudarsan Nanda, A note on a theorem on a nonlinear complementarity problem, ibid 27 (19783) 161–163.
44. J. Parida and B. Sahoo, On the complex nonlinear complementarity problem, ibid 14 (1976a), 129–136.
45. ______, Existence theory for the complex nonlinear complementarity problem, ibid 14 (1976b), 417–435.
46. A. Ravindran, A computer routine for quadratic and linear programming problems [H], Commun. *ACM* 15(9) (1972) 818–820.
47. ______, A comparison of the primal-simplex and complementary pivot methods for linear programming, *Naval Res. Logistics Q.* 20(1) (1973) 96–100.
48. S. Reich, some remarks concerning contraction mappings, *Canad. Math. Bull.* 14 (1971) 121–124, MR 45 # 1145.
49. B. E. Rhoades, A comparison of various definitions of contractive mappings, *Trans. Amer. Math. Soc.* 226 (1977) 257–289.
50. H. Scarf, The approximation of fixed points of a continuous map. *SIAM J. Appl. Math.* 15 (1967), 328–343.
51. R. Saigal, On the class of complementary cones and Lemkes' algorithm, *SIAM J. Appl. Math.* 23 (1972) 46–60.
52. ______, A note on a special linear complementarity problem, *Opsearch* 7(3) (1970) 175–183.
53. S. Simons, Variational inequalities via the Hann-Banach theorem, *Arch. Math.* 31 (1978) 482–490.
54. L. T. Watson, Solving nonlinear complementarity problem by homotopy methods, *SIAM J. Control Optim.* 17(1) (1979) 36.

Chapter 5
A Derivative for Semipreinvex Functions and Its Applications in Semipreinvex Programming

Y. X. Zhao, S. Y. Wang, L. Coladas Uria, and S. K. Mishra

Abstract A directional derivative concept is introduced to develop Fritz–John and Kuhn–Tucker conditions for the optimization of general semipreinvex functions. The relationship between the optimization problem and the corresponding semiprevariational inequality problem is also shown.

5.1 Introduction

Because of its importance in optimization theory, the concept of convexity has been generalized in many ways to explore the extent to which results obtained for classical convex functions can be extended to more general classes of functions [1, 5, 7, 6, 4, 3, 9, 2, 8, 10, 12, 14, 11, 13, 15]. In [11] the concept of semipreinvexity was introduced,[1] the local minima of semipreinvex functions were shown to be global minima, and a theorem of the alternative was proved. In the development of a Fritz–John type condition for inequality-constrained minimization, Yang and Chen

Y. X. Zhao
Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, e-mail: zyx@amss.ac.cn

S. Y. Wang
Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, e-mail: sywang@amss.ac.cn

L. Coladas Uria
Department of Statistics and Operations Research, Faculty of Mathematics, University of Santiago de Compostela,15782 Santiago de Compostela, Spain, e-mail: luis.coladas@usc.es

S. K. Mishra
Department of Mathematics, Faculty of Science, Banaras Hindu University, Varanasi 221 005, India, e-mail: bhu.skmishra@gmail.com

[1] There is a typographical error in Yang and Chen's equation (3'), in which $f(y+\alpha\tau(y,x,\alpha)$ should be $f(x+\alpha\tau(y,x,\alpha))$

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_5,

employed a concept of differentiability—arcwise directional differentiability—that cannot usefully be applied to all semipreinvex functions. Here we introduce a concept of differentiability that is in keeping with the spirit of semipreinvexity. Using this concept, we prove the general Fritz–John type and Kuhn–Tucker type optimality conditions. In addition, we show the relationship between the minimization problem for semipreinvex functions and a generalized variational inequality problem.

5.2 Preliminaries

Definition 5.1.

(i) A subset K of $\mathbb{R}^n$ is said to be *semiconnected* with respect to a function $\tau : K \times K \times [0,1] \to \mathbb{R}^n$ (hereinafter, τ-*semiconnected*) iff $\forall x,y \in K$ and $\alpha \in [0,1]$, $x+\alpha\tau(y,x,\alpha) \in X$ and

$$\lim_{\alpha \downarrow 0} \alpha\tau(y,x,\alpha) = 0. \tag{5.1}$$

(ii) A real function f defined on a τ-semiconnected subset K of $\mathbb{R}^n$ is said to be *semipreinvex* with respect to τ (hereinafter, τ-*semipreinvex*) iff $\forall (x,y,\alpha) \in K \times K \times [0,1]$,

$$f(x+\alpha\tau(y,x,\alpha)) \leq (1-\alpha)f(x) + \alpha f(y). \tag{5.2}$$

Remark 5.1. Yang and Chen [11] included condition (5.1) in the definition of semipreinvexity, not in the definition of semiconnectedness. However, because this condition makes no mention of the function f, our practice here seems more rational.

Remark 5.2. Semipreinvex functions constitute a proper subset of convexlike functions and a proper superset of preinvex functions (and of arcwise convex functions). A function f is said to be convex on K if $\forall x,y \in K$, the line segment $L(x,y)$ between x and y belongs to K, and f is convex on $L(x,y)$. Preinvexity requires only that $\forall x,y \in K$, there is some point $x+\tau(y,x) \in K$ such that f is convex on $L(x,x+\tau(y,x)) \subseteq K$. Convexlikeness requires even less, that $\forall x,y \in K$ there is a (not necessarily connected) family Π of points $P(x,y;\alpha) \in K$, $\alpha \in [0,1]$, such that f is convex on Π with weights α; that is, $f(P(x,y;\alpha)) \leq (1-\alpha)f(x) + \alpha f(y)$. Semipreinvexity is like convexlikeness, except that it additionally requires that $P(x,y;\alpha) \to x$ as $\alpha \to 0$ (condition (5.1)).

Example 5.1. Consider $K = \{z = (x(z),y(z)) \in \mathbb{R}^2 \mid (x)^2 + (y)^2 = 1, y \geq 0, y \neq 1\}$. For $z \in K$, let $\theta z = \cos^{-1} x(z)$, and for $\theta \in [0,\pi]\backslash\{\pi/2\}$, let $z(\theta) = (\cos\theta, \sin\theta)$. Consider $f : K \to \mathbb{R}$ such that

$$f(z) = \begin{cases} \theta z, & x(z) > 0; \\ \pi - \theta z, & x(z) < 0, \end{cases}$$

and $\tau : K \times K \times [0,1] \to \mathbb{R}^2$ such that

$$\tau(z2, z1, \alpha) = \begin{cases} \frac{z((1-\alpha)\theta z1+\alpha\theta z2)-z1}{\alpha}, & x(z1), x(z2) < 0, \alpha \neq 0; \\ \frac{z((1-\alpha)\theta z1+\alpha\theta z2)-z1}{\alpha}, & x(z1), x(z2) < 0, \alpha \neq 0; \\ \frac{z((1-\alpha)\theta z1+\alpha(\pi-\theta z2))-z1}{\alpha}, & x(z1) < 0, x(z2) > 0, \alpha \neq 0; \\ \frac{z((1-\alpha)\theta z1+\alpha(\pi-\theta z2))-z1}{\alpha}, & x(z1) > 0, x(z2) < 0, \alpha \neq 0; \\ z1, & \alpha = 0. \end{cases}$$

Then K is τ-semiconnected and f is τ-semipreinvex on K.

Definition 5.2. Let f be a real function on a τ-semiconnected set $K \subseteq \mathbb{R}^n$, $(x,y) \in K \times K$, and $A = \{\alpha_i\} \subset [0,1]$, such that $\lim_{i\to\infty}\alpha_i = 0$. If

$$\xi(f,x,y,A,\tau) := \lim_{\alpha_i \to 0} \alpha_i^{-1}[f(x+\alpha_i\tau(y,x,\alpha_i)) - f(x)]$$

exists, then we call it a *point sequence derivative of f* at x with respect to τ. For given f, x, and y, the set of all such derivatives (the *point sequence derivative set*) is denoted by $M(f,x,y,\tau)$.

Example 5.2. With the premises of Definition 5.2, let $K = [1,3] \cup [-6,-2]$, $f(x) = 1/x$, and

$$\tau(y,x,\alpha) = \begin{cases} y - x, & x,y \in [1,3], \alpha \in [0,1] \text{ rational}; \\ 3 - x, & x,y \in [1,3], \alpha \in [0,1] \text{ irrational}; \\ y - x, & x,y \in [-6,-2]; \\ -2 - x, & x \in [-6,-2], y \in [1,3]; \\ 1 - x, & x \in [1,3], y \in [-6,-2]. \end{cases}$$

Then, it is easy to show that K is τ-semiconnected and $M(f,2,1,\tau) = \{-\frac{1}{4}, \frac{1}{4}\}$.

Semipreinvex functions are convexlike, therefore the following theorem of the alternative holds (see [6] for a proof for general convexlike functions).

Lemma 5.1. *Let $f_1, f_2, \ldots, f_p : K \to \mathbb{R}$ be τ-semipreinvex on K. Then either (i) or (ii) holds, but not both:*

(i) $\exists x \in K$ *such that,* $\forall i \in \{1,\ldots,p\}$, $f_i(x) < 0$.
(ii) $\exists \lambda \in \mathbb{R}^p_+ \setminus \{\not{\vdash}\}$ *such that,* $\forall x \in K$, $\lambda^T(f_1(x),\ldots,f_p(x)) \geq 0$.

Lemma 5.2. *If f is τ-semipreinvex on $K \subseteq \mathbb{R}^n$ and $(x,y) \in K \times K$, then* $\forall \xi(f,x,y,A,\tau) \in M(f,x,y,\tau)$, $f(y) - f(x) \geq \xi(f,x,y,A,\tau)$.

Proof. Let $\xi(f,x,y,A,\tau) \in M(f,x,y,\tau)$, where $A = \{\alpha_i\} \subset [0,1]$. Because f is τ-semipreinvex, $f(x+\alpha_i\tau(y,x,\alpha_i)) \leq (1-\alpha_i)f(x)+\alpha_i f(y)$; that is,

$$f(y)-f(x) \geq \frac{f(x+\alpha_i\tau(y,x,\alpha_i))-f(x)}{\alpha_i}.$$

This lemma follows from Definition 5.2 upon taking the limit $\alpha_i \to 0$. □

5.3 Fritz–John and Kuhn–Tucker Results

We consider the problem

$$\text{(P)} \qquad \min_{x\in S} f(x),$$

where $K \subseteq \mathbb{R}^n$ is τ-semiconnected, $f : K \to \mathbb{R}$ is τ-semipreinvex on K, and $S = \{x \in K \mid g_i(x) \leq 0\ \forall i \in \{1,\ldots,m\}\}$, where $g_1,\ldots,g_m : K \to \mathbb{R}$ are all τ-semipreinvex on K. In relation to this problem, we define the following conditions that a point $\bar{x} \in S$ may satisfy.

Condition R: $\forall x \in S$, there exist (i) $A\bar{x}x = \{\alpha_i\} \subset [0,1]$ such that $\lim_{i\to\infty}\alpha_i = 0$ and (ii) $\xi(f,\bar{x},x,A\bar{x}x,\tau) \in M(f,\bar{x},x,\tau)$, $\xi(g_1,\bar{x},x,A\bar{x}x,\tau) \in M(g_1,\bar{x},x,\tau),\ldots,$ and $\xi(g_m,\bar{x},x,A\bar{x}x,\tau) \in M(g_m,\bar{x},x,\tau)$.

Condition C: $\bar{x}$ satisfies Condition R and there exist real numbers $\lambda(\bar{x}),\mu_1(\bar{x}),\ldots,\mu_m(\bar{x}) \geq 0$, not all zero, such that $\forall x \in S$,

$$\lambda(\bar{x})\xi(f,\bar{x},x,A\bar{x}x,\tau)+\sum_{i=1}^{m}\mu_i(\bar{x})\xi(g_i,\bar{x},x,A\bar{x}x,\tau) \geq 0,$$

$$\sum_{i=1}^{m}\mu_i(\bar{x})g_i(\bar{x}) = 0.$$

Theorem 5.1 (Fritz–John conditions). *With K, f, and g_i as above, let $\bar{x}$ be a solution of problem (P) that satisfies Condition R. Then $\bar{x}$ satisfies Condition C.*

Proof. Because $\bar{x}$ is a solution for problem (P), there is no $x \in K$ such that

$$f(x)-f(\bar{x}) < 0 \quad \text{and} \quad g_i(x) < 0\ \forall i \in \{1,\ldots,m\}.$$

By Lemma 5.1, there exist real numbers $\lambda,\mu_1,\ldots,\mu_m \geq 0$, not all zero, such that, $\forall x \in K$,

$$\lambda(f(x)-f(\bar{x}))+\sum_{i=1}^{m}\mu_i g_i(x) \geq 0. \tag{5.3}$$

Putting $x = \bar{x}$, it follows that $\sum_{i=1}^{m}\mu_i g_i(\bar{x}) \geq 0$; but also, $\sum_{i=1}^{m}\mu_i g_i(\bar{x}) \leq 0$ because $\mu_1,\mu_2,\ldots,\mu_m \geq 0$ and $g_1(\bar{x}),g_2(\bar{x}),\ldots,g_m(\bar{x}) \leq 0$. Hence, $\sum_{i=1}^{m}\mu_i g_i(\bar{x}) = 0$. Because

K is τ-semiconnected, $\bar{x}+\alpha\tau(\bar{x},x,\alpha) \in K\ \forall(x,\alpha) \in S\times[0,1]$. Hence, by inequality (5.3),

$$\lambda(f(\bar{x}+\alpha\tau(\bar{x},x,\alpha)) - f(\bar{x})) + \sum_{i=1}^{m} \mu_i g_i(\bar{x}+\alpha\tau(\bar{x},x,\alpha)) \geq 0.$$

Therefore, $\forall\alpha \in (0,1]$,

$$\lambda\frac{f(\bar{x}+\alpha\tau(\bar{x},x,\alpha)) - f(\bar{x})}{\alpha} + \sum_{i=1}^{m} \mu_i \frac{g_i(\bar{x}+\alpha\tau(\bar{x},x,\alpha)) - g_i(\bar{x})}{\alpha} \geq 0.$$

If $A\bar{x}x$, $\xi(f,\bar{x},x,A\bar{x}x,\tau)$, $\xi(g_1,\bar{x},x,A\bar{x}x,\tau),\ldots,\xi(g_m,\bar{x},x,A\bar{x}x,\tau)$ are as guaranteed by condition R, then letting $\alpha \to 0^+$ on the set $A\bar{x}x$ leads to the desired result that $\bar{x}$ satisfies condition C with $\lambda(\bar{x}) = \lambda$, $\mu_1(\bar{x}) = \mu_1,\mu_2,\ldots,\mu_m(\bar{x}) = \mu_m$. □

Theorem 5.2 (Kuhn–Tucker necessary conditions). *With K, f, and g_i as above, let $\bar{x}$ be a solution of problem (P) that satisfies condition R. If $\exists x' \in S$ such that $g_i(x') < 0\ \forall i \in \{1,\ldots,m\}$, then $\bar{x}$ satisfies condition C with $\lambda(\bar{x}) > 0$.*

Proof. By Theorem 5.1, there exist real numbers $\lambda,\mu_1,\ldots,\mu_m \geq 0$, not all zero, such that $\forall x \in S$,

$$\lambda\xi(f,\bar{x},x,A\bar{x}x,\tau) + \sum_{i=1}^{m} \mu_i\xi(g_i,\bar{x},x,A\bar{x}x,\tau) \geq 0, \tag{5.4}$$

$$\sum_{i=1}^{m} \mu_i g_i(\bar{x}) = 0. \tag{5.5}$$

Suppose $\lambda = 0$. Then by (5.4), $\sum_{i=1}^{m} \mu_i\xi(g_i,\bar{x},x',A\bar{x}x',\tau) \geq 0$, where not all μ_i are zero. Because $g_i(x') - g_i(\bar{x}) \geq \xi(g_i,\bar{x},x',A\bar{x}x',\tau)$ (by Lemma 5.2), it follows that $\sum_{i=1}^{m} \mu_i(g_i(x') - g_i(\bar{x})) \geq 0$, which, together with (5.5), shows that

$$\sum_{i=1}^{m} \mu_i g_i(x') \geq 0.$$

This obviously contradicts the premise that $g_i(x') < 0\ \forall i \in \{1,\ldots,m\}$. Hence, $\lambda > 0$, which is the desired result. □

Theorem 5.3 (Kuhn–Tucker sufficient conditions). *With K, f, and g_i as above, if $\bar{x} \in S$ satisfies condition C with $\lambda(\bar{x}) > 0$, then $\bar{x}$ is a solution of problem (P).*

Proof. By Lemma 5.2, $\forall x \in S$, $f(x) - f(\bar{x}) \geq \xi(f,\bar{x},x,A\bar{x}x,\tau)$, and $g_i(x) - g_i(\bar{x}) \geq \xi(g_i,\bar{x},x,A\bar{x}x,\tau)\ \forall i \in \{1,\ldots,m\}$, which, together with condition C, lead to

$$\lambda(\bar{x})(f(x) - f(\bar{x})) + \sum_{i=1}^{m} \mu_i(\bar{x})g_i(x) \geq 0. \tag{5.6}$$

$\sum_{i=1}^{m} \mu_{\bar{x}i}g_i(x) \leq 0$ for all $x \in S$ (because $\mu_{\bar{x}1},\ldots,\mu_{\bar{x}m} \geq 0$) and $\lambda(\bar{x}) > 0$, therefore it follows from (5.6) that $f(x) - f(\bar{x}) \geq 0\ \forall x \in S$, which is the desired result. □

Remark 5.3. Theorem 5.3 can be strengthened by weakening condition C because it is not necessary for $A\bar{x}x$ to be the same for all the point sequence derivatives $\xi(f,\bar{x},x,A\bar{x}x,\tau)$, $\xi(g_1,\bar{x},x,A\bar{x}x,\tau),\ldots,\xi(g_m,\bar{x},x,A\bar{x}x,\tau)$, and neither is it necessary for the coefficients $\lambda(\bar{x}),\mu_1(\bar{x}),\ldots,\mu_m(\bar{x})$ to be the same for all $x \in S$.

Remark 5.4. If the sequence A in Definition 5.2 is replaced by the interval $A' = (0,\alpha']$ for some $\alpha' \in (0,1]$, and τ is such that $\sigma : [0,\alpha'] \to K,\ \alpha \mapsto x + \alpha\tau(y,x,\alpha)$, is a differentiable curve starting at x, then $\xi(f,x,y,A',\tau)$ becomes an arcwise directional derivative, Theorem 5.1 becomes Theorem 3 of [11], and Theorem 5.2 becomes the corresponding Kuhn–Tucker type theorem. There may also be situations in which $\lim_{\alpha\to 0}\alpha^{-1}[f(x+\alpha\tau(y,x,\alpha)) - f(x)]$ exists even though σ is not a differentiable curve.

Remark 5.5. Consider the multiobjective problem

$$\text{(MP)} \qquad \min_{x\in S} f(x),$$

and

$$(\text{TP}_\lambda) \qquad \min_{x\in S} \lambda^T f(x),$$

where $\lambda \in \mathbb{R}^p_+\backslash\{0\}$, K and S are as in problem (P), and all the components of $f = (f_1,f_2,\ldots,f_p) : K \to \mathbb{R}^p$ are τ-semipreinvex on K. By a result obtained in [16], any solution for TP_λ is a weakly proper solution for MP. Using this result, Theorems 5.1–5.3 can easily be generalized to the corresponding conditions on weakly proper solutions for MP. For brevity we omit the details here.

5.4 The Semiprevariational Inequality Problem for Point Sequence Derivative

Consider the unconstrained versions of problems (P) and MP where $S = K$ and denote them by UP and UMP, respectively. The corresponding scalar and vector semiprevariational inequality problems for point sequence derivatives are defined as follows.

(SPVI) To find $\bar{x} \in K$ such that $\forall x \in K$, $\nexists\xi(f,\bar{x},x,A\bar{x}x,\tau) \in M(f,\bar{x},x,\tau)$ such that $\xi(f,\bar{x},x,A\bar{x}x,\tau) < 0$.

(VSPVI) Assume that $\forall(i,x,y) \in \{1,\ldots p\} \times K \times K, M(f_i,x,y,\tau) \neq \emptyset$. To find $\bar{x} \in K$ such that $\forall x \in K$; there exists at least one $\xi(f_i,\bar{x},x,A\bar{x}x,\tau) \geq 0$.

Theorem 5.4. *If $f : K \to \mathbb{R}$ is τ-semipreinvex and continuous on K, then*

(i) If $\bar{x}$ is a solution for problem UP, then $\bar{x}$ is also a solution for problem SPVI.
(ii) If $\exists\, \bar{x} \in K$ such that $\forall x \in K$, $\exists\xi(f,\bar{x},x,A\bar{x}x,\tau) \geq 0$, then $\bar{x}$ is a solution of problems UP and SPVI, respectively.

Proof.

(i) If $M(f,\bar{x},x,\tau)$ is empty for all $x \in K$, there is nothing needing proof. Otherwise, let x be such that $M(f,\bar{x},x,\tau)$ is nonempty, and consider $\xi(f,\bar{x},x,A\bar{x}x,\tau) \in M(f,\bar{x},x,\tau)$. Because $\bar{x}$ is a solution for problem UP, $\forall \alpha \in [0,1]$,

$$f(\bar{x}+\alpha\tau(x,\bar{x},\alpha)) \geq f(\bar{x}).$$

Therefore, $\xi(f,\bar{x},x,A\bar{x}x,\tau) \geq 0$ by Definition 5.2, which leads to part (i).

(ii) The result follows immediately from Lemma 5.2 and part (i). □

Remark 5.6. Theorem 5.4 shows that if f is τ-semipreinvex and $\forall x \in K, M(f,\bar{x},x,\tau)$ contains at least one nonnegative member, then all its members are nonnegative.

Remark 5.7. In [17], semiprevariational inequality problems are defined, assuming stronger differentiability conditions, for functions defined on, and mapping to, more general ordered spaces.

Theorem 5.5. *If all the components of $f = (f_1, f_2, \ldots, f_p) : K \to \mathbb{R}^p$ are τ-semipreinvex on K, then*

(i) If $\bar{x}$ is a weakly proper solution of problem UMP such that $\forall (i,x) \in \{1,\ldots,p\} \times K$, $M(f_i,\bar{x},x,\tau)$ is nonempty, then $\bar{x}$ solves problem VSPVI.

(ii) If $\bar{x}$ is a solution of problem VSPVI, then it is a weakly proper solution of problem UMP.

Proof.

(i) By the assumptions, it is clear that $\forall x \in K$, (a) $f(x) - f(\bar{x}) \notin$-intR^p_+ and (b) $\forall i \in \{1,\ldots,p\}, \exists \xi(f_i,\bar{x},x,A\bar{x}x,\tau) \in M(f_i,\bar{x},x,\tau)$. By (a), $\forall (x,\alpha) \in K \times (0,1]$,

$$\frac{f(\bar{x}+\alpha\tau(x,\bar{x},\alpha)) - f(\bar{x})}{\alpha} \notin -\text{int} R^p_+.$$

Letting $\alpha \to 0^+$ on the set $A\bar{x}x$ leads to part (i).

(ii) By the definition of problem VSPVI, for arbitrarily given $x \in K$, there exist $i \in \{1,\ldots,p\}$ and $\xi(f_i,\bar{x},x,A\bar{x}x,\tau) \in M(f_i,\bar{x},x,\tau)$ such that $\xi(f_i,\bar{x},x,A\bar{x}x,\tau) \geq 0$. The desired result follows immediately from Lemma 5.2. □

References

1. Craven, B.D., Glover, B.M.: Invex functions and duality. *Austral. Math. Soc.* **39**, 1–20 (1985)
2. Ng, C.T., Nikodem, K.: On approximately convex functions. *J. Math. Anal. Appl.* **118**, 103–108 (1993)
3. Martin, D.H.: The essence of invexity. *J. Optim. Theory Appl.* **47(1)**, 65–76 (1985)
4. Luo, H.Z., Xu, Z.K.: Note on characterizations of prequasi-invex functions. *J. Optim. Theory Appl.* **120(2)**, 429–439 (2004)

5. Fu, J.Y., Wang, Y.H.: Arcwise connected cone-convex functions and mathematical programming. *J. Optim. Theory Appl.* **118(2)**, 339–352 (2003)
6. Hayashi, M., Komiya, H.: Perfect duality for convexlike programs. *J. Optim. Theory Appl.* **38**, 179–189 (1982)
7. Hanson, M.A.: On ratio invexity in mathematical programming. *J. Math. Anal. Appl.* **205**, 330–336 (1997)
8. Rueda, N.G., Hanson, M.A.: Optimality criteria in mathematical programming involving generalized invexity. *J. Math. Anal. Appl.* **130**, 375–385 (1988)
9. Mohan, S.R., Neogy, S.K.: Note on invex sets and preinvex functions. *J. Math. Anal. Appl.* **189**, 901–908 (1995)
10. Weir, T.: Pre-invex functions in multiple objective optimization. *J. Math. Anal. Appl.* **136**, 29–38 (1988)
11. Yang, X.Q., Chen, G.Y.: A class of nonconvex functions and pre-variational inequalities. *J. Math. Anal. Appl.* **169**, 359–373 (1992)
12. Xu, Z.K.: Necessary conditions for suboptimization over the weakly efficient set associated to generalized invex multiobjective programming. *J. Math. Anal. Appl.* **201**, 502–515 (1996)
13. Yang, X.M., Yang, X.Q., Teo, K.L.: Characterizations and applications of prequasi-invex functions. *J. Optim. Theory Appl.* **110(3)**, 645–668 (2001)
14. Xu, Z.K.: On invexity-type nonlinear programming problems. *J. Optim. Theory Appl.* **80(1)**, 135–148 (1994)
15. Pini, R., Singh, C.: A survey of recent 1985–1995 advances in generalized convexity with applications to duality theory and optimality conditions. *Optimization* **39**, 311–360 (1997)
16. Lin, C.Y., Dong, J.L.: *The Methods and Theory for Multi-objective Optimization* (in Chinese). pp. 62–63. Jilin Education Press, Changchun, China (1992)
17. Lai, H.C.: Optimality conditions for semi-preinvex programming. *Taiwanese J. Math.* **1(4)**, 389–404 (1997)
18. Yamashita, N., Taji, K., Fukushima, M.: Unconstrained optimization reformulations of variational inequality problems. *J. Optim. Theory Appl.* **92(3)**, 439–456 (1997)
19. Chang, S.S., Zhang, C.J.: On a class of generalized variational inequalities and quasi-variational inequalities. *J. Math. Anal. Appl.* **179**, 250–259 (1993)

Chapter 6
Proximal Proper Saddle Points in Set-Valued Optimization

C. S. Lalitha and R. Arora

Abstract A new notion of saddle point, namely proximal proper saddle point is introduced in terms of a Lagrangian map associated with a set-valued optimization problem for which an existence criterion is obtained. It is also observed that this saddle point is not related to some of the well-known saddle points. A saddle point type optimality criterion is derived for optimality in terms of proximal proper minimizers.

6.1 Introduction

The study of set-valued optimization has received a great deal of attention in the recent past due to its application in various fields such as economics, game theory, and differential inclusions (see Aubin and Frankowska [2], Klein and Thompson [8]). Set-valued maps are involved in problems where the existence or uniqueness of a solution is not guaranteed. These maps are also used at various instances in nonsmooth analysis, for example, tangents, cones, subgradients, and inverses of functions are all set-valued maps. The concept of set-valued maps is also used by Zangwill [14] to discuss the convergence of algorithms to determine optimal solutions of nonlinear programming problems.

The optimal solution for a set-valued optimization problem is often considered in terms of efficiency. To avoid some of the undesirable efficient points, attempts have been made by various researchers to refine the notion of efficiency. The concept

C. S. Lalitha
Department of Mathematics, University of Delhi South Campus, Benito Jaurez Road,
New Delhi 110 021, India, e-mail: cslalitha@rediffmail.com

R. Arora
Department of Mathematics, University of Delhi, Delhi 110 007, India
e-mail: ruchiiq@yahoo.com

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_6,

of proper efficiency was introduced in different forms by Kuhn and Tucker [9], Geoffrion [6], Borwein [4], and Benson [3] for vector-valued optimization problems.

Later, this study was also extended for set-valued optimization problems (see Li [12] and Rong and Wu [13]). Lagrangian saddle points and saddle point optimality criteria play a crucial role in the study of optimization problems. Li and Chen [12] characterized weak efficiency whereas Li [12] characterized Benson proper efficiency in terms of saddle points of Lagrangian maps associated with a constrained set-valued optimization problem.

The notion of proximal proper efficiency has been introduced by Lalitha and Arora [10] where it has been observed that proximal proper efficiency refines Borwein proper efficiency and is independent of Benson proper efficiency. Extending this comparative study, it is seen that proximal proper efficiency is independent of the notion of superefficiency as well. Motivated by this independent nature of proximal proper efficiency, a new notion of proximal proper saddle point is introduced in this chapter. Finally, a saddle point type optimality criterion is obtained for a constrained set-valued optimization problem in terms of proximal proper efficiency.

The chapter is organized into five sections. Section 6.2 presents some definitions and results used in the chapter. Section 6.3 deals with a comparative study of the notion of proximal proper efficiency with some of the well-known notions of proper efficiency. In Section 6.4, necessary and sufficient conditions for the existence of a proximal proper saddle point are established in terms of the Lagrangian map. Section 6.5 comprises the characterization of proximal proper efficiency in terms of the proximal proper saddle point.

6.2 Preliminaries

A set C in R^n is said to be a cone if $\lambda c \in C$ for any $c \in C$, $\lambda \geq 0$, $\lambda \in R$. A cone C is said to be convex if $C + C \subseteq C$ and pointed if $C \cap (-C) = \{0\}$. The dual cone of C, denoted by C^*, is defined as

$$C^* := \{d \in R^n : \langle d, c \rangle \geq 0, \; \forall \, c \in C\}$$

and strict dual cone C^{*0} of C is defined as

$$C^{*0} := \{d \in R^n : \langle d, c \rangle > 0, \; \forall \, c \in C \setminus \{0\}\}.$$

A set B in R^n is said to be locally star-shaped at $\bar{x} \in B$ if for any $x \in B$, there exists a real number $a(x,\bar{x})$, $0 < a(x,\bar{x}) \leq 1$ such that

$$(1-\lambda)\bar{x} + \lambda x \in B \quad \text{for} \;\; 0 < \lambda \leq a(x,\bar{x}).$$

B is said to be locally star-shaped if it is locally star-shaped at each of its points. From the definition, it is obvious that if $a(x,\bar{x}) = 1$ for every $x, \bar{x} \in B$, then the set B

is a convex set. Thus every convex set is locally star-shaped but the converse may not necessarily be true. The set $B = \{(x,y) \in R^2 : x^2 \leq |y|\} \cup \{(x,y) \in R^2 : y = 0\}$ is locally star-shaped at the origin but is not convex in any neighborhood containing the origin.

A set $B \subseteq R^n$ is said to be a C-closed set if $B+C$ is a closed set. The contingent cone or the Bouligand tangent cone to B at $\bar{y} \in B$ is defined as

$$T(B,\bar{y}) := \{d \in R^n : \exists\, t_j \downarrow 0, d_j \to d \text{ such that } \bar{y} + t_j d_j \in B\}.$$

For a closed set B in R^n and $x \notin B$, let $\bar{y}$ be the projection of x onto B; that is,

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

The vector $x-\bar{y}$ is a proximal normal direction to B at $\bar{y}$ and any nonnegative multiple of such a vector is a proximal normal to B at $\bar{y}$. The set of all proximal normals to B at $\bar{y}$ forms the proximal normal cone to B at $\bar{y}$ and is denoted by $N_P(B,\bar{y})$. For more details, refer to Clarke et al. [5]. The following theorem characterizes a proximal normal vector in the form of an inequality called proximal normal inequality.

Lemma 6.1 ([5, Proposition 1.5]). *A vector ξ belongs to $N_P(B,\bar{y})$ if and only if there exists $\sigma = \sigma(\xi,\bar{y}) \geq 0$ such that*

$$\langle \xi, y-\bar{y} \rangle \leq \sigma \|y-\bar{y}\|^2 \qquad \forall\, y \in B.$$

If B is a convex set then the above inequality becomes

$$\langle \xi, y-\bar{y} \rangle \leq 0 \qquad \forall\, y \in B.$$

A point $\bar{y} \in B$ is said to be a minimizer of B if $(B-\bar{y}) \cap (-C) = \{0\}$ whereas it is a maximizer of B if $(B-\bar{y}) \cap C = \{0\}$. The set of minimizers is denoted by $\mathrm{Min}[B,C]$ and the set of all maximizers is denoted by $\mathrm{Max}[B,C]$. A minimizer $\bar{y}$ of a closed set B is called a proximal proper minimizer of B if $N_P(B+C,\bar{y}) \cap (-C^{*0}) \neq \phi$, where C is a closed convex pointed cone in R^n and B is a C-closed set. Similarly, a maximizer $\bar{y}$ of B is a proximal proper maximizer of B if $N_P(B-C,\bar{y}) \cap C^{*0} \neq \phi$ where B is a $-C$-closed set. We denote the set of proximal proper minimizers by $\mathrm{Pr}[B,C]$ and the set of proximal proper maximizers by $\mathrm{Pr}^0[B,C]$.

6.3 Proximal Proper Efficiency and Its Relation with Other Notions of Proper Efficiency

Several notions of proper efficiency have been studied in the literature. In this section proximal proper efficiency is compared with some well-known notions of proper efficiency. We first recall the notions of Borwein [4] and Benson [3] proper efficiency.

A point $\bar{y} \in B$ is said to be a Borwein proper minimizer of B if $T(B+C-\bar{y}) \cap (-C) = \{0\}$. The set of Borwein proper minimizers is denoted by $\mathrm{Bor}[B,C]$. The relationship between the proximal proper minimizer and Borwein proper minimzer has been established by Lalitha and Arora [10] as

$$\mathrm{Pr}[B,C] \subseteq \mathrm{Bor}[B,C].$$

A point $\bar{y} \in B$ is said to be a Benson proper minimizer of B if cl cone $(B+C-\bar{y}) \cap (-C) = \{0\}$. The set of Benson proper minimizers is denoted by $\mathrm{Ben}[B,C]$.

Furthermore, we compare proximal proper efficiency with another notion of proper efficiency, namely superefficiency [1].

A point $\bar{y} \in B$ is said to be a superefficient point of B if there exists $M > 0$ such that cl cone $(B-\bar{y}) \cap (U-C) \subseteq MU$ where U is a closed unit ball. The set of all superefficient points of B is denoted by $\mathrm{SE}[B,C]$.

In view of the relationships established among the above-mentioned notions of proper efficiency in [7], the following conclusion can be made

$$\mathrm{SE}[B,C] \subseteq \mathrm{Ben}[B,C] \subseteq \mathrm{Bor}[B,C]. \tag{6.1}$$

The following two examples illustrate the relation of proximal proper efficiency with other notions of proper efficiency, stated above. In the first example, a proximal proper minimizer is shown to be neither Benson proper efficient nor superefficient.

Example 6.1. Let $B = \{(x,0,0) : x \leq 0\} \cup \{(x,0,1-x) : x \geq 1\}$, $C = \{(x,y,z) : x^2 + y^2 \leq z^2, z \geq 0\}$. It can be seen that $\bar{y} = (0,0,0) \in \mathrm{Pr}[B,C]$ as $N_P(B+C,\bar{y}) = \{(x,y,z) : x^2+y^2 \leq z^2, x \geq 0, z \leq 0\}$ whose intersection with $-(C^{*0})$ is nonempty. However, it can be seen that there does not exist any $M > 0$ such that the intersection of the cl cone $(B-\bar{y}) = \{(x,0,z) : -x \leq z \leq 0, x \geq 0\} \cup \{(x,0,0) : x \leq 0\}$ with $(U-C)$ is contained in MU where U is the closed unit ball. Hence $\bar{y} \notin \mathrm{SE}[B,C]$. Also, it can be seen that $\bar{y} \notin \mathrm{Ben}[B,C]$ as the cl cone $(B+C-\bar{y}) \cap (-C) = \{(x,0,-x) : x \geq 0\}$. However, it may be observed here that $\bar{y} \in \mathrm{Bor}[B,C]$.

The following example provides a point that is a superefficient point and hence a Benson and Borwein proper minimizer that is not a proximal proper minimizer.

Example 6.2. Let $C = R^3_+$ and $B = \{(x,y,z) : x^2+y^2 \leq 4z^2, z \geq 0\} \cup \{(x,y,z) : y^2 \leq x+z, z \leq 0\}$. Because $N_P(B+C,\bar{y}) = \{0\}$ for $\bar{y} = (0,0,0)$, it follows that $\bar{y} \notin \mathrm{Pr}[B,C]$. Because cl cone $(B-\bar{y}) = B$, it may be noted that for a closed unit ball U, cl cone $(B-\bar{y}) \cap (U-C) \subseteq MU$ for any $M > 1$; that is, $\bar{y} \in \mathrm{SE}[B,C]$. Also by (6.1), $\bar{y}$ belongs to each of the sets, $\mathrm{Ben}[B,C]$ and $\mathrm{Bor}[B,C]$.

The comparison made so far can be summarized in Figure 6.1.

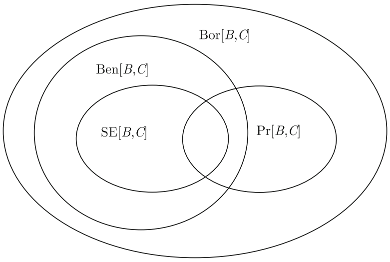

Fig. 6.1 Relations among certain notions of proper efficiency.

6.4 Proximal Proper Saddle Points

In the sequel, we assume that $X = R^m$, $Y = R^n$, $Z = R^p$; $C \subseteq Y$, and $D \subseteq Z$ are closed convex pointed cones with nonempty interiors and $F : X \to 2^Y$ and $G : X \to 2^Z$ are set-valued maps with nonempty values.

We consider the following set-valued optimization problem.

$$\text{(VP)} \quad \text{Min } F(x)$$
$$\text{subject to } \; G(x) \cap (-D) \neq \phi.$$

The set of feasible solutions of (VP) is

$$S = \{x \in X : G(x) \cap (-D) \neq \phi\}$$

and the image set of S under F is given by $F(S) = \bigcup_{x \in S} F(x)$.

Let $\mathscr{L}(Z,Y)$ be the space of continuous linear operators from Z to Y and $\mathscr{L}_+(Z,Y)$ be defined as

$$\mathscr{L}_+(Z,Y) = \{T \in \mathscr{L}(Z,Y) : T(D) \subseteq C\}.$$

The Lagrangian map associated with the set-valued problem (VP), is the set-valued map $L : X \times \mathscr{L}_+(Z,Y) \to 2^Y$ given as

$$L(x,T) = F(x) + T(G(x)).$$

Based on different notions of proper efficiency, the following notions of proper saddle points have been defined by various authors.

Definition 6.1. A pair $(\bar{x},\bar{T}) \in X \times \mathscr{L}_+(Z,Y)$ is said to be

(i) A *proper saddle point* of L in terms of Benson proper efficiency [11] if

$$L(\bar{x},\bar{T}) \cap \mathrm{Ben}\left[\bigcup_{x\in X} L(x,\bar{T}),C\right] \cap \mathrm{Ben}^0\left[\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(\bar{x},T),C\right] \neq \phi,$$

where $\bar{y} \in \mathrm{Ben}^0[B,C]$ if $-\bar{y} \in \mathrm{Ben}[-B,C]$ for any set $B \subseteq R^p$

(ii) A *super saddle point* [1] of L if

$$L(\bar{x},\bar{T}) \cap \mathrm{SE}\left[\bigcup_{x\in X} L(x,\bar{T}),C\right] \cap \mathrm{SE}^0\left[\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(\bar{x},T),C\right] \neq \phi,$$

where $\bar{y} \in \mathrm{SE}^0[B,C]$ if $-\bar{y} \in \mathrm{SE}[-B,C]$ for any set $B \subseteq R^p$

We now introduce the notion of proximal proper saddle point for the Lagrangian map $L(x,T)$. We assume throughout that for each fixed $T \in \mathscr{L}_+(Z,Y)$, the set $\bigcup_{x\in X} L(x,T)$ is a C-closed set in Y and for each fixed $x \in X$, the set $\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(x,T)$ is $-C$-closed in Y.

Definition 6.2. A pair $(\bar{x},\bar{T}) \in X \times \mathscr{L}_+(Z,Y)$ is said to be a *proximal proper saddle point* of L if

$$L(\bar{x},\bar{T}) \cap \mathrm{Pr}\left[\bigcup_{x\in X} L(x,\bar{T}),C\right] \cap \mathrm{Pr}^0\left[\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(\bar{x},T),C\right] \neq \phi.$$

The relations analogous to those presented in Figure 6.1 also hold for the corresponding proper saddle points. The following example illustrates this fact.

Example 6.3. Let $X = R$, $Y = R^2$, $Z = R$, $C = R^2_+$, and $D = R_+$. Define $F : X \to 2^Y$ by

$$F(x) = \begin{cases} \{(2x,1)\} & \text{if } x < 0 \\ \{(0,0)\} & \text{if } x = 0 \\ \{(0,t) : 0 \leq t \leq x\} & \text{if } x > 0, \end{cases}$$

and $G : X \to 2^Z$ by $G(x) = \{x\}$. The Lagrangian $L : X \times \mathscr{L}_+(Z,Y) \to 2^Y$ is given by $L(x,T) = F(x) + T(G(x))$ where $T \in \mathscr{L}_+(Z,Y)$ is of the form $T(x) = (\alpha x, \beta x)$. As $T \in \mathscr{L}_+(Z,Y)$ therefore $\alpha \geq 0, \beta \geq 0$. It can be seen that $(\bar{x},\bar{T})$ is a proximal proper saddle point for $\bar{x} = 0, \bar{T}(x) = (x,0)$. However, it can be seen that $(\bar{x},\bar{T})$ is not a proper saddle point in the sense of Benson proper efficiency as

$$(0,0) \notin \mathrm{Ben}\left[\bigcup_{x\in X} L(x,\bar{T}),C\right].$$

It is also evident that $(\bar{x},\bar{T})$ is not a super saddle point as

$$(0,0) \notin \mathrm{SE}\left[\bigcup_{x\in X} L(x,\bar{T}),C\right].$$

In order to obtain characterizations for proximal proper saddle points, we require the following results.

Lemma 6.2. *For* $(\bar{x},\bar{T}) \in X \times \mathscr{L}_+(Z,Y)$ *if*

$$\bar{y}+\bar{T}(\bar{z}) \in \mathrm{Max}\left[\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(\bar{x},T),C\right] \tag{6.2}$$

for some $\bar{y} \in F(\bar{x}), \bar{z} \in G(\bar{x})$ *then*

$$\bar{y}+\bar{T}(\bar{z}) \in \mathrm{Max}\left[\bigcup_{T\in\mathscr{L}_+(Z,Y)} S(T),C\right]$$

where $S(T) = \bar{y}+T(\bar{z})$.

Proof. As $S(T)-\bar{y}-\bar{T}(\bar{z}) \in F(\bar{x})+T(G(\bar{x}))-\bar{y}-\bar{T}(\bar{z}) = L(\bar{x},T)-\bar{y}-\bar{T}(\bar{z})$ therefore on using (6.2), we have

$$\left(\bigcup_{T\in\mathscr{L}_+(Z,Y)} S(T)-\bar{y}-\bar{T}(\bar{z})\right)\cap C = \{0\}$$

and hence the result.

The next two propositions are used to obtain necessary conditions for the existence of proximal proper saddle points.

Proposition 6.1. *For* $(\bar{x},\bar{T}) \in X \times \mathscr{L}_+(Z,Y)$ *if*

$$\bar{y}+\bar{T}(\bar{z}) \in \mathrm{Pr}^0\left[\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(\bar{x},T),C\right] \tag{6.3}$$

for some $\bar{y} \in F(\bar{x}), \bar{z} \in G(\bar{x})$ *then* $\bar{z} \in -D, \bar{T}(\bar{z}) = 0$ *and* $G(\bar{x}) \subseteq -D$.

Proof. A proximal proper maximizer is a maximizer, therefore it follows from Lemma 6.2 that

$$\left(\bigcup_{T\in\mathscr{L}_+(Z,Y)} T(\bar{z})-\bar{T}(\bar{z})\right)\cap C = \{0\}.$$

Using the separation theorem, it can be established that $\bar{z} \in -D$ and $G(\bar{x}) \subseteq -D$ which follows on the lines of Proposition 6.1 of Li [11].

We now claim that $\bar{T}(\bar{z}) = 0$. If this were not true, we would have

$$\langle \psi, -\bar{T}(\bar{z}) \rangle > 0 \; \forall \; \psi \in C^{*0}, \tag{6.4}$$

where $-\bar{T}(\bar{z}) = \bar{T}(-\bar{z}) \in C$ as $\bar{T} \in \mathscr{L}_+(Z,Y)$. From (6.3), it follows that there exists ξ in C^{*0} such that

$$\xi \in N_P\left[\bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T) - C, \bar{y} + \bar{T}(\bar{z}) \right].$$

If we define $g : \mathscr{L}_+(Z,Y) \to 2^Y$ as $g(T) = -L(\bar{x},T)$, it can be seen that $g(\mu T_1 + (1-\mu)T_2) = \mu g(T_1) + (1-\mu) g(T_2)$, for $T_1, T_2 \in \mathscr{L}_+(Z,Y)$ and $\mu \in [0,1]$. Thus g is a C-convexlike set-valued map and therefore $\bigcup_{T \in \mathscr{L}_+(Z,Y)} g(T) + C$ is a convex set in Y which further implies that $\bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T) - C$ is a convex set in Y. Now from Lemma 6.1, we have

$$\langle \xi, y' - \bar{y} - \bar{T}(\bar{z}) \rangle \leq 0 \; \forall \; y' \in \bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T) - C.$$

As $\bar{y} \in \bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T) - C$, we get $\langle \xi, -\bar{T}(\bar{z}) \rangle \leq 0$ which is a contradiction to (6.4). Hence $\bar{T}(\bar{z}) = 0$.

Proposition 6.2. *For* $(\bar{x}, \bar{T}) \in X \times \mathscr{L}_+(Z,Y)$*, if condition* (6.3) *of Proposition* 6.1 *holds for some* $\bar{y} \in F(\bar{x}), \bar{z} \in G(\bar{x})$ *and* $F(\bar{x})$ *is* $-C$*-closed then* $\bar{y} \in \mathrm{Pr}^0[F(\bar{x}), C]$.

Proof. As

$$(F(\bar{x}) - \bar{y}) \cap C \subseteq \left(\bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T) - \bar{y} + \bar{T}(\bar{z}) \right) \cap C = \{0\},$$

therefore $\bar{y} \in \mathrm{Max}\,[F(\bar{x}), C]$. In view of Proposition 6.1, we have $\bar{T}(\bar{z}) = 0$ therefore $F(\bar{x}) - C \subseteq \bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T) - C$ and hence

$$N_P\left[\bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T) - C, \bar{y} + \bar{T}(\bar{z}) \right] \cap C^{*0} \subseteq N_P(F(\bar{x}) - C, \bar{y}) \cap C^{*0}.$$

Because

$$\bar{y} + \bar{T}(\bar{z}) \in \mathrm{Pr}^0\left[\bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T), C \right],$$

by definition, we have

$$N_P\left[\bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T) - C, \bar{y} + \bar{T}(\bar{z}) \right] \cap C^{*0} \neq \phi.$$

Hence $N_P(F(\bar{x}) - C, \bar{y}) \cap C^{*0} \neq \phi$; that is, $\bar{y} \in \mathrm{Pr}^0[F(\bar{x}), C]$.

Remark 6.1. In Example 6.2, it can be seen that for the proximal proper saddle point $(\bar{x},\bar{T})$ of L where $\bar{x}=0$ and $\bar{T}(x)=(x,0)$, there exists $\bar{y}=(0,0)\in F(\bar{x}),\bar{z}=0\in G(\bar{x})$ such that $\bar{T}(\bar{z})=0, G(\bar{x})\subseteq -R_+$ and

$$\bar{y}\in \mathrm{Pr}\left[\bigcup_{x\in X} L(x,\bar{T}),C\right]\cap \mathrm{Pr}^0[F(\bar{x}),C].$$

From the above two propositions, we have the following theorem which gives the necessary conditions for the existence of a proximal proper saddle point.

Theorem 6.1. *If $(\bar{x},\bar{T})$ is a proximal proper saddle point of L and $F(\bar{x})$ is $-C$-closed, then there exist $\bar{y}\in F(\bar{x})$, $\bar{z}\in G(\bar{x})$ such that*

(i) $\bar{T}(\bar{z})=0.$
(ii) $G(\bar{x})\subseteq -D.$
(iii) $\bar{y}\in \mathrm{Pr}\left[\bigcup_{x\in X} L(x,\bar{T}),C\right]\cap \mathrm{Pr}^0[F(\bar{x}),C].$

In the next theorem, we establish that conditions (i)–(iii) in Theorem 6.1 are also sufficient for the existence of proximal proper saddle points for the Lagrangian map.

Theorem 6.2. *If conditions* (i)–(iii) *of Theorem* 6.1 *hold for some $\bar{x}\in X$, $\bar{y}\in F(\bar{x})$, $\bar{z}\in G(\bar{x})$, and $\bar{T}\in \mathscr{L}_+(Z,Y)$, where $F(\bar{x})$ is $-C$-closed, then $(\bar{x},\bar{T})$ is a proximal proper saddle point of L.*

Proof. From condition (i), it is obvious that

$$\bar{y}=\bar{y}+\bar{T}(\bar{z})\in L(\bar{x},\bar{T}). \tag{6.5}$$

On using condition (ii), we have

$$\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(\bar{x},T)=F(\bar{x})+\bigcup_{T\in\mathscr{L}_+(Z,Y)} T(G(\bar{x}))\subseteq F(\bar{x})-C. \tag{6.6}$$

As $\bar{y}\in \mathrm{Max}\ [F(\bar{x}),C]$, it follows that

$$\bar{y}\in \mathrm{Max}\left[\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(\bar{x},T),C\right].$$

Again from (6.2) and Lemma 6.1, it is obvious that

$$N_P(F(\bar{x})-C,\bar{y})\subseteq N_P\left[\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(\bar{x},T)-C,\bar{y}\right]$$

which implies that

$$\bar{y}\in \mathrm{Pr}^0\left[\bigcup_{T\in\mathscr{L}_+(Z,Y)} L(\bar{x},T),C\right]$$

as $\bar{y} \in \mathrm{Pr}^0[F(\bar{x}),C]$. From Theorem 6.1, condition (iii), and the above relation, it follows that $(\bar{x},\bar{T})$ is a proximal proper saddle point of L.

6.5 Optimality Criteria in Terms of Proximal Proper Efficient Saddle Point

We first introduce the concept of proper efficiency in terms of proximal normal cone for the problem (VP).

Definition 6.3. A point $\bar{x} \in S$ is called a *proximal proper minimal solution* of (VP), if $F(\bar{x}) \cap \mathrm{Pr}[F(S),C] \neq \phi$ where $F(S)$ is C-closed.

Definition 6.4. A point $(\bar{x},\bar{y})$ is said to be a *proximal proper minimizer* of (VP) if $\bar{y} \in F(\bar{x}) \cap \mathrm{Pr}[F(S),C]$ where $F(S)$ is C-closed.

The next theorem gives the sufficient optimality criteria for proximal proper efficiency of (VP).

Theorem 6.3. *If $(\bar{x},\bar{T})$ is a proximal proper saddle point of L and $F(\bar{x})$ and $F(S)$ are $-C$-closed and C-closed sets, respectively, then $\bar{x}$ is a proximal proper minimal solution of* (VP).

Proof. From Theorem 6.1, we have $\bar{z} \in G(\bar{x}) \subseteq -D$ which implies $\bar{x}$ is a feasible solution of (VP). Again from Theorem 6.1, we have $0 \in \bar{T}(G(\bar{x}))$, which implies

$$F(S) - \bar{y} \subseteq \bigcup_{x\in S}(F(x) + \bar{T}(G(x))) - \bar{y} \subseteq \bigcup_{x\in X} L(x,\bar{T}) - \bar{y}. \tag{6.7}$$

From condition (iii) of Theorem 6.1, it follows that $\bar{y} \in \mathrm{Min}[\bigcup_{x\in X} L(x,\bar{T}),C]$ which along with the above relation yields $\bar{y} \in \mathrm{Min}[F(S),C]$. Again from condition (iii) of Theorem 6.1, we have

$$N_P\left[\bigcup_{x\in X} L(x,\bar{T}) + C, \bar{y}\right] \cap (-C^{*0}) \neq \phi.$$

Now from (6.7) and Lemma 6.1, it follows that $\xi \in N_P(F(S)+C,\bar{y})$ for some $\xi \in -C^{*0}$. Therefore $\bar{y} \in \mathrm{Pr}[F(S),C]$ and hence $\bar{y}$ is a proximal proper minimal solution of (VP).

Remark 6.2. In Example 6.3, the feasible set is $S = -R_+$ and $(0,0) \in F(\bar{x}) \cap \mathrm{Pr}[F(S),C]$. Thus $\bar{x} = 0$ is a proximal proper minimal solution of (VP).

We say that (VP) satisfies the generalized Slater's constraint qualification if there exists x' such that $G(x') \cap (-\mathrm{int}D) \neq \phi$. We recall that a set-valued map $F : A \to 2^Y$, $A \subseteq X$, is said to be C-convex on A, if A is a convex set and for all $x,\ u \in A$, $y \in F(x)$, $v \in F(u)$, and $\lambda \in (0,1)$,

$$(1-\lambda)v+\lambda y \in F((1-\lambda)u+\lambda x)+C,$$

where $C \subseteq Y$ is a closed convex pointed cone. Also F is said to be C-convexlike on A if for all $x,u \in A, y \in F(x), v \in F(u)$, and $\lambda \in (0,1)$,

$$(1-\lambda)v+\lambda y \in F(A)+C.$$

It is obvious that F is C-convex if and only if epi F is a convex set and that F is C-convexlike if and only is $F(A)+C$ is a convex set. The following lemma is the alternative theorem that has been proved in [10] for a more general class of set-valued maps, namely C-semilocally convexlike set-valued maps. Every C-semilocally convexlike map is C-convexlike, therefore the following is a particular case of Theorem 4.1 in [10].

Lemma 6.3. *If the map F is a C-convexlike map on X such that $F(X)$ is C-closed, then exactly one of the following systems is consistent.*

(i) $\exists\, x \in X, F(x) \cap (-\text{int}C) \neq \phi$.
(ii) $\exists\, \varphi \in C^* \setminus \{0\},\ \forall\, y \in F(X), \langle \varphi, y \rangle \geq 0$.

We now have the necessary optimality criteria under the generalized Slater's constraint qualification and cone convexlike assumptions.

Theorem 6.4. *Let $(\bar{x},\bar{y})$ be a proximal proper minimizer of* (VP) *such that $F(\bar{x})$ is $-C$-closed and the following conditions hold.*

(i) F is C-convexlike on S and (F,G) is $(C \times D)$-convexlike on X.
(ii) (VP) satisfies the generalized Slater's constraint qualification.
(iii) $G(\bar{x}) \subseteq -D$, $\bar{y} \in \text{Pr}^0[F(\bar{x}),C]$.

Then there exists $\bar{T} \in \mathscr{L}_+(Z,Y)$ such that $(\bar{x},\bar{T})$ is a proximal proper saddle point of L.

Proof. Because $(\bar{x},\bar{y})$ is a proximal proper minimizer of (VP) $\bar{y} \in F(\bar{x}) \cap \text{Pr}[F(S),C]$. Thus there exists $\xi \in -C^{*0}$ such that $\xi \in N_P(F(S)+C,\bar{y})$. Because F is C-convexlike on S, the set $F(S)+C$ is convex and hence by Lemma 6.1, we have

$$\langle \xi, y' - \bar{y} \rangle \leq 0\ \forall\, y' \in F(S)+C.$$

$F(S) \subseteq F(S)+C$, thus it follows that

$$\langle \xi, y' - \bar{y} \rangle \leq 0\ \forall\, y' \in F(S).$$

Define $\psi = -\xi$; then $\psi \in C^{*^0}$ and

$$\langle \psi, y' - \bar{y} \rangle \geq 0\ \forall\, y' \in F(S).$$

Proceeding on the lines of Theorem 5.1 of Li [11], it can be seen that

$$\langle \psi, y' - \bar{y} \rangle \geq 0\ \forall\, y' \in F(x)+\bar{T}(G(x))\ \forall\, x \in X. \tag{6.8}$$

We now claim that

$$\bar{y} \in \mathrm{Min}\left[\bigcup_{x\in X} L(x,\bar{T}), C\right].$$

Otherwise there exists $d \neq 0$ such that

$$d \in \left(\bigcup_{x\in X} L(x,\bar{T}) - \bar{y}\right) \cap (-C);$$

that is, there exists $y \in F(x) + \bar{T}(G(x))$ for some $x \in X$ such that $d = y - \bar{y}$. As $\psi \in C^{*0}$, we have $\langle \psi, d\rangle < 0$ which implies $\langle \psi, y - \bar{y}\rangle < 0$, a contradiction to (6.8). Thus

$$\bar{y} \in \mathrm{Min}\left[\bigcup_{x\in X} L(x,\bar{T}), C\right].$$

As (F,G) is $(C\times D)$-convexlike on X, the set-valued map $F + \bar{T}(G)$ is C-convexlike on X. Thus $F(X)+\bar{T}(G(X))+C$ is a convex set; that is, $\bigcup_{x\in X} L(x,\bar{T})+C$ is a convex set. As $\langle -\psi, c\rangle \leq 0$ for every $c \in C$, from (6.8), it follows that

$$\langle -\psi, y' - \bar{y}\rangle \leq 0 \ \forall\, y' \in \bigcup_{x\in X} L(x,\bar{T}) + C.$$

Therefore by Lemma 6.1, we have

$$\xi = -\psi \in N_P\left[\bigcup_{x\in X} L(x,\bar{T}) + C, \bar{y}\right].$$

As $\xi \in -C^{*0}$, it follows that

$$\bar{y} \in \mathrm{Pr}\left[\bigcup_{x\in X} L(x,\bar{T}), C\right].$$

Hence by Theorem 6.2, we conclude that $(\bar{x},\bar{T})$ is a proximal proper saddle point of L.

Remark 6.3. The following example illustrates that the condition of cone convexlikeness made in Theorem 6.4 cannot be relaxed.

Example 6.4. Let $X = R$, $Y = R^2$, $Z = R$, $C = R^2_+$, $D = R_+$. Define $F : X \to 2^Y$ by

$$F(x) = \begin{cases} \{(x,1)\} & \text{if } x < 0 \\ \{(x,t) : -x \leq t \leq x\} & \text{if } 0 \leq x \leq 1 \\ \{(-t,0) : t > 0\} & \text{if } x > 1, \end{cases}$$

and $G : X \to 2^Z$ by $G(x) = \{x\}$. For $\bar{x} = 0, \bar{y} = (0,0), (\bar{x},\bar{y})$ is a proximal proper minimizer of (VP) as $\bar{y} \in \mathrm{Pr}[F(S),C]$. Clearly, (VP) satisfies generalized Slater's constraint qualification, $G(\bar{x}) \subseteq -D$ and $\bar{y} \in \mathrm{Pr}^0[F(\bar{x}),C]$. However, F is not

C-convexlike on S as $F(S)+C$ is not a convex set. We now show that there does not exist any $\bar{T} \in \mathscr{L}_+(Z,Y)$ such that $(\bar{x},\bar{T})$ is a proximal proper saddle point of L. Every linear continuous operator $T \in \mathscr{L}_+(Z,Y)$ is from $T(x) = (\alpha x, \beta x), (\alpha,\beta) \in R^2$ and for $T \in \mathscr{L}_+(Z,Y), T(D) \subseteq C$, therefore $(\alpha,\beta) \in R^2_+$. For $T \in \mathscr{L}_+(Z,Y)$, the Lagrangian is defined as

$$L(x,T) = \begin{cases} \{(x+\alpha x, 1+\beta x)\} & \text{if } x < 0 \\ \{(x+\alpha x, t+\beta x) : -x \leq t \leq x\} & \text{if } 0 \leq x \leq 1 \\ \{(-t+\alpha x, \beta x) : t > 0\} & \text{if } x > 1. \end{cases}$$

Here

$$\mathrm{Pr}^0\left[\bigcup_{T \in \mathscr{L}_+(Z,Y)} L(\bar{x},T), C\right] = L(\bar{x},T) = \{\bar{y}\}.$$

Next, we show

$$\bar{y} \notin \mathrm{Pr}\left[\bigcup_{x \in X} L(x,T), \bar{y}\right];$$

that is,

$$N_P\left[\bigcup_{x \in X} L(x,T) + C, \bar{y}\right] \cap (-C^{*0}) = \phi$$

for any $T \in \mathscr{L}_+(Z,Y)$. If $\beta = 0$ and $\alpha \geq 0$ in $T(x)$ then $(-t+2\alpha, 0) \in L(2,T) \subseteq \bigcup_{x \in X} L(x,T)$ for any $t > 0$ which implies that no vector in $-C^{*^0}$ is in $N_P\left[\bigcup_{x \in X} L(x,T) + C, \bar{y}\right]$. If $\beta > 0$ and $\alpha \geq 0$ in $T(x)$ then

$$(-(1+\alpha)/\beta, 0) \in L(-1/\beta, T) \subseteq \bigcup_{x \in X} L(x,T)$$

which again implies that no vector in $-C^{*0}$ is in $N_P\left[\bigcup_{x \in X} L(x,T) + C, \bar{y}\right]$. Hence $(\bar{x},T)$ is not a proximal proper saddle point for any $T \in \mathscr{L}_+(Z,Y)$.

6.6 Conclusions

The concept of proximal proper saddle point is used to obtain an optimality criterion for a constrained set-valued optimization problem. Various notions of saddle points have been studied in the literature ([1, 11]). The independence of proximal proper efficiency and hence proximal saddle points with other saddle points makes its study significant and worthwhile. The observation is further enhanced by the fact that proper efficient points that cannot be characterized by other notions of saddle points ([1, 11]), have a characterization in terms of proximal proper saddle point.

References

1. A. Mehra, Super efficiency in vector optimization with nearly convexlike set valued maps, *J. Math. Anal.* **276** (2002), 815–832.
2. J.P. Aubin, and H. Frankowska, Set-Valued Analysis, Birkhaüser, Boston, 1990.
3. H.P. Benson, An improved definition of proper efficiency for vector minimization with respect to cones, *J. Math. Anal. Appl.* **71** (1979), 232–241.
4. J.M. Borwein, Proper efficient points for maximization with respect to cones, *SIAM J. Cont. Optim.* **15** (1977), 57–63.
5. F.H. Clarke, Y.S. Ledyaev, R.J. Stern, and P.R. Wolenski, Nonsmooth Analysis and Control Theory, Springer Verlag, New York, 1998.
6. A.M. Geoffrion, Proper efficiency and the theory of vector maximization, *J. Math. Anal. Appl.* **22** (1968), 618–630.
7. A. Guerragio, E. Molho, and A. Zaffaroni, On the notion of proper efficiency in vector optimization, *J. Optim. Theory Appl.* **82** (1994), 1–21.
8. K. Klein, and A.C. Thompson, Theory of Correspondences, John Wiley, New York, 1984.
9. H.W. Kuhn, and A.W. Tucker, Nonlinear programming, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, CA, 1951.
10. C.S. Lalitha, and R. Arora, Proper efficiency for minimization with respect to proximal normal cones, *Bull. Aust. Math. Soc.* **71** (2005), 215–224.
11. Z.F. Li, Benson proper efficiency in the vector optimization of set-valued maps, *J. Optim. Theory Appl.* **98**(3) (1998), 623–649.
12. Z.F. Li, and G.Y. Chen, Lagrangian multipliers, saddle points and duality in vector optimization of set-valued maps, *J. Math. Anal. Appl.* **215** (1997), 297–316.
13. W.D. Rong, and Y.N. Wu, Characterizations of super efficiency in cone-convexlike vector optimization with set-valued maps, *Math. Meth. Oper. Res.* **48** (1998), 247–258.
14. W.I. Zangwill, Nonlinear Programming: A Unified Approach, Prentice–Hall, Eagelwood Cliffs, NJ, 1969.

Chapter 7
Metric Regularity and Optimality Conditions in Nonsmooth Optimization

Anulekha Dhara and Aparna Mehra

Abstract The concept of metric regularity and its role in deriving the optimality conditions for optimization problems is not new. This chapter presents the notion of metric regularity and explores the relationship between a modified version of the well-known basic constraint qualification with that of metric regularity. We also study its application in obtaining the Karush–Kuhn–Tucker optimality conditions for nonsmooth optimization problems with set inclusion and abstract constraints by converting the constrained problem into an unconstrained problem.

7.1 Introduction

A lot of work has been done in the literature to study optimality conditions, well known as the Karush–Kuhn–Tucker (KKT) optimality conditions, for nonsmooth optimization problems. The common approach taken to establish these conditions primarily involves two stages. First deriving the Fritz–John optimality conditions wherein the Lagrange multipliers associated with the subdifferentials of the objective function and the constraint functions of the problem are not all zeroes. In turn, it may result in the multiplier being associated with the subdifferential of the objective function being zero. But in an optimization problem, the objective function is pivotal and hence the optimality conditions are expected to keep up this perspective. But the Fritz–John optimality conditions may fail to satisfy this aspect. In order to ensure that the Lagrange multiplier associated with the subdifferential of the objective function is nonzero, in the second stage some conditions known as the

Anulekha Dhara
Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110 016, India, e-mail: anulekha_iitd@yahoo.co.in

Aparna Mehra
Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110 016, India, e-mail: apmehra@maths.iitd.ac.in

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_7,

constraint qualification are imposed on the constraints consequently leading to the KKT optimality conditions.

Although the aforementioned scheme is widely adopted, there is yet another approach to deriving the KKT optimality condition. The main idea in the latter is to somehow convert the constrained optimization problem into an appropriate unconstrained optimization problem and replace the constraint qualification with the notion of metric regularity. A great deal of work [1, 4, 6–11, 16] has been reported in the literature with respect to metric regularity and its relationship to various constraint qualifications and also how it can be used as a tool to derive the KKT optimality conditions. The metric regularity helps to express the tangential approximation or the normal cone of the feasible region in terms of the functions and the sets involved in the constraints. Here, we work along these lines to establish the KKT optimality conditions for a nonsmooth optimization problem involving set inclusion and abstract constraints.

This article is organized as follows. Some notations and preliminary results are quoted in Section 7.2, and Section 7.3 presents certain well-known constraint qualifications and studies a relationship between modified basic constraint qualification and metric regularity. Section 7.4 deals with our primary objective of obtaining the KKT optimality conditions for a nonsmooth optimization problem via metric regularity. We end this chapter with some concluding remarks in Section 7.5.

7.2 Notations and Preliminaries

Let $S \subseteq \mathfrak{R}^n$. We denote the convex hull of S and closure of S by *co* S and *cl* S, respectively. Let $\mathscr{B}_{\mathfrak{R}^n}$ denote the open unit ball in $\mathfrak{R}^n$ and $x_0 + \varepsilon\mathscr{B}_{\mathfrak{R}^n}$ be the open ball of radius ε centered at $x_0 \in \mathfrak{R}^n$. For $(x_1, y_1),\ (x_2, y_2) \in \mathfrak{R}^l \times \mathfrak{R}^k$, $\langle (x_1, y_1), (x_2, y_2) \rangle = \langle x_1, x_2 \rangle + \langle y_1, y_2 \rangle$.

Let $\phi : \mathfrak{R}^n \to \bar{\mathfrak{R}}$. The *domain* of ϕ is defined as *dom* $\phi = \{x \in \mathfrak{R}^n : \phi(x) < +\infty\}$ and its *epigraph* is given by *epi* $\phi = \{(x, r) \in \mathfrak{R}^n \times \mathfrak{R} : \phi(x) \leqq r\}$. ϕ is said to be *Lipschitz* on $S \subseteq \mathfrak{R}^n$ with Lipschitz constant $K > 0$ if

$$|\phi(x_1) - \phi(x_2)| \leqq K\|x_1 - x_2\|, \qquad \forall\, x_1, x_2 \in S,$$

and *Lipschitz* around $x_0 \in \mathfrak{R}^n$ if there exist $K_{x_0} > 0$ and a neighborhood $\mathscr{N}(x_0)$ of x_0 such that

$$|\phi(x_1) - \phi(x_2)| \leqq K_{x_0}\|x_1 - x_2\|, \qquad \forall\, x_1, x_2 \in \mathscr{N}(x_0).$$

The *distance function*, $d_S : \mathfrak{R}^n \to \mathfrak{R}$, to S is defined as

$$d_S(x_0) = d(x_0; S) = \inf_{x \in S} \|x - x_0\|.$$

d_S is a Lipschitz function on $\mathfrak{R}^n$ with Lipschitz constant 1. For a convex set S, d_S is a convex function. Also for a closed set S, $x_0 \notin S$ implies $d_S(x_0) > 0$, and the infimum in the above expression is attained say at $x^* \in S$; then $x^* = P(x_0; S)$, the *projection* of x_0 on S. In particular, for $S = \{x^*\}$,

$$d(x_0; x^*) = \|x_0 - x^*\|,$$

the distance between the points x_0 and x^*.

The *indicator function*, $\delta_S : \mathfrak{R}^n \to \bar{\mathfrak{R}}$, to S is defined as

$$\delta_S(x_0) = \begin{cases} 0, & x_0 \in S \\ +\infty, & x_0 \notin S. \end{cases}$$

δ_S is a lower semicontinuous function for a closed set S.

Next we present the notion of lower semicontinuity of a multifunction (multi-valued function) and certain other results which are used in the subsequent sections.

Let $G : \mathfrak{R}^n \rightrightarrows \mathfrak{R}^m$ be a multifunction. The *graph* of G is defined as

$$grp\ G = \{(x,y) \in \mathfrak{R}^n \times \mathfrak{R}^m : y \in G(x)\}.$$

G is said to be *lower semicontinuous* at $(x_0, y_0) \in grp\ G$ if for any neighborhood $\mathcal{N}(y_0)$ of y_0, there exists a neighborhood $\mathcal{N}(x_0)$ of x_0 such that

$$G(x) \cap \mathcal{N}(y_0) \neq \emptyset, \qquad \forall\, x \in \mathcal{N}(x_0).$$

Proposition 7.1 (Proposition 1, [8]). *$G : \mathfrak{R}^n \rightrightarrows \mathfrak{R}^m$ is a lower semicontinuous multifunction at $(x_0, y_0) \in grp\ G$ if and only if*

$$\lim_{(x,y)\to(x_0,y_0)} d(y; G(x)) = 0.$$

Proposition 7.2 (Theorem 2.1, [1]). (Ekeland's Variational Principle) *Let $\phi : \mathfrak{R}^n \to \bar{\mathfrak{R}}$ be a lower semicontinuous function with $x_0 \in dom\ \phi$ and $S \subseteq \mathfrak{R}^n$ be a closed set. Assume that for $\varepsilon > 0$,*

$$\phi(x_0) \leqq \inf_{x\in S} \phi(x) + \varepsilon.$$

Then for every $\lambda > 0$ there exists $x_\lambda \in S$ such that

$$\begin{aligned} d(x_\lambda; x_0) &\leqq \lambda, \\ \phi(x_\lambda) &\leqq \phi(x_0), \\ \phi(x_\lambda) &\leqq \phi(x) + \lambda^{-1}\varepsilon\, d(x; x_\lambda), \qquad \forall\, x \in S. \end{aligned}$$

Proposition 7.3 (Proposition 2.4.3, [2]). *Let $\phi : \mathfrak{R}^n \to \mathfrak{R}$ be Lipschitz on $S \subseteq \mathfrak{R}^n$ with Lipschitz constant $K > 0$. Assume that ϕ attains a minimum over $F \subseteq S$ at $x_0 \in F$. Then, for every $K' \geqq K$, the function $\phi + K' d_F$ attains a minimum over S at x_0.*

Because we aim at working in a nonsmooth scenario, it is natural for us to talk about the notion of subdifferentiability. In the literature, various types of subdifferentiability have been discussed. In our present study we use the proximal and the basic (or limiting) subdifferentials and also a relation of the coderivative with that of the basic subdifferential. For details on proximal and basic (or limiting) subdifferentials one may refer to Clarke et al. [3]. The main reason for including these subdifferentials in our work is that the concept of metric regularity has been well studied [9, 10] under generalized subdifferentials and coderivatives such as approximate subdifferentials, Fréchet subdifferentials, and coderivatives. The following definitions and results are taken from [3].

Definition 7.1. Let $\phi : \mathfrak{R}^n \to \bar{\mathfrak{R}}$ and $x_0 \in dom\ \phi$. A vector $\xi \in \mathfrak{R}^n$ is called a *proximal subgradient* of ϕ at x_0 if and only if there exist $\sigma > 0$ and $\delta > 0$ such that

$$\phi(x) - \phi(x_0) \geqq \langle \xi, x - x_0 \rangle - \sigma \|x - x_0\|^2, \qquad \forall\, x \in x_0 + \delta \mathscr{B}_{\mathfrak{R}^n}.$$

The collection of proximal subgradients of ϕ at x_0, called a *proximal subdifferential* to ϕ at x_0, denoted by $\partial_P \phi(x_0)$, is always a convex set. However, it is not necessarily nonempty and closed/open. Observe that if ϕ is twice differentiable then $\partial_P \phi(x_0) = \{\nabla \phi(x_0)\}$, and if $x_0 \notin dom\ \phi$ then, $\partial_P \phi(x_0) = \emptyset$. Even for $x_0 \in dom\ \phi$, $\partial_P \phi(x_0)$ may turn out to be empty. To move away from such a trivial scenario, we have the concept of basic or limiting subdifferentials.

Definition 7.2. Let $\phi : \mathfrak{R}^n \to \bar{\mathfrak{R}}$ and $x_0 \in dom\ \phi$. The *basic* or the *limiting subdifferential* of ϕ at x_0 is given by

$$\partial_L \phi(x_0) = \left\{ \lim_{k \to \infty} \xi_k : \xi_k \in \partial_P \phi(x_k),\ x_k \to^{\phi} x_0 \right\},$$

where $x_k \to^{\phi} x_0$ means $x_k \to x_0$ and $\phi(x_k) \to \phi(x_0)$.

The basic subdifferential is closed but not necessarily convex. From the definition it is obvious that for a point $x_0 \in dom\ \phi$, $\partial_P \phi(x_0) \subseteq \partial_L \phi(x_0)$. At points $x_0 \notin dom\ \phi$, $\partial_L \phi(x_0) = \emptyset$. For a locally Lipschitz function, $\partial_L \phi(x_0)$ is always nonempty, and for a convex function it coincides with the convex subdifferential of Rockafellar [14]. Moreover, for a Lipschitz function ϕ on an open convex set $S \subseteq \mathfrak{R}^n$ with Lipschitz constant $K > 0$, if $\xi \in \partial_P \phi(x) \cup \partial_L \phi(x)$, then $\|\xi\| \leqq K$. In this case, $cl\ co\ \partial_L \phi(x_0)$ coincides with $\partial_C \phi(x_0)$, the Clarke subdifferential [2] of ϕ at x_0. Also, for $S \subseteq \mathfrak{R}^n$ if we take $\phi = \delta_S$, it leads to the notion of normal cones.

Definition 7.3. A vector $\xi \in \mathfrak{R}^n$ is called *proximal normal* to $S \subseteq \mathfrak{R}^n$ at $x_0 \in S$ if and only if there exist $\sigma > 0$ and $\delta > 0$ such that

$$\langle \xi, x - x_0 \rangle \leqq \sigma \|x - x_0\|^2, \qquad \forall\, x \in S \cap (x_0 + \delta \mathscr{B}_{\mathfrak{R}^n}).$$

The set of all proximal normals forms a cone called the *proximal normal cone* to S at x_0 and is denoted by $N_P(x_0; S)$. The proximal normal cone is a convex cone that may neither be open nor closed and may sometimes reduce to a trivial scenario containing only the zero element.

Definition 7.4. The *basic* or the *limiting normal cone* to $S \subseteq \mathfrak{R}^n$ at $x_0 \in S$ is given by

$$N_L(x_0;S) = \left\{ \lim_{k\to\infty} \xi_k : \xi_k \in N_P(x_k;S),\ x_k \to^S x_0 \right\},$$

where $x_k \to^S x_0$ means the sequence $\{x_k\} \subseteq S$ converges to $x_0 \in S$.

The basic normal cone is a closed cone but unlike the proximal normal cone it may not be convex. Moreover, at any boundary point of the set, the basic normal cone is always nontrivial. In terms of the indicator function to the set S, $N_P(x_0;S) = \partial_P \delta_S(x_0)$ and $N_L(x_0;S) = \partial_L \delta_S(x_0)$. The proximal subdifferential can also be expressed in terms of the proximal normal cone as

$$\partial_P \phi(x_0) = \{\xi \in \mathfrak{R}^n : (\xi,-1) \in N_P((x_0,\phi(x_0)); epi\ \phi)\}.$$

Similarly, the basic subdifferential is expressed in terms of the basic normal cone as

$$\partial_L \phi(x_0) = \{\xi \in \mathfrak{R}^n : (\xi,-1) \in N_L((x_0,\phi(x_0)); epi\ \phi)\}.$$

Observe that for $\alpha > 0$ such that $(\xi,-\alpha) \in N_P((x_0,\phi(x_0)); epi\ \phi)$ implies $\xi/\alpha \in \partial_P \phi(x_0)$. Below we state a result relating the normal cones to the subdifferentials of the distance function.

Proposition 7.4. *Let $S \subseteq \mathfrak{R}^n$ be a closed set. Then for $x_0 \in S$,*

$$\partial_P d_S(x_0) = N_P(x_0;S) \cap \mathscr{B}_{\mathfrak{R}^n} \quad \text{and} \quad \partial_L d_S(x_0) = N_L(x_0;S) \cap \mathscr{B}_{\mathfrak{R}^n},$$
$$N_P(x_0;S) = \bigcup_{\lambda \geqq 0} \lambda\, \partial_P d_S(x_0) \quad \text{and} \quad N_L(x_0;S) = \bigcup_{\lambda \geqq 0} \lambda\, \partial_L d_S(x_0),$$

and for $x_0 \notin S$, the projection of x_0 on S, $P(x_0;S) = \{x^*\}$ and

$$\partial_P d_S(x_0) = \partial_L d_S(x_0) = \left\{ \frac{x_0 - x^*}{\|x_0 - x^*\|} \right\}$$

and $\partial_P d_S(x_0) \subset N_P(x^*;S)$.

Observe that for $x_0 \notin S$, if $\xi \in \partial_P d_S(x_0) \cup \partial_L d_S(x_0)$ then $\|\xi\| = 1$. Because we desire to obtain the KKT optimality conditions, we require the sum and the chain (or composition) rules. Here, we state the calculus rules for the basic subdifferentials. For their detailed proofs one may refer to [3].

Theorem 7.1 (Sum Rule). *If $\phi_i : \mathfrak{R}^n \to \mathfrak{R},\ i = 1,2$ with one of the functions as Lipschitz, then*

$$\partial_L(\phi_1 + \phi_2)(x_0) \subseteq \partial_L \phi_1(x_0) + \partial_L \phi_2(x_0).$$

Theorem 7.2 (Chain rule). *Let $\phi(x) = (h \circ g)(x)$, where $g : \Re^n \to \Re^m$ is Lipschitz around x_0 and $h : \Re^m \to \bar{\Re}$ is Lipschitz around $g(x_0)$, then*

$$\partial_L \phi(x_0) \subseteq \{\partial_L(\mu g)(x_0) : \mu \in \partial_L h(g(x_0))\}.$$

We end this section with the notion of coderivative and its relation to the basic subdifferential taken from Rockafellar and Wets [15].

Definition 7.5. Let $\phi : \Re^n \to \Re^m$. The *coderivative* of ϕ at $x_0 \in \Re^n$ is the multi-valued function $D^*\phi(x_0) : \Re^m \rightrightarrows \Re^n$ defined by

$$\xi \in D^*\phi(x_0)(\mu) \quad \Leftrightarrow \quad (\xi, -\mu) \in N_L((x_0, \phi(x_0)); grp\ \phi).$$

If ϕ is continuous at x_0, then for any $\mu \in \Re^m$,

$$D^*\phi(x_0)(\mu) = \partial_L(\mu\phi)(x_0).$$

7.3 Constraint Qualifications and Metric Regularity

The main objective of this work is to establish the KKT optimality conditions for optimization problem (P),

$$\begin{aligned} &\inf f(x) \\ &\text{subject to } g(x) \in D \\ &\qquad\qquad x \in C, \end{aligned}$$

where $f : \Re^n \to \Re$ and $g : \Re^n \to \Re^m$ are, respectively, Lipschitz and continuous on $\Re^n$, and $D \subseteq \Re^m$ and $C \subseteq \Re^n$ are closed sets. Let the feasible set be denoted by $F = \{x \in \Re^n : g(x) \in D,\ x \in C\}$ which is closed. As already discussed in the beginning, to develop the KKT optimality conditions for (P), one has to invoke appropriate constraint qualification. Some of the well-known constraint qualifications corresponding to F that are studied in the literature are as follows.

(SCQ) Slater constraint qualification: There exists $x_0 \in C$ such that $g(x_0) \in int\ D$.
(BCQ) Basic constraint qualification at $x_0 \in F$: If

$$\mu \in N_L(g(x_0); D), \qquad 0 \in \partial(\mu g)(x_0) + N_L(x_0; C) \Rightarrow \mu = 0.$$

(RCQ) Robinson's constraint qualification at $x_0 \in F$: For g is differentiable at x_0 and continuous on C if

$$0 \in int\{\nabla g(x_0)(C - x_0) - (D - g(x_0))\}.$$

For the next two constraint qualifications, let $D = \Re^m_-$, $C = \Re^n$, and g be differentiable. Denote the active index set by $I(x_0) = \{i \in \{1, \ldots, m\} : g_i(x_0) = 0\}$.

$(MFCQ)$ Mangasarian Fromovitz constraint qualification: There exists $d \in \Re^n$ such that $\nabla g_i(x_0)d < 0,\ i \in I(x_0)$.

(ACQ) Abadie constraint qualification: For convex function $g_i,\ i = 1,\dots,m$, if

$$T(x_0;F) = \{d \in \Re^n : \nabla g_i(x_0)d \leqq 0,\ i \in I(x_0)\},$$

where $T(x_0;F)$ is the convex tangent cone.

In the differentiable convex scenario, the relation between the above constraint qualifications [5] is

$$(SCQ)/(RCQ) \Leftrightarrow (MFCQ) \Leftrightarrow (BCQ) \Leftrightarrow (ACQ).$$

It is well known that the notion of metric regularity is closely related to these constraint qualifications under particular scenarios. Work has been reported in [11, 16] relating the above constraint qualifications with metric regularity for inequality constraints. To know more about it we move on to define the concept of metric regularity for (P).

Definition 7.6 (Definition 2.1, [9]). Let $x_0 \in S \subseteq \Re^n$ and $G : \Re^n \rightrightarrows \Re^m$ be a multi-valued function with $(x_0,y_0) \in grp\ G$. G is said to be *metrically regular* at (x_0,y_0) with respect to S if there exist $\gamma \geqq 0$ and $\varepsilon > 0$ such that

$$d(x;S \cap G^{-1}(y)) \leqq \gamma d(y;G(x))$$

for all $x \in S \cap (x_0 + \varepsilon\mathscr{B}_{\Re^n})$ and $y \in y_0 + \varepsilon\mathscr{B}_{\Re^m}$. If $S = \Re^n$, G is said to be metrically regular at $(x_0,y_0) \in grp\ G$.

Metric regularity has been characterized in terms of Lipschitz conditions and open coverings in [12, 13, 15]. Here we study the relation between the modified basic constraint qualification and metric regularity by working along the lines of Jourani and Thibault (Theorem 2.3, [9]) for which we need the following result. The proof is along the lines of Theorem 3.2 established by Borwein [1].

Lemma 7.1. *Let $G : \Re^n \rightrightarrows \Re^m$ be lower semicontinuous at $(x_0,y_0) \in grp\ G$. If G is not metrically regular at (x_0,y_0) with respect to S, then there exist sequences $x_k \to x_0$, $y_k \to y_0$ and $\alpha_k \downarrow 0$ such that*

(i) $x_k \in S$.
(ii) $y_k \notin G(x_k)$.
(iii) $d(y_k;G(x_k)) \leqq d(y_k;G(x)) + \alpha_k d(x;x_k),\ \forall\, x \in S$.

Proof. Obviously, the hypothesis implies there exist sequences $c_k \to x_0$ and $y_k \to y_0$ such that $c_k \in S$ and

$$d(c_k;S \cap G^{-1}(y_k)) > kd(y_k;G(c_k)).$$

Define the function $\phi_k(x) = d(y_k; G(x))$ on S. Observe that for every $x \in S$, $\phi_k(x) \geqq 0$ and thus

$$\phi_k(x^*) \leqq \phi_k(x) + \varepsilon_k, \qquad \forall\, x \in S$$

for $x^* = c_k$ and $\varepsilon_k = \phi_k(x^*)$. Applying Ekeland's variational principle (Proposition 7.2), for $\lambda_k = \min\{k\varepsilon_k, \sqrt{\varepsilon_k}\}$, and writing $\alpha_k = \lambda_k^{-1}\varepsilon_k$, there exists $x_k \in S$ such that

$$\begin{aligned} d(x_k; x^*) &\leqq \lambda_k, \\ \phi_k(x_k) &\leqq \phi_k(x^*), \\ \phi_k(x_k) &\leqq \phi_k(x) + \alpha_k d(x; x_k), \qquad \forall\, x \in S. \end{aligned}$$

By lower semicontinuity of G at (x_0, y_0) and Proposition 7.1,

$$\varepsilon_k = \phi_k(c_k) = d(y_k; G(c_k)) \downarrow 0 \text{ as } k \to \infty.$$

This implies $\lambda_k \downarrow 0$, and hence $\alpha_k \downarrow 0$. Therefore, we have constructed sequences $x_k \to x_0$, $y_k \to y_0$, and $\alpha_k \downarrow 0$ such that $x_k \in S$ and

$$d(y_k; G(x_k)) \leqq d(y_k; G(x)) + \alpha_k d(x; x_k), \qquad \forall\, x \in S,$$

leading to (iii). Furthermore, because $d(x_k; c_k) \leqq \lambda_k$, we have

$$d(x_k; c_k) \leqq k d(y_k; G(c_k)) < d(c_k; S \cap G^{-1}(y_k))$$

which implies $x_k \notin S \cap G^{-1}(y_k)$. But $x_k \in S$, consequently $y_k \notin G(x_k)$, thus proving the result. □

Observe that if the multivalued function $G : \Re^n \rightrightarrows \Re^m$ is defined as $G(x) = -g(x) + D$, then

$$G^{-1}(y) = \{x \in \Re^n : y + g(x) \in D\},$$

and hence $G^{-1}(0) = g^{-1}(D)$. We say (P) is *metrically regular* at x_0 with respect to C and D if and only if G is metrically regular at $(x_0, 0)$ with respect to C. It is important to note that $d(y; G(x)) = d(y + g(x); D)$.

Theorem 7.3. *Let $C \subseteq \Re^n$ and $D \subseteq \Re^m$ be closed and $g : \Re^n \to \Re^m$ be Lipschitz around x_0 with Lipschitz constant $K_{x_0} > 0$. Suppose that the modified basic constraint qualification ($\widetilde{BCQ}$), that is,*

$$\mu \in \partial_L d_D(g(x_0)),\ 0 \in \partial_L(\mu g)(x_0) + K' \partial_L d_C(x_0) \Rightarrow \mu = 0,$$

where $K' \geqq K_{x_0}$, holds at x_0. Then (P) is metrically regular at x_0 with respect to C and D.

Proof. Suppose (P) is not metrically regular at x_0 with respect to C and D, which implies G is not metrically regular at $(x_0, 0)$ with respect to C. Therefore, by

Lemma 7.1, there exist sequences $x_k \to x_0$, $y_k \to 0$, and $\alpha_k \downarrow 0$ such that (i)–(iii) hold. Clearly (i) and (iii) imply that x_k minimizes the function $d(y_k;G(x)) + \alpha_k d(x;x_k)$ over C. By Proposition 7.3, x_k minimizes the function ϕ_k over $\mathfrak{R}^n$ where

$$\begin{aligned}\phi_k(x) &= d(y_k;G(x)) + (K' + \alpha_k)d(x;C) + \alpha_k d(x;x_k)\\ &= d(y_k + g(x);D) + (K' + \alpha_k)d(x;C) + \alpha_k d(x;x_k),\end{aligned}$$

where $K' = \max\{1, K_{x_0}\}$. Consequently $0 \in \partial_L \phi_k(x_k)$, which in view of the sum rule (Theorem 7.1) along with the chain rule (Theorem 7.2) implies

$$0 \in \bigcup_{\mu \in \partial_L d_D(h(x_k,y_k))} \partial_L(\mu g)(x_k) + (K' + \alpha_k)\partial_L d_C(x_k) + \alpha_k \mathscr{B}_{\mathfrak{R}^n},$$

where $h(x,y) = y + g(x)$. Hence, there exist $\mu_k \in \partial_L d_D(h(x_k,y_k))$, $\xi_k^1 \in \partial_L(\mu_k g)(x_k)$, $\xi_k^2 \in \partial_L d_C(x_k)$, and $b_k \in \mathscr{B}_{\mathfrak{R}^n}$ such that

$$0 = \xi_k^1 + (K' + \alpha_k)\xi_k^2 + \alpha_k b_k.$$

Because g is Lipschitz around x_0, and h and $(d_D \circ h)$ are Lipschitz around $(x_0, 0)$ and $h(x_0, y_0)$, respectively, by virtue of boundedness of the limiting subdifferentials for Lipschitz functions, the sequences above have convergent subsequences. Without loss of generality, let $\mu_k \to \mu \in \partial_L d_D(g(x_0))$, $\xi_k^1 \to \xi^1 \in \partial_L(\mu g)(x_0)$, and $\xi_k^2 \to \xi^2 \in \partial_L d_C(x_0)$. Therefore, as $k \to \infty$, $\alpha_k \downarrow 0$ and hence

$$0 = \xi^1 + K'\xi^2.$$

$y_k \notin G(x_k)$, hence $h(x_k,y_k) = y_k + g(x_k) \notin D$. By Proposition 7.4, for every k, $\|\mu_k\| = 1$, where $\mu_k \in \partial_L d_D(h(x_k,y_k))$. Therefore, $\|\mu\| = 1$; that is, $\mu \neq 0$, thereby contradicting the assumption $(\widetilde{BCQ})$. □

The following result is an immediate consequence of the above theorem by taking $g(x) = x$ in (P).

Corollary 7.1. *Let $x_0 \in C \cap D$. Suppose that*

$$(-\partial_L d_C(x_0)) \cap \partial_L d_D(x_0) = \{0\}.$$

Then the multivalued function $G(x) = -x + D$ is metrically regular at $(x_0, 0)$ with respect to C.

From Theorem 7.3, we found that $(\widetilde{BCQ})$ acts as a sufficient condition for metric regularity of (P). We present below a simple example to show that the converse in general may fail to hold.

Example 7.1. Let $C = \mathfrak{R}$, $D = \mathfrak{R}_- \subset \mathfrak{R}$, and $g(x) = |x|$. Then $F = \{x \in \mathfrak{R} : |x| \leqq 0\} = \{0\}$. Because g is a convex function, $\partial_P g(0) = \partial g(0) = [-1,1]$. Observe that neither (BCQ) nor $(\widetilde{BCQ})$ is satisfied but the condition of metric regularity holds at $(0,0)$ for $\gamma = 1$ and $\varepsilon > 0$.

It is important to note that g is a nonsmooth function in the above example. But for the case of differentiable function g, we state a result from Rockafellar and Wets [15] that characterizes a particular form of metric regularity in terms of (BCQ). For this purpose we may take note of the following.

If we take $y_0 = 0$ and $y = 0$ in the definition of metric regularity, the inequality becomes

$$d(x; C \cap G^{-1}(0)) \leqq \gamma d(0; G(x))$$

for all $x \in C \cap (x_0 + \varepsilon \mathscr{B}_{\mathfrak{R}^n})$ which can be rewritten as

$$d(x; C \cap G^{-1}(0)) \leqq \gamma (d(0; G(x)) + d(x; C))$$

for all $x \in x_0 + \varepsilon \mathscr{B}_{\mathfrak{R}^n}$. In particular, taking $G(x) = -g(x) + D$, we have

$$d(x; C \cap g^{-1}(D)) \leqq \gamma (d(g(x); D) + d(x; C))$$

for all $x \in x_0 + \varepsilon \mathscr{B}_{\mathfrak{R}^n}$.

Theorem 7.4. *Let $C = \mathfrak{R}^n$, $D \subseteq \mathfrak{R}^m$ be closed and $g : \mathfrak{R}^n \to \mathfrak{R}^m$ be differentiable. For $\gamma \geqq 0$, consider the following metric regularity of (P)*

$$d(x; g^{-1}(D)) \leqq \gamma d(g(x); D)$$

for all $x \in x_0 + \varepsilon \mathscr{B}_{\mathfrak{R}^n}$ (a particular case of metric regularity with $y_0 = y = 0$). Then the necessary and sufficient condition for this metric regularity to hold at x_0 is that

$$\mu \in N_L(g(x_0); D), \qquad \mu \nabla g(x_0) = 0 \Rightarrow \mu = 0.$$

The latter relation is clearly the (BCQ).

7.4 Optimality Conditions

In this section we establish the optimality condition for problem (P) via metric regularity. While establishing the optimality conditions, we invoke the calculus rules of the basic subdifferentials. Recall from Section 7.2 that for the chain rule to hold we require Lipschitz conditions on both g as well as h. In the theorem to follow, we make use of metric regularity to derive the chain rule for the basic subdifferential without imposing any Lipschitz condition on either g or h. The proof can be worked out along the lines of Theorem 3.1 of [9] and is presented here for completeness.

Theorem 7.5. *Let $\phi(x) = (h \circ g)(x)$ where $g : \mathfrak{R}^n \to \mathfrak{R}^m$ is continuous around x_0 and $h : \mathfrak{R}^m \to \bar{\mathfrak{R}}$ is lower semicontinuous around $g(x_0)$ with $x_0 \in dom\ \phi$. Consider the function $G : \mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R} \to \mathfrak{R}^m$ defined as $G(x, y, r) = g(x) - y$ and let $D = \mathfrak{R}^n \times epi\ h$. Assume that the multivalued function $H(x, y, r) = -(x, y, r) + D$ is* *metrically regular at* $(x_0, g(x_0), h(g(x_0)), 0) = (x_0, y_0, r_0, 0)$ *with respect to* $G^{-1}(0)$. *Then*

$$\partial_L(h \circ g)(x_0) \subseteq \bigcup_{\mu \in \partial_L h(y_0)} \partial_L(\mu g)(x_0).$$

Proof. Define $\phi(x) = (h \circ g)(x) = \inf\{r : G(x,y,r) = 0,\ (x,y,r) \in D\}$. Let $\xi \in \partial_L \phi(x_0)$. By Definition 7.2, there exist sequences $x_k \to^{\phi} x_0$ and $\xi_k \in \partial_P \phi(x_k)$ such that $\xi_k \to \xi$. As $\xi_k \in \partial_P \phi(x_k)$, there exist $\sigma_k > 0$ and $\delta_k > 0$,

$$\phi(x) - \phi(x_k) \geqq \langle \xi_k, x - x_k \rangle - \sigma_k \|x - x_k\|^2, \qquad \forall\, x \in x_k + \delta_k \mathscr{B}_{\mathfrak{R}^n},$$

which implies x_k is a local minimizer of the function $\phi(x) - \langle \xi_k, x \rangle + \sigma_k \|x - x_k\|^2$. Therefore, $(x_k, g(x_k), h(g(x_k)))$ is a local minimizer of the function

$$r - \langle \xi_k, x \rangle + \sigma_k \|(x,y,r) - (x_k, g(x_k), h(g(x_k)))\|^2$$

on $G^{-1}(0) \cap D$. Observe that this function is Lipschitz with Lipschitz constant $K = 1 + \|\xi_k\| + 2\sigma_k > 0$. By Proposition 7.3 $(x_k, g(x_k), h(g(x_k)))$ is a local minimizer of the function

$$r - \langle \xi_k, x \rangle + \sigma_k \|(x,y,r) - (x_k, g(x_k), h(g(x_k)))\|^2 + Kd((x,y,r); G^{-1}(0) \cap D).$$

Because $\xi_k \to \xi$, there exists $M > 0$ such that $\|\xi_k\| \leqq M$. Also by assumption, H is metrically regular at $(x_0, g(x_0), h(g(x_0)), 0)$ with respect to $G^{-1}(0)$ which implies there exist some constant $\gamma \geqq 0$ such that $(x_k, g(x_k), h(g(x_k)))$ minimizes the function

$$\begin{aligned} &r - \langle \xi_k, x \rangle + \sigma_k \|(x,y,r) - (x_k, g(x_k), h(g(x_k)))\|^2 \\ &\quad + \gamma(d((x,y,r); D) + d((x,y,r); G^{-1}(0))). \end{aligned}$$

Observe that $G^{-1}(0) = grp\ g \times \mathfrak{R}$. Also by hypothesis, $D = \mathfrak{R}^n \times epi\ h$. Thus $(x_k, g(x_k), h(g(x_k)))$ is a local minimizer of the function

$$\begin{aligned} &r - \langle \xi_k, x \rangle + \sigma_k \|(x,y,r) - (x_k, g(x_k), h(g(x_k)))\|^2 \\ &\quad + \gamma(d((y,r); epi\ h) + d((x,y); grp\ g)). \end{aligned}$$

Therefore,

$$\begin{aligned} \gamma(d((y,r); epi\ h) + d((x,y); grp\ g)) \geqq\ & \langle (\xi_k, 0, -1), (x,y,r) - (x_k, g(x_k), h(g(x_k))) \rangle \\ & - \sigma_k \|(x,y,r) - (x_k, g(x_k), h(g(x_k)))\|^2 \end{aligned}$$

for every $(x,y,r) \in (x_k, g(x_k), h(g(x_k))) + \delta_k \mathscr{B}_{\mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R}}$. By Definition 7.1,

$$(\xi_k, 0, -1) \in \partial_P(\gamma(d((g(x_k), h(g(x_k))); epi\ h) + d((x_k, g(x_k)); grp\ g))).$$

Applying Definition 7.2,

$$\begin{aligned}(\xi,0,-1) \in & \partial_L(\gamma(d((y_0,r_0);epi\ h)+d((x_0,y_0);grp\ g))) \\ & \subseteq \{0\}\times N_L((y_0,r_0);epi\ h)+N_L((x_0,y_0);grp\ g)\times\{0\}.\end{aligned}$$

This containment implies there exist $(\mu_1,\tau_1)\in N_L((y_0,r_0);epi\ h)$ and $(\xi_2,\mu_2)\in N_L((x_0,y_0);grp\ g)$ such that

$$(\xi,0,-1)=(0,\mu_1,\tau_1)+(\xi_2,\mu_2,0)$$

thereby leading to $(\mu,-1)\in N_L((y_0,r_0);epi\ h)$ and $(\xi,-\mu)\in N_L((x_0,y_0);grp\ g)$ which, by relation between the subdifferential and normal cone along with Definition 7.5, implies

$$\mu\in\partial_L h(g(x_0)) \quad \text{and} \quad \xi\in D^*g(x_0)(\mu)=\partial_L(\mu g)(x_0).$$

Hence proving the requisite result. □

Using this result, we can obtain the optimality conditions for (P) in terms of the basic subdifferentials and basic normal cone but for that we need the following corollary which can be worked out on the lines of Corollary 3.4 in [9] using Definition 7.5.

Corollary 7.2. *Let $x_0\in F=g^{-1}(D)\cap C$. Assume that the multivalued function $G(x)=-g(x)+D$ is metrically regular at $(x_0,0)$ with respect to C. Then*

$$\partial_L\delta_F(x_0)\subseteq\bigcup_{\mu\in N_L(g(x_0);D)}\partial_L(\mu g)(x_0)+N_L(x_0;C).$$

Next we present the KKT optimality conditions for (P) using the notion of metric regularity.

Theorem 7.6. *Let $x_0\in\mathfrak{R}^n$ be an optimal solution of (P). Assume that (P) is metrically regular at x_0 with respect to C and D. Then there exist $\mu\in N_L(g(x_0);D)$ such that*

$$0\in\partial_L f(x_0)+\partial_L(\mu g)(x_0)+N_L(x_0;C).$$

Proof. Observe that x_0 is an optimal solution of the problem

$$\inf_{x\in\mathfrak{R}^n}(f+K_f d_F)(x),$$

where $K_f>0$ is the Lipschitz constant of f. Then by the optimality condition

$$\begin{aligned}0 &\in \partial_L(f+K_f d_F)(x_0) \\ &\subseteq \partial_L f(x_0)+N_L(x_0;F)=\partial_L f(x_0)+\partial_L\delta_F(x_0),\end{aligned}$$

where the feasible set $F = g^{-1}(D) \cap C$. Therefore by Corollary 7.2, there exist $\mu \in \partial_L \delta_D(g(x_0)) = N_L(g(x_0); D)$ such that

$$0 \in \partial_L f(x_0) + \partial_L(\mu g)(x_0) + N_L(x_0; C)$$

as desired. □

We end this section by presenting an example.

Example 7.2. Consider the following optimization problem

$$\min x$$
$$\text{subject to } |x| \leqq 0.$$

Here the necessary optimality condition holds at the point $x_0 = 0$. Observe that the notion of metric regularity is satisfied at x_0 as shown in Example 7.1. Now if the constraint $|x| \leqq 0$ is replaced by $x^2 \leqq 0$, $x_0 = 0$ remains the optimal point. However, in this case the necessary optimality condition does not hold inasmuch as metrical regularity fails at x_0.

7.5 Conclusions

In this work we have attempted to present the optimality conditions from a different perspective wherein the concept of metric regularity acts as the main tool. As seen, under the assumption of metric regularity, one can obtain the KKT optimality conditions for optimization problems without going into the intricacy of first establishing the Fritz–John optimality conditions. For this aim, herein problem (P) is converted into an unconstrained problem using the distance function and then the resultant unconstrained problem is dealt with directly. We conclude the chapter with an important observation. As discussed in Section 7.3, there are various constraint qualifications implying metric regularity under different scenarios. It has been observed that *ACQ* is the most general constraint qualification under a convex differentiability setting, and in the case where the constraints are in the form of inequalities, *ACQ* is equivalent to metric regularity [11]. Also *RCQ* is related to metric regularity under differentiability. In our study we have concentrated only on the relationship between the $(\widetilde{BCQ})$ and metric regularity. It would indeed be interesting to look into exploring the relationships of the other constraint qualifications with metric regularity under a more generalized subdifferentiability setting.

References

1. Borwein, J.M.: Stability and regularity points of inequality systems. *J. Optim. Theory Appl.* **48**, 9–52 (1986).

2. Clarke, F.H.: *Optimization and Nonsmooth Analysis.* Wiley-Interscience, New York (1983).
3. Clake, F.H., Ledyaev, Y.S., Stern, R.J., Wolenski, P.R.: Nonsmooth analysis and control theory. In: Axler, S., Gehring, F.W., Ribet, K.A. (eds.) *Graduate Texts in Mathematics*, **178**. Springer, New York (1998).
4. Cominetti, R.: Metric regularity, tangent sets and second-order optimality conditions. *Appl. Math. Optim.* **21**, 265–287 (1990).
5. Hiriart-Urruty, J.-B., Lemarechal, C.: *Convex Analysis and Minimization Algorithms (I).* Springer-Verlag, New York (1993).
6. Jourani, A.: Metric regularity and second-order necessary optimality conditions for minimization problems under inclusion constraints. *J. Optim. Theory Appl.* **81**, 97–120 (1994).
7. Jourani, A.: Constraint qualifications and Lagrange multipliers in nondifferentiable programming problems. *J. Optim. Theory Appl.* **81**, 533–548 (1994).
8. Jourani, A., Thibault, L.: Approximate subdifferential and metric regularity: the finite dimensional case. *Math. Program.* **47**, 203–218 (1990).
9. Jourani, A., Thibault, L.: Metric regularity and subdifferential calculus in Banach spaces. *Set-Valued Anal.* **3**, 87–100 (1995).
10. Jourani, A., Thibault, L.: Metric regularity for strongly compactly Lipschitzian mappings. *Nonlinear Anal.* **24**, 229–240 (1995).
11. Li, W.: Abadie's constraint qualification, metric regularity, and error bounds for differentiable convex inequalities. *SIAM J. Optim.* **7**, 966–978 (1997).
12. Mordukhovich, B.S.: Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions. *Trans. Amer. Math. Soc.* **340**, 1–35 (1993).
13. Mordukhovich, B.S.: *Variational Analysis and Generalized Differentiation. I: Basic Theory*. Springer-Verlag, Berlin (2006).
14. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton, NJ (1970).
15. Rockafellar, R.T., Wets, R.J-B.: *Variational Analysis.* Springer-Verlag, Berlin (1998).
16. Zheng, X.Y., Ng, K.F.: Metric regularity and constraint qualifications for convex inequalities on Banach spaces. *SIAM J. Optim.* **14**, 757–772 (2004).

Chapter 8
An Application of the Modified Subgradient Method for Solving Fuzzy Linear Fractional Programming Problem

Pankaj Gupta and Mukesh Kumar Mehlawat

Abstract We present an application of the "modified subgradient method" to solve a fuzzy linear fractional programming problem. We concentrate on a linear fractional programming problem in which both the right-hand side and the technological coefficients are fuzzy numbers. We compare efficiency of the proposed solution method with the well-known "fuzzy decisive set method" in terms of the number of iterations taken to reach the optimal solution. A numerical illustration is provided for the purpose.

8.1 Introduction

Mathematical programming is used extensively in facilitating managerial decisions in a large number of domains. An important class of mathematical programming problems is fractional programming which has attracted the attention of many researchers in the past. For the state of the art in the theory, methods, and applications in fractional programming, we refer the reader to Stancu-Minasian [12]. The main reason for interest in fractional programming stems from the fact that mathematical programming models could better fit real-world problems if we consider optimization of the ratio between the physical and/or economic quantities.

In classical problems of mathematical programming generally, and in linear fractional programming in particular, the coefficients of the problem are assumed to be exactly known. However, in practice this assumption is seldom satisfied by the great majority of real-life problems. The modeling of input data inaccuracy can be made by means of fuzzy set theory. The concept of decision making in a fuzzy

Pankaj Gupta
Department of Operational Research, University of Delhi, Delhi, India,e-mail: pgupta@or.du.ac.in

Mukesh Kumar Mehlawat
Department of Operational Research, University of Delhi, Delhi, India,
e-mail: mukesh0980@yahoo.com

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_8,

environment was first proposed by Bellman and Zadeh [1]. Subsequently, Tanaka, Okuda, and Asai [15] made use of this concept in mathematical programming. The use of fuzzy set theory concepts in fractional programming has been discussed, for example, in the works of Chakraborty and Gupta [2], Dutta, Rao, and Tiwari [3], and Stancu-Minasian and Pop [13].

Li and Chen [5] presented a fuzzy linear fractional programming model with fuzzy coefficients. Sakawa and Kato [8] formulated multiple objective linear fractional programming problems (MOLFPP) with block angular structure involving fuzzy numbers and through the introduction of extended Pareto optimality concepts, an interactive fuzzy satisfying method was presented. In another paper, Sakawa and Kato [9] presented an interactive satisfying method for structured MOLFPP with fuzzy numbers in the objective functions and in the constraints. The authors changed the fuzzy problem into a deterministic one using fuzzy intervals. Sakawa and Nishizaki [10] presented an interactive method for solving two-level linear fractional programming problems with fuzzy parameters. Sakawa, Nishizaki, and Hitaka [11] developed an interactive fuzzy programming method for multilevel 0–1 programming problems with fuzzy parameters through genetic algorithms. Mehra, Chandra, and Bector [6] proposed the concepts of an (α,β)-acceptable optimal solution and (α,β)-acceptable optimal value of a fuzzy linear fractional programming problem with fuzzy coefficients, and developed a method to compute them. Recently, Pop and Stancu-Minasian [14] proposed a method of solving the fully fuzzified linear fractional programming problems, where all the parameters and variables are triangular fuzzy numbers.

The purpose of this chapter is to provide an efficient solution methodology to solve a linear fractional programming problem (LFPP) in which both the right-hand side and the technological coefficients are fuzzy numbers with linear membership functions. The fuzzy problem is first converted into an equivalent crisp problem which is a max–min problem of finding a solution that satisfies the constraints and the goal with the maximum degree. The idea for this approach is due to Bellman and Zadeh [1]. The crisp problem, obtained in such a manner, can be nonlinear (even nonconvex), where the nonlinearity arises in constraints. We present an application of the “modified subgradient method” [4] by proposing a similar solution technique for the LFPP. We compare the proposed solution methodology with the fuzzy decisive set method [7]. With the help of a numerical illustration, we show that the proposed method is more effective from the point of view of the number of iterations required for obtaining the desired compromise solution.

This chapter is organized as follows. In Section 8.2, we discuss the LFPP in which both the right-hand side and the technological coefficients are fuzzy numbers. Application of modified subgradient method to fuzzy linear fractional programming problem (FLFPP) is presented in Section 8.3. In Section 8.4, the modified subgradient method and fuzzy decisive set method are used to solve numerical illustrations. This section also contains a discussion on the effectiveness of the proposed method over the fuzzy decisive set method. Finally in Section 8.5, we submit our concluding remarks.

8.2 LFPP with Fuzzy Technological Coefficients and Fuzzy Right-Hand Side Numbers

In this section, we consider the LFPP with fuzzy technological coefficients and fuzzy right-hand side numbers.

$$\text{(P)}\quad \max\ z = \frac{\sum_{j=1}^{n} c_j x_j + c_0}{\sum_{j=1}^{n} d_j x_j + d_0}$$

$$\text{subject to}\quad \sum_{j=1}^{n} \widetilde{a}_{ij} x_j \le \widetilde{b}_i, \qquad 1 \le i \le m,$$

$$x_j \ge 0, \qquad 1 \le j \le n,$$

where $c_0, (c_j)_{j=1,2,\ldots,n}$ and $d_0, (d_j)_{j=1,2,\ldots,n}$ are the coefficients of the linear fractional objective function.

$$(\widetilde{a}_{ij})_{i=1,2,\ldots,m}^{j=1,2,\ldots,n} \quad \text{and} \quad (\widetilde{b}_i)_{i=1,2,\ldots,m}$$

are the technological coefficients and the right-hand side of the linear constraints, respectively. $(x_j)_{j=1,2,\ldots,n}$ are the decision variables and it is required that at least one $x_j > 0$. Furthermore, it is assumed that the denominator of the objective function in (P) is strictly positive for any x_j in the feasible region.

Let us define fuzzy numbers $\widetilde{a}_{ij}$ and $\widetilde{b}_i$ with the following linear membership functions.

$$\mu_{\widetilde{a}_{ij}}(x) = \begin{cases} 1 & \text{if } x < a_{ij}, \\ (a_{ij} + d_{ij} - x)/d_{ij} & \text{if } a_{ij} \le x < a_{ij} + d_{ij}, \\ 0 & \text{if } x \ge a_{ij} + d_{ij}, \end{cases}$$

and

$$\mu_{\widetilde{b}_i}(x) = \begin{cases} 1 & \text{if } x < b_i, \\ (b_i + p_i - x)/p_i & \text{if } b_i \le x < b_i + p_i, \\ 0 & \text{if } x \ge b_i + p_i, \end{cases}$$

where $x \in R$, $d_{ij} > 0$ for $i = 1,2,\ldots,m; j = 1,2,\ldots,n$, and $p_i > 0$ for $i = 1,2,\ldots,m$. For defuzzification of the problem (P), we first fuzzify its objective function. For this, we first calculate the lower (z_l) and upper (z_u) bounds of the optimal objective values. The optimal values z_l and z_u can be defined by solving the following standard linear fractional programming problems.

$$\text{(P1)}\;\; z_1 = \max \frac{\sum_{j=1}^{n} c_j x_j + c_0}{\sum_{j=1}^{n} d_j x_j + d_0}$$

$$\text{subject to} \quad \sum_{j=1}^{n} (a_{ij} + d_{ij}) x_j \leq b_i, \qquad 1 \leq i \leq m,$$

$$x_j \geq 0, \qquad 1 \leq j \leq n,$$

$$\text{(P2)}\;\; z_2 = \max \frac{\sum_{j=1}^{n} c_j x_j + c_0}{\sum_{j=1}^{n} d_j x_j + d_0}$$

$$\text{subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \leq b_i + p_i, \qquad 1 \leq i \leq m,$$

$$x_j \geq 0, \qquad 1 \leq j \leq n,$$

$$\text{(P3)}\;\; z_3 = \max \frac{\sum_{j=1}^{n} c_j x_j + c_0}{\sum_{j=1}^{n} d_j x_j + d_0}$$

$$\text{subject to} \quad \sum_{j=1}^{n} (a_{ij} + d_{ij}) x_j \leq b_i + p_i, \qquad 1 \leq i \leq m,$$

$$x_j \geq 0, \qquad 1 \leq j \leq n,$$

and

$$\text{(P4)}\;\; z_4 = \max \frac{\sum_{j=1}^{n} c_j x_j + c_0}{\sum_{j=1}^{n} d_j x_j + d_0}$$

$$\text{subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \qquad 1 \leq i \leq m,$$

$$x_j \geq 0, \qquad 1 \leq j \leq n.$$

Let $z_l = \min(z_1, z_2, z_3, z_4)$ and $z_u = \max(z_1, z_2, z_3, z_4)$. The objective function of the problem (P) takes values between z_l and z_u whereas technological coefficients take values between a_{ij} and $a_{ij} + d_{ij}$ and the right-hand side numbers take values between b_i and $b_i + p_i$. Here, it is assumed that the linear fractional crisp problems P1, P2, P3, and P4 have finite optimal values.

Then, the fuzzy set of optimal values, G, which is a subset of R^n is defined as

$$\mu_G(x) = \begin{cases} 0 & \text{if } \dfrac{\sum_{j=1}^{n} c_j x_j + c_0}{\sum_{j=1}^{n} d_j x_j + d_0} < z_l, \\ \left(\dfrac{\sum_{j=1}^{n} c_j x_j + c_0}{\sum_{j=1}^{n} d_j x_j + d_0} - z_l \right) \Big/ (z_u - z_l) & \text{if } z_l \le \dfrac{\sum_{j=1}^{n} c_j x_j + c_0}{\sum_{j=1}^{n} d_j x_j + d_0} < z_u, \\ 1 & \text{if } \dfrac{\sum_{j=1}^{n} c_j x_j + c_0}{\sum_{j=1}^{n} d_j x_j + d_0} \ge z_u. \end{cases} \tag{8.1}$$

The fuzzy set of the ith constraint, C_i, which is a subset of R^m, is defined as

$$\mu_{C_i}(x) = \begin{cases} 0 & \text{if } b_i < \sum_{j=1}^{n} a_{ij} x_j, \\ \left(b_i - \sum_{j=1}^{n} a_{ij} x_j \right) \Big/ \left(\sum_{j=1}^{n} d_{ij} x_j + p_i \right) & \text{if } \sum_{j=1}^{n} a_{ij} x_j \le b_i \\ & < \sum_{j=1}^{n} (a_{ij} + d_{ij}) x_j + p_i, \\ 1 & \text{if } b_i \ge \sum_{j=1}^{n} (a_{ij} + d_{ij}) x_j + p_i. \end{cases} \tag{8.2}$$

Using the definition of the fuzzy decision proposed by Bellman and Zadeh [1], we have

$$\mu_D(x) = \min(\mu_G(x), \min_i(\mu_{C_i}(x))).$$

The optimal fuzzy decision is a solution of the problem

$$\max(\mu_D(x)) = \max_{x \ge 0} \min(\mu_G(x), \min_i(\mu_{C_i}(x))).$$

Consequently, the problem (P) reduces to the following optimization problem.

$$\begin{aligned} \text{(P5)} \quad & \max \lambda \\ \text{subject to} \quad & \lambda \le \mu_G(x), \\ & \lambda \le \mu_{C_i}(x), \quad 1 \le i \le m, \\ & x_j \ge 0, \quad 1 \le j \le n, \\ & 0 \le \lambda \le 1. \end{aligned}$$

Using (8.1) and (8.2), the problem P5 can be written as

$$\text{(P6)} \quad \max \lambda$$

$$\begin{aligned}\text{subject to} \quad & \lambda (z_u - z_l) \left(\sum_{j=1}^{n} d_j x_j + d_0 \right) - \left(\sum_{j=1}^{n} c_j x_j + c_0 \right) \\ & + z_l \left(\sum_{j=1}^{n} d_j x_j + d_0 \right) \leq 0, \\ & \sum_{j=1}^{n} (a_{ij} + \lambda d_{ij}) x_j + \lambda p_i - b_i \leq 0, \qquad 1 \leq i \leq m, \\ & x_j \geq 0, \qquad 1 \leq j \leq n, \\ & 0 \leq \lambda \leq 1. \end{aligned}$$

It may be noted that in general the problem P6 is not convex. Therefore the solution of this problem requires the special approach adopted for solving general nonconvex optimization problems.

8.3 Modified Subgradient Method to Solve the FLFPP

In this section, we present an application of the modified subgradient method [4] to solve the FLFPP. The modified subgradient method can be applied for solving a large class of nonconvex and nonsmooth constrained optimization problems. This method is based on construction of the dual problem by using a sharp Lagrangian function and has many advantages [4]. Some of them are the following.

- The zero duality gap property is proved for a sufficiently large class of problems.
- The value of the dual function strictly increases at each iteration.
- The method does not use any penalty parameters.
- The presented method has a natural stopping criterion.

For details of the modified subgradient method, we refer the reader to [4].

We use the following notations.

- k is the number of iterations.
- (u^k, c^k) is a vector of Lagrange multipliers at the kth iteration.
- x^k is a minimizer of the Lagrange function $L(x, u^k, c^k)$.
- H is the upper bound for the values of the dual function.

For applying the subgradient method to the problem P6, we first formulate it with equality constraints by using slack variables P_0 and $P_i, i = 1, 2, \ldots, m$. Then, problem P6 can be written as

$$\text{(P7)}\quad \max \lambda = -\min(-\lambda)$$

$$\text{subject to}\quad g_0(x,\lambda,P_0) = \lambda(z_u - z_l)\left(\sum_{j=1}^{n} d_j x_j + d_0\right) - \left(\sum_{j=1}^{n} c_j x_j + c_0\right) + z_l\left(\sum_{j=1}^{n} d_j x_j + d_0\right) + P_0 = 0,$$

$$g_i(x,\lambda,P_i) = \sum_{j=1}^{n} (a_{ij} + \lambda d_{ij}) x_j + \lambda p_i - b_i + P_i = 0, \qquad 1 \le i \le m,$$

$$x \ge 0, \qquad P_0, P_i \ge 0, \qquad 1 \le i \le m,$$

$$0 \le \lambda \le 1.$$

For this problem, we define the following set

$$S = \{(x,\lambda,P) | x = (x_1,\ldots,x_n), P = (P_0,P_1,\ldots,P_m), \underline{x}_j \le x_j \le \overline{x}_j, P_0, P_i \ge 0, \lambda \in [0,1]\}.$$

Here, $\underline{x}_j = \min\{x_j^*, x_j^{**}, x_j^{***}, x_j^{****}\}$ and $\overline{x}_j = \max\{x_j^*, x_j^{**}, x_j^{***}, x_j^{****}\}$, where x_j^*, x_j^{**}, x_j^{***}, x_j^{****} are the optimal values of x_j for the problems $P1$, $P2$, $P3$, and $P4$, respectively. It may be pointed out that the lower and upper bounds on x_j are incorporated in order to achieve a faster convergence rate. Also, we take $g(x,\lambda,P) = (g_0, g_1, \ldots, g_m)$.

Consider the following equivalent form of the problem P7 as the primal problem:

$$(P7)\text{-1 Minimize } P' = \min_{(x,\lambda,P)\in S} -\lambda$$

$$\text{subject to}\quad g(x,\lambda,P) = 0.$$

For the problem $P7$-1, we introduce the following augmented Lagrangian function,

$$L(x,\lambda,P,u,c) = -\lambda + c\|g(x,\lambda,P)\| - u^T g(x,\lambda,P),$$

where $(x,\lambda,P) \in S, u \in R^{m+1}$ and $c \ge 0$.

The above-defined augmented Lagrangian is termed sharp Lagrangian because of the presence of the augmenting function $\|g(x,\lambda,P)\|$, where $\|\cdot\|$ is the norm function.

The augmented Lagrangian associated with the problem $P7$ can be written as

$$L(x,\lambda,P,u,c)$$

$$= -\lambda + c\left[\left(\lambda(z_u - z_l)\left(\sum_{j=1}^{n} d_j x_j + d_0\right) - \left(\sum_{j=1}^{n} c_j x_j + c_0\right)\right.\right.$$

$$+z_l\left(\sum_{j=1}^{n} d_j x_j + d_0\right) + P_0\Bigg)^2 + \sum_{i=1}^{m}\left(\sum_{j=1}^{n}(a_{ij}+\lambda d_{ij})x_j + \lambda p_i - b_i + P_i\right)^2\Bigg]^{1/2}$$

$$-u_0\left(\lambda(z_u - z_l)\left(\sum_{j=1}^{n} d_j x_j + d_0\right) - \left(\sum_{j=1}^{n} c_j x_j + c_0\right) + z_l\left(\sum_{j=1}^{n} d_j x_j + d_0\right) + P_0\right)$$

$$-\sum_{i=1}^{m} u_i\left(\sum_{j=1}^{n}(a_{ij}+\lambda d_{ij})x_j + \lambda p_i - b_i + P_i\right)$$

Let the following dual function be considered.

$$H(u,c) = \min_{(x,\lambda,P)\in S} L(x,\lambda,P,u,c)$$

for $u \in R^{m+1}$ and $c \geq 0$.

We now consider the following dual problem to the problem $P7$-1,

$$(DP7)-1 \ \text{ Maximize } \ P'' = \max_{(u\in R^{m+1}, c\geq 0)} H(u,c).$$

It can be shown on similar lines to those in [4] that under the assumptions of continuity of $g(x,\lambda,P)$ and that S is compact and a feasible solution exists to $P7$-1, we have Minimize $P' =$ Maximize P'' and there exists a solution to $DP7$-1.

Also, assuming that Minimize $P' =$ Maximize P'' and for some $\bar{u} \in R^{m+1}, \bar{c} \geq 0$,

$$\min_{(x,\lambda,P)\in S} L(x,\lambda,P,\bar{u},\bar{c}) = -\bar{\lambda} + \bar{c}\|g(\bar{x},\bar{\lambda},\bar{P})\| - \bar{u}^T g(\bar{x},\bar{\lambda},\bar{P}),$$

it can be shown on similar lines to those in [4] that $(\bar{x},\bar{\lambda},\bar{P})$ is a solution to $P7$-1/$P7$ and $(\bar{u},\bar{c})$ is a solution to $DP7$-1 if and only if

$$g(\bar{x},\bar{\lambda},\bar{P}) = 0.$$

The maximization of the dual function $H(u,c)$ by using the subgradient method gives us the optimal value of the primal problem $P7$-1/$P7$.

We now present the following algorithm that solves the dual problem using approximation of the necessary and sufficient condition $g(\bar{x},\bar{\lambda},\bar{P}) = 0$.

Initialization Step. Choose a vector (u_0^1, u_i^1, c^1) with $c^1 \geq 0$ and go to main step. Let $k = 1$.

Main Step.

Step 1. Given (u_0^k, u_i^k, c^k), solve the following subproblem.

$$\min\left\{-\lambda + c\left[\left(\lambda(z_u - z_l)\left(\sum_{j=1}^{n} d_j x_j + d_0\right) - \left(\sum_{j=1}^{n} c_j x_j + c_0\right)\right.\right.\right.$$

$$\left.\left. + z_l\left(\sum_{j=1}^{n} d_j x_j + d_0\right) + P_0\right)^2 + \sum_{i=1}^{m}\left(\sum_{j=1}^{n}(a_{ij} + \lambda d_{ij})x_j + \lambda p_i - b_i + P_i\right)^2\right]^{1/2}$$

$$-u_0\left(\lambda(z_u - z_l)\left(\sum_{j=1}^{n} d_j x_j + d_0\right) - \left(\sum_{j=1}^{n} c_j x_j + c_0\right) + z_l\left(\sum_{j=1}^{n} d_j x_j + d_0\right) + P_0\right)$$

$$\left. - \sum_{i=1}^{m} u_i\left(\sum_{j=1}^{n}(a_{ij} + \lambda d_{ij})x_j + \lambda p_i - b_i + P_i\right)\right\}$$

subject to $\quad (x, \lambda, P) \in S,\ u = (u_0, u_1, \ldots, u_m) \in R^{m+1}, \qquad c \geq 0.$

Let (x^k, λ^k, P^k) be any solution. If $\|g(x^k, \lambda^k, P^k)\|$ becomes sufficiently small, then stop. (u_0^k, u_i^k, c^k) is a solution to the dual problem and because of the zero gap property, (x^k, λ^k, P^k) is a solution to the problem P7. Hence, (x^k, λ^k) is a solution to the problem P6. Otherwise, go to step 2.

Step 2. Let

$$u_0^{k+1} = u_0^k - s^k\left(\lambda(z_u - z_l)\left(\sum_{j=1}^{n} d_j x_j + d_0\right) - \left(\sum_{j=1}^{n} c_j x_j + c_0\right)\right.$$

$$\left. + z_l\left(\sum_{j=1}^{n} d_j x_j + d_0\right) + P_0\right)$$

$$u_i^{k+1} = u_i^k - s^k\left(\sum_{j=1}^{n}(a_{ij} + \lambda d_{ij})x_j + \lambda p_i - b_i + P_i\right), \qquad 1 \leq i \leq m$$

$$c^{k+1} = c^k + (s^k + \varepsilon^k)\|g(x^k, \lambda^k, P^k)\|,$$

where s^k and ε^k are positive scalar stepsizes. Replace k by $k+1$ and repeat step 1.

The following stepsize formulas can be used for generating iterations of the modified subgradient method.

$$s^k = \frac{\alpha_k(H_k - H(u^k, c^k))}{5\|g(x^k, \lambda^k, P^k)\|^2}, \qquad \varepsilon^k = \beta s^k, \qquad \beta > 0,$$

where H_k is an approximation to the optimal dual value, $0 < \alpha_k < 2$, $0 < \varepsilon^k < s^k$.

It has been been proved in [4] that the new iterations of the modified subgradient method with sharp Lagrangian strictly improve the objective function value for all choices of the stepsizes s^k and ε^k within permissible limits.

8.4 Numerical Illustrations

We first briefly describe the fuzzy decisive set method. This method is based on the idea that for a fixed value of λ, the problem P6 is a linear programming problem. Obtaining the optimal solution λ^* to the problem P6 is equivalent to determining the maximum value of λ so that the feasible set is nonempty. The algorithm of this method for the problem P6 is presented below.

Algorithm.

Step 1. Set $\lambda = 1$ and test whether a feasible set satisfying the constraints of the problem P6 exists using phase one of the two-phase simplex method. If a feasible set exists, set $\lambda = 1$. Otherwise, set $\lambda^L = 0$ and $\lambda^R = 1$ and go to the next step.

Step 2. For the value of $\lambda = (\lambda^L + \lambda^R)/2$, update the values of λ^L and λ^R using the bisection method as follows.

$$\begin{aligned} \lambda^L &= \lambda \quad \text{if feasible set is nonempty for } \lambda. \\ \lambda^R &= \lambda \quad \text{if feasible set is empty for } \lambda. \end{aligned}$$

Consequently, for each λ, test whether a feasible set of the problem P6 exists and determine the maximum value λ^* satisfying the constraints of the problem P6.

Next, we present a numerical illustration.

Example. We consider the following linear fractional programming problem.

$$\begin{aligned} \text{(LFPP)} \quad & \max \frac{5x_1 + 3x_2}{2x_1 + x_2 + 3} \\ \text{subject to} \quad & \tilde{2}x_1 + \tilde{3}x_2 \leq \tilde{4} \\ & \tilde{1}x_1 + \tilde{2}x_2 \leq \tilde{3} \\ & x_1, x_2 \geq 0, \end{aligned}$$

which take fuzzy parameters. We consider

$$(a_{ij}) = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}, \qquad (d_{ij}) = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \Rightarrow (a_{ij} + d_{ij}) = \begin{bmatrix} 3 & 5 \\ 2 & 3 \end{bmatrix}$$

$$(b_i) = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \qquad (p_i) = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \Rightarrow (b_i + p_i) = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$$

For solving the above problem, we must solve the following four subproblems.

$$z_1 = \max \frac{5x_1 + 3x_2}{2x_1 + x_2 + 3}$$

$$\begin{aligned} \text{subject to} \quad 3x_1 + 5x_2 &\leq 4, \\ 2x_1 + 3x_2 &\leq 3, \\ x_1, x_2 &\geq 0, \end{aligned}$$

$$z_2 = \max \frac{5x_1 + 3x_2}{2x_1 + x_2 + 3}$$

$$\begin{aligned} \text{subject to} \quad 2x_1 + 3x_2 &\leq 7, \\ 1x_1 + 2x_2 &\leq 5, \\ x_1, x_2 &\geq 0, \end{aligned}$$

$$z_3 = \max \frac{5x_1 + 3x_2}{2x_1 + x_2 + 3}$$

$$\begin{aligned} \text{subject to} \quad 3x_1 + 5x_2 &\leq 7, \\ 2x_1 + 3x_2 &\leq 5, \\ x_1, x_2 &\geq 0, \end{aligned}$$

and

$$z_4 = \max \frac{5x_1 + 3x_2}{2x_1 + x_2 + 3}$$

$$\begin{aligned} \text{subject to} \quad 2x_1 + 3x_2 &\leq 4, \\ 1x_1 + 2x_2 &\leq 3, \\ x_1, x_2 &\geq 0. \end{aligned}$$

Optimal solutions of these subproblems are given in Table 8.1.

Table 8.1 Optimal solutions

	Objective Function Value	x_1	x_2
z_1	1.176471	1.33333	0
z_2	1.75	3.5	0
z_3	1.521739	2.33333	0
z_4	1.42857	2	0

By using these optimal values, the LFPP can be reduced to the following equivalent nonlinear programming problem.

$$\max\ \lambda$$
$$\text{subject to}\quad \lambda \le \frac{\dfrac{5x_1+3x_2}{2x_1+x_2+3}-1.176471}{1.75-1.176471},$$
$$\lambda \le \frac{4-2x_1-3x_2}{x_1+2x_2+3},$$
$$\lambda \le \frac{3-x_1-2x_2}{x_1+x_2+2},$$
$$0 \le \lambda \le 1,$$
$$x_1, x_2 \ge 0;$$

that is,

$$\text{(NLP)}\quad \max \lambda$$
$$\text{subject to}\quad (2.647058-1.147058\lambda)x_1+(1.823529-0.573529\lambda)x_2 \ge 3.529413+1.720587\lambda,$$
$$(2+\lambda)x_1+(3+2\lambda)x_2 \le 4-3\lambda,$$
$$(1+\lambda)x_1+(2+\lambda)x_2 \le 3-2\lambda,$$
$$0 \le \lambda \le 1,$$
$$x_1, x_2 \ge 0.$$

We first solve problem NLP by using the fuzzy decisive set method.
For $\lambda = 1$, the constraints of the problem NLP can be written as

$$1.5x_1+1.25x_2 \ge 5.25,$$
$$3x_1+5x_2 \le 1,$$
$$2x_1+3x_2 \le 1,$$
$$x_1,\ x_2 \ge 0.$$

Using phase-I, we obtain: $x_1 = 0.33333$, $x_2 = 0$. The above feasible set is empty, therefore by taking $\lambda^L = 0$ and $\lambda^R = 1$, the new value of $\lambda = (0+1)/2 = 1/2$ is used.

For $\lambda = 1/2 = 0.5$, the constraints of the problem NLP can be written as

$$2.073529x_1+1.5367645x_2 \ge 4.3897065,$$
$$2.5x_1+4x_2 \le 2.5,$$
$$1.5x_1+2.5x_2 \le 2,$$
$$x_1, x_2 \ge 0.$$

Using phase-I, we obtain: $x_1 = 1$, $x_2 = 0$. The above feasible set is again empty, therefore by taking $\lambda^L = 0$ and $\lambda^R = 1/2$, the new value of $\lambda = (0+1/2)/2 = 1/4$ is used.

For $\lambda = 1/4 = 0.25$, the constraints of the problem NLP can be written as

$$\begin{aligned} 2.3602935x_1 + 1.68014675x_2 &\geq 3.95955975, \\ 2.25x_1 + 3.5x_2 &\leq 3.25, \\ 1.25x_1 + 2.25x_2 &\leq 2.5, \\ x_1, x_2 &\geq 0, \end{aligned}$$

Using phase-I, we obtain: $x_1 = 1.44444$, $x_2 = 0$. The above feasible set is empty, therefore by taking $\lambda^L = 0$ and $\lambda^R = 1/4$, the new value of $\lambda = (0+1/4)/2 = 1/8$ is used.

For $\lambda = 1/8 = 0.125$, the constraints of the problem NLP can be written as

$$\begin{aligned} 2.50367575x_1 + 1.751837875x_2 &\geq 3.744486375, \\ 2.125x_1 + 3.25x_2 &\leq 3.625, \\ 1.125x_1 + 2.125x_2 &\leq 2.75, \\ x_1, x_2 &\geq 0. \end{aligned}$$

Using phase-I, we obtain: $x_1 = 1.705882$, $x_2 = 0$. The above feasible set is non-empty, therefore by taking $\lambda^L = 1/8$ and $\lambda^R = 1/4$, the new value of $\lambda = (1/8 + 1/4)/2 = 3/16$ is used.

The values of λ obtained in the next 22 iterations are presented in Table 8.2.

Table 8.2 Values of λ

Iterations	λ	Iterations	λ
4	0.18375	15	0.184020996
5	0.15625	16	0.183990478
6	0.171875	17	0.183975219
7	0.1796875	18	0.183982849
8	0.18359375	19	0.183979034
9	0.18359375	20	0.183980941
10	0.185546875	21	0.183981895
11	0.184570312	22	0.18398237
12	0.184082031	23	0.183982133
13	0.18383789	24	0.183982014
14	0.18395996	25	0.183982014

Consequently, we obtain the optimal value of λ at the 25th iteration by using the fuzzy decisive set method and the optimal solution is $x_1^* = 1.578792$, $x_2^* = 0$.

Now, we demonstrate the proposed solution method on the same problem. The stopping criteria used is taken as $\|g(x^k, \lambda^k, P^k)\| \leq 10^{-5}$ and $\varepsilon^k = 0.95s^k$.

We consider the following form of the problem NLP.

$$\begin{aligned}
&\max \lambda = -\min(-\lambda)\\
&\qquad 1.72058\lambda - (2.647058 - 1.147058\lambda)x_1 - (1.823529 - 0.573529\lambda)x_2\\
&\qquad\qquad + 3.529413 + P_0 = 0,\\
\text{subject to}\quad &(2+\lambda)x_1 + (3+2\lambda)x_2 + 3\lambda - 4 + P_1 = 0,\\
&(1+\lambda)x_1 + (2+\lambda)x_2 + 2\lambda - 3 + P_2 = 0,\\
&0 \le \lambda \le 1,\\
&x_1, x_2, P_0, P_1, P_2 \ge 0,
\end{aligned}$$

where P_0, P_1, and P_2 are slack variables.

Using the proposed method, the following optimal solution of the NLP is obtained at the end of second iteration, $x_1^* = 1.578792, x_2^* = 0$, and $\lambda^* = 0.183982$. A summary of the computational results is presented in Table 8.3.

Table 8.3 Summary of the Computational Results

k	u_0^k	u_1^k	u_2^k	c^k
1	0	0	0	0
2	−0.052370468	−0.048341839	−0.026856559	0.148518765

x_1^k	x_2^k	H	$H(u^k, c^k)$	$\|g(x^k, \lambda^k, P^k)\|$	s^k
1.33333	0	0.2	−1	4.72654023	0.016114
1.578792	0	0.2	−0.1839821	3.7×10^{-7}	

For the sake of completeness, we also present the details of the computational procedure.

The augmented Lagrangian function for the problem NLP is written as

$$\begin{aligned}
L(x,\lambda,P,u,c) = &-\lambda + c[(1.72058\lambda - (2.647058 - 1.147058\lambda)x_1\\
&- (1.823529 - 0.573529\lambda)x_2 + 3.529413 + P_0)^2\\
&+ ((2+\lambda)x_1 + (3+2\lambda)x_2 + 3\lambda - 4 + P_1)^2\\
&+ ((1+\lambda)x_1 + (2+\lambda)x_2 + 2\lambda - 3 + P_2)^2]^{1/2}\\
&- u_0(1.72058\lambda - (2.647058 - 1.147058\lambda)x_1\\
&- (1.823529 - 0.573529\lambda)x_2 + 3.529413 + P_0)\\
&- u_1((2+\lambda)x_1 + (3+2\lambda)x_2 + 3\lambda - 4 + P_1)\\
&- u_2((1+\lambda)x_1 + (2+\lambda)x_2 + 2\lambda - 3 + P_2).
\end{aligned}$$

Consider the initial vector $(u_0^1, u_1^1, u_2^1, c^1) = (0,0,0,0)$ and solve the following subproblem using LINGO software.

$$\min L(x,\lambda,P,0,0) = -\lambda$$
$$\text{subject to} \quad 0 \le \lambda \le 1,$$
$$1.33333 \le x_1 \le 3.5,$$
$$0 \le x_2 \le 0.$$

The optimal solution of the above subproblem is obtained as

$$x_1 = 1.33333, \qquad x_2 = 0, \qquad \lambda = 1, \qquad P_0 = P_1 = P_2 = 0.$$

At this solution

$$g_1(x^1,\lambda^1,P^1) = 3.249998, \qquad g_2(x^1,\lambda^1,P^1) = 2.99999,$$
$$g_3(x^1,\lambda^1,P^1) = 1.66666,$$

and hence,

$$\|g(x^1,\lambda^1,P^1)\| = ((3.249998)^2 + (2.99999)^2 + (1.66666)^2)^{1/2} = 4.72654023.$$

Because $\|g(x^1,\lambda^1,P^1)\| \nleq 10^{-5}$, we calculate new values of the Lagrange multipliers $(u_0^2, u_1^2, u_2^2, c^2)$ by using Step 2 of the proposed method.
Let

$$\alpha_1 = 1.5, \qquad H(u^1,c^1) = -1, \qquad H_1 = 0.2.$$

Then

$$5\|g(x^1,\lambda^1,P^1)\|^2 = 5 \times 22.3401825 = 111.7009127,$$
$$s^1 = \frac{\alpha_1(H_1 - H(u^1,c^1))}{5\|g(x^1,\lambda^1,P^1)\|^2} = \frac{1.5(0.2+1)}{111.7009127} = 0.016114.$$

Hence,

$$u_0^2 = 0 - (0.016114)(3.249998) = -0.052370468,$$
$$u_1^2 = 0 - (0.016114)(2.99999) = -0.048341839,$$
$$u_2^2 = 0 - (0.016114)(1.66666) = -0.026856559.$$

Also,

$$\varepsilon^1 = 0.0153083,$$
$$c^2 = 0 + (0.0314223)(4.72654023) = 0.148518765.$$

The solution at second iteration is obtained as

$$x_1 = 1.578792, \qquad x_2 = 0, \qquad \lambda^* = 0.1839821,$$
$$P_0 = P_1 = 0, \qquad P_2 = 0.7627742.$$

At this solution

$$g_1(x^2,\lambda^2,P^2) = 2.542575 \times 10^{-7}, \qquad g_2(x^2,\lambda^2,P^2) = -2.32377 \times 10^{-7},$$
$$g_3(x^2,\lambda^2,P^2) = -1.32377 \times 10^{-7},$$

and hence,

$$\begin{aligned}&\|g(x^2,\lambda^2,P^2)\|\\ &\quad = ((2.542575 \times 10^{-7})^2 + (-2.32377 \times 10^{-7})^2 + (-1.32377 \times 10^{-7})^2)^{1/2}\\ &\quad = 3.7 \times 10^{-7}.\end{aligned}$$

Because $\|g(x^2,\lambda^2,P^2)\| < 10^{-5}$, the solution $x_1^* = 1.578792$, $x_2^* = 0$, and $\lambda^* = 0.1839821$ is the optimal solution to the problem NLP. This means that, the vector (x_1^*, x_2^*) is a solution to the problem LFPP which has the best membership grade λ^*.

It may be noted that the optimal value of λ obtained at the second iteration of the modified subgradient method is approximately equal to the optimal value of λ calculated at the 25th iteration of the fuzzy decisive set method. Hence, the proposed solution method that can solve nonconvex optimization problems with binary restrictions can be considered as an efficient solution procedure for the problem under consideration.

8.5 Concluding Remarks

We have studied a fuzzy linear fractional programming problem in which both the right-hand side and the technological coefficients are fuzzy numbers. Bellman and Zadeh's approach has been used for converting the fuzzy linear fractional programming problem into an equivalent crisp problem. The constraints in problem P6 are generally not convex, thus the problem may be solved either by the fuzzy decisive set method, which is presented by Sakawa and Yana [7], or by using some linearization methods. There are some disadvantages in using these methods. The fuzzy decisive set method takes a long time to solve the problem. On the other hand, the linearization methods increase the number of the constraints. Here, we have presented the modified subgradient method and used it for solving the defuzzified problem P6. This method is based on the duality theory using the augmented Lagrangian function termed the sharp Lagrangian. An illustrative numerical example has been solved to demonstrate the proposed method and also to compare the effectiveness of the proposed method with that of the fuzzy decisive set method from the point of view of the number of iterations required for obtaining the desired optimal solution.

The present solution approach can also be extended to develop applications of the modified subgradient method for solving fuzzy linear fractional programming problems with fuzzy coefficients and for solving fully fuzzified linear fractional programming problems.

References

1. Bellman, R., Zadeh, L. A.: Decision making in a fuzzy environment. *Manage. Sci.* 17, 141–164 (1970).
2. Chakraborty, M., Gupta, S.: Fuzzy mathematical programming for multi-objective linear fractional programming problem. *Fuzzy Sets Syst.* 125, 335–342 (2002).
3. Dutta, D., Rao, J. R., Tiwari, R. N.: Fuzzy approaches for multiple criteria linear fractional optimization: a comment. *Fuzzy Sets Syst.* 54, 347–349 (1993).
4. Gasimov, R. N.: Augmented Lagrangian duality and nondifferentiable optimization methods in nonconvex programming. *J. Global Optim.* 24, 187–203 (2002).
5. Li, D. F., Chen, S.: A fuzzy programming approach to fuzzy linear fractional programming with fuzzy coefficients. *J. Fuzzy Math.* 4, 829–834 (1996).
6. Mehra, A., Chandra, S., Bector, C. R.: Acceptable optimality in linear fractional programming with fuzzy coefficients. *Fuzzy Optim. Decision Making* 6, 5–16 (2007).
7. Sakawa, M., Yana, H.: Interactive decision making for multi-objective linear fractional programming problems with fuzzy parameters. *Cybern. Syst.* 16, 377–397 (1985).
8. Sakawa, M., Kato, K.: Interactive decision-making for multiobjective linear fractional programming problems with block angular structure involving fuzzy numbers. *Fuzzy Sets Syst.* 97, 19–31 (1998).
9. Sakawa, M., Kato, K.: An interactive fuzzy satisficing method for structured multiobjective linear fractional programs with fuzzy numbers. *European J. Oper. Res.* 107, 575–589 (1998).
10. Sakawa, M., Nishizaki, I.: Interactive fuzzy programming for cooperate two-level linear fractional programming problems with multiple decision makers. *Fuzzy Sets Syst.* 115, 93–103 (2000).
11. Sakawa, M., Nishizaki, I., Hitaka, M.: Interactive fuzzy programming for multi-level 0-1 programming problems with fuzzy parameters through genetic algorithms. *Fuzzy Sets Syst.* 117, 95–111 (2001).
12. Stancu-Minasian, I. M.: A sixth bibliography of fractional programming. *Optimization* 55, 405–428 (2006).
13. Stancu-Minasian, I. M., Pop, B.: On a fuzzy set approach to solving multiple objective linear fractional programming problem. *Fuzzy Sets Syst.* 134, 397–405 (2003).
14. Pop, B., Stancu-Minasian, I. M.: A method of solving fully fuzzified linear fractional programming problems. *J. Appl. Math. Comput.* 27, 227–242 (2008).
15. Tanaka, H., Okuda, T., Asai, K.: On fuzzy mathematical programming. *J. Cybern.* 3, 37–46 (1984).

Chapter 9
On Sufficient Optimality Conditions for Semi-Infinite Discrete Minmax Fractional Programming Problems Under Generalized V-Invexity

S. K. Mishra, Kin Keung Lai, Sy-Ming Guu, and Kalpana Shukla

Abstract In this chapter, we establish many global nonparametric sufficient optimality conditions under various generalized $V-\rho$-invexity assumptions for a minmax fractional programming problem with infinitely many nonlinear inequality and equality constraints.

9.1 Introduction

Consider the following semi-infinite discrete minmax fractional programming problem.

$$\text{Minimize} \quad \max_{1 \leq i \leq p} \frac{f_i(x)}{g_i(x)} \tag{P}$$

S. K. Mishra
Department of Mathematics, Faculty of Science, Banaras Hindu University, Varanasi 221 005, India, e-mail: shashikant.dr@gmail.com

Kin Keung Lai
Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, e-mail: mskklai@cityu.edu.hk

Sy-Ming Guu
College of Management, Yuan-Ze University, 135 Yuan-Tung Road, Chung-Li, Taoyuan, 32003, Taiwan, e-mail: iesmguu@saturn.yzu.edu.tw

Kalpana Shukla
Department of Mathematics, Faculty of Science, Banaras Hindu University, Varanasi 221 005, India, e-mail: bhukalpanabhu@gmail.com

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_9,

subject to

$$G_j(x,t) \leq 0,\ \ \forall\, t \in T_j,\ j \in \underline{q}$$

$$H_k(x,s) = 0,\ \ \forall\, s \in S_k,\ k \in \underline{r}$$

$$x \in X,$$

where p, q and r are positive integers, X is an open convex subset of R^n (n-dimensional Euclidean space), for each $j \in \underline{q} \equiv \{1,\ldots,q\}$ and $k \in \underline{r}$, T_j and S_k are compact subset of complete metric spaces, for each $i \in \underline{p}$, f_i and g_i are real-valued functions defined on X, for each $j \in \underline{q}$, $z \to G_j(z,t)$ is a real-valued function defined on X for all $t \in T_j$, for each $k \in \underline{r}$, $z \to H_k(z,s)$ is a real-valued function defined on X for all $s \in S_k$, for each $j \in \underline{q}$, $t \to G_j(x,t)$ is a continuous real-valued function defined on T_j, for each $k \in \underline{r}$, $s \to H_k(x,s)$ is a continuous real-valued function defined on S_k, $\forall\, x \in X$, and for each $i \in \underline{p}$, $g_i(x) > 0$ for all x satisfying the constraints of (P).

Nonlinear programming problems such as (P) with a finite number of constraints (i.e., when G_j are independent of t, and the functions H_k are independent of s) are well known in the literature of mathematical programming problems as generalized fractional programming problems or minmax fractional programming problems. Minmax fractional programming problems are important, as they contain standard nonlinear programming problems and fractional programming problems as special cases. They are also important for their modeling capabilities (these problems have been used in economics, financial planning, multiobjective decision theory, and facility location problems) where these problems have been the subject for the last three decades. See, for example, [10].

A mathematical programming problem with a finite number of variables and infinitely many constraints is called a semi-infinite programming problem. Problems of this kind have been utilized for the modeling and analysis of many theoretical and concrete real-world practical problems. Semi-infinite programming problems have found relevance in engineering design (design of earthquake-resistant structures, design of control systems, digital filters, and electronic circuits; see [14]), boundary value problems, defect minimization for operator equations, geometry, random graphs, graphs related to Newton flows, wavelet analysis, semidefinite programming, geometric programming, and optimal control problems, the reader may see [14]. For details on semi-infinite programming and its applications, optimality conditions, duality theorems, and numerical algorithms, see [4, 5, 7, 8, 12, 13].

At this point, it is important to note that duality theory and generalized convexity (see [10, 11]) which have been playing significant roles in optimization theory are missing in the area of semi-infinite programming. In fact, at present very few papers are present in the literature dealing with semi-infinite programming problems and any class of generalized convexity. See, for example, [14].

In this chapter, we discuss a number of global nonparametric sufficient optimality conditions for (P) under a variety of generalized $V-\rho$-invexity assumptions. In Section 9.2, we present necessary definitions and preliminaries required in the chapter. We recall the necessary optimality conditions for (P) due to [14]. In Section 9.3, we give our results on global nonparametric sufficient optimality conditions for (P) under a variety of generalized $V-\rho$-invexity assumptions. This work extends some earlier works of [14] to $V-\rho$-invex functions.

9.2 Preliminaries

Definition 9.1. Let f be a real-valued differentiable function defined on X. Then f is said to be *invex at y* if there exists a function $\eta : X \times X \to R^n$ such that for each $x \in X$,

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

where

$$\nabla f = \left(\frac{\partial f(y)}{\partial y_1}, \ldots, \frac{\partial f(y)}{\partial y_n} \right)$$

is the gradient of f at y, and $\langle a,\ b \rangle$ denotes the inner product of the vectors a and b; f is said to be η-invex on X if the above inequality holds for all $x, y \in X$.

This generalization of the concept of convexity was given by [6]; the term invex (for invariant convex) was coined by [2]. For details on invex functions and their applications in optimization, see [10].

However, the major difficulty is that the invex problems require the same kernel function $\eta(x,\ u)$ for the objective and the constraints. This requirement turns out to be a severe restriction in applications. Because of this restriction, pseudo-linear multiobjective problems ([1]) and certain nonlinear multiobjective fractional programming problems require separate treatment as far as optimality and duality properties are concerned. In order to avoid this restriction, [9] introduced the notion of V-invexity for a vector function $f = (f_1, f_2, \ldots, f_p)$ and discussed its applications to a class of constrained multiobjective optimization problems. We now give the definitions of [9] as follows.

Definition 9.2. A vector function $f : X \to R^p$ is said to be *V-invex* if there exist functions $\eta : X \times X \to R^n$ and $\alpha_i : X \times X \to R^+ - \{0\}$ such that for each $x, \bar{x} \in X$ and for $i = 1, 2, \ldots, p$,

$$f_i(x) - f_i(\bar{x}) \geq \langle \alpha_i(x,\bar{x}) \nabla f_i(\bar{x}), \eta(x,\bar{x}) \rangle.$$

For $p = 1$ and $\bar{\eta}(x,\bar{x}) = \alpha_i(x,\bar{x})\eta(x,\bar{x})$ the above definition reduces to the usual definition of invexity given by [6]. For more details about V-invex functions and vector optimization, the reader is referred to [11].

Definition 9.3. A vector function $f : X \to R^p$ is said to be $V-\rho$-*invex* if there exist functions $\eta : X \times X \to R^n$, $\alpha_i : X \times X \to R^+ - \{0\}$ and $\rho_i \in R$ such that for each $x, \bar{x} \in X$ and for $i = 1, 2, \ldots, p$,

$$f_i(x) - f_i(\bar{x}) \geq \langle \alpha_i(x,\bar{x}) \nabla f_i(\bar{x}),\ \eta(x,\bar{x}) \rangle + \rho_i \|x - \bar{x}\|^2.$$

Definition 9.4. A vector function $f : X \to R^p$ is said to be $V-\rho$-*pseudoinvex* if there exist functions $\eta : X \times X \to R^n$, $\beta_i : X \times X \to R^+ - \{0\}$ and $\rho_i \in R$ such that for each $x, \bar{x} \in X$ and for $i = 1, 2, \ldots, p$,

$$\left\langle \sum_{i=1}^{p} \nabla f_i(\bar{x}),\ \eta(x,\bar{x}) \right\rangle + \rho_i \|x - \bar{x}\|^2 \geq 0 \Rightarrow \sum_{i=1}^{p} \beta_i(x,\bar{x}) f_i(x) \geq \sum_{i=1}^{p} \beta_i(x,\bar{x}) f_i(\bar{x}).$$

Definition 9.5. A vector function $f : X \to R^p$ is said to be $V-\rho$-*quasi-invex* if there exist functions $\eta : X \times X \to R^n$, $\delta_i : X \times X \to R^+ - \{0\}$ and $\rho_i \in R$ such that for each $x, \bar{x} \in X$ and for $i = 1, 2, \ldots, p$,

$$\sum_{i=1}^{p} \delta_i(x,\bar{x}) f_i(x) \leq \sum_{i=1}^{p} \delta_i(x,\bar{x}) f_i(\bar{x}) \Rightarrow \left\langle \sum_{i=1}^{p} \nabla f_i(\bar{x}),\ \eta(x,\bar{x}) \right\rangle + \rho_i \|x - \bar{x}\|^2 \leq 0.$$

Definition 9.6. A vector function $f : X \to R^p$ is said to be $V-\rho$-*prestrictly quasi-invex* if there exist functions $\eta : X \times X \to R^n$, $\delta_i : X \times X \to R^+ - \{0\}$ and $\rho_i \in R$ such that for each $x, \bar{x} \in X$ and for $i = 1, 2, \ldots, p$,

$$\sum_{i=1}^{p} \delta_i(x,\bar{x}) f_i(x) < \sum_{i=1}^{p} \delta_i(x,\bar{x}) f_i(\bar{x}) \Rightarrow \left\langle \sum_{i=1}^{p} \nabla f_i(\bar{x}),\ \eta(x,\bar{x}) \right\rangle + \rho_i \|x - \bar{x}\|^2 \leq 0.$$

We need the Dinkelbach-type [3] indirect approach with the help of the following auxiliary parametric problem.

$$\underset{x \in F}{\text{Minimize}} \max_{1 \leq i \leq p} \{ f_i(x) - \lambda g_i(x) \}, \tag{Pλ}$$

where λ is a parameter. It is easy to see that this problem is equivalent to (P) if λ is chosen to be the optimal value of (P). Note that if λ^* is the optimal value of (P) and $v(\lambda)$ the optimal value of Pλ for any fixed $\lambda \in R$ such that Pλ has an optimal solution, then the following hold.

1. If x^* is an optimal solution of (P) then it is an optimal solution of Pλ^* and $v(\lambda^*) = 0$.
2. If P$\bar{\lambda}$ has an optimal solution $\bar{x}$ for some $\bar{\lambda} \in R$ with $v(\bar{\lambda}) = 0$, then $\bar{x}$ is an optimal solution of (P) and $\bar{\lambda} = \lambda^*$.

We need the following definitions and the Abadie constraint qualification in the sequel.

Definition 9.7. The tangent cone to the feasible set F of (P) at $\bar{x} \in F$ is the set

$$T(F, \bar{x}) = \big\{h \in R^n : h = \lim_{n\to\infty} t_n(x^n - \bar{x}) \text{ such that } x^n \in F,$$
$$\times \lim_{n\to\infty} x^n = \bar{x}, \textit{ and } t_n > 0, \forall\, n = 1, \ldots\big\}.$$

Definition 9.8. Let $\bar{x} \in F$. The linearizing cone at $\bar{x}$ for (P) is the set defined by

$$C(\bar{x}) \equiv \big\{h \in R^n : \langle \nabla G_j(\bar{x}, t), h\rangle \le 0, \forall\, t \in \hat{T}_j(\bar{x}),$$
$$j \in \underline{q}, \langle \nabla H_k(\bar{x}, s), h\rangle = 0, \forall\, s \in S_k(\bar{x}), k \in \underline{r}\big\},$$

where $\hat{T}_j(\bar{x}) \equiv \{t \in T_j : \ G_j(\bar{x}, t) = 0\}$.

Definition 9.9. The problem (P) satisfies the generalized Abadie constraint qualification at a given point $\bar{x} \in F$ if we have $C(\bar{x}) \subseteq T(F, \bar{x})$.

Lemma 9.1 (5). *Let A and B be compact sets in R^n and C an arbitrary set in R^n. Suppose that the set cone $(B) +$ span (C) is the conic hull of B (i.e., the smallest convex cone containing B) and span (C) is the linear hull of C (i.e., the smallest subspace containing C). Then either the system*

$$\begin{cases} \langle a, z\rangle < 0, & \forall\, a \in A, \\ \langle b, z\rangle \le 0, & \forall\, b \in B, \\ \langle c, z\rangle = 0, & \forall\, c \in C, \end{cases}$$

has a solution $z \in R^n$, or there exist integers μ, v_0, and v with $0 \le v_0 \le v \le n+1$, such that there exist μ points $a^i \in A$, v_0 points $b^m \in B$, $v - v_0$ points $c^m \in C$, μ nonnegative numbers u_i, with $u_i > 0$ for at least one $i \in \underline{\mu}$ and v real numbers v_m with $v_m > 0$ for $m \in \underline{v_0}$, such that

$$\sum_{i=1}^{\mu} u_i a^i + \sum_{m=1}^{v_0} v_m b^m + \sum_{m=v_0+1}^{v} v_m c^m = 0$$

but never both.

Zalmai and Zhang [14] *have shown that the optimality of $x^* \in F$ for (P) implies the inconsistency of a certain semi-infinite system of linear inequalities and equalities.*

Lemma 9.2. *Let x^* be an optimal solution for (P) and $\lambda^* = \max_{1\le i\le p} f_i(x^*)/g_i(x^*)$ for each $i \in \underline{p}$, let f_i and g_i be continuously differentiable at x^*, for each $j \in \underline{q}$, let the function $z \to G_j(z, t)$ be continuously differentiable at x^* for all $t \in T_j$, and for each $k \in \underline{r}$, let the function $z \to H_k(z, t)$ be continuously differentiable at x^* for all $s \in S_k$. If the generalized Abadie constraint qualification holds at x^*, then the system*

$$\begin{cases} \langle \nabla f_i(x^*) - \lambda^* \nabla g_i(x^*),\, z \rangle < 0, & i \in p(x^*), \\ \langle \nabla G_j(x^*,\, t),\, z \rangle \leq 0, & \forall\, t \in \hat{T}_j(x^*),\ j \in \underline{q}, \\ \langle \nabla H_k(x^*,\, s),\, z \rangle = 0, & \forall\, s \in S_k,\ k \in \underline{r}, \end{cases}$$

has no solution $z \in R^n$ *where* $p(x^*) \equiv \{j \in \underline{p} : f_j(x^*)/g_j(x^*) = \max_{1 \leq i \leq p} f_i(x^*)/g_i(x^*)\}$.

Lemma 9.3 (Necessary optimality conditions). *Let* $x^* \in F$ *and* $\lambda^* = \max_{1 \leq i \leq p} f_i(x^*)/g_i(x^*)$, *for each* $i \in \underline{p}$, *let* f_i *and* g_i *be continuously differentiable at* x^*, *for each* $j \in \underline{q}$, *let the function* $z \to G_j(z,\ t)$ *be continuously differentiable at* x^* *for all* $t \in T_j$, *and for each* $k \in \underline{r}$, *let the function* $z \to H_k(z,\ s)$ *be continuously differentiable at* x^* *for all* $s \in S_k$. *If* x^* *is an optimal solution for (P), if the generalized Abadie constraint qualification holds at* x^* *and if the set cone*$\{\nabla G_j(x^*,\ t) : t \in \hat{T}_j(x^*),\ j \in \underline{q}\} + span\{\nabla H_k(x^*,\ s) : s \in S_k, k \in \underline{r}\}$ *is closed, then there exist* $u^* \in U \equiv \left\{u \in R^p : u \geq 0, \sum_{i=1}^{p} u_i = 1\right\}$ *and integers* v_0^* *and* v^* *with* $0 \leq v_0^* \leq v^* \leq n+1$, *such that there exist* v_0^* *indices* j_m, *with* $1 \leq j_m \leq q$, *together with* v_0^* *points* $t^m \in \hat{T}_{j_m}(x^*)$, $m \in \underline{v}_0^*$, $v^* - v_0^*$ *indices* k_m, *with* $1 \leq k_m \leq r$, *together with* $v^* - v_0^*$ *points* $s^m \in S_{k_m}(x^*)$, $m \in \underline{v}^* \backslash \underline{v}_0^*$ *and* v^* *real numbers* v_m^* *with* $v_m^* > 0$ *for* $m \in \underline{v}_0^*$ *with the property that*

$$\sum_{i=1}^{p} u_i^* [\nabla f_i(x^*) - \lambda^* \nabla g_i(x^*)] + \sum_{m=1}^{v_0^*} v_m^* \nabla G_{j_m}(x^*,\, t^m)$$

$$+ \sum_{m=v_0^*+1}^{v^*} v_m^* \nabla H_{k_m}(x^*,\, s^m) = 0$$

$$u_i^* [f_i(x^*) - \lambda^* g_i(x^*)] = 0, \qquad i \in \underline{p}.$$

The following parameter-free version of the above necessary optimality conditions is given in [14].

Lemma 9.4 (Parameter-free necessary optimality conditions). *Let* $x^* \in F$ *and let the functions* f_i *and* $g_i, i \in \underline{p}$, $z \to G_j(z,\, t)$ *and* $z \to H_k(z,\, s)$ *be continuously differentiable at* x^*, *for all* $t \in T_j$, $s \in S_k$, $j \in \underline{q}$, *and* $k \in \underline{r}$. *If* x^* *is an optimal solution for (P), if the generalized Abadie constraint qualification holds at* x^*, *and if the set cone*

$$\{\nabla G_j(x^*,\, t) : t \in \hat{T}_j(x^*),\ j \in \underline{q}\} + span\{\nabla H_k(x^*,\, s) : s \in S_k, k \in \underline{r}\}$$

is closed, then there exist $u^* \in U \equiv \{u \in R^p : u \geq 0, \sum_{i=1}^{p} u_i = 1\}$ *and integers* v_0 *and* v *with* $0 \leq v_0 \leq v \leq n+1$, *such that there exist* v_0 *indices* j_m, *with* $1 \leq j_m \leq q$, *together with* v_0 *points* $t^m \in \hat{T}_{j_m}(x^*)$, $m \in \underline{v}_0$, $v - v_0$ *indices* k_m, *with* $1 \leq k_m \leq r$, *together with* $v - v_0$ *points* $s^m \in S_{k_m}(x^*)$, $m \in \underline{v} \backslash \underline{v}_0$ *and* v *real numbers* v_m^* *with* $v_m^* > 0$ *for* $m \in \underline{v}_0$ *with the property that*

$$\sum_{i=1}^{p} u_i^* [\Gamma(x^*, u^*)\nabla f_i(x^*) - \Phi(x^*, u^*)\nabla g_i(x^*)]$$

$$+ \sum_{m=1}^{v_0} v_m^* \nabla G_{j_m}(x^*, t^m) + \sum_{m=v_0+1}^{v} v_m^* \nabla H_{k_m}(x^*, s^m) = 0$$

$$u_i^* [\Gamma(x^*, u^*)\nabla f_i(x^*) - \Phi(x^*, u^*)\nabla g_i(x^*)] = 0, \qquad i \in \underline{p}$$

$$\max_{1 \le i \le p} \frac{f_i(x^*)}{g_i(x^*)} = \frac{\Phi(x^*, u^*)}{\Gamma(x^*, u^*)},$$

where $\Phi(x^*, u^*) = \sum_{i=1}^{p} u_i^* f_i(x^*)$ *and* $\Gamma(x^*, u^*) = \sum_{i=1}^{p} u_i^* g_i(x^*)$.

[14] is needed in the proofs of the sufficient optimality conditions in the next section. The proof of the lemma is straightforward and omitted.

Lemma 9.5. *For each* $x \in X$, $\Psi(x) \equiv \max_{1 \le i \le p} \frac{f_i(x)}{g_i(x)}$.

9.3 Sufficient Optimality Conditions

In this section, we present a number of sufficient optimality conditions in which various generalized $V - \rho$-invexity assumptions are imposed.

Theorem 9.1. *Let* $x^* \in F$ *and assume that* $f_i(x^*) \ge 0$, $g_i(x^*) > 0$, $i \in \underline{p}$, *that the functions* f_i, g_i, $i \in \underline{p}$, $z \to G_j(z, t)$, *and* $z \to H_k(z, s)$ *are differentiable at* x^* *for all* $t \in T_j$, $s \in S_k$, $j \in \underline{q}$, *and* $k \in \underline{r}$, *and that there exist* $u^* \in U$ *and integers* v_0 *and* v *with* $0 \le v_0 \le v \le n+1$, *such that there exist* v_0 *indices* j_m, *with* $1 \le j_m \le q$, *together with* v_0 *points* $t^m \in \hat{T}_{j_m}(x^*)$, $m \in \underline{v}_0$, $v - v_0$ *indices* k_m, *with* $1 \le k_m \le r$, *together with* $v - v_0$ *points* $s^m \in S_{k_m}(x^*)$, $m \in \underline{v} \backslash \underline{v}_0$, *and* v *real numbers* v_m^* *with* $v_m^* > 0$ *for* $m \in \underline{v}_0$ *with the property that*

$$\sum_{i=1}^{p} u_i^* [\Gamma(x^*, u^*)\nabla f_i(x^*) - \Phi(x^*, u^*)\nabla g_i(x^*)] + \sum_{m=1}^{v_0} v_m^* \nabla G_{j_m}(x^*, t^m)$$

$$+ \sum_{m=v_0+1}^{v} v_m^* \nabla H_{k_m}(x^*, s^m) = 0 \tag{9.1}$$

$$\max_{1 \le i \le p} \frac{f_i(x^*)}{g_i(x^*)} = \frac{\Phi(x^*, u^*)}{\Gamma(x^*, u^*)}. \tag{9.2}$$

Assume furthermore that any one of the following conditions holds.

1. a. For each $i \in \underline{p}$, f_i is $V - \bar{\rho}_i$-invex and $-g_i$ is $V - \tilde{\rho}_i$-invex at x^*.
   b. The function $z \to G_{j_m}(z, t^m)$ is $V - \hat{\rho}_m$-quasi-invex at x^* for each $m \in \underline{v}_0$.

c. The function $z \to v_m^* H_{k_m}(z, s^m)$ is $V - \breve{\rho}_m$-quasi-invex at x^* for each $m \in \underline{v} \backslash \underline{v}_0$.
d. $\rho^* + \sum_{m=1}^{v_0} v_m^* \hat{\rho}_m + \sum_{m=v_0+1}^{v} \breve{\rho}_m \geq 0$.

2. a. For each $i \in \underline{p}$, f_i is $V - \bar{\rho}_i$-invex and $-g_i$ is $V - \tilde{\rho}_i$-invex at x^*.
b. The function $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z, t^m)$ is $V - \hat{\rho}$-quasi-invex at x^* for each $m \in \underline{v}_0$.
c. The function $z \to v_m^* H_{k_m}(z, s^m)$ is $V - \breve{\rho}_m$-quasi-invex at x^* for each $m \in \underline{v} \backslash \underline{v}_0$.
d. $\rho^* + \hat{\rho} + \sum_{m=v_0+1}^{v} \breve{\rho}_m \geq 0$.

3. a. For each $i \in \underline{p}$, f_i is $V - \bar{\rho}_i$-invex and $-g_i$ is $V - \tilde{\rho}_i$-invex at x^*.
b. The function $z \to G_{j_m}(z, t^m)$ is $V - \hat{\rho}_m$-quasi-invex at x^* for each $m \in \underline{v}_0$.
c. The function $z \to \sum_{m=v_0+1}^{v} v_m^* H_{k_m}(z, s^m)$ is $V - \breve{\rho}$-quasi-invex at x^* for each $m \in \underline{v} \backslash \underline{v}_0$.
d. $\rho^* + \sum_{m=1}^{v_0} v_m^* \hat{\rho}_m + \breve{\rho} \geq 0$.

4. a. For each $i \in \underline{p}$, f_i is $V - \bar{\rho}_i$-invex and $-g_i$ is $V - \tilde{\rho}_i$-invex at x^*.
b. The function $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z, t^m)$ is $V - \hat{\rho}$-quasi-invex at x^* for each $m \in \underline{v}_0$.
c. The function $z \to \sum_{m=v_0+1}^{v} v_m^* H_{k_m}(z, s^m)$ is $V - \breve{\rho}$-quasi-invex at x^* for each $m \in \underline{v} \backslash \underline{v}_0$.
d. $\rho^* + \hat{\rho} + \breve{\rho} \geq 0$;

5. a. For each $i \in \underline{p}$, f_i is $V - \bar{\rho}_i$-invex and $-g_i$ is $V - \tilde{\rho}_i$-invex at x^*.
b. The function $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z, t^m) + \sum_{m=v_0+1}^{v} v_m^* H_{k_m}(z, s^m)$ is $V - \hat{\rho}$-quasi-invex at x^*.
c. $\rho^* + \hat{\rho} \geq 0$.

Then x^* is an optimal solution for (P).

Proof. Let x be an arbitrary feasible solution for (P).

(a) From feasibility, it follows that $G_{j_m}(x, t^m) \leq 0 = G_{j_m}(x^*, t^m)$, for $t^m \in \hat{T}_{j_m}(x^*)$ for each $m \in \underline{v}_0$. Because $\beta_{j_m}(x, x^*) > 0$, for each $m \in \underline{v}_0$, we get

$$\beta_{j_m}(x, x^*) G_{j_m}(x, t^m) \leq \beta_{j_m}(x, x^*) G_{j_m}(x^*, t^m).$$

Function $z \to G_{j_m}(z, t^m)$ is $V - \hat{\rho}_m$-quasi-invex at x^* for each $m \in \underline{v}_0$, therefore we have

$$\langle \nabla G_{j_m}(x^*, t^m),\ \eta(x, x^*) \rangle + \hat{\rho}_m \|x - x^*\|^2 \leq 0.$$

As $v_m^* > 0$ for each $m \in \underline{v}_0$, the above inequalities yield

$$\left\langle \sum_{m=1}^{v_0} v_m^* \nabla G_{j_m}(x^*, t^m),\ \eta(x, x^*) \right\rangle + \sum_{m=1}^{v_0} v_m^* \hat{\rho}_m \|x - x^*\|^2 \leq 0. \tag{9.3}$$

Similarly, by feasibility and $V-\breve{\rho}_m$-quasi-invexity of the function $z \to v_m^* H_{k_m}(z,\ s^m)$ at x^* for each $m \in \underline{v}\backslash\underline{v}_0$, we have

$$\left\langle \sum_{m=v_0+1}^{v} v_m^* \nabla H_{k_m}(x^*,\ s^m),\ \eta\,(x,\ x^*) \right\rangle + \sum_{m=v_0+1}^{v} \breve{\rho}_m \|x-x^*\|^2 \leq 0. \tag{9.4}$$

Because $u^* \geq 0,\ \Phi(x^*,\ u^*) \geq 0$, and $\Gamma(x^*,\ u^*) > 0$, we have

$$\sum_{i=1}^{p} u_i^* [\Gamma(x^*,\ u^*) f_i(x) - \Phi(x^*,\ u^*) g_i(x)]$$

$$= \sum_{i=1}^{p} u_i^* \{\Gamma(x^*,\ u^*)[f_i(x) - f_i(x^*)] - \Phi(x^*,\ u^*)[g_i(x) - g_i(x^*)]\}$$

(by the definitions of Φ and Γ)

$$\geq \sum_{i=1}^{p} u_i^* \{\langle \Gamma(x^*,\ u^*)\alpha(x,\ x^*)\nabla f_i(x^*) - \Phi(x^*,\ u^*)\delta(x,\ x^*)\nabla g_i(x^*),\ \eta(x,\ x^*)\rangle$$

$$+ [\Gamma(x^*,\ u^*)\bar{\rho}_i + \Phi(x^*,\ u^*)\tilde{\rho}_i]\|x-x^*\|^2 l\}$$

(by (a) of condition (1))

$$= -\left\langle \sum_{i=1}^{v_0} v_m^* \nabla G_{j_m}(x^*,\ t^m) + \sum_{m=v_0+1}^{v} v_m^* \nabla H_{k_m}(x^*,\ s^m),\ \eta(x,\ x^*) \right\rangle$$

$$+ \sum_{i=1}^{p} u_i^* [\Gamma(x^*,\ u^*)\bar{\rho}_i + \Phi(x^*,\ u^*)\tilde{\rho}_i]\|x-x^*\|^2 \quad \text{(by (9.1)}$$

$$\geq \left(\rho^* + \sum_{m=1}^{v_0} v_m^* \hat{\rho}_m + \sum_{m=v_0+1}^{v} v_m^* \breve{\rho}_m \right) \|x-x^*\|^2 \quad \text{(by (9.3) and (9.4))}$$

$$\geq 0 \quad \text{(by (d) of condition (1))}. \tag{9.5}$$

From (9.2), (9.5), and Lemma 9.1, we have

$$\Psi(x^*) = \frac{\Phi(x^*,\ u^*)}{\Gamma(x^*,\ u^*)} \leq \frac{\sum_{i=1}^{p} u_i^* f_i(x^*)}{\sum_{i=1}^{p} u_i^* g_i(x^*)} \leq \max_{u \in U} \frac{\sum_{i=1}^{p} u_i^* f_i(x^*)}{\sum_{i=1}^{p} u_i^* g_i(x^*)} = \Psi(x).$$

Thus, x^* is an optimal solution for (P).

(b) From feasibility and the fact that $v_m^* > 0$ for each $m \in \underline{v}_0$, we have

$$\sum_{m=1}^{v_0} v_m^* \beta_{j_m}(x,\ x^*) G_{j_m}(x,\ t^m) \leq \sum_{m=1}^{v_0} v_m^* \beta_{j_m}(x,\ x^*) G_{j_m}(x^*,\ t^m),$$

which by $V-\hat{\rho}$-quasi-invexity of the function $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z, t^m)$ at x^* for each $m \in \underline{v}_0$, yields

$$\left\langle \sum_{m=1}^{v_0} v_m^* \nabla G_{j_m}(x^*, t^m),\ \eta(x, x^*) \right\rangle + \hat{\rho}\|x-x^*\|^2 \leq 0. \tag{9.6}$$

The rest of the proof is similar to part (a): in place of inequality (9.3); in part we use (9.6) and get (9.5) which leads to the conclusion that x^* is an optimal solution for (P).

The proofs of parts (c)–(e) are similar to the parts (a) and (b).

Let the functions $z \to \Theta(z, x^*, u^*)$ and $z \to \Xi(z, x^*, u^*)$ defined for fixed x^* and u^* on X by

$$\Theta(z, x^*, u^*) = \Gamma(x^*, u^*) f_i(z) - \Phi(x^*, u^*) g_i(z), \qquad i \in \underline{p}, \Xi$$

$$(z, x^*, u^*) = \sum_{i=1}^{p} u_i^* [\Gamma(x^*, u^*) f_i(z) - \Phi(x^*, u^*) g_i(z)].$$

Theorem 9.2. *Let $x^* \in F$ and assume that $f_i(x^*) \geq 0$, $g_i(x^*) > 0$, $i \in \underline{p}$, that the functions f_i, g_i, $i \in \underline{p}$, $z \to G_j(z, t)$ and $z \to H_k(z, s)$ are differentiable at x^* for all $t \in T_j$, $s \in S_k$, $j \in \underline{q}$, and $k \in \underline{r}$, and that there exist $u^* \in U$ and integers v_0 and v with $0 \leq v_0 \leq v \leq n+1$, such that there exist v_0 indices j_m, with $1 \leq j_m \leq q$, together with v_0 points $t^m \in \hat{T}_{j_m}(x^*)$, $m \in \underline{v}_0$, $v - v_0$ indices k_m, with $1 \leq k_m \leq r$, together with $v - v_0$ points $s^m \in S_{k_m}(x^*)$, $m \in \underline{v} \backslash \underline{v}_0$ and v real numbers v_m^* with $v_m^* > 0$ for $m \in \underline{v}_0$ such that* 9.1 *and* 9.2 *hold. Assume furthermore that any one of the following conditions holds.*

6. a. $z \to \Xi(z, x^*, u^*)$ *is* $V-\rho_i$*-pseudo-invex at* x^*.
   b. *The function* $z \to G_{j_m}(z, t^m)$ *is* $V-\hat{\rho}_m$*-quasi-invex at* x^* *for each* $m \in \underline{v}_0$.
   c. *The function* $z \to H_{k_m}(z, s^m)$ *is* $V-\breve{\rho}_m$*-quasi-invex at* x^* *for each* $m \in \underline{v} \backslash \underline{v}_0$.
   d. $\bar{\rho} + \sum_{m=1}^{v_0} v_m^* \hat{\rho}_m + \sum_{m=v_0+1}^{v} \breve{\rho}_m \geq 0$.
7. a. $z \to \Xi(z, x^*, u^*)$ *is* $V-\rho_i$*-pseudo-invex at* x^*.
   b. *The function* $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z, t^m)$ *is* $V-\hat{\rho}$*-quasi-invex at* x^* *for each* $m \in \underline{v}_0$.
   c. *The function* $z \to v_m^* H_{k_m}(z, s^m)$ *is* $V-\breve{\rho}_m$*-quasi-invex at* x^* *for each* $m \in \underline{v} \backslash \underline{v}_0$.
   d. $\bar{\rho} + \hat{\rho} + \sum_{m=v_0+1}^{v} \breve{\rho}_m \geq 0$.
8. a. $z \to \Xi(z, x^*, u^*)$ *is* $V-\rho_i$*-pseudo-invex at* x^*.
   b. *The function* $z \to G_{j_m}(z, t^m)$ *is* $V-\hat{\rho}_m$*-quasi-invex at* x^* *for each* $m \in \underline{v}_0$.
   c. *The function* $z \to \sum_{m=v_0+1}^{v} v_m^* H_{k_m}(z, s^m)$ *is* $V-\breve{\rho}$*-quasi-invex at* x^* *for each* $m \in \underline{v} \backslash \underline{v}_0$.
   d. $\bar{\rho} + \sum_{m=1}^{v_0} v_m^* \hat{\rho}_m + \breve{\rho} \geq 0$.
9. a. $z \to \Xi(z, x^*, u^*)$ *is* $V-\rho_i$*-pseudo-invex at* x^*.
   b. *The function* $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z, t^m)$ *is* $V-\hat{\rho}$*-quasi-invex at* x^* *for each* $m \in \underline{v}_0$.

c. *The function* $z \to \sum_{m=v_0+1}^{v} v_m^* H_{k_m}(z,\, s^m)$ *is* $V - \breve{\rho}$*-quasi-invex at* x^* *for each* $m \in \underline{v} \backslash \underline{v}_0$.
d. $\bar{\rho} + \hat{\rho} + \breve{\rho} \geq 0$.

10. a. $z \to \Xi(z,\, x^*,\, u^*)$ *is* $V - \rho_i$*-pseudo-invex at* x^*.
b. *The function* $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z,\, t^m) + \sum_{m=v_0+1}^{v} v_m^* H_{k_m}(z,\, s^m)$ *is* $V - \hat{\rho}$*-quasi-invex at* x^*.
c. $\bar{\rho} + \hat{\rho} \geq 0$.

Then x^* *is an optimal solution for (P).*

Proof. (f) Let x be an arbitrary feasible solution for (P). Due to the assumptions (b) and (c), 9.3 and 9.4 remain valid. Using 9.1, we get

$$\left\langle \sum_{i=1}^{p} u_i^* [\Gamma(x^*,\, u^*) \nabla f_i(x^*) - \Phi(x^*,\, u^*) \nabla g_i(x^*)],\, \eta(x,\, x^*) \right\rangle$$

$$\geq \left(\sum_{m=1}^{v_0} v_m^* \hat{\rho}_m + \sum_{m=v_0+1}^{v} \breve{\rho}_m \right) \|x - x^*\|^2 \geq -\bar{\rho} \|x - x^*\|^2$$

$$\Rightarrow \Xi(x, x^*, u^*) \geq \Xi(x^*, x^*, u^*) = 0$$

which leads to the conclusion that x^* is an optimal solution for (P).

The proofs of other parts are similar to part (f).

Theorem 9.3. *Let* $x^* \in F$ *and assume that* $f_i(x^*) \geq 0,\ g_i(x^*) > 0,\ i \in \underline{p}$, *that the functions* $f_i,\ g_i,\ i \in \underline{p},\ z \to G_j(z,\, t)$ *and* $z \to H_k(z,\, s)$ *are differentiable at* x^* *for all* $t \in T_j,\ s \in S_k,\ j \in \underline{q}$, *and* $k \in \underline{r}$, *and that there exist* $u^* \in U$ *and integers* v_0 *and* v *with* $0 \leq v_0 \leq v \leq n+1$, *such that there exist* v_0 *indices* j_m, *with* $1 \leq j_m \leq q$, *together with* v_0 *points* $t^m \in \hat{T}_{j_m}(x^*)$, $m \in \underline{v}_0$, $v - v_0$ *indices* k_m, *with* $1 \leq k_m \leq r$, *together with* $v - v_0$ *points* $s^m \in S_{k_m}(x^*)$, $m \in \underline{v} \backslash \underline{v}_0$ *and* v *real numbers* v_m^* *with* $v_m^* > 0$ *for* $m \in \underline{v}_0$ *such that* 9.1 *and* 9.2 *hold. Assume furthermore that any one of the following conditions is satisfied.*

11. a. $z \to \Xi(z,\, x^*,\, u^*)$ *is* $V - \rho_i$*-prestrictly quasi-invex at* x^*.
b. *The function* $z \to G_{j_m}(z,\, t^m)$ *is* $V - \hat{\rho}_m$*-quasi-invex at* x^* *for each* $m \in \underline{v}_0$.
c. *The function* $z \to v_m^* H_{k_m}(z,\, s^m)$ *is* $V - \breve{\rho}_m$*-quasi-invex at* x^* *for each* $m \in \underline{v} \backslash \underline{v}_0$.
d. $\bar{\rho} + \sum_{m=1}^{v_0} v_m^* \hat{\rho}_m + \sum_{m=v_0+1}^{v} \breve{\rho}_m \geq 0$.

12. a. $z \to \Xi(z,\, x^*,\, u^*)$ *is* $V - \rho_i$*-prestrictly quasi-invex at* x^*.
b. *The function* $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z,\, t^m)$ *is* $V - \hat{\rho}$*-quasi-invex at* x^* *for each* $m \in \underline{v}_0$.
c. *The function* $z \to v_m^* H_{k_m}(z,\, s^m)$ *is* $V - \breve{\rho}_m$*-quasi-invex at* x^* *for each* $m \in \underline{v} \backslash \underline{v}_0$.
d. $\bar{\rho} + \hat{\rho} + \sum_{m=v_0+1}^{v} \breve{\rho}_m \geq 0$.

13. a. $z \to \Xi(z,\, x^*,\, u^*)$ *is* $V - \rho_i$*-prestrictly quasi-invex at* x^*.
b. *The function* $z \to G_{j_m}(z,\, t^m)$ *is* $V - \hat{\rho}_m$*-quasi-invex at* x^* *for each* $m \in \underline{v}_0$.

c. The function $z \to \sum_{m=v_0+1}^{v} v_m^* H_{k_m}(z, s^m)$ *is* $V - \breve{\rho}$*-quasi-invex at* x^* *for each* $m \in \underline{v}\backslash\underline{v}_0$.

d. $\bar{\rho} + \sum_{m=1}^{v_0} v_m^* \hat{\rho}_m + \breve{\rho} \geq 0$.

14. a. $z \to \Xi(z, x^*, u^*)$ *is* $V - \rho_i$*-prestrictly quasi-invex at* x^*.

b. The function $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z, t^m)$ *is* $V - \hat{\rho}$*-quasi-invex at* x^* *for each* $m \in \underline{v}_0$.

c. The function $z \to \sum_{m=v_0+1}^{v} v_m^* H_{k_m}(z, s^m)$ *is* $V - \breve{\rho}$*-quasi-invex at* x^* *for each* $m \in \underline{v}\backslash\underline{v}_0$.

d. $\bar{\rho} + \hat{\rho} + \breve{\rho} \geq 0$.

15. a. $z \to \Xi(z, x^*, u^*)$ *is* $V - \rho_i$*-prestrictly quasi-invex at* x^*.

b. The function $z \to \sum_{m=1}^{v_0} v_m^* G_{j_m}(z, t^m) + \sum_{m=v_0+1}^{v} v_m^* H_{k_m}(z, s^m)$ *is* $V - \hat{\rho}$*-quasi-invex at* x^*.

c. $\bar{\rho} + \hat{\rho} \geq 0$.

Then x^* *is an optimal solution for (P).*

Proof. (k) Let x be an arbitrary feasible solution for (P). Due to the assumptions (b) and (c), 9.3 and 9.4 remain valid. Using 9.1, we get

$$\left\langle \sum_{i=1}^{p} u_i^* [\Gamma(x^*, u^*) \nabla f_i(x^*) - \Phi(x^*, u^*) \nabla g_i(x^*)], \; \eta(x, x^*) \right\rangle$$

$$\geq \left(\sum_{m=1}^{v_0} v_m^* \hat{\rho}_m + \sum_{m=v_0+1}^{v} \breve{\rho}_m \right) \|x - x^*\|^2 > -\bar{\rho} \|x - x^*\|^2$$

$$\Rightarrow \Xi(x, x^*, u^*) \geq \Xi(x^*, x^*, u^*) = 0, \quad \text{(by assumption (a))}$$

which leads to the conclusion that x^* is an optimal solution for (P).

The proofs of other parts are similar to part (k).

References

1. Chew, K.L., Choo, E.U.: Pseudolinearity and efficiency. *Math. Program.* **28**, 226–239 (1984)
2. Craven, B.D.: Invex functions and constrained local minima, *Bull. Austral. Math. Soc.* **24** 357–366 (1981)
3. Dinkelbach, W.: On nonlinear fractional programming. *Manage. Sci.* **13** 492–498 (1967)
4. Fiaclco, A.V., Kortanek K.O. (Eds.): *Semi-infinite Programming and Applications. Lecture Notes in Economics and Mathematical Systems*, Vol. 215, Springer, Berlin (1983)
5. Goberna, M.A., Lopez, M.A. (Eds.): *Semi-infinite Programming-Recent Advances*. Kluwer, Dordrecht (2001)
6. Hanson, M.A.: On sufficiency of the Kuhn-Tucker conditions, *J. Math. Anal. Appl.* **80** 545–550 (1981)
7. Hettich, R. (eds.): *Semi-Infinite Programming. Lecture Notes in Control and Information Sciences*, Vol. 15, Springer, Berlin (1979)
8. Hettich R., Kortanek, K.O.: Semi-infinite programming: Theory, methods and applications. *SIAM Rev.* **35** 380–429 (1993)

9. Jeyakumar V., Mond, B.: On generalized convex mathematical programming, *J. Austral. Math. Soc. Ser. B* **34** 43–53 (1992)
10. Mishra, S.K., Giorgi, G.: *Invexity and Optimization*. Springer, Berlin (2008)
11. Mishra, S.K., Wang, S.Y., Lai, K.K.: *V-Invex Functions and Vector Optimization*. Springer, New York (2008)
12. Reemtsen, R., Ruckmann, J.J. (Eds.): *Semi-Infinite Programming*, Kluwer, Boston (1998)
13. Weber, G.W.: Generalized semi-infinite optimization: Theory and applications in optimal control and discrete optimization. *J. Stat. Manage. Syst.* **5** 359–388 (2002)
14. Zalmai, G.J., Zhang, Q.: Global nonparametric sufficient optimality conditions for semi-infinite discrete minimax fractional programming problems involving generalized (η,ρ)-invex functions. *Numer. Funct. Anal. Optim.* **28** (1–2) 173–209 (2007)

Chapter 10
Ekeland-Type Variational Principles and Equilibrium Problems

Qamrul Hasan Ansari and Lai-Jiu Lin

Abstract In this survey chapter we present different forms of Ekeland's variational principle involving τ-functions, τ-functions and fitting functions, and Q-functions, respectively. The equilibrium version of Ekeland-type variational principle is also presented. We give some equivalences of our variational principles with the Caristi–Kirk-type fixed point theorem for multivalued maps, Takahashi minimization theorem, and some other related results. As applications of our results, we derive the existence results for solutions of equilibrium problems and fixed point theorems for multivalued maps. The results of this chapter extend and generalize many results that recently appeared in the literature.

10.1 Introduction

The following variational principle was discovered by Ekeland in 1972 [23] (see also, [24, 25]), now known as Ekeland's variational principle (in short, EVP).

Theorem 10.1 (Ekeland's Variational Principle). *Let X be a complete metric space, and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower semicontinuous, and bounded below functional. Then for any given $\varepsilon > 0$, for any $x_0 \in X$ such that $f(x_0) > \inf_{x \in X} f(x)$ and for every $\lambda > 0$, there exists $\bar{x} \in X$ such that*

$$\begin{cases} f(\bar{x}) \leq f(x_0), \\ d(x_0, \bar{x}) \leq \lambda, \\ f(\bar{x}) < f(y) + (\varepsilon/\lambda) d(\bar{x}, y) \quad \text{for all } y \in X, \ y \neq \bar{x}. \end{cases}$$

Qamrul Hasan Ansari
Department of Mathematics, Aligarh Muslim University, Aligarh 202 002, India,
e-mail: qhansari@gmail.com

Lai-Jiu Lin
Department of Mathematics, National Changhua University for Education, Changhua 50058, Taiwan, ROC, e-mail: maljlin@math.ncue.edu.tw

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_10,

It provides an approximate minimizer of a bounded below lower semicontinuous function in a given neighborhood of a point. This localization property is very useful and explains the importance of this result. EVP is among the most important results obtained in nonlinear analysis. It has appeared as one of the most useful tools for solving problems from optimization, optimal control theory, game theory, nonlinear equations, dynamical systems, and so on. See, for example, [5–7, 22, 26, 35, 54, 55, 67] and the references therein. Since the discovery of EVP, there have also appeared many extensions or equivalent formulations of EVP. See, for example, [1, 7, 13, 30, 34, 38–41, 44, 45, 49–51, 54–56, 59, 61, 64, 68, 70] and the references therein.

In 1981, Sullivan [59] established that the validity of an EVP statement on a metric space (X,d) is equivalent to the completeness of (X,d). In 1982, McLinden [51] showed how EVP, or more precisely the augmented form of it provided by Rockafellar [58], can be adapted to extremum problems of minimax type. Many famous results, namely, the Krasnosel'skii–Zabrjeko, the Caristi–Kirk fixed point theorem [16], the Petal theorem, and the Daneš drop theorem, were discovered around the same time but independently of each other. It was later found that all these results were equivalent to EVP. In 1986, Penot [57] proved that EVP is equivalent to some of the most famous results in nonlinear analysis, namely, Caristi's fixed point theorem [15] and the petal theorem. In the same paper, some areas of applications of a geometric result of Daneš [20] known as the drop theorem are mentioned. For further details on the equivalences of these results, we refer to [21, 34, 43, 38, 55, 62] and the references therein.

Aubin and Frankowska [7] established the following form of Ekeland's variational principle which is equivalent to Theorem 10.1.

Theorem 10.2 ([7]). *Let (X,d) be a complete metric space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper, bounded below, and lower semicontinuous functional. Let $\hat{x} \in Dom(f)$ and $e > 0$ be fixed. Then there exists $\bar{x} \in X$ such that*

(a) $f(\bar{x}) - f(\hat{x}) + ed(\hat{x},\bar{x}) \leq 0$.
(b) $f(\bar{x}) < f(x) + ed(x,\bar{x})$ *for all* $x \in X \setminus \{\bar{x}\}$.

In 1996, Kada, Suzuki, and Takahashi [45] introduced the concept of a w-distance defined on a metric space and extended EVP, the minimization theorem, and Kirk–Caristi fixed point theorem for a w-distance. Suzuki [61] introduced a more general concept than w-distance, called τ-distance, and established EVP for τ-distance. He also extended most of the results of [45] for τ-distance. For further results involving τ-functions, we refer to [64] and references therein. It seems that the concept of τ-distance is a little more complicated, therefore Lin and Du [49] introduced the concept of a τ-function which is an extension of a w-distance but independent of τ-distance. They established a generalized EVP for lower semicontinuous from above functions and with a τ-function. They also derived the minimization theorem, nonconvex equilibrium theorem, common fixed point theorem for a family of multivalued maps, and flower petal theorem.

Recently, Lin and Du [50] introduced another kind of function, called the fitting function, which is also more general than the w-distance. They established a variant of the generalized EVP and maximal element theorem involving τ-functions and fitting functions. They provided some equivalent formulations of this theorem. They also gave another type of EVP and maximal element theorem involving a τ-function.

In [1], we introduced the concept of a Q-function defined on a quasi-metric space that generalizes the notion of a τ-function and a w-distance. We established Ekeland-type variational principles, one in the setting of quasi-metric spaces with a Q-function but without any lower semicontinuity assumption on the underlying function and the other in the setting of complete quasi-metric spaces with a Q-function. The equilibrium version of an Ekeland-type variational principle in the setting of quasi-metric spaces with a Q-function is also presented. We proved some equivalences of our variational principles with a Caristi–Kirk-type fixed point theorem for multivalued maps, Takahashi minimization theorem, and some other related results. As applications of our results, we derived the existence results for solutions of equilibrium problems and fixed point theorems for multivalued maps.

Moreover, EVP was also considered and studied in a more general setting, for example, in the setting of $\mathscr{F}$-topological spaces [30]. In the recent past, it was also studied in the setting of locally convex spaces and uniform spaces; See, for example, [40–42] and the references therein. Very recently, Hamel [40] and Jing-Hui [44] proved the equivalence of EVP, Phelp's lemma, and Daneš drop theorem in the setting of locally convex spaces.

Investigations of equilibrium states of a system play a central role in such diverse fields as economics, mechanics, biology, and the social sciences. Now there are a number of general mathematical problems suggested for modeling and studying various kinds of equilibria. Many researchers were and are considering these problems in order to obtain existence and uniqueness results and propose solution methods. The mathematical equilibrium problem (in short, EP), which is to find an element $\bar{x}$ of a set K such that

$$F(\bar{x},y) \geq 0 \quad \text{for all } y \in K,$$

where $F : K \times K \to \mathbb{R}$ is a bifunction such that $F(x,x) = 0$ for all $x \in K$, seems the most general problem and includes other equilibrium types such as optimization, saddle point, fixed point, complementarity, and variational inequality ones. In this general form, EP was first considered by Nikaido and Isoda [53] as an auxiliary problem to establish existence results for Nash's equilibrium points in noncooperative games. This transformation allows one to extend various iterative methods, which were proposed for saddle point problems, for the case of EP. In EP theory, the key contribution was made by Ky Fan, whose new existence results contained the original techniques which became a basis for most further existence theorems in topological spaces. Within the context of calculus of variations, motivated mainly by the works of Stampacchia, there arises the work of Brézis, Nirenberg, and Stampacchia [14] establishing a more general result than that in [29]. The equilibrium problem is a unified model of several problems, for example, the optimization problem, saddle point problem, Nash equilibrium problem, variational inequality problem, nonlinear complementarity problem, fixed point problem, and so on. In the last decade, it emerged as a new research direction in nonlinear analysis, optimization, optimal control, game theory, mathematical economics, among others. Most of the results on the existence of solutions of equilibrium problems are studied in the setting of topological vector spaces by using some kind of fixed point (Fan–Browder type-fixed point) theorem or KKM-type theorem. But it was after the work of Blum and Oetteli [12] that many mathematicians started to study the EP again. For further details on equilibrium problems, we refer to [4, 9–12, 14, 17, 18, 27, 28, 32, 33, 36, 37, 69] and the references therein.

In 1993, Oettli and Théra [54] started the study of the equilibrium version of Ekeland's variational principle. By using this kind of variational principle, they first gave the existence of a solution of an equilibrium problem in the setting of complete metric spaces. They also showed that their existence result for a solution of the equilibrium problem is equivalent to the Ekeland-type variational principle for bifunctions, Caristi–Kirk fixed point theorem for multivalued maps [16], and a maximal element theorem. It was further studied by Hamel [39] and Park [56] in a more general setting. They also proved the equivalences of several problems and equilibrium version of EVP. Recently, Bianchi, Rassay, and Pini [11] established the following extended form of the equilibrium version of EVP but in the setting of finite-dimensional spaces.

Theorem 10.3 ([11]). *Let K be a nonempty closed subset of $\mathbb{R}^n$ and $F : K \times K \to \mathbb{R}$ be a bifunction such that the following conditions hold.*

(i) For all $x \in K$, $F(x,x) = 0$.
(ii) For all $x \in K$, $F(x,\cdot)$ is bounded below and lower semicontinuous.
(iii) For all $x,y,z \in K$, $F(x,z) \leq F(x,y) + F(y,z)$.

Then, for every $\varepsilon > 0$ and for every $x_0 \in K$, there exists $\bar{x} \in K$ such that

$$\begin{cases} F(x_0,\bar{x}) + \varepsilon \|x_0 - \bar{x}\| \leq 0, \\ F(\bar{x},y) + \varepsilon \|\bar{x} - y\| \quad \text{for all } y \in K,\ y \neq \bar{x}. \end{cases}$$

Motivated by the concept of an ε-solution, introduced in [46], of EP and system of equilibrium problems (in short, SEP), Bianchi et. al. [11] studied the existence of solutions of EP and SEP without any kind of convexity assumption on either set or bifunction involved in the formulation of EP or SEP but in the setting of finite-dimension spaces.

In this survey chapter we present some results from [1, 49, 50]. In particular, we present different forms of Ekeland's variational principle involving τ-functions, τ-functions and fitting functions, and Q-functions, respectively. The equilibrium version of Ekeland-type variational principle is also presented. We give some equivalences of our variational principles with the Caristi–Kirk-type fixed point theorem for multivalued maps, Takahashi minimization theorem, and some other related results. As applications of our results, we derive the existence results for solutions of equilibrium problems and fixed point theorems for multivalued maps. The results of this chapter extend and generalize many results that have appeared recently in the literature.

10.2 τ-Function, Ekeland-Type Variational Principle, and Some Related Results

Throughout the chapter, unless otherwise specified, we denote by $\mathbb{N}$ the set of all natural numbers, $\mathbb{R}$ the set of all real numbers, and $\mathbb{R}_+ = [0,\infty)$. The family of all subsets of X is denoted by 2^X.

The following concept of w-distance is introduced by Kada, Suzuki, and Takahashi [45].

Definition 10.1 ([45]). Let (X,d) be a metric space. A function $p : X \times X \to \mathbb{R}_+$ is said to be a *w-distance* on X if the following conditions are satisfied.

($w1$) For all $x,y,z \in X$, $p(x,z) \le p(x,y) + p(y,z)$.
($w2$) For all $x \in X$, $p(x,\cdot) : X \to \mathbb{R}_+$ is lower semicontinuous.
($w3$) For any $\varepsilon > 0$, there exists $\delta > 0$ such that $p(z,x) \le \delta$ and $p(z,y) \le \delta$ imply $d(x,y) \le \varepsilon$.

It can be easily checked that every metric is a w-distance. Several examples and properties of w-distances are given in [45]. Kada et. al. [45] established the nonconvex minimization theorem, Caristi's fixed point theorem, Ekeland's variational principle, and some fixed point theorems involving a w-distance. The w-distance is further studied and used in [60, 65].

This section deals with the concepts and results that appeared in [49].

Throughout the section, unless specified otherwise, we assume that (X,d) is a metric space and $\varphi : (-\infty,\infty] \to (0,\infty)$ is a nondecreasing function.

An extended real-valued function $f : X \to (-\infty,\infty]$ is said to be

(i) *Lower semicontinuous from above* (in short, lsca) *at* $x_0 \in X$ [19] if for any sequence $\{x_n\}$ in X with $x_n \to x_0$ and $f(x_1) \ge f(x_2) \ge \cdots \ge f(x_n) \ge \cdots$ imply that $f(x_0) \le \lim_{n\to\infty} f(x_n)$.
(ii) *Upper semicontinuous from below* (in short, uscb) *at* $x_0 \in X$ if for any sequence $\{x_n\}$ in X with $x_n \to x_0$ and $f(x_1) \le f(x_2) \le \cdots \le f(x_n) \le \cdots$ imply that $f(x_0) \ge \lim_{n\to\infty} f(x_n)$.

The function f is said to be lsca (respectively, uscb) on X if f is lsca (respectively, uscb) at every point of X. The function f is said to be proper if $f \not\equiv \infty$.

It is obvious that the lower (respectively, upper) semicontinuity implies the lower (respectively, upper) semicontinuity from above (respectively, below), but the reverse is not true (see [19, Example 1.3]).

We introduce the following concept of a τ-function which is different from the definition of a τ-distance studied in [61]. It generalizes the concept of w-distance.

Definition 10.2 ([49]). A function $p : X \times X \to [0,\infty)$ is said to be a τ-*function* if the following conditions hold.

($\tau1$) For all $x,y,z \in X$, $p(x,z) \le p(x,y) + p(y,z)$.
($\tau2$) If $x \in X$ and $\{y_n\}$ in X with $\lim_{n\to\infty} y_n = y$ and $p(x,y_n) \le M$ for some $M = M(x) > 0$ then $p(x,y) \le M$.

($\tau 3$) For any sequence $\{x_n\}$ in X with $\limsup_{n\to\infty}\{p(x_n,x_m) : m > n\} = 0$, and if there exists a sequence $\{y_n\}$ in X such that $\lim_{n\to\infty} p(x_n,y_n) = 0$, then $\lim_{n\to\infty} d(x_n,y_n) = 0$.
($\tau 4$) For $x,y,z \in X$, $p(x,y) = 0$ and $p(x,z) = 0$ imply $y = z$.

It is known that if p is a w-distance on $X \times X$, then for every $x,y,z \in X$, $p(x,y) = 0$ and $p(x,z) = 0$ imply $y = z$. See, for example, [45, 67].

If $p(x,\cdot)$ is lower semicontinuous for each $x \in X$, then condition ($\tau 2$) holds.

Remark 10.1. Every w-distance is a τ-function.

Indeed, let p be a w-distance on $X \times X$. Clearly, ($\tau 1$) and ($\tau 4$) hold. If $x \in X$ and $\{y_n\}$ in X with $\lim_{n\to\infty} y_n = y$ such that $p(x,y_n) \le M$ for some $M = M(x) > 0$, then by ($w2$), $p(x,y) \le \underline{\lim}_{n\to\infty} p(x,y_n) \le M$. Therefore ($\tau 2$) holds. Let $\{x_n\}$ be a sequence in X with $\limsup_{n\to\infty}\{p(x_n,x_m) : m > n\} = 0$ and there exists $\{y_n\}$ in X such that $\lim_{n\to\infty} p(x_n,y_n) = 0$. For any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $p(x_n,x_{n+1}) \le \delta/2$ and $p(x_n,y_n) < \delta/2$ whenever $n \ge n_0$. So $p(x_n,y_{n+1}) \le p(x_n,x_{n+1}) + p(x_{n+1},y_{n+1}) < \delta$ whenever $n \ge n_0$. Then by ($w3$), $d(x_{n+1},y_{n+1}) < \varepsilon$ whenever $n \ge n_0$. Hence $\lim_{n\to\infty} d(x_n,y_n) = 0$ and ($\tau 3$) holds. Therefore, p is a τ-function on $X \times X$.

Lemma 10.1 ([49, Lemma 2.1]). *Let p be a τ-function on $X \times X$. If a sequence $\{x_n\}$ in X with* $\limsup_{n\to\infty}\{p(x_n,x_m) : m > m\} = 0$*, then $\{x_n\}$ is a Cauchy sequence in X.*

Lemma 10.2 ([49, Lemma 2.2]). *Let $f : X \to (-\infty,\infty]$ be a function and p be a τ-function on $X \times X$. For each $x \in X$, let*

$$S(x) = \{y \in X : y \neq x,\ p(x,y) \le \varphi(f(x))(f(x) - f(y))\}.$$

If $S(x)$ is nonempty for some $x \in X$, then for each $y \in S(x)$, we have $f(y) \le f(x)$ and $S(y) \subseteq S(x)$.

Lemma 10.3 ([50, Lemma 2.2]). *Suppose that the function $p : X \times X \to [0,\infty)$ satisfies the conditions ($\tau 1$) and ($\tau 4$) and the function $q : X \times X \to (-\infty,\infty]$ satisfies $q(x,x) \ge 0$ for all $x \in X$ and $q(x,z) \le q(x,y) + q(y,z)$ for all $x,y,z \in X$. For each $x \in X$, let $G : X \to 2^X$ be defined by*

$$G(x) = \{y \in X : y \neq x, p(x,y) + q(x,y) \le 0\}.$$

If $G(x)$ is nonempty for some $x \in X$, then for each $y \in G(x)$, we have $q(x,y) \le 0$ and $G(y) \subseteq G(x)$.

Remark 10.2. For any function $f : X \to (-\infty,\infty]$, the bifunction $q : X \times X \to (-\infty,\infty]$ defined by $q(x,y) = f(y) - f(x)$ satisfies the conditions of Lemma 10.3.

We present an intersection result involving a τ-function and a function which is proper lsca and bounded below. This result plays a key role in the proof of the main result of this section.

Proposition 10.1 ([49, Proposition 2.1]). *Let $f : X \to (-\infty, \infty]$ be a proper lsca and bounded below function and p be a τ-function on $X \times X$. For each $x \in X$, let $S(x)$ be the same as in Lemma* 10.2. *If $\{x_n\}$ is a sequence in X such that $S(x_n)$ is nonempty and $x_{n+1} \in S(x_n)$ for all $n \in \mathbb{N}$, then there exists $x_0 \in X$ such that $x_n \to x_0$ and $x_0 \in \bigcap_{n=1}^{\infty} S(x_n)$.*

Moreover, if $f(x_{n+1}) \leq \inf_{z \in S(x_n)} f(z) + \frac{1}{n}$ for all $n \in \mathbb{N}$, then $\bigcap_{n=1}^{\infty} S(x_n)$ contains precisely one point.

By applying Proposition 10.1, we obtain the following generalization of Ekeland's variational principle for lower semicontinuous from above functions and involving a τ-function.

Theorem 10.4 (Generalized Ekeland's Variational Principle [49, Theorem 2.1]). *Let $f : X \to (-\infty, \infty]$ be a proper lsca and bounded below function and p be a τ-function on $X \times X$. Then there exists $v \in X$ such that $p(v,x) > \varphi(f(v))(f(v) - f(x))$ for all $x \in X$ with $x \neq v$.*

Proof. On the contrary, assume that for each $x \in X$, there exists $y \in X$ with $y \neq x$ such that $p(x,y) \leq \varphi(f(x))(f(x) - f(y))$. For each $x \in X$, let $S(x)$ be the same as in Lemma 10.2. Then $S(x) \neq \emptyset$ for all $x \in X$. Because f is proper, there exists $u \in X$ with $f(u) < \infty$. We define inductively a sequence $\{u_n\}$ in X, starting with $u_1 = u$. Then choose $u_2 \in S(u_1)$ such that $f(u_2) \leq \inf_{x \in S(u_1)} f(x) + 1$. Suppose that $u_n \in X$ is known; then choose $u_{n+1} \in S(u_n)$ such that $f(u_{n+1}) \leq \inf_{x \in S(u_n)} f(x) + 1/n$. From Proposition 10.1, there exists $x_0 \in X$ such that $\bigcap_{n=1}^{\infty} S(u_n) = \{x_0\}$. By Lemma 10.2, $S(x_0) \subseteq \bigcap_{n=1}^{\infty} S(u_n) = \{x_0\}$ and hence $S(x_0) = \{x_0\}$, which is a contradiction. Therefore there exists $v \in X$ such that $p(v,x) > \varphi(f(v))(f(v) - f(x))$ for all $x \in X$ with $x \neq v$.

As a first application of the generalized Ekeland's variational principle, we derive the following generalized Caristi's (common) fixed point theorem for a family of multivalued maps.

Theorem 10.5 (Generalized Caristi's Common Fixed Point Theorem for a Family of Multivalued Maps [49, Theorem 2.2]). *Let p and f be the same as in Theorem* 10.4. *Let I be any index set and for each $i \in I$; let $T_i : X \to 2^X$ be a multivalued map with nonempty values such that for each $x \in X$. There exists $y = y(x,i) \in T_i(x)$ with*

$$p(x,y) \leq \varphi(f(x))(f(x) - f(y)). \tag{10.1}$$

Then there exists $v \in X$ such that $v \in \bigcap_{i \in I} T_i(v)$; that is, the family of multivalued maps $\{T_i\}_{i \in I}$ has a common fixed point in X, and $p(v,v) = 0$.

Proof. From Theorem 10.4, there exists $v \in X$ such that $p(v,x) > \varphi(f(v))(f(v) - f(x))$ for all $x \in X$ with $x \neq v$. We claim that $v \in \bigcap_{i \in I} T_i(v)$ and $p(v,v) = 0$.

By the hypothesis, for each $i \in I$, there exists $w(v,i) \in T_i(v)$ such that $p(v,w(v,i)) \leq \varphi(f(v))(f(v) - f(w(v,i)))$. Then $w(v,i) = v$ for each $i \in I$. Indeed, if $w(v,i_0) \neq v$

for some $i_0 \in I$, then $p(v,w(v,i_0)) \leq \varphi(f(v))(f(v) - f(w(v,i_0))) < p(v,w(v,i_0))$, which leads to a contradiction. Hence $v = w(v,i) \in T_i(v)$ for all $i \in I$. Because $p(v,v) \leq \varphi(f(v))(f(v) - f(v)) = 0$, we obtain $p(v,v) = 0$.

If for each $i \in I$, T_i is a single-valued map, then the following result can be easily derived from the above theorem.

Corollary 10.1 (Generalized Caristi's Common Fixed Point Theorem for a Family of Single-Valued Maps [49, Corollary 2.1]). *Let p and f be the same as in Theorem* 10.4. *Let I be any index set and for each $i \in I$, let $g_i : X \to X$ be a single-valued map such that $p(x,g_i(x)) \leq \varphi(f(x))(f(x) - f(g_i(x)))$ for all $x \in X$. Then there exists $v \in X$ such that $g_i(v) = v$ for each $i \in I$ and $p(v,v) = 0$.*

Remark 10.3.

(a) Corollary 10.1 implies Theorem 10.5.
Indeed, under the hypothesis of Theorem 10.5, for each $x \in X$, there exists $y(x,i) \in T_i(x)$ such that $p(x,y(x,i)) \leq \varphi(f(x))(f(x) - f(y(x,i)))$. For each $i \in I$, we set $g_i(x) = y(x,i)$. Then g_i is a single-valued map from X into itself satisfying $p(x,g_i(x)) \leq \varphi(f(x))(f(x) - f(g_i(x)))$ for all $x \in X$. By Corollary 10.1, there exists $v \in X$ such that $v = g_i(v) \in T_i(v)$ for each $i \in I$ and $p(v,v) = 0$.

(b) Theorem 10.5 implies Theorem 10.4.
Indeed, suppose that for each $x \in X$, there exists $y \in X$ with $y \neq x$ such that $p(x,y) \leq \varphi(f(x))(f(x) - f(y))$. Then for each $x \in X$, we can define a multivalued mapping $T : X \to 2^X \setminus \{\emptyset\}$ by

$$T(x) = \{y \in X : y \neq x,\ p(x,y) \leq \varphi(f(x))(f(x) - f(y))\}.$$

By Theorem 10.5, T has a fixed point $v \in X$; that is, $v \in T(v)$. But $v \notin T(v)$, a contradiction.

In the rest of the section, unless specified otherwise, we assume that (X,d), p, f, and φ are the same as in Theorem 10.4 and I is any index set.

As a second application of Theorem 10.4, we present the following nonconvex maximal element theorem for a family of multivalued maps.

Theorem 10.6 (Nonconvex Maximal Element Theorem for a Family of Multivalued Maps [49, Theorem 2.3]). *For each $i \in I$, let $T_i : X \to 2^X$ be a multivalued map. Assume that for each $(x,i) \in X \times I$ with $T_i(x) \neq \emptyset$, there exists $y = y(x,i) \in X$ with $y \neq x$ such that* (10.1) *holds. Then there exists $v \in X$ such that $T_i(v) = \emptyset$ for each $i \in I$.*

Remark 10.4. Theorem 10.6 implies Theorem 10.4.

Indeed, suppose that for each $x \in X$, there exists $y \in X$ with $y \neq x$ such that $p(x,y) \leq \varphi(f(x))(f(x) - f(y))$. For each $x \in X$, define a multivalued map by

$$T(x) = \{y \in X : y \neq x,\ p(x,y) \leq \varphi(f(x))(f(x) - f(y))\}.$$

Then $T(x) \neq \emptyset$ for all $x \in X$. But from Theorem 10.6, there exists $v \in X$ such that $T(v) = \emptyset$, a contradiction.

The following result is the generalization of the nonconvex minimization theorems studied in [45, 66].

Theorem 10.7 (Generalized Takahashi's Nonconvex Minimization Theorem [49, Theorem 3.1]). *Suppose that for any $x \in X$ with $f(x) > \inf_{z\in X} f(z)$, there exists $y \in X$ with $y \neq x$ such that* (10.1) *holds. Then there exists $v \in X$ such that $f(v) = \inf_{z\in X} f(z)$.*

Remark 10.5. Theorem 10.7 implies Theorem 10.4.

Indeed, suppose that for each $x \in X$, there exists $y \in X$ with $y \neq x$ such that $p(x,y) \leq \varphi(f(x))(f(x) - f(y))$. Then, by Theorem 10.7, there exists $v \in X$ such that $f(v) = \inf_{x\in X} f(x)$. By our supposition, there exists $w \in X$ with $w \neq v$ such that $p(v,w) \leq \varphi(f(v))(f(v) - f(w)) \leq 0$. Hence $p(v,w) = 0$ and $f(v) = f(w) = \inf_{x\in X} f(x)$. There exists $z \in X$ with $z \neq w$ such that $p(w,z) \leq \varphi(f(w))(f(w) - f(z)) \leq 0$. So we also have $p(w,z) = 0$ and $f(v) = f(w) = f(z) = \inf_{x\in X} f(x)$. Because $p(v,z) \leq p(v,w) + p(w,z) = 0$, $p(v,z) = 0$. By condition (τ4), we have $w = z$, which leads to a contradiction.

Remark 10.6. [45, Theorem 1] and [62, Theorem 5] are special cases of Theorem 10.7.

We derive the following nonconvex minimax theorem from Theorem 10.4.

Theorem 10.8 (Nonconvex Minimax Theorem [49, Theorem 3.2]). *Let $F : X \times X \to (-\infty,\infty]$ be a function such that it is proper and lsca and bounded in the first argument. Suppose that for each $x \in X$ with*

$$\{u \in X : F(x,u) > \inf_{a\in X} F(a,u)\} \neq \emptyset,$$

there exists $y = y(x) \in X$ with $y \neq x$ such that

$$p(x,y) \leq \varphi(F(x,w))(F(x,w) - F(y,w)) \quad \textit{for all } w \in X. \tag{10.2}$$

Then $\inf_{x\in X} \sup_{y\in X} F(x,y) = \sup_{y\in X} \inf_{x\in X} F(x,y)$.

Remark 10.7. The convexity assumptions on the sets or on the bifunctions are essential in many existing general topological minimax theorems. McLinden [51] obtained some applications of Ekeland's variational principles to minimax problems in the setting of Banach spaces. The results in [51] are patterned after Rockafellar's augmented version of Ekeland's variational principle, in which additional information of subgradient type is extracted from the basic Ekeland's inequality. Note that the assumption and conclusion of Theorem 10.8 are different from those studied in [51]. Ansari, Lin, and Su [3] and Lin [48] studied minimax theorems for a family of multivalued mappings in locally convex topological vector spaces. Certain convexity assumptions are assumed in [3, 48] and the references therein.

The following result provides the existence of a solution of EP without any kind of convexity assumption on the underlying set and bifunction.

Theorem 10.9 (Nonconvex Equilibrium Theorem [49, Theorem 3.3]). *Let F and φ be the same as in Theorem* 10.8*. Suppose that for each $x \in X$ with $\{u \in X : F(x,u) < 0\} \neq \emptyset$, there exists $y = y(x) \in X$ with $y \neq x$ such that* (10.2) *holds. Then there exists $v \in X$ such that $F(v,y) \geq 0$ for all $y \in X$.*

Example 10.1. Let $X = [0,1]$ and $d(x,y) = |x-y|$. Then (X,d) is a complete metric space. Let a and b be positive real numbers with $a \geq b$. Let $F : X \times X \to \mathbb{R}$ be defined by $F(x,y) = ax - by$. It is easy to see that for each $y \in X$, the function $x \mapsto F(x,y)$ is a proper lsc and bounded below function on X and $F(1,y) \geq 0$ for all $y \in X$. In fact, $F(x,y) \geq 0$ for all $x \in [(b/a),1]$ and all $y \in X$. Note that for each $x \in [0,(b/a))$, $F(x,y) = ax - by < 0$ for all $y \in [(a/b)x,1]$. Hence $\{u \in X : F(x,u) < 0\} \neq \emptyset$ for all $x \in [0,(b/a))$. For any $x \geq y$, $x,y \in X$, we have $x - y = (1/a)\{(ax-bu) - (ay-bu)\}$ for all $u \in X$. Define a nondecreasing function $\varphi : [0,+\infty) \to [0,+\infty)$ by $\varphi(t) = (1/a)$. Hence $d(x,y) \leq \varphi(F(x,u))(F(x,u) - F(y,u))$ for all $x \geq y$, $x,y,u \in X$. By Theorem 10.9, there exists $v \in X$ such that $F(v,y) \geq 0$ for all $y \in X$.

Remark 10.8. Oettli and Théra [54] and Park [56] gave some equilibrium formulations of Ekeland's variational principles. But note that, in [54], the author assumed that

(a) $F(x,z) \leq F(x,y) + F(y,z)$ for all $x,y,z \in X$;
(b) For any $x \in X$, $F(x,\cdot) : X \to (-\infty,\infty]$ is lower semicontinuous;
(c) There exists $x_0 \in X$ such that $\inf_{y \in X} F(x_0,y) > -\infty$;

and, in [54], the authors assumed that $F(x,x) = 0$ for any $x \in X$ in addition to conditions (a) and (b). So Theorem 10.9 is different from the one obtained in [54].

Theorem 10.10 ([49, Theorem 3.4]). *For each $i \in I$, let $T_i : X \to 2^X$ be a multivalued map with nonempty values, $g_i, h_i : X \times X \to \mathbb{R}$ be functions and $\{a_i\}$ and $\{b_i\}$ be families of real numbers. Assume that the following conditions hold.*

(i) For each $(x,i) \in X \times I$, there exists $y = y(x,i) \in T_i(x)$ such that $g_i(x,y) \geq a_i$ and $p(x,y) \leq \varphi(f(x))(f(x) - f(y))$.
(ii) For each $(u,i) \in X \times I$, there exists $w = w(u,i) \in T_i(u)$ such that $h_i(u,w) \leq b_i$ and $p(u,w) \leq \varphi(f(u))(f(u) - f(w))$.

Then there exists $x_0 \in T_i(x_0)$ such that $g_i(x_0,x_0) \geq a_i$ and $h_i(x_0,x_0) \leq b_i$ for all $i \in I$ and $p(x_0,x_0) = 0$.

Remark 10.9.

(a) In Theorem 10.10, if $g_i = h_i = F_i$ and $a_i = b_i = c_i$, then there exists $x_0 \in T_i(x_0)$ such that $F_i(x_0,x_0) = c_i$ for all $i \in I$ and $p(x_0,x_0) = 0$.
(b) In (a), if $T_i(x) = X$ for all $x \in X$, then there exists $x_0 \in X$ such that $F_i(x_0,x_0) = c_i$ for all $i \in I$ and $P_i(x_0,x_0) = 0$.
(c) [2, Theorem 3.1] is a special case of Theorem 10.10.

Remark 10.10. Theorem 10.10 implies Theorem 10.4.

Indeed, assume that for each $x \in X$, there exists $y \in X$ with $y \neq x$ such that $p(x,y) \leq \varphi(f(x))(f(x) - f(y))$. Define a multivalued map $T : X \to 2^X \setminus \{\emptyset\}$ by

$T(x) = \{y \in X : y \neq x\}$ and a function $F : X \times X \to \mathbb{R}$ by $F(x,y) = \chi_{T(x)}(y)$, where χ_A is the characteristic function for an arbitrary set A. Note that $y \in T(x) \Leftrightarrow F(x,y) = 1$. Thus for each $x \in X$, there exists $y \in X$ such that $F(x,y) = 1$ and $p(x,y) \leq \varphi(f(x))(f(x) - f(y))$. By Remark 10.9(c) with $c = 1$, there exists $x_0 \in X$ such that $F(x_0,x_0) = 1$ and $p(x_0,x_0) = 0$. Hence we have $x_0 \in T(x_0)$. This is a contradiction.

Definition 10.3. Let (X,d) be a metric space and $a,b \in X$. Let $\kappa : X \to (0,\infty)$ be a function and p be a w-distance on X. The (p,κ)*-flower petal* $P_\varepsilon(a,b)$ (in short, $P_\varepsilon(a,b,\kappa)$) associated with $\varepsilon \in (0,\infty)$ and $a,b \in X$ is the set

$$P_\varepsilon(a,b,\kappa) = \{x \in X : \varepsilon p(a,x) \leq \kappa(a)(p(b,a) - p(b,x))\}.$$

Obviously, if the w-distance p with $p(a,a) = 0$, then $P_\varepsilon(a,b,\kappa)$ is nonempty.

Lemma 10.4 ([49, Lemma 4.1]). *Let $\varepsilon > 0$ and p be a w-distance on X. Suppose that there exists $u \in X$ such that $f(u) < \infty$ and $p(u,u) = 0$. Then there exists $v \in X$ such that*

(i) $\varepsilon p(u,v) \leq \varphi(f(u))(f(u) - f(v))$.
(ii) $\varepsilon p(v,x) > \varphi(f(v))(f(v) - f(x))$ for all $x \in X$ with $x \neq v$.

Finally, as an application of our results mentioned above, we establish a generalized flower petal theorem.

Theorem 10.11 (Generalized Flower Petal Theorem [49, Theorem 4.1]). *Let M be a proper complete subset of a metric space (X,d) and $a \in M$. Let p be a w-distance on X with $p(a,a) = 0$. Suppose that $b \in X \setminus M$, $p(b,M) = \inf_{x\in M} p(b,x) \geq r$ and $p(b,a) = s > 0$ and there exists a function $\kappa : X \to (0,\infty)$ satisfies $\kappa(x) = \varphi(p(b,x))$ for some nondecreasing function $\varphi : (-\infty,\infty] \to (0,\infty)$. Then for each $\varepsilon > 0$, there exists $v \in M \cap P_\varepsilon(a,b,\kappa)_1$ such that $P_\varepsilon(v,b,\kappa)_1 \cap (M\setminus\{v\}) = \emptyset$. Moreover, $p(a,v) \leq \varepsilon^{-1}\kappa(a)(s-r)$.*

Remark 10.11. Under the assumptions of Theorem 10.11, we cannot verify $v \in P_\varepsilon(v,b,\kappa)_1$, but if we assume that the w-distance p with $p(x,x) = 0$ for all $x \in X$, then for each $\varepsilon > 0$, there exists $v \in M \cap P_\varepsilon(a,b,\kappa)_1$ such that $P_\varepsilon(v,b,\kappa)_1 \cap M = \{v\}$.

In Theorem 10.11, if $k(x) = 1$ for all $x \in X$, we can obtain the primitive flower petal theorem [57].

10.3 Fitting Function, Ekeland-Type Variational Principle, and Some Equivalent Results

Let X be a nonempty set and "$\lesssim$" be a quasi-order (preorder or pseudo-order, i.e., a reflexive and transitive relation) on X. Then $(X,\lesssim)$ is called a *quasi-ordered set*. Let (X,d) be a metric space with a quasi-order $\lesssim$. A nonempty subset M of X is said to be $\lesssim$ *complete* if every nondecreasing Cauchy sequence in M converges. An element v in X is called a *maximal element* of X if there is no element x of X, different from v, such that $v \lesssim x$; that is, $v \lesssim w$ for some $w \in X$ implies that $v = w$.

Throughout the section, unless otherwise specified, we assume that (X,d) is a metric space.

The following concept of a fitting function defined on a topological space is introduced in [50].

Definition 10.4 ([50]). Let X be a topological space. A bifunction $q : X \times X \to (-\infty,\infty]$ is said to be a *fitting function* if the following conditions are satisfied.

(i) For all $x,y,z \in X$, $q(x,z) \le q(x,y) + q(y,z)$.
(ii) For all $x \in X$, $q(x,\cdot)$ is lower semicontinuous.

Obviously, if q_1 and q_2 are fitting functions and $\alpha \ge 0$, then αq_1 and $q_1 + q_2$ are also fitting functions.

Let us give some examples of fitting functions.

Example 10.2. Any w-distance is a fitting function. In particular, a metric is a fitting function.

Example 10.3. Let X be a topological space and $f : X \to (-\infty,\infty]$ be a lower semicontinuous function. Then the function $q : X \times X \to (-\infty,\infty]$ defined by $q(x,y) = f(y) - f(x)$ is a fitting function.

Example 10.4. Let (X,d) be a metric space and $T : X \to X$ be a continuous map. Then it is easy to verify that the functions $q_i : X \times X \to [0,\infty)$, $i = 1,2,3$, defined by

$$q_1(x,y) = \max\{d(x,y), d(x,Ty)\},$$
$$q_2(x,y) = \max\{d(Tx,y), d(Tx,Ty)\}$$

and

$$q_3(x,y) = \max\{d(x,y), d(x,Ty), d(Tx,y), d(Tx,Ty)\}$$

are fitting functions.

Example 10.5. Let $(X,\|\cdot\|)$ be a normed vector space, $f : X \to [0,\infty)$ be any function, $g : X \to [0,\infty)$ be a lower semicontinuous function, and $a \ge 0$. Then the function $q : X \times X \to [0,\infty)$ defined by

$$q(x,y) = a\|x-y\| + f(x) + g(y)$$

is a fitting function. In particular, any constant function on $X \times X$ and the function $q : X \times X \to [0,\infty)$ defined by $q(x,y) = \|x\| + \|y\|$ are fitting functions.

Example 10.6. Let $X = \mathbb{R}$ with the metric $d(x,y) = |x-y|$, $0 < a < b$, and $c \ge 0$. Define the function $q : X \times X \to [0,\infty)$ by

$$q(x,y) = \max\{a(y-x), b(x-y)\} + c.$$

Then q is nonsymmetric. It is easy to see that q is a fitting function.

Example 10.7. Let $X = [0,\infty)$ with the metric $d(x,y) = |x-y|$ and $M = [a,b]$ be a compact subset of X, where $a < b$. Hence M is complete. Let $m > n > 0$. Define the function $q : X \times X \to \mathbb{R}$ by

$$q(x,y) = mx - ny.$$

Then $q(x,x) \geq 0$ for all $x \in X$ and q is a fitting function. Moreover, $\inf_{y\in M} q(x,y) > -\infty$ for each $x \in M$.

The following result is a variant of generalized Ekeland's variational principle and maximal element theorem for τ-functions and fitting functions in the setting of $\lesssim$ complete metric spaces.

Theorem 10.12 ([50, Theorem 2.1]). *Let p be a τ-function such that $p(x,\cdot)$ is lower semicontinuous for each $x \in X$ and $q : X \times X \to (-\infty,\infty]$ be a fitting function such that $q(x,x) \geq 0$ for all $x \in X$. Define a binary relation $\lesssim_{(p,q)}$ on X by*

$$x \lesssim_{(p,q)} y \quad \Leftrightarrow \quad x = y \quad \text{or} \quad p(x,y) + q(x,y) \leq 0.$$

Suppose that there exists a nonempty subset M of X such that

(i) M is $\lesssim_{(p,q)}$ complete.
(ii) There exists $u \in M$ such that $\inf_{y\in M} q(u,y) > -\infty$.

For each $x \in M$, let $S_M : M \to 2^M$ be defined by

$$S_M(x) = \{y \in M : x \lesssim_{(p,q)} y\}.$$

Then

(a) $\lesssim_{(p,q)}$ is a quasi-order induced by p and q.
(b) There exists $v \in M$ such that

(1) v is a maximal element of M.
(2) $S_M(v) = \{v\}$.
(3) $p(v,x) + q(v,x) > 0$ for all $x \in M$ with $x \neq v$.

The following result immediately follows from Theorem 10.12.

Corollary 10.2 ([50, Corollary 2.1]). *Let p be a τ-function such that $p(x,\cdot)$ is lower semicontinuous for each $x \in X$ and $f : X \to (-\infty,\infty]$ be a lower semicontinuous function. Define a binary relation $\lesssim_{(p,f)}$ on X by*

$$x \lesssim_{(p,f)} y \quad \Leftrightarrow \quad x = y \ \text{ or } \ p(x,y) \leq f(x) - f(y).$$

Suppose that there exists a nonempty subset M of X such that

$(iii)_1$ M is $\lesssim_{(p,f)}$ complete.
$(iv)_1$ There exists $u \in M$ such that $f(u) - \inf_{y\in M} f(y) < \infty$.

For each $x \in M$, let $S_M : M \to 2^M$ be defined by

$$S_M(x) = \{y \in M : x \precsim_{(p,f)} y\}.$$

Then

(a) $\precsim_{(p,f)}$ is a quasi-order induced by p, and f.
(b) There exists $v \in M$ such that

(1) v is a maximal element of M.
(2) $S_M(v) = \{v\}$.
(3) $p(v,x) > f(v) - f(x)$ for all $x \in M$ with $x \neq v$.

Theorem 10.13 ([50, Theorem 2.2]). *Let p and q be the same as in Theorem* 10.12. *Let $\varepsilon > 0$ and $\lambda > 0$ be given. Define a binary relation $\precsim_{(\varepsilon,\lambda,p,q)}$ on X by*

$$x \precsim_{(\varepsilon,\lambda,p,q)} y \quad \Leftrightarrow \quad x = y \quad or \quad \varepsilon\lambda^{-1} p(x,y) + q(x,y) \leq 0.$$

Suppose that there exists a nonempty subset M of X such that

$(iii)_2$ M is $\precsim_{(\varepsilon,\lambda,p,q)}$ complete.
$(iv)_2$ There exists $u \in M$ such that $p(u,u) = q(u,u) = 0$ and $\inf_{y \in M} q(u,y) \geq -\varepsilon$.

For each $x \in M$, let $S_M : M \to 2^M$ be defined by

$$S_M(x) = \{y \in M : x \precsim_{(\varepsilon,\lambda,p,q)} y\}.$$

Then

(a) $\precsim_{(\varepsilon,\lambda,p,q)}$ is a quasi-order induced by ε, λ, p, and q.
(b) There exists $v \in M$ such that

(1) v is a maximal element of M.
(2) $S_M(v) = \{v\}$.
(3) $p(u,v) \leq \lambda$.
(4) $-\varepsilon \leq q(u,v) \leq 0$.
(5) $\varepsilon\lambda^{-1} p(u,v) + q(u,v) \leq 0$.
(6) $\varepsilon\lambda^{-1} p(v,x) + q(v,x) > 0$ for all $x \in M$ with $x \neq v$.

We now present some equivalent formulations of Theorem 10.12.

Theorem 10.14 ([50, Theorem 3.1]). *Under the same assumption of Theorem* 10.4, *the following statements are equivalent.*

(i) *There exists $z \in M$ such that $S_M(z) = \{z\}$.*
(ii) (Maximal Element Theorem). *There exists a maximal element in M.*
(iii) (Generalized Ekeland's Variational Principle). *There exists $v \in M$ such that $p(v,x) + q(v,x) > 0$ for all $x \in M$ with $x \neq v$.*

(iv) (Generalized Caristi's Common Fixed Point Theorem for a Family of Multivalued Maps). *Let I be an index set. For each $i \in I$, let $T_i : M \to 2^X$ be a multivalued map with nonempty values such that for each $x \in M$, there exists $y = y(x,i) \in T_i(x)$ with $x \lesssim_{(p,q)} y$. Then $\{T_i\}_{i\in I}$ has a common fixed point in M.*

(v) (Generalized Caristi's Common Fixed Point Theorem for a Family of Single-Valued Maps). *Let I be an index set. For each $i \in I$, let $T_i : M \to X$ be a single-valued map such that $x \lesssim_{(p,q)} T_i(x)$ for all $x \in M$. Then $\{T_i\}_{i\in I}$ has a common fixed point in M.*

(vi) (Common Fixed Point Theorem for a Family of Multivalued Maps). *Let I be an index set. For each $i \in I$, let $T_i : M \to 2^X$ be a multivalued map with nonempty values such that for each $x \in M$ with $x \notin T_i(x)$, there exists $y = y(x,i) \in M$ with $y \neq x$ such that $x \lesssim_{(p,q)} y$. Then $\{T_i\}_{i\in I}$ has a common fixed point in M.*

(vii) (Common Stationary Point Theorem). *Let I be an index set. For each $i \in I$, let $T_i : M \to 2^X$ be a multivalued map with nonempty values such that for each $x \in M$. $x \lesssim_{(p,q)} y$ for all $y \in T_i(x)$. Then $\{T_i\}_{i\in I}$ has a common stationary point x_0 in M; that is, $T_i(x_0) = \{x_0\}$ for each $i \in I$.*

(viii) (Nonconvex Maximal Element Theorem for a Family of Multivalued Maps). *Let I be an index set. For each $i \in I$, let $T_i : M \to 2^X$ be a multivalued map. Suppose that for each $(x,i) \in M \times I$ with $T_i(x) \neq \emptyset$, there exists $y = y(x,i) \in M$ with $y \neq x$ such that $x \lesssim_{(p,q)} y$. Then there exists $x_0 \in M$ such that $T_i(x_0) = \emptyset$ for all $i \in I$.*

We establish another type of Ekeland's variational principle involving the τ-function and a maximal element theorem that are different from Theorem 10.12 in general.

Theorem 10.15 ([50, Theorem 4.1]). *Let $f : X \to (-\infty,\infty]$ be a lsca function, $\varphi : (-\infty,\infty] \to (0,\infty)$ be a nondecreasing function, and p be a τ-function on X. Define a binary relation $\lesssim_{(p,f,\varphi)}$ on X by*

$$x \lesssim_{(p,f,\varphi)} y \quad x = y \quad or \quad p(x,y) \leq \varphi(f(x))(f(x) - f(y)).$$

Suppose that there exists a nonempty subset M of X such that

(i) M is $\lesssim_{(p,f,\varphi)}$ complete.
(ii) There exists $u \in M$ such that $f(u) < \infty$.
(iii) f is bounded below on M.

For each $x \in M$, let $S_M : M \to 2^M$ be defined by

$$S_M(x) = \{y \in M : x \lesssim_{(p,f,\varphi)} y\}.$$

Then

(a) $\lesssim_{(p,f,\varphi)}$ is a quasi-order induced by p, q, and φ;
(b) There exists $v \in M$ such that

(1) v is a maximal element of M.
(2) $S_M(v) = \{v\}$.
(3) $p(v,x) > \varphi(f(v))(f(v) - f(x))$ for all $x \in M$ with $x \neq v$.

Remark 10.12.

(a) Theorem 10.4 is a special case of Theorem 10.15.
(b) Theorem 10.15 is different from Theorem 10.12 in the following ways.

(1) In Theorem 10.12, we assumed that p is a τ-function such that $p(x,\cdot)$ is lower semicontinuous for all $x \in X$, but in Theorem 3.4, we only assumed that p is a τ-function such that $p(x,\cdot)$ need not be lower semicontinuous for all $x \in X$.
(2) In Theorem 10.15, we assumed that $f : X \to (-\infty,\infty]$ is a lsca function and $\varphi : (-\infty,\infty] \to (0,\infty)$ is a nondecreasing function. So the function $q : X \times X \to (-\infty,\infty]$ defined by $q(x,y) = \varphi(f(x))(f(y) - f(x))$ does not satisfy conditions (ii) and (iii) of Theorem 10.12 in general.
(3) The quasi-orders in Theorem 10.12 and Theorem 10.15 are different in general.
(4) The conclusion "$p(v,x) > \varphi(f(v))(f(v) - f(x))$ for all $x \in M$ with $x \neq v$" in Theorem 10.15 is different from the conclusion "$p(v,x) + q(v,x) > 0$ for all $x \in M$ with $x \neq v$" in Theorem 10.12 in general.

Theorem 10.16 ([50, Theorem 4.2]). *Let (X,d) be a complete metric space, $f : X \to (-\infty,\infty]$ be a proper lower semicontinuous and bounded below function and $\varphi : (-\infty,\infty] \to (0,\infty)$ be a nondecreasing function. Let p be a τ-function and $\varepsilon > 0$. Suppose that there exists $u \in X$ such that $p(u,\cdot)$ is lower semicontinuous, $f(u) \leq \inf_{x\in X} f(x) + \varepsilon$ and $p(u,u) = 0$. Then there exists $v \in X$ such that*

(a) $p(u,v) \leq \varphi(f(u))$.
(b) $f(v) \leq \inf_{x\in X} f(x) + \varepsilon$.
(c) $0 \leq f(u) - f(v) \leq \varepsilon$.
(d) $\varepsilon p(u,v) \leq \varphi(f(u))(f(u) - f(v))$.
(e) $\varepsilon p(v,x) > \varphi(f(v))(f(v) - f(x))$ for all $x \in X$ with $x \neq v$.

We have the following nonconvex minimax theorem involving a τ-function and a fitting function in the setting of complete metric spaces.

Theorem 10.17 (Nonconvex Minimax Theorem [50, Theorem 6.1]). *Let (X,d) be a complete metric space, p be a τ-function such that $p(x,x) = 0$, and $p(x,\cdot)$ is lower semicontinuous for all $x \in X$, and $q : X \times X \to (-\infty,\infty]$ be a fitting function such that $q(x,x) \leq 0$ for all $x \in X$ and there exists $u \in X$ such that $\inf_{y\in X} q(u,y) > -\infty$. Suppose that for each $x \in X$ with $\{u \in X : q(x,u) > \inf_{a\in X} q(a,u)\} \neq \emptyset$, there exists $y = y(x) \in X$ with $y \neq x$ such that $p(x,y) + q(x,y) \leq 0$. Then $\inf_{x\in X} \sup_{y\in X} q(x,y) = \sup_{y\in X} \inf_{x\in X} q(x,y)$.*

Finally, we have an existence result for a solution of EP involving the fitting function and in the setting of complete metric spaces.

Theorem 10.18 (Nonconvex Equilibrium Theorem [50, Theorem 6.4]). *Let (X,d), p, and q be the same as in Theorem* 10.17. *Suppose that for each $x \in X$ with $\{u \in X : q(x,u) < 0\} \neq \emptyset$, there exists $y = y(x) \in X$ with $y \neq x$ such that $p(x,y) + q(x,y) \leq 0$. Then there exists $x_0 \in X$ such that $q(x_0,y) \geq 0$ for all $y \in X$.*

10.4 *Q*-Functions, Ekeland-Type Variational Principle, and Related Results

The concept of a quasi-metric space generalizes the concept of a metric space by lifting the symmetry condition. For the quasi-metric space (X,d), the concepts of Cauchy sequences, convergent sequences, and completeness can be defined in the same manner as in the setting of metric spaces.

Throughout the section, unless otherwise specified, we assume that X is a quasi-metric space with the quasi-metric d.

We introduce the concept of a Q-function on a quasi-metric space X.

Definition 10.5 ([1]). A function $q : X \times X \to \mathbb{R}_+$ is called a *Q-function* on X if the following conditions are satisfied.

(Q1) For all $x,y,z \in X$, $q(x,z) \leq q(x,y) + q(y,z)$.
(Q2) If $x \in X$ and $\{y_n\}_{n\in\mathbb{N}}$ is a sequence in X such that it converges to a point y (with respect to the quasi-metric) and $q(x,y_n) \leq M$ for some $M = M(x) > 0$, then $q(x,y) \leq M$.
(Q3) For any $\varepsilon > 0$, there exists $\delta > 0$ such that $q(x,y) \leq \delta$ and $q(x,z) \leq \delta$ imply $d(y,z) \leq \varepsilon$.

Remark 10.13. If (X,d) is a metric space and in addition to (Q1)–(Q3), the following condition is also satisfied,

(Q4) For any sequence $\{x_n\}_{n\in\mathbb{N}}$ in X with $\lim_{n\to\infty} \sup\{q(x_n,x_m) : m > n\} = 0$, and if there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ in X such that $\lim_{n\to\infty} q(x_n,y_n) = 0$, then $\lim_{n\to\infty} d(x_n,y_n) = 0$,

then the definition of a Q-function reduces to the definition of a τ-function. We have seen in Section 10.2 that every w-distance is a τ-function. In fact, if we consider (X,d) as a metric space and replace (Q2) by the following condition,

(Q5) For any $x \in X$, the function $q(x,\cdot) : X \to \mathbb{R}_+$ is lower semicontinuous,

then the definition of a Q-function becomes the definition of a w-distance. It is easy to see that if $q(x,\cdot)$ is lower semicontinuous, then (Q2) holds. Hence, it is obvious that every w-function is a τ-function and every τ-function is a Q-distance but converse assertions do not hold.

Example 10.8.

(a) Let $X = \mathbb{R}$. Suppose $d : X \times X \to \mathbb{R}_+$ is defined as

$$d(x,y) = \begin{cases} 0 & \text{if } x = y \\ |y| & \text{otherwise} \end{cases}$$

and $q : X \times X \to \mathbb{R}_+$ is defined as

$$q(x,y) = |y| \quad \text{for all } x,y \in X.$$

Then it is easy to see that d is a quasi-metric on X and q is a Q-function on X. But q is neither a τ-function nor a w-distance.

(b) Let $X = [0,1]$. Suppose $d : X \times X \to \mathbb{R}_+$ is defined as

$$d(x,y) = \begin{cases} y - x & \text{if } y \geq x \\ 2(x-y) & \text{otherwise} \end{cases}$$

and $q : X \times X \to \mathbb{R}_+$ is defined as

$$q(x,y) = |x - y| \quad \text{for all } x,y \in X.$$

Then q is a Q-function on X. However q is neither a τ-function nor a w-distance, because (X,d) is not a metric space.

Remark 10.14. Let (X,d) be a quasi-metric space and q be a Q-function on X. If $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ is a nondecreasing and subadditive function such that $\eta(0) = 0$, then $\eta \circ q$ is a Q-function on X.

We present some properties of a Q-function that are similar to the properties of a w-distance.

Lemma 10.5 ([1, Lemma 2.1]). *Let $q : X \times X \to \mathbb{R}_+$ be a Q-function on X and, $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ be sequences in X. Let $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\beta_n\}_{n\in\mathbb{N}}$ be sequences in $\mathbb{R}_+$ such that they converge to 0, and let $x,y,z \in X$. Then the following conditions hold.*

(i) If $q(x_n,y) \leq \alpha_n$ and $q(x_n,z) \leq \beta_n$ for all $n \in \mathbb{N}$, then $y = z$. In particular, if $q(x,y) = 0$ and $q(x,z) = 0$, then $y = z$.
(ii) If $q(x_n,y_n) \leq \alpha_n$ and $q(x_n,y) \leq \beta_n$ for all $n \in \mathbb{N}$, then $\{y_n\}_{n\in\mathbb{N}}$ converges to y.
(iii) If $q(x_n,x_m) \leq \alpha_n$ for all $n,m \in \mathbb{N}$ with $m > n$, then $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence.
(iv) If $q(y,x_n) \leq \alpha_n$ for all $n \in \mathbb{N}$, then $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence.
(v) If $q_1,q_2,\ldots,q_n$ are Q-functions on X, then $q(x,y) = \max\{q_1(x,y), q_2(x,y),\ldots, q_n(x,y)\}$ is also a Q-function on X.

The proof of this lemma lies on the lines of the proof of Lemma 1 in [45] and therefore we omit it.

Definition 10.6. A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be *lower monotone* if for any sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ converging to some point $x \in X$ and satisfying $f(x_{n+1}) \leq f(x_n)$ for all $n \in \mathbb{N}$, we have $f(x) \leq f(x_n)$ for each $n \in \mathbb{N}$.

Remark 10.15. Note that the lower monotonicity is slightly weaker than lower semicontinuity. In other words, every lower monotone function is lower semicontinuous but the converse is not true in general.

Remark 10.16. From the definitions of lower monotone and lsca of a function, it is clear that every lower monotone function is lsca but the converse may not be true in general. But, if f is bounded below then both concepts are equivalent.

Indeed, let $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ be a sequence such that it converges to some point x and $f(x_{n+1}) \leq f(x_n)$ for all $n \in \mathbb{N}$. We claim that $f(x) \leq f(x_n)$ for all $n \in \mathbb{N}$, if f is lsca and bounded below.

Because f is bounded below and $f(x_{n+1}) \leq f(x_n)$ for all $n \in \mathbb{N}$, $\lim_{n\to\infty} f(x_n)$ exists. Let $r = \lim_{n\to\infty} f(x_n) = \inf_{n\in\mathbb{N}} f(x_n)$, then $f(x_n) \geq r$ for all $n \in \mathbb{N}$. The lsca of f implies that $f(x) \leq \lim_{n\to\infty} f(x_n) = r$ and thus $f(x) \leq r \leq f(x_n)$ for all $n \in \mathbb{N}$.

Definition 10.7. Let X be an ordered space with an ordering $\preccurlyeq$ on X.

(i) The ordering $\preccurlyeq$ on X is called a *quasi-order* on X if it is a reflexive and transitive relation.
(ii) A sequence $\{x_n\}_{n\in\mathbb{N}}$ in X is called *decreasing* (with respect to $\preccurlyeq$) if $x_{n+1} \preccurlyeq x_n$ for all $x \in \mathbb{N}$.
(iii) The quasi-order $\preccurlyeq$ on X is said to be *lower closed* if for every $x \in X$, the section $S(x) = \{y \in X : y \preccurlyeq x\}$ is lower closed; that is, if $\{x_n\}_{n\in\mathbb{N}} \subseteq S(x)$ is decreasing with respect to $\preccurlyeq$ and convergent to $\tilde{x} \in X$ with respect to the quasi-metric on X, then $\tilde{x} \in S(x)$.

Definition 10.8. Let (X,d) be a quasi-metric with a quasi-order $\preccurlyeq$ on X. For any $x \in X$, the set $S(x) = \{y \in X : y \preccurlyeq x\}$ is said to be $\preccurlyeq$*-complete* if every decreasing (with respect to $\preccurlyeq$) Cauchy sequence in $S(x)$ converges in it.

We present two generalizations of the Ekeland-type variational principle for a Q-function, one in the setting of incomplete quasi-metric spaces and the other in the setting of complete quasi-metric spaces.

Theorem 10.19 ([1, Theorem 3.1]). *Let (X,d) be a quasi-metric space (not necessarily complete), $q : X \times X \to \mathbb{R}_+$ a Q-function on X, $\varphi : (-\infty,\infty] \to (0,\infty)$ a nondecreasing function and $f : X \to \mathbb{R} \cup \{+\infty\}$ a proper and bounded below function. Define a quasi-order $\preccurlyeq$ on X as*

$$y \preccurlyeq x \quad \Leftrightarrow \quad x = y \quad \text{or} \quad q(x,y) \leq \varphi(f(x))(f(x) - f(y)). \tag{10.3}$$

Suppose that there exists $\hat{x} \in X$ such that $\inf_{x\in X} f(x) < f(\hat{x})$ and $S(\hat{x}) = \{y \in X : y \preccurlyeq \hat{x}\}$ is $\preccurlyeq$-complete. Then there exists $\bar{x} \in X$ such that

(a) $q(\hat{x},\bar{x}) \leq \varphi(f(\hat{x}))(f(\hat{x}) - f(\bar{x}))$
(b) $q(\bar{x},x) > \varphi(f(\bar{x}))(f(\bar{x}) - f(x))$ *for all $x \in X$, $x \neq \bar{x}$.*

Remark 10.17. In Theorem 10.19, we did not assume any kind of lower semi-continuity. Instead, we assumed that the set $S(\hat{x})$ is $\preccurlyeq$-complete.

The following result is a simplified form of Theorem 10.19.

Theorem 10.20 ([1, Theorem 3.2]). *Let (X,d) be a complete quasi-metric space, $q : X \times X \to \mathbb{R}_+$ a Q-function on X, $\varphi : (-\infty,\infty] \to (0,\infty)$ a nondecreasing function and $f : X \to \mathbb{R} \cup \{+\infty\}$ a proper, lsca, and bounded below function. Assume that there exists $\hat{x} \in X$ such that $\inf_{x\in X} f(x) < f(\hat{x})$, then there exists $\bar{x} \in X$ such that*

(a) $q(\hat{x},\bar{x}) \leq \varphi(f(\hat{x}))(f(\hat{x}) - f(\bar{x}))$
(b) $q(\bar{x},x) > \varphi(f(\bar{x}))(f(\bar{x}) - f(x))$ for all $x \in X$, $x \neq \bar{x}$.

Remark 10.18.

(a) Theorems 10.19 and 10.20 extend and generalize Theorem 2.1, Theorem 2.1 in [19], Theorem 1.1 in [24], Theorem 1 in [25], Theorem 3 in [45], and Theorem 3 in [56]; See also the references therein.
(b) In [41], Hamel established an Ekeland-type variational principle (similar to Theorem 10.20) in the setting of uniform spaces generated by a family of quasi-metrics. He proved his results for sequentially lower monotone functions and for a quasi-metric. The above Theorems 10.19 and 10.20 are proved for lsca functions which are more general than lower monotone functions, and for a Q-function. As shown in the examples below that the concept of a Q-function and a quasi-metric are not comparable, the results of this section are different from those considered in [41].

Example 10.9. Let $(X, \|\cdot\|)$ be a normed space. Then a function $q : X \times X \to \mathbb{R}_+$ defined as $q(x,y) = \|y\|$ for all $x,y \in X$, is a Q-function. But it is not a quasi-metric on X.

Example 10.10. Let $X = \mathbb{R}$. Define a function $d : X \times X \to \mathbb{R}_+$ as

$$d(x,y) = \begin{cases} 0 & \text{if } x = y \\ |x| & \text{otherwise.} \end{cases}$$

Then d is a quasi-metric on X but it is not a Q-function. We remark that every metric d is a Q-function.

Corollary 10.3. *Let X, q, f, and φ be the same as in Theorem* 10.20. *Let $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing and subadditive function such that $\eta(0) = 0$. Assume that there exists $\hat{x} \in X$ such that $\inf_{x\in X} f(x) < f(\hat{x})$; then there exists $\bar{x} \in X$ such that*

(a) $\eta(q(\hat{x},\bar{x})) \leq \varphi(f(\hat{x}))(f(\hat{x}) - f(\bar{x}))$
(b) $\eta(q(\bar{x},x)) > \varphi(f(\bar{x}))(f(\bar{x}) - f(x))$ for all $x \in X$, $x \neq \bar{x}$.

Remark 10.19. Very recently, Bosch, Garcia, and Garcia [13] established a result similar to Corollary 10.3 but for a Minkowski gauge and in the setting of locally complete spaces. In addition to our assumptions on η, they also assumed that it is continuous. Therefore, the main result in [13] and Corollary 10.3 are not comparable.

We now present the Caristi–Kirk-type fixed point theorem, Takahashi's minimization theorem, and an equilibrium version of the Ekeland-type variational principle for a Q-function in the setting of complete quasi-metric spaces. We also establish the equivalences among these results and Theorem 10.20.

Theorem 10.21 ([1, Theorem 4.1]). *Let (X,d) be a complete quasi-metric space, $q : X \times X \to \mathbb{R}_+$ a Q-function on X, $\varphi : (-\infty,\infty] \to (0,\infty)$ a nondecreasing function, and $f : X \to \mathbb{R} \cup \{+\infty\}$ a proper, lsca, and bounded below function. Then the following statements are equivalent reformulations of Theorem* 10.20.

(i) (Caristi–Kirk-Type Fixed Point Theorem). *Let $T : X \to 2^X$ be a multivalued map with nonempty values. If the condition*

$$\text{for all } y \in T(x) : q(x,y) \leq \varphi(f(x))(f(x) - f(y))$$

is satisfied, then T has an invariant point in X; that is, there exists $\bar{x} \in X$ such that $\{\bar{x}\} = T(\bar{x})$.
If the condition

$$\text{there exists } y \in T(x) : q(x,y) \leq \varphi(f(x))(f(x) - f(y))$$

is satisfied, then T has a fixed point in X, that is, there exists $\bar{x} \in X$ such that $\bar{x} \in T(\bar{x})$.

(ii) (Takahashi's Minimization Theorem). *Assume that for each $\hat{x} \in X$ with $\inf_{z\in X} f(z) < f(\hat{x})$, there exists $x \in X$ such that*

$$x \neq \hat{x} \quad \text{and} \quad q(\hat{x},x) \leq \varphi(f(\hat{x}))(f(\hat{x}) - f(x)).$$

Then there exists $\bar{x} \in X$ such that $f(\bar{x}) = \inf_{y\in X} f(y)$.

(iii) (Equilibrium Version of Ekeland-Type Variational Principle). *Let $F : X \times X \to \mathbb{R} \cup \{+\infty\}$ be a function satisfying the following conditions.*

(E1) For all $x,y,z \in X$, $F(x,z) \leq F(x,y) + F(y,z)$.
(E2) For each fixed $x \in X$, the function $F(x,\cdot) : X \mapsto \mathbb{R} \cup \{+\infty\}$ is proper and lsca.
(E3) There exists $\hat{x} \in X$ such that $\inf_{x\in X} F(\hat{x},x) > -\infty$.

Then, there exists $\bar{x} \in X$ such that

(aa) $\varphi(f(\hat{x}))F(\hat{x},\bar{x}) + q(\hat{x},\bar{x}) \leq 0$.
(bb) $\varphi(f(\bar{x}))F(\bar{x},x) + q(\bar{x},x) > 0$ *for all $x \in X$, $x \neq \bar{x}$.*

Remark 10.20.

(a) Hamel [41] proved similar results to Theorem 10.21 for sequentially lower monotone functions and in the setting of uniform spaces generated by a family of quasi-metrics. Theorem 10.21 is proved for lsca functions that are more general than lower monotone functions, and for a Q-function. As we have seen above that Q functions and quasi-metrics are not comparable, the results of this paper are different from those considered in [41].

(b) Theorem 10.21(i) generalizes Theorem 2.2 in [8], Theorem $(2.1)'$ in [15], and a result in [16].
(c) Theorem 10.21 (ii) extends and generalizes Theorem 1 in [45] and [66].
(d) Theorem 10.21 (iii) generalizes Theorem 10.4.

Corollary 10.4 ([1, Corollary 4.1]). *Let X, q, f, and φ be the same as in Theorem* 10.21. *Let $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing and subadditive function such that $\eta(0) = 0$ and let $T : X \to 2^X$ be a multivalued map with nonempty values. If for all $x \in X$, there is a $y \in T(x)$ satisfying*

$$\eta(q(x,y)) \leq \varphi(f(x))(f(x) - f(y)),$$

then T has a fixed point in X.

Remark 10.21.

(a) Corollary 10.4 generalizes Theorem 4.2 in [31] in the following ways.

 (i) X is a complete quasi-metric space in Corollary 10.4 and it is a complete metric space in Theorem 4.2 in [31].
 (ii) f is bounded and lsca in Corollary 10.4 and is bounded below and lower semicontinuous in Theorem 4.2 in [31].
 (iii) In Corollary 10.4, η is not necessarily continuous.

(b) Corollary 10.4 also generalizes and extends Theorem 3.17 in [47] in several ways.

By using the same arguments as in the proof of Theorem 10.5, a common fixed point theorem can be easily derived for a family of multivalued maps similar to Theorem 10.5. The proof is straightforward, therefore we do not mention it here.

As a particular case of Theorem 10.21(iii), we derive the following result by taking $\varphi(f(x)) = 1/\varepsilon$ for all $x \in X$ and for any given $\varepsilon > 0$.

Corollary 10.5 (Equilibrium Version of Ekeland-Type Variational Principle [1, Corollary 4.2]). *Let (X,d) be a complete quasi-metric space and $q : X \times X \to \mathbb{R}_+$ be a Q-function on X. Let $F : X \times X \to \mathbb{R}$ be a function satisfying the following conditions.*

(E1) For all $x,y,z \in X$, $F(x,z) \leq F(x,y) + F(y,z)$.
(E2) For each fixed $x \in X$, the function $F(x,\cdot) : X \mapsto \mathbb{R}$ is lsca and bounded below.

Then, for any $\varepsilon > 0$ and for any $\hat{x} \in X$, there exists $\bar{x} \in X$ such that

(aa) $F(\hat{x},\bar{x}) + \varepsilon q(\hat{x},\bar{x}) \leq 0$.
(bb) $F(\bar{x},x) + \varepsilon q(\bar{x},x) > 0 \quad$ for all $x \in X$, $x \neq \bar{x}$.

Remark 10.22. Corollary 10.5 can be seem as an extension of Theorem 2.1 of Bianchi, Rassay, and Pini [11] to the setting of complete quasi-metric spaces and for a Q-function.

We present some existence results for a solution of equilibrium problems without convexity assumptions.

Let K be a nonempty subset of a metric space X and let $F : K \times K \to \mathbb{R}$ be a real-valued function. Recall the *equilibrium problem* (in short, EP) of finding $\bar{x} \in K$ such that

$$F(\bar{x}, y) \geq 0 \quad \text{for all } y \in K.$$

Definition 10.9. Let K be a nonempty subset of a metric space X, $F : K \times K \to \mathbb{R}$ a real-valued function and q a Q-function on X. Let $\varepsilon > 0$ be given. A point $\bar{x}$ is called an ε-*solution* of EP if

$$F(\bar{x}, y) + \varepsilon q(\bar{x}, y) \geq 0 \quad \text{for all } y \in K. \tag{10.4}$$

It is called a *strictly* ε-*solution* of EP if the inequality in (10.4) is strict for all $x \neq y$.

We note that Corollary 10.5(bb) gives the existence of a strict ε-solution of EP for any $\varepsilon > 0$.

We establish the existence of a solution of the equilibrium problem without any convexity assumption.

Theorem 10.22 ([1, Theorem 5.1]). *Let K be a nonempty compact subset of a complete metric space X and q be a Q-function on X. Let $F : K \times K \to \mathbb{R}$ be a real-valued function satisfying the following conditions.*

($E1$) For all $x, y, z \in K$, $F(x, z) \leq F(x, y) + F(y, z)$.
($E2$) For each fixed $x \in K$, the function $F(x, \cdot) : K \mapsto \mathbb{R}$ is lsca and bounded below.
($E3$) For each fixed $y \in K$, the function $F(\cdot, y) : K \mapsto \mathbb{R}$ is upper semicontinuous.

Then there exists a solution $\bar{x} \in K$ of EP.

On the lines of Theorem 4.1 in [11] we can easily derive the existence results for a solution of EP when K is not necessarily compact.

Remark 10.23.

(a) Theorem 10.22 generalizes Proposition 3.2 in [11] in the following ways.

  (i) In Theorem 10.22, we did not assume that $F(x, x) = 0$ for all $x \in X$.
  (ii) In Theorem 10.22, $F(x, \cdot)$ is lsca, whereas it is lower semicontinuous in [11].

(b) We notice that the product of n complete quasi-metric spaces (X_i, d_i) is a complete quasi-metric space (X, d), where $X = \prod_{i=1}^{n} X_i, d(x, y) = \max_{1 \leq i \leq n} \{d_1(x_1, y_1), \ldots, d_n(x_n, y_n)\}$, $x = (x_1, x_2, \ldots, x_n) \in X$ and $y = (y_1, y_2, \ldots, y_n) \in X$. By Lemma 10.5(v), $q(x, y) = \max_{1 \leq i \leq n}\{q_1(x_1, y_1), \ldots, q_n(x_n, y_n)\}$ is a Q-function on X, where q_i is a Q-function on X_i for all $i = 1, 2, \ldots, n$. Therefore, Theorem 2.2 in [11] can be easily extended for complete quasi-metric spaces and Q-functions as Theorem 10.22.

Definition 10.10. Let (X,d) be a complete quasi-metric space and q be a Q-function on X. We say that $x_0 \in X$ *satisfies Condition (A)* if and only if every sequence $\{x_n\} \subseteq X$ satisfying $F(x_0,x_n) \leq 1/n$ for all $n \in \mathbb{N}$ and $F(x_n,x)+(1/n)q(x_n,x) \geq 0$ for all $x \in X$ and $n \in \mathbb{N}$, has a convergent subsequence.

This definition is introduced and considered by Oettli and Théra [54] in the setting of a complete metric space.

The following result provides the existence of a solution of EP under condition (A) but without the compactness assumption.

Theorem 10.23 ([1, Theorem 5.2]). *Let (X,d) be a complete quasi-metric space, q a Q-function on X, and $F : X \times X \to \mathbb{R}$ satisfy conditions $(E1)$–$(E2)$ of Corollary* 10.4 *and be upper semicontinuous in the first argument. If some $x_0 \in X$ satisfies Condition (A), then there exists a solution $\bar{x} \in X$ of EP.*

Remark 10.24.

(a) Theorem 10.16 extends Theorem 6(a) in [54] for a Q-function and in the setting of complete quasi-metric spaces.
(b) In Theorems 10.22 and 10.23, we have not assumed that $F(x,x) = 0$ for all $x \in X$. This assumption, some kind of convexity condition on the underlying function F, and convexity structure on the underlying set K are required in almost all the results on the existence of a solution of EP appearing in the literature; see, for example, [4, 9–12, 14, 17, 18, 27, 28, 32, 33, 37] and the references therein. But in Theorems 10.19 and 10.20, neither any kind of convexity condition is required on the function F nor a convexity structure on the set K. Therefore, the results of this section are new in the literature.

The following theorem provides the equivalence among the equilibrium version of the Ekeland-type variational principle, equilibrium problem, Caristi–Kirk-type fixed point theorem and Oettli and Théra-type theorem.

Theorem 10.24 ([1, Theorem 5.3]). *Let (X,d) be a complete quasi-metric space and $q : X \times X \to \mathbb{R}_+$ be a Q-function on X. Let $F : X \times X \to \mathbb{R}$ be a function satisfying the conditions $(E1)$ and $(E2)$ of Corollary* 10.5. *Then the following statements are equivalent.*

(i) (Equilibrium Form of Ekeland-Type Variational Principle). *For every $\hat{x} \in X$, there exists $\bar{x} \in X$ such that*

$$\bar{x} \in \hat{S} := \{x \in X : F(\hat{x},x) + q(\hat{x},x) \leq 0, \ \ x \neq \hat{x}\}$$

and

$$F(\bar{x},x) + q(\bar{x},x) > 0 \quad \textit{for all } x \in X \textit{ and } x \neq \bar{x}.$$

(ii) (Existence of Solutions of EP). *Assume that*

$$\begin{cases} \textit{For every } \tilde{x} \in \hat{S}, \textit{ there exists } x \in X \\ \textit{Such that } x \neq \tilde{x} \textit{ and } F(\tilde{x},x) + q(\tilde{x},x) \leq 0. \end{cases}$$

Then there exists $\bar{x} \in \hat{S}$ such that $F(\bar{x},x) \geq 0$ for all $x \in X$.

(iii) (Caristi–Kirk-Type Fixed Point Theorem). *Let $T : X \to 2^X$ be a multivalued mapping such that*

$$\begin{cases} \text{For every } \tilde{x} \in \hat{S}, \text{ there exists } x \in T(\tilde{x}) \text{ satisfying} \\ F(\tilde{x},x) + q(\tilde{x},x) \leq 0. \end{cases}$$

Then there exists $\bar{x} \in \hat{S}$ such that $\bar{x} \in T(\bar{x})$.

(iv) (Oettli and Théra-Type Theorem). *Let D be a subset of X such that*

$$\begin{cases} \text{For every } \tilde{x} \in \hat{S} \backslash D, \text{ there exists } x \in X \\ \text{Such that } x \neq \tilde{x} \text{ and } F(\tilde{x},x) + q(\hat{x},x) \leq 0. \end{cases}$$

Then there exists $\bar{x} \in \hat{S} \cap D$.

The proof of this theorem lies on the lines of the proof of Theorem 5 in [54] and therefore we omit it.

Remark 10.25. A generalization of Nadler's fixed point theorem [52] to the complete quasi-metric spaces with a Q-function is established in [1].

References

1. Al-Homidan, S., Ansari, Q.H., Yao, J.-C.: Some generalizations of Ekeland-type variational principle with applications to equilibrium problems and fixed point theory. *Nonlinear Anal.* **69**, 126–139 (2008).
2. Amemiya, M., Takahashi, W.: Fixed point theorems for fuzzy mappings in complete metric spaces. *Fuzzy Sets Syst.* **125**, 253–260 (2002).
3. Ansari, Q.H., Lin, L.J., Su, L.B.: Systems of simultaneous generalized vector quasi-equilibrium problems and their applications. *J. Optim. Theory Appl.* **127(1)**, 27–44 (2005).
4. Ansari, Q.H., Wong, N.C., Yao, J.C.: The existence of nonlinear inequalities. *Appl. Math. Lett.* **12(5)**, 89–92 (1999).
5. Aubin, J.-P.: *Mathematical Methods of Game and Economic Theory*. North-Holland, Amsterdam, New York, Oxford (1979).
6. Aubin, J.-P., Ekeland, I.: *Applied Nonlinear Analysis*. John Wiley, New York (1984).
7. Aubin, J.-P., Frankowska, H.: *Set-Valued Analysis*. Birkhäuser, Boston (1990).
8. Bae, J.S.: Fixed point theorems for weakly contractive multivalued maps. *J. Math. Anal. Appl.* **284**, 690–697 (2003).
9. Bianchi, M., Schaible, S.: Generalized monotone bifunctions and equilibrium problems. J. Optim. Theory Appl. **90**, 31–43 (1996).
10. Bianchi, M., Schaible, S.: Equilibrium problems under generalized convexity and generalized monotonicity. *J. Global Optim.* **30**, 121–134 (2004).
11. Bianchi, M., Kassay, G., Pini, R.: Existence of equilibria via Ekeland's principle. *J. Math. Anal. Appl.* **305**, 502–512 (2005).
12. Blum, E., Oettli, W.: From optimization and variational inequalities to equilibrium problems. *Math. Student* **63** (1–4), 123–145 (1994).
13. Bosch, C., Garcia, A., Garcia, C.L.: An extension of Ekeland's variational principle to locally complete spaces. *J. Math. Anal. Appl.* **328**, 106–108 (2007).

14. Brézis, H., Nirenberg, L., Stampacchia, G.: A remark on Ky Fan's minimax principle. *Boll. Un. Mat. Ital.* **6**, 293–300 (1972).
15. Caristi, J.: Fixed point theorems for mappings satisfying inwardness conditions. *Trans. Amer. Math. Soc.* **215**, 241–251 (1976).
16. Caristi, J., Kirk, W.A.: Geometric ixed point theory and inwardness conditions. In: *The Geonetry of Metric and Linear Spaces*, Lecture Notes in Mathematics, **490**, pp. 74–83. Springer-Verlag, Berlin (1975).
17. Chadli, O., Chbani, Z., Riahi, H.: Equilibrium problems with generalized monotone bifunctions and applications to variational inequalities. *J. Optim. Theory Appl.* **105**, 299–323 (2000).
18. Chadli, O., Wong, N.C., Yao, J.C.: Equilibrium problems with applications to eigenvalue problems. *J. Optim. Theory Appl.* **117**, 245–266 (2003).
19. Chen, Y., Cho, Y.J., Yang, L.: Note on the results with lower semi-continuity. *Bull. Korean Math. Soc.* **39** (4), 535–541 (2002).
20. Daneš, J.: A geometric theorem useful in nonlinear analysis. *Boll. Un. Mat. Ital.* **6**, 369–372 (1972).
21. Daneš, J.: Equivalence of some geometric and related results of nonlinear functional analysis. *Commentat. Math. Univ. Carolinae* **26**, 443–454 (1985).
22. De Figueiredo, D.G.: *The Ekeland Variational Principle with Applications and Detours*. Tata Institute of Fundamental Research, Bombay (1989).
23. Ekeland, I.: Sur les problèms variationnels. *C. R. Acad. Sci. Paris* **275**, 1057–1059 (1972).
24. Ekeland, I.: On the variational principle. *J. Math. Anal. Appl.* **47**, 324–353 (1974).
25. Ekeland, I.: On convex minimization problems. *Bull. Amer. Math. Soc.* **1(3)**, 445–474 (1979).
26. Facchinei, F., Pang, J.-S.: *Finite Dimensional Variational Inequalities and Complementarity Problems*, I. Springer-Verlag, New York (2003).
27. Fakhar, M., Zafarani, J.: Generalized equilibrium problems for quasimonotone and pseudomonotone bifunctions. *J. Optim. Theory Appl.* **123**, 349–364 (2004).
28. Fakhar, M., Zafarani, J.: Equilibrium problems in the quasimonotone case. *J. Optim. Theory Appl.* **126**, 125–136 (2005).
29. Fan, K.: A minimax inequlity and applications, In: Shisha, O. (ed.) *Inequality III*, pp. 103–113. Academic Press, New York (1972).
30. Fang, J.-X.: The variational principle and fixed point theorems in certain topological spaces. *J. Math. Anal. Appl.* **202**, 398–412 (1996).
31. Feng, Y.Q., Liu, S.Y.: Fixed point theorems for multivalued contractive mappings and multivalued Caristi type mappings. *J. Math. Anal. Appl.* **317**, 103–112 (2006).
32. Flores-Bazán, F.: Existence theorems for generalized noncoercive equilibrium problems: The quasi-convex case. *SIAM J. Optim.* **11**, 675–690 (2000).
33. Flores-Bazán, F.: Existence theory for finite-dimensional pseudomonotone equilibrium problems. *Acta Appl. Math.* **77**, 249–297 (2003).
34. Georgiev, P.G.: The strong Ekeland variational principle, the strong drop theorem and applications. *J. Math. Anal. Appl.* **131**, 1–21 (1988).
35. Göfert, A., Tammer, Chr., Riahi, H., Zălinescu, C.: *Variational Methods in Partially Ordered Spaces*. Springer-Verlag, New York (2003).
36. Hadjisavvas, H., Komlósi, S., Schaible, S.: *Handbook of Generalized Convexity and Generalized Monotonicity*. Springer-Verlag, Berlin (2005).
37. Hadjisavvas, H., Schaible, S.: From scalar to vector equilibrium problems in quasimonotone case. *J. Optim. Theory Appl.* **96**, 297–309 (1998).
38. Hamel, A.: Remarks to an equivalent formulation of Ekeland's variational principle. *Optimization* **31** (3), 233–238 (1994).
39. Hamel, A.: Equivalents to Ekeland's variational principle in $\mathscr{F}$-type topological spaces. *Report of the Institute of Optimization and Stochastics* 9, Martin Luther University, Halle-Wittenberg (2001).
40. Hamel, A.: Phelps' lemma, Daneš' drop theorem and Ekeland principle in locally convex spaces. *Proc. Amer. Math. Soc.* **131**, 3025–3038 (2003).

41. Hamel, A.: Equivalents to Ekeland's variational principle in uniform spaces. *Nonlinear Anal.* **62**(5), 913–924 (2005).
42. Isac. G.: Nuclear cones in product spaces, Pareto efficiency and Ekeland-type variational principles in locally convex spaces. *Optimization* **53**, 253–268 (2004).
43. Isac, G.: Ekeland principle and nuclear cones: A geometrical aspect. *Math. Computat. Modell.* **26** (11), 111–116 (1997).
44. Jing-Hui, Q.: Ekeland's variational principle in locally complete spaces. *Math. Nachr.* **257**, 55–58 (2003).
45. Kada, O., Suzuki, T., Takahashi, W.: Nonconvex minimization theorems and fixed point theorems in complete metric spaces. *Math. Japonica* **44(2)**, 381–391 (1996).
46. Kas, P., Kassay, G., Sensoy, Z.B.: On generalized equilibrium points. *J. Math. Anal. Appl.* **296**, 619–633 (2004).
47. Khamsi, M.A., Kirk, W.A.: *An Introduction to Metric Spaces and Fixed Point Theory*. John Willy & Sons, New York (2001).
48. Lin, L.J.: System of coincidence theorems with applications. *J. Math. Anal. Appl.* **285**, 408–418 (2003).
49. Lin, L.J., Du, W.S.: Ekeland's variational principle, minimax theorems and existence of nonconvex equilibria in complete metric spaces. *J. Math. Anal. Appl.* **323**, 360–370 (2006).
50. Lin, L.J., Du, W.S.: On maximal element theorems, variants of Ekeland's variational principle and their applications. *Nonlinear Anal.* **68**, 1246–1262 (2008).
51. McLinden, L.: An application of Ekeland's theorem to minimax problems. *Nonlinear Anal. Theory, Meth. Appl.* **6**, 189–196 (1982).
52. Nadler, Jr. S.B.: Multivalued contraction mappings. *Pacific J. Math.* **30**, 475–488 (1969).
53. Nikaido, H., Isoda, K.: Note on noncooperative convex games. *Pacific J. Math.* **5**, 807–815 (1955).
54. Oettli, W., Théra, M.: Equivalents of Ekeland's principle. *Bull. Austral. Math. Soc.* **48**, 385–392 (1993).
55. Park, S.: Equivalent formulations of Ekeland's variational principle for approximate solutions of minimization problems and their applications. In: S.P. Singh, V.M. Sehgal, and J.H.W. Burry (Eds.) *Operator Equations and Fixed Point Theorems*, pp. 55–68. MSRI-Korea Publishers 1, Soul (1986).
56. Park, S.: On generalizations of the Ekeland-type variational principles. *Nonlinear Anal.* **39**, 881–889 (2000).
57. Penot, J.-P.: The drop theorem, the petal theorem and Ekeland's variational principle. *Nonlinear Anal. Theory, Meth. Appl.* **10**, 813–822 (1986).
58. Rockafellar, R.T.: Directionally Lipschitz functions and subdifferential calculus. *Proc. London Math. Soc.* **39**, 331–355 (1979).
59. Sullivan, F.: A characterization of complete metric spaces. *Proc. Amer. Math. Soc.* **83**, 345–346 (1981).
60. Shioji, N., Suzuki, T., Takahashi, W.: Contractive mappings, Kannan mappings and metric completeness. *Proc. Amer. Math. Soc.* **126**, 3117–3124 (1998).
61. Suzuki, T.: Generalized distance and existence theorems in complete metric spaces. *J. Math. Anal. Appl.* **253**, 440–458 (2001).
62. Suzuki, T.: On Downing-Kirk's theorem. *J. Math. Anal. Appl.* **286**, 453–458 (2003).
63. Suzuki, T.: Generalized Caristi's fixed point theorems by Bae and others. *J. Math. Anal. Appl.* **302**, 502–508 (2005).
64. Suzuki, T.: The strong Ekeland variational principle. *J. Math. Anal. Appl.* **320**, 787–794 (2006).
65. Suzuki, T., Takahashi, W.: Fixed point theorems and characterizations of metric completeness. *Top. Meth. Nonlinear Anal.* **8**, 371–382 (1996).
66. Takahashi, W.: Existence theorems generalizing fixed point theorems for multivalued mappings. In: Baillon, J.-B., Théra, M. (Eds.) *Fixed Point Theory and Applications*, pp. 397–406. Pitman Research Notes in Mathematics, **252**, Longman, Harlow (1990).

67. Takahashi, W.: *Nonlinear Functional Analysis*. Yokohama Publishers, Yokohama (2000).
68. Tataru, D.: Viscosity solutions of Hamilton-Jacobi equations with unbounded nonlinear terms. *J. Math. Anal. Appl.* **163**, 345–392 (1992).
69. Yuan, G.X.-Z.: *KKM Theory and Applications in Nonlinear Analysis*. Marcel Dekker, New York (1999).
70. Zhong, C.-K.: On Ekeland's variational principle and a minimax theorem. *J. Math. Anal. Appl.* **205**, 239–250 (1997).

Chapter 11
Decomposition Methods Based on Augmented Lagrangians: A Survey

Abdelouahed Hamdi and Shashi K. Mishra

Abstract In this chapter, we provide a non-exhaustive account of decomposition algorithms for solving structured large scale convex and non-convex optimization problems with major emphasis on several splitting approaches based on the classical or modified augmented Lagrangian functions. This study covers last 40 years of research on theoretical properties of augmented Lagrangians.

11.1 Augmented Lagrangian Method

In 1968–1969, Powell [77] and Hestenes [46] introduced independently (under different forms) a new algorithm to solve nonlinear problems with equality constraints. Later on, Haarhoff and Buys [36] proposed a similar iterative scheme based on the same idea. The name "*multiplier method*" is due to Hestenes. To the following nonlinear programming problem

$$\min_{h_i(x)=0,\ i=\overline{1,p}} f(x) \tag{11.1}$$

by combining ideas from primal–dual methods and from penalty methods, Hestenes added a penalty term to the ordinary Lagrangian to obtain a penalized Lagrangian. His penalized Lagrangian can be written:

$$\mathscr{L}_{\mathscr{A}}(x,v,k) = f(x) + \sum_{i=1}^{p} v_i h_i(x) + \frac{k}{2}\sum_{i=1}^{p} h_i^2(x), \tag{11.2}$$

Abdelouahed Hamdi
Kuwait University, College of Science, Department of Mathematics and Computer Science, P.O. Box 5969, Safat 13060, State of Kuwait, e-mail: whamdi99@yahoo.fr

Shashi K. Mishra
Department of Mathematics, Banaras Hindu University, Varanasi, India
e-mail: bhu.skmishra@gmail.com

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_11,

where $v = (v_1, \ldots, v_p)$ is the Lagrange multiplier vector and $k > 0$ is an arbitrary penalty factor (in this formulation, all the constraints are supposed to be equality constraints). Hestenes proposed to solve a sequence of unconstrained minimization problems using the penalized Lagrangian. The sth unconstrained minimization is

$$\min_x \mathscr{L}_{\mathscr{A}}(x, v^s, k), \tag{11.3}$$

where v^s is the current estimate of the Lagrange multiplier vector. After each minimization, v^s is updated by the following formula,

$$v_i^{s+1} = v_i^s + kh_i(x^s), \qquad i = \overline{1,p}, \tag{11.4}$$

where x^s solves (11.3). The goal of this iteration is to solve the dual problem. An important characteristic is that the penalty parameter k should not increase to infinity, to avoid ill-conditioning which is usually associated with penalty methods. The dual iteration tends to converge rapidly, making the algorithm very efficient. The numerical experiment confirmed the efficiency of the method, but the first three articles cited above offered a limited theoretical analysis.

The quadratically penalized Lagrangian (11.2) used by Hestenes was in fact studied earlier by Arrow, Hurwicz, and Uzawa ([2], 1958). They analyzed the properties of the penalized Lagrangian saddle point and suggested a gradient algorithm to localize this saddle point directly.

In 1970, a large number of new results on the multiplier method appeared. Some of them concentrated on analytical properties of the penalized Lagrangians and others focused on multiplier-type algorithms based on these Lagrangians.

The penalized Lagrangian of Arrow and Solow dealt with equality constraints only. In 1970, Rockafellar [79] introduced a similar Lagrangian for inequality constraints, followed in 1972 by a more detailed analysis of the duality aspects of his Lagrangian [80]. Rockafellar integrated the penalized Lagrangians quadratically without the theory of generalized Lagrangians. In another article in 1972 [78], Rockafellar gave an extension of Hestenes' method to manipulate inequality constraints, using the Lagrangian of [79]. He proved the global convergence of the algorithm in the case of a convex programming problem, and also demonstrated that the convergence was preserved with an inexact minimization of the subproblem. However, his stopping test was not applied numerically. Rockafellar [78] gave conditions under which convex programs are solvable using the quadratically penalized Lagrangian, which he baptized augmented Lagrangian.

Buys' thesis in 1972 [14] contains an extensive analysis of the multiplier methods. Using the generalized Lagrangian introduced in 1970 by Rockafellar, Buys generalized Hestenes' algorithm to manipulate inequality constraints directly, without using slack variables (the same algorithm given by Rockafellar at the same time). Buys proved local convergence in the nonconvex case under second-order sufficient conditions. He also showed that the unconstrained minimizations need not be exact. In his book, Luenberger [56] briefly explores the dual aspect of the multiplier method and also gives an interpretation that the dual iteration is a gradient iteration to maximize the dual function. Rockafellar [80], in 1974, gave an extension to the unique theoretical result of Arrow, Gould, and Howe ([1], 1973), concerning saddle points in nonconvex programming, which states that under the second-order sufficient optimality conditions and under the strict complementarity condition, one can show the existence of a saddle point for a certain class of generalized Lagrangians. Rockafellar, by introducing the notions of quadratic increase and stability, has shown that, by using an augmented duality (replace the ordinary Lagrangian by the augmented Lagrangian), the saddle points exist and the dual problem need not be constrained in the case of problems with inequality constraints.

A summary of the augmented Lagrangian method with all its variants up to 1976 with 66 references can be found in [8].

In the case of inequality constraints,

$$\min_{g_i(x)\leq 0,\ i=\overline{1,p}} f(x), \tag{11.5}$$

it is straightforward to build the augmented Lagrangian via some slack variables to get:

$$\mathscr{L}_{\mathscr{A}}(x,v,k) = f(x) + \sum_{i=1}^{p} \begin{cases} v_i\, g_i(x) + \frac{k}{2}[g_i(x)]^2 & \text{if } kg_i(x)+v_i \geq 0 \\ -\frac{1}{2k}v_i^2 & \text{otherwise.} \end{cases} \tag{11.6}$$

In 1975, Pierre and Lowe [71] published the first book entirely dedicated to the study of mathematical programming approaches via augmented Lagrangians. They devoted two chapters to the numerical study of the method, comparing its performance to that of other methods.

Having presented the multiplier method and the augmented Lagrangian function, we show, in the next section, the evolution of this primal–dual approach.

11.2 Extensions of Augmented Lagrangian Methods

Since 1970, Miele et al. ([63–65]) have considered modifications of the multiplier method where only approximate minimizations were required in the rest of the unconstrained problems. They also have proposed modifications of the dual iteration and have given some theoretical results concerning their modifications. Numerical experiments have been carried out by Miele et al. [63] indicating the efficiency of their propositions. To solve a program with equality constraints, Miele and his associates have proposed, in a first paper, two classes of algorithms with the common characteristic to determine the Lagrange multipliers v^*. Their idea meets that of Fletcher [28] and Martensson [62], to minimize the error on the optimality condition satisfaction. The first class uses an ordinary gradient method to solve the unconstrained problem, and itself has two variants. In the first one, the factor k is selected, and is fixed all along the iterations. On the other hand, the second variant is based on the idea to satisfy the constraints, which induces the following characterization of k,

$$k = \frac{P(x)}{\|\nabla P(x)\|^2},$$

with $P(x) = \|h(x)\|^2$.

The second class also comprises two variants with a common step, that of determining v^* in the following way.

$$\nabla h(x)'\nabla h(x)v^* + \nabla h(x)'[\nabla f(x) + kh(x)\nabla h(x)'] - h(x) = 0.$$

The first variant uses a constant penalty factor, chosen a priori, whereas the second one uses the same update expression for k that is used in the first class of algorithms cited above.

In general, these two classes of algorithms target the simultaneous reduction of the objective function, the constraints, and the augmented Lagrangian function. In 1972, Miele et al. [65] proposed other modifications, emphasizing the possibility of a single gradient or modified Newton iteration for the unconstrained subproblems. One can say that the works of Miele et al. from 1970–1972 are the first numerical results concerning the augmented Lagrangian method and its variants.

In 1972, Kort and Bertsekas [51] introduced a class of penalty functions for the cases of equalities and inequalities in the convex case. The Hestenes–Powell–Rockafellar method was a particular case of the algorithm proposed by Kort and Bertsekas. These last two have shown the global convergence in the convex case and have given a geometric interpretation of the method.

Fletcher ([27–29]) and Lill [53] have proposed and analyzed a variant of the multiplier method using a dual iteration similar to that of Miele. But when Miele alternates his dual step with one or more cycles of a descent algorithm, Fletcher replaces the dual variable by another one, dependent on the primal variable x. This approach has an important theoretical motivation but Fletcher's corresponding algorithm turned out to be more difficult to implement than the original multiplier method. The difficulty is that the dual iteration requires a matrix inversion. Different from that of the multiplier, Fletcher's method requires a single unconstrained minimization provided a penalty coefficient large enough has been chosen a priori. This approach was been studied by Martensson in 1973 [62] and then in 1975 by Mukai and Polak [66].

Bertsekas [7], and at the same time Buys in his thesis, have given a first result concerning the linear convergence of the multiplier method. Bertsekas [7] has also suggested modifying the dual iteration by using alternate choices of the dual step. He showed that this modification accelerates the convergence for convex problems. More generally, the optimal step[1] (giving the best convergence rate) depends on the structure of the eigenvalues of the dual function Hessian.

Rupp [83] has also studied some properties of the convergence of the Hestenes–Powell–Rockafellar algorithm. In [84], he proved local convergence with a modified

[1] It is the parameter α in $v^{s+1} = v^s + \alpha_s h(x^{s+1})$.

Newton step, using second-order sufficient optimality conditions. In [82] and [81], Rupp extended the method to the infinite-dimension case for an optimal control problem, proving also the local convergence and deducting similar results on the convergence rate.

Mangasarian [57] introduced a general class of Lagrangians. Following Powell's [77] formulation, Mangasarian expresses his Lagrangian in terms of shifted penalty functions. Mangasarian concentrated his study on local results for non-convex problems.

Arrow, Gould, and Howe [1] proposed a class of generalized Lagrangians similar to the class treated by Kort [50]. They concentrated on the theoretical properties of the Lagrangian, especially the properties of the saddle points, but they did not propose any algorithm.

Other authors have proposed variants of the basic method. We mention Tripathi and Narendra [96] who each analyzed and reported some applications of the multiplier method.

In 1976, Kort and Bertsekas [51] and Kort [50] showed new results on the augmented Lagrangian method, based on a large class of generalized Lagrangians. These results are interesting from the point of view of convergence rate, which varies according to the penalty functions used. For certain penalty functions, the augmented Lagrangians

$$\mathscr{L}_{\mathscr{A}}(x,v,z,k) = f(x) + \sum_{i=1}^{m} v_i h_i(x) + \sum_{i=1}^{m} P_I[h_i(x),0,k] + \sum_{i=m+1}^{p} z_i h_i(x) + P_E[h_i(x),0,k],$$

are twice continuously differentiable, which is not the case for the quadratic augmented Lagrangian of Rockafellar.

As examples of penalty functions that they proposed, one finds

$$P_I(t,0,\lambda) = \begin{cases} \frac{\lambda}{2}t^2 + t^\rho, & \text{if } t \geq 0, \\ 0, & \text{if } t \leq 0, \end{cases}$$

and

$$P_I(t,0,\lambda) = \begin{cases} \frac{\lambda}{2}t^2, & \text{if } t \geq 0, \\ 0, & \text{if } t \leq 0, \end{cases}$$

and for the case of equality constraints

$$P_E(t,\lambda) = \frac{\lambda}{2}t^2,$$

$$P_E(t,\lambda) = \frac{\lambda}{\rho}t^\rho, \qquad \rho > 1,$$

and

$$P_E(t,\lambda) = \lambda(\cosh(t) - 1).$$

It should be noted that Kort [50] and Kort and Bertsekas [51] have invested mostly in the convex case. Their method was clearly presented as a combination of a penalty method and a primal–dual method.

The augmented Lagrangian algorithm has the particularity of being easy to implement, which has pushed a large number of researchers to design techniques and modifications to accelerate the convergence of this method. For example, Betts [12] proposed an algorithm characterized by the estimation of the multipliers, which is done through the resolution of a quadratic problem, issued from the minimization of K–K–T conditions, and the unconstrained minimization is accelerated by a Newton–Raphson-type extrapolation procedure. This approach inconveniently the needs to identify the active constraints at each iteration.

Still in 1977, Bertsekas [9] proposed the multiplier method with partial elimination of the constraints. This is equivalent to solving a subproblem with part of the constraints that are easy to deal with. Bertsekas showed the convergence of his approach for the case of equality constraints. Also, he applied his method and the multiplier method to solve problems whose objective functions suffer from ill-conditioning or nondifferentiability. His application is equivalent to an approximation procedure based on the use of the augmented Lagrangian method to solve the intermediary problem in the approximation. For more details, see [9]. In 1979, Bertsekas [10] again considered the idea of Buys [14] which was to maximize the dual augmented function by a second-order approach (the case of equality constraints). He showed the convergence of the second-order multiplier method based on Newton's method, but did not require the choice of an initial multiplier close enough to the Lagrange multiplier v^* provided the penalty coefficient k was large enough. A similar study was done before Bertsekas by Tapia [91], in 1977, who gave a deep analysis of the different formulations of the dual iteration within the multipliers method. Tapia was interested in primal–dual quasi-Newtonian approaches to improve Rupp's [82] 1975 result, who used a Newtonian approach. Tapia obtained a Q-superlinear convergence by solving the unconstrained minimizations by a quasi-Newtonian approach, and by using different ways to update the sequence of multipliers. He also obtained a Q-quadratic convergence for his diagonal approach which, in summary, uses a Newton or quasi-Newton step in the minimization of the augmented Lagrangian function. Byrd [15] extended Tapia's [91] results by using two Newton steps in the multiplier diagonal approach.

In 1978, Glad [32] presented an approach similar to that of Tapia [91], which updated the dual variables at each iteration of the unconstrained minimization approach used. Glad used a quasi-Newtonian approach of the BFGS type[2] for the minimization of the augmented Lagrangian and he obtained a superlinear convergence.

Another variant of the quadratic augmented Lagrangian algorithm is due to Betts [13] who, this time, exploited the primal–dual approach to find a stationary point of the augmented Lagrangian, and the standard penalty approach which seeks to satisfy the constraints by augmenting the value of the penalty coefficient. Betts proposed

[2] Broyden–Fletcher–Goldfarb–Shanno.

combining these two approaches as follows. First determine an approximation of a stationary point of the augmented Lagrangian by a projected-gradient-type method, and then use this point as an initial point for the minimization of the quadratic term that was added to the ordinary Lagrangian. This approach also depends on the resolution of a quadratic program to estimate the multipliers.

In 1980, Boggs and Tolle [6] proposed a variety of augmented Lagrangians having the particularity of being concave and quadratic in v. They developed duality results that lie on the estimation of the Lagrange multipliers as a function of x and of a penalty parameter, the maximum of their modified Lagrangian in v. Their development links the theory of modified Lagrangians to the exact penalties.

Also, from the numerical point of view, Coope and Fletcher [20] have studied and tested a variant of the augmented Lagrangian algorithm which comprises two main steps: the first one is a quadratic (linear) approximation of the augmented Lagrangian function (respectively, of the constraints) and the second one is an inexact minimization of the augmented Lagrangian. If the first step does not satisfy the corresponding test, then one moves to the second step and alternates between them until an appropriate stop test is satisfied.

Still in 1980, we cite the work of Polak and Tits [72] who proposed a technique to limit the increase in the penalty coefficient in the multiplier method. They relied on the Mukai–Polak [66] idea in the framework of the multiplier algorithm with the multiplier function.

Di Pillo and Grippo [21, 22] have designed an approach for the case of equality constraints that they later extended to the inequality case. This approach is based on consideration of the augmented Lagrangian of Hestenes–Powell to which they added the first-order necessary optimality condition in terms of penalty. They showed that, under certain hypotheses, a local solution of the constrained problem and the associated multiplier can be determined by a single unconstrained minimization of the new Lagrangian in (x,v), for finite values of k and without the need for matrices inversion, as is the case in the approaches of Fletcher [29], Mukai and Polak [66], and Martensson [62]. Similarly, Lucidi [55] introduced a twice-continuously differentiable augmented Lagrangian that is also based on first-order necessary optimality conditions.

11.3 Nonquadratic Augmented Lagrangians

Following the work of Arrow, Gould, and Howe [1], Kort [50], and Kort and Bertsekas [51], a certain number of researchers further developed the study of the nonquadratic proximal approaches. Among them, one may cite Bertsekas [11] and the references therein, and more recently Polyak [73], Eckstein [23], Teboulle [95], and Tseng and Bertsekas [97]. These authors were motivated by the fact that nonquadratic proximal regularization improves the convergence rate (see [3]) and consequently their performances. Moreover, the augmented Lagrangians are, in the convex case, the result of a dual proximal regularization, thus these works gave birth to a quite rich variety of augmented Lagrangians. Also, one of the contributions of the notion of nonquadratic proximal regularization is to produce twice-differentiable augmented Lagrangians[3] (in the case of inequality constraints), a property that is not verified in the case of quadratic regularization.

In 1990, Censor and Zenios [16], presented a proximal point algorithm based on the use of a class of functions, called Bregman functions, in the convex differentiable case. A Bregman function ψ (a function that satisfies a certain number of properties [22]) allows us, to define a kind of "*distance*" D_ψ:

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

$D_\psi(x,y)$ can be interpreted as the vertical distance between x and the tangent plane to ψ at y. (For more details, see [22].) The functional $D_\psi(x,y)$ replaces the quadratic term of the proximal regularization and the corresponding modified Lagrangian for problem (11.5) is given by:

$$L_{\mathrm{Breg}}(x,v,k) = f(x) + k\sum_{i=1}^{m} \psi_i^+\left(k^{-1}g_i(x) + \psi_i'(v_i)\right) - \psi_i^+\left(\psi_i'(v_i)\right),$$

where $\psi_i^+(z) = \sup_{t\geq 0}\{z\,t - \psi_i(t)\}$. In the same spirit of regularization, Teboulle [94], [95], and Iusem and Teboulle [47] replaced the functional D_ψ by the ϕ-divergence functional d_ϕ, in the form[4]

$$d_\phi(x,y) = \sum_i y_i \phi(x_i/y_i),$$

where ϕ is a real variable function, $\mathscr{C}^2$, strictly convex and verifying other properties. As examples of these functions, one may cite

$$\phi(t) = t\log(t) - t + 1, \quad \phi(t) = -\log(t) + t - 1, \qquad \phi(t) = t + t^{-1} - 2,$$
$$\phi(t) = (\sqrt{t} - 1)^2.$$

The associated augmented Lagrangian is given by

$$L_{div}(x,v,k) = f(x) + k\sum_{i=1}^{m} \psi_i^*(k^{-1}g_i(x)),$$

where $\psi_i^*(z) = \sup_t\{z\ t - \psi_i(t)\}$. For more details, we refer the reader to [95] and [94].

Also, we can cite a paper of Auslender, Teboulle, and Ben-Tiba ([3] and [4], 1999), where the authors introduced a new nonquadratic proxlike quasi-distance $d(u,v) = \sum_{j=1}^n v_j^2 \phi(u_j/v_j), \forall\, u,v > 0$, where

---

[3] A big advantage inasmuch as one can use Newton-type methods.

[4] Note that d_ϕ is not a distance.

$$\phi_{\nu,\mu}(t) = \begin{cases} \frac{\nu}{2}(t-1)^2 + \mu\,(t - \log(t) - 1) & \text{if } t > 0 \\ \infty & \text{otherwise} \end{cases} \tag{11.7}$$

where their augmented Lagrangian is given by

$$L_k(u,v) = f(x) + k^{-1} \sum_{i=1}^{m} u_i^2 \phi_{\nu,\mu}^{*}[k(p_i + g_i(x))/u_i].$$

These new proxlike algorithms were applied to solve unconstrained convex minimization and also variational inequality problems. And the dual application of such proxlike schemes gave birth to a C^∞ new augmented Lagrangian. This result is still worthy because it enables the use of efficient Newton-type methods to solve the subproblems. In parallel, we also cite the work of Roman Polyak on the nonlinear rescaling principle to solve nonlinear programming problems. Polyak's results can initiate a general scheme to generate modified augmented Lagrangians ([74–76]). Many results related to the nonlinear rescaling principle as primal–dual algorithms, complexity, rate of convergence, numerical efficiency, and numerical comparisons can be found in ([74–77]).

11.4 Augmented Lagrangian and Decomposition Methods

The resolution of a large nonconvex problem by an augmented Lagrangian-type method has been done in a certain number (less than ten) of works. But unfortunately, the nonseparable character of the augmented Lagrangian, due to mixed products in the quadratic term $\| \sum_{i=1}^{p} h_i(x_i) \|^2$, has pushed researchers into trying to overcome this obstacle which deprives us of the advantages of the augmented Lagrangian method. Many approaches have been proposed; among them, one may cite that of Stephanopoulos and Westerberg [89] who considered a model of problems frequently found in the conception of engineering systems, in the form of a large system composed of many interconnected subsystems. Their idea is based on a first-order approximation of the nonseparable quadratic term. Watanabe, Nishimura, and Matsubara [100] have added variables to the augmented Lagrangian minimization problem in such a way that the nonseparable term is replaced by the minimum of a separable function. Their decomposition approach comprises three levels and it should be noted that the introduction of variables increases the method complexity. Also, in the context of solving the duality gap problem in nonconvex programming, Bertsekas [10] proposed in 1979 a convexification procedure different from those exposed so far inasmuch as it is not based on the direct utilization of the augmented Lagrangian. His new procedure, applied to problem (11.1) with $h(x) = (h_1(x), \ldots, h_p(x))$, preserves its separability. It is based on the following local minimization,

$$\varphi_k(z) = \min_{h(x)=0} \left\{ f(x) + \frac{k}{2}\|x - z\|^2 \right\},$$

where z is an estimate of the solution of (IP) in a neighborhood of x^*. Bertsekas [10] showed that x^* is the solution of the following problem,

$$\min_{z \in \Re^n} \varphi_k(z).$$

Thanks to the penalty coefficient k, we have a local convexity that allows us to use primal–dual methods, in particular that of decomposition by prices.
The method is done in three steps. First, solve

$$g_k(u,z) = \min_{x \in \Re^n} \left\{ f(x) + \frac{k}{2}\|x-z\|^2 + u^\top h(x) \right\},$$

then calculate $\bar{g}_k(z) = \max_u g_\lambda(u,z)$, and finally minimize $\bar{g}_\lambda(z)$ with respect to z. The minimization in the first step is separable. Indeed, it can be written

$$\min_{x \in \Re^n} \left\{ \sum_{i=1}^{p} f_i(x_i) + \frac{k}{2}\|x_i - z_i\|^2 + v_i' h_i(x_i) \right\}.$$

Even the problem in the third step is decomposable.

In the same spirit of this decomposition, but passing by an augmented Lagrangian, Tanikawa and Mukai [90] considered the approach given by Bertsekas, which is quite interesting in its simplicity, but suffers enormously from the fact that it operates in three steps, which makes it expensive. And following the path of Fletcher [30], who introduced the notion of the multiplier function to eliminate the multiplier update step in the augmented Lagrangian method, Tanikawa and Mukai have combined this procedure with that of Bertsekas. That is, by choosing an estimate z of the solution x^* of problem (IP), the Lagrangian corresponding to the following problem,

$$\min_{h(x)=0} \left\{ f(x) + \frac{k}{2}\|x-z\|^2 \right\},$$

can be written

$$L_k(x, v(x), z) = f(x) + \frac{k}{2}\|x-z\|^2 + v(x)^\top h(x),$$

where $v(\cdot)$ denotes the multiplier function defined by

$$v(x) = \arg\min \left\{ \|\nabla_x L(x,v)\|^2 : v \in \Re^m \right\},$$

and that satisfies $v(x^*) = v^*$.

One can see that because $v(x)$ depends on x, then $L_k(x, v(x))$ is not separable. From the fact that z approaches x^* and $v(x)$ approaches $v(x^*)$, one may think that $v(z)$ approaches $v(x^*)$. Immediately, the new Lagrangian

$$L_k(x,v(z),z) = \sum_{i=1}^{p} \left\{ f_i(x_i) + \frac{k}{2}\|x_i - z_i\|^2 + v(z)'h_i(x_i) \right\},$$

is separable. But, following Bertsekas' idea, $\bar{z}$ minimizing $L_k(x,v(z),z)$ in x is a good estimate of x^* if and only if we penalize $L_k(x,v(z),z)$. That is, we must add a quadratic term $\beta h(x)'M(x)h(x)$, where $\beta > 0$ and $M(x)$ is a positive-definite matrix. This solution deteriorates the separability, which brought Tanikawa and Mukai to propose a penalization of the form:

$$L_{k,\beta}(x,v(z),z) = \sum_{i=1}^{p} \left\{ f_i(x_i) + \frac{k}{2}\|x_i - z_i\|^2 + v(z)'h_i(x_i) \right\} + \beta h(z)'M(z)h(x),$$

which, this time, conserves the separability of the modified Lagrangian. Indeed,

$$L_{k,\beta}(x,v(z),z) = \sum_{i=1}^{p} \left\{ f_i(x_i) + \frac{k}{2}\|x_i - z_i\|^2 + [v(z)' + \beta h(z)'M(z)]h_i(x_i) \right\},$$

can be written in the form $L_{k,\beta}(x,v(z),z) = \sum_{i=1}^{p} L^i_{k,\beta}(x_i,v(z),z_i)$ with

$$L^i_{k,\beta}(x_i,v(z),z_i) = f_i(x_i) + \frac{k}{2}\|x_i - z_i\|^2 + [v(z)' + \beta h(z)'M(z)]h_i(x_i).$$

In conclusion, the proposed algorithm consists of the minimization of $L_{k,\beta}(x,v(z),z)$ in x, and then in z.

The multiplier function $v(x)$ is given by

$$v(x) = -\left[\frac{\partial h(x)}{\partial x}\frac{\partial h(x)'}{\partial x}\right]^{-1} \frac{\partial h(x)'}{\partial x}\nabla f(x).$$

We remark that if we take $M(x) = \frac{1}{2}I$ where I is the identity matrix and $k = 0$, then we find the expression of the usual augmented Lagrangian. Tanikawa and Mukai have shown that $L_{k,\beta}(x,v(z),z)$ is locally convex and have given a bound on the error $\|\bar{z} - x^*\|$ in terms of the error $\|z - x^*\|$, where $\bar{z}$ minimizes $L_{k,\beta}(x,v(z),z)$ in x. This bound decreases considerably if we take

$$M(x) = \left[\frac{\partial h(x)}{\partial x}\frac{\partial h(x)'}{\partial x}\right]^{-1} \quad \text{and} \quad \beta = k.$$

Moreover, the convergence is linear.

We also cite Tatjewski and Engelmann [93] who considered the work of Tanikawa and Mukai but for the case of inequality constraints. For more details see [92].

Differently from what has been proposed, Cohen [19] proposed replacing the original problem to solve by a sequence of auxiliary problems built around auxiliary functionals called "kernels". Cohen [19] proposed this approach in the general framework of decomposition.

Also, for certain problems of the form

$$\min f(x) + h(z)$$
$$Mx = z,\ x \in \Re^n, \tag{11.8}$$

the application of the multiplier method by alternating the minimizations with respect to x and to z gives birth to an algorithm that is favorable to decomposition, called the multiplier alternating directions algorithm. This approach has been studied extensively and sufficient conditions for convergence have been obtained by using the theory of maximal monotone operators (Lions and Mercier [54], Eckstein and Bertsekas [25]) or the theory of saddle points (Glowinski, Fortin, and Le Tallec [31], [33]). The link with the proximal point algorithm has been studied by Rockafellar [78]. It has also been shown [23] that the method is a particular case of the Douglas–Rachford method for the computation of the zeros of the sum of maximal monotone operators (see [17, 99, 98]). Many other decomposition schemes, such as the Han and Lou algorithm, the partial inverse method of Spingarn [88], [87], and the block decomposition method for convex programming of Golshtein ([34], 1986), are also particular cases of the Douglas–Rachford method (see the thesis of Eckstein for all the proofs). The multiplier alternating directions method also known as ADM can be summarized as follows.

$$x^{s+1} \in \arg\min_x \left\{ f(x) + \langle v^s, Mx\rangle + \frac{k^s}{2}\|Mx - z^s\|^2 \right\}$$

$$z^{s+1} \in \arg\min_z \left\{ h(z) - \langle v^s, z\rangle + \frac{k^s}{2}\|Mx^{s+1} - z\|^2 \right\}$$

$$v^{s+1} = v^s + k^s(Mx^{s+1} - z^{s+1}).$$

We note that this last approach, in the convex case, is in fact the application of the Douglas–Rachford method to find a zero of a sum of two maximal monotone operators [31]. Chen and Teboulle [18] proposed a decomposition method of the same type as the alternating directions, where a quadratic proximal term replaces the penalty term of the augmented Lagrangian (it can be called predictor–corrector ADM); that is,

$$p^{s+1} = v^s + k^s(Mx^s - z^s)$$

$$x^{s+1} \in \arg\min_x \left\{ f(x) + \langle p^{s+1}, Mx\rangle + \frac{1}{2k^s}\|x - x^s\|^2 \right\}$$

$$z^{s+1} \in \arg\min_z \left\{ h(z) - \langle p^{s+1}, z\rangle + \frac{1}{2k^s}\|z - z^s\|^2 \right\}$$

$$v^{s+1} = v^s + k^s(Mx^{s+1} - z^{s+1}).$$

In nonconvex programming, we may cite an approach due to Shin ([86] 1992), which applies a recursive quadratic programming method (RQPM). In other words, he applied a descent method whose direction is the solution of the following quadratic problem.

$$(PQ) \qquad \begin{cases} \min & \sum_{i=1}^{p} \frac{1}{2} d_i' B_i d_i + \nabla f_i(x_i)' d_i \\ st & \sum_{i=1}^{p} h_i(x_i) + \nabla h_i(x_i)' d_i = 0, \end{cases}$$

where B is a diagonal positive-definite matrix formed only by the positive terms of the Hessian of the objective function $\sum_{i=1}^{p} f_i(x_i)$.

Shin has used a gradient method to solve the separable dual problem associated with *PQ*. His approach has been applied to solve control problems.

We recall that the augmented Lagrangian method with partial constraints elimination can be seen as a decomposition mean.[5] This idea has been exploited by Ferris and Mangasarian ([26], 1991). The latter have treated a convex program with a strongly convex objective function.

We also cite the work of Mulvey and Ruszczyński ([68], 1992) and Ruszczynski ([85], 1995) who proposed a decomposition method for linear problems of the form

$$\begin{aligned} \min \sum_{i=1}^{p} c_i' x_i & \\ st \quad \sum_{i=1}^{p} Q_i x_i = q&, \\ A_i x_i = b_i, \quad i = \overline{1,p}& \end{aligned}$$

based on the use of a quadratic augmented Lagrangian. The main idea is to linearize the nonseparable term as was done earlier by Stephanopoulos and Westerberg in 1975 [89]. Still, Mulvey and Ruszczyński ([67], 1995) proposed an approach using successive separable approximations of the nonseparable quadratic term, which, in a certain way, meets the idea of Stephanopoulos and Westerberg. This can be interpreted by a type *IP* problem as follows.

Minimize the function

$$L_A^i(x_i, \bar{x}, v_i, k) = f_i(x_i) + h_i(x_i) v_i + \frac{k}{2} \left\| h_i(x_i) + \sum_{j \neq i} h_j(\bar{x}_j) \right\|^2,$$

in x_i by a nonlinear Jacobi method according to the terminology of Bertsekas and Tsitsiklis, then update for all $i = \overline{1,p}$

$$\begin{cases} \bar{x}_i^{s+1} = \bar{x}_i^{s} + \tau(x_i^{s+1} - \bar{x}_i^{s}), \quad \tau > 0, \\ v_i^{s+1} = v_i^s + k^s \sum_{i=1}^{p} h_i(x_i^{s+1}). \end{cases}$$

Kiwiel, Rosa, and Ruszczyński ([48], 1995), proposed a proximal approach that alternates a linearization of the objective functions of the separable problem

[5] That is, we will have to solve problems of lower dimension.

constituted by the sum of two convex functions. That is, for the unconstrained minimization of the nonseparable augmented Lagrangian associated with problem IP, the approach in the case of linear constraints ($\sum_{i=1}^{p} h_i(x_i) = Ax = \sum_{i=1}^{p} A_i x_i$) can be summarized as follows.

The problem

$$\arg\min_{x_i \in S_i} \underbrace{\sum_{i=1}^{p} f_i(x_i) + \sum_{i=1}^{p} v_i^s h_i(x_i)}_{F(x)} + \underbrace{\frac{k^s}{2}\left|\sum_{i=1}^{p} h_i(x_i)\right|^2}_{G(x)},$$

is solved in the following manner. Let $w_G^0 \in \partial G(z_G^0)$, $\tilde{G}_1(\cdot) = G(z_G^0) + \langle w_G^0, \cdot - z_G^0\rangle$

$$z_F^s := \arg\min_{x \in S}\left\{\sum_{i=1}^{p} f_i(x_i) + \sum_{i=1}^{p} v_i^s h_i(x_i) + \tilde{G}_k(x) + \frac{1}{2\beta}\sum_{i=1}^{p}|x_i - x_i^s|^2\right\}; \qquad (11.9)$$

then set: $w_F^s = -w_G^s - \beta^{-1}(z_F^s - x^s)$, $\tilde{F}_s(\cdot) = F(z_F^s) + \langle w_F^s, \cdot - z_F^s\rangle$, and determine

$$z_G^s := \arg\min_{x \in S}\left\{\frac{k^s}{2}\left|\sum_{i=1}^{p} h_i(x_i)\right|^2 + \tilde{F}_s(x) + \frac{1}{2\beta}\|x - x^s\|^2\right\}, \qquad (11.10)$$

put: $w_G^s = -w_F^s - \beta^{-1}(z_G^s - x^s)$, $\tilde{G}_{s+1}(\cdot) = G(z_G^s) + \langle w_G^s, \cdot - z_G^s\rangle$ and repeat the procedure.

The problem (11.9) is decomposable and (11.10), in this case, is easy to solve because it is quadratic. Here, we note that this approach is interesting for the case of monotropic problems, however, for the case of nonlinear constraints, the second minimization obtained is a difficult problem and, moreover, it is notl decomposable.

Kontogiorgis and Meyer [49] proposed an extension to ADM where the penalty parameter was replaced by a positive-semidefinite matrix. Their iterative scheme to solve problem (11.8) can be presented as follows.

$$\begin{cases} x^{s+1} \in \operatorname{Arg\,min}_x \; f(x) + \langle u^s, H^\top Mx\rangle + 0.5\|Mx - z^s\|_{H^\top}^2. \\ z^{s+1} \in \operatorname{Arg\,min}_z \; h(z) + \langle u^s, H^\top z\rangle + 0.5\|Mx^{s+1} - z\|_{H^\top}^2 \\ u^{s+1} = u^s + (Mx^{s+1} - z^{s+1}). \end{cases}$$

Kontogiorgis and Meyer [49] obtained, in the convex case, a convergence result without assuming strict convexity.

In this part, we consider the following optimization problem,

$$\min_{x \in A} f(x), \qquad (11.11)$$

where A is a nonempty closed subspace of $\Re^n$. Studying this class of problems is motivated by the fact that there are many ways to transform separable convex models to the above form (11.11), where generally the coupling constraints between the different subsystems can be represented by a subspace of the product space of the copies of the primal and dual variables. Problem (11.11) can be rewritten as

$$\min_{x\in\Re^n}\{f(x)+\chi_A(x)\}, \tag{11.12}$$

where χ_A stands for the characteristic function of A.

If we assume further that $A\cap \mathrm{int}(\mathrm{dom}(f))\neq\emptyset$, then the necessary optimality condition for (11.11) is equivalent to the search for a zero of the sum of two operators ∂f and $\partial\chi_A$. As we know that $\partial\chi_A(x)=A^{\perp}$ for $x\in A$, we conclude that solving (11.11) is equivalent to solving a problem of the form (11.13) with[6] $T=\partial f$; that is,

$$\text{Find } x\in A \text{ and } y\in B=A^{\perp} \quad \text{such that } \ y\in Tx;$$

that is,[7]

$$\text{Find } (x,y)\in A\times B\cap Gr(T). \tag{11.13}$$

To this class of problems, Spingarn [87, 88] proposed a general scheme to solve by applying the proximal point algorithm to a new operator introduced by Spingarn, called the partial inverse T_A where its graph is defined as

$$Gr(T_A)=\{(x_A+y_B,y_A+x_B)\ \ :\ \ y\in Tx\},\quad \text{with } x_A=Proj_A(x).$$

The obtained algorithm is known as the partial inverse of Spingarn which can be summed up by these steps.

Partial Inverse Method of Spingarn (PIMS)

`Step 1.` Initialize: $x^0\in A,\ y^0\in B$ and $\lambda_0>0,\ s=0$.
`Step 2.` Determine: $u^s,\ v^s$ such that $x^s+y^s=u^s+v^s$ and

$$\frac{1}{\lambda_s}v_A^s+v_B^s\in T\left(u_A^s+\frac{1}{\lambda_s}u_B^s\right)$$

`Step 3.` If $(v^s,w^s)\in A\times B$, **Stop.**
Else: Go to 4.
`Step 4.` Update: $x^{s+1}=u_A^s,\ y^{s+1}=v_B^s,\ s\longleftarrow s+1$ and go back to 2.

The convergence of this method was studied by Spingarn [87, 88]. However, Step 2 turns to be difficult to implement in practice except for the case where $\lambda=1$. To this problem, Mahey, Oualibouch, and Pham [58] introduced the notion of scaled proximal decomposition on a graph of a maximal monotone operator, that induces an implementable algorithm similar to PIMS and where the convergence was accelerated by the mean of the proximal parameter λ.

To this goal, let us define the notion of proximal decomposition on a given graph of a maximal monotone operator T in $\Re^n$.

---

[6] ∂f is maximal monotone because f is convex, lower semicontinuous, and proper.

[7] $Gr(T)=\{(x,y)\in\Re^n\times\Re^n\ :\ y\in Tx\}$ denotes the graph of the operator T.

Definition 11.1. We call a scaled proximal decomposition of $(x,y) \in \Re^n \times \Re^n$ on the graph of T, the *unique couple* $(v,w) \in \Re^n \times \Re^n$ satisfying:

$$x + \lambda y = v + \lambda w \quad \text{and} \quad (v,w) \in Gr(T).$$

If we directly apply the scheme of this proximal decomposition to (11.13) with $T = \partial f$, we obtain:

$$v^s = (\lambda \partial f + I)^{-1}(x^s + \lambda y^s) \Longleftrightarrow x^s + \lambda y^s \in \lambda \partial f(v^s) + v^s$$

$$\Longleftrightarrow 0 \in \lambda \partial f(v^s) + v^s - (x^s + \lambda y^s)$$

$$\Longleftrightarrow 0 \in \partial f(v^s) + \frac{v^s - (x^s + \lambda y^s)}{\lambda}$$

$$\Longleftrightarrow v^s \in \operatorname{Arg}\min_z \left\{ f(z) + \frac{1}{2\lambda} \|z - x^s - \lambda y^s\|^2 \right\}$$

and

$$w^s = \frac{x^s + \lambda y^s - v^s}{\lambda}.$$

Thus, (11.13) can be solved by the (SPDG)[8]

Scaled Proximal Decomposition on the Graph (SPDG)

`Step 1.` Initialize: $x^0 \in A$ and $y^0 \in A^{\perp}$, $\lambda > 0$, $s = 0$.
`Step 2.` Proximal decomposition:
determine: v^k, w^k such that $v^s = \arg\min_{z \in \Re^n} \{ f(z) + (1/2\lambda) \|z - x^s - \lambda y^s\|^2 \}$
$w^s = \frac{1}{\lambda}(x^s + \lambda y^s - v^s)$
`Step 3.` Test: If $(v^s, w^s) \in A \times B$, **Stop.**
Else: Go to 4.
`Step 4.` Update: $x^{s+1} = v^s_A$, $y^{s+1} = w^s_B$, $s \longleftarrow s+1$ and go back to 2.

Mahey, Oualibouch, and Pham [58] showed that if ∂f is strongly monotone with coefficient ρ and Lipschitzian with constant L, then the SPDG converges linearly with a rate of convergence given by

$$\tau(\lambda) = \sqrt{1 - \frac{2\lambda\rho}{(1 + \lambda L)^2)}}.$$

For further results about the efficiency of SPDG and other related decomposition schemes see [58–60].

[8] (SPDG) scaled proximal decomposition on the graph.

If the objective is separable, that is,

$$f(x) = \sum_{i=1}^{p} f_i(x_i),$$

then, the proximal decomposition step in SPDG can be rewritten

$$v^s = \arg\min_{z_i} \left\{ \sum_{i=1}^{p} f_i(z_i) + \frac{1}{2\lambda} \sum_{i=1}^{p} |z_i - x_i^s - \lambda y_i^s|^2 \right\}, \quad \text{and}$$

$$w_i^s = \frac{1}{\lambda}(x_i^s + \lambda y_i^s - v_i^s), \qquad \forall i \in \overline{1,p}$$

which is equivalent to:

$$v_i^s = \arg\min_{z_i} \left\{ f_i(z_i) + \frac{1}{2\lambda} |z_i - x_i^s - \lambda y_i^s|^2 \right\}, \quad \text{and}$$

$$w_i^s = \frac{1}{\lambda}(x_i^s + \lambda y_i^s - v_i^s), \qquad \forall i \in \overline{1,p}$$

which shows clearly that SPDG is favorable for the solution of large-scale separable convex problems. The reader may refer to the thesis of A. Ouorou [69] for many applications in telecommunication problems.

We presented in this survey PIMS and SPDG which are not based on augmented Lagrangians to motivate the decomposition approaches developed by Hamdi, Mahey, and Dussault [38] known as SALA which stands for separable augmented Lagrangian algorithms. The departure point comes from the application of PIMS or SPDG to a dual problem of a separable convex model.

$$\min \left\{ f(x) = \sum_{i=1}^{p} f_i(x_i) \quad : \; g(x) = \sum_{i=1}^{p} g_i(x_i) = 0, \; x_i \in S_i, \; i = \overline{1,p} \right\} \tag{11.14}$$

and let its dual problem:

$$\max_u \; h(u) = \sum_{i=1}^{p} h_i(u) = \min \sum_{i=1}^{p} f_i(x_i) + u \sum_{i=1}^{p} g_i(x_i). \tag{11.15}$$

To make problem (11.15) separable, we propose using a copy of each dual variable; in other words we consider the equivalent model:

$$\max \; \sum_{i=1}^{p} h_i(u_i) = \sum_{i=1}^{p} \min\{f_i(x_i) + u_i g_i(x_i)\}.$$

$$u_i = u, \; i = \overline{1,p}$$

Setting $A = \{u \in \Re^p : u_1 = u_2 = \cdots = u_p\}$, (11.15) becomes

$$\max_{u \in A} h(u). \tag{11.16}$$

The optimality conditions are resumed now to find:

$$(w^s, y^s) \in A \times A^\perp \quad \text{such that } y^s \in \partial(-h(w^s)),$$

and by applying SPDG, we get the following iterative scheme:

$$u^s := \arg\max_u \left\{ h(u) - \frac{1}{2\lambda}\|u - w^s - \lambda y^s\|^2 \right\} \tag{11.17}$$

$$v^s := \frac{1}{\lambda}(w^s + \lambda y^s - u^s). \tag{11.18}$$

(11.17) is equivalent to solve for all $i = \overline{1,p}$.

$$\max_{u_i} \left\{ h_i(u_i) - \frac{1}{2\lambda}(u_i - w_i^s - \lambda y_i^s)^2 \right\}. \tag{11.19}$$

And because in our case the projection can be done explicitly so that $(y_A)_i = [Proj_A(y)]_i = 1/p\sum_{i=1}^p y_i$ and according to the fact that $\Re^p = A \oplus A^\perp$, we show easily that $(y_{A^\perp})_i = [Proj_{A^\perp}(y)]_i = y_i - (1/p)\sum_{i=1}^p y_i$, and finally the obtained algorithm can be resumed as follows:

$$\begin{cases} x_i^{s+1} := \arg\min_{x_i \in S_i} \left\{ f_i(x_i) + u^s g_i(x_i) + \frac{\lambda}{2} g_i^2(x_i) + \lambda y_i^s g_i(x_i) \right\} \\ u^{s+1} = u^s + \frac{1}{p}\sum_{i=1}^p \lambda g_i(x_i^{s+1}) \\ y_i^{s+1} = -g_i(x_i^s) + \frac{1}{p}\sum_{i=1}^p g_i(x_i^{s+1}). \end{cases}$$

One may observe that the subproblem is equivalent to miniming an augmented Lagrangian. Indeed, by adding constant terms to the expression giving x_i^{s+1}, that is,

$$x_i^{s+1} := \arg\min_{x_i \in S_i} \left\{ f_i(x_i) + u^s g_i(x_i) + \underbrace{u^s y_i^s} + \frac{\lambda}{2}\left\{ |g_i(x_i)|^2 + 2y_i^s g_i(x_i) + \underbrace{y_i^{s2}} \right\} \right\},$$

which is equivalent to

$$x_i^{s+1} := \arg\min_{x_i \in S_i} \underbrace{\left\{ f_i(x_i) + (g_i(x_i) + y_i^s)u^s + \frac{\lambda}{2}|g_i(x_i) + y_i^s|^2 \right\}}_{L_A(x_i, y_i^s, u^s, \lambda)},$$

where $L_A(x_i, y_i^s, u^s, \lambda)$ is the associated augmented Lagrangian to the following problem

$$\text{minimiser} \quad f_i(x_i)$$
$$g_i(x_i) + y_i = 0, \qquad i = \overline{1,p}.$$

and by denoting $r^{s+1} = \sum_{i=1}^{p} g_i(x_i^{s+1})$, the above algorithm becomes:

$$\begin{cases} x_i^{s+1} := \arg\min_{x_i \in S_i} \left\{ f_i(x_i) + (g_i(x_i) + y_i^s)u^s + \frac{\lambda}{2}|g_i(x_i) + y_i^s|^2 \right\} \\ u^{s+1} = u^s + \frac{\lambda}{p} r^{s+1} \\ y_i^{s+1} = -g_i(x_i^{s+1}) + \frac{1}{p} r^{s+1}. \end{cases}$$

The convergence is guaranteed by the obtained results for SPDG in the convex case.

The above algorithm is equivalent to the algorithm SALA developed for solving convex and nonconvex separable models. SALA was proposed to avoid any linear approximation making the nonseparable augmented Lagrangian favorable to decomposition and to parallel computing. To compensate for this drawback, Hamdi, Mahey, and Dussault [38] proposed an iterative scheme that can be derived from the resource directive subproblems associated with the coupling constraints. The main idea of this algorithm is to add an allocation vector to the constraints to apply a classical augmented Lagrangian algorithm with partial elimination of the constraints that induces a separable augmented Lagrangian. We present the algorithm SALA for the large block separable nonlinear constrained optimization problem (11.14).

Define an allocation vector $y = (y_1, \ldots, y_p)$ with $y_i \in \Re^m$ such that

$$y \in A = \left\{ (y_1, \ldots, y_p) \in \Re^{mp} | \sum_{i=1}^{p} y_i = 0 \right\}$$

and we get an equivalent problem

$$(\overline{SEP}) \qquad \begin{cases} \min & \sum_{i=1}^{p} f_i(x_i) \\ \text{such that} & g_i(x_i) + y_i = 0, \;\; i = \overline{1,p} \\ & \sum_{i=1}^{p} y_i = 0 \\ & x_i \in S_i, \;\; i = \overline{1,p} \end{cases}$$

The expression of the augmented Lagrangian function associated with the problem $\overline{SEP}$ is written

$$L(x,y,v,k) = \sum_{i=1}^{p} f_i(x_i) + \sum_{i=1}^{p} \langle v_i, g_i(x_i) + y_i \rangle + \frac{k}{2} \sum_{i=1}^{p} \|g_i(x_i) + y_i\|^2.$$

By using the multiplier method we get the following steps:

Separable Augmented Lagrangian Algorithm (SALA)

`Step 1.` Initialize: $v^0 \geq 0,\ v^0 \in V,\ \beta \geq 1,\ \varepsilon_0 > 0,\ y^0 \ :\ \sum_{i=1}^{m} y_i^0 = 0,\ k^0 > 0,$ $s = 0.$

`Step 2.` Determine: $\forall i = \overline{1,p}$

$$x_i^{s+1} = \arg\min_{x_i \in S_i} f_i(x_i) + \langle v^s, g_i(x_i) + y_i^s \rangle + \frac{k^s}{2} \|g_i(x_i) + y_i^s\|^2.$$

`Step 3.` Update:

$$\begin{cases} v^{s+1} = v^s + \frac{k^s}{p} r^{s+1}, \;\; r^{s+1} = \sum_{i=1}^{p} g_i(x_i^{s+1}) \\ y_i^{s+1} = -g_i(x_i^{s+1}) + \frac{1}{p} r^{s+1}, \; \forall i = \overline{1,p} \\ k^{s+1} = \beta k^s, \end{cases}$$

and go back to 2.

For the convergence analysis of SALA for solving convex and nonconvex problems, we refer to Hamdi, Mahey, and Dussault [38]. The version of SALA for an inequality constrained problem is given in [41]. In the case equality constraints, another version of SALA was developed by Hamdi and Mahey ([39], 2000), where a diagonal scheme is used. In other terms, the algorithm called DSALA can be seen as an Arrow–Hurwicz scheme version of SALA, where it alternates one iteration in the primal space followed by one iteration in the dual space. DSALA offers a possibility to use second-order updates for the multipliers which increases the efficiency of this primal–dual decomposition method.[9]

Diagonal Separable Augmented Lagrangian Algorithm (DSALA)

`Step 1.` Initialize: $x^0, \; \beta \geq 1, \; \varepsilon, \; \varepsilon_0 > 0, \; \sum_{i=1}^{m} y_i^0 = 0, \; \lambda_0 > 0, \; s = 0$
`Step 2.` Determine:

$$\forall i = \overline{1,p} \;\; x_i^{s+1} = x_i^s - [H_i^s]^{-1} \nabla_x L_A^i(x_i^s, y_i^s, u^s, \lambda_s)$$

`Step 3.` Compute the residual $r^{s+1} = \sum_{i=1}^{p} g_i(x_i^{s+1})$
If : $\|r^{s+1}\|_\infty < \varepsilon$ stop.
Else : go to step 4
`Step 4.` Update:

$$\begin{cases} u^{s+1} = u^s + \frac{\lambda_s}{p} r^{s+1}, \\ y_i^{s+1} = \theta_i(x_i^{s+1}, y_i^s, u^s, \lambda_s), \;\; i = 1, \ldots, p \\ \lambda_{k+1} = \beta \lambda_s, \qquad \text{and return to step 2} \end{cases}$$

with

$$\theta_i(x_i^{s+1}, y_i^s, u^s, \lambda_s) = -g_i(x_i^{s+1}) + \sum_{i=1}^{p} \frac{g_i(x_i^{s+1})}{p}.$$

[9] H_i^k stands for the Hessian matrix of the augmented Lagrangian.

One of the drawbacks of the SALA algorithm, when applied to the inequality constrained problem, is that the corresponding augmented Lagrangian function is not twice differentiable even if the cost and constraint functions are, and then we cannot apply the second-order methods, for instance, Newton or Newton-type methods, to minimize the augmented Lagrangians. This problem of the lack of smoothness of augmented Lagrangians was studied thoroughly by Zibulevsky and Ben-Tal ([101], 1997), Polyak ([75, 76]), and by Auslender, Teboulle and Ben-Tiba ([3, 4] 1999). But in the context of decomposition methods, we can cite the seminal paper of Auslender and Teboulle [5], where the authors used their new proxlike kernels [3] to develop a decomposition scheme based on a new C^∞ nonquadratic augmented Lagrangian. Their algorithm called the entropic proximal decomposition can be presented as follows.

Entropic Proximal Decomposition Algorithm (EPDA)

`Step 1.` Given ϕ defined in (11.7),

$$(v_j^0, p_j^0, y_j^0) \in \Re^{d_j} \times \Re^m \times \Re^m \quad \text{for } j = 1,\dots,n.\ (u^0, w^0) \in \Re^{mn}_+ \ \text{and } \lambda_0 > 0.$$

`Step 2.` Compute:

$$p_j^{s+1} = y_j^s + (2\theta)^{-1}\lambda_s(w^s - u_j^s), \qquad j = 1,\dots,n.$$

`Step 3.` For each $j = 1,\dots,n$, find

$$v_j^{s+1} = \arg\min\{L^s_{\lambda_s,j}(u^s, v_j) \ : \ v_j \in \Re^{d_j}\}.$$

`Step 4.` Compute

$$\begin{cases} w_i^{s+1} = w_i^s(\phi^*)'(-\lambda_s \sum_{j=1}^n p_{ij}^{s+1}/w_i^s),\ i = 1,\dots,m, \\ u_{ij}^{s+1} = u_{ij}^s(\phi^*)'(\lambda_s(g_{ij}(x^{s+1}) + p_{ij}^{s+1})/u_{ij}^s),\ i = 1,\dots,m, \\ y_j^{s+1} = y_j^s + (2\theta)^{-1}\lambda_s(w^{s+1} - u_j^{s+1}),\ j = 1,\dots,n, \end{cases}$$

where their augmented Lagrangian is given by

$$L^s_{\lambda_s,j}(u^s, v_j) = f_j(x_j) + \lambda_s^{-1}\sum_{i=1}^m (u_{ij}^s)^2\phi^*(\lambda_s(p_{ij}^{s+1} + g_{ij}(x_j))/u_{ij}^s)$$
$$+ \theta\lambda_s^{-1}\|x_j - x_j^s\|^2.$$

The above entropic proximal decomposition algorithm can be seen as a proximal multiplier-type method. Such technique was developed later and studied by Hamdi in [40] in the case of algorithms of type SALA. In the same period Kyonno and Fukushima ([52], 2000) extended the method of Chen and Teboulle by replacing the quadratic proximal terms by Bregman proximal distances.[10] Their algorithm called

[10] It cannot be considered as an augmented Lagrangian-based decomposition approach.

the nonlinear predictor-corrector proximal multiplier method (NPCPMM) solves problem (11.8) and can be summed up as follows.

Nonlinear Predictor Corrector Proximal Multiplier Method (NPCPMM)

`Step 1.` Given a Bregman function ψ and ϕ, $\mu > 0$, $\nu > 0$, choose (x^0, z^0, y^0) and $\varepsilon \in (0, \bar{c}/2)$, where

$$\bar{c} = \min\left\{\frac{\sqrt{\nu}}{2}, \frac{\sqrt{\mu}}{2\|M\|}\right\}$$

and let $s = 0$.

`Step 2.` Choose a_s, b_s and c_s such that $\varepsilon < c_s < \bar{c} - \varepsilon$ and compute: $p^{s+1} = y^s + c_s(Mx^s - z^s)$.

`Step 3.` Find (x^{s+1}, γ^{s+1}) and (z^{s+1}, δ^{s+1}) such that

$$\begin{cases} \gamma^{s+1} \in \partial_{a_s} f^s(x^{s+1}), & \gamma^{s+1} + c_s^{-1}\{\nabla\phi(x^{s+1}) - \nabla\phi(x^s)\} = 0, \\ \delta^{s+1} \in \partial_{b_s} h^s(z^{s+1}), & \delta^{s+1} + c_s^{-1}\{\nabla\psi(x^{s+1}) - \nabla\psi(z^s)\} = 0, \end{cases}$$

where $f^s(x) = f(x) + \langle p^{s+1}, Mx\rangle$ and $h^s(z) = h(z) - \langle p^{s+1}, z\rangle$.

`Step 4.` Let $y^{s+1} = y^s + c_s(Mx^{s+1} - z^{s+1})$, and go back to Step 1.

Some similar algorithms have been applied to block-structured linear problems (see [34], [37]) but it is important to emphasize the role of the constraint structure in applying a decomposition procedure and its influence on the performance of the resulting algorithm.

In 2005, Hamdi [40] proposed a proximal multiplier version of SALA offering more numerical stability but still sensitive to the proximal and the penalty parameters of the modified algorithm. The proposed regularization was applied first in the primal space, secondly in the dual space, and finally in both spaces of the same time. Stable numerical results were obtained. In 2006, using the nonlinear principle of Polyak, Hamdi [41] proposed a generalization of SALA to overcome the nonsmoothness of the augmented Lagrangian when applied to solve inequality constrained separable problems. The algorithm can be seen as a direct application of SALA with a modified Lagrangian. The modified Lagrangian is the classical Lagrangian of the rescaled problem. The main iterations can be summarized as follows.

Separable Augmented Lagrangian Algorithm (φ SALA)

`Step 1.` Select $\varphi \in \Phi$[11], $u^0 \in \Re^{mp}_{++}$ where $u_j \in V$, $j = \overline{1,m}$, $\lambda > 0$, $y^0 = (y_1, \ldots, y_m) \,:\, \sum_{j=1}^m y^0_{ij} = 0$, $j = \overline{1,m}$ and $\lambda^0_j = \lambda(u^0_j)^{-1}$.

`Step 2.` Determine: for any $i = \overline{1,p}$

$$x_i^{s+1} := \arg\min_{x_i \in \Re^{n_i}} \left\{ f_i(x_i) - \sum_{j=1}^{m} \frac{1}{\lambda_j^s} u_j^s \varphi(\lambda_j^s(g_{ij}(x_i) + y_{ij}^s)) \right\}.$$

[11] Φ denotes the class of functions φ satisfying some properties as in [74].

Step 4. Update and go back to step 2:

$$\begin{cases} y_{ij}^{s+1} = -g_{ij}(x_i^{s+1}) + \delta_j^{s+1} \quad i = \overline{1,p}\,, \quad j = \overline{1,m}. \\ u_j^{s+1} = u_j^s \varphi'(\lambda_j^s \delta_j^{s+1}), \\ \lambda_j^{s+1} = \lambda (u_j^{s+1})^{-1}. \end{cases}$$

Recently, in 2003 Gueye, Dussault, and Mahey [35] worked on a modification of SALA to solve large-scale convex programs with separable structure by using a multidimensional scaling matrix Λ instead of the penalty parameter λ in SALA.

Let $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ be a diagonal positive-definite matrix . Then the constraints can be transformed to the equivalent ones as follows.

$$\sum_{j=1}^{p} \Lambda g_j(x_j) = 0,$$

and as it was done in SALA, an allocation vector $y = (y_1, y_2, \ldots, y_p)$ with $y_j \in \mathbb{R}^{mp}$ is defined and we get an equivalent problem:

$$(PE) \quad \begin{cases} \min \sum_{j=1}^{p} f_j(x_j) \\ \text{subject to } \Lambda g_j(x_j) + y_j = 0, & j = \overline{1,p} \\ x_j \in S_j, \sum_{j=1}^{p} y_j = 0, & j = \overline{1,p}, \end{cases}$$

and finally, the following algorithm (SALAMS) is obtained:

Separable Augmented Lagrangian Algorithm with Multidimensional Scaling (φ SALAMS)

Step 1. Select:

$$\Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_m\}, \qquad \Lambda_0 > 0, \qquad u^0 \geq 0, \qquad \varepsilon_0 > 0, \qquad y^0 \in A^{\perp},$$
$$\text{and} \quad s = 0$$

Step 2. Determine:

$$\forall j = \overline{1,p} \qquad x_j^{s+1} = \arg\min_{x_j} f_j(x_j) + u^s(\Lambda^s g_j(x_j) + y_j^s) + \frac{1}{2}\|\Lambda^s g_j(x_j) + y_j^s\|^2$$

Step 3. Calculate the residual

$$r^{s+1} = \sum_{j=1}^{p} g_j(x_j^{s+1}) :$$

and update:

$$\begin{cases} u^{s+1} = u^s + \frac{1}{p}\Lambda^t r^{s+1} \\ y_j^{s+1} = -\Lambda^s g_j(x_j^{s+1}) + \frac{1}{p}\Lambda^s r^{s+1}, \qquad \forall j = \overline{1,p} \end{cases}$$

and return to step 2.

Their convergence analysis was done only for affine constraints.

To finish this survey, we cite some of the latest works on decomposition schemes based on augmented Lagrangians. A. Ouorou [70] in 2002 proposed a Gauss–Seidel method to solve

$$\begin{array}{l} \min\ [F(x) = \sum_{j=1}^{n} F_j(x_j)] \\ \text{s.t.} \\ \qquad Ax \geq b,\ x \in \Re^n, \end{array} \tag{11.20}$$

where each $F_j : \Re \to]-\infty, +\infty]$ is a closed proper convex function, A is a $p \times n$ matrix having no zero row, and $b \in \Re^m$. Let A^j and A_i denote, respectively, the column j and the row i of matrix A. By rewriting the dual problem as follows,

$$\max\{b^\top y - F^*(z) \ : A^\top y = z,\ \ y \in \Re^m_+,\ z \in \Re^n.\}, \tag{11.21}$$

the original problem has been replaced by a game with two players where the criteria were two different quadratic augmented Lagrangian functions defined for the primal and the dual problems. A primal–dual algorithm has then been proposed to reach an equilibrium of this game to solve problem (11.20) with equality constraints and its dual. In 2008, Hamdi and Al-Saud [42] proposed a generalization of the proposed scheme in [70] by using generalized augmented Lagrangian functions. The twin augmented Lagrangian decomposition method is based on the use of an augmented Lagrangian in the dual and primal spaces as follows.

$$L_D(y,z,x;\lambda) := b^\top y - F^*(z) + \langle x, z - A^\top y\rangle - \frac{\lambda}{2}\|z - A^\top y\|^2, \tag{11.22}$$

where λ is a positive penalty parameter, and L_D denotes the augmented Lagrangian in the dual space. Here, we considered x as a vector of multipliers associated with the constraints $z - A^\top y = 0$.

In the primal space, we rescaled the linear constraints according to Polyak's nonlinear rescaling principle to get the following primal augmented Lagrangian.

$$L_P(x,y,z;\lambda) := \sum_{j=1}^{n} F_j(x_j) - z^\top x + b^\top y + \frac{1}{\lambda}\sum_{i=1}^{m} y_i[\lambda(A_i x - b_i) - \psi(\lambda(A_i x - b_i))]. \tag{11.23}$$

and the Gauss–Seidel scheme can be summed as follows.

Twin Augmented Lagrangian (φ T)

Step 1. Initialize: $x^0 \in \Re^n$, $y^0 \in \Re^m_+$ and $\lambda > 0$. Define $z^0 := A^\top y^0$ and set $s = 0$.

Step 2. Compute for each $j = 1, \dots, n$,

$$x_j^{s+1} :\longrightarrow \min_{x_j} F_j(x_j) - \langle z_j^s, x_j \rangle + \frac{1}{\lambda} \sum_{i=1}^{m} y_i^s [\lambda (A_i x_{[j]} - b_i) - \psi(\lambda (A_i x_{[j]} - b_i))],$$

Step 3. Compute for each $j = 1, \dots, n$,

$$z_j^{s+1} \to \max_{z_j} \left\{ -F_j^*(z_j) + \langle x_j^{s+1}, z_j \rangle - \frac{\lambda}{2} (z_j - (A^\top y)_j)^2 \right\}.$$

Step 4. Compute for each $i = 1, \dots, m$,

$$y_i^{s+1} = \max \left\{ 0, \; \frac{1}{\lambda \|A_i\|^2} [b_i - \langle A_i, \, x^{s+1} \rangle + \lambda \langle A_i, \, z^{s+1} - A_{[i]}^\top y_{[i]}^s] \right\}.$$

Set $s = s + 1$ and go back to step 2.

11.5 Conclusion

In this chapter, we propose a nonexhaustive survey about decomposition algorithms for solving structured large-scale convex or nonconvex optimization problems. We focus particularly on all splitting approaches based on the classical augmented Lagrangian functions or on any other modified Lagrangians in general. This study covers the last 40 years of research in nonlinear programming, where the researchers studied the nice theoretical properties of augmented Lagrangians and at the same time were interested in solutions to remedy the main drawbacks related to the corresponding methods. The first one concerns the fact that augmented Lagrangians and/or many other modified (or generalized) Lagrangians are no longer separable even when the original problem is separable, which is not favorable to decomposition schemes. The second drawback is the fact that these Lagrangians are only differentiable once even when the problem's data allow for higher differentiability, disabling the application of efficient Newton-type methods. In fact such a lack of continuity in the second derivative can significantly slow down the rate of convergence of these algorithms and thus cause algorithmic failure. The authors would like to underline the existence of many works regarding some decomposition schemes based on augmented Lagrangians for variational inequality problems ([61, 43–45]) and references therein.

References

1. K.J. Arrow, F.J. Gould, S.M. Howe: A general saddle point result for constrained optimization. *Math. Program.*, 2, 2 (1973), pp. 225–234.
2. K.J. Arrow, L. Hurwicz, H. Uzawa: *Studies in Linear and Nonlinear Programming*, Ed. Arrow-Hurwicz, Uzawa, 1958, Stanford University Press, USA.
3. A. Auslender, M. Teboulle, S. Ben-Tiba: A logarithmic-quadratic proximal method for variational inequalities. *Comput. Optim. Appl.*, 12 (1999), pp. 31–40.
4. A. Auslender, M. Teboulle, S. Ben-Tiba: Interior proximal and multiplier methods based on second order homogeneous functionals. *Math. Ope. Res.*, 24 (1999), pp. 645–668.
5. A. Auslender, M. Teboulle: Entropic proximal decomposition methods for convex programs and variationnal inequalities. *Math. Program.*, Ser. A 91 (2001), pp. 33–47.
6. P.T. Boggs, J.W. Tolle: Augmented Lagrangians which are quadratic in the multiplier. *J.O.T.A*, 21 (1980), pp. 17–26.
7. D.P. Bertsekas: On the method of multipliers for convex programming. *IEEE Trans. Automatic Control*, AC 20 (1973), pp. 385–388.
8. D.P. Bertsekas: Multiplier methods: A survey. *Automatica*, 12 (1976), pp. 133–145.
9. D.P. Bertsekas: Approximation procedures based on the method of multipliers. *J.O.T.A*, 23 (1977), pp. 487–510.
10. D.P. Bertsekas: Convexification procedures and decomposition methods for non convex optimization problems. *J.O.T.A*, 29 (1979), pp. 169–197.
11. D.P. Bertsekas: *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, New York, 1982.
12. J.T. Betts: An accelerated multiplier method for nonlinear programming. *J.O.T.A*, 21, 2 (1977), pp. 137–227.
13. J.T. Betts: A gradient projection-multiplier method for nonlinear programming. *J.O.T.A*, 24, 4 (1978), pp. 523–548.
14. J.D. Buys: Dual algorithms for constrained optimization problems. Doctoral dissertation, University of Leiden, 1972.
15. R.H. Byrd: Local convergence of the diagonalized method of multipliers. *J.O.T.A*, 26, 4 (1978), pp. 485–500.
16. Y. Censor, S.A. Zenios: The proximal minimization with D-function. *J.O.T.A*, 73 (1992), pp. 451–464.
17. G. H. Chen, R.T. Rockafellar: Convergence rates in forward-backward splitting methods in Lagrangian optimization. *S.I.A.M. J. Optim.*, 7 (1997), pp. 421–444.
18. G. Chen, M. Teboulle: A proximal-based decomposition method for convex minimization problems. *Math. Program.*, 64 (1994), pp. 81–101.
19. G. Cohen: Décomposition et coordination en optimisation déterministe, différentiable et non différentiable. Thése d'Etat, Univ Paris Dauphine, 1984, France.
20. I.D. Coope, R. Fletcher: Some numerical experience with a globally convergent algorithm for nonlinear constrained optimization. *J.O.T.A*, 32, 4 (1980), pp. 1–16.
21. G. Di Pillo, L. Grippo: A new class of augmented Lagrangians in nonlinear programming. *S.I.A.M. J. Control Optim.*, 17, 5 (1979), pp. 618–628.
22. G. Di Pillo, L. Grippo: A new augmented Lagrangians function for inequality constraints in nonlinear programming problems. *J.O.T.A*, 36, 4 (1982), pp. 495–519.
23. J.E. Eckstein: Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming. Working Paper 91-004, Harvard Business School (1990).
24. J.E. Eckstein: Parallel alternating direction multiplier decomposition of convex programs. *J.O.T.A*, 80 (1994), pp. 39–62.
25. J.E. Eckstein, D.P. Bertsekas: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.*, 55 (1992), pp. 293–318.
26. M.C. Ferris, O.L. Mangasarian: Parallel constraint distribution. *S.I.A.M. J. Optim.*, 1, 4 (1991), pp. 487–500.

27. R. Fletcher: *Methods for Linear Programming, Integer and Nonlinear Programming*. Ed. J. Abadie, North-Holland, Amsterdam, 1970.
28. R. Fletcher: *A Class of Methods for Nonlinear Programming with Termination and Convergence*. North-Holland, Amsterdam, 1970.
29. R. Fletcher: *Methods Related to Lagrangians Functions. Numerical Methods for Constrained Optimization*, Ed. P.E. Gill and W. Murray, 1974, Academic Press, New York, pp. 219–235.
30. R. Fletcher: An ideal penalty function for constrained optimization. *Nonlinear programming 2*, Ed. O.L. Mangasarian, R.R. Meyer, S.M. Robinson, Academic Press, New York, (1975), pp. 121–163.
31. M. Fortin, R. Glowinski: *Méthodes de Lagrangien augmenté*. Dunod, Paris, 1982.
32. A. Glad: Properties of updating methods for the multipliers in augmented Lagrangians. *J.O.T.A*, 25, 4 (1978), pp. 181–193.
33. R. Glowinski, P. Le Tallec: *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. SIAM Studies in Applied Mathematics, Philadelphia, 1989.
34. E.G. Golshtein: The block method of convex programming, *Soviet Math. Dokl.*, 33 (1986), pp. 584–587.
35. O.M. Gueye, J.P. Dussault, P. Mahey: Separable augmented Lagrangian algorithm with multidimensional scaling for monotropic programming. *J.O.T.A*, 127–2 (2005), pp. 329–345.
36. P.C. Haarhoff, J.D. Buys: A new method for the optimization of a nonlinear function subject to nonlinear constraints. *Comput. J.*, 13 (1970), pp. 178–184.
37. C.D. Ha: Decomposition methods for structured convex programming, Ph.D. Thesis, Univ. Wisconsin-Madison, 1980.
38. A. Hamdi, P. Mahey, J.P. Dussault: A new decomposition method in non convex programming via separable augmented Lagrangians. in *Recent Advances in Optimization*, Gritzmann, Horst, Sachs, and Tichatschke Eds, Lecture Notes in Economics and Mathematical systems, Springer, New York, 452 (1997), pp. 90–104.
39. A. Hamdi, P. Mahey: Separable diagonalized multiplier method for decomposing nonlinear programs. *Comput. Appl. Math.*, 19, 1 (2000), pp. 1–29.
40. A. Hamdi: Two-level primal-dual proximal decomposition technique to solve large scale optimization problems. *Appl. Math. Comput.*, 106–3 (2005), pp. 921–938.
41. A. Hamdi: Decomposition for structured convex programs with smooth multiplier methods. *Appl. Math. Comput.*, 169/1 (2006), pp. 218–241.
42. A. Hamdi, H. Al-Saud: A Re-scaled twin augmented Lagrangian method for saddle point seeking. *Nonlinear Anal.: Theory, Meth. Appl.*, 69, 12, 15 December 2008, pp. 4796–4802.
43. D. Han, Wenyu Sun: New decomposition methods for solving variational inequality problems. *Math. Comput. Model.*, 37 (2003), pp. 405–418.
44. B.S. He, H. Yang, L.Z. Liao: Decomposition method for a class of monontone variational inequality problem. *J.O.T.A*, 103–3 (1999), pp. 603–622.
45. B.S. He, H. Yang, L.Z. Liao, D. Han: A new inexact alternating directions method for monotone variational inequalities. *Math. Program.*, 92–1 (2002), pp. 103–118.
46. M.R. Hestenes: Multiplier and gradient methods. *J.O.T.A*, 4 (1969), pp. 303–320.
47. A.N. Iusem, M. Teboulle: Convergence analysis of non quadratic proximal methods for convex and linear programming. *Math. Oper. Res.*, 20, 3 (1995), pp. 657–677.
48. K.C. Kiwiel, C.H. Rosa, A. Ruszczyński: *Decomposition via Alternating Linearization*. WP-95-051 June 1995 IIASA, Austria.
49. S. Kontogiorgis, R.R. Meyer: A variable-penalty alternating direction method for convex optimization. Technical report 95–18 (1995).
50. B.W. Kort: Rate of convergence of the method of multipliers with inexact minimization. *Nonlinear Programming 2*, Ed. O.L. Mangasarian, R. Meyer, S.M. Robinson, Academic Press, New York (1975), pp. 193–214.
51. B.W. Kort, D.P. Bertsekas: Combined primal-dual and penalty methods for convex programming. *S.I.A.M. J. Control Optim.*, 14 (1976), pp. 268–294.
52. M. Kyono, M. Fukushima: Nonlinear proximal decomposition method for convex programming. *J.O.T.A*, 106, 2 (2000), pp. 357–372.

53. S.A. Lill: Generalization of an exact method for solving equality constrained problems to deal with inequality constrains. *Numerical Methods for Nonlinear Optimization*, Ed. F.A. Lootsma, 1972, Academic Press, New York, pp. 383–393.
54. P.L. Lions, B. Mercier: Splitting algorithms for the sum of two nonlinear operators. *S.I.A.M. J. Numer. Anal.*, 16, 6 (1979), pp. 964–979.
55. S. Lucidi: New results on a class of exact augmented Lagrangians. *J.O.T.A*, 58, 2 (1988), pp. 259–282.
56. D.G. Luenberger: *Introduction to Linear and Nonlinear Programming*. Addison-Wesley, Reading, MA, 2nd Ed 1984.
57. O.L. Mangasarian: Unconstrained Lagrangians in nonlinear programming. *S.I.A.M. J. Control*, 13 (1975), pp. 772–791.
58. P. Mahey, S. Oualibouch, D.T. Pham: Proximal decomposition on the graph of a maximal monotone operator. *S.I.A.M. J. Optim.*, 5, 2 (1995), pp. 454–466.
59. P. Mahey: Separable augmented Lagrangians for the decomposition of large convex programs. *Investigación Operativa*, 5, 1 (1995), pp. 1–26.
60. P. Mahey, J.P. Dussault, A. Benchakroun, A. Hamdi: *Adaptive Scaling and Convergence Rates of a Separable Augmented Lagrangian Algorithm*. Lecture Notes in Economics and Mathematical Systems (1999) Ed. by V.H. Nguyen, J.J. Strodiot and P. Tossings. Vol. 523 (1999), Springer, New York, pp. 59–67.
61. S. Makler-Scheimberg, V.H. Nguyen, J.J. Strodiot: A family of perturbed parallel decomposition methods for variational inequalities. *J.O.T.A*, 89, 2 (1996), pp. 423–452.
62. K. Martensson: A new approach to constrained function optimization. *J.O.T.A*, 12, 6 (1973), pp. 531–554.
63. A. Miele, E.E. Cragg, R.R. Iver, A.V. Levy: Use of the augmented penalty function in mathematical programming. Part I, *J.O.T.A*, 8 (1971), pp. 115–130.
64. A. Miele, E.E. Cragg, A.V. Levy: Use of the augmented penalty function in mathematical programming. Part II, *J.O.T.A*, 8 (1971), pp. 131–153.
65. A. Miele, P.E. Mosely, A.V. Levy, G.M. Coggings: On the method of multipliers for mathematical programming problems. *J.O.T.A*, 10, 1 (1972), pp. 1–33.
66. H. Mukai, E. Polak: A quadratically convergent primal-dual algorithm with global convergence properties for solving optimization problems with equality constraints. *Math. Program.*, 9, 3 (1975), pp. 336–349.
67. J.M. Mulvey, A. Ruszczynski: A new scenario decomposition method for large-scale stochastic optimization. *Oper. Res.*, 43, 3 (1995), pp. 477–490.
68. J.M. Mulvey, A. Ruszczyński: A diagonal quadratic approximation method for large scale linear programs. *Oper. Res. Lett.*, 12 (1992), pp. 205–215.
69. A. Ouorou: Décomposition proximale des problèmes de multiflots à caractère convexe. Application aux problèmes de routage dans les réseaux de communication. Doctorat de l'université Blaise Pascal, Clermont-Ferrand, 1995.
70. A. Ouorou: A Primal-dual algorithm for monotropic programming and its application to network optimization. *Comput. Optim. Appl.*, 15 (2002), pp. 125–143.
71. D.A. Pierre, M.J. Lowe: *Mathematical Programming via Augmented Lagrangians. Applied Mathematics and Computation*, Addison-Wesley, Reading, MA, 1975.
72. E. Polak, A.L. Tits: A globally convergent, implementable multiplier method with automatic penalty limitation. *Appl. Math. Optim.*, 6 (1980), pp. 335–360.
73. R. Polyak: Modified barrier function (theory and methods). *Math. Program.*, 54 (1992), pp. 177–222.
74. R. Polyak: Log-sigmoid multipliers method in constrained optimization. *Ann. Oper. Res.*, 101 (2001), pp. 427–460.
75. R. Polyak: Nonlinear rescaling multiplier method as interior quadratic prox. *Math. Program.*, (2003).
76. R. Polyak, I. Griva: Primal-dual nonlinear rescaling methods for convex optimization. *J.O.T.A*, 122–1 (2004), pp. 111–156.
77. M.J.D. Powell: *A Method for Nonlinear Constraints in Minimization Problems*. (R. Fletcher, ed.) Academic Press, New York, 1969.

78. R.T. Rockafellar: The multiplier method of Hestenes and Powell applied to convex programming. *J.O.T.A*, 12 (1973), pp. 555–562.
79. R.T. Rockafellar: A dual approach to solving nonlinear programming problems by unconstrained optimization. *Math. Prog.*, 5 (1973), pp. 354–373.
80. R.T. Rockafellar: Augmented Lagrange multiplier function and duality in non convex programming. *S.I.A.M. J. Control*, 12 (1974), pp. 268–285.
81. R.D. Rupp: Approximation of the classical isopemetric problem. *J.O.T.A*, 9 (1972), pp. 251–264.
82. R.D. Rupp: A method for solving a quadratic optimal control. Ibid, 9 (1975), pp. 238–250.
83. R.D. Rupp: A nonlinear optimal control minimization technique. *Trans. Amer. Math. Soc.*, 78 (1973), pp. 357–381.
84. R.D. Rupp: On the combination of the multiplier method of Hestenes and Powell with Newtons's method. SUNY, Rep, Albany, NY 1973.
85. A. Ruszczynski: On convergence of an augmented Lagrangian decomposition method for sparse convex optimization. *Math. Oper. Res.*, 20, 3 (1995), pp. 634–656.
86. L. Shin: Complete decomposition algorithm for non-convex separable optimization problems and applications. *Automatica*, 28, 6 (1992), pp. 1249–1254.
87. J.E. Spingarn: Applications of the method of partial inverse to convex programming decomposition. *Math Prog*, 32 (1985), pp. 199–223.
88. J.E. Spingarn: Partial inverse of a monotone operator. *Appl. Math. Optim.*, 10 (1983), pp. 247–265.
89. G. Stephanopoulos, A.W. Westerberg: The use of Hestenes's method of multipliers to resolve dual gaps in engineering system optimization. *J.O.T.A*, 15 (1975), pp. 285–309.
90. A. Tanikawa, H. Mukai: A new technique for non convex primal-dual decomposition of a large-scale separable optimization problem. *IEEE Trans. Autom. Control*, AC-30, 2, 1985, pp. 133–143.
91. R.A. Tapia: Diagonalized multiplier methods and quasi-Newton methods for constrained minimization. *J.O.T.A*, 22 (1977), pp. 135–194.
92. P. Tatjewski: New dual-type decomposition algorithm for non convex separable optimization problems. *Automatica*, 25, 2 (1989), pp. 233–242.
93. P. Tatjewski, B. Engelmann: Two-level primal-dual decomposition technique for large-scale non convex optimization problems with constraints. *J.O.T.A*, 64, 1 (1990), pp. 183–205.
94. M. Teboulle: Entropic proximal mappings with application to nonlinear programming. *Math. Oper. Res.*, 17 (1992), 670–690.
95. M. Teboulle: Convergence of proximal-like algorithms. *S.I.A.M. J. Optim.*, 7–4 (1997), pp. 1069–1083.
96. S.S. Tripathi, K.S. Narendra: Constrained optimization problems using multiplier methods. *J.O.T.A*, 9 (1972), pp. 59–70.
97. P. Tseng, D.P. Bertsekas: On the convergence of the exponential multiplier method for convex programming. *Math., Program.*, 60 (1993), pp. 1–19.
98. P. Tseng, Alternating: Projection-proximal methods for convex programming and variational inequalities. *S.I.A.M. J. Optim.*, 7 (1997), pp. 951–965.
99. P. Tseng: A modified forward-backward splitting method for maximal monotone mappings. *S.I.A.M. J. Control Optim.*, 38, pp. 431–446 (1998).
100. N. Watanabe, Y. Nishimura, M. Matsubara: Decomposition in large system optimization using the method of multipliers. *J.O.T.A*, 25 (1978), pp. 181–193.
101. M. Zibulevsky, A. Ben-Tal: Penalty-barrier methods for convex programming problems. *S.I.A.M. J. Optim.*, 7 (1997), pp. 347–366.

Chapter 12
Second-Order Symmetric Duality with Generalized Invexity

S. K. Padhan and C. Nahak

Abstract A pair of second-order symmetric dual programs such as Wolfe-type and Mond–Wier-type are considered and appropriate duality results are established. Second-order $\rho-(\eta,\theta)$-bonvexity and $\rho-(\eta,\theta)$-boncavity of the kernel function are studied. It is also observed that for a particular kernel function, both these pairs of programs reduce to general nonlinear problem introduced by Mangasarian. Many examples and counterexamples are illustrated to justify our work.

12.1 Introduction

The study of second-order duality is useful due to the computational advantage over first-order duality as it gives bounds for the value of the objective function when approximations are used (see [7], [8], [10]). Symmetric duality in nonlinear programming in which the dual of dual is primal was introduced by Dorn [5]. Subsequently Dantzig Eisenberg, and Cottle [4] and Mond [11] significantly developed the notion of symmetric duality. Motivated by the concept of second and higher duality in nonlinear programming problems introduced by Mangasarian [8], several researchers [1, 9, 11, 12] have been working in this field. Mond [10] established Mangasarian's duality relations assuming rather simple inequalities for the objective and constant function. Bector and Chandra [3] called the functions satisfying these inequalities bonvex/boncave. Mond [10] has further studied second-order symmetric dual programs.

In this chapter we study second-order symmetric duality for Wolfe and Mond–Weir-type problems under $\rho-(\eta,\theta)$-bonvexity and $\rho-(\eta,\theta)$-boncavity

S. K. Padhan
Indian Institute of Technology, Kharagpur, India, e-mail: sarojpadhan@gmail.com

C. Nahak
Indian Institute of Technology, Kharagpur, India, e-mail: cnahak@maths.iitkgp.ernet.in

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_12,

assumptions, respectively. Different duality results (weak, strong, converse) are established. Many examples and counterexamples are discussed to support the work.

12.2 Notation and Preliminaries

Let $\mathbb{R}^n$ denote the n-dimensional Euclidean space. Let $f(x,y)$ be a twice-differentiable function in $\mathbb{R}^n \times \mathbb{R}^m$ and $f_x(\bar{x},\bar{y})$ denote the gradient of f with respect to x at $(\bar{x},\bar{y})$; $f_y(\bar{x},\bar{y})$ is defined similarly. Also let $f_{xx}(\bar{x},\bar{y})$ and $f_{yy}(\bar{x},\bar{y})$ denote the $n \times n$ and $m \times m$ symmetric Hessian matrices at $(\bar{x},\bar{y})$, respectively. The symbol z^T stands for the transpose of a vector z.

Definition 12.1 ([6]). A twice-differentiable function f defined on a set $S \subseteq \mathbb{R}^n$ is said to be η-*convex* at $\bar{x} \in S$ if there exists $\eta(x,\bar{x})$ defined on $S \times S$ such that for all $p \in \mathbb{R}^n$,

$$f(x) - f(\bar{x}) \geq [\eta(x,\bar{x})]^T[\nabla f(\bar{x}) + \nabla^2 f(\bar{x})p] - \frac{1}{2}p^T \nabla^2 f(\bar{x})p, \qquad \forall x \in S.$$

Definition 12.2 ([6]). A twice-differentiable function f defined on a set $S \subseteq \mathbb{R}^n$ is said to be η-*pseudo-convex* at $\bar{x} \in S$ if there exists $\eta(x,\bar{x})$ defined on $S \times S$ such that for all $p \in \mathbb{R}^n$,

$$[\eta(x,\bar{x})]^T[\nabla f(\bar{x}) + \nabla^2 f(\bar{x})p] \geq 0 \Rightarrow f(x) \geq f(\bar{x}) - \frac{1}{2}p^T \nabla^2 f(\bar{x})p, \forall x \in S.$$

A twice differentiable function is η-concave and η-pseudo-concave if $-f$ is η-convex and η-pseudoconvex, respectively.

Definition 12.3. Let $f(x,y)$ be a twice-differentiable function on $\mathbb{R}^n \times \mathbb{R}^m$. f is said to be second-order $\rho - (\eta,\theta)$-convex at $u \in \mathbb{R}^n$, for fixed v, with respect to $\eta,\ \theta$ if there exist $\eta,\ \theta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, and $\rho \in \mathbb{R}$ such that

$$f(x,v) - f(u,v) \geq [\eta(x,u)]^T[f_u(u,v) + f_{uu}(u,v)r] - \frac{1}{2}r^T f_{uu}(u,v)r + \rho\|\theta(x,u)\|^2,$$
$$\forall x,\ r \in \mathbb{R}^n.$$

It follows that every η-convex function is $\rho - (\eta,\theta)$-bonvex but the converse is not true, which follows from the following counterexample (12.1).

Example 12.1. Let $f : [0,2\pi] \times [0,2\pi] \longrightarrow \mathbb{R}$ be defined by

$$f(x,y) = -2x^2 - 2x - 2y - \sin^2 x.$$

The above function is not η-convex but $\rho - (\eta,\theta)$-bonvex for $\eta(x,u) = -\frac{1}{2}\sin^2 u - u - x - 1, \qquad \theta(x,u) = \sqrt{x^2 + ux + x + u + 12}$, and $\rho = -100$.

Definition 12.4. Let $f(x,y)$ be a twice-differentiable function in $\mathbb{R}^n \times \mathbb{R}^m$. f is said to be second-order $\rho-(\eta,\theta)$-pseudo-bonvex at $u \in \mathbb{R}^n$, for fixed v, with respect to $\eta,\ \theta$ if there exist $\eta,\ \theta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $\rho \in \mathbb{R}$ such that

$$[\eta(x,u)]^T [f_u(u,v) + f_{uu}(u,v)r] \geq 0$$

$$\Rightarrow f(x,v) - f(u,v) + \frac{1}{2} r^T f_{uu}(u,v)r - \rho\|\theta(x,u)\|^2 \geq 0, \forall x,\ r \in \mathbb{R}^n.$$

It follows that every $\rho-(\eta,\theta)$-bonvex function is $\rho-(\eta,\theta)$-pseudo-bonvex but the converse is not true, which follows from counterexample (12.2).

Example 12.2. Let $f : [0,2\pi] \times [0,2\pi] \longrightarrow \mathbb{R}$ be defined by

$$f(x,y) = \sin^2 x + 235x + 3y.$$

The above function is not $\rho-(\eta,\theta)$-bonvex but $\rho-(\eta,\theta)$-pseudo-bonvex for $\eta(x,u) = x+u+1$, $\theta(x,u) = \sqrt{8\sin^2 u + 34u}$, and $\rho = -7$.

Definition 12.5. Let $f(x,y)$ be a twice differentiable function in $\mathbb{R}^n \times \mathbb{R}^m$. f is said to be second-order $\rho-(\eta,\theta)$-boncave at $y \in \mathbb{R}^m$, for fixed x, with respect to $\eta,\ \theta$ if there exist $\eta,\ \theta : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ and $\rho \in \mathbb{R}$ such that

$$f(x,v) - f(x,y) \leq [\eta(v,y)]^T [f_y(x,y) + f_{yy}(x,y)p] - \frac{1}{2} p^T f_{yy}(x,y)p + \rho\|\theta(v,y)\|^2,$$

$$\forall v,\ p \in \mathbb{R}^m.$$

It follows that every η-concave function is $\rho-(\eta,\theta)$-boncave but the converse is not true, which follows from counterexample (12.3).

Example 12.3. Let $f : [0,2\pi] \times [0,2\pi] \longrightarrow \mathbb{R}$ be defined by

$$f(x,y) = 5x + 90y + \sin^2 y.$$

The above function is not η-concave but $\rho-(\eta,\theta)$-boncave for $\eta(v,y) = -\sin^2 y - 2y - 1$, $\theta(v,y) = \sqrt{\sin^2 v + 60\sin^2 y + 30v + 65y + 75}$, and $\rho = 3$.

Definition 12.6. Let $f(x,y)$ be a twice-differentiable function in $\mathbb{R}^n \times \mathbb{R}^m$. f is said to be *second-order* $\rho-(\eta,\theta)$*-pseudo-boncave* at $y \in \mathbb{R}^m$, for fixed x, with respect to $\eta,\ \theta$ if there exist $\eta,\ \theta : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ and $\rho \in \mathbb{R}$ such that

$$-[\eta(v,y)]^T [f_y(x,y) + f_{yy}(x,y)p] \geq 0$$

$$\Rightarrow f(x,y) - f(x,v) - \frac{1}{2} p^T f_{yy}(x,y)p + \rho\|\theta(v,y)\|^2 \geq 0, \qquad \forall v,\ p \in \mathbb{R}^m.$$

It follows that every η-boncave function is $\rho-(\eta,\theta)$-pseudo-boncave but the converse is not true, which follows from counterexample (12.4).

Example 12.4. Let $f : [0,2\pi] \times [0,2\pi] \longrightarrow \mathbb{R}$ be defined by

$$f(x,y) = -2x - 90y - \sin^2 y.$$

The above function is not η-boncave but $\rho - (\eta,\theta)$-pseudo-boncave for $\eta(v,y) = \sin^2 y + 2y + 1$, $\theta(v,y) = \sqrt{\sin^2 v + 60\sin^2 y + 30v + 65y + 75}$, and $\rho = -3$.

12.3 Wolfe-Type Symmetric Duality

In this section, we consider the following pair of second-order Wolfe-type problems and establish the weak, strong, and converse duality results.

Primal (WP) Minimize $M(x,y,p) = f(x,y) - y^T f_y(x,y) - y^T f_{yy}(x,y)p - \frac{1}{2}p^T f_{yy}(x,y)p$

$$\text{subject to } f_y(x,y) + f_{yy}(x,y)p \leq 0, \tag{12.1}$$

$$x \geq 0. \tag{12.2}$$

Dual (WD) Maximize $N(u,v,r) = f(u,v) - u^T f_u(u,v) - u^T f_{uu}(u,v)r - \frac{1}{2}r^T f_{uu}(u,v)r$

$$\text{subject to } f_u(u,v) + f_{uu}(u,v)r \geq 0, \tag{12.3}$$

$$v \geq 0. \tag{12.4}$$

Theorem 12.1 (Weak Duality). *Let (x,y,p) and (u,v,r) be feasible solutions of WP and WD respectively. Let*

(i) $f(x,y)$ be $\rho_1 - (\eta_1,\theta_1)$-bonvex in x for fixed v.
(ii) $f(x,y)$ be $\rho_2 - (\eta_2,\theta_2)$-boncave in v for fixed x.
(iii) $\eta_1(x,u) + u \geq 0$, $\eta_2(v,y) + y \geq 0$ and $\rho_1\|\theta_1(x,u)\|^2 - \rho_2\|\theta_2(v,y)\|^2 \geq 0$.

Then $M(x,y,p) \geq N(u,v,r)$; that is, $\inf WP \geq \sup WD$.

Proof. From (i) and (ii) we have

$$f(x,v) - f(u,v) \geq \eta_1(x,u)^T[f_u(u,v) + f_{uu}(u,v)r] - \frac{1}{2}r^T f_{uu}(u,v)r + \rho_1\|\theta_1(x,u)\|^2 \tag{12.5}$$

$$f(x,v) - f(x,y) \leq \eta_2(v,y)^T[f_y(x,y) + f_{yy}(x,y)p] - \frac{1}{2}p^T f_{yy}(x,y)p + \rho_2\|\theta_2(v,y)\|^2 \tag{12.6}$$

Subtracting Equation (12.6) from Equation (12.5) we get

$$f(x,y) - f(u,v) \geq \eta_1(x,u)^T[f_u(u,v) + f_{uu}(u,v)r] - \frac{1}{2}r^T f_{uu}(u,v)r + \rho_1\|\theta_1(x,u)\|^2$$

$$- \eta_2(v,y)^T[f_y(x,y) + f_{yy}(x,y)p] + \frac{1}{2}p^T f_{yy}(x,y)p - \rho_2\|\theta_2(v,y)\|^2$$

$$\Rightarrow \left[f(x,y) - y^T f_y(x,y) - y^T f_{yy}(x,y)p - \frac{1}{2}p^T f_{yy}(x,y)p\right]$$

$$- \left[f(u,v) - u^T f_u(u,v) - u^T f_{uu}(u,v)r - \frac{1}{2}r^T f_{uu}(u,v)r\right]$$

$$\geq (\eta_1(x,u)^T + u)[f_u(u,v) + f_{uu}(u,v)r] - (\eta_2(v,y)^T + y)$$

$$[f_y(x,y) + f_{yy}(x,y)p] + \rho_1\|\theta_1(x,u)\|^2 - \rho_2\|\theta_2(v,y)\|^2$$

$$\geq 0 \quad \text{(by (iii) and Equations (12.1), (12.3))}$$

$$\Rightarrow M(x,y,p) \geq N(u,v,r); \ \text{ that is, } \inf WP \geq \sup WD.$$

□

Theorem 12.2 (Strong Duality). *Let $f(x,y)$ be a thrice-differentiable function and $(\bar{x},\bar{y},\bar{p})$ be a local optimal solution for WP. If*

(i) $f_{yy}(\bar{x},\bar{y})$ is nonsingular,
(ii) $(f_{yy}(\bar{x},\bar{y})\bar{p})_y\bar{p} = 0 \Rightarrow \bar{p} = 0$,

then

(I) $(\bar{x},\bar{y},\bar{r} = 0)$ is feasible for WD, and
(II) $M(\bar{x},\bar{y},\bar{p}) = N(\bar{x},\bar{y},\bar{r})$.

Furthermore, if the weak duality theorem (12.1) holds between the primal WP and the dual WD, then $(\bar{x},\bar{y},\bar{p} = 0)$ and $(\bar{x},\bar{y},\bar{r} = 0)$ are global optimal solutions for WP and WD, respectively.

Proof. Let

$$L_p = \alpha\left[f(x,y) - y^T f_y(x,y) - y^T f_{yy}(x,y)p - \frac{1}{2}p^T f_{yy}(x,y)p\right]$$

$$+ \beta^T[f_y(x,y) + f_{yy}(x,y)p] - \gamma^T x,$$

where $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}^m$, and $\gamma \in \mathbb{R}^n$. Because $(\bar{x},\bar{y},\bar{p})$ is a local optimal solution for (WP), by the Fritz–John optimality conditions [2], there exist $\bar{\alpha} \in \mathbb{R}$, $\bar{\beta} \in \mathbb{R}^m$, and $\bar{\gamma} \in \mathbb{R}^n$ such that

$$\frac{\partial L_p}{\partial x} = \bar{\alpha}[f_x(\bar{x},\bar{y})] + f_{yx}(\bar{x},\bar{y})[\bar{\beta} - \bar{\alpha}\bar{y}] + (f_{yy}(\bar{x},\bar{y})\bar{p})_x\left[\bar{\beta} - \bar{\alpha}\bar{y} - \frac{1}{2}\bar{\alpha}\bar{p}\right] - \bar{\gamma} = 0, \quad (12.7)$$

$$\frac{\partial L_p}{\partial y} = f_{yy}(\bar{x},\bar{y})[\bar{\beta} - \bar{\alpha}\bar{y} - \bar{\alpha}\bar{p}] + (f_{yy}(\bar{x},\bar{y})\bar{p})_y\left[\bar{\beta} - \bar{\alpha}\bar{y} - \frac{1}{2}\bar{\alpha}\bar{p}\right] = 0, \quad (12.8)$$

$$\frac{\partial L_p}{\partial p} = f_{yy}(\bar{x},\bar{y})[\bar{\beta} - \bar{\alpha}(\bar{y}+\bar{p})] = 0, \quad (12.9)$$

$$\bar{\beta}^T\frac{\partial L_p}{\partial \beta} = \bar{\beta}^T[f_y(\bar{x},\bar{y}) + f_{yy}(\bar{x},\bar{y})\bar{P}] = 0, \quad (12.10)$$

$$\bar{\gamma}^T\frac{\partial L_p}{\partial \gamma} = \bar{\gamma}^T\bar{x} = 0, \quad (12.11)$$

$$(\bar{\alpha},\bar{\beta},\bar{\gamma}) \geq \mathbf{0}, \qquad (\bar{\alpha},\bar{\beta},\bar{\gamma}) \neq \mathbf{0}. \quad (12.12)$$

Using hypothesis (i) in (12.9) we get

$$\bar{\beta} = \bar{\alpha}(\bar{y}+\bar{p}). \quad (12.13)$$

If $\bar{\alpha} = 0$, from (12.13) $\bar{\beta} = 0$ which contradicts (12.12). Hence

$$\bar{\alpha} > 0. \quad (12.14)$$

Substituting (12.13) and (12.14) in (12.8) we have

$$(f_{yy}(\bar{x},\bar{y})\bar{p})_y\bar{p} = 0. \quad (12.15)$$

So

$$\bar{p} = 0 \qquad \text{(by hypothesis (ii))}. \quad (12.16)$$

Using (12.12)–(12.14) and (12.16) in (12.7) we obtain

$$f_x(\bar{x},\bar{y}) \geq 0. \quad (12.17)$$

Again from (12.13), (12.14), and (12.16) we get

$$\bar{y} \geq 0. \quad (12.18)$$

Applying (12.13), (12.14), and (12.16) in (12.10) we have

$$\bar{y}^T f_y(\bar{x},\bar{y}) = 0. \quad (12.19)$$

Multiplying $\bar{x}$ on (12.7), and applying (12.11), (12.13), (12.14), and (12.16) we get

$$\bar{x}^T f_x(\bar{x},\bar{y}) = 0. \tag{12.20}$$

Therefore from (12.17)–(12.20) we see that $(\bar{x},\bar{y},\bar{r}=0)$ is feasible for WD and

$$M(\bar{x},\bar{y},\bar{p}=0) = N(\bar{x},\bar{y},\bar{r}=0)$$

Also, by Theorem (12.1), $(\bar{x},\bar{y},\bar{p}=0)$ and $(\bar{x},\bar{y},\bar{r}=0)$ are global optimal solutions for WP and WD, respectively. □

Theorem 12.3 (Converse Duality). *Let $f(u,v)$ be a thrice-differentiable function and $(\bar{u},\bar{v},\bar{r})$ be a local optimal solution for WD. If*

(i) $f_{uu}(\bar{u},\bar{v})$ *is nonsingular,*
(ii) $(f_{uu}(\bar{u},\bar{v})\bar{r})_u\bar{r} = 0 \Rightarrow \bar{r} = 0,$

then

(I) $(\bar{u},\bar{v},\bar{p}=0)$ *is feasible for WP, and*
(II) $M(\bar{u},\bar{v},\bar{p}) = N(\bar{u},\bar{v},\bar{r})$.

Moreover, if the weak duality theorem (12.1) holds between the primal WP and the dual WD, then $(\bar{u},\bar{v},\bar{p}=0)$ and $(\bar{u},\bar{v},\bar{r}=0)$ are global optimal solutions for WP and WD, respectively.

Proof. The proof is similar to that of Theorem (12.2). □

12.4 Mond–Wier-Type Symmetric Duality

In this section, we consider the following pair of second-order Mond–Wier-type problems and establish weak, strong, and converse duality theorems.

$$\text{Primal (MWP)} \quad \text{Minimize } F(x,y,p) = f(x,y) - \frac{1}{2}p^T f_{yy}(x,y)p$$

$$\text{subject to } f_y(x,y) + f_{yy}(x,y)r \leq 0, \tag{12.1}$$

$$y^T f_y(x,y) + y^T f_{yy}(x,y)p \geq 0, \tag{12.2}$$

$$x \geq 0. \tag{12.3}$$

$$\text{Dual (MWD)} \quad \text{Maximize } G(u,v,r) = f(u,v) - \frac{1}{2}r^T f_{uu}(u,v)r$$

$$\text{subject to } f_u(u,v) + f_{uu}(u,v)p \geq 0, \tag{12.4}$$

$$u^T f_u(u,v) + u^T f_{uu}(u,v)r \leq 0, \tag{12.5}$$

$$v \geq 0. \tag{12.6}$$

Theorem 12.4 (Weak Duality). *Let (x,y,p) and (u,v,r) be feasible solutions of MWP and MWD, respectively. Let*

(i) $f(x,y)$ be $\rho_1-(\eta_1,\theta_1)$-bonvex in x for fixed v,
(ii) $f(x,y)$ be $\rho_2-(\eta_2,\theta_2)$-boncave in v for fixed x, and
(iii) $\eta_1(x,u)+u\geq 0$, $\eta_2(v,y)+y\geq 0$ and $\rho_1\|\theta_1(x,u)\|^2-\rho_2\|\theta_2(v,y)\|^2\geq 0$.

Then $F(x,y,p)\geq G(u,v,r)$; that is, $\inf MWP \geq \sup MWD$.

Proof. From (12.3) and hypothesis (iii), we have

$$\begin{aligned}\eta_1^T(x,u)[f_u(u,v)+f_{uu}(u,v)p] &\geq -u^T f_u(u,v)-u^T f_{uu}(u,v)r\\ &\geq 0. \qquad \text{(by (12.3))}\end{aligned} \tag{12.7}$$

Now hypothesis (i) gives

$$f(x,v)-f(u,v)+\frac{1}{2}r^T f_{uu}(u,v)r-\rho_1\|\theta_1(x,u)\|^2\geq 0. \tag{12.8}$$

Again from (12.1) and hypothesis (iii), we get

$$\begin{aligned}-\eta_2^T(v,y)[f_y(x,y)+f_{yy}(x,y)r] &\geq y^T f_y(x,y)+y^T f_{yy}(x,y)p\\ &\geq 0. \qquad \text{(by (12.2))}\end{aligned} \tag{12.9}$$

Now hypothesis (ii) gives

$$f(x,y)-f(x,v)-\frac{1}{2}p^T f_{yy}(x,y)p+\rho_2\|\theta_2(v,y)\|^2\geq 0. \tag{12.10}$$

From (12.8), (12.10), and hypothesis (iii), we obtain

$$F(x,y,p)\geq G(u,v,r); \quad \text{that is} \;\; \inf MWP\geq \sup MWD \qquad \square$$

Theorem 12.5 (Strong Duality). *Let $f(x,y)$ be a thrice-differentiable function and $(\bar{x},\bar{y},\bar{p})$ be a local optimal solution for MWP. If*

(i) $f_{yy}(\bar{x},\bar{y})$ is nonsingular,
(ii) $(f_{yy}(\bar{x},\bar{y})\bar{p})_y$ is positive or negative definite, and
(iii) $f_y(\bar{x},\bar{y})+f_{yy}(\bar{x},\bar{y})\bar{p}\neq 0$,

then

(I) $\bar{p}=0$, $(\bar{x},\bar{y},\bar{r}=0)$ is feasible for MWD, and
(II) $F(\bar{x},\bar{y},\bar{p})=G(\bar{x},\bar{y},\bar{r})$.

Furthermore, if the weak duality theorem (12.4) holds between the primal MWP and the dual MWD, then $(\bar{x},\bar{y},\bar{p})$ and $(\bar{x},\bar{y},\bar{r})$ are global optimal solutions for MWP and MWD, respectively.

Proof. The proof follows from Theorem 5 of [6].

Theorem 12.6 (Converse Duality). *Let $f(u,v)$ be a thrice-differentiable function and $(\bar{u},\bar{v},\bar{r})$ be a local optimal solution for MWD. If*

(i) $f_{uu}(\bar{u},\bar{v})$ *is nonsingular,*
(ii) $(f_{uu}(\bar{u},\bar{v})\bar{r})_u$ *is positive or negative definite, and*
(iii) $(f_u(\bar{u},\bar{v})+f_{uu}(\bar{u},\bar{v})\bar{r} \neq 0,$

then

(I) $\bar{r}=0,\ (\bar{u},\bar{v},\bar{p}=0)$ *is feasible for MWP, and*
(II) $F(\bar{u},\bar{v},\bar{p}) = G(\bar{u},\bar{v},\bar{r})$.

Moreover, if the weak duality theorem (12.4) holds between the primal MWP and the dual MWD, then $(\bar{u},\bar{v},\bar{p})$ and $(\bar{u},\bar{v},\bar{r})$ are global optimal solutions for MWP and MWD, respectively.

Proof. The proof follows from Theorem 6 of [6].

12.5 Conclusion

(a) First-order symmetric duality

If $p=r=0$, then the above Wolfe-type programs reduce to the first-order symmetric dual programs of Dantzig, Eisenberg, and Cottle [4]. Similarly, Mond–Wier-type programs reduce to the first-order symmetric dual programs of Mond and Weir [12].

(b) Second order symmetric duality

If we take $f(x,y) = f(x)+y^T g(x)$, then MWP reduces to the following nonlinear program.

$$\text{Minimize } f(x)+y^T g(x)$$

$$\text{subject to } g(x) \leq 0, \tag{12.1}$$

$$x \geq 0. \tag{12.2}$$

Also, the dual MWD becomes

$$\text{Maximize } f(u)+v^T g(u) - \frac{1}{2} r^T [f_{uu}(u)+(v^T g(u))_{uu}]r$$

$$\text{subject to } f_u(u)+(v^T g(u))_u + f_{uu}(u)r + (v^T g(u))_{uu} r \geq 0, \tag{12.3}$$

$$u^T f_u(u)+u^T(v^T g(u))_u + u^T f_{uu}(u)r + u^T(v^T g(u))_{uu} r \geq 0, \tag{12.4}$$

$$v \geq 0. \tag{12.5}$$

Thus we obtain a pair of nonlinear program problems studied by Mangasarian [8] and Mond and Weir [12], respectively, in the second- and higher-order dual sense.

References

1. Ahmad I, Husain Z (2004) Non-differentiable second-order symmetric duality in multi-objective programming problems. *J. Math. Anal. Appl.*, 18:721–728.
2. Bazaraa MS, Goode JJ (1973) On symmetric duality in nonlinear programming. *Oper. Res.*, 21:1–9
3. Bector CR, Chandra S (1987) Generalized invexity and higher order duality for fractional programming. *Opsearch*, 24:143–154
4. Dantzig GB, Eisenberg E, Cottle RW (1965) Symmetric dual nonlinear programs. *Pacific J. Math.*, 15:809–812
5. Dorn WS (1960) A symmetric dual theorem for quadratic programs. *J. Oper. Res. Soc. Japan*, 2:93–97
6. Gulati TR, Ahmad I, Husain I (2001) Second-order symmetric duality with generalized convexity. *Opsearch*, 38:210–222
7. Jeyakumar V (1986) p-convexity and second order duality. *Utilitas Math.*, 29:71–85
8. Mangasarian OL, Second and higher order duality in nonlinear programming. *J. Math. Anal. Appl.*, 51:607–620
9. Mishra SK (2005) Non-differentiable higher-order symmetric duality in mathematical programming with generalized invexity. *European J. Oper. Res.*, 167:28–34
10. Mond B (1974) Second order duality for nonlinear programs. *Opsearch*, 11:90–99
11. Mond B (1965) A symmetric dual theorem for nonlinear programs. *Quart. Appl. Math.*, 23:265–269
12. Mond B, Weir T (1991) Symmetric duality for nonlinear multiobjective programming in: S. Kumar (Ed.). *Recent Devlopment in Mathematical Programming*, Gordon and Breach Science, London, 137–153

Chapter 13
A Dynamic Solution Concept to Cooperative Games with Fuzzy Coalitions

Surajit Borkotokey

Abstract The problem of distribution of payoffs through negotiation among the players in a cooperative game with fuzzy coalitions is considered. It is argued that this distribution is influenced by satisfaction of the players in regard to better performance and success within a cooperative endeavour. As a possible alternative to static solutions where this point is ignored, a framework concerning the players' satisfactions upon receiving an allocation of the worth is studied. A solution of the negotiation process is defined and the corresponding convergence theorem is established.

13.1 Introduction

A cooperative game with side payments can be completely characterized by a real-valued set function v, called the characteristic function defined over the set of all possible coalitions. This characteristic function assigns to each coalition a non-negative real number called its worth. In the literature, worth of a coalition is interpreted either as the minimum payoff, the members of the coalition assure themselves or as the maximum payoff they expect to achieve by forming it. Thus, depending upon its interpretation, worth can be treated as either an upper or a lower bound of the actual amount (payoff) achieved after the players form coalitions. A solution is a rational distribution of the payoffs to the individual players.[1]

In crisp games, a subset of the players' set can be viewed as a coalition with full participation and a coalition structure would represent a partition of the players' set. However, this idea is not very interesting while dealing with practical situations. There are numerous instances [1, 5, 4, 22], where players would prefer to participate partially in the coalitions. We call them fuzzy coalitions. Moreover, it is possible to

Surajit Borkotokey
Department of Mathematics, Dibrugarh University, Assam, India, e-mail: surajitbor@yahoo.com

[1] A solution is indeed a vector in $\mathbb{R}^n$, n being the number of players involved in the game.

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_13,

have a player who wants to participate simultaneously in different coalitions with varied participation rates. In order to model such situations, Aubin [1], Butnariu [5], Branzei, Dimitrov, and Tijs [4] and Tsurumi, Tanino, and Inuiguchi [22] among others extended the notion of crisp games into fuzzy games and obtained interesting results similar to their crisp counterparts.

Two approaches of solution concepts in cooperative games are available in the literature, namely the static approach and dynamic approach. Solutions such as core, Shapley value, compromise value, and the like are static in nature. They do not address the process of coalition formation and the bargaining about how to split the surplus rationally among the players. In crisp games this problem has been well studied by many researchers such as Dickemann [7], Ray and Vohra [18–20], Tohme and Sandholm [21], and so on. Most of these works are inclined to obtain a final coalition structure along with a payoff vector after executing a chain of bargaining and negotiation protocols. Agent (Player) negotiation is an iterative process through which a joint decision is made by two or more agents (players) to reach a mutually acceptable agreement. It is worth mentioning the observations made by Carmichael [6] who wrote that, in crisp cooperative game theory, agreements are binding by definition. Wage bargaining between an employer and a labour union is an example of cooperative bargaining because the outcome of the game is a legally binding contract if they can agree. It is a threat outcome if they do not, where a player's threat outcome or threat utility is the best alternative outcome in the event of no agreement. In cooperative games, players have an incentive to make agreements that are worth making and that they won't regret. As stated by Friedman [8] it is therefore, "natural to focus attention on what players ought, in some sense, to agree on." The restrictions on the bargaining outcome follow from such commonsense observations.

1. Individual rationality: Players won't agree to anything less than they could get by not reaching an agreement.
2. Group rationality: Players should agree on something on which they cannot jointly improve.
3. Anonymity or symmetry: The solution should not depend on the labelling of the players.
4. Transformation invariance or invariance to equivalent utility representations: The solution shouldn't change if either player's utility function is altered in a linear way.
5. Independence of irrelevant alternatives: If the number or range of possible outcomes is restricted but this doesn't affect the threat point and the previous solution is still available, the outcome should not change.

Individual rationality indicates that the players won't agree to any outcome that gives them a lower payoff than their payoff if there is no agreement. Group rationality implies that the negotiated outcome should be Pareto efficient. In geometric terms these restrictions mean that the outcome of bargaining must lie on the contract curve or payoff possibility frontier (sometimes known as the utility increments frontier). The third restriction implies that when the players' utility functions and their threat utilities are the same they receive equal shares. That is, any asymmetries in the final payoff should only be attributable to differences in their utility functions or their threat outcomes. The outcome should be independent of interpersonal comparisons of utilities. Restriction 4 means that the solution is independent of the units in which utility is measured. For example, if the bargain is over money and one player's utility for money doubles, this shouldn't change the monetary outcome but whatever the player gets he will simply value it twice as much. With these restrictions imposed Nash [17] showed that there is a unique solution to the bargaining problem known as the Nash bargaining solution. The Nash bargaining solution is the outcome which maximises the product of the players' gains from any agreement. This product is known as the Nash product (refer to [6]).

The situation, however, complicates when we consider cooperative games with fuzzy coalitions. Lai and Lin [10] have studied agent negotiation in e-business by deploying fuzzy constraints. Luo et al. [14] developed a fuzzy constraint-based model for agent negotiation in a trading environment. However, all these approaches are noncooperative in nature. Agent negotiation in fuzzy coalitional games in terms of optimum benefit to all the players is indeed a challenging task. Although many dynamic learning models have been developed for crisp cases in the recent past, little research has been done in a fuzzy setting. The protocols developed so far for crisp games cannot be directly extended to their fuzzy counterparts because

1. The crisp system is finite in the sense that the set of possible alternatives for bargaining is finite, although this assumption in a fuzzy environment would make it oversimplified.
2. In a fuzzy setting, a partition of the grand coalition is meaningless as the players can form multiple coalitions simultaneously.

So, in a dynamic crisp system, reaching a certain allocation requires the two a priori unrelated processes on the part of the players: coalition formation and bargaining on the distribution of payoffs simultaneously; on the other hand, in a fuzzy system, we can consider the second process only: bargaining on the distribution of payoffs after all the players offer their memberships in a particular coalition. However, by awarding some binding incentives, the players can be encouraged to make further coalitions.

In the case of static solution concepts such as that of the Shapley values defined by Tsurumi, Tanino, and Inuiguchi [22] and Li and Zhang [12] for cooperative games with fuzzy coalitions, the resulting allocations to the players within a coalition are a rational distribution of the worth of that coalition. Such a solution is entirely dependent on the rate of players' participation. However, this idea does not encourage the players to form further coalitions and hence does not influence the process of forming coalitions. In order to ensure that the players are motivated to form fuzzy coalitions, they need some binding incentives. There are instances where satisfaction of individuals of high social position and acceptance is emphasised in enhancing the worth of the corresponding fuzzy coalition. Organizations influencing public opinion such as news channels, magazines, and the like usually employ individuals having already acquired societal acceptance. It is easily comprehended that these individuals do not usually contribute much to the production process, however, their presence is enough for a binding coalition and consequently a better return. Thus every possible measure is taken to keep those people in the organization. In order to do so, we need to revisit the process of distribution of payoffs in a dynamic environment. Moreover, performance of the current and the next coalition would be enhanced if each player were satisfied with the payoffs he or she gets (or expects to get) from the current coalition. In the case of a task allocation in a work group, for example, if we identify worth with the gain achieved as a result of the success of the project then it would be a function of the following well-known parameters.

1. Goal achievement
2. Timeliness of the players
3. Quality of the performed task
4. Team collaboration level
5. Individual contribution level

The interested reader may see [9] for an in-depth discussion. Among the above five parameters, the first four parameters are criteria for better performance of a player and the fifth parameter emphasizes the participation rate. Performance of an individual is influenced by her physiological and psychological states and traits along with her ability, training, and education [9]. Lim and Zain Mohamed [13] have observed that doubts often arise about what and who actually determine project success. They have explored the issues from different perspectives of people looking at the project. They proposed to classify project success into two categories: the macro and micro viewpoints and suggested that two criteria are sufficient to determine the macro viewpoint of project success: completion and satisfaction. Whereas the completion criterion alone is enough to determine the micro viewpoint of project success, a successful endeavour would result in enhanced worth so there is a need to explore the factors affecting project success. These factors will be equally important for a more general setup addressing other games as well. Thus the problem of finding a solution to such a game is not restricted to distribution of the payoff according to the participation rate but also encouraging the players to perform well by focussing on their satisfaction levels as well. A third observation in this regard is about the cooperation among the players. The static solutions do not in general reflect the cooperation among the players explicitly. However, in modelling human behaviour, likings and dislikings among the players play an important role. For a meaningful coalition, the members need to be supportive of one another. Therefore, it is natural to expect all the players to be satisfied at par with a possible solution vector in a cooperative environment. Thus an aggregated satisfaction value over a particular payoff to an individual player within the coalition can be derived to address all the aforementioned aspects. A solution in this paradigm should be such that every player is almost equally satisfied.

In this chapter, we have considered the problem of distribution of payoffs among the players in an n-person cooperative game with fuzzy or partial participation.

The objective of our study is to provide a systematic treatment of satisfaction level as a basis for negotiation among rational agents, who can participate in various coalitions with a varied rate of memberships simultaneously. We assume that the worth of a fuzzy coalition is evolved dynamically, as opposed to its static behaviour considered in the literature so far.

Initially the players would announce their memberships in a fuzzy coalition and disclose their demands as well. If the total payoff demanded is achievable from the worth, no negotiation will be required. If, however, the demands are not met, they have to revise their requirements and a mediator propose an offer initially based on their rates of participation. Upon receiving the proposal offered by the mediator, each of the participating players provides the membership value of his satitsisfaction. On the basis of this information the mediator will update his belief and propose the next offer and the process will continue until a stopping condition is met. Thus the mediator would offer alternative proposals to the players judging their reactions to the previous offers. We have developed a stopping rule and proposed the process of updating the belief of the mediator by use of a suitable probability measure towards the possible reactions of the players upon different offers. Furthermore, a similarity relation is defined to measure the similarity between the satisfaction levels of the individual players over a single proposal. The negotiation strategy is so designed that the mediator would propose only offers (possible solutions) for which the similarity value would be maximum at each stage of the negotiation process. What we have also kept in mind is that, in the negotiation process, each of the players has a single motive: maximizing the individual payoff, which is well represented by the monotonic increasing functions characterising the fuzzy sets of their satisfactions. However, negotiation asks a player to accommodate the desires and views of all the other players. This suggests that an appropriate negotiation process should restrain the players from claiming irrational demands and it should reward those who are more open in forming coalitions. Our model shows that the negotiation process thus defined speeds up for cooperating players. We provide an example to show the usefulness of our proposed model.

13.2 Preliminaries

In this section we give the needed definitions and results from [1–23], those used in this chapter. A fuzzy set is characterized by a membership function from the universal set to $[0,1]$. Thus, without loss of generality, we denote the fuzzy sets here by their membership functions. We consider the class of fuzzy games defined by Azrieli and Lehrer [2]. This class seems to be more general than the other existing classes and includes the class of crisp games as a subclass. Its interpretation, however, is rather different. A fuzzy subset of a crisp set X is a function from X to $[0,1]$, assigning every element of X a membership between 0 and 1. Let N be a finite set representing the types of agents in a large population. There is a continuum of agents of each type and $Q_i \geq 0$ is the size of type i $(i = 1,2,\ldots,n)$ agents. The entire population is, therefore, represented by a nonnegative vector $Q = (Q_1, \dots, Q_n)$, and possible coalitions are identifed with the vectors that are (coordinatewise) smaller than Q. Thus formalizing the notion, we have the following.

For every nonnegative vector $Q \in \mathbb{R}^n$, let $F(Q)$ be the box given by

$$F(Q) = \{c \in \mathbb{R}^n : 0 \leq c \leq Q\}. \tag{13.1}$$

The point Q is interpreted as the 'grand coalition' in the fuzzy sense, and every $c \in F(Q)$ is a possible fuzzy coalition. For every $Q \geq 0 : Q \in \mathbb{R}^n$, a fuzzy game is a pair (Q, v) such that

(i) $Q \in \mathbb{R}^n$ and $Q \geq 0$.
(ii) $v : F(Q) \to \mathbb{R}^+ \cup 0$ is bounded and satisfies $v(0) = 0$

where $0 \in F(Q)$ is the zero vector signifying 0-size of all types of players. Thus if c_i represents the amount of agents of type i $(i = 1, 2, \dots, n)$ that participate in a coalition, then the total worth of $c = (c_1, c_2, \dots, c_n)$ is given by the real number $v(c)$ [see [2] for more details].

This model has another interpretation due to Azrieli and Lehrer [2]. Assume that for every i $(i = 1, 2, 3, \dots, n)$, the amount of resources available for agent i is $Q_i \geq 0$ (this can be time, money, etc.). Each agent can choose to invest any fraction of her resources $c_i \leq Q_i$ in a joint project. Note that a fuzzy coalition in Aubin's [1] sense is given by a membership function from N to $[0,\ 1]$, however, the two approaches are equivalent in the following sense.

If for every $c \in F(Q)$, c_i $(0 \leq c_i \leq Q_i)$ is the amount of resources that agent i invests, then we can uniquely define a function $S_c^Q : N \to [0, 1]$ as follows.

$$S_c^Q(i) = \begin{cases} \frac{c_i}{Q_i} & \text{if } Q_i \neq 0 \text{ and } c_i \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{13.2}$$

The function S_c^Q can be interpreted as the membership function for a possible fuzzy coalition in Aubin's sense pertaining to c in $F(Q)$. Thus under this interpertation, every $c \in F(Q)$ corresponds to a unique fuzzy coalition S_c^Q in membership function form and vice versa. The support of c denoted by $Supp(\ c)$ is the set $\{i \in N \mid c_i > 0\}$. The following definitions are important.

Definition 13.1 ([1]). A core solution to the game (Q, v) is a vector $x \in \mathbb{R}^n$ such that:

(i) $x \cdot Q \ = \ v(Q)$ and
(ii) $x \cdot c \ = \ v(c) \ \forall c \in F(Q)^2$.

Definition 13.2. The minimum deal index of a fuzzy game (Q, v) with respect to a fuzzy coalition c is the vector $x(i, c) \in \mathbb{R}^n$ such that

$$x(i,c) = \begin{cases} v(0|c_i) + \frac{c_i}{\sum c_i}[v(c) - \sum_i v(0|c_i)] & \text{if } i \in \text{Supp } c \\ 0 & \text{otherwise,} \end{cases} \tag{13.3}$$

[2] For two vectors x and y in $\mathbb{R}^n$, $x \cdot y$ represents the inner product $\sum_{i=1}^n x_i \cdot y_i$.

where $v(0|c_i) = v(0,\dots,0,c_i,0,\dots,0)$, and $v(c).c_i$ may be interpreted as the proportion of resources of the ith component in $v(c)$.

Remark 13.1. The minimum deal index is a core solution.

A vector of payoffs $x = (x_1, x_2, \dots, x_n)$, one for each player, is called an allocation. An allocation is *feasible for coalition c* if $\sum_{i=1}^{n} x_i^c \leq v(c)$.

The following theorem establishes an independence of the choice of participation by the players in various fuzzy coalitions.

Theorem 13.1. *If $Q_i (> 0)$ represents the total resource of player i in the game (Q, v), and it is exhausted in forming a finite number of coalitions c^j, $j = 1,2,\dots,$ $m < \infty$, (i.e., $\sum_{j=1}^{n} c_i^j = Q_i \ \forall i \in N$), then the membership distributions of individual players are linearly independent.*

Proof. Let for $\alpha_i \in \mathbb{R}$, $i \in N$,

$$\sum_{i=1}^{n} \alpha_i (c_i^1, c_i^2, \dots, c_i^m) = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i \cdot c_i^j = 0 \ \ : j = 1,2,\dots,m \Rightarrow \sum_{j=1}^{m} \sum_{i=1}^{n} \alpha_i \cdot c_i^j = 0$$

$$\Rightarrow \sum_{i=1}^{n} \alpha_i \cdot Q_i = 0 \Rightarrow \alpha_i = 0, \ \text{ as } Q_i s \text{ are all positive.}$$

■

13.3 Our Model

We now turn to our model. For our game, let us assume that $v(c)$ represents the payoff that the players achieve after forming the coalition c. The negotiation process among the players for a suitable allocation of the worth incorporates the satisfaction values of each player upon the previous offers. Consequently a solution to such a game is an allocation $x^c \in \mathbb{R}^n$, *feasible for c* which exhausts $v(c)$. We have adopted solution axioms for our game similar to those given by Nash [17] for crisp games and which we mentioned in the introduction. Formally we have the following.

Definition 13.3 (Nash bargaining solution in the fuzzy sense). A *Nash bargaining solution in the fuzzy sense* is a feasible allocation $x^c \in \mathbb{R}^n$ which exhausts $v(c)$; that is, for each c we must have $\sum_{i=1}^{n} x_i^c = v(c)$ and satisfy the following axioms.

1. Individual rationality: Players won't agree to anything less than they could get by not reaching an agreement.
2. Group rationality: Players should agree on something they cannot jointly improve on. Here improvement is essentially measured in terms of an aggregated satisfaction value.
3. Transformation invariance: The solution should not change if any player's participation is altered in a linear way.

4. Independent of irrelevant alternatives: If the number of range of possible outcomes is restricted but the previous solution is still available, the outcome should not change.

All the players would offer their participation in various fuzzy coalitions among them. The negotiation takes place for one coalition at a time. A coalition structure $\{c^j\}_{j=1}^m$ in a fuzzy sense is a set of all m fuzzy coalitions formed by the players by offering their resources. Thus the players may form the fuzzy coalitions in the coalition structure all at one time or one after another. In either case, at time $t = 0$, all the players announce their respective initial demands (i.e., the share of the payoff) they initially aspire to get with respect to a fuzzy coalition c say. Realized payoffs depend on the compatibility of the demands of the players.

13.3.1 The Allocation Process

The allocation at time $t = 1$ is made as follows. If the demands within the coalition c are feasible, each member of c will receive his or her demand. Otherwise they will be asked to construct individually the fuzzy sets of their satisfaction over the payoffs. Consequently the mediator will start offering proposals at each t until a stopping condition is attained. In general, upon receipt of an offer at time t, the players would react by announcing their level of satisfaction. The mediator, unaware of the actual fuzzy sets of satisfaction of the players instead updates her beliefs (fuzzy sets of the satisfaction of the players as she believes them to be) from the preceding information of the players. She then proposes the next offer so that the satisfaction degrees of the players as she believes are closer than the previous ones. When all the players are almost equally satisfied up to a desired exception, that is, when the similarity among all the players' satisfaction is maximum, the negotiation stops and the corresponding proposal would be the required solution with respect to the coalition.

Definition 13.4 (Similarity function). The *similarity function* among n fuzzy sets μ_i on the negotiated issues, denoted by $\mathtt{Sim}(\mu_1, \mu_2, \ldots, \mu_n) : \mathbb{R}^n \to [0,1]$ is defined as

$$\mathtt{Sim}(\mu_1, \mu_2, \ldots, \mu_n)(z) = 1 - \frac{1}{n}\sqrt{\sum_{\substack{i,j=1 \\ i>j}}^{n} (\mu_i(z) - \mu_j(z))^2}. \qquad (13.4)$$

Let $S\_P_i(\cdot)$ denote the membership function representing the fuzzy set of satisfaction the ith player is to define according to his aspirations. However, the players need not construct well-formulated fuzzy sets of satisfaction. All they require to announce, rationally, is the degree $S\_P_i(z)$ of their satisfaction over a particular offer z at each time period. Let $D^0 = \{z \mid \sum_{i=1}^n z_i = v(c); z_i \geq 0\}$ be the set of all feasible allocations and $D_i^0 = \{(z \| z_i) \mid \sum_{\substack{j=1 \\ j\neq i}}^n z_i = v(c); z_i \geq 0\}$ where $(z \| z_i) = (z_1, z_2, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n)$. The ith component $P_i^{t+1}(z \mid x^t)$ of the conditional probality vector $P^{t+1}(z \mid x^t)$ of choosing an offer z at time $t+1$ by the mediator from D^0 for player i given that an offer x^t has been offered in time t, is defined as

$$P_i^{t+1}(z \mid x^t) = \left(\frac{\sum\limits_{\substack{j=1 \\ j \neq i}}^{n} c_j}{\sum\limits_{j=1}^{n} c_j} \right) \frac{\underbrace{\int\int \cdots \int}_{D_i^0} (S\_P_i(x^t) + \mu_i^t(z)) \prod\limits_{\substack{j=1 \\ j \neq i}}^{n} dz_j}{\underbrace{\int\int \cdots \int}_{D^0} (S\_P_i(x^t) + \mu_i^t(z)) \prod\limits_{j=1}^{n} dz_j}, \tag{13.5}$$

where μ_i^t is the belief function of the mediator towards player i updated at time t and defined as follows.

Definition 13.5 (Belief function). Given the previous belief (i.e., at time $(t-1)$), $\mu_i^{t-1}(\cdot)\ \forall i \in N$ and the previous offers $x^0, x^1, \ldots, x^{t-1}$, the *belief function* $\mu_i^t(\cdot)$ at time t, is defined as

$$\mu_i^t(z) = \begin{cases} P_i^t(z \mid x^t) \cdot S\_P_i(x^{t-1}) \vee \mu_i^{t-1}(z)\ \forall z \in D^0 & \text{if } z \neq x^k,\ k = 0,1,2,\ldots,(t-1) \\ S\_P_i(x^k) & \text{if } z = x^k,\ k = 0,1,2,\ldots,(t-1), \end{cases} \tag{13.6}$$

with $\mu_i^0(\cdot)$ as the initial belief suitably chosen by the mediator.

Definition 13.6 (Set of feasible proposals). Given the beliefs $\mu_i^{t+1}(\cdot)$ updated for each player by the mediator, at time $t+1$, the latest offer $x^t = x^t$ at time t and the membership values $S\_P_i(x^t)$, the set of feasible proposals for time $t+1$ is constructed as

$$D^{t+1} = \left\{ z \mid \sum_{i=1}^{n} z_i = v(c); z_i \geq v(c_i|0);\ \mathtt{Sim}(\mu_1^{t+1}, \ldots, \mu_n^{t+1})(z) \right.$$
$$\left. \leq \mathtt{Sim}(S\_P_1, \ldots, S\_P_n)(x^t) \right\}. \tag{13.7}$$

Definition 13.7 (Expected allocation proposal). Given D^{t+1}, the set of feasible proposals for time $t+1$ an *expected allocation proposal* x^{t+1} is defined as

$$x^{t+1} = \arg_z \left(\max_{z \in D^{t+1}} \mathtt{Sim}(\mu_i^{t+1}, \ldots, \mu_i^{t+1})(z) \right). \tag{13.8}$$

Definition 13.8 (Solution of the negotiation process). Given a sequence $x^t \in D^t$ of feasible allocations, a solution of the negotiation process is a feasible allocation $x^c \in \mathbb{R}^n$ which satisfies the following,

(a) $\sum_{i=1}^{n} x_i^c = v(c)$

(b) $\texttt{Sim}(S\_P_1,\dots,S\_P_n)(x^c) \geq \texttt{Sim}(S\_P_1,\dots,S\_P_n)(x^t) \qquad \forall x^t \in D^t.$ (13.9)

Definition 13.9 (Bargaining cost). Amount payable by the *i*th player to the mediator because of her efforts towards achieving an agreement among the players is termed the *bargaining cost*. Formally, we define the *bargaining cost function* for our model as $r(i,t) = (x_i^t(1 - S\_P_i(x^t)))/(c_i v(c)) \cdot \phi(t)$, where $\phi(t)$ is the cost due to the delay in time.

Assumption 13.10. In a negotiation process, each player has to pay the mediator a bargaining cost which increases with delay in time and irrationality in satisfaction of the *i*th player. Furthermore, if the demands made at the begining of the game are feasible, there will be no bargaining cost.

Note that the purpose of taking Assumption 13.10 is to ensure that the players announce their demands rationally and sensibly. There may be a case with a very low satisfaction on a very high payoff with a low level of participation in the coalition (we call it irrational) by a particular player whereas the others are genuinely dissatisfied with their own payoffs as well. Yet, we will get a solution as the similarity remains at a high among them. Such a solution would never serve our purpose of inducing satisfaction of the players. So, the bargaining cost function is so designed that the players are debarred from adopting such irrational expectations.

13.3.2 Protocol

Step 1. Each player i will announce his or her demand for participating in the coalition c.

1. If all these demands together form a feasible allocation for c, GOTO step 5.
2. Else Player i will revise his aspiration and design a fuzzy set of satisfaction represented by a membership function, denoted as $S\_P_i(\cdot)$.
3. The mediator offers the initial proposal.
4. The players will announce rates of their satisfaction.

Step 2. Every player's degree of satisfaction at the tth stage is used by the mediator to update her belief for the $(t+1)$th stage. She will then choose an expected allocation proposal from a possible set of alternatives.

Step 3. Player i will announce his degree of satisfaction upon receiving the offer made by the mediator.

Step 4. A stopping rule is tested:
'Whether $\texttt{Sim}(S\_P_1,\dots,S\_P_n)(x^{t+1}) \leq \texttt{Sim}(S\_P_1,\dots,S\_P_n)(x^t)$?'

1. If the condition is met take $t = t+1$ and GOTO step 2.
2. If the condition is not met, continue.

Step 5. The proposal x^t offered in time t is a solution of the allocation process and the process terminates.

Remark 13.2. Note that here we have not mentioned what the initial offer should be. This offer can be anything, however, in order to speed up the negotiation process, the mediator may start with a core solution, or the minimum deal index. Similarly, the fuzzy sets representing initial beliefs (one for each of the players) can be justifiably considered as monotonic increasing functions having nonzero supports on the initial offer.

13.3.3 Main Theorem

Before stating and proving the main theorem, we make the following observation.

Observation 13.11. D^t, the set of feasible proposals at time t defined by Equation [13.7] is closed, bounded, and monotonic decreasing with respect to time t.

Theorem 13.2. *There exists a solution of the negotiation process.*

Proof. We consider two cases.

Case I. $D^t = \emptyset$ for some $\infty > t > 0$.

Here an expected allocation proposal in stage $t-1$ would be trivially a solution of the negotiation process (given by the stopping rule).

Case II. $D^t \neq \emptyset, \ \forall t$

In this case, D^t gets reduced at each stage so that the similarity among the satisfaction levels of the players as believed by the mediator gets closer to 1. Thus the negotiation process will converge on an allocation, say $x^* \in \cap_{t=1}^{\infty} D^t$. Formally, $x_t \to x^*$ as $t \to \infty$ and

$$\sum_{\substack{i,j=1 \\ i\neq j}}^{n} \{\mu_i^t(x^*) - \mu_j^t(x^*)\}^2 \leq \sum_{\substack{i,j=1 \\ i\neq j}}^{n} \{\mu_i^t(x^t) - \mu_j^t(x^t)\}^2. \tag{13.10}$$

Continuity of $S\_P_i(\cdot)$ implies that $S\_P_i(x^t) \to S\_P_i(x^*)$ and $P_i^t(x^* \mid x^t) \to 1$ such that $\mu_i^t(x^*) \to S\_P_i(x^*)$ as $t \to \infty$ for all i.

(13.10)$\Rightarrow$

$$\sum_{\substack{i,j=1 \\ i\neq j}}^{n} \{\mu_i^t(x^*) - \mu_j^t(x^*)\}^2 \leq \sum_{\substack{i,j=1 \\ i\neq j}}^{n} \{\mu_i^t(x^t) - \mu_j^t(x^t)\}^2$$

$$\leq \sum_{\substack{i,j=1 \\ i\neq j}}^{n} \{S\_P_i(x^{t-1}) - S\_P_j(x^{t-1})\}^2 \ \ \forall t.$$

Taking limits on the left, in particular, we get

$$\sum_{\substack{i,j=1 \\ i\neq j}}^{n} \{S\_P_i(x^*) - S\_P_j(x^*)\}^2 \leq \sum_{\substack{i,j=1 \\ i\neq j}}^{n} \{S\_P_i(x^t) - S\_P_j(x^t)\}^2 \;\; \forall x^t \in D^t$$

Thus x^* is the required solution of the negotiation process. ■

Theorem 13.3. *A solution of the negotiation process is a Nash bargaining solution in the fuzzy sense.*

Definition 13.12. A coalition c is said to satisfy the maximum cooperation criterion with respect to an allocation x^* if

$$\texttt{Sim}(S\_P_1,\ldots,S\_P_n)(x^*) \leq \texttt{Sim}(S\_P_1,\ldots,S\_P_n)(x) \qquad \forall x \in D^t \text{ and } \forall t.$$

Corollary 13.1. *If a coalition satisfies the maximum cooperation criterion with respect to the allocation x^*, then x^* is the unique solution of the negotiation process.*

Proof. From the definition of maximum cooperation criterion we have $x^* \in \cap_0^\infty D^t$ so that $\cap_0^\infty D^t \neq \emptyset$. The rest follows from the second part of Theorem 13.3. ■

Theorem 13.4. *At each stage of the negotiation process, the set of possible solutions to the game is getting smaller. This reduces, at each stage, the labour of seeking alternative solutions. Thus the negotiation process converges rapidly.*

Theorem 13.5. *When there is a solution, as a result of agreement among the players, no further improvement of the distribution gives a better solution with increased satisfaction to all the players including the mediator simultaneously, a Pareto optimality condition.*

13.4 An Example

We consider a practical problem of allocating payoffs among the members of a design and production engineering project of an organization which hires groups of experts from different desciplines. Mich, Fedrizzi, and Garigliano [16], in their paper, presented a general strategy for the management of decision making under uncertainty in industry, and its application to the specific problem of route generation in an electric engineering company. They have pointed out that a major problem in the manufacturing industry is integration of the design and production engineering process. Motivated by their paper, we have considered the subjectivity inevitably present in design and production engineering processes of a manufacturing industry. There is a strong link between these two phases, because the characteristics of a design determine the manufacturing process needed and, vice versa, features of the production cycle act as constraints on the acceptable designs. In the system, we have a scientific engineering team (SET) such that

SET $= \{G_1, G_2, \ldots, G_n\}$ where for example,
$G_1 =$ Group of production engineers
$G_2 =$ Group of designers
$G_3 =$ Group of accountants
$G_4 =$ Group of personal mangers, and so on

Each expert group has realistically specific competencies and local goals (rates of participation and demands) that they ask in order to define the design and production vectors. Thereafter, every group has given a membership of their participation (fuzzy coalition) in the project. We assume that the worth of the coalition is the amount generated for the payment to the members after completion of the work. This amount is indeed a suitable percentage of the surplus incurred after all the production expenses including salary to the members of the hiring authority have been made. A solution to the above problem can be termed as a goal vector, where each of the groups should attain its own specific goals. In practice, it is evident that the members of G_i account for their participation in the project according to their requirement in it and the success of the project would depend upon their joint endeavour. Therefore, special emphasis is given to the satisfaction of the members. Thus, solving the production–design problem is indeed equivalent to solving the cooperative game with fuzzy coalitions, where a solution vector means a goal vector which should contain the specific goal of each expert group (fuzzy coalition). The mediator here, is a representative appointed by the hiring organization.

In our example, for simplicity, we took only three groups of experts, namely, G_1, G_2, and G_3 representing three individual players. Let $Q = (1,1,1)$ represent total resource vector of the players. For a particular project, the players offered their participation as $c = (0.2, 0.4, 0.5)$. Now, with the terminologies used in Definition [13.2], we have the worth $v(c) = 420$, $v(0|c_1) = 20$, $v(0|x_2) = 80$, $v(0|x_3) = 100$, and $D^0 = \{z = (z_1, z_2, z_3) \in \mathbb{R}^3 \mid z_1 + z_2 + z_3 = 420\}$

The players announced their demands as (100, 200, 300). This is not feasible. So the negotiation took place. The mediator started with the minimum deal index as the initial offer.

Thus $x^0 = (60, 160, 200)$.

Satisfaction rates were announced by the members at $t = 0 : (0.85, 0.8, 0.95)$. Based on the satisfaction rates, one can easily infer that a solution should not be far from the minimum deal index. Thus the mediator constructed the initial beliefs as follows.

For any $z = (z_1, z_2, z_3) \in D^0$.

$$\mu_1^0(z) = \begin{cases} 0 & \text{if } z_1 < 50 \\ \frac{z_1 - 50}{370} & \text{if } 50 < z_1 < 420 \\ 1 & \text{if } z_1 \geq 420 \end{cases} \tag{13.11}$$

$$\mu_2^0(z) = \begin{cases} 0 & \text{if } z_2 < 150 \\ \frac{z_2 - 150}{270} & \text{if } 150 < z_2 < 420 \\ 1 & \text{if } z_2 \geq 420 \end{cases} \tag{13.12}$$

$$\mu_3^0(z) = \begin{cases} 0 & \text{if } z_3 < 210 \\ \frac{z_3 - 210}{210} & \text{if } 210 < z_3 < 420\,. \\ 1 & \text{if } z_3 \geq 420 \end{cases} \tag{13.13}$$

Consequently, the bargining cost function $r(i,t)$ is defined as follows.

$$r(i,t) = \frac{x_i^t(1 - S\_P_i(x^t))}{c_i v(c)} \cdot \phi(t),$$

where $\phi(t)$ is defined as: $\phi(t) = t + 1$.
Second Offer made by the mediator: $x^1 = (70.4039, 161.0930, 188.5031)$.
Satisfaction rates announced by the members at $t = 1$: $(0.9, 0.85, 0.9)$.
Test condition was checked.
As $\texttt{Sim}(S\_P_1, \ldots, S\_P_n)(x^1) = 0.005 \leq \texttt{Sim}(S\_P_1, \ldots, S\_P_n)(x^0) = 0.035$, the mediator would continue the negotiation process.
Third Offer made by the mediator: $x^2 = (84.26, 163.58, 172.16)$.
Satisfaction rates announced by the members at $t = 2$: $(0.95, 0.9, 0.85)$.
As $\texttt{Sim}(S\_P_1, S\_P_2, S\_P_3)(x^2) = 0.02 > \texttt{Sim}(S\_P_1, S\_P_2, S\_P_3)(x^1) = 0.005$, so the stopping condition is met and x^1 is the required solution of the negotiation process.
The bargaining cost of the game is computed as $(0.1678, 0.5753, 0.4488)$.

13.5 Conclusion

We have developed a model to solve dynamically a cooperative game with fuzzy coalitions. The static solutions to such games incorporate only the rates of player participation, however, we have argued that player satisfaction is an essential component of determining payoffs to participating players. We have given three different aspects for incorporating satisfaction in finding a solution acceptable to all players. Satisfaction can provide binding incentives to the players for participating in the coalitions. Similarly there are situations where despite low participation in the coalition, an individual has to be paid a handful of payoff in order to make it more worthwhile. We have given an example of such a situation in the introduction. Thirdly, likings and dislikings among the players over different issues in a coalition influence its worth. Thus it is expected that the solution vector is so designed that all the players are equally satisfied. We have developed an algorithm for the purpose and validated it by means of a simple example. A solution of the negotiation process is proposed and its existence is proved. To speed up the negotiation process and also to avoid irrational demands by the players, we have proposed the notion of a bargaining cost to be paid by the players to the mediator. The success of the model depends primarily on generating the initial beliefs by the mediator which represent the basic human characteristics of increased satisfaction over enhanced payoff. However, in order to model more complex characteristics where enhanced payoff will result in no further increase in satisfaction, or a slight increase (decrease) of payoff may boost (fade away) satisfaction of the players, the simple monotonic increasing function signifying the initial belief will not be appropriate. Thus, at a later stage, we propose to incorporate those variants in the model to make it more compatible with real-life situations.

Acknowledgements The author is highly grateful to Dr. S.K. Mishra, BHU, India for his constant support and suggestions for completion of this work.

References

1. Aubin, J.P.: *Mathematical Methods of Game and Economic Theory* (rev. ed.)., North-Holland, Amsterdam (1982).
2. Azrieli, Y., Lehrer, E.: On some families of cooperative fuzzy games, *Int. J. Game Theory*, **36**, 1–15 (2007)
3. Borkotokey, S.: Cooperative games with fuzzy coalitions and fuzzy characteristic functions, *Fuzzy Sets Syst.*, **159**, 138–151 (2008)
4. Branzei, R., Dimitrov, D., Tijs, S.: *Models in Cooperative Game Theory: Crisp, Fuzzy and Multichoice Games*, Lecture Notes in Economics and Mathematical Systems, Springer, **556**, Berlin (2004)
5. Butnariu, D.: Stability and Shapley value for an n-persons fuzzy game, *Fuzzy Sets Syst.* **4**, 63–72 (1980)
6. Carmichael, F.: *A Guide to Game Theory*, Pearson Education Limited, Prentice Hall, Upper Saddle River (2005)
7. Dieckmann, T.: Dynamic coalition formation and the core, *J. Econ. Behav. Org.*, **49**, 3, 363–380 (2002)
8. Friedman, J.W.: *Game Theory with Applications to Economics*, Oxford University Press, New York (1986)
9. Furnham, A., Forde, L., Kirsti, K.: Personality and work motivation, *Personality Individual Differences*, **26**, 1035–1043 (1999)
10. Lai, K.R., Lin, M.W.: Modeling agent negotiation via fuzzy constraints in e-business, *Comput. Intell.*, **20**, 4, (2004)
11. Lehrer, E.: Allocation processes in cooperative games, *Int. J. Game Theory*, **31**, 651–654 (2002)
12. Li, S., Zhang, Q.: A simplified expression of the Shapley function for fuzzy game, *European J. Oper. Res.*, **196**, 234–245 (2009)
13. Lim C.S., Zain Mohamed M.: Criteria of project success: An exploratory re-examination, *Int. J. Project Manage.*, **17**, 243–248 (1999)
14. Luo, X., Jennings, N.R., Shadbolt, H., Leung, F., Lee, J.H.M.: A fuzzy constraint based model for bilateral multi-issue negotiations in semi competitive environments, *Artif. Intell.*, **148**, 53–102 (2003)
15. Mares, M., Vlach, M.: Fuzzy coalitional structures, *Math. Soft Comput.* XIII, **1**, 59–70 (2006)
16. Mich, L., Fedrizzi, M., Garigliano, R.: *Negotiation and Conflict Resolution in Production Engineering Through Source Control, Fuzzy Logic and Soft Computing* (Ed.), Advances in Fuzzy Systems-Applications and Theory **4**, World Scientific, Singapore, 181–188 (1995)

17. Nash, J.F.: The bargaining problem, *Econometrica*, **18**, 155–162 (1950)
18. Ray, D., Vohra, R.: Equilibrium binding agreements. *J. Econ. Theory*, **73**, 30–78 (1997)
19. Ray, D., Vohra, R.: A theory of endogenous coalition structure, *Games Econ. Behav.*, **26**, 286–336 (1999)
20. Ray, D., Vohra, R.: Coalitional power and public goods. *J. Pol. Econ.* **109**, 1355–1384 (2001)
21. Tohme, F., Sandholm, T.: Coalition formation process with belief revision among bounded self interested agents, (Source: Internet, Open access article).
22. Tsurumi, M., Tanino, T., Inuiguchi, M.: Theory and methodology–A Shapley function on a class of cooperative fuzzy games, *European J. Oper. Res.*, **129**, 596–618 (2001)
23. Yager, R.: Multiagent negotiation using linguistically expressed mediation rules, *Group Decision Negotiation*, **16** 1–23 (2007)

Chapter 14
Characterizations of the Solution Sets and Sufficient Optimality Criteria via Higher-Order Strong Convexity

Pooja Arora, Guneet Bhatia, and Anjana Gupta

Abstract In this chapter we introduce four new generalized classes of strongly convex functions, namely strongly pseudoconvex type I and type II of order m and strongly quasiconvex type I and type II of order m. Characterizations of the set of strict minimizers of order m via strong convexity of order m are derived. Sufficient optimality conditions for higher-order efficient solutions for a vector optimization problem are presented. Some mixed duality results are also established.

14.1 Introduction

The theory of vector optimization is a problem of continuing interest in defining and characterizing its solution. Several solution concepts of vector optimization problems have emerged in the literature in an urge to obtain more satisfactory representation of such points. The concept of local minimizer of higher order in nonlinear programming originated from the study of iterative numerical methods. Auslender [1] derived necessary and sufficient optimality conditions for isolated local minima of orders 1 and 2 for the optimization problem,

$$\text{minimize}\{f(x) : x \in C\},$$

Pooja Arora
Department of Operational Research, University of Delhi, Delhi 110 007, India
e-mail: poojaarora82@rediffmail.com

Anjana Gupta
Department of Mathematics, MAIT, GGSIPU, Delhi, India
e-mail: guptaanjana2003@yahoo.co.in

Guneet Bhatia
Department of Mathematics, University of Delhi, Delhi 110 007, India
e-mail: guneet172@yahoo.co.in

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_14, © Springer Science+Business Media, LLC 2011

where $f : R^n \to R$ is a locally Lipschitz function and the feasible set C is a closed subset of R^n.

Studniarski [5] extended the definition of Auslender [1] for isolated local minima of orders 1 and 2 to the isolated local minima of order m, a positive integer, under the assumption that C is any subset of R^n which is not necessarily closed. He also derived the optimality conditions for the above problem by means of lower and upper Dini directional derivatives of f. Ward [3] renamed the isolated minima of order m as the strict local minimizer of order m for the scalar optimization problem and characterized it using tangent cones. Jimenez [2] extended the ideas of Ward [3] to define the notion of strict local efficient solution of order m for the vector minimization problem.

It is worthwhile to note that the notions of convexity and generalized convexity play a crucial role in optimization theory. In this chapter we use strong convexity of order m to derive a characterization for the set of strict minimizers of order m for a scalar optimization problem. Furthermore, four new generalizations of strongly convex functions of order m are employed to develop sufficient optimality conditions and to establish mixed duality results for a vector optimization problem.

The chapter has been organized as follows. In Section 14.2, we consider strong convexity of higher order for differentiable functions and its characterization in terms of gradient vector. This characterization leads us to the four new generalizations of strong convexity of order m, viz. strongly pseudoconvex type I and type II functions of order m and strongly quasiconvex type I and type II functions of order m. Examples are presented to illustrate the relationship between these new classes and the existing notions of pseudoconvexity and quasiconvexity. In Section 14.3, we characterize the set of strict minimizers of order m for a scalar optimization problem via strong convexity of order m. In Section 14.4, we study the concept of efficient solution of order m for a vector optimization problem. Optimality conditions are derived. A dual is proposed and mixed duality results are established in Section 14.5.

14.2 Strongly Convex Function

This section starts with the concepts of a strongly convex function of order m. Let X be an open convex subset of R^n equipped with the Euclidean norm $\|\cdot\|$.

Definition 14.1 ([4]). A function $f : X \to R$ is said to be a *strongly convex function* of order m if there exists a constant $c > 0$ such that for any $x, y \in X$ and $t \in [0,1]$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - ct(1-t)\|x-y\|^m$$

For $m = 2$, the function is referred to as *strongly convex* in the ordinary sense [7]. Strong convexity of any order implies convexity but the converse is not true, in general [4].

Remark 14.1 ([4]). If each f_i, $i = 1,2,\ldots,p$ is strongly convex of order m on X, then for $t_i > 0$, $i = 1,2,\ldots,p$, $\sum_{i=1}^{p} t_i f_i$ and $\max_{1\leq i\leq p} f_i$ are also strongly convex of order m on X.

The following theorem gives the characterization of a strongly convex function in terms of its gradient vector.

Theorem 14.1 ([4]). *Suppose $f : X \to R$ is continuously differentiable on X. Then, f is a strongly convex function of order m on X if and only if there exists a constant $c > 0$ such that*

$$f(y) \geq f(x) + \nabla f(x)^t (y-x) + c\|y-x\|^m, \qquad \forall\, x,y \in X. \tag{14.1}$$

The above characterization leads us to the following new classes of functions.

Definition 14.2. A differentiable function $f : X \to R$ is said to be *strongly pseudoconvex type I of order m on X* if there exists a constant $c > 0$ such that for any $x,y \in X$

$$\nabla f(x)^t (y-x) \geq 0 \Rightarrow f(y) \geq f(x) + c\|y-x\|^m.$$

Remark 14.2. Every strongly pseudoconvex type I function of order m is pseudoconvex. However, the converse of the above statement may not be true. For example,

$$f(x) = \begin{cases} x^4 & \text{if } x \leq 0 \\ 0 & \text{if } x > 0 \end{cases}$$

is pseudoconvex but is not strongly pseudoconvex type I of any order, as for $x = 0$, $y > 0$ we have $\nabla f(x)^t (y-x) = 0$, however, $f(y) \geq f(x) + c\|y-x\|^m$ is not true, for any $c > 0$.

Definition 14.3. A differentiable function $f : X \to R$ is said to be *strongly pseudoconvex type II of order m on X* if there exists a constant $c > 0$ such that for any $x,y \in X$

$$\nabla f(x)^t (y-x) + c\|y-x\|^m \geq 0 \Rightarrow f(y) \geq f(x).$$

Remark 14.3. Every strongly pseudoconvex type II function of order m is pseudoconvex. However, the converse of the above statement may not be true. For example, $f(x) = -x$ is pseudoconvex but is not strongly pseudoconvex type II of any order; as for $y = 1$, $x = 0$ we have $f(y) < f(x)$ but $\nabla f(x)^t (y-x) + c\|y-x\|^m > 0$, for any $c > 1$.

Definition 14.4. A differentiable function $f : X \to R$ is said to be strongly quasiconvex type I of order m on X if there exists a constant $c > 0$ such that for $x,y \in X$,

$$f(y) \leq f(x) \Rightarrow \nabla f(x)^t (y-x) + c\|y-x\|^m \leq 0$$

Remark 14.4. Every strongly quasiconvex type I function of order m is quasiconvex. The converse of the above statement may not hold. The function $f(x) = \sqrt{1-x^2}$,

$x \in [0,1]$ is quasiconvex but is not strongly quasiconvex type I of any order. As for $y = 1/2$, $x = 0$, we have $f(y) \leq f(x)$ but $\nabla f(x)^t(y-x) + c\|y-x\|^m \leq 0$ does not hold for any $c > 0$.

Definition 14.5. A differentiable function $f : X \to R$ is said to be *strongly quasiconvex type II of order m on X* if there exists a constant $c > 0$ such that for $x, y \in X$

$$f(y) \leq f(x) + c\|y-x\|^m \Rightarrow \nabla f(x)^t(y-x) \leq 0.$$

Remark 14.5. Every strongly quasiconvex type II function of order m is quasiconvex.

14.3 Characterization of Solution Sets

In this section, strong convexity of order m is employed to derive a characterization of the set of strict minimizers of order m for the following scalar optimization problem,

$$\text{(P)} \quad \text{minimize } f(x)$$
$$\text{subject to } x \in S,$$

where S is a convex subset of R^n and f is a real-valued differentiable function defined on an open subset X of R^n.

We now extend the notion of strict minimizer of order m defined in [3] as follows.

Definition 14.6. Let $m \geq 1$ be an integer. A point $x^0 \in S$ is said to be a *strict minimizer of order m* for (P) if there exists a positive number $\alpha > 0$ such that

$$f(x) \geq f(x^0) + \alpha\|x-x^0\|^m, \qquad \forall\, x \in S.$$

We denote the set of all strict minimizers of order m as $\bar{S}$. Throughout the section, we assume the set $\bar{S}$ to be nonempty.

Definition 14.7 ([4]). A map $F : X \to R^n$ is *strongly monotone of order m* on X if there exists $\alpha > 0$ such that, for $x, y \in X$,

$$(F(y) - F(x))^t(y-x) \geq \alpha\|y-x\|^m.$$

For $m = 2$, the map is referred to as *strongly monotone* [7].

Remark 14.6 ([4]). It is evident that if f is differentiable and strongly convex of order m on X then ∇f is strongly monotone of order m on X, where $\alpha = 2c$.

Lemma 14.1. *If f is differentiable and strongly convex of order m on X and $\bar{x}, \bar{y} \in \bar{S}$, then*

$$\nabla f(\bar{x})^t(\bar{y}-\bar{x}) = \nabla f(\bar{y})^t(\bar{x}-\bar{y}) = 0.$$

Proof. Because $\bar{x}, \bar{y} \in \bar{S}$, we obtain

$$\nabla f(\bar{x})^t(\bar{y}-\bar{x}) \geq 0 \quad \text{and} \quad \nabla f(\bar{y})^t(\bar{x}-\bar{y}) \geq 0. \tag{14.2}$$

As f is strongly convex of order m, then it follows from Remark 14.6 that ∇f is strongly monotone of order m. Hence, the inequalities in (14.2) imply, respectively, that

$$\nabla f(\bar{y})^t(\bar{y}-\bar{x}) \geq 0 \quad \text{and} \quad \nabla f(\bar{x})^t(\bar{x}-\bar{y}) \geq 0. \tag{14.3}$$

Combining the inequalities in (14.2) and (14.3) yields the required conclusion.

The following theorem gives a characterization of the set of strict minimizers of order m.

Theorem 14.2. *If f is differentiable and strongly convex of order m on X and $\bar{x} \in \bar{S}$, then $\bar{S} = \tilde{S} = S'$, where*

$$\tilde{S} = \{x \in S : \nabla f(x)^t(\bar{x}-x) = 0\}, \tag{14.4}$$

$$S' = \{x \in S : \nabla f(x)^t(\bar{x}-x) \geq 0\}. \tag{14.5}$$

Proof. Let $x \in \bar{S}$; then as $\bar{x} \in \bar{S}$, it follows from Lemma 14.1 that $\nabla f(x)^t(\bar{x}-x) = 0$. Thus, $x \in \tilde{S}$ which implies that $\bar{S} \subseteq \tilde{S}$. Conversely, if $x \in \tilde{S}$ then $\nabla f(x)^t(\bar{x}-x) = 0$. Because f is strongly convex of order m, we have

$$f(\bar{x}) \geq f(x).$$

Because $\bar{x} \in \bar{S}$ and the above inequality holds, we have

$$f(\bar{x}) = f(x).$$

Then it follows that $\tilde{S} \subseteq \bar{S}$ and hence, $\tilde{S} = \bar{S}$. It is obvious from (14.4) that the inclusion $\bar{S} \subseteq S'$ holds. Assume that $x \in S'$ and $\nabla f(x)^t(\bar{x}-x) \geq 0$. In as much as f is strongly convex of order m, we have

$$f(\bar{x}) \geq f(x),$$

$\bar{x} \in \bar{S}$ and the above inequality holds, therefore we have

$$f(\bar{x}) = f(x).$$

Thus, $S' \subseteq \bar{S}$ and hence $\bar{S} = S'$.

Theorem 14.3. *If f is differentiable and strongly convex of order m on X and $\bar{x} \in \bar{S}$, then $\bar{S} = S^* = \hat{S}$, where*

$$S^* = \{x \in S : \nabla f(\bar{x})^t(x-\bar{x}) = \nabla f(x)^t(\bar{x}-x)\}, \tag{14.6}$$

$$\hat{S} = \{x \in S : \nabla f(\bar{x})^t(x-\bar{x}) \leq \nabla f(x)^t(\bar{x}-x)\}. \tag{14.7}$$

Proof. We first show $\bar{S} \subseteq S^*$. Let $x \in \bar{S}$; then it follows from Lemma 14.1 that

$$\nabla f(x)^t(\bar{x}-x) = \nabla f(\bar{x})^t(x-\bar{x}) = 0.$$

Thus, $x \in S^*$. The inclusion $S^* \subseteq \hat{S}$ follows trivially.

Let $x \in \hat{S}$; then it follows from (14.7) that

$$\nabla f(\bar{x})^t(x-\bar{x}) \leq \nabla f(x)^t(\bar{x}-x). \tag{14.8}$$

Because $\bar{x} \in \bar{S}$ we have

$$\nabla f(\bar{x})^t(\bar{y}-\bar{x}) \geq 0;$$

then it follows from (14.8) that

$$\nabla f(x)^t(\bar{x}-x) \geq 0.$$

Because f is strongly convex of order m, we have

$$f(\bar{x}) \geq f(x).$$

Because $\bar{x} \in \bar{S}$, the above inequality yields that

$$f(\bar{x}) = f(x).$$

Thus, $x \in \bar{S}$ and therefore, $\hat{S} \subseteq \bar{S}$.

14.4 Optimality Conditions

In this section, we present the necessary and sufficient conditions for the following vector optimization problem (VP) to possess an efficient solution of order m

$$\text{(VP)} \quad \text{Minimize } f(x) = (f_1(x), \ldots, f_p(x))$$
$$\text{subject to } g_j(x) \leq 0, \qquad j = 1,2,\ldots,q,$$

where $f_i, g_j : X \to R$, $i = 1,2,\ldots,p$, $j = 1,2,\ldots,q$ are real-valued differentiable functions.

Let $S = \{x \in X : g_j(x) \leq 0, j = 1,2,\ldots,q\}$ be the set of all feasible solutions for (VP).

Definition 14.8. A point $x^0 \in S$ is said to be an *efficient solution* for (VP) if

$$f(x) \nleq f(x^0), \qquad \forall\, x \in S;$$

that is, there exists no $x \in S$, such that

$$f_i(x) \leqq f_i(x^0) \quad \forall\, i = 1,2,\ldots,p,\; i \neq j$$
$$f_j(x) < f_j(x^0) \quad \text{for some } \; j.$$

Definition 14.9. Let $m \geq 1$ be an integer. A point $x^0 \in S$ is said to be an *efficient solution* of order m for (VP) if there exists an $\alpha \in \operatorname{int} R^p_+$ such that

$$f(x) \nleq f(x^0) + \alpha \parallel x - x^0 \parallel^m, \qquad \forall\, x \in S;$$

that is, there exists no $x \in S$, such that

$$f_i(x) \leqq f_i(x^0) + \alpha_i \parallel x - x^0 \parallel^m, \qquad \forall\, i = 1,2,\ldots,p,\; i \neq j,$$
$$f_j(x) < f_j(x^0) + \alpha_j \parallel x - x^0 \parallel^m, \qquad \text{for some } \; j.$$

Remark 14.7. For any integer $m \geq 1$, if x is an efficient solution of order m for (VP), then it is also an efficient solution for (VP). But the converse is not necessarily true as can be seen by the following example,

$$\text{minimize} f(x) = (f_1(x), f_2(x))^t, \qquad x \in R$$

$$f_1(x) = \begin{cases} -x^2 \sin\left(\frac{1}{x}\right) - x^2, & \text{if } \; x \neq 0, \\ 0 & \text{if } \; x = 0, \end{cases}$$

and $f_2(x) = (f_1(x))^2$. It is easy to verify that $x^0 = 0$ is an efficient solution for (VP) but it is not an efficient solution of order m, as for any integer m, $\alpha \in \operatorname{int} R^2_+$ and $\varepsilon > 0$ we can choose a positive integer n, $n > 1/4((2/\varepsilon\pi) + 1)$ and $x = (2/(4n-1)\pi)$ such that $x \in (-\varepsilon, \varepsilon)$ and $f(x) < \alpha|x|^m$.

Theorem 14.4 (Karush–Kuhn–Tucker-type necessary optimality conditions). *Suppose $x^0 \in S$ is an efficient solution of order m and the functions f_i, $i = 1,2,\ldots,p$ and g_j, $j = 1,2,\ldots,q$ are differentiable at x^0. Let an appropriate constraint qualification* [6] *hold at x^0; then there exist $\lambda^0 \in R^p_+$, $\mu^0 \in R^q_+$ such that*

$$\sum_{i=1}^{p} \lambda_i^0 \nabla f_i(x^0) + \sum_{j=1}^{q} \mu_j^0 \nabla g_j(x^0) = 0 \tag{14.9}$$

$$\mu_j^0 g_j(x^0) = 0, \qquad j = 1,2,\ldots,q \tag{14.10}$$

$$\lambda^{0t} e = 1, \quad \text{where } \; e = (1,\ldots,1) \in R^p_+ \tag{14.11}$$

The following results present sufficient conditions for the existence of a higher-order efficient solution.

Theorem 14.5. *Let the conditions* (14.9)–(14.11) *be satisfied at $x^0 \in S$. Suppose f_i, $i = 1,2,\ldots,p$ are strongly convex of order m on X and $\mu_j^0 g_j$, $j = 1,2,\ldots,q$ are* *strongly quasiconvex type I of order m on X. Then x^0 is an efficient solution of order m for* (VP).

Proof. Because $f_i, i = 1, 2, \ldots, p$ are strongly convex of order m therefore there exist constants $c_i > 0$, $i = 1, 2, \ldots, p$ such that

$$f_i(x) - f_i(x^0) \geq (x - x^0)^t \nabla f_i(x^0) + c_i \parallel x - x^0 \parallel^m .$$

As $\lambda_i \geq 0$, $i = 1, 2, \ldots, p$ therefore from the above inequality we have

$$\sum_{i=1}^{p} \lambda_i^0 f_i(x) - \sum_{i=1}^{p} \lambda_i^0 f_i(x^0) \geq (x - x^0)^t \sum_{i=1}^{p} \lambda_i^0 \nabla f_i(x^0) + \sum_{i=1}^{p} \lambda_i^0 c_i \parallel x - x^0 \parallel^m \quad (14.12)$$

Now for $x \in S$, $g_j(x) \leq 0$, $\forall\ j = 1, 2, \ldots, q$. As $\mu_j^0 \geq 0$, $\forall\ j = 1, 2, \ldots, q$, therefore we have

$$\mu_j^0 g_j(x) \leq 0 = \mu_j^0 g_j(x^0).$$

$\mu_j^0 g_j$, $j = 1, 2, \ldots, q$ are strongly quasiconvex type I of order m, therefore there exist constants $c_j > 0$, $j = 1, 2, \ldots, q$ such that

$$(x - x^0)^t \nabla \mu_j^0 g_j(x^0) + c_j \parallel x - x^0 \parallel^m \leq 0, \forall j = 1, 2, \ldots, q$$

or

$$(x - x^0)^t \sum_{j=1}^{q} \mu_j^0 \nabla g_j(x^0) + \sum_{j=1}^{q} c_j \parallel x - x^0 \parallel^m \leq 0. \quad (14.13)$$

Adding (14.12) and (14.13) and using (14.9), we get

$$\sum_{i=1}^{p} \lambda_i^0 f_i(x) - \sum_{i=1}^{p} \lambda_i^0 f_i(x^0) \geq \left(\sum_{i=1}^{p} \lambda_i^0 c_i + \sum_{j=1}^{q} c_j \right) \parallel x - x^0 \parallel^m$$

which implies that

$$\lambda^t (f(x) - f(x^0)) \geq a \parallel x - x^0 \parallel^m,$$

where $a = \sum_{i=1}^{p} \lambda_i^0 c_i + \sum_{j=1}^{q} c_j$. This implies that

$$\lambda^t [f(x) - f(x^0) - c \parallel x - x^0 \parallel^m] \geq 0, \quad (14.14)$$

where $c = ae$, in as much as $\lambda^t e = 1$. It follows from (14.14) that there exists $c \in \text{int } R_+^p$ such that for all $x \in S$,

$$f(x) \not\leq f(x^0) + c \parallel x - x^0 \parallel^m,$$

thereby implying that x^0 is an efficient solution of order m for (VP).

Theorem 14.6. *Let the conditions* (14.9)–(14.11) *be satisfied at $x^0 \in S$. Suppose $\lambda^{0t} f$ is strongly quasiconvex type II of order m on X and $\mu^{0t} g$ is strongly pseudoconvex type II of order m on X. Then x^0 is an efficient solution of order m.*

Proof. Let us assume on the contrary that x^0 is not an efficient solution of order m for (VP). Then for every $\alpha \in \text{int}\, R^p_+$ we have

$$f(x) \leq f(x^0) + \alpha \parallel x - x^0 \parallel^m .$$

Because $\lambda^{0t} e = 1$, we therefore have

$$\begin{aligned} \lambda^{0t} f(x) &\leqq \lambda^{0t} f(x^0) + \lambda^{0t} \alpha \parallel x - x^0 \parallel^m \\ &= \lambda^{0t} f(x^0) + c \parallel x - x^0 \parallel^m, \quad \text{where } c = \lambda^{0t}\alpha > 0. \end{aligned}$$

As $\lambda^{0t} f$ is strongly quasiconvex type II of order m at x^0, the above relation implies that

$$(x - x^0)^t \nabla \lambda^{0t} f(x^0) \leq 0.$$

Then it follows from (14.9) that

$$(x - x^0)^t \nabla \mu^{0t} g(x^0) \geq 0$$

or

$$(x - x^0)^t \nabla \mu^{0t} g(x^0) + c \parallel x - x^0 \parallel^m \geq 0, \qquad \forall\, c > 0.$$

Because $\mu^{0t} g$ is strongly pseudoconvex type II of order m, the above inequality yields that

$$\mu^{0t} g(x) \geq \mu^{0t} g(x^0) = 0$$

which is not possible. Hence x^0 is an efficient solution of order m.

14.5 Mixed Duality

In this section, we develop the duality relationship between (VP) and its mixed dual under generalized convexity assumptions.

Let the index set $Q = \{1, 2, \ldots, q\}$ be partitioned into two disjoint subsets K and J such that $Q = K \cup J$. The mixed dual for (VP) is given by

$$\begin{aligned} \text{(VD)} \quad \text{maximize} \quad & f(u) + \mu_J g_J(u) e \\ \text{subject to} \quad & \lambda^t \nabla f(u) + \mu^t \nabla g(u) = 0 \\ & \mu_k g_k(u) \geq 0, \qquad k \in K \\ & \lambda \geq 0,\, \lambda^t e = 1,\, \mu \geq 0,\, e = (1, \ldots, 1)^t \in R^p. \end{aligned}$$

Let $S_D = \{(u, \lambda, \mu) \,|\, \lambda^t \nabla f(u) + \mu^t \nabla g(u) = 0,\ \mu_k g_k(u) \geq 0,\ k \in K,\ \lambda \geq 0,\ \lambda^t e = 1,\ \mu \geq 0\}$ be the feasible set of (VD).

Theorem 14.7 (Weak Duality). *Let x and (u, λ, μ) be feasible for* (VP) *and* (VD)*, respectively. Suppose $\left(\sum_{i=1}^p \lambda_i f_i + \sum_{j \in J} \mu_j g_j\right)(\cdot)$ is strongly pseudoconvex type I of* *order m at u and* $\sum_{k\in K}\mu_k g_k(\cdot)$ *is strongly quasiconvex type I of order m at u; then the following does not hold*

$$f_i(x) < f_i(u) + \sum_{j\in J}\mu_j g_j(u), \qquad \forall\, i = 1,2,\ldots,p.$$

Proof. Because (u,λ,μ) is feasible for (VD),

$$\sum_{i=1}^{p}\lambda_i \nabla f_i(u) + \sum_{j=1}^{q}\mu_j \nabla g_j(u) = 0 \tag{14.15}$$

and

$$\mu_k g_k(u) \geq 0, \qquad k \in K.$$

As $x \in S, g_k(x) \leq 0, k \in K$, and also $\mu_k \geq 0, k \in K$, therefore

$$\sum_{k\in K}\mu_k g_k(x) \leq \sum_{k\in K}\mu_k g_k(u).$$

$\sum_{k\in K}\mu_k g_k(\cdot)$ is strongly quasiconvex type I of order m; from the above inequality it thus follows that

$$(x-u)^t \sum_{k\in K}\mu_k \nabla g_k(x) + c \parallel x-u \parallel^m \leq 0.$$

Using (14.15) we have

$$(x-u)^t \left[\sum_{i=1}^{p}\lambda_i \nabla f_i(u) + \sum_{j\in J}\mu_j \nabla g_j(u)\right] - c \parallel x-u \parallel^m \geq 0$$

or

$$(x-u)^t \left[\sum_{i=1}^{p}\lambda_i \nabla f_i(u) + \sum_{j\in J}\mu_j \nabla g_j(u)\right] \geq 0.$$

$\left(\sum_{i=1}^{p}\lambda_i f_i + \sum_{j\in J}\mu_j g_j\right)(\cdot)$ is strongly pseudoconvex type I of order m at u, therefore

$$\sum_{i=1}^{p}\lambda_i f_i(x) + \sum_{j\in J}\mu_j g_j(x) \geq \sum_{i=1}^{p}\lambda_i f_i(u) + \sum_{j\in J}\mu_j g_j(u) + c \parallel x-u \parallel^m$$

$$\sum_{i=1}^{p}\lambda_i f_i(x) + \sum_{j\in J}\mu_j g_j(x) \geq \sum_{i=1}^{p}\lambda_i f_i(u) + \sum_{j\in J}\mu_j g_j(u).$$

Using $g_j(x) \leq 0$, $\mu_j \geq 0$, $j \in J$, we have

$$\lambda^t(f(x) - f(u) - \sum_{j\in J}\mu_j g_j(u)) \geq 0,$$

hence the result follows.

Theorem 14.8 (Weak Duality). *Let x and (u,λ,μ) be feasible for* (VP) *and* (VD)*, respectively. Suppose $(\sum_{i=1}^{p}\lambda_i f_i + \sum_{j\in J}\mu_j g_j)(\cdot)$ is strongly pseudoconvex type II of order m at u and $\sum_{k\in K}\mu_k g_k(\cdot)$ is strongly quasiconvex type II of order m at u. Then the following does not hold.*

$$f_i(x) < f_i(u) + \sum_{j\in J}\mu_j g_j(u), \qquad \forall\, i = 1,2,\ldots,p$$

Proof. (u,λ,μ) is feasible for (VD), therefore

$$\sum_{i=1}^{p}\lambda_i \nabla f_i(u) + \sum_{j=1}^{q}\mu_j \nabla g_j(u) = 0 \tag{14.16}$$

and

$$\mu_k g_k(u) \geq 0, \qquad k \in K.$$

Moreover, because x is feasible for (VP), $g_k(x) \leq 0,\ k \in K$. Also $\mu_k \geq 0,\ k \in K$; then from the above inequality, it follows that

$$\sum_{k\in K}\mu_k g_k(x) \leq \sum_{k\in K}\mu_k g_k(u)$$

or

$$\sum_{k\in K}\mu_k g_k(x) \leq \sum_{k\in K}\mu_k g_k(u) + c \parallel x-u \parallel^m, \qquad \forall\, c > 0.$$

Because $\sum_{k\in K}\mu_k g_k(\cdot)$ is strongly quasiconvex type II of order m,

$$(x-u)^t \sum_{k\in K}\mu_k \nabla g_k(x) \leq 0.$$

Using (14.16) we have

$$(x-u)^t\left[\sum_{i=1}^{p}\lambda_i \nabla f_i(u) + \sum_{j\in J}\mu_j \nabla g_j(u)\right] + c \parallel u-x \parallel^m \geq 0, \qquad \forall c > 0.$$

As $(\sum_{i=1}^{p}\lambda_i f_i + \sum_{j\in J}\mu_j g_j)(\cdot)$ is strongly pseudoconvex type II of order m at u, we have

$$\sum_{i=1}^{p}\lambda_i f_i(x) + \sum_{j\in J}\mu_j g_j(x) \geq \sum_{i=1}^{p}\lambda_i f_i(u) + \sum_{j\in J}\mu_j g_j(u).$$

Using $g_j(x) \leq 0$, $\mu_j \geq 0$, $j \in J$, we have

$$\lambda^t(f(x) - f(u) - \sum_{j\in J}\mu_j g_j(u)) \geq 0;$$

that is,

$$f_i(x) < f_i(u) + \sum_{j\in J}\mu_j g_j(u), \qquad \forall\, i = 1,2,\ldots,p$$

cannot hold.

Theorem 14.9 (Strong Duality). *Suppose x^0 is an efficient solution of order m for (VP) and constraint qualification* [6] *holds at x^0. Then there exist $\lambda^0 \in R^p_+$ and $\mu^0 \in R^q_+$ such that (x^0, λ^0, μ^0) is feasible for (VD). Furthermore, if the conditions of either Theorem* 14.7 *or* 14.8 *hold, then (x^0, λ^0, μ^0) is weak efficient for (VD).*

Proof. The proof follows from Theorem 14.4 and the weak duality theorem.

Acknowledgments

The authors are thankful to Professor (Mrs.) Davinder Bhatia for the constant support and encouragement throughout the preparation of this chapter.

References

1. A. Auslender, Stability in mathematical programming with non-differentiable data, *SIAM J. Control Optim.* 22 (1984), 239–254.
2. B. Jimenez, Strict efficiency in vector optimization, *J. Math. Anal. Appl.* 265 (2002), 264–284.
3. D.E. Ward, Characterization of strict local minima and necessary conditions for weak sharp minima, *J. Optim. Theory Appl.* 80 (1994), 551–571.
4. G.H. Lin, M. Fukushima, Some exact penalty results for non-linear programs and mathematical programs with equilibrium constraints, *J. Optim. Theory Appl.* 118 (2003), 67–80.
5. M. Studniarski, Necessary and sufficient conditions for isolated local minima of nonsmooth functions, *SIAM J. Control Optim.* 24 (1986), 1044–1049.
6. O.L. Mangasarian, *Nonlinear programming, corrected reprint of the 1969 original*, Classical Appl. Math., vol. 10, Society for Industrial and Applied Mathematics, SIAM, Philadelphia, PA, 1994.
7. S. Karamardian, S. Schaible, Seven kinds of monotone maps, *J. Optim. Theory Appl.* 66 (1990), 37–46.

Chapter 15
Variational Inequalities and Optimistic Bilevel Programming Problem Via Convexifactors

Bhawna Kohli

Abstract In this chapter, we introduce ∂_∞^*-pseudo-convex and ∂_∞^*-quasi-convex functions on the lines of Dutta and Chandra [13] in terms of convexifactors and their recession cone and utilize them to establish interrelations between solutions of the Bilevel programming problem and Stampacchia and Minty-type variational inequalities defined in terms of convexifactors and their recession cone. We also establish existence results for these variational inequalities.

15.1 Introduction

The variational inequality problem (VIP) has received extensive attention in recent years due to its applications in the field of economics, management sciences, and so on. In fact many equilibrium problems in economics, game theory, mechanics, traffic analysis, and so on can be transformed into variational inequality problems. It was first introduced by Hartman and Stampacchia [18] in 1966 in their seminal paper. Later on it was extended to vector variational inequality problems (VVIP) by Giannessi [14] in 1980. Since then a great deal of research started in the area of VVIP as a consequence of a lot of inclination of researchers towards vector optimization. Many researchers have contributed in this direction including Chen [5], Giannessi [17, 16, 15], Yang, and Teo [24], Mishra and Wang [21], Chinaie et al. [6], Rezaie and Zafarani [23] and so on.

The study of the bilevel programming problem (BLPP) has been motivated by its importance both in theoretical and practical (real-world) applications. The bilevel programming problem was introduced to the research community in the 1970s. Since then intensive investigation of these problems began in both theoretical and practical applications such as economics, engineering, and medicine among

Bhawna Kohli
Department of Mathematics, University of Delhi, Delhi 110 007, India
e-mail: bhawna_kohli@rediffmail.com

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_15,

others. BLPP studies the two combined optimization problems. In this, variables of the first (or leader's) problem are the parameters of the second (or follower's) problem and the optimal solution of the latter is needed to calculate the objective function value of the former. Due to its immense importance and interesting nature it has been studied by many eminent researchers such as Dempe [9, 8], Bard [3, 2], Outrata [22], Ye and Ye [25], Ye and Zhu [26], and most recently by Dempe, Dutta, and Mordukhovich [10]. For practical applications and recent developments on BLPP one can see Bard [3] and Dempe [9].

In this chapter we introduce the Minty-type variational inequality problems (MTVIP) and Stampacchia-type variational inequality problem (STVIP) in terms of convexifactors and their recession cone, and study relationships between solutions of these problems and the bilevel programming problem. Although the bulk of the literature on variational inequalities is based on the rigid assumption of generalized monotonicity, in this chapter we have made an attempt to move without it. For that we have introduced ∂^*_∞-pseudo-convex and ∂^*_∞-quasi-convex functions on the lines of Dutta and Chandra [13] in terms of convexifactors and their recession cone. We establish existence results for the two problems using the KKM lemma in terms of convexifactors. These convexifactors are important tools of nonsmooth analysis introduced by Demyanov [11] in 1994 and further studied by Jeyakumar and Luc [19], Dutta and Chandra [13, 12], and Li and Zhang [20] among others. These are recent generalizations of subdifferentials that are subsets of many well-known subdifferentials such as those of Clarke and Michel Penot. Hence our results are sharper than those using other subdifferentials.

The chapter comprises four sections. In Section 15.2, we give some basic definitions and results and also introduce the notion of ∂^*_∞-pseudo-convex and ∂^*_∞-quasi-convex functions. The bilevel programming problem is discussed in Section 15.3. Section 15.4 deals with MTVIP and STVIP and relations between solutions of these problems and the bilevel programming problem. We also establish existence results for Minty-type and Stampacchia-type variational inequalities in this section.

15.2 Preliminaries

This chapter focuses on finite-dimensional spaces. We begin by defining upper and lower Dini derivatives as follows.

Let $F : \mathbb{R}^{n_1} \to \mathbb{R} \cup \{\pm\infty\}$ be an extended real-valued function and let $x \in \mathbb{R}^{n_1}$ where $F(x)$ is finite. Then the upper and lower Dini derivatives of F at x in the direction v are defined, respectively, by

$$(F)^+_d(x,v) = \limsup_{t\to 0^+} \frac{F(x+tv)-F(x)}{t}$$

and

$$(F)^-_d(x,v) = \liminf_{t\to 0^-} \frac{F(x+tv)-F(x)}{t}$$

Dini derivatives may be finite as well as infinite. In particular if F is locally Lipschitz both the upper and lower Dini derivatives are finite.

For any set $A \subset \mathbb{R}^{n_1}$, the closure, convex hull, and the closed convex hull of A are denoted, respectively, by $\overline{A}$, CoA, and $\overline{CoA}$.

We now give the definitions of convexifactors [12].

Definition 15.1. Let $F : \mathbb{R}^{n_1} \to \mathbb{R} \cup \{\pm\infty\}$ be an extended real-valued function and let $x \in \mathbb{R}^{n_1}$ where $F(x)$ is finite.

(i) F is said to admit an *upper convexifactor* (*UCF*) $\partial^u F(x)$ at x if $\partial^u F(x) \subseteq \mathbb{R}^{n_1}$ is a closed set and

$$(F)_d^-(x,v) \leq \sup_{x^* \in \partial^u F(x)} \langle x^*, v \rangle, \quad \text{for all } v \in \mathbb{R}^{n_1}.$$

(ii) F is said to admit a *lower convexifactor* (*LCF*) $\partial_l F(x)$ at x if $\partial_l F(x) \subseteq \mathbb{R}^{n_1}$ is a closed set and

$$(F)_d^+(x,v) \geq \inf_{x^* \in \partial_l F(x)} \langle x^*, v \rangle, \quad \text{for all } v \in \mathbb{R}^{n_1}.$$

(iii) F is said to admit a convexifactor (CF) $\partial^* F(x)$ at x if $\partial^* F(x)$ is both a UCF and LCF of F at x.

It may be noted that convexifactors are not necessarily convex or compact [12, 13, 19]. Because of these relaxations convexifactors can be easily applied to a large class of nonsmooth functions.

We now state the following mean value theorem given by Jeyakumar and Luc [19].

Theorem 15.1. *Let $a, b \in \mathbb{R}^{n_1}$, and let $F : \mathbb{R}^{n_1} \to \mathbb{R}$ be a continuous function. Assume that, for each $x \in (a,b)$, $\partial^u F(x)$ and $\partial_l F(x)$ are, respectively, the upper and lower convexifactors of F at x. Then there exist $c \in (a,b)$ and a sequence $\{x_k\} \subset co(\partial^u F(c)) \cup co(\partial_l F(c))$ such that*

$$F(b) - F(a) = \lim_{k \to \infty} \langle x_k, b - a \rangle.$$

Definition 15.2 (Recession Cone of A). Let $A \subset \mathbb{R}^{n_1}$ be any nonempty set. The recession cone of A, denoted by A_∞ is defined as

$$A_\infty = \left\{ \lim_{i \to \infty} t_i a_i, a_i \in A, \{t_i\} \downarrow 0 \right\}.$$

A is bounded if and only if its recession cone is trivial.

Definition 15.3. A set-valued map $\Gamma : \mathbb{R}^{n_1} \to 2^{\mathbb{R}^{n_1}}$ is called a *KKM-map* if for every finite subset $\{u_1, u_2, \ldots, u_n\}$ of $\mathbb{R}^{n_1}$ its convex hull

$$co(\{u_1, u_2, \ldots, u_n\}) \subset \bigcup_{i=1}^{n} \Gamma(u_i).$$

We now state the following generalized Fan's KKM lemma from Rezaie and Zafarani [23] for the finite-dimensional case.

Lemma 15.1. *Let $\Gamma, \widehat{\Gamma} : A \subset \mathbb{R}^{n_1} \to 2^{\mathbb{R}^{n_1}}$ be two set-valued mappings, where A is a nonempty subset of $\mathbb{R}^{n_1}$ such that the following are satisfied.*

(i) $\widehat{\Gamma}(x) \subseteq \Gamma(x)$, for all $x \in A$.
(ii) $\widehat{\Gamma}$ is a KKM-map,
(iii) $\Gamma(x)$ is closed for all $x \in A$ and is bounded for at least one $x \in A$.

Then $\bigcap_{x\in A} \Gamma(x) \neq \phi$.

We now give the definition of ∂^*_∞-pseudo-convex function and ∂^*_∞-quasi-convex function on the lines of Dutta and Chandra [13].

Let $F : \mathbb{R}^{n_1} \to \mathbb{R}$ be a real-valued function and let $\bar{x} \in \mathbb{R}^{n_1}$.

We assume that F admits convexifactor $\partial^* F(\bar{x})$.

Definition 15.4. The function F is said to be ∂^*_∞-pseudo-convex at $\bar{x}$ if

$$F(x) < F(\bar{x}) \Rightarrow \langle \xi, x - \bar{x} \rangle < 0, \quad \text{for all } \xi \in \partial^* F(\bar{x}) \cup (\partial^* F(\bar{x}))_\infty \setminus \{0\}.$$

Remark 15.1.

(i) If F is a differentiable function then $\partial^* F(\bar{x}) = \{\nabla F(\bar{x})\}$ and the above definition reduces to the definition of pseudo-convex function.
(ii) If $\partial^* F(\bar{x})$ is bounded then the above definition reduces to the definition of ∂^*-pseudo-convex function introduced by Dutta and Chandra [13].
(iii) If F is a locally Lipschitz function and $\partial^* F(\bar{x}) = \partial^c F(\bar{x})$, where $\partial^c F(\bar{x})$ is the Clarke generalized gradient, then the above definition reduces to the definition of ∂^c-pseudo-convex function defined by Bector, Chandra, and Dutta [4].

Definition 15.5. The function F is said to be ∂^*_∞-quasi-convex at $\bar{x}$ if

$$F(x) \leq F(\bar{x}) \Rightarrow \langle \xi, x - \bar{x} \rangle \leq 0, \quad \text{for all } \xi \in \partial^* F(\bar{x}) \cup (\partial^* F(\bar{x}))_\infty \setminus \{0\}.$$

Remark 15.2.

(i) If F is a differentiable function then $\partial^* F(\bar{x}) = \{\nabla F(\bar{x})\}$ and the above definition reduces to the definition of quasi-convex function.
(ii) If F is a locally Lipschitz function and $\partial^* F(\bar{x}) = \partial^c F(\bar{x})$ where $\partial^c F(\bar{x})$ is the Clarke generalized gradient then the above definition reduces to the definition of ∂^c-quasi-convex function defined by Bector, Chandra, and Dutta [4].

15.3 Bilevel Programming Problem

In this section we study the bilevel programming problem given as follows.

$$(BLPP) \quad \min_{x,y} F(x,y)$$
$$\text{subject to } \; G_j(x,y) \leq 0, j \in J, y \in \psi(x),$$

where for each $x \in \mathbb{R}^{n_1}$, $\psi(x)$ is the set of optimal solutions to the following optimization problem

$$\min_y f(x,y)$$
$$\text{subject to } g_i(x,y) \le 0, \qquad i \in I,$$

where $F, f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$, $G_j : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$, $j \in J = \{1,2,\ldots,m_2\}$ and $g_i : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$, $i \in I = \{1,2,\ldots,m_1\}$; n_i and m_i, $i = 1,2$ are integers with $n_i \ge 1$ and $m_i \ge 0$. $f(\cdot,\cdot)$ and $g_i(\cdot,\cdot)$, $i \in I$ are continuous convex and

$$\psi(x) = \underset{y}{\operatorname{argmin}}\{f(x,y) : g(x,y) \le 0\}.$$

So the idea is that the lower-level decision maker, or the follower minimizes his objective function based on the leader's choice x and returns the solution $y = y(x)$ to the leader who then uses it to minimize her objective function. If the optimal solution of the lower-level problem is uniquely determined for all $x \in \mathbb{R}^{n_1}$ then the problem BLPP is well defined. However, if there are multiple solutions to the lower-level problem for a given x, then the upper-level objective becomes a set-valued map. In order to overcome this difficulty, two different solution concepts have been considered in the literature, namely the optimistic solution and the pessimistic solution.

In this chapter we have focused on the optimistic approach only. According to this approach the leader assumes the cooperation of the follower in the sense that the follower will in any case take an optimal solution which is a best one from the leader's point of view. This leads to the following optimistic bilevel programming problem (OBLPP).

$$\text{(OBLPP)} \quad \min_x \varphi_0(x), \quad x \in \mathbb{R}^{n_1}$$
$$\text{where } \varphi_0(x) = \min_y \{F(x,y) : G_j(x,y) \le 0, j \in J, y \in \psi(x)\}$$
$$\text{and } \psi(x) \text{ is the set of optimal solutions to the lower level problem}$$
$$\min_y f(x,y)$$
$$\text{subject to } g_i(x,y) \le 0, \qquad i \in I.$$

[10] A point $\bar{x} \in \mathbb{R}^{n_1}$ is called a local optimistic solution of the bilevel programming problem if $\bar{y} \in \psi(\bar{x}), \bar{x} \in \mathbb{R}^{n_1}, F(\bar{x},\bar{y}) = \varphi_0(\bar{x})$, and there is a number $\varepsilon > 0$ such that $\varphi_0(x) \ge \varphi_0(\bar{x})$, for all $x \in \mathbb{R}^{n_1}, \| x - \bar{x} \| < \varepsilon$.

To obtain the necessary optimality conditions for the optimistic bilevel programming problem, we follow the value function approach initiated by Outrata [22] according to which the OBLPP can be converted into a single-level mathematical programming problem with the help of the value function of the lower-level problem given by

$$V(x) = \min_y \{f(x,y) : g_i(x,y) \le 0, i \in I, y \in \mathbb{R}^{n_2}\}.$$

Then the reformulated optimistic bilevel programming problem (ROBLPP) is given as

$$
\begin{aligned}
\text{(ROBLPP)}\quad & \min_{x,y} F(x,y) \\
& \text{subject to} \\
& f(x,y) - V(x) \le 0, \\
& g_i(x,y) \le 0, \qquad i \in I, \qquad G_j(x,y) \le 0, \qquad j \in J, \\
& (x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}.
\end{aligned}
$$

Let $X \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ denote the feasible set for ROBLPP; that is,

$$X = \{(x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} | f(x,y) - V(x) \le 0, g_i(x,y) \le 0, i \in I, G_j(x,y) \le 0, j \in J\}.$$

However, the price to pay in this reformulation is that ROBLPP is nonsmooth even for smooth initial data.

Remark 15.3 ([11]). Note that ROBLPP is globally equivalent to OBLPP, and local optimal solutions to OBLPP are always locally optimal to ROBLPP.

15.4 Variational Inequality Problem

In this section we consider the following types of variational inequality problems and study relationships between the solution of BLPP and associated VIP and their existence results.

We now consider the following Minty-type variational inequality problem MTVIP and Stampacchia-type variational inequality problem STVIP in terms of convexifactors and their recession cone. These can be regarded as nonsmooth versions of variational inequalities.

Let C be a nonempty convex subset of X.

(MTVIP) Find $(\bar{x},\bar{y}) \in C$ such that for all $(x,y) \in C$ there exists

$$
\begin{aligned}
&\xi \in \partial^* F(x,y) \cup (\partial^* F(x,y))_\infty \setminus \{0\} \text{ such that} \\
&\langle \xi, (x,y) - (\bar{x},\bar{y}) \rangle \ge 0.
\end{aligned}
$$

(STVIP) Find $(\bar{x},\bar{y}) \in C$ such that for all $(x,y) \in C$ there exists

$$
\begin{aligned}
&\xi \in \partial^* F(\bar{x},\bar{y}) \cup (\partial^* F(\bar{x},\bar{y}))_\infty \setminus \{0\} \text{ such that} \\
&\langle \xi, (x,y) - (\bar{x},\bar{y}) \rangle \ge 0.
\end{aligned}
$$

Theorem 15.2. *Let F admit a convexifactor $\partial^* F(x,y)$ at (x,y). If $(\bar{x},\bar{y}) \in C$ is an optimal solution of BLPP and F is ∂^*_∞-quasi-convex, then $(\bar{x},\bar{y})$ is a solution of MTVIP.*

Proof. Because $(\bar{x},\bar{y}) \in C$ is an optimal solution of BLPP, by Remark 15.1 it is an optimal solution of ROBLPP.

We have

$$F(x,y) \geq F(\bar{x},\bar{y}), \quad \text{for all } (x,y) \in C.$$

F is ∂^*_∞-quasi-convex, therefore we have

$$\langle \xi, (x,y) - (\bar{x},\bar{y}) \rangle \geq 0, \quad \text{for all } \xi \in \partial^* F(x,y) \cup (\partial^* F(x,y))_\infty \setminus \{0\},$$

which implies $(\bar{x},\bar{y}) \in C$ is a solution of MTVIP.

We now show that under the assumption of ∂^*_∞-pseudo-convexity on F the converse of the above theorem holds.

Theorem 15.3. *Let F admit a convexifactor $\partial^* F(x,y)$ at (x,y). If $(\bar{x},\bar{y}) \in C$ is a solution of MTVIP and F is ∂^*_∞-pseudo-convex, then $(\bar{x},\bar{y})$ is an optimal solution of BLPP.*

Proof. Let $(\bar{x},\bar{y}) \in C$ be a solution of MTVIP.

Then for all $(x,y) \in C$, there exists $\xi \in \partial^* F(x,y) \cup (\partial^* F(x,y))_\infty \setminus \{0\}$ such that $\langle \xi, (x,y) - (\bar{x},\bar{y}) \rangle \geq 0$.

Consider any sequence $\{\alpha_n\} \downarrow 0$ with $\alpha_n \in (0,1]$.

Because C is convex, $(x_n,y_n) = (\bar{x},\bar{y}) + \alpha_n((x,y) - (\bar{x},\bar{y})) \in C$.

As $(\bar{x},\bar{y}) \in C$ is a solution of MTVIP and $(x_n,y_n) \in C$, there exists $\xi_n \in \partial^* F(x_n,y_n) \cup (\partial^* F(x_n,y_n))_\infty \setminus \{0\}$ such that

$$\langle \xi_n, (x_n,y_n) - (\bar{x},\bar{y}) \rangle \geq 0 \tag{15.1}$$

Inasmuch as we have

$$\begin{aligned} 0 &= \langle \xi_n, ((x_n,y_n) - (x_n,y_n)) \rangle \\ &= \alpha_n \langle \xi_n, (x,y) - (x_n,y_n) \rangle + (1-\alpha_n) \langle \xi_n, (\bar{x},\bar{y}) - (x_n,y_n) \rangle \end{aligned}$$

(15.1) gives

$$\langle \xi_n, (x,y) - (x_n,y_n) \rangle \geq 0 \tag{15.2}$$

Case (i). If $\partial^* F(x_n,y_n)$ is bounded we assume that $\xi_n \to \xi$. Because $\partial^* F$ is closed, $\xi_n \in \partial^* F(x_n,y_n), \xi_n \to \xi$, and $(x_n,y_n) \to (\bar{x},\bar{y})$ as $n \to \infty$ we have $\xi \in \partial^* F(\bar{x},\bar{y})$.

Case (ii). If $\partial^* F(x_n,y_n)$ is unbounded we assume that $\lim_{n\to\infty} \|\xi_n\| = \infty$ and

$$\lim_{n\to\infty} \frac{\xi_n}{\|\xi_n\|} = \xi_0 \in (\partial^* F(\bar{x},\bar{y}))_\infty \setminus \{0\}$$

because as $n \to \infty, (x_n,y_n) \to (\bar{x},\bar{y})$.

Therefore for any $(x,y) \in C$, there exist $\xi \in \partial^* F(\bar{x},\bar{y}) \cup (\partial^* F(\bar{x},\bar{y}))_\infty \setminus \{0\}$ such that $\langle \xi, (x,y) - (\bar{x},\bar{y}) \rangle \geq 0$.

As F is ∂^*_∞-pseudo-convex we have

$$F(x,y) \geq F(\bar{x},\bar{y}), \text{ for all } (x,y) \in C.$$

Hence $(\bar{x},\bar{y})$ is an optimal solution of BLPP.

In the next theorem we use the mean value theorem to prove that an optimal solution of BLPP is a solution of MTVIP.

Theorem 15.4. *Let F be a continuous function that admits bounded, convex convexifactor $\partial^* F(x,y)$ at (x,y). If $(\bar{x},\bar{y}) \in C$ is an optimal solution of BLPP then $(\bar{x},\bar{y})$ is a solution of MTVIP.*

Proof. Suppose on the contrary that $(\bar{x},\bar{y}) \in C$ is not a solution of MTVIP; then for all $\xi \in \partial^* F(x,y), (x,y) \in C$, we have

$$\langle \xi, (x,y) - (\bar{x},\bar{y}) \rangle < 0. \tag{15.3}$$

By mean value theorem 15.1 there exist $z \in ((x,y),(\bar{x},\bar{y}))$ and $\xi' \in \partial^* F(z)$ such that

$$F(\bar{x},\bar{y}) - F(x,y) = \langle \xi', (\bar{x},\bar{y}) - (x,y) \rangle, \tag{15.4}$$

where $z = (x,y) + \lambda((\bar{x},\bar{y}) - (x,y)), \lambda \in (0,1)$.

Because (15.3) is true for all $(x,y) \in C$ and as C is convex it is true in particular for z, therefore by (15.4) we have

$$\langle \xi', (\bar{x},\bar{y}) - (x,y) \rangle > 0$$

and hence $F(\bar{x},\bar{y}) - F(x,y) > 0$ for all $(x,y) \in C$.

Which implies $(\bar{x},\bar{y})$ is not an optimal solution of BLPP, which is a contradiction.

Theorem 15.5. *Let F admit a convexifactor $\partial^* F(\bar{x},\bar{y})$ at $(\bar{x},\bar{y})$. If $(\bar{x},\bar{y}) \in C$ is an optimal solution of BLPP and F is ∂^*_∞-quasi-convex then $(\bar{x},\bar{y})$ is a solution of STVIP.*

Proof. Because $(\bar{x},\bar{y}) \in C$ is an optimal solution of BLPP, by Remark 15.1 it is an optimal solution of ROBLPP.

We have

$$F(x,y) \geq F(\bar{x},\bar{y}), \quad \text{for all } (x,y) \in C. \tag{15.5}$$

Let $(\bar{x}(\lambda),\bar{y}(\lambda)) = (\bar{x},\bar{y}) + \lambda((x,y) - (\bar{x},\bar{y})), \lambda \in [0,1]$. Then $(\bar{x}(\lambda),\bar{y}(\lambda)) \in C$ as C is convex.

Replacing (x,y) by $(\bar{x}(\lambda),\bar{y}(\lambda))$ in (15.5) we get

$$F(\bar{x}(\lambda),\bar{y}(\lambda)) \geq F(\bar{x},\bar{y}), \quad \text{for all } (\bar{x}(\lambda),\bar{y}(\lambda)) \in C.$$

Because F is ∂^*_∞-quasi-convex we have

$$\langle \xi_\lambda, (x,y) - (\bar{x},\bar{y}) \rangle \geq 0, \quad \text{for all } \xi_\lambda \in \partial^* F(\bar{x}(\lambda),\bar{y}(\lambda)).$$

Case (i). If $\partial^* F$ is locally bounded at $(\bar{x},\bar{y})$, there exist a neighbourhood of $(\bar{x},\bar{y})$ and a constant $k' > 0$ such that for each (x',y') in this neighbourhood and $\xi \in \partial^* F(x',y')$ we have $\|\xi\| \leq k'$.

As $(\bar{x}(\lambda),\bar{y}(\lambda)) \to (\bar{x},\bar{y})$ when $\lambda \to 0^+$, thus for $\lambda > 0$ small enough $\|\xi_\lambda\| \leq k'$. Without loss of generality we may assume that $\xi_\lambda \to \xi'$

Because $\partial^* F$ is closed, $\xi' \in \partial^* F(\bar{x},\bar{y})$, therefore for $(\bar{x},\bar{y}) \in C$, there exists $\xi' \in \partial^* F(\bar{x},\bar{y})$ such that $\langle \xi',(x,y)-(\bar{x},\bar{y})\rangle \geq 0$, for all $(x,y) \in C$.

Case (ii). If $\partial^* F$ is unbounded we assume that

$$\lim_{\lambda\to\infty} \|\xi_\lambda\| = \infty \text{ and } \lim_{\lambda\to\infty} \frac{\xi_\lambda}{\|\xi_\lambda\|} = \xi_0 \in (\partial^* F(\bar{x},\bar{y}))_\infty \setminus \{0\}$$

because as $\lambda \to 0^+, (\bar{x}(\lambda),\bar{y}(\lambda)) \to (\bar{x},\bar{y})$.

Therefore for any $(x,y) \in C$, there exist $\xi \in \partial^* F(\bar{x},\bar{y}) \cup (\partial^* F(\bar{x},\bar{y}))_\infty \setminus \{0\}$ such that $\langle \xi,(x,y)-(\bar{x},\bar{y})\rangle \geq 0$.

Hence $(\bar{x},\bar{y})$ is a solution of STVIP.

Theorem 15.6. *Let F admit a convexifactor $\partial^* F(\bar{x},\bar{y})$ at $(\bar{x},\bar{y})$. If $(\bar{x},\bar{y}) \in C$ is a solution of STVIP and F is ∂^*_∞-pseudo-convex then $(\bar{x},\bar{y})$ is an optimal solution of BLPP.*

Proof. Suppose on the contrary that $(\bar{x},\bar{y}) \in C$ is not an optimal solution of BLPP; then there exists $(x,y) \in C$ such that

$$F(x,y) < F(\bar{x},\bar{y}). \tag{15.6}$$

Because F is ∂^*_∞-pseudo-convex we have

$$\langle \xi,(x,y)-(\bar{x},\bar{y})\rangle < 0, \quad \text{for all } \xi \in \partial^* F(\bar{x},\bar{y}) \cup (\partial^* F(\bar{x},\bar{y}))_\infty \setminus \{0\}$$

which is in contradiction to the fact that $(\bar{x},\bar{y})$ is a solution of STVIP.

Theorem 15.7. *Let F be a continuous function that admits a bounded convex convexifactor $\partial^* F(\bar{x},\bar{y})$ at $(\bar{x},\bar{y})$. Suppose that F is ∂^*_∞-pseudo-convex. Then $(\bar{x},\bar{y}) \in C$ is a solution of STVIP if and only if it is a solution of MTVIP.*

Proof. Let $(\bar{x},\bar{y}) \in C$ be a solution of STVIP.

Then for all $(x,y) \in C$, there exists $\xi \in \partial^* F(\bar{x},\bar{y})$ such that

$$\langle \xi,(x,y)-(\bar{x},\bar{y})\rangle \geq 0. \tag{15.7}$$

Because F is ∂^*_∞-pseudo-convex we have

$$F(x,y) \geq F(\bar{x},\bar{y}). \tag{15.8}$$

Let $(x(\lambda),y(\lambda)) = (x,y) + \lambda((\bar{x},\bar{y})-(x,y)), \lambda \in [0,1]$. Then $(x(\lambda),y(\lambda)) \in C$ as C is convex.

Replacing (x,y) by $(x(\lambda),y(\lambda))$ in (15.8) we get

$$F(x(\lambda),y(\lambda)) \geq F(\bar{x},\bar{y}), \quad \text{for all } (x(\lambda),y(\lambda)) \in C. \tag{15.9}$$

By mean value theorem 15.1, there exist $z_\alpha \in ((x(\lambda),y(\lambda)),(\bar{x},\bar{y}))$ and $\xi_\alpha \in \partial^* F(z_\alpha)$ such that

$$\begin{aligned} F(\bar{x},\bar{y}) - F(x(\lambda),y(\lambda)) &= \langle \xi_\alpha, (\bar{x},\bar{y}) - (x(\lambda),y(\lambda))\rangle \\ &= (1-\lambda)\langle \xi_\alpha, (\bar{x},\bar{y}) - (x,y)\rangle, \end{aligned} \tag{15.10}$$

where $z_\alpha = (x,y) + \alpha((\bar{x},\bar{y}) - (x,y)), \alpha \in [0,1]$.

Using (15.9) and (15.10), there exist $z_\alpha \in ((x(\lambda),y(\lambda)),(\bar{x},\bar{y}))$, and $\xi_\alpha \in \partial^* F(z_\alpha)$ such that $\langle \xi_\alpha, (x,y) - (\bar{x},\bar{y})\rangle \geq 0$.

Because $\partial^* F$ is assumed to be bounded and $z_\alpha \to (x,y)$ when $\alpha \to 0^+$, for $\alpha > 0$ small enough $\|\xi_\alpha\| \leq k'$ for a constant $k' > 0$.

Without loss of generality we may assume that $\xi_\alpha \to \xi'$.

$\partial^* F$ is closed, $\xi' \in \partial^* F(x,y)$, therefore for $(\bar{x},\bar{y}) \in C$, there exists $\xi' \in \partial^* F(x,y)$ such that $\langle \xi', (x,y) - (\bar{x},\bar{y})\rangle \geq 0$, for all $(x,y) \in C$.

Hence $(\bar{x},\bar{y})$ is a solution of MTVIP.

Conversely suppose that $(\bar{x},\bar{y}) \in C$ is a solution of MTVIP.

Now proceeding on the similar lines of Theorem 15.3 we get that for any $(x,y) \in C$, there exists $\xi \in \partial^* F(\bar{x},\bar{y})$ such that

$$\langle \xi, (x,y) - (\bar{x},\bar{y})\rangle \geq 0.$$

Hence $(\bar{x},\bar{y})$ is a solution of STVIP.

Now we prove existence results for MTVIP and STVIP.

Theorem 15.8. *Let C be a nonempty, compact, and convex subset of X. Let $F : C \to \mathbb{R}$ admit a convexifactor $\partial^* F(x,y)$ at (x,y). Then, MTVIP has a solution.*

Proof. Let set-valued mapping $\Gamma = \widehat{\Gamma} : C \subset X \to 2^C$ be such that

$$\Gamma(x,y) = \left\{ \begin{array}{c} (\bar{x},\bar{y}) \in C : \exists \xi \in \partial^* F(x,y) \cup (\partial^* F(x,y))_\infty \setminus \{0\} : \\ \langle \xi, (x,y) - (\bar{x},\bar{y})\rangle \geq 0, \forall (x,y) \in C \end{array} \right\}.$$

We first prove that Γ is a KKM map.

Suppose that Γ is not a KKM map.

Then there exists a finite set $\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\} \subset C$, $t_i \geq 0$, $i = 1,2,\ldots n, \sum_{i=1}^n t_i = 1$ such that

$$(x,y) = \sum_{i=1}^n t_i(x_i,y_i) \notin \bigcup_{i=1}^n \Gamma(x_i,y_i).$$

Then there exists $\xi' \in \partial^* F(x,y) \cup (\partial^* F(x,y))_\infty \setminus \{0\}$ such that

$$\langle \xi', (x_i,y_i) - (x,y) \rangle < 0, \quad \text{for all } \ i = 1,2,\ldots,n.$$
$$\Rightarrow \sum_{i=1}^{n} t_i \langle \xi', (x_i,y_i) - (x,y) \rangle < 0$$

which is equivalent to

$$0 = \xi' \left\langle \sum_{i=1}^{n} t_i (x_i,y_i) - \sum_{i=1}^{n} t_i (x,y) \right\rangle$$

which is a contradiction.

Hence $\Gamma = \widehat{\Gamma}$ is a KKM map.

Next we show Γ is closed for every $(x,y) \in C$.

Let $\{(x_n,y_n)\}$ be a sequence in $\Gamma(x,y)$ such that $(x_n,y_n) \to (x_0,y_0) \in C$, to show $(x_0,y_0) \in \Gamma$.

$(x_n,y_n) \in \Gamma(x,y)$, therefore there exists $\xi_n \in \partial^* F(x,y) \cup (\partial^* F(x,y))_\infty \setminus \{0\}$ such that

$$\langle \xi_n, (x,y) - (x_n,y_n) \rangle \geq 0, \quad \text{for all } \ n. \tag{15.11}$$

Case (i). If $\partial^* F$ is bounded then we assume that $\xi_n \to \xi$. Because $\partial^* F$ is closed, $\xi_n \in \partial^* F(x_n,y_n), \xi_n \to \xi$, and $(x_n,y_n) \to (x_0,y_0)$ as $n \to \infty$ we have $\xi \in \partial^* F(x_0,y_0)$.

Then from (15.11) we get $\langle \xi, (x,y) - (x_0,y_0) \rangle \geq 0$ which implies $(x_0,y_0) \in \Gamma(x,y)$.

Case (ii). If $\partial^* F$ is unbounded we assume that

$$\lim_{n\to\infty} \|\xi_n\| = \infty \quad \text{and} \quad \lim_{n\to\infty} \frac{\xi_n}{\|\xi_n\|} = \xi_0 \in (\partial^* F(x_0,y_0))_\infty \setminus \{0\}.$$

Then proceeding as above we get $(x_0,y_0) \in \Gamma(x,y)$.

Hence $\Gamma(x,y)$ is closed.

Furthermore because C is bounded it follows that $\Gamma(x,y)$ is bounded for each $(x,y) \in C$. Therefore by Lemma 15.1, we have $\bigcap_{(x,y)\in C} \Gamma(x,y) \neq \phi$.

Hence any $(\bar{x},\bar{y}) \in C$ in this intersection is a solution of MTVLI.

Corollary 15.1. *In Theorem* 15.8*, if we assume* F *to be* ∂^*_∞*-pseudo-convex, then BLPP has an optimal solution.*

Proof. The proof follows by Theorem 15.3.

Theorem 15.9. *Let* C *be a nonempty, compact, and convex subset of* X*. Let* $F : C \to \mathbb{R}$ *admit a convexifactor* $\partial^* F(\bar{x},\bar{y})$ *at* $(\bar{x},\bar{y})$*. Then STVIP has a solution.*

Proof. Let set-valued mapping $\Gamma = \widehat{\Gamma} : C \subset X \to 2^C$ be such that

$$\Gamma(x,y) = \left\{ \begin{array}{c} (\bar{x},\bar{y}) \in C : \exists \xi \in \partial^* F(\bar{x},\bar{y}) \cup (\partial^* F(\bar{x},\bar{y}))_\infty \setminus \{0\} : \\ \langle \xi, (x,y) - (\bar{x},\bar{y}) \rangle \geq 0, \forall (x,y) \in C \end{array} \right\}.$$

Now proceeding on the lines of Theorem 15.8 we arrive at the desired result.

Corollary 15.2. *In Theorem* 15.9*, if we assume F to be* ∂^*_∞*-pseudo-convex, then BLPP has an optimal solution.*

Proof. The proof follows by Theorem 15.6.

References

1. Bard, J.F.: *Practical Bilevel Optimization, Algorithms and Applications*, Kluwer Academic, The Netherlands (1998)
2. Bard, J.F.: Optimality conditions for the bilevel programming problem, *Naval Res. Logistics Quart.* **31**, 13–26 (1984)
3. Bard, J.F.: Some properties of the bilevel programming problem, *J. Optim. Theory Appl.* **68**, 371–378 (1991)
4. Bector, C.R., Chandra, S., and Dutta, J.: *Principles of Optimization Theory*, Narosa (2005)
5. Chen, G.Y.: Existence of solutions for a vector variational inequality, an extension of Hartman-Stampacchia theorem, *J. Optim. Theory Appl.* **74**, 445–456 (1992)
6. Chinaie, M., Jabarootian, T., Rezaie, M., and Zafarani, J.: Minty's lemma and vector variational-like inequalities, *J. Global Optim.* **40**, 463–473 (2008).
7. Dempe, S.: Foundations of Bilevel Programming, Kluwer Academic, Boston (2002)
8. Dempe, S.: A necessary and sufficient optimality condition for bilevel programming problem, *Optimization* **25**, 341–354 (1992)
9. Dempe, S.: First order necessary optimality conditions for general bilevel programming problems, *J. Optim. Theory Appl.* **95**, 735–739 (1997)
10. Dempe, S., Dutta, J., and Mordukhovich, B.S.: New necessary optimality conditions in optimistic bilevel programming, *Optimization* **56**, 577–604 (2007)
11. Demyanov, V.F.: Convexification and concavification of positively homogeneous function by the same family of linear functions, Report 3,208,802 Universita di Pisa (1994)
12. Dutta, J., and Chandra, S.: Convexifactors, generalized convexity and optimality conditions, *J. Optim. Theory Appl.* **113**, 41–65 (2002)
13. Dutta, J., and Chandra, S.: Convexifactors, generalized convexity and vector optimization, *Optimization* **53**, 77–94 (2004)
14. Giannessi, F.: Theorems of the alternative, quadratic programs and complementarity problems, In: R. W. Cottle, F. Giannessi and J. L. Lions (Eds.), *Variational Inequalities and Complementarity Problems* (John Wiley and Sons: New York), 151–186 (1980)
15. Giannessi, F.: *On Minty Variational Principle, New Trends in Mathematical Programming*, Kluwer Academic, Dordrecht, The Netherlands (1997)
16. Giannessi, F.: On Minty Variational Principle. In: F. Giannessi, S. Komlosi and T. Rape-sak (Eds.), New Trends in Mathematical Programming (Kluwer Academic, Dordrecht, The Netherlands) 93–99 (1998)
17. Giannessi, F.: *Vector Variational Inequalities and Vector Equilibria* (Kluwer Academic, Dordrecht, The Netherlands) (2000).
18. Hartman, P., and Stampacchia, G.: On some nonlinear elliptic differential functional equations, *Acta Math.* **115**, 153–188 (1966)
19. Jeyakumar, V., and Luc, D.T.: Nonsmooth calculus, maximality and monotonicity of convexificators, *J. Optim. Theory Appl.* **101**, 599–621 (1999)
20. Li, X.F., and Zhang, J.Z.: Necessary optimality conditions in terms of convexificators in lipschitz optimization, *J. Optim. Theory Appl.* **131**, 429–452 (2006)
21. Mishra, S.K., and Wang, S.Y.: Vector variational-like inequalities and nonsmooth vector optimization problems, *Nonlinear Anal.* **64**, 1939–1945 (2005)
22. Outrata, J.V.: On necessary optimality conditions for stackelberg problems, *J. Optim. Theory Appl.* **76**, 306–320 (1993)

23. Rezaie, M., and Zafarani, J.: Vector optimization and variational-like inequalities, *J. Global Optim.* **43**, 47–66 (2009)
24. Yang, X.M., Yang, X.Q., and Teo, K.L., Some remarks on the minty vector variational inequality, *J. Optim. Theory Appl.* **121** (1), 193–201 (2004)
25. Ye, J.J., and Ye, X.Y.: Necessary optimality conditions for optimization problems with variational inequality constraints, *Math. Oper. Res.* **22**, 977–997 (1997)
26. Ye, J.J., and Zhu, D.L.: Optimality conditions for bilevel programming problems, *Optimization* **33**, 9–27 (1995)

Chapter 16
On Efficiency in Nondifferentiable Multiobjective Optimization Involving Pseudo d-Univex Functions; Duality

J. S. Rautela and Vinay Singh

Abstract In this chapter we introduce the concepts of KT-pseudo d-univex-I, KT-pseudo d-univex-II, and FJ-pseudo d-univex-II functions. The main objective of introducing these functions is to establish characterizations for efficient solutions to nondifferentiable multiobjective programming problems. Moreover, characterizations for efficient solutions by Fritz–John optimality conditions are also obtained. Furthermore, the Mond–Weir type dual problem is studied and weak, strong, and converse duality results are established involving the aforementioned class of functions.

16.1 Introduction

The search for solutions to mathematical programming problems has been carried out through the study of optimality conditions and of the properties of the functions that are involved, as well as through the study of dual problems. In the case of optimality conditions, it is customary to use critical points of the Kuhn–Tucker or Fritz–John [13] types. In the case of the kinds of functions employed in mathematical programming problems, to make the results more applicable in practice, the tendency has been to replace convex functions with more general ones, with the objective of obtaining a solution through an optimality condition.

With the introduction of the invex function, Hanson [10], Craven [4], and Craven and Glover [5] established the equivalence between a global minimum of a scalar function and a stationary point; and this also characterizes the invex functions

J. S. Rautela
Department of Mathematics, Faculty of Applied Sciences and Humanities, Echelon Institute of Technology, Faridabad 121 002, India, e-mail: sky_dreamz@rediffmail.com

Vinay Singh
Department of Mathematics, Faculty of Science, Banaras Hindu University, Varanasi 221 005, India, e-mail: vinaybhu1981@gmail.com

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4_16,

(see [5] and [3]). In 1992, Bector, Suneja, and Gupta [7] introduced a more general class known as univex functions. Unifying the approach of Antczak [1] and Bector, Suneja, and Gupta [7], Mishra, Wang, and Lai [15] introduced the concept of d-univexity.

Consider the following constrained multiobjective programming problem (CMP).

$$\begin{aligned}&\text{(CMP) Minimize } (f_1(x),\ldots,f_p(x))\\&\text{subject to: } g_j(x) \leqq 0,\\&x \in S \subseteq R^n,\end{aligned}$$

where $f_i : S \subseteq R^n \to R,\ i = 1,2,\ldots,p,\ g_j : S \subseteq R^n \to R, j = 1,2,\ldots,m$ are nondifferentiable functions on the open set $S \subseteq R^n$.

Definition 16.1. A feasible point, $\bar{x}$, is said to be an *efficient solution* of (CMP) if there does not exist another feasible point, x, such that $f(x) \leq f(\bar{x})$.

There appears a more general concept, namely weakly efficient solution of (CMP).

Definition 16.2. A feasible point $\bar{x}$ is said to be a *weakly efficient solution* of (CMP) if there does not exist another feasible point x, such that $f(x) < f(\bar{x})$.

Following the same lines as scalar problems, the outline is to obtain classes of functions that make up constrained multiobjective problems, such that any class of functions characterized by having every critical point as an efficient solution of (CMP) must be equivalent to these classes of functions. So it is a question of extending, amongst others, the kind of KT-invex functions introduced by Martin [12], as well as the results obtained by him. And in order to do that, we use Fritz–John and Kuhn–Tucker vector critical points as optimality conditions for multiobjective problems. In Section 16.2, we introduce new kinds of functions based on generalized invexity for multiobjective programming problems with constraints, in which joint properties of the objective function and the functions involved in the constraints of the problem are formed. Along similar lines, Osuna-Gomez, Beato-Moreno, and Rufian-Lizana [17] studied weakly efficient solutions for (CMP), and provided a new class of functions which extended the KT-invex functions [12] to the multiobjective case and proved that they are characterized by all Kuhn–Tucker vector critical points being weakly efficient solutions. In section 16.2, we introduce the concepts of KT-pseudo d-univex-I, KT-pseudo d-univex-II, and FJ-pseudo d-univex-II function. In Section 16.3 of this chapter, we extend this study to efficient points. Finally, in Section 16.4, we conclude by studying the duality of the (CMP) and the Mond–Weir-type dual problems.

16.2 Optimality Conditions. KT/FJ-Pseudo d-Univexity

Let $f : S \subseteq R^n \to R^p$ be a nondifferentiable function on open set S and $b_0 : S \times S \to R_+, \phi_0 : R \to R$ with $a < 0 \Rightarrow \phi_0(a) \leq 0$, $\eta : X \times X \to R^n$ is a vector-valued function. First we recall some known results.

Definition 16.3 ([15]). Let f be a nondifferentiable function on the open set S. Then the vector function f is said to be *d-univex with respect to* b_0, ϕ_0 and η at $u \in X$ if there exist b_0, ϕ_0, and η, such that $\forall x \in X$,

$$b_0(x,u)\phi_0[f(x) - f(u)] \geqq f'(u;\eta(x,u)),$$

where $f'(u;\eta(x,u))$ denotes the directional derivative of f in the direction $\eta(x,u)$,

$$f'(u;\eta(x,u)) = \lim_{\lambda \to 0^+} \frac{f(u+\lambda\eta(x,u)) - f(u)}{\lambda}.$$

The class of univex functions is defined by Bector, Suneja, and Gupta [7]. For further studies on the univex function, see Mishra and Giorgi [14]. Next we generalize the class of pseudo d-univex functions, to introduce d-univex-I, to distinguish it from a new class which we introduce and designate as pseudo d-univex-II.

Definition 16.4. Let f be a nondifferentiable function on the open set S. Then the vector function f is said to be *pseudo d-univex-I* with respect to b_0, ϕ_0, and η at $u \in X$ if there exist b_0, ϕ_0, and η, such that $\forall x \in X$,

$$b_0(x,u)\phi_0[f_i(x) - f_i(\bar{x})] < 0 \Rightarrow f_i'(\bar{x};\eta(x,\bar{x})) < 0.$$

Definition 16.5. Let f be a nondifferentiable function on the open set S. Then the vector function f is said to be *pseudo d-univex-II* with respect to b_0, ϕ_0, and η at $u \in X$ if there exist b_0, ϕ_0, and η, such that $\forall x \in X$,

$$b_0(x,u)\phi_0[f_i(x) - f_i(\bar{x})] \leq 0 \Rightarrow f_i'(\bar{x};\eta(x,\bar{x})) < 0.$$

Proposition 16.1. *Let f be a nondifferentiable function on the open set S. If the vector function f is pseudo d-univex-II, then the vector function f is pseudo d-univex-I.*

Definition 16.6. Problem (CMP) is said to be KT-pseudo d-univex-I with respect to b_0, ϕ_0, and η at $u \in X$ if there exist b_0, ϕ_0, and η, such that $\forall x_1, x_2 \in X$,

$$b_0(x_1,x_2)\phi_0[f_i(x_1) - f_i(x_2)] < 0 \Rightarrow \begin{cases} f_i'(x_2;\eta(x_1,x_2)) < 0 \\ g_j'(x_2;\eta(x_1,x_2)) \leqq 0, \forall j \in I(x_2), \end{cases}$$

where $I(x_2) = \{j = 1,\ldots,m : g_j(x_2) = 0\}$.

Definition 16.7. A feasible point $\bar{x}$ for (CMP) is said to be a Fritz–John vector critical point FJVCP, if there exist $\lambda \in R^p, \mu \in R^m$, such that

$$\lambda^T f'(\bar{x};\eta(x,\bar{x})) + \mu^T g'(\bar{x};\eta(x,\bar{x})) = 0 \quad (16.1)$$

$$\mu^T g(\bar{x}) = 0 \quad (16.2)$$

$$(\lambda,\mu) \geqq 0, \quad (\lambda,\mu) \neq 0. \quad (16.3)$$

This is equivalent to saying that there exists $(\lambda, \mu_I) \geq 0$ such that

$$\lambda^T f'(\overline{x}; \eta(x, \bar{x})) + \mu_I^T g_I'(\overline{x}; \eta(x, \bar{x})) = 0, \; I = I(\overline{x}) = \{j = 1, \ldots, m : g_j(\overline{x}) = 0\}. \tag{16.4}$$

The same occurs in the following definition, but in this case with $(\lambda, \mu_I) \geq 0$, $\lambda \neq 0$.

Definition 16.8. A feasible point $\overline{x}$ for (CMP) is said to be a Kuhn–Tucker vector critical point KTVCP, if there exist $\lambda \in R^p, \mu \in R^m$, such that

$$\lambda^T f'(\overline{x}; \eta(x, \bar{x})) + \mu^T g'(\overline{x}; \eta(x, \bar{x})) = 0 \tag{16.5}$$

$$\mu^T g(\overline{x}) = 0 \tag{16.6}$$

$$\mu \geqq 0 \tag{16.7}$$

$$\lambda \geq 0. \tag{16.8}$$

The following, Chankong and Haimes [6] and Kanniappan [11] type results are needed in the sequel.

Theorem 16.1. *If $\overline{x}$ is an efficient solution of (CMP), then $\overline{x}$ is a FJVCP.*

In a similar manner to Kanniappan [11] and Gulati and Talaat [9] type following the Kuhn–Tucker optimality result also needed for efficient solutions of (CMP), for which we need to take on a constraint qualification.

Theorem 16.2. *If $\overline{x}$ is an efficient solution of (CMP), and a constraint qualification is satisfied at $\overline{x}$, then $\overline{x}$ is KTVCP.*

The following Osuna, Gomez, Beato-Moreno, and Rufian-Lizana [17] type results are needed.

Theorem 16.3. *Every KTVCP is a weakly efficient solution of (CMP) if and only if problem (CMP) is KT-pseudo d-univex-I.*

For the study of efficient points of (CMP) from the condition of optimality of Kuhn–Tucker vector critical points, we need a new kind of function, one that is contained in the KT-pseudo d-univex-I class, and which we present below.

Definition 16.9. Problem (CMP) is said to be KT-pseudo d-univex-II with respect to b_0, ϕ_0, and η at $u \in X$ if there exist b_0, ϕ_0, and η, such that $\forall x_1, x_2 \in X$,

$$b_0(x_1, x_2)\phi_0[f(x_1) - f(x_2)] \leq 0 \Rightarrow \begin{cases} f'(x_2; \eta(x_1, x_2)) < 0 \\ g_j'(x_2; \eta(x_1, x_2)) \leqq 0, \; \forall j \in I(x_2), \end{cases}$$

where $I(x_2) = \{j = 1, \ldots, m : g_j(x_2) = 0\}$.

In the same way, for the study of efficient points from the Fritz–John optimality condition we need a new kind of function which we designate as FJ-pseudo d-invex-II.

Definition 16.10. Problem (CMP) is said to be *FJ-pseudo d-univex-II* with respect to b_0, ϕ_0, and η at $u \in X$ if there exist b_0, ϕ_0, and η, such that $\forall x_1, x_2 \in X$,

$$b_0(x_1,x_2)\phi_0[f(x_1) - f(x_2)] \leq 0 \Rightarrow \begin{cases} f(x_2;\eta(x_1,x_2)) < 0 \\ g'_j(x_2;\eta(x_1,x_2)) < 0, \ \forall j \in I(x_2), \end{cases}$$

where $I(x_2) = \{j = 1,\ldots,m : g_j(x_2) = 0\}$.

16.3 Characterization of Efficient Solutions

The Kuhn–Tucker optimality condition is necessary for a point to be an efficient solution for (CMP), as we have already seen. Let us also observe that under KT-pseudo d-univex-II, the Kuhn–Tucker optimality condition is sufficient for a point to be an efficient solution. But moreover, KT-pseudo d-univex-II is a necessary condition, as we demonstrate below.

Theorem 16.4. *Every KTVCP is an efficient solution of (CMP) if and only if (CMP) is KT-pseudo d-univex*-II.

Proof.

(i) Let $\overline{x}$ be a KTVCP and (CMP) KT-pseudo d-univex-II. We have to prove that $\overline{x}$ is an efficient solution of (CMP), and to do so let us suppose that it is not. Then there exists a feasible point x such that

$$f(x) - f(\overline{x}) \leq 0, \qquad \forall\, i = 1,\ldots,p.$$

Because $b_0 \geq 0$, $a < 0 \Rightarrow \phi_0(a) \leq 0$, so that above inequality yields

$$b_0(x,\overline{x})\phi_0[f(x) - f(\overline{x})] \leq 0.$$

Because (CMP) is KT-pseudo d-univex-II, there exist b_0, ϕ_0, and η such that

$$\begin{cases} f'(\overline{x};\eta(x,\overline{x})) < 0 \\ g'_I(\bar{x};\eta(x,\overline{x})) \leqq 0, \ I = I(\bar{x}). \end{cases} \tag{16.9}$$

On the other hand, $\overline{x}$ is a KTVCP; then $\exists(\bar{\lambda},\bar{\mu}) \geq 0$, $\bar{\lambda} \neq 0$ such that

$$\bar{\lambda}^T f'(\bar{x};\eta(x,\bar{x})) + \bar{\mu}_I^T g'_I(\overline{x};\eta(x,\bar{x})) = 0, \ I = I(\bar{x}). \tag{16.10}$$

$\bar{\lambda} \geq 0$, $\bar{\mu}_I \geqq 0$, therefore from (16.9), it follows that

$$\begin{cases} \bar{\lambda}^T f'(\overline{x};\eta(x,\overline{x})) < 0 \\ \bar{\mu}_I^T g'_I(\overline{x};\eta(x,\overline{x})) \leqq 0, \end{cases}$$

and then

$$\bar{\lambda}^T f'(\bar{x};\eta(x,\bar{x})) + \bar{\mu}_I^T g_I'(\bar{x};\eta(x,\bar{x})) < 0,$$

which is a contradiction to (16.10), and therefore, $\bar{x}$ is an efficient solution of (CMP).

(ii) Let us suppose that there exist two feasible points x and $\bar{x}$ such that

$$f(x) - f(\bar{x}) \leq 0,$$

inasmuch as otherwise (CMP) would be KT-pseudo d-univex-II, and the result would be proved. This means that $\bar{x}$ is not an efficient solution, and by using the initial hypothesis, $\bar{x}$ is not a KTVCP; that is,

$$\bar{\lambda}^T f'(\bar{x};\eta(x,\bar{x})) + \bar{\mu}_I^T g_I'(\bar{x};\eta(x,\bar{x})) = 0$$

has no solution $\bar{\lambda} \geq 0,\ \mu_I \geqq 0$. Therefore, by Motzkin's theorem, the system

$$f'(\bar{x};\eta(x,\bar{x})) < 0$$
$$g_I'(\bar{x};\eta(x,\bar{x})) \leqq 0,$$

has no solution $\eta(x,\bar{x}) \in R^n$. In consequence, (CMP) is the KT-pseudo d-univex-II. ■

We state the following theorem without proof, as its proof follows similar lines to those in the proof of Theorem 16.4

Theorem 16.5. *Every FJVCP is an efficient solution of (CMP) if and only if (CMP) is FJ-pseudo d-univex-II.*

16.4 Duality

Let us move on to the study of duality. In order to do so, we tackle duality between the multiobjective problem (CMP), and two associated problems of Mond–Weir [16] type, defined for the multiobjective case by Egudo and Hanson [8], but with the difference that the complementarity condition has been restricted and that of the nonnegativity of (λ,μ) has been extended. Note that $\lambda > 0$ is not necessary. Let us begin with the first problem DM1 as the dual of (CMP), and formulated thus:
(DM1)

$$\text{Maximize } f(u)$$

subject to:

$$\lambda^T f'(u;\eta(x,u)) + \mu^T g'(u;\eta(x,u)) = 0 \tag{16.11}$$

$$\mu_j g_j(u) = 0, \qquad j = 1,\ldots,m \tag{16.12}$$

$$\mu \geqq 0 \tag{16.13}$$

$$\lambda \geq 0 \tag{16.14}$$

$$u \in S \subseteq R^n.$$

We need a new class of pseudo d-univex function, which differs slightly from those already defined and which we present below.

Definition 16.11. The pair of functions (f,g) is said to be KT-pseudo d-univex-II with respect to b_0, ϕ_0, and η at $u \in X$ if there exist b_0, ϕ_0, and η, such that $\forall x_1, x_2 \in X$,

$$b_0(x_1,x_2)\phi_0[f(x_1)-f(x_2)] \leq 0 \Rightarrow \begin{cases} f'(x_2;\eta(x_1,x_2)) < 0 \\ g'_I(x_2;\eta(x_1,x_2)) \leqq 0 \end{cases}$$

with $I = I(x_2) = \{j = 1,\ldots,m : g_j(x_2) = 0\}$.

Theorem 16.6 (Weak Duality). *Let x be a feasible point for (CMP), and (u,λ,μ) a feasible point for DM1. If (f,g) is KT-pseudo d-univex-*II *on S, then $f(x) \leq f(u)$ is not verified.*

Proof. Let us suppose that (f,g) is KT-pseudo d-univex-II with respect to a vector function η. Let x be a feasible point for (CMP) and (u,λ,μ) a feasible point for DM1, such that $f(x) \leq f(u)$; if not, the result would be proved. Then, there exist $\lambda \in R^p, \mu \in R^m$ such that

$$\lambda^T f'(u;\eta(x,u)) + \mu^T g'(u;\eta(x,u)) = 0$$
$$\mu_j g_j(u) = 0, \quad j = 1,\ldots,m$$
$$\mu \geqq 0$$
$$\lambda \geq 0;$$

that is,

$$\lambda^T f'(u;\eta(x,u)) + \mu_I^T g'_I(u;\eta(x,u)) = 0 \tag{16.15}$$

with $(\lambda,\mu_I) \geq 0,\ \lambda \neq 0,\ I = I(u) = \{j = 1,\ldots,m : g_j(u) = 0\}$. Because $f(x) \leq f(u)$ from the KT-pseudo d-univex-II of (f,g) there exist b_0, ϕ_0, and η such that

$$b_0(x,u)\phi_0[f(x)-f(u)] \leq 0$$

it follows that

$$\begin{cases} f'(u;\eta(x,u)) < 0 \\ g'_I(u;\eta(x,u)) \leqq 0,\ I = I(\bar{x}) \end{cases}$$

and by multiplying by (λ,μ_I) we have

$$\lambda^T f'(u;\eta(x,u)) + \mu_I^T g'_I(u;\eta(x,u)) < 0,$$

which is a contradiction to (16.15), and therefore, $f(x) \leq f(u)$ is not verified. ∎

This weak duality result allows us to prove the strong duality, as follows.

Theorem 16.7 (Strong Duality). *Let (f, g) be KT-pseudo d-univex-II on S. If $\bar{x}$ is an efficient solution for (CMP) and a constraint qualification is satisfied at $\bar{x}$, then there exist λ, μ, such that $(\bar{x}, \lambda, \mu)$ is an efficient solution of DM1.*

Proof. Let us suppose that $\bar{x}$ is an efficient solution for (CMP). From Theorem 16.2, $\bar{x}$ is KTVCP; that is, $\exists(\lambda, \mu) \geq 0, \quad \lambda \neq 0$ such that

$$\lambda^T f'(\bar{x}; \eta(x, \bar{x})) + \mu^T g'(\bar{x}; \eta(x, \bar{x})) = 0$$
$$\mu^T g_j(\bar{x}) = 0.$$

Because $g(\bar{x}) \leqq 0,$ and $\mu^T g_j(\bar{x}) = 0$, it follows that

$$\mu_j g_j(\bar{x}) = 0, \qquad j = 1, \ldots, m.$$

Then, $\bar{x}$ is a feasible point for DM1, and the weak duality theorem $f_i(\bar{x}) \leq f_i(u)$ is not verified, where u is a feasible point for DM1. Therefore, $\bar{x}$ is an efficient solution of DM1. ∎

The converse result is also verified, as we demonstrate below.

Theorem 16.8 (Converse Duality). *Let (CMP) be KT-pseudo d-univex-II, and $\bar{x}$ a feasible point for (CMP). If $(\bar{x}, \lambda, \mu)$ is a feasible point for DM1, then $\bar{x}$ is an efficient solution of (CMP).*

Proof. Let us suppose that $\bar{x}$ is a feasible point for (CMP). If $(\bar{x}, \lambda, \mu)$ is a feasible point for DM1, then

$$\lambda^T f'(\bar{x}; \eta(x, \bar{x})) + \mu^T g'(\bar{x}; \eta(x, \bar{x})) = 0$$
$$\mu_j g_j(\bar{x}) = 0, \qquad j = 1, \ldots, m$$

and therefore $\bar{x}$ is a KTVCP. Because (CMP) is KT-pseudo d-univex-II, from Theorem 16.4 it follows that $\bar{x}$ is an efficient solution of (CMP). ∎

Similarly, we can obtain duality results for (CMP) and the following dual problem.

(DM2) Maximize $f(u)$

subject to:

$$\lambda^T f'(u; \eta(x, u)) + \mu^T g'(u; \eta(x, u)) = 0 \tag{16.16}$$
$$\mu_j g_j(u) = 0, \qquad j = 1, \ldots, m \tag{16.17}$$
$$(\lambda, \mu) \geq 0 \tag{16.18}$$
$$u \in S \subseteq R^n.$$

For this, we introduce the following definition.

Definition 16.12. The pair of the functions (f,g) is said to be *FJ-pseudo d-univex*-II *on S* with respect to b_0, ϕ_0, and η at $u \in X$ if there exist b_0, ϕ_0, and η, such that $\forall x_1, x_2 \in X$,

$$b_0(x_1,x_2)\phi_0[f(x_1) \leq f(x_2)] \Rightarrow \begin{cases} f'(x_2;\eta(x_1,x_2)) < 0 \\ g'(x_2;\eta(x_1,x_2)) < 0 \end{cases}$$

with $I(x_2) = \{j = 1,\ldots,m : g_j(x_2) = 0\}$.

The relationship between this class of function and the FJ-pseudo d-univexity-II of (CMP) is as follows.

Theorem 16.9 (Weak Duality). *Let x be a feasible point for (CMP), and* (u,λ,μ) *a feasible point for DM*2*. If* (f,g) *is FJ-pseudo d-univex*-II *on S then* $f(x) \leq f(u)$ *is not verified.*

Theorem 16.10 (Strong Duality). *Let* (f,g) *be FJ-pseudo d-univex*-II *on S. If* $\overline{x}$ *is an efficient solution for (CMP), then there exist* λ, μ *such that* $(\overline{x},\lambda,\mu)$ *is an efficient solution of DM*2*.*

Theorem 16.11 (Converse Duality). *Let (CMP) be FJ-pseudo d-univex*-II*, and* $\overline{x}$ *a feasible point for (CMP). If* $(\overline{x},\lambda,\mu)$ *is a feasible point for DM*2 *then* $\overline{x}$ *is an efficient solution of (CMP).*

References

1. T. Antczak, Multiobjective programming under d-invexity, *European*, *J. Oper. Res.* 137 (2002) 28–36.
2. M. Arana-Jimenez, A. Rufian-Lizana, R. Osuna-Gomez, G. Ruiz-Garzon, Pseudoinvexity, optimality conditions and efficiency in multiobjective problems; duality, *Nonlinear Anal.* 68 (2008) 24–34.
3. A. Ben-Israel, B. Mond, What is invexity?, *J. Aust. Math. Soc. Ser. B* 28 (1986) 1–9.
4. B. D. Craven, Invex functions and constraint local minima, *J. Aust. Math. Soc. Ser. B* 24 (1981) 357–366.
5. B. D. Craven, B. M. Glover, Invex functions and duality, *J. Aust. Math. Soc. Ser. A* 39 (1985) 1–20.
6. V. Chankong, Y. Y. Haimes, *Multiobjective Decision Making: Theory and Methodology*, North-Holland, New York, 1983.
7. C. R. Bector, S. K. Suneja, S. Gupta, Univex functions and univex nonlinear programming, in: *Proceedings of the Administrative Sciences Association of Canada*, 1992, pp. 115–124.
8. R. R. Egudo, M. A. Hanson, Multiobjective duality with invexity, *J. Math. Anal. Appl.* 126 (1987) 469–477.
9. T. R. Gulati, N. Talaat, Sufficiency and duality in nondifferentiable multiobjective programming, *Opsearch* 28 (2) (1991) 73–87.
10. M. A. Hanson, On sufficiency of Kuhn-Tucker conditions, *J. Math. Anal. Appl.* 80 (1981) 545–550.
11. P. Kanniappan, Necessary conditions for optimality of nondifferentiable convex multiobjective programming, *J. Optim. Theory Appl.* 40 (1983) 167–174.

12. D. M. Martin, The essence of invexity, *J. Optim. Theory Appl.* 47 (1) (1985) 65–76.
13. O. L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York, 1969.
14. S. K. Mishra, G. Giorgi, *Invexity and Optimization, Nonconvex Optimization and Its Applications*, Vol. 88, Springer-Verlag, Berlin, 2008.
15. S. K. Mishra, S. Y. Wang, K. K. Lai, Nondifferentiable multiobjective programming under generalized d-univexity, *European J. Oper. Res.* 160 (2005) 218–226.
16. B. Mond, T. Weir, Generalized concavity and duality, in: S. Schaible, W. T. Ziemba (Eds), *Generalized Concavity in Optimization and Economics*, Academic Press, New York, 1981, pp. 263–279.
17. R. Osuna-Gomez, A. Beato-Moreno, A. Rufian-Lizana, Generalized convexity in multiobjective programming, *J. Math. Anal. Appl.* 233 (1999) 205–220.

Index

S. K. Mishra, *Topics in Nonconvex Optimization,* Springer Optimization and Its Applications 50,
DOI 10.1007/978-1-4419-9640-4,