



HAL
open science

Scalarization and stability in multi-objective optimization

Moslem Zamani

► **To cite this version:**

Moslem Zamani. Scalarization and stability in multi-objective optimization. General Mathematics [math.GM]. Université d'Avignon; University of Teheran, 2016. English. NNT : 2016AVIG0414 . tel-01388148

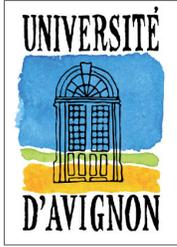
HAL Id: tel-01388148

<https://theses.hal.science/tel-01388148>

Submitted on 26 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITÉ D'AVIGNON
ET DES PAYS DE VAUCLUSE

UNIVERSITY OF AVIGNON AND VAUCLUSE
UNIVERSITY OF TEHRAN



THESIS

presented at University of Tehran and University of Avignon and Vaucluse
of the requirements for the degree of Doctor of Philosophy

SPECIALITY : Applied Mathematics-Optimization

Doctoral School (ED 536)

Laboratory of Mathematics of Avignon (EA 2151)

*Scalarization and stability in multi-objective
optimization*

by

Moslem Zamani

Submitted publicly on July 12th 2016 in front of a committee composed by :

Dinh The Luc (University of Avignon)	Thesis Supervisor
Majid Soleimani-damaneh (University of Tehran)	Thesis Supervisor
Giovanni Crespi (University of Insubria)	Reviewer
Masoud Zarepisheh (Memorial Sloan-Kettering Cancer Center)	Reviewer& Member
Mohammad B. Asadi (University of Tehran)	Member
Nezam Mahdavi-Amiri (Sharif University of Technology)	Member
Alberto Seeger (University of Avignon)	Member

July 2016

Abstract

In this thesis, three crucial questions arising in multi-objective optimization are investigated. First, the existence of properly efficient solutions via scalarization tools is studied. A basic theorem credited to Benson is extended from the convex case to the general case. Some further scalarization techniques are also discussed. The second part of the thesis is devoted to robustness. Various notions from the literature are briefly reviewed. Afterwards, a norm-based definition given by Georgiev, Luc and Pardalos is generalized to nonlinear multi-objective optimization. Necessary and sufficient conditions for robust solutions under appropriate assumptions are given. Relationships between new robustness notion and some known ones are highlighted. Two kinds of modifications in the objective functions are dealt with and relationships between the weak/proper/robust efficient solutions of the problems, before and after the perturbation, are established. Finally, we discuss the sensitivity analysis and stability in parametrized multi-objective optimization. Strict semi-differentiability of set-valued mappings of feasible sets and feasible values is proved under appropriate assumptions. Furthermore, some sufficient conditions for semi-differentiability of efficient sets and efficient values are presented. Finally, pseudo-Lipschitz continuity of aforementioned set-valued mappings is investigated.

Résumé

Cette thèse porte sur trois questions qui se posent en optimisation multi-objectif. Dans un premier temps, nous étudions l'existence de solutions efficaces via des techniques de scalarisation. On étend le théorème de Benson du cas convexe à un cas général. De plus, nous examinons d'autres techniques de scalarisation. Dans un second temps, nous abordons la question de robustesse. Nous examinons les concepts proposés dans la littérature sur le sujet. On étend au cas d'optimisation multi-objectif non-linéaire la définition de Georgiev et ses collaborateurs. Quelques conditions nécessaires et suffisantes pour obtenir une solution robuste moyennant des hypothèses appropriées sont données. Les relations entre cette notion de robustesse et certaines définitions mentionnées sont mises en évidence. Deux types de modifications des fonctions objectif sont traités et les relations entre les solutions faibles/propres/ robustes efficaces sont établies. Le dernier chapitre est consacré à l'analyse de sensibilité et de stabilité en optimisation multi-objectif paramétrée. On montre sous des conditions faibles que la multi-application de l'ensemble des solutions réalisables et des valeurs réalisables sont strictement semi-différentiables. On donne quelques conditions suffisantes pour la semi-différentiabilité de l'ensemble efficace et des valeurs efficaces. De plus, nous étudions la pseudo-Lipschitz continuité des multi-applications ci dessus citées.

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgments

I would like to express my deepest gratitude to my supervisors, Prof. Dinh The Luc and Prof. Majid Soleimani-damaneh, for their guidance and continuous support throughout my doctorate study. Their patience and helpful assistance helped me finish this thesis. Finally, I would like to thank my parents for understanding, love and support.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

Abstract	iii
Résumé	v
Preface	3
Originality	7
1 Preliminaries	9
1.1 Multi-objective optimization	9
1.2 Tangent cones and normal cones	11
1.3 Set-valued mappings	16
1.3.1 Continuity	16
1.3.2 Differentiability	19
1.4 Clarke subdifferential	23
2 Scalarization: Extending two important theorems	29
2.1 Introduction	29
2.2 Benson’s method	30
2.3 Direction-based scalarization	36
2.4 Compromise Programming	37
3 Robustness	43
3.1 Introduction	43
3.1.1 Stochastic multi-objective optimization	44

3.1.2	Robustness: an introduction	45
3.2	Robustness	51
3.3	Robustness radius	63
3.4	Comparison with other notions	64
3.5	Modification of the objective function	69
4	Stability and sensitivity analysis	77
4.1	Introduction	77
4.2	Parametric multi-objective optimization	79
4.3	Pseudo-Lipschitz continuity and semi-differentiability of the efficient set and the Pareto frontier	90
	Bibliography	99

List of Figures

1-1	A set related to Example 1.1	14
1-2	Tangent cones in Example 1.1	14
1-3	Upper and lower semicontinuity of set-valued mappings	17
3-1	Image of f in Example 3.1	47

THIS PAGE INTENTIONALLY LEFT BLANK

Preface

Decision making is as old as mankind on earth and everyone somehow faces this issue everyday. In this problem, Decision Maker (DM) has some alternatives among which he/she wishes to choose one with the most beneficial achievement. In most cases, DM is faced with conflicting objectives that make the choice of best alternative complicated. To clarify this let us give an example. Consider a person who is going to buy a car and their criteria are fuel consumption and power. It goes without saying that a car with high power burns more gas. There is no single alternative that simultaneously satisfies the two criteria. How can DM make a choice?

As in a decision making problem, one is supposed to select merely one alternative; he/she has to trade-off among objectives and opts for an alternative. However, it would be profitable to distinguish alternatives (feasible set) which can be candidates. Edgeworth¹ is among the first to address this issue. He considered multicriteria economic decision making and regarded an alternative as an eligible candidate if there exists no feasible solution for which an improvement in one objective does not lead to deteriorating at least one of the remaining objectives. Vilfredo Pareto² in his famous book³ studied this problem and called these candidates as efficient points.

Nowadays, in the literature, decision making problems are classified and treated according to the nature of the feasible set. If the feasible set is finite and given explicitly in the beginning of the solution process, it is handled by multicriteria decision analysis (MCDA); otherwise, it is treated by multi-objective optimization. However, both paradigms could be applied to make a successful decision. MCDA is out of scope

¹Philosopher and political economist, 1845- 1926

²Engineer, sociologist, economist, political scientist, and philosopher, 1848- 1923

³Manuale di Economia Politica, Societa Editrice Libreria, Milano, Italy, 1906

of our work here and we refer the reader to the book [27] where special techniques like AHP and TOPSIS are introduced to handle these problems.

Multi-objective optimization is a part of mathematical programming dealing with the minimization of some conflicting objective functions over a feasible set. Multi-objective optimization problems occur in many fields including economics, management, multicriteria design optimization, water resource planning, medicine, etc. Up to now a great deal of work has been devoted to multi-objective optimization, from both theoretical and numerical aspects.

Numerous approaches have been proposed to solve multi-objective optimization problems including scalarization techniques, interactive methods and so on. Scalarization is a traditional approach to solve multi-objective optimization problems. By scalarization methods, one solves a single objective optimization problem corresponding to a given multi-objective optimization problem whose optimal solutions can be efficient. Strictly speaking, a single objective optimization problem is formulated related to the multi-objective optimization via a real-valued scalarizing function, typically being a function of the objective functions, auxiliary scalar or vector variables, and/or scalar or vector parameters. In addition, the feasible set of the new problem may be restricted by new constraint functions associated with the objective functions of the multi-objective optimization problem and/or new variables are introduced [24, 26].

Scalarization methods have been employed like an engine in interactive methods and approaches for estimating efficient sets. Moreover, heuristic methods such as genetic algorithm, ant colony method, etc., have been widely developed for solving multi-objective optimization problems in recent decades [18].

One of the solution concepts which plays an important role in multi-objective optimization, from both theoretical and practical points of view, is the proper efficiency notion. This concept was first introduced by Geoffrion [31] to eliminate the efficient solutions which cannot be characterized by weighted-sum method and those points turn out to have unbounded trade-offs. However, similar concepts were also proposed before and after Geoffrion. Kuhn and Tucker [48] introduced a notion for proper effi-

ciency in differentiable multi-objective optimization which is equivalent to Geoffrion's definition under convexity. Benson [9] and Henig [42] proposed two different notions for proper efficiency in general vector optimization which are equivalent to Geoffrion's definition in multi-objective optimization.

Some scalarization methods are able to find properly efficient solutions and to prove their existence. Geoffrion [31] established that any optimal solution of the weighted sum method with a strictly positive combination of objective functions is a properly efficient solution. Benson [8] introduced a new scalarization method, usually referred to as Benson's method. It is well-known under convexity that if Benson's method corresponding to a given multi-objective optimization problem is unbounded, then there exists no properly efficient solution [8]. Moreover, Kaliszewski [46] introduced a scalarization method that is able to produce efficient solutions with a given bounded trade-off.

An important factor involved with multi-objective optimization problems is uncertainty. The uncertainty may arise from estimation of parameters in the model, error of machine, structure of problem and so on. Indeed, some parameters are often unknown at the beginning of solving a multi-objective optimization problem.

Three common approaches addressing this subject in the literature are stochastic optimization, robust optimization and stability/sensitivity analysis. Stochastic optimization and robust optimization incorporate uncertainty in the model. Nevertheless, in stability and sensitivity analysis methods first one solves the problem regardless of uncertainty, and then analyzes continuity and differentiability of the optimal value or the optimal set with respect to an uncertain set.

This thesis is organized as follows. In Chapter 1, some preliminaries on multi-objective optimization and set-valued mappings are given. Chapter 2 is devoted to scalarization methods. The purpose of this chapter is to check the existence of properly efficient solutions by scalarization techniques. In Chapter 3, we study the concept of robustness in multi-objective optimization. We extend Georgiev et al.'s definition [32] to nonlinear case and provide some theoretical results concerning this notion. Chapter 4 is dedicated to stability and sensitivity analysis in multi-objective

optimization. We investigate pseudo-Lipschitz continuity and semi-differentiability of set-valued mappings corresponding to the efficient sets and efficient values of parametric multi-objective optimization problems.

This is author's quick glance at the subject of the thesis. More detailed introduction is provided at the beginning of each chapter containing more information about the topic treated there. More history about multi-objective optimization can be found in [74]. In addition, [75] provides an overview of some works on this subject.

The results of this thesis are essentially published in the following works:

1. Soleimani-damaneh, M., and Zamani, M. On Benson's scalarization in multiobjective optimization. *Optimization Letters*, in press.
2. Zamani, M., Soleimani-damaneh, M., and Kabgani, A. Robustness in nonsmooth nonlinear multi-objective programming. *European Journal of Operational Research* 247, 2 (2015), 370-378.
3. Luc, D. T., Soleimani-damaneh, M., and Zamani, M. Stability and sensitivity analysis in multi-objective optimization, in preparation.

Originality

I certify that this work contains no material which has been accepted for any degree or diploma in my name, in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person, except where due reference has been made in the text.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 1

Preliminaries

This chapter contains some background materials on set-valued mappings and multi-objective optimization. Moreover, some preliminary results are provided which will be used later. We also provide notations that will be used throughout the thesis. In most cases, we follow the notations from the books [17, 24, 47] according to the subject.

In the present work, we consider problems in finite dimensional Euclidean spaces only. The notation $\|\cdot\|$ stands for Euclidean norm and both $\langle a, b \rangle$ and $a^T b$ denote inner product of $a, b \in \mathbf{R}^n$. For a set $X \subseteq \mathbf{R}^n$, we use the notations $int(X)$, $cl(X)$, $co(X)$ and $pos(X)$ for the interior, the closure, the convex hull and the convex conic hull of X , respectively. We use the notation $x \xrightarrow{X} \bar{x}$ to show that x tends to \bar{x} while $x \in X$. Moreover, the notation $t \downarrow 0$ means that t goes to 0 from above. In matrix spaces, the notation $\|\cdot\|$ stands for Frobenius norm.

1.1 Multi-objective optimization

We consider the following general multi-objective optimization problem (MOP):

$$\begin{aligned} \min & f(x) \\ \text{s.t.} & x \in X. \end{aligned}$$

The set of feasible solutions of this problem is a closed set $X \subseteq \mathbf{R}^n$ and $f : X \rightarrow \mathbf{R}^p$ is a continuous vector-valued function.

The solution notion for (MOP) is defined with respect to an ordering cone which is used for ordering the criterion space \mathbf{R}^p . We use the natural ordering cone defined by

$$\mathbf{R}_{\geq}^p = \{x \in \mathbf{R}^p : x_j \geq 0, j = 1, 2, \dots, p\}.$$

Utilizing this ordering cone, a feasible solution $\bar{x} \in X$ is called an efficient solution of (MOP) if

$$\left(f(\bar{x}) - \mathbf{R}_{\geq}^p\right) \cap f(X) = \{f(\bar{x})\}.$$

Likewise, we say $\bar{x} \in X$ is a weakly efficient solution of (MOP) if

$$\left(f(\bar{x}) - \mathbf{R}_{>}^p\right) \cap f(X) = \emptyset,$$

where $\mathbf{R}_{>}^p = \{x \in \mathbf{R}^p : x_j > 0, j = 1, 2, \dots, p\}$. We will use the symbol \mathbf{R}_{\geq}^p to denote the set $\mathbf{R}_{\geq}^p \setminus \{0\}$. Throughout the thesis, the notations \leq , \leq and $<$ stand for the following orders on \mathbf{R}^p :

$$x \leq y \iff y - x \in \mathbf{R}_{\geq}^p,$$

$$x \leq y \iff y - x \in \mathbf{R}_{\geq}^p,$$

$$x < y \iff y - x \in \mathbf{R}_{>}^p.$$

All above-mentioned orders are transitive. Nevertheless, only the order \leq is reflexive and antisymmetric which induces a partial order on \mathbf{R}^p .

A solution concept which plays an important role in multi-objective optimization, from both theoretical and practical points of view, is proper efficiency. This notion was introduced to eliminate the points which cannot be characterized by weighted-sum method and those points turning out to have unbounded trade-offs [31]. There are different definitions for proper efficiency in the literature; see [40, 65]. We use the following ones.

Definition 1.1. A feasible solution $\bar{x} \in X$ is called properly efficient in the Geoffrion's sense, if it is efficient and there is a real number $M > 0$ such that for all $i \in \{1, 2, \dots, p\}$ and $x \in X$ satisfying $f_i(x) < f_i(\bar{x})$ there exists an index $j \in \{1, 2, \dots, p\}$ such that $f_j(x) > f_j(\bar{x})$ and

$$\frac{f_i(\bar{x}) - f_i(x)}{f_j(x) - f_j(\bar{x})} \leq M.$$

Definition 1.2. A feasible solution $\bar{x} \in X$ is called properly efficient in the Benson's sense, if

$$cl\left(\text{pos}(f(X) + \mathbf{R}_{\geq}^p - f(\bar{x}))\right) \cap (-\mathbf{R}_{\geq}^p) = \{0\}.$$

Definition 1.3. A feasible solution $\bar{x} \in X$ is called properly efficient in the sense of Henig if $(f(\bar{x}) - C) \cap f(X) = \{f(\bar{x})\}$, for some convex pointed closed cone C satisfying $\mathbf{R}_{\geq}^p \setminus \{0\} \subseteq \text{int}(C)$.

These three definitions have been proven to be equivalent for the natural cone; see [65]. Hereafter, the set of efficient solutions, the set of weakly efficient and the set of properly efficient solutions are denoted by X_E , X_{WE} and X_{PE} , respectively. The set X_{WE} is closed but X_E and X_{PE} are not necessarily closed. Also, setting $Y = f(X)$, the set of nondominated points, denoted by Y_N , is defined by $Y_N = f(X_E)$, and the set of properly nondominated points, denoted by Y_{PN} , is defined by $Y_{PN} = f(X_{PE})$. Similarly, the set of weakly nondominated points, denoted by Y_{WN} , is defined by $Y_{WN} = f(X_{WE})$.

The set Y is called \mathbf{R}_{\geq}^p -closed, if $Y + \mathbf{R}_{\geq}^p$ is closed. Similarly, Y is called \mathbf{R}_{\geq}^p -convex, if $Y + \mathbf{R}_{\geq}^p$ is convex. If X is a convex set and f_1, f_2, \dots, f_p are convex functions, then Y is \mathbf{R}_{\geq}^p -convex.

1.2 Tangent cones and normal cones

We start this section with a definition of set convergence. The proof of forthcoming statements in this section can be found in [17, 47, 59].

Definition 1.4. Let $\{C_t\}_{t \in T}$ be a family of subsets of \mathbf{R}^n , where $T \subseteq \mathbf{R}^q$. Suppose that $\bar{t} \in cl(T)$. The upper limit and lower limit, denoted by $\limsup_{t \rightarrow \bar{t}} C_t$ and

$\liminf_{t \rightarrow \bar{t}} C_t$ respectively, are defined as follows:

$$\begin{aligned}\limsup_{t \rightarrow \bar{t}} C_t &= \{x : \exists \{t_n\} \subseteq T, \exists \{x_n\}; x_n \in C_{t_n}, t_n \rightarrow \bar{t}, x_n \rightarrow x\}, \\ \liminf_{t \rightarrow \bar{t}} C_t &= \{x : \forall \{t_n\} \subseteq T, t_n \rightarrow \bar{t}, \exists \{x_n\}; x_n \in C_{t_n}, x_n \rightarrow x\}.\end{aligned}$$

These limits are known as Kuratowski-Painlevé upper and lower limits, respectively. Both limits are closed sets. It is easy to see that the lower limit is a subset of the upper limit.

The main idea behind the definition of a tangent cone is to obtain a set involving all directions which one can move along with in the set. Some important applications of tangent cone in optimization theory is its vital role in deriving optimality conditions and specifying admissible directions in numerical methods. Elements of tangent cone are known as variations in calculus of variation. Tangent cone plays a role similar to tangent space in differential geometry. As in optimization, encountering nonsmoothness even for differentiable problems, the notion of tangent space is not practicable. There are various definitions of tangent cone in the literature . Here, we just mention some of them. We refer the reader to [34] and the comprehensive book [47] for more detailed information about tangent cones.

Definition 1.5. *The Bouligand tangent cone to X at $\bar{x} \in X$, denoted by $T_X(\bar{x})$, is defined as*

$$T_X(\bar{x}) := \limsup_{t \downarrow 0} \frac{X - \bar{x}}{t}.$$

In other words, $d \in T_X(\bar{x})$, if there are sequences $\{x_n\} \subseteq X$ and $t_n \downarrow 0$ such that $t_n^{-1}(x_n - \bar{x}) \rightarrow d$, or equivalently, if there are sequences $d_n \rightarrow d$ and $t_n \downarrow 0$ with $\{\bar{x} + t_n d_n\} \subseteq X$.

The Bouligand tangent cone is a nonempty closed cone, but not necessarily convex. When X is a differentiable manifold, the Bouligand tangent cone at a given point coincides with the tangent space. Moreover, the Bouligand tangent cone is closure of admissible directions as X is convex.

Definition 1.6. *Let $\bar{x} \in X$. The adjacent cone to X at $\bar{x} \in X$, denoted by $T_X^a(\bar{x})$, is*

defined by

$$T_X^a(\bar{x}) := \liminf_{t \downarrow 0} \frac{X - \bar{x}}{t}.$$

In other words, $d \in T_X^a(\bar{x})$, if for every sequence $t_n \downarrow 0$, there is a sequence $d_n \rightarrow d$ with $\{\bar{x} + t_n d_n\} \subseteq X$. Put differently, $d \in T_X^a(\bar{x})$ if and only if there is a function $\phi : [0, \epsilon) \rightarrow X$ such that $\phi(0) = \bar{x}$ and $\phi'_+(0) = d$.

Adjacent cone is nonempty, closed and a subset of the Bouligand tangent cone.

Definition 1.7. The Clarke tangent cone to X at $\bar{x} \in X$ is defined by

$$T_X^c(\bar{x}) := \liminf_{t \downarrow 0, x \xrightarrow{X} \bar{x}} \frac{X - x}{t}.$$

In other words, $d \in T_X^c(\bar{x})$ if and only if for each sequence $x \xrightarrow{X} \bar{x}$ along with a sequence $t_n \downarrow 0$ there is a sequence $d_n \rightarrow d$ such that $\{x_n + t_n d_n\} \subseteq X$.

the Clarke tangent cone is closed and convex. Moreover, it is a subset of adjacent cone. In Section 1.4, we will give an equivalent definition for the Clarke tangent cone. The following example demonstrates that inclusion relation among tangent cones may be strict.

Example 1.1. Let $X = cl(\{(x_1, x_2) : x_2 = \sin(x_1^{-1}), x_1 > 0\}) \cup pos\left(\begin{bmatrix} -2 \\ -1 \end{bmatrix}\right)$ be given.

At the origin, we have

$$T_X(0) = \{(d_1, d_2) : d_1 \geq 0\} \cup pos\left(\begin{bmatrix} -2 \\ -1 \end{bmatrix}\right),$$

$$T_X^a(0) = pos\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}\right) \cup pos\left(\begin{bmatrix} -2 \\ -1 \end{bmatrix}\right),$$

$$T_X^c(0) = \{0\}.$$

Figures 1-1 and 1-2 depict X and corresponding tangent cones, respectively.

As we can see from the above example, it is generally impossible to compute the Clarke tangent cone if the Bouligand tangent or the adjacent cone is known at a

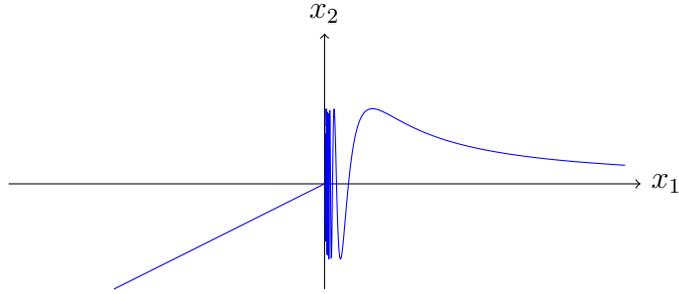
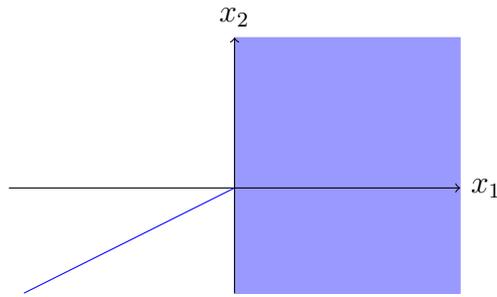
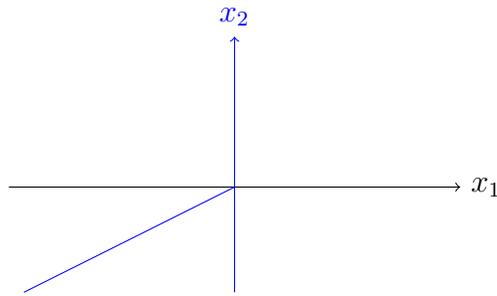


Figure 1-1 – A set related to Example 1.1



(a) $T_X(0)$



(b) $T_X^c(0)$

Figure 1-2 – Tangent cones in Example 1.1

given point. For instance, tangent cone may be the whole space while the Clarke tangent cone is equal to zero. However, the following interesting theorem addresses a connection between the Bouligand tangent cone around a point and Clarke tangent cone at that point.

Theorem 1.1. *Let $\bar{x} \in X$ be a given point and let X be locally closed around it. Then,*

$$T_X^c(\bar{x}) = \liminf_{x \xrightarrow{X} \bar{x}} T_X(x).$$

Let K be a cone. The polar cone of K is defined and denoted by $K^\circ := \{\nu :$

$\langle \nu, d \rangle \leq 0, \forall d \in K$. The polar cone of a subspace equals the orthogonal subspace. In the rest of this section, we pay attention to the concept of normal cones. Normal cone is a counterpart of normal bundle in differential manifolds.

Definition 1.8. *Let $\bar{x} \in X$ be given and let X be locally closed around it. The Fréchet normal cone to X at \bar{x} , denoted by $\hat{N}_X(\bar{x})$, is defined as*

$$\hat{N}_X(\bar{x}) = \{\nu \in \mathbf{R}^n : \limsup_{x \xrightarrow{X} \bar{x}} \frac{\langle \nu, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0\}.$$

In other words, $\nu \in \hat{N}_X(\bar{x})$ if and only if

$$\langle \nu, x - \bar{x} \rangle \leq o(\|x - \bar{x}\|), \quad \forall x \in X,$$

where $o(\|x - \bar{x}\|)/\|x - \bar{x}\| \rightarrow 0$, when $x \xrightarrow{X} \bar{x}$. Moreover, the limiting normal cone to X at \bar{x} , written as $N_X(\bar{x})$, is defined by

$$N_X(\bar{x}) := \limsup_{x \xrightarrow{X} \bar{x}} \hat{N}_X(x).$$

The two above-mentioned cones are both closed, but, in general, only the Fréchet normal cone is convex. Moreover, the Fréchet normal cone is the polar cone of the Bouligand tangent cone. The closure of convex hull of limiting normal cone is called the Clarke normal cone and is denoted by $N_X^c(\bar{x})$. Clarke normal cone is the polar cone of Clarke tangent cone.

Theorem 1.2. *Let $\hat{N}_\epsilon(\bar{x}) := \{\nu : \limsup_{x \xrightarrow{X} \bar{x}} \frac{\langle \nu, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq \epsilon\}$. If X is locally closed around \bar{x} , then we have*

$$N_X(\bar{x}) := \limsup_{x \xrightarrow{X, \epsilon \downarrow 0} \bar{x}} \hat{N}_X(x). \quad (1.1)$$

The preceding theorem states that the Fréchet normal cone can be replaced by $\hat{N}_\epsilon(\bar{x})$ in the definition of limiting normal cone. Mordukhovich [59] introduced this set to define his normal cone by the formula (1.1). However, in finite-dimensional

spaces, Mordukhovich's normal cone coincides with the limiting normal cone; see [59] for more details.

Definition 1.9. *Let X be locally closed around \bar{x} . X is said to be regular at \bar{x} if $\hat{N}_X(\bar{x}) = N_X(\bar{x})$.*

It is easily seen that X is regular at \bar{x} , if and only if $T_X(\bar{x}) = T_X^c(\bar{x})$.

1.3 Set-valued mappings

Let $\Gamma : X \rightrightarrows \mathbf{R}^p$ be a set-valued mapping, i.e., Γ sends each $x \in X$ to a subset of \mathbf{R}^p . We call the set $\text{gph}\Gamma := \{(x, y) : x \in X, y \in \Gamma(x)\}$ as graph of Γ . Moreover, domain of Γ is defined as $\text{dom}(\Gamma) := \{x \in X : \Gamma(x) \neq \emptyset\}$.

In this section, we review some notions about continuity and differentiability of set-valued mappings. The reader can find the proof of the forthcoming statements given in this section in [47]. Unfortunately, there are no unified terminology in this area.

1.3.1 Continuity

Definition 1.10. *The set-valued mapping Γ is said to be closed at $\bar{x} \in X$, if for each convergent sequence $\{(x_n, y_n)\} \subseteq \text{gph}\Gamma$ with $(x_n, y_n) \rightarrow (\bar{x}, \bar{y})$, we have $\bar{y} \in \Gamma(\bar{x})$. In other words, $\limsup_{x \rightarrow \bar{x}} \Gamma(x) \subseteq \Gamma(\bar{x})$.*

It is clear from the definition that if Γ is closed at \bar{x} , then the set $\Gamma(\bar{x})$ is also closed.

Definition 1.11. *We say that Γ is lower semicontinuous at $\bar{x} \in X$, if $\liminf_{x \rightarrow \bar{x}} \Gamma(x) \supseteq \Gamma(\bar{x})$, i.e., for each open set U in \mathbf{R}^p with $\Gamma(\bar{x}) \cap U \neq \emptyset$, there exists a neighborhood O of \bar{x} such that*

$$\Gamma(x) \cap U \neq \emptyset, \quad \forall x \in O.$$

In addition, Γ is upper semicontinuous at $\bar{x} \in X$, if for each open set U containing $\Gamma(\bar{x})$, there exist a neighborhood O of \bar{x} with

$$\Gamma(x) \subseteq U, \quad \forall x \in O.$$

We say that Γ is continuous at \bar{x} , if it is both upper semicontinuous and lower semicontinuous at the point. It is easily seen that upper semicontinuity of Γ at \bar{x} and closedness of $\Gamma(\bar{x})$ imply closedness of Γ at the point. However, the converse is not necessarily true; for instance, consider $\Gamma : \mathbf{R} \rightrightarrows \mathbf{R}$ given by

$$\Gamma(x) = \begin{cases} \{0, x^{-1}\}, & x \neq 0, \\ 0, & x = 0 \end{cases}$$

at $\bar{x} = 0$. However, if for some neighborhood O of \bar{x} and each $x \in O$, $\Gamma(x)$ is compact and if Γ is closed at \bar{x} , then it is upper semicontinuous at the point; see [54].

Lower semicontinuity and upper semicontinuity are equivalent to continuity when Γ is single-valued. Of course they are distinct when being set-valued; see Figure 1-3 depicting two set-valued mappings from \mathbf{R} to the power set of \mathbf{R} .

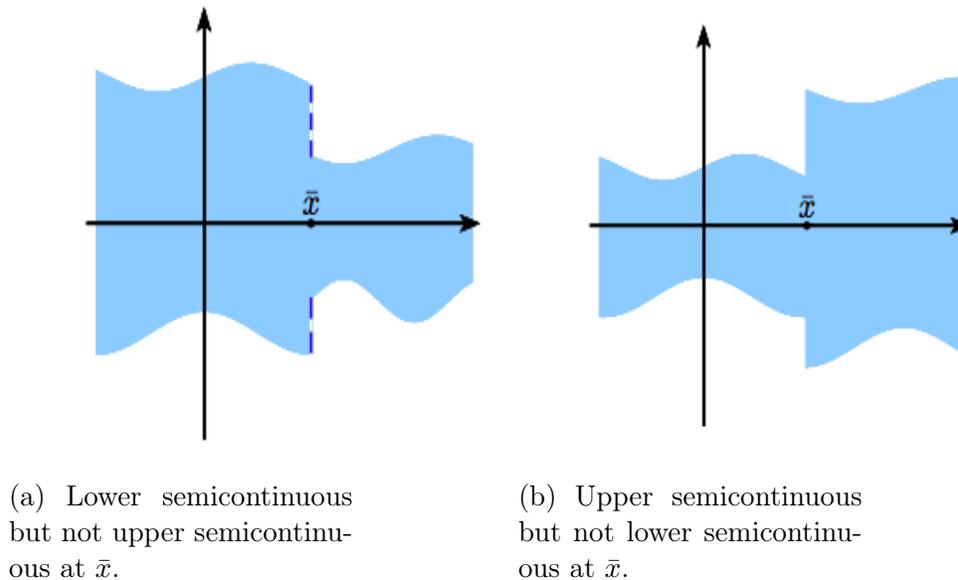


Figure 1-3 – Upper and lower semicontinuity of set-valued mappings

Some scholars consider definition of closedness for upper semicontinuity of set-valued mappings. Γ is closed (lower semicontinuous or upper semicontinuous), if it is closed (lower semicontinuous or upper semicontinuous) at each point of its domain. Moreover, we call a set-valued mapping Γ closed-valued (convex-valued), if for each

$x \in X$, the set $\Gamma(x)$ is closed (convex).

Now, we recall some notions about Lipschitz continuity for set-valued mappings that are extensions of Lipschitz continuity of functions.

Definition 1.12. *The set-valued mapping Γ is called upper locally Lipschitz at $\bar{x} \in X$, if there exist a constant k and a neighborhood O of \bar{x} such that*

$$\Gamma(x) \subseteq \Gamma(\bar{x}) + k\|x - \bar{x}\|B, \quad \forall x \in O,$$

where B is the closed unit ball in \mathbf{R}^p .

It is easy to see that upper locally Lipschitz continuity at a given point implies closedness at the point, provided the value of set-valued mapping is closed at this point. Nonetheless, a set-valued mapping may be upper locally Lipschitz while it is not upper semicontinuous. The following example clarifies the issue.

Example 1.2. *Let $\Gamma : \mathbf{R} \rightrightarrows \mathbf{R}_{\geq}^2$ be given by $\Gamma(x) = \{(y_1, y_2) : y_2 y_1 \geq x, y_1 \geq 1\}$. It is easily seen that Γ is upper locally Lipschitz at $\bar{x} = 2$, though the open set $U = \{(y_1, y_2) : y_2 y_1 > 2 - y_1^{-1}, y_1 > 0.5\}$ does not involve $\Gamma(x)$ for each $x < 2$.*

Compactness and upper locally Lipschitz continuity of set-valued mapping imply upper semicontinuity. In some important cases, it is of interest to investigate the behaviour of set-valued mappings around a point in their graph; for example differentiability of set-valued mappings. Due to this fact, pseudo-Lipschitz continuity (also called Aubin property and Lipschitz-like property) was proposed. Aubin [4] first introduced this notion which plays an essential role in analysis of set-valued mappings; see [64].

Definition 1.13. *The set-valued mapping Γ is said to be pseudo-Lipschitz continuous at $(\bar{x}, \bar{y}) \in \text{gph}\Gamma$, if there are neighborhoods O of \bar{x} and U of \bar{y} such that*

$$\Gamma(x_1) \cap U \subseteq \Gamma(x_2) + k\|x_1 - x_2\|B, \quad \forall x_1, x_2 \in O,$$

for a constant k .

Γ is lower semicontinuous at \bar{x} , if it is pseudo-Lipschitz continuous at (\bar{x}, \bar{y}) for each $\bar{y} \in \Gamma(\bar{x})$. On the contrary, a pseudo-Lipschitz set-valued mapping may be neither upper semicontinuous nor closed at \bar{x} ; see Example 1.2. Moreover, pseudo-Lipschitz continuity at (\bar{x}, \bar{y}) for each $\bar{y} \in \Gamma(\bar{x})$ may not imply upper Lipschitz continuity at \bar{x} , and vice versa. However, pseudo-Lipschitz continuity at (\bar{x}, \bar{y}) for each $\bar{y} \in \Gamma(\bar{x})$ implies upper Lipschitz continuity, closedness and upper semi continuity at \bar{x} under additional assumptions such as closedness, compactness, etc. We omit the proofs as they are straightforward. The next theorem states a handy criterion for checking pseudo-Lipschitz continuity.

Theorem 1.3. *Let $\text{gph}\Gamma$ be locally closed around $(\bar{x}, \bar{y}) \in \text{gph}\Gamma$. Γ is pseudo-Lipschitz continuous at (\bar{x}, \bar{y}) if and only if*

$$(\nu, 0) \in N_{\text{gph}\Gamma}(\bar{x}, \bar{y}) \Rightarrow \nu = 0.$$

The preceding theorem is known as Mordukhovich criterion in the literature. In this section, we only recalled some notions about continuity of set-valued mapping which will be used in our work. More concepts, including continuity notions defined by Pompeiu-Hausdorff distance, can be very utile and meaningful in some situations. For a deeper discussion of these notions and their relationships with stated concepts, we refer the reader to [47, 62, 64].

1.3.2 Differentiability

In this subsection, we review some concepts about differentiability of set-valued mappings, most of which are motivated by geometric interpretation of derivative of single-valued functions. first, we introduce the concept of contingent derivatives introduced by Aubin [3]. It is a kind of standard derivative notion of set-valued mappings and it also has been used widely in the literature. Some scholars use the term graphical derivative instead of contingent derivative.

Definition 1.14. *Let $(\bar{x}, \bar{y}) \in \text{gph}\Gamma$ be given. A set-valued mapping $D\Gamma(\bar{x}, \bar{y}) : \mathbf{R}^n \rightrightarrows$*

\mathbf{R}^p is called contingent derivative of Γ at (\bar{x}, \bar{y}) , if

$$gphD\Gamma(\bar{x}, \bar{y}) = T_{gph\Gamma}(\bar{x}, \bar{y}).$$

As mentioned before, contingent derivative is an extension of derivative for single-valued functions. If Γ is single-valued and differentiable, then we have $D\Gamma(\bar{x}, \Gamma(\bar{x}))(d) = \nabla\Gamma(\bar{x})d$, for each $d \in \mathbf{R}^n$. When Γ is pseudo-Lipschitz continuous at (\bar{x}, \bar{y}) , $dom(D\Gamma(\bar{x}, \bar{y})) = \mathbf{R}^n$. While this not true when Γ is not pseudo-Lipschitz continuous. Furthermore, contingent derivative enjoys the following property:

$$D\Gamma(\bar{x}, \bar{y})(\bar{d}) = \limsup_{d \rightarrow \bar{d}, t \downarrow 0} \frac{\Gamma(\bar{x} + td) - \bar{y}}{t}. \quad (1.2)$$

The last formula is an extension of the upper Dini derivative notion for set-valued mappings. As the Bouligand tangent cone can be intractable in some cases (may be whole space or non-convex), other tangent cones have also been used to define new notions of derivative.

Definition 1.15. Let $(\bar{x}, \bar{y}) \in gph\Gamma$ be given. A set-valued mapping $D_{adj}\Gamma(\bar{x}, \bar{y}) : \mathbf{R}^n \rightrightarrows \mathbf{R}^p$ is said to be adjacent derivative of Γ at (\bar{x}, \bar{y}) , if

$$gphD_{adj}\Gamma(\bar{x}, \bar{y}) = T_{gph\Gamma}^a(\bar{x}, \bar{y}).$$

In addition, we say Γ is Proto-differentiable at (\bar{x}, \bar{y}) , if $T_{gph\Gamma}(\bar{x}, \bar{y}) = T_{gph\Gamma}^a(\bar{x}, \bar{y})$.

Rockafellar [63] introduced the concept of Proto-differentiability to take advantage of derivability. He also established the Proto-differentiability of a class of set-valued mappings corresponding to some parametric systems (under some mild conditions). Likewise, Clarke derivative is defined as follows.

Definition 1.16. Let $(\bar{x}, \bar{y}) \in gph\Gamma$ be given. A set-valued mapping $D^c\Gamma(\bar{x}, \bar{y}) : \mathbf{R}^n \rightrightarrows \mathbf{R}^p$ is called Clarke derivative of Γ at (\bar{x}, \bar{y}) , if

$$gphD^c\Gamma(\bar{x}, \bar{y}) = T_{gph\Gamma}^c(\bar{x}, \bar{y}).$$

Moreover, Γ is said to be Clarke differentiable at (\bar{x}, \bar{y}) , if $T_{gph\Gamma}(\bar{x}, \bar{y}) = T_{gph\Gamma}^c(\bar{x}, \bar{y})$.

It is easily seen that Clarke differentiability implies Proto-differentiability. In the following example, we calculate the aforementioned derivatives at a given point.

Example 1.3. Let $\Gamma : \mathbf{R} \rightrightarrows \mathbf{R}$ be given by

$$\Gamma(x) = \begin{cases} \{t : t \leq x \sin(\ln x)\}, & x > 0 \\ \{t : t \leq 0\}, & x \leq 0. \end{cases}$$

At $(\bar{x}, \bar{y}) = (0, 0)$, we have $D\Gamma(\bar{x}, \bar{y})(1) = \{e : e \leq 1\}$, $D_{adj}\Gamma(\bar{x}, \bar{y})(1) = \{e : e \leq -1\}$ and $D^c\Gamma(\bar{x}, \bar{y})(1) = \emptyset$.

Relation (1.2) provides a motivation for defining the following concept.

Definition 1.17. A set-valued mapping $D_{low}\Gamma(\bar{x}, \bar{y}) : \mathbf{R}^n \rightrightarrows \mathbf{R}^p$ is called lower Dini derivative of Γ at $(\bar{x}, \bar{y}) \in gph\Gamma$ if for each $\bar{d} \in \mathbf{R}^n$,

$$D_{low}\Gamma(\bar{x}, \bar{y})(\bar{d}) = \liminf_{d \rightarrow \bar{d}, t \downarrow 0} \frac{\Gamma(\bar{x} + td) - \bar{y}}{t}.$$

If $T_{gph\Gamma}(\bar{x}, \bar{y}) = gphD_{low}\Gamma(\bar{x}, \bar{y})$, then Γ is said to be semi-differentiable at (\bar{x}, \bar{y}) ; equivalently, if for each $(d, e) \in T_{gph\Gamma}(\bar{x}, \bar{y})$, $d_n \rightarrow \bar{d}$ and $t_n \downarrow 0$, there exists $e_n \rightarrow e$ such that $\{(\bar{x} + t_n d_n, \bar{y} + t_n e_n)\} \subseteq \Gamma$.

As mentioned above, the semi-differentiability is an extension of the lower Dini derivative. The concept of semi-differentiability was introduced by Penot [60]. He also established semi-differentiability of some parametric systems. It is obvious from definitions that semi-differentiability implies Proto-differentiability. The converse of this statement may not hold, in general. The following example demonstrates this.

Example 1.4. Let $\Gamma : \mathbf{R} \rightrightarrows \mathbf{R}$ be given by

$$\Gamma(x) = \begin{cases} \{\sqrt{x}\}, & x > 0 \\ \{0\}, & x \leq 0. \end{cases}$$

At $(\bar{x}, \bar{y}) = (0, 0)$, we have $D_{adj}\Gamma(\bar{x}, \bar{y})(0) = \{e : e \geq 0\}$, but $D_{low}\Gamma(\bar{x}, \bar{y})(0) = \emptyset$.

The following proposition provides a sufficient condition under which the converse of the above statement holds.

Proposition 1.1. *Let $(\bar{x}, \bar{y}) \in \text{gph}\Gamma$ be given and let Γ be pseudo-Lipschitz continuous at the point. If Γ is Proto-differentiable at (\bar{x}, \bar{y}) , then it is also semi-differentiable at the point.*

If a set-valued mapping Γ is convex-valued and pseudo-Lipschitz continuous at a given point, then it is semi-differentiable [20]. In the remainder of this section, we pay attention to the concept of strictly lower Dini derivative proposed by Thibault [73].

Definition 1.18. *A set-valued mapping $D_{s-low}\Gamma(\bar{x}, \bar{y}) : \mathbf{R}^n \rightrightarrows \mathbf{R}^p$ is said to be strictly lower Dini derivative of Γ at $(\bar{x}, \bar{y}) \in \text{gph}\Gamma$ if for each \bar{d} ,*

$$D_{s-low}\Gamma(\bar{x}, \bar{y})(\bar{d}) = \liminf_{\substack{d \rightarrow \bar{d}, t \downarrow 0 \\ (x, y) \xrightarrow{\text{gph}\Gamma} (\bar{x}, \bar{y})}} \frac{\Gamma(x + td) - y}{t}.$$

In addition, if $T_{\text{gph}\Gamma}(\bar{x}, \bar{y}) = \text{gph}D_{s-low}\Gamma(\bar{x}, \bar{y})$, then we say Γ is strictly semi-differentiable at (\bar{x}, \bar{y}) . In other words, if for each $(d, e) \in T_{\text{gph}\Gamma}(\bar{x}, \bar{y})$, $d_n \rightarrow \bar{d}$, $(x_n, y_n) \xrightarrow{\text{gph}\Gamma} (\bar{x}, \bar{y})$ and $t_n \downarrow 0$, then there exists $e_n \rightarrow e$ such that $\{(x_n + t_n d_n, y_n + t_n e_n)\} \subseteq \Gamma$.

Rockafellar [2] established Proto-differentiability of a parametric system under Mangasarian-Fromovitz constraint qualification and Amahroq et al. [2] proved its strict semi-differentiability under the same condition. They also investigated strict semi-differentiability of general parametric systems involving intersection. It is obvious that strict semi-differentiability at a given point implies both semi-differentiability and Clarke differentiability. however, the converse does not necessarily hold. The following proposition introduces sufficient conditions under which strict semi-differentiability is derived from Clarke differentiability.

Proposition 1.2. *Let $(\bar{x}, \bar{y}) \in \text{gph}\Gamma$ be given and let Γ be pseudo-Lipschitz continuous at the point. If Γ is Clarke differentiable at (\bar{x}, \bar{y}) , then it is also strictly semi-differentiable at the point.*

As the last notion of the subsection, we address s-derivative. It was introduced by Shi under the name TP-derivative [67]. Taa called it s-derivative [69].

Definition 1.19. *A set-valued mapping $D^s\Gamma : \mathbf{R}^n \rightrightarrows \mathbf{R}^p$ is called s-derivative of Γ at $(\bar{x}, \bar{y}) \in \text{gph}\Gamma$, if $D_s\Gamma(\bar{d})$ is consisted of all e such that there are $(d_n, e_n) \rightarrow (d, e)$ and $\{t_n\} \subseteq \mathbf{R}_{\geq}$ with $\{(\bar{x} + t_n d_n, \bar{y} + t_n e_n)\} \subseteq \text{gph}\Gamma$ and $t_n d_n \rightarrow 0$.*

It is easy to see that the graph of s-derivative contains that of contingent derivative. In addition, all of the derivatives presented so far will be equal when the graph of the set-valued mapping in question is convex.

We conclude the section by the celebrated Kakutani fixed-point theorem.

Theorem 1.4. *(Kakutani fixed-point theorem) Let $\Gamma : C \rightrightarrows C$ be a given set-valued mapping. Suppose that C is compact and convex and $\Gamma(x)$ is convex and nonempty for each $x \in C$. If graph of Γ is closed, then there exists $\bar{x} \in C$ satisfying $\bar{x} \in C(\bar{x})$.*

The proof of Theorem 1.4 can be found in [29]. It is worth mentioning that John Nash used the Kakutani fixed-point theorem to prove the existence of equilibrium in finite impartial games.

1.4 Clarke subdifferential

Throughout this work, we consider locally Lipschitz data when we are talking about nonsmooth data. To handle nonsmoothness, we choose the Clarke generalized gradient. This tool was introduced by Clarke in 1973 in his Ph.D. thesis and since then has found considerable applications in both optimization and nonsmooth analysis. It shares many classical theorems of real analysis such as mean value theorem, implicit function theorem and so on. We only provide the definitions and the results that will be needed in the sequel. See [16, 17] for more detailed information on this subject. According to Rademacher's theorem, we know every locally Lipschitz function on \mathbf{R}^n is almost everywhere differentiable in the sense of Lebesgue measure. The definition of generalized gradient is as follows.

Definition 1.20. Let $f : \mathbf{R}^n \rightarrow \mathbf{R}^p$ be a locally Lipschitz function. The generalized gradient of f at \bar{x} , denoted by $\partial f(\bar{x})$, is defined by

$$\partial f(\bar{x}) := \text{co}\left\{\lim_{n \rightarrow \infty} \nabla f(x_n) : x_n \rightarrow \bar{x}, x_n \notin X, x_n \notin X_f\right\},$$

where X is any set with zero Lebesgue measure and X_f is the set of points at which f is not differentiable.

Definition 1.21. Let $f : \mathbf{R}^n \rightarrow \mathbf{R}^p$. We say f is strictly differentiable at \bar{x} if there is $\nabla f(\bar{x}) \in \mathbf{R}^{p \times n}$ such that

$$\lim_{x, y \rightarrow \bar{x}} \frac{f(x) - f(y) - \langle \nabla f(\bar{x}), x - y \rangle}{\|x - y\|} = 0.$$

It is readily seen that continuously differentiable functions are strictly differentiable on their domains. Generalized gradient enjoys nice properties. We list some of them in the following propositions.

Proposition 1.3. Let $f : \mathbf{R}^n \rightarrow \mathbf{R}^p$ be a locally Lipschitz function with modulus k near \bar{x} .

- 1) $\partial f(\bar{x})$ is a nonempty, convex and compact set.
- 2) For every $\nu \in \partial f(\bar{x})$ we have $\|\nu\| \leq k$.
- 3) The set-valued mapping $\partial f(\cdot)$ is upper semi-continuous and closed.
- 4) f is strictly differentiable at \bar{x} if and only if $\partial f(\bar{x})$ is a singleton.

The next proposition provides more properties when f is from \mathbf{R}^n to \mathbf{R} .

Proposition 1.4. Suppose that $f : \mathbf{R}^n \rightarrow \mathbf{R}$ is a locally Lipschitz function.

- 1) If f is convex, then generalized gradient coincides with subgradient in the sense of classic convex analysis.
- 2) If \bar{x} is a local extremum, then $0 \in \partial f(\bar{x})$.

3)

$$\limsup_{x \rightarrow \bar{x}, t \downarrow 0} \frac{f(x + td) - f(x)}{t} = \max_{\nu \in \partial f(\bar{x})} \langle \nu, d \rangle.$$

We denote the value $\max_{\nu \in \partial f(\bar{x})} \langle \nu, d \rangle$ by $f^\circ(\bar{x}; d)$.

Definition 1.22. Let $f : \mathbf{R}^n \rightarrow \mathbf{R}$ be a locally Lipschitz function. f is called regular at \bar{x} , if for each $d \in \mathbf{R}^n$,

$$f'(\bar{x}; d) := \lim_{t \downarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t} = \max_{\nu \in \partial f(\bar{x})} \langle \nu, d \rangle.$$

We call $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ regular at a given point, if its components are regular at the point.

We say that a function is regular, if it is regular at each point in its domain. The class of regular functions is vast; for example, convex functions and strictly differentiable functions are regular. Furthermore, sum and pointwise maximum of a finite number of regular functions are also regular. However, negative scalar multiplication does not necessarily preserve regularity; consider $f(x) = \|x\|$. the next proposition gives some calculus rules.

Proposition 1.5. Let $f_1, \dots, f_k : \mathbf{R}^n \rightarrow \mathbf{R}$ be locally Lipschitz functions.

- 1) For each scalar λ , $\partial \lambda f_1(x) = \lambda \partial f_1(x)$.
- 2) $\partial(\sum_{i=1}^k f_i(x)) \subseteq \sum_{i=1}^k \partial f_i(x)$.
- 3) Suppose f is given by $f(x) = \max_{i=1}^k f_i(x)$. Then, $\partial f(x) \subseteq \text{co}(\cup_{i \in I(x)} \partial f_i(x))$, where $I(x) = \{i : f_i(x) = f(x)\}$.

The inclusions given in properties (2) and (3) hold as equality when f_i are regular.

Theorem 1.5. (Mean Value Theorem) Let $f : \mathbf{R}^n \rightarrow \mathbf{R}$ be a locally Lipschitz function. Then, for each $x, y \in \mathbf{R}^n$, there exists $z \in (x, y)$ such that

$$f(y) - f(x) \in \langle \partial f(z), y - x \rangle.$$

It is easy to see that if $f : \mathbf{R}^n \rightarrow \mathbf{R}^p$, then $\partial f(x) \subseteq \begin{bmatrix} \partial f_1(x) \\ \cdot \\ \partial f_p(x) \end{bmatrix}$. Nonetheless, equality does not hold necessarily; for instance, consider $f(x) = \begin{bmatrix} |x| \\ -|x| \end{bmatrix}$. Then, $\partial f(0) = \text{co}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right)$, while $\begin{bmatrix} \partial f_1(0) \\ \partial f_2(0) \end{bmatrix} = \text{co}\left(\begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$.

Theorem 1.6. (*Implicit Function Theorem*) *Let $f : \mathbf{R}^n \times \mathbf{R}^p \rightarrow \mathbf{R}^p$ be a locally Lipschitz function and let $f(\bar{x}, \bar{y}) = 0$. Suppose that each $M \in \partial_y f(\bar{x}, \bar{y})$ is invertible. Then, there exist a neighborhood O of \bar{x} and a Lipschitz function $g : O \rightarrow \mathbf{R}^p$ such that $g(\bar{x}) = \bar{y}$ and*

$$f(x, g(x)) = 0, \quad \forall x \in O.$$

In the above theorem, $\partial_y f(\bar{x}, \bar{y}) = \{M : \exists N; [N \ M] \in \partial f(\bar{x}, \bar{y})\}$. A related question to the foregoing theorem is the following amazing open problem.

Question 1.1. *Let $f : \mathbf{R}^{(n+p)} \rightarrow \mathbf{R}^p$ be locally Lipschitz. Suppose that for every $M \in \partial f(\bar{x})$, M has full rank. Is there some $n \times (n+p)$ matrix, N , such that $\begin{bmatrix} M \\ N \end{bmatrix}$ is invertible for each $M \in \partial f(\bar{x})$.*

The answer to Question 1.1 for some classes of functions is affirmative; see [21] and the references therein. The response to this query may lead to extending important results about Lipschitz manifolds and implicit function theorems. Consequently, some results in sensitivity analysis could be extended. Now, let us see the implicit function theorem in a more general setting [45, 62].

Theorem 1.7. *Let $f : \mathbf{R}^q \times \mathbf{R}^n \rightarrow \mathbf{R}^p$ be locally Lipschitz. Suppose that the set-valued mapping Γ is defined by $\Gamma(u) = \{x \in C : f(u, x) \in K\}$, where $K \subseteq \mathbf{R}^p$ is a closed and convex cone and $C \subseteq \mathbf{R}^n$ is closed and convex. If $0 \in \text{int}(M(T(\bar{x})) + K)$ for every $M \in \partial_x f(\bar{u}, \bar{x})$, then Γ is pseudo-Lipschitz continuous at (\bar{u}, \bar{x}) .*

Now, we characterize Clarke tangent cones and Clarke normal cone by generalized gradient notion. Let d_X be the distance function defined by $d_X(y) = \inf\{\|y - x\| : x \in X\}$. The distance function is Lipschitz with modulus 1. Thanks to the distance function, Clarke tangent cone can be expressed as $T_X^c(\bar{x}) = \{d : d_X^\circ(\bar{x}; d) \leq 0\}$ and $N_X^c(\bar{x}) = cl\left(pos(\partial d_X(\bar{x}))\right)$.

Lemma 1.1. *The vector d belongs to Clarke tangent cone X at $\bar{x} \in X$ if and only if for each sequence $x_n \xrightarrow{X} \bar{x}$ and $t_n \downarrow 0$, there exist subsequences $\{x_{n_k}\}$ and $d_{n_k} \rightarrow d$ such that $\{x_{n_k} + t_{n_k}d_{n_k}\} \subseteq X$.*

Proof. Due to the facts addressed before the lemma, we only need to prove the "only if part". Let d fulfill the stated property. From the definition of generalized gradient, there are sequences $y_n \rightarrow \bar{x}$ and $t_n \downarrow 0$ such that

$$\lim_{n \rightarrow \infty} \frac{d_X(y_n + t_n d) - d_X(y_n)}{t_n} = d_X^\circ(\bar{x}; d).$$

Of course, $\{y_n\}$ converges to \bar{x} . Consider a sequence $\{x_n\}$ with $\|y_n - x_n\| \leq d_X(y_n) + t_n/n$. Therefore, there is a sequence $d_{n_k} \rightarrow d$ such that $\{x_{n_k} + t_{n_k}d_{n_k}\} \subseteq X$. As a result,

$$\begin{aligned} d_X(y_{n_k} + t_{n_k}d) &\leq d_X(x_{n_k} + t_{n_k}d_{n_k}) + \|y_{n_k} - x_{n_k}\| + t_{n_k}\|d - d_{n_k}\| \\ &\leq d_X(y_{n_k}) + t_{n_k}\|d - d_{n_k}\| + \frac{t_{n_k}}{n_k}. \end{aligned}$$

These imply $d_X^\circ(\bar{x}; d) = 0$, leading to $d \in T_X^c(\bar{x})$. □

In most practical problems, sets are characterized explicitly by functions. The next theorem provides estimations for tangent cones and normal cones in these situations.

Theorem 1.8. *Let $X \subseteq \mathbf{R}^n$ be given by $X = \{x : g(x) \leq 0\}$, where $g : \mathbf{R}^n \rightarrow \mathbf{R}^m$ is a locally Lipschitz function. Suppose that $\bar{x} \in X$ and $I(\bar{x}) := \{i : g_i(\bar{x}) = 0\}$. If*

$0 \notin \text{co}(\cup_{i \in I(\bar{x})} \partial g_i(\bar{x}))$, then

$$\{d : g_i^\circ(\bar{x}; d) \leq 0, \forall i \in I(\bar{x})\} \subseteq T_X^c(\bar{x}),$$

$$N_X^c(\bar{x}) \subseteq \text{co}(\{\partial g_i(\bar{x}), \forall i \in I(\bar{x})\}).$$

In addition to the above-mentioned assumptions, if g is regular at \bar{x} , then equality holds in both inclusions, X is regular at \bar{x} and we have $T_X(\bar{x}) = T_X^c(\bar{x})$.

Proof. Let h be given by $h(x) = \max_{1 \leq i \leq m} g_i(x)$. Thus, $X = \{x : h(x) \leq 0\}$ and the theorem follows from Theorem 10.42 in [17] and Proposition 1.5. \square

We close this chapter by the extended Farkas Lemma for semi-infinite linear systems.

Theorem 1.9. [37] *Let T be a given index set. Then, exactly one of the following two statements is true:*

(i) *The following system has a solution.*

$$\begin{cases} a_t^T x \leq 0, & t \in T \\ a^T x > 0. \end{cases}$$

(ii) $a \in \text{cl}(\text{pos}(\{a_t : t \in T\}))$.

Chapter 2

Scalarization: Extending two important theorems

2.1 Introduction

An important question in multi-objective optimization is the investigation of existence of proper efficient solutions. Scalarization methods are not only powerful tools to generate efficient solutions, but also provide valuable information about proper and improper efficient solutions. Here, we only concentrate on the applications of scalarization tools related to properly efficient solutions. We refer the reader to [24, 26, 56] for studying more aspects of these methods.

By scalarization methods, one formulates a single objective optimization problem corresponding to a given multi-objective optimization problem and he/she studies the relationships between the solutions of the two problems. It turns out in multi-objective optimization that any efficient solution of a given multi-objective optimization problem can be characterized as a solution of certain single objective optimization problems [24, 26].

Scalarization techniques have numerous applications in interactive algorithms [57], estimating Pareto frontier [56], obtaining solutions with bounded trade-offs [46], etc.

In this chapter, we review some popular scalarization methods, though most of them are closely related. The reader is referred to the book [26] for more details. In

the chapter, we consider the following general multi-objective optimization problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{2.1}$$

where X is a subset of \mathbf{R}^n and $f : X \rightarrow \mathbf{R}^p$.

2.2 Benson's method

This section deals with the Benson's method as a popular scalarization technique. Benson's method [8] gives an examination of the existence of efficient and properly efficient solutions for multi-objective optimization problems. This technique employs l_1 -norm to check the efficiency of a given point x^0 by solving the following single-objective optimization problem:

$$\begin{aligned} \max \quad & \sum_{k=1}^p l_k \\ \text{s.t.} \quad & f_k(x^0) - l_k - f_k(x) = 0, \quad k = 1, 2, \dots, p, \\ & l_k \geq 0, \quad k = 1, 2, \dots, p, \\ & x \in X. \end{aligned} \tag{2.2}$$

A linear version of the above model was studied by Ecker and Kouada [23]. Some useful examples clarifying Model (2.2) can be found in Benson [8] and Giannessi et al. [33].

The vector $x^0 \in X$ is efficient if and only if the optimal value of Problem (2.2) is zero; see [9, 24]. If (\hat{x}, \hat{l}) is an optimal solution of Model (2.2), then \hat{x} is an efficient solution; see [9, 24]. The natural question that arises here is: what happens when Problem (2.2) is unbounded? Benson [8] answered this question under convexity assumptions. He proved the following result. See also Theorem 4.16 in [24].

Theorem 2.1. *[9, 24] Assume that f_k , $k = 1, 2, \dots, p$, are convex functions and X*

is a convex set. If Problem (2.2) is unbounded, then $X_{PE} = \emptyset$.

As can be seen from the above theorem and discussion, any optimal solution of Problem (2.2) yields an efficient solution. Furthermore, in many cases the unboundedness of Problem (2.2) shows that no properly efficient solution exists. In the following, we prove that this important result holds in all cases (without any assumption).

Theorem 2.2. *If Problem (2.2) is unbounded, then $X_{PE} = \emptyset$.*

Proof. If $x^0 \in X$ is efficient, then Problem (2.2) has a finite optimal value equal to zero (see [8, 24]). Therefore, due to the assumption, we have $x^0 \in X \setminus X_E$.

To the contrary, assume that there exists $\hat{x} \in X_{PE}$. Since $\hat{x} \in X_{PE}$, due to the Henig proper efficiency, there exists a convex and pointed cone C , such that $\mathbf{R}_{\geq}^p \subseteq \text{int}C$ and

$$(f(\hat{x}) - C \setminus \{0\}) \cap f(X) = \emptyset.$$

According to $\mathbf{R}_{\geq}^p \subseteq \text{int}C$, we have $e_i \in \text{int}C$, for $i = 1, 2, \dots, m$ (vector e_i denotes the i th unit vector in \mathbf{R}^p). Thus, there exists $r_i > 0$ such that $B(e_i; r_i) = \{y \in \mathbf{R}^p : \|y - e_i\| < r_i\} \subseteq C$.

Now, consider the following system, in which α , θ , and y are variables:

$$\begin{cases} f(x^0) - \alpha e_i = f(\hat{x}) - \theta(e_i + y), \\ \|y\| < r_i, \\ \alpha > 0, \theta > 0. \end{cases} \quad (2.3)$$

Set $\bar{\alpha}_i := 2r_i^{-1}\|f(\hat{x}) - f(x^0)\|$ and $\bar{y} := \frac{r_i}{2} \times \frac{f(\hat{x}) - f(x^0)}{\|f(\hat{x}) - f(x^0)\|}$. The vector

$$(\alpha, \theta, y) = (\bar{\alpha}_i, \bar{\alpha}_i, \bar{y})$$

is a solution of the system (2.3), and $e_i + y \in B(e_i; r_i)$. Thus,

$$f(x^0) - \bar{\alpha}_i e_i - f(\hat{x}) \in -C.$$

Also, clearly we have

$$\forall \alpha > \bar{\alpha}_i, \quad -(\alpha - \bar{\alpha}_i)e_i \in -C \setminus \{0\}.$$

Therefore, for each $\alpha > \bar{\alpha}_i$, we have $f(x^0) - \alpha e_i - f(\hat{x}) \in -C \setminus \{0\}$, because C is a convex cone. This implies $f(x^0) - \alpha e_i \in f(\hat{x}) - C \setminus \{0\}$. Hence, $f(x^0) - \alpha e_i \notin f(X)$, due to Henig proper efficiency.

So far, we proved that for each e_i , $i = 1, 2, \dots, m$, there exists a positive scalar $\bar{\alpha}_i > 0$, such that

$$\forall \alpha > \bar{\alpha}_i, \quad f(x^0) - \alpha e_i \notin f(X). \quad (2.4)$$

Now, defining $\bar{\alpha} := \max_{1 \leq i \leq m} \bar{\alpha}_i$, we show that for any $d \in \mathbf{R}_{\geq}^p$ satisfying $\sum_{j=1}^m d_j = 1$ and any $\alpha > \bar{\alpha}$, we have $f(x^0) - \alpha d \notin f(X)$. The cone $-C$ is convex and also, for each i , we have $f(x^0) - \bar{\alpha} e_i - f(\hat{x}) \in -C$. Therefore,

$$\sum_{i=1}^m d_i (f(x^0) - \bar{\alpha} e_i - f(\hat{x})) \in -C.$$

This implies

$$f(x^0) - \bar{\alpha} d - f(\hat{x}) \in -C.$$

Furthermore, we have $d \in C$, which implies

$$-(\alpha - \bar{\alpha})d \in -C \setminus \{0\}, \quad \forall \alpha > \bar{\alpha}.$$

Thus,

$$f(x^0) - \alpha d - f(\hat{x}) \in -C \setminus \{0\}, \quad \forall \alpha > \bar{\alpha}.$$

Therefore, due to the proper efficiency of \hat{x} , we get

$$f(x^0) - \alpha d \notin f(X), \quad \forall \alpha > \bar{\alpha}.$$

Thus, $\bar{\alpha}$ provides an upper bound for the objective function of Problem (2.2). This contradicts the unboundedness assumption on Problem (2.2), and completes the proof. □

It is worth mentioning that the converse of Theorem 2.2 does not hold necessarily. The following example illustrate this point.

Example 2.1. *Consider this multi-objective optimization problem.*

$$\begin{aligned} & \min (x_1, x_2) \\ & \text{s.t. } x_1 x_2 = 1, \\ & \quad x \in -\mathbf{R}_{>}^2. \end{aligned}$$

It is easy to see that the above problem does not have any properly efficient solution. However, the Benson's problem for each x^0 belonging to feasible set is bounded.

Now, we consider the following single-objective problem which has been studied by Guddat et al. [39] as a hybrid scalarization method; see also [24]:

$$\begin{aligned} & \min \sum_{k=1}^p \lambda_k f_k(x) \\ & \text{s.t. } f_k(x) \leq f_k(x^0), \quad k = 1, 2, \dots, p, \\ & \quad x \in X. \end{aligned} \tag{2.5}$$

In this problem, $\lambda_1, \lambda_2, \dots, \lambda_m$ are nonnegative fixed scalars. This problem is an extension of Problem (2.2). Setting $\lambda_i = 1$, for $i = 1, 2, \dots, p$, Model (2.5) leads to Model (2.2).

The vector $x^0 \in X$ is efficient if and only if $x^0 \in X$ is an optimal solution of Problem (2.5); see [24, 39]. It is not difficult to see that the unboundedness of Problem (2.5) implies the unboundedness of Problem (2.2). Therefore, by Theorem 1, we get the following corollary about the problem of Guddat et al. [39].

Corrolary 2.1. *If Problem (2.5) is unbounded, then $X_{PE} = \emptyset$.*

Under the \mathbf{R}_{\geq}^m -closedness and \mathbf{R}_{\geq}^m -convexity assumptions, Benson [8] proved that, the unboundedness of Problem (2.2) implies $X_E = \emptyset$. Due to the above discussion, it can be seen that, under \mathbf{R}_{\geq}^m -closedness and \mathbf{R}_{\geq}^m -convexity assumptions, the unboundedness of Problem (2.5) implies $X_E = \emptyset$ as well. Although, in Theorem 2.2, we omitted the \mathbf{R}_{\geq}^m -convexity assumption for investigating X_{PE} , the following example shows that one cannot omit this assumption for X_E .

Example 2.2. *Let $X = \{(x_1, x_2) : x_1 < 0, x_2 \leq 0, x_2 \geq \frac{1}{x_1}\} \cup \{(0, 0)\}$ and $f_1(x) = x_1, f_2(x) = x_2$. Hence, $Y = X$ is \mathbf{R}_{\geq}^2 -closed, while it is not \mathbf{R}_{\geq}^2 -convex. Considering $x^0 = (0, 0)$, Problem (2.2) is unbounded while $X_E = \{(x_1, x_2) \in X : x_2 = \frac{1}{x_1}\}$.*

It is worth mentioning that Theorem 2.2 holds for any ordering cone. Let $C \subseteq \mathbf{R}^p$ be a convex, closed and pointed cone with $\text{int}(C) \neq \emptyset$. We call $\bar{x} \in X$ to be an efficient solution of Problem (2.1) with respect to C , if $f(X) \cap (f(\bar{x}) - C) = \{f(\bar{x})\}$. Likewise, $\bar{x} \in X$ is called a weakly efficient solution of Problem (2.1) with respect to C if $f(X) \cap (f(\bar{x}) - \text{int}(C)) = \emptyset$.

The point $\bar{x} \in X$ is called a properly efficient solution of (2.1) in the Benson's sense with respect to C , if

$$\text{cl}(\text{pos}(f(X) + C - f(\bar{x}))) \cap (-C) = \{0\}.$$

That is equivalent to the existence of a pointed, closed and convex cone K with $C \setminus \{0\} \subseteq \text{int}(K)$ such that $f(X) \cap (f(\bar{x}) - K) = \{f(\bar{x})\}$ (proper efficiency in the Henig's sense).

Consider the following scalarization corresponding to Problem (2.1) with respect to C :

$$\begin{aligned} & \min \lambda^T f(x) \\ & \text{s.t. } f(x_0) - f(x) \in C \\ & \quad x \in X, \end{aligned} \tag{2.6}$$

in which $\lambda \in C^\circ$ is a fixed nonzero vector and x_0 is a given feasible solution. If x_0 is an efficient solution of Problem (2.1), then the optimal value of the preceding problem is zero.

Theorem 2.3. *If $X_{PE} \neq \emptyset$ with respect to C , then for any $x_0 \in X$, Problem (2.6) has a finite optimal value.*

Proof. To the contrary, assume that there exists $x_0 \in X$ such that Problem (2.6) is unbounded. Because of $X_{PE} \neq \emptyset$, there is $\hat{x} \in X_{PE}$. According to Henig proper efficiency, there exists a closed, convex and pointed cone K , such that $C \setminus \{0\} \subseteq \text{int}(K)$ and

$$\left(f(X) - f(\hat{x}) \right) \cap (-K) = \{0\}. \quad (2.7)$$

As Problem (2.6) is unbounded, there is a sequence $\{x_n\} \subseteq X$ such that

$$0 \neq f(x_0) - f(x_n) \in C \subseteq K, \quad \lambda^T f(x_n) \leq -n, \quad \forall n \in \mathbb{N}.$$

The sequence $\frac{f(x_n) - f(x_0)}{\|f(x_n) - f(x_0)\|}$ is bounded in R^n , and so, without loss of generality we may assume that this sequence converges to a nonzero $d \in -C$ due to the closedness of C . Moreover, it is trivial that $\|f(x_n) - f(x_0)\| \rightarrow \infty$, as λ is fixed in R^n . So, $\frac{f(x_n) - f(\hat{x})}{\|f(x_n) - f(x_0)\|} \rightarrow d$. Furthermore, $C \setminus \{0\} \subseteq \text{int}(K)$, which leads to $-d \in \text{int}(K)$. Therefore, for n_0 sufficiently large, we have

$$\frac{f(x_n) - f(\hat{x})}{\|f(x_n) - f(x_0)\|} \in -K, \quad \forall n \geq n_0.$$

This contradicts (2.7), and completes the proof. □

2.3 Direction-based scalarization

Another popular scalarization method is direction-based scalarization [26]. The direction-based scalarization problem is as follows:

$$\begin{aligned}
 & \min t \\
 & \text{s.t. } f(x) \leq a + tr \\
 & x \in X,
 \end{aligned} \tag{2.8}$$

where $a \in \mathbf{R}^p$ and $r \in \mathbf{R}_{\geq}^p$. This is also known as the Pascoletti-Serafini scalarization in the literature. In addition, ϵ -constraint scalarization is a special case of this method [26].

Proposition 2.1. (i) *If (\bar{x}, \bar{t}) is an optimal solution of Problem (2.8), then \bar{x} is a weakly efficient solution of Problem (2.1).*

(ii) *If \bar{x} is an efficient solution of Problem (2.1), then $(\bar{x}, 0)$ is an optimal solution of the following problem for every $r \in \mathbf{R}_{\geq}^p$:*

$$\begin{aligned}
 & \min t \\
 & \text{s.t. } f(x) \leq f(\bar{x}) + tr \\
 & x \in X.
 \end{aligned}$$

Proof. See [26]. □

Proposition 2.2. *If Problem (2.8) is unbounded for $a \in \mathbf{R}^p$ and $r \in \mathbf{R}_{\geq}^p$, then multi-objective optimization Problem (2.1) does not have any properly efficient solution.*

Proof. As Problem (2.8) is unbounded, there is some $\hat{x} \in X$ such that the following

problem is unbounded:

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & f(x) \leq f(\hat{x}) + tr \\ & x \in X. \end{aligned}$$

Thus, the following problem is unbounded:

$$\begin{aligned} \max \quad & \sum_{i=1}^p l_i \\ \text{s.t.} \quad & f(x) + l = f(\hat{x}) \\ & x \in X, l \geq 0. \end{aligned}$$

By virtue of Theorem 2.2, the properly efficient set of multi-objective optimization problem (2.1) is empty. \square

2.4 Compromise Programming

Compromise programming is related to some useful techniques in multi-objective programming which are seeking for solutions as close as possible to an utopia point [24, 30]. We call the vector $y^U \in \mathbf{R}^p$ an utopia point, if $y_i^U < \min_{x \in X} f_i(x)$, for each i . The general form of a compromise programming problem corresponding to (MOP) is $\{\min d(f(x), y^U) \text{ s.t. } x \in X\}$, in which d is a measure function from $\mathbf{R}^p \times \mathbf{R}^p$ to $[0, +\infty)$. A popular measure function, which has been widely used in the literature, is defined by

$$d(\lambda, y) = \|\lambda \odot y\|_q,$$

for each $(\lambda, y) \in \mathbf{R}^p \times \mathbf{R}^p$, in which q is a positive integer,

$$\lambda \odot y = (\lambda_1 y_1, \lambda_2 y_2, \dots, \lambda_q y_q)$$

and

$$\|\lambda \odot y\|_q = \left(\sum_{j=1}^p |\lambda_j y_j|^q \right)^{\frac{1}{q}}.$$

Considering a $\lambda \in \mathbf{R}_{>}^p$, the set of best approximations of the ideal point measured by $\|\cdot\|_q$ is defined by

$$A(\lambda, q, Y) = \left\{ \bar{y} \in Y : \|\lambda \odot (\bar{y} - y^U)\|_q = \min_{y \in Y} \|\lambda \odot (y - y^U)\|_q \right\},$$

in which $Y = f(X)$. Now, the set of best approximations of y^U considering all positive weights is defined by

$$A(Y) = \bigcup_{\lambda \in \text{ri}(\Delta)} \bigcup_{1 \leq q < \infty} A(\lambda, q, Y),$$

where Δ stands for the standard simplex. The following result has been proved by Gearhart [30], and it has been addressed by Ehrgott [24] and Sawaragi et al. [65].

Theorem 2.4. *If Y is \mathbf{R}_{\geq}^m -closed, then $A(Y) \subseteq Y_{PN}$.*

Gearhart [30] used Theorem 2.4 to show that the set of compromise solutions is dense in the set of efficient points. Another proof for this theorem has been given by Ehrgott [24] and Sawaragi et al. [65]. The proof given in [24, 65] is completely different to that provided in [30], though the proof given in [24, 65] is not correct. In a part of the proof, the authors claimed that

$$\|\lambda \odot (\hat{\mathbf{y}} - \frac{d}{\beta_k} - y^U)\|_q < \|\lambda \odot (\hat{\mathbf{y}} - y^U)\|_q, \quad \forall k$$

implies

$$\lim_{k \rightarrow +\infty} \|\lambda \odot (\hat{\mathbf{y}} - \frac{d}{\beta_k} - y^U)\|_q < \|\lambda \odot (\hat{\mathbf{y}} - y^U)\|_q.$$

(See p. 188 of Ehrgott [24].) Clearly, this assertion is not correct. In fact, $<$ is converted to \leq after operating limit. This makes the given proof in [24, 65] incorrect.

In the following, we provide another proof for Theorem A. Furthermore, it is shown that the \mathbf{R}_{\geq}^m -closedness assumption is redundant and it is removed. This assumption has been considered in the three publications [24, 30, 65]. In fact, Theorem

A establishes that, under \mathbf{R}_{\geq}^m -closedness assumption, each compromise solution is properly efficient. In the following, we show that this holds in the general case.

Theorem 2.5. $A(Y) \subseteq Y_{PN}$.

Proof. Without loss of generality, we assume that $y^U = 0$. Thus, for every $y \in Y$, we have $y > 0$. Considering $\bar{y} \in A(Y)$, there are $\lambda \in \Lambda^0$ and $q \in [1, \infty)$ such that \bar{y} is an optimal solution (a minimizer) of the following optimization problem:

$$\min \sum_{i=1}^m \lambda_i^q y_i^q \quad (2.9a)$$

$$s.t. \ y \in Y. \quad (2.9b)$$

We define a real-valued function $f : \mathbf{R}_{\geq}^p \rightarrow \mathbf{R}$, by $f(y) = \sum_{i=1}^p \lambda_i^q y_i^q$. This function has the following properties on $\mathbf{R}_{>}^p$ (and immediately on Y):

- (i) $\nabla f(y) > 0, \ \forall y \in \mathbf{R}_{>}^p$,
- (ii) $f(y + y') \geq f(y) + f(y'), \ \forall y, y' \in \mathbf{R}_{>}^p$.

The proof of property (i) is straightforward. So, we prove only property (ii). For two positive scalars y_i and y'_i , we have $(y_i + y'_i)^q \geq (y_i)^q + (y'_i)^q$. Therefore,

$$\begin{aligned} & (\lambda_i)^q (y_i + y'_i)^q \geq (\lambda_i)^q (y_i)^q + (\lambda_i)^q (y'_i)^q \quad \forall i = 1, 2, \dots, m \\ \Rightarrow & \sum_{i=1}^m (\lambda_i)^q (y_i + y'_i)^q \geq \sum_{i=1}^m (\lambda_i)^q (y_i)^q + \sum_{i=1}^m (\lambda_i)^q (y'_i)^q \\ \Rightarrow & f(y + y') \geq f(y) + f(y'). \end{aligned}$$

Now, to the contrary, assume that $\bar{y} \notin Y_{PN}$. Thus, according to Benson's proper efficiency definition, there are sequences $\{\beta_k\} \subseteq (0, +\infty)$, $\{y_k\} \subseteq Y$ and $\{d_k\} \subseteq \mathbf{R}_{\geq}^p$ such that

$$\lim_{k \rightarrow \infty} \beta_k (y_k + d_k - \bar{y}) = -d, \quad (2.10)$$

for some nonzero $d \in \mathbf{R}_{\geq}^p$.

Note that $\{\beta_k\}$ is a sequence of positive real numbers. Therefore, this sequence is either bounded or unbounded. We consider these two possible cases and show that in both cases we have a contradiction.

Case 1. $\{\beta_k\}$ is bounded.

In this case, $\{\beta_k\}$ has a convergent subsequence. For simplicity, we denote this subsequence by $\{\beta_k\}$ again. So, $\beta_k \rightarrow \beta$, for a $\beta \geq 0$.

Assume that $\beta = 0$. There exists an i such that $d_i > 0$. Thus, by (2.10), there exists $k_0 \in \mathbf{N}$ such that

$$\forall k \geq k_0, \quad \beta_k(y_{ik} + d_{ik} - \bar{y}_i) < -\frac{d_i}{2}.$$

Hence, for each $k \geq k_0$, we get

$$y_{ik} < -\frac{d_i}{2\beta_k} + \bar{y}_i.$$

This implies $y_{ik} \rightarrow -\infty$. This yields a contradiction, because $y > 0$ for each $y \in Y$.

Now, assume that $\beta > 0$. According to (2.10), there exists a nonzero $d^0 \in \mathbf{R}_{\geq}^p$ such that

$$\lim_{k \rightarrow \infty} y_k + d_k - \bar{y} = -d^0. \quad (2.11)$$

Since $f(\frac{d^0}{2}) > 0$, there exists an $\bar{\epsilon} > 0$ such that

$$\forall \epsilon \in (0, \bar{\epsilon}], \quad f\left(\frac{d^0}{2}\right) > \left((1 + \epsilon)^q - 1\right) f(\bar{y}). \quad (2.12)$$

Due to $\bar{y} > 0$ and (2.11), there exists $k_1 \in \mathbf{N}$ such that for each $k \geq k_1$, we have

$$y_k + d_k - \bar{y} < -\frac{d^0}{2} + \bar{\epsilon}\bar{y}.$$

Hence, we get

$$0 < y_k + \frac{d^0}{2} < (1 + \bar{\epsilon})\bar{y},$$

which implies

$$f(y_k) + f\left(\frac{d^0}{2}\right) \leq f\left(y_k + \frac{d^0}{2}\right) < (1 + \bar{\epsilon})^q f(\bar{y}), \quad (2.13)$$

because of property (ii) and the fact that f is strictly increasing. Therefore, for each $k \geq k_1$, we get

$$f(y_k) < (1 + \bar{\epsilon})^q f(\bar{y}) - f\left(\frac{d^0}{2}\right) < f(\bar{y}),$$

because of (2.12). This contradicts the optimality of \bar{y} for the optimization problem (3.21).

Case 2. $\{\beta_k\}$ is unbounded.

Without loss of generality, we assume that $\beta_k \rightarrow \infty$. Thus, from (2.10), we have

$$\lim_{k \rightarrow \infty} y_k + d_k - \bar{y} = 0.$$

Since \bar{y} is an optimal solution of Problem (3.21) and $d_k \geq 0$, by property (ii), we have $f(y_k + d_k) \geq f(\bar{y})$. Furthermore, f is a continuously differentiable function on $\mathbf{R}_{>}^p$. Therefore, by the mean value theorem, we have

$$0 \leq f(y_k + d_k) - f(\bar{y}) = \nabla f(z_k)^T (y_k + d_k - \bar{y}),$$

in which $z_k = \alpha\bar{y} + (1 - \alpha)(y_k + d_k)$ for an $\alpha \in (0, 1)$. Thus,

$$\nabla f(z_k)^T \left(\beta_k (y_k + d_k - \bar{y}) \right) \geq 0. \quad (2.14)$$

Since $y_k + d_k \rightarrow \bar{y}$, we have $z_k \rightarrow \bar{y}$. Therefore,

$$0 \leq \lim_{k \rightarrow \infty} \nabla f(z_k)^T \left(\beta_k (y_k + d_k - \bar{y}) \right) = \nabla f(\bar{y})^T (-d).$$

This is a contradiction, because $\nabla f(\bar{y}) > 0$ and $d \in \mathbf{R}_{\geq}^p \setminus \{0\}$ gives $\nabla f(\bar{y})^T (-d) < 0$.

Hence, in both possible cases we had a contradiction, and the proof is complete. □

Corrolary 2.2. *If $f(x^0) \in A(Y)$, then the optimal value of Problem (2.2) is zero. Also, x^0 is an optimal solution of Problem (2.5).*

The converse of the above corollary is not correct necessarily, i.e., there exists some $x^0 \in X$ for which the optimal value of Problem (2.2) is zero, while $f(x^0) \notin A(Y)$. To show this, it is sufficient to consider an improperly efficient point. But if Y is \mathbf{R}_{\geq}^m -closed, there exists the utopia point and the optimal value of Problem (2.2) is zero, then $f(x^0)$ is the limit of some outcomes belonging to $A(Y)$. This results from the fact that, under the considered assumptions, the set of compromise solutions is dense in the set of efficient points; see [24, 30].

Chapter 3

Robustness

3.1 Introduction

Due to perturbations and partial knowledge, in most practical optimization problems we are faced with uncertainty. Popular approaches for dealing with uncertainty are stochastic optimization, robust optimization, stability and sensitivity analysis. Each approach has its own advantages. Stability and sensitivity analysis are used to analyze a solution set. In the first step, we neglect uncertainty in the model, and stability and sensitivity analysis tools are utilized after obtaining an optimal solution. They are also called post-optimal analysis techniques in the literature. We investigate robustness in the present chapter and stability and sensitivity analysis in the next chapter.

Unlike stability and sensitivity analysis, stochastic optimization and robust optimization entangle uncertainty in the first step. These approaches find solution(s) while uncertainty is involved in objective function and constraints. Consider the following uncertain multi-objective optimization problem:

$$\begin{aligned} \min & f(u, x) \\ \text{s.t.} & g(u, x) \leq 0 \\ & x \in X, u \in U, \end{aligned} \tag{3.1}$$

where x and u are decision variable and uncertain parameter, respectively. In addition, $f : U \times X \rightarrow \mathbf{R}^p$, $g : U \times X \rightarrow \mathbf{R}^m$ and $U \subseteq \mathbf{R}^q$. Moreover, $X \subseteq \mathbf{R}^n$ and $p \geq 2$. The set U is called an uncertain set and its members are called as scenarios. To handle the problem, we are faced with two challenges. First, which decision variable vector is admissible? Second, which admissible decision variable vector is efficient? One practical method to address these questions is to apply scalarization techniques in order to reduce the uncertain multi-objective problem to an uncertain single objective problem. Then, one can apply the rich theory of stochastic optimization and robust optimization. The reader is referred to great books [7, 66] for details about stochastic optimization and robust optimization in single objective optimization. However, we merely mention methods that sustain the multi-objective nature of the problem.

3.1.1 Stochastic multi-objective optimization

In stochastic optimization, data of Problem (3.1) are assumed to have a statistical structure. We mean that there are probability measure P and σ -algebra $F \subseteq 2^U$ such that (U, F, P) is a probability space. Moreover, for every $x \in X$, the functions $f_i(\cdot, x)$ and $g_i(\cdot, x)$ are measurable (random variables).

We review two intuitive ways for managing the problem called multi-objective method and stochastic method. Suppose that the feasible set is a fixed set \bar{X} ; for instance, $\bar{X} = \{x \in X : P(g_i(u, x) \leq 0) \geq 1 - \epsilon, \forall i\}$ for some tolerance $\epsilon \in (0, 1)$.

In multi-objective approach corresponding to each objective function f_i , a vector

$$(z_i^{(1)}(x), \dots, z_i^{(i_s)}(x)) := (\mathbf{H}^1(f_i(u, x)), \dots, \mathbf{H}^s(f_i(u, x))),$$

is defined. Each \mathbf{H}^i can be expectation value, risk measure, etc. According to the transformation, \bar{x} is called efficient if it is an efficient solution of the following multi-objective problem:

$$\begin{aligned} \min & (z_1^{(1)}(x), \dots, z_1^{(1_s)}(x), \dots, z_p^1(x), \dots, z_p^{(p_s)}(x)) \\ \text{s.t.} & x \in \bar{X}. \end{aligned}$$

Stochastic methods use some aggregation function $A : \mathbf{R}^p \rightarrow \mathbf{R}$ and transforms the stochastic multi-objective optimization to the following stochastic single objective optimization:

$$\begin{aligned} \min & A(f_1(u, x), \dots, f_p(u, x)) \\ \text{s.t.} & x \in \bar{X}, \end{aligned}$$

and manage this problem. More details and various stochastic methods can be seen in the two survey papers [1, 41].

3.1.2 Robustness: an introduction

Robust optimization has been introduced by Soyster [68] in 1973; However, it has attracted attention in the late 1990's. The works of Ben-tal, El Ghaoui and co-workers introduced robust optimization as a strong tool to handle uncertainty. Unlike stochastic optimization, the uncertain set U can be any arbitrary set. However, it is common to regard uncertain set as a compact set. The reader is referred to [7, 10] for more information.

Consider problem (3.1) with $p = 1$. A common approach to deal with it is to manage the optimization problem,

$$\begin{aligned} \min & (\sup_{u \in U} f(u, x)) \\ \text{s.t.} & g(u, x) \leq 0, \forall u \in U, \end{aligned} \tag{3.2}$$

called robust counterpart. As the problem signifies, we try to find a point that hedges against all scenarios and simultaneously minimizes the worst case of the objective function with respect to U .

The robust counterpart problem, in general, may be intractable. For example, consider the case that U is not finite, which is a semi-infinite programming problem. However, for some practical cases robust counterpart is manageable. For instance, in the linear case, when uncertain set is an ellipsoid or polyhedral the robust counter-

part will be conic programming or linear programming, respectively. Thus, robust counterpart enjoys not only nice properties, but also there are polynomial-time algorithms to solve it; see [7, 10] for more details. Semi-infinite programming tools have also been used to handle robust counterpart; see [36] and the references therein.

In the approach by robust counterpart we consider only feasible points that satisfy all constraints for all scenarios. In certain cases these feasible points hardly exist, and therefore the methods that disregard rare scenarios are often proposed to solve the problem (light robustness and some other notions go in this direction). We refer the reader to [38] and reference therein for more discussions on other robust approaches.

Robustness in single objective optimization has received considerable attention in the past decades, but robustness in multi-objective optimization has been considered seriously in recent years; see [19, 25, 35, 36]. This notion has been studied from different standpoints. In the rest of this section, we review some concepts of robustness in multi-objective optimization.

Deb and Gupta [19] introduced two definitions for robustness. In the first definition, they call an efficient solution robust if it is an efficient solution of the mean of all objective functions. In the second definition, the objectives do not change, but a constraint is added which restricts the absolute difference between the mean and the original objective values. There, the efficient solutions of the modified problem are called robust efficient.

Consider the following uncertain multi-objective optimization problem,

$$\begin{aligned} \min \quad & f(u, x) \\ \text{s.t.} \quad & x \in X, u \in U. \end{aligned}$$

where $f : U \times X \rightarrow \mathbf{R}^p$ is continuous. We only care about uncertainty in the objective function and the feasible set is not subject to uncertainty. Otherwise, we can specify a set X by the single objective robust optimization methods, such as hedging against all scenarios and so on. If U is singleton, then the problem reduces to a deterministic multi-objective problem. We denote the above problem by $P(u)$ for a given $u \in U$.

We follow the terms and the notations used in [44].

Definition 3.1. [44] *The feasible solution $\bar{x} \in X$ is called a flimsily robust (weakly) efficient solution if it is a (weakly) efficient solution of $P(u)$ for some $u \in U$.*

Definition 3.2. [44] *We say $\bar{x} \in X$ is highly robust (weakly) efficient if it is a (weakly) efficient solution of $P(u)$ for every $u \in U$.*

It is easy to see that highly robust (weakly) efficiency implies flimsily robust (weakly) efficiency. It is likely that highly robust efficient solution may not exist even for finite uncertain set U . The following example demonstrates this.

Example 3.1. *Let $X = \{x_1, x_2, x_3\}$ and $U = \{u_1, u_2\}$. Suppose that the values of $f : U \times X \rightarrow \mathbf{R}^2$ are given as follows:*

$f(u_1, x_1)$	$f(u_1, x_2)$	$f(u_1, x_3)$	$f(u_2, x_1)$	$f(u_2, x_2)$	$f(u_2, x_3)$
(1, 1)	(3.5, 2)	(3.2, 2.8)	(6, 1.5)	(2.6, 0.7)	(5.1, 1.2)

It is readily seen that x_1 and x_2 are flimsily robust efficient while there is no highly robust efficient solution (see Figure 3-1).

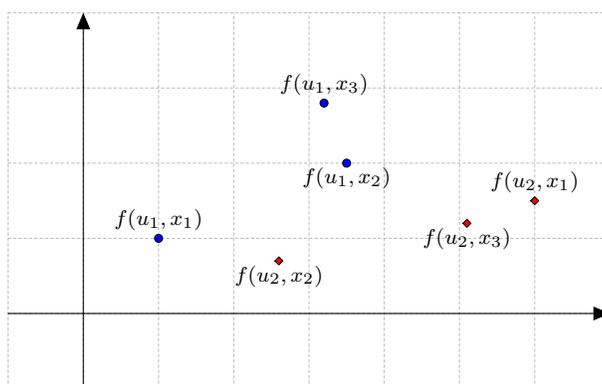


Figure 3-1 – Image of f in Example 3.1

Goberna et al. [35] consider uncertain linear multi-objective optimization, and provide some necessary and sufficient conditions for highly robust weak efficiency.

Ehrgott et al. [25] extended the worst case robustness notion from the single objective optimization to the multi-objective case based on the idea of set order.

Definition 3.3. *The decision variable $\bar{x} \in X$ is called a set-based minmax robust efficient solution if there is no $\hat{x} \in X$ such that*

$$f_U(\hat{x}) \subseteq f_U(\bar{x}) - \mathbf{R}_{\geq}^p,$$

where $f_U(x) = \{f(u, x) : u \in U\}$. Likewise, $\bar{x} \in X$ is called a set-based minmax robust weakly efficient solution if there is no $\hat{x} \in X$ such that

$$f_U(\hat{x}) \subseteq f_U(\bar{x}) - \mathbf{R}_{>}^p.$$

For the case $p = 1$ and under the compactness of U , these two concepts are equivalent to the nonexistence of $\hat{x} \in X$ with $\sup_{u \in U} f(u, \hat{x}) < \sup_{u \in U} f(u, \bar{x})$, which is exactly the definition of robustness in single objective optimization. In Example 3.1, all admissible points are set-based minmax robust efficient.

The following propositions, by taking advantage of two scalarization techniques, introduce some methods to obtain a set-based minmax robust (weakly) efficient solution. Their proofs can be found in [25].

Proposition 3.1. *Let \bar{x} be an optimal solution of the following problem:*

$$\begin{aligned} & \min(\sup_{u \in U} \langle \lambda, f(u, x) \rangle) \\ & \text{s.t. } x \in X, \end{aligned}$$

where $\lambda \in \mathbf{R}_{>}^p$ ($\lambda \in \mathbf{R}_{\geq}^p$). Then \bar{x} is a set-based minmax robust (weakly) efficient solution.

Proposition 3.2. *suppose that \bar{x} is an optimal solution of the following problem:*

$$\begin{aligned} & \min(\sup_{u \in U} f_i(u, x)) \\ & \text{s.t. } f_j(u, x) \leq \epsilon_j \quad j \neq i \quad \forall u \in U, \\ & x \in X, \end{aligned}$$

where $\epsilon \in \mathbf{R}^p$. Then \bar{x} is a set-based minmax robust weakly efficient solution. In addition, if \bar{x} is the unique optimal solution, then it is a set-based minmax robust efficient solution.

Unlike deterministic multi-objective optimization, the aforementioned methods cannot produce all set-based minmax robust (weakly) efficient solutions even for the convex case; see [25] for some counterexamples.

In general there is no relationship between set-based minmax robust efficiency and flimsily robust efficiency. Strictly speaking, a feasible point may be set-based minmax robust efficient while it is not flimsily robust efficient, and vice versa. Similarly highly robust efficiency and set-based minmax robust efficiency have no connection. To clarify this point, we take another look at Example 3.1. x_3 is a set-based minmax robust efficient point while it is neither flimsily robust efficient nor highly robust efficient.

Based on the idea of set-based minmax robust efficiency, the concept of hull-based minmax robust efficiency has been proposed by Bokrantz et al. [11].

Definition 3.4. *The decision variable $\bar{x} \in X$ is called hull-based minmax robust efficient if there is no $\hat{x} \in X$ such that*

$$f_U(\hat{x}) \subseteq \text{co}(f_U(\bar{x})) - \mathbf{R}_{\geq}^p.$$

It is established that hull-based minmax robust efficiency implies set-based minmax robust efficiency [11]. However, the converse does not hold necessarily; for instance, in Example 3.1, x_3 is set-based minmax robust efficient but not hull-based minmax robust efficient. Another notion suggested for robustness in multi-objective optimization is point-based minmax robust efficiency, which deals with efficient solutions of a deterministic multi-objective optimization problem for which each objective function is worst case of the original objective on the uncertain set.

Definition 3.5. [28, 50] If \bar{x} is an efficient solution of the following problem:

$$\begin{aligned} \min f_U^{\max}(x) \\ x \in X, \end{aligned}$$

where $f_U^{\max}(x) := \begin{bmatrix} \sup_{u \in U} f_1(u, x) \\ \vdots \\ \sup_{u \in U} f_p(u, x) \end{bmatrix}$, then it is called a point-based minmax robust efficient solution.

This concept was introduced by Kuroiwa et al. [50] and Fliege et al. [28] almost simultaneously. Moreover, Fliege and Werner [28] applied it to portfolio optimization.

Note that if $\sup_{u \in U} f_i(u, x) = \max_{u \in U} f_i(u, x)$ for all $x \in X$ and each $i \in \{1, \dots, p\}$, then every point-based minmax weakly robust efficient solution is set-based minmax robust weakly efficient; see Theorem 4.11 of [25]. The converse does not hold in general. However, Ehrgott et al. [25] pointed out that if $U = U_1 \times \dots \times U_p$ and the objective functions f_1, \dots, f_p are independent of each other with respect to the uncertain set, namely if $f((u_1, \dots, u_p), x) := \begin{bmatrix} f_1(u_1, x) \\ \vdots \\ f_p(u_p, x) \end{bmatrix}$, then point-based minmax robust (weak) efficiency is equivalent to set-based minmax robust (weak) efficiency.

Other concepts have also been proposed for robustness; we listed here some of them. For example, on account of the definition of set-based minmax robust efficiency Ide et al. [43, 47] introduced order-robust definition by using orders for set. Schöbel et al. [44] extended light robustness to multi-objective optimization. Pourkarimi et al. introduced two new concepts for robustness in linear multi-objective optimization [61]. We refer reader to the survey paper [44].

In a recent work, Georgiev et al. [32] have defined the robustness for linear multi-objective optimization problems from a different point of view. They considered a perturbation standpoint, and defined an efficient solution as a robust solution if it remained efficient for small perturbations of the coefficients of the objective functions.

They also studied their definition considering different kinds of perturbations, including changing the objectives' coefficients and adding a new objective function. They obtained necessary and sufficient conditions and presented various nice properties of the robust solutions in the linear case. Goberna et al. [35,36] extended Georgiev et al.'s definition for linear multi-objective optimization problems under perturbations of the coefficients of both the objective functions and constraints.

In the rest of our work here, we extend the definition given by Georgiev et al. to nonlinear multi-objective optimization. We show that, under the compactness of the feasible set or convexity, the set of robust efficient solutions is a subset of the set of properly efficient solutions. Some necessary and sufficient conditions for robust solutions with respect to the Bouligand tangent cone and non-ascent directions, under appropriate assumptions, are given. A robustness radius is calculated. The relationships between the robustness notion considered in this sense and some mentioned definitions are highlighted. Two kinds of modifications in the objective functions are dealt with and the relationships between the weak/proper/robust efficient solutions of the problems, before and after the perturbation, are established. Some examples, to clarify the theoretical results, are given.

In Section 3.2, robustness is defined, its relationship with proper efficiency is established, and some necessary and sufficient conditions are proved. Section 3.3 is devoted to the robustness radius calculation. Section 3.4 contains some results on connections between the newly and previously defined robustness definitions. In Section 3.5, we study some alterations of the objective functions that preserve weak/proper/robust efficiency.

3.2 Robustness

We start this section by introducing the concept of robust solution in nonlinear multi-objective optimization. This definition extends Definition 3.1 in Georgiev et al. [32].

Consider the following multi-objective optimization problem:

$$\begin{aligned} \min f(x) \\ \text{s.t. } x \in X, \end{aligned} \tag{3.3}$$

where $X \subseteq \mathbf{R}^n$ is nonempty and $f : X \rightarrow \mathbf{R}^p$ is locally Lipschitz.

Definition 3.6. *Let $\bar{x} \in X$ be an efficient solution of Problem (3.3). \bar{x} is called a robust efficient solution if there exists $\epsilon > 0$ such that for any $p \times n$ matrix C with $\|C\| < \epsilon$, the vector \bar{x} is an efficient solution of*

$$\begin{aligned} \min f(x) + Cx \\ \text{s.t. } x \in X. \end{aligned}$$

Note that the notation $\|\cdot\|$ stands for the Frobenius norm in matrix spaces. However, since all the norms are equivalent in finite dimension, the definition of robust is independent of the choice of norm. In the above definition, ϵ is called radius of robustness. Thanks to mean value Theorem 1.5, we have the following result.

Proposition 3.3. *Let \bar{x} be a robust efficient solution with radius ϵ . Then for any Lipschitz function $h : \mathbf{R}^n \rightarrow \mathbf{R}^p$ with modulus less than ϵ , \bar{x} is an efficient solution of the following problem:*

$$\begin{aligned} \min f(x) + h(x) \\ \text{s.t. } x \in X. \end{aligned}$$

The following theorem presents a nice property of robust efficient solutions. It states that the set of robust efficient solutions is a subset of properly efficient solutions under the compactness of the feasible set.

Theorem 3.1. *Let X be compact. If \bar{x} is a robust efficient solution of Problem (3.3), then \bar{x} is a proper efficient solution of Problem (3.3).*

Proof. Suppose \bar{x} is not a properly efficient solution. Then, there exist $\{x_i\} \subseteq X$, increasing sequence $\{M_i\}$ of positive real numbers, and $k \in \{1, \dots, p\}$, such that

$M_i \rightarrow +\infty$,

$$f_k(x_i) < f_k(\bar{x}), \quad \forall i, \quad (3.4)$$

and

$$\frac{f_k(\bar{x}) - f_k(x_i)}{f_j(x_i) - f_j(\bar{x})} > M_i, \quad \text{for each } j \in \{1, \dots, p\} \text{ with } f_j(x_i) > f_j(\bar{x}). \quad (3.5)$$

Since X is compact, without loss of generality, we may assume that $\{x_i\}$ converges to some $\hat{x} \in X$. Also, we define $Q_i = \{j : f_j(x_i) > f_j(\bar{x})\}$. This set is nonempty because \bar{x} is efficient. Without loss of generality, by choosing an appropriate subsequence, Q_i is a constant set for all i indices. So, we denote it by Q . Two cases may occur for \hat{x} ; either it is equal to \bar{x} or not. We consider these two possible cases and get a contradiction in each case.

Due to the robustness of \bar{x} , there exists some $\epsilon > 0$ such that \bar{x} is an efficient solution of Problem (3.4) for any matrix $C_{p \times n}$ with $\|C\| < \epsilon$. Let $\hat{x} \neq \bar{x}$. We can choose the matrix with property $\|\tilde{C}_{p \times n}\| < \epsilon$ such that

$$\tilde{C}^j(\hat{x} - \bar{x}) < -2\delta, \quad \forall j \in Q, \quad (3.6)$$

$$\tilde{C}^j = 0, \quad \forall j \in \{1, \dots, p\} \setminus Q, \quad (3.7)$$

for some $\delta > 0$ (\tilde{C}^j denotes the j th row of \tilde{C}). Since f is bounded on X , from (3.5), we have $f_j(x_i) \rightarrow f_j(\bar{x})$ for each $j \in Q$ as $i \rightarrow +\infty$. Therefore, for sufficiently large i values, we have $f_j(x_i) - f_j(\bar{x}) - \delta < 0$. Hence, by (3.6), for sufficiently large i values, we get

$$f_j(x_i) + \tilde{C}^j x_i < f_j(\bar{x}) + \tilde{C}^j \bar{x} - \delta < f_j(\bar{x}) + \tilde{C}^j \bar{x}, \quad \forall j \in Q. \quad (3.8)$$

Also, by (3.7) and due to the definition of Q , for sufficiency large i values, we have

$$f_j(x_i) + \tilde{C}^j x_i \leq f_j(\bar{x}) + \tilde{C}^j \bar{x}, \quad \forall j \in \{1, \dots, p\} \setminus Q. \quad (3.9)$$

Inequalities (3.8) and (3.9) contradict the robustness of \bar{x} .

Now, we consider the latter case, $\bar{x} = \hat{x}$. We assume that the sequence $\left\{ \frac{x_i - \bar{x}}{\|x_i - \bar{x}\|} \right\}$

converges to some nonzero vector d . We choose $\tilde{C}_{p \times n}$ satisfying $\|\tilde{C}_{p \times n}\| < \epsilon$ and

$$\tilde{C}^j d < -2\delta, \quad \forall j \in Q, \quad (3.10)$$

$$\tilde{C}^j = 0, \quad \forall j \in \{1, \dots, p\} \setminus Q, \quad (3.11)$$

for some $\delta > 0$. Assume that L_j is the Lipschitz constant of f_j on a neighborhood of \bar{x} . By (3.5), for sufficiently large i values, we get

$$f_j(x_i) - f_j(\bar{x}) < \frac{L_k \|\bar{x} - x_i\|}{M_i} < \delta \|\bar{x} - x_i\|. \quad (3.12)$$

Therefore, by (3.10)-(3.12) we get inequalities (3.8) and (3.9) in this case as well. These contradict the robustness of \bar{x} and the proof is complete. \square

The converse of the above theorem does not hold necessarily, even for the linear case; see Example 3.2 of [32].

The following example shows that the compactness assumption of X in Theorem 3.1 is essential.

Example 3.2. Consider the multi-objective optimization problem,

$$\begin{aligned} & \min (-x, x^3) \\ & \text{s.t. } x \in \mathbf{R}. \end{aligned}$$

It is not difficult to see that $\bar{x} = 1$ is a robust efficient solution (consider $\epsilon = 0.1$), while the problem does not have any properly efficient solution.

Now, we are going to provide a characterization of robust efficient solutions with respect to the non-ascent directions of the objective function and the Bouligand tangent cone of the feasible set.

Definition 3.7. The vector $d \in \mathbf{R}^n$ is called a non-ascent direction of f at \bar{x} if $d^T \xi \leq 0$, for each $\xi \in \partial f_i(\bar{x})$ and each $i \in \{1, 2, \dots, p\}$. $G(\bar{x})$ denotes the set of all non-ascent directions of f at \bar{x} .

The following theorem presents a necessary condition for robustness.

Theorem 3.2. *If \bar{x} is a robust efficient solution of Problem (3.3), then $T_X(\bar{x}) \cap G(\bar{x}) = \{0\}$.*

Proof. We argue by giving a contradiction. Suppose that $0 \neq d \in G(\bar{x}) \cap T_X(\bar{x})$. By robustness of \bar{x} , there exists an $\epsilon > 0$ such that \bar{x} is an efficient solution of problem (3.4) for any matrix $C_{p \times n}$, with $\|C\| < \epsilon$. We choose a matrix $\tilde{C}_{p \times n}$ such that

$$\|\tilde{C}\| < \epsilon \quad \text{and} \quad \tilde{C}d < -2\delta e \quad (3.13)$$

for some $\delta > 0$ (e is a column vector with all components being equal to one). Since $d \in T_X(\bar{x})$,

$$\exists(\{x_i\} \subseteq X, t_i \downarrow 0); \frac{x_i - \bar{x}}{t_i} \rightarrow d. \quad (3.14)$$

Therefore, from (3.13) and (3.14), for sufficiently large i , we have

$$\tilde{C}\left(\frac{x_i - \bar{x}}{t_i}\right) < -\delta e, \quad (3.15)$$

which implies $\tilde{C}x_i + t_i\delta e < \tilde{C}\bar{x}$. Using the mean value theorem (Theorem 1.5), for each i , we have

$$f(x_i) = f(\bar{x}) + \xi_i^T(x_i - \bar{x}), \quad (3.16)$$

where ξ_i is an $n \times p$ matrix whose j th column belongs to $\partial f_j(\tilde{x}_i^j)$, for some $\tilde{x}_i^j \in (\bar{x}, x_i)$.

Thus,

$$\begin{aligned} f(x_i) + \tilde{C}x_i + t_i\delta e &< f(\bar{x}) + \xi_i^T(x_i - \bar{x}) + \tilde{C}\bar{x} \\ \Rightarrow f(x_i) + \tilde{C}x_i + t_i(\delta e - \xi_i^T\left(\frac{x_i - \bar{x}}{t_i}\right)) &< f(\bar{x}) + \tilde{C}\bar{x}. \end{aligned}$$

Since $\tilde{x}_i^j \rightarrow \bar{x}$, as $i \rightarrow +\infty$, and f is locally Lipschitz at \bar{x} , by Proposition 1.3 the sequence $\{\xi_i\}$ is bounded. Hence, $\xi_i \rightarrow \xi$, for some $\xi \in \partial f(\bar{x})$, because of Proposition 1.3. Thus, $\xi^T d \leq 0$. Therefore, for sufficiently large i values, $\xi_i^T\left(\frac{x_i - \bar{x}}{t_i}\right) < \delta e$. Thus, we get

$$f(x_i) + \tilde{C}x_i < f(\bar{x}) + \tilde{C}\bar{x},$$

which contradicts the robustness of \bar{x} , and completes the proof. \square

The condition given in the above theorem is necessary for robustness and it is not sufficient, in general. The following example clarifies this.

Example 3.3. Consider the multi-objective optimization problem,

$$\begin{aligned} \min & (f_1(x), f_2(x)) \\ \text{s.t. } & x \in \mathbf{R}, \end{aligned}$$

where

$$\begin{aligned} f_1(x) &:= x, \\ f_2(x) &:= \begin{cases} -x & |x| < 1, \\ -x^{(\frac{1}{3})} & |x| \geq 1. \end{cases} \end{aligned}$$

Let $\bar{x} = 2$. At this point we have $T_X(\bar{x}) = \mathbf{R}$ and $G(\bar{x}) = \{0\}$. It is not difficult to see that $\bar{x} = 2$ is an efficient solution of the above problem, while for any $\epsilon > 0$ it is not an efficient solution of

$$\begin{aligned} \min & (f_1(x), f_2(x) + \frac{\epsilon}{2}x) \\ \text{s.t. } & x \in \mathbf{R}, \end{aligned}$$

because for each $\epsilon > 0$, by setting $x_\epsilon = \min\{-125, \frac{-1}{\epsilon^3}\}$, we have $f_1(x_\epsilon) < f_1(2)$ and $f_2(x_\epsilon) \leq f_2(2)$.

As shown by the above example, the necessary condition given in Theorem 3.2 may not be sufficient for robustness, in general. Theorem 3.3 establishes that this condition is sufficient under convexity assumption.

Theorem 3.3. Let X be a closed and convex set and f_i ($i = 1, \dots, p$) be convex. Assume that \bar{x} is an efficient solution of Problem (3.3). \bar{x} is a robust efficient solution of Problem (3.3) if and only if $T_X(\bar{x}) \cap G(\bar{x}) = \{0\}$.

Proof. The “only if” part is derived from Theorem 3.2. For “if” part, suppose that \bar{x} is not a robust efficient solution. Thus, there exist a sequence $\{C_i\}$ of $p \times n$

matrices and a sequence $\{x_i\} \subseteq X$ such that $C_i \rightarrow 0$, and

$$f(x_i) + C_i x_i \leq f(\bar{x}) + C_i \bar{x}. \quad (3.17)$$

Set

$$d_i := \frac{x_i - \bar{x}}{\|x_i - \bar{x}\|}. \quad (3.18)$$

Two cases may occur for the sequence $\{x_i\}$. Either it has a subsequence convergent to \bar{x} or it does not have any subsequence convergent to \bar{x} . We consider these two possible cases and we get a contradiction in each case.

In the first case, without loss of generality, we assume that $x_i \rightarrow \bar{x}$. From the convexity of f , for any $\xi \in \partial f(\bar{x})$, we have,

$$f(x_i) \geq f(\bar{x}) + \xi^T(x_i - \bar{x}), \quad (3.19)$$

with ξ being an $n \times p$ matrix whose j th column belongs to $\partial f_j(\bar{x})$. Therefore, due to (3.17), we have

$$\|x_i - \bar{x}\|^{-1} (\xi^T(x_i - \bar{x}) + C_i(x_i - \bar{x})) \leq 0. \quad (3.20)$$

Without loss of generality, we can assume that $d_i \rightarrow d$, for some $d \in \mathbf{R}^p$ with $\|d\| = 1$ and it is obvious that $d \in T_X(\bar{x})$. Moreover, from (3.20) we conclude that $d \in G(\bar{x})$. Thus $d \in G(\bar{x}) \cap T_X(\bar{x})$. This gives a contradiction.

Now, we consider the other case: $\{x_i\}$ does not have any subsequence convergent to \bar{x} . Therefore, without loss of generality, there exists a $r > 0$ such that $\|x_i - \bar{x}\| > r$. On the other hand, $d_i \rightarrow d$ for some nonzero $d \in T_X(\bar{x})$. Since X is convex and closed, for each i , we have

$$\begin{aligned} td_i + \bar{x} &\in X, \quad \forall t \in [0, r], \\ td + \bar{x} &\in X, \quad \forall t \in [0, r]. \end{aligned}$$

Thus, $0 \neq d \in T_X(\bar{x})$. Suppose that $\{t_i\}$ is a sequence of scalars in $[0, r]$ that converges

to zero. By convexity of f and due to (3.17) and (3.18), we get

$$\begin{aligned} f(\bar{x} + t_i d_i) &\leq \left(1 - \frac{t_i}{\|x_i - \bar{x}\|}\right) f(\bar{x}) + \frac{t_i}{\|x_i - \bar{x}\|} f(x_i) \\ &\leq f(\bar{x}) + \frac{t_i}{\|x_i - \bar{x}\|} C_i (\bar{x} - x_i). \end{aligned}$$

Since $C_i \rightarrow 0$, from the convexity of f and the above statement, we have $\xi^T d \leq 0$, with ξ being an $n \times p$ matrix whose j th column belongs to $\partial f_j(\bar{x})$. Therefore, $0 \neq d \in G(\bar{x}) \cap T_X(\bar{x})$. This gives a contradiction and completes the proof. \square

In the rest of this section, we consider a multi-objective optimization problem whose feasible set is defined by some constraint functions. Consider

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i = 1, 2, \dots, m, \end{aligned} \tag{3.21}$$

where $f : \mathbf{R}^n \rightarrow \mathbf{R}^p$ is the objective function (i.e., $f(x) = (f_1(x), \dots, f_p(x))$) and the g_j functions define the constraints. Hereafter, whenever we use the Clarke subdifferential for the g_j functions, we assume that these functions are locally Lipschitz.

For a feasible point \bar{x} , the index set $I(\bar{x})$ is defined by

$$I(\bar{x}) = \{j \in \{1, 2, \dots, m\} : g_j(\bar{x}) = 0\}.$$

In the following, we are going to provide a characterization of robust efficient solutions of Problem (3.21). The following constraint qualification (CQ) helps us in the sequel.

Definition 3.8. *We say that constraint qualification (CQ) holds at \bar{x} , if*

$$0 \notin \text{co} \left\{ \bigcup_{j \in I(\bar{x})} \partial g_j(\bar{x}) \right\}.$$

Theorem 3.4. *If \bar{x} is a robust efficient solution which satisfies (CQ), then*

$$\text{pos} \left(\bigcup_{i=1}^p \partial f_i(\bar{x}) \right) + \text{pos} \left(\bigcup_{i \in I(\bar{x})} \partial g_i(\bar{x}) \right) = \mathbf{R}^n.$$

Proof. For simplicity, we set $I_{\bar{x}} = \text{pos}\left(\bigcup_{i=1}^p \partial f_i(\bar{x})\right) + \text{pos}\left(\bigcup_{i \in I(\bar{x})} \partial g_i(\bar{x})\right)$. It can be seen that under the assumptions of the theorem and Theorem 1.8,

$$\{d : g_i^\circ(\bar{x}; d) \leq 0, \forall i \in I(\bar{x})\} \subseteq T_X(\bar{x}).$$

Therefore, according to Theorem 3.2, the system below has no solution $d \in \mathbf{R}^n$:

$$\begin{aligned} \xi^T d &\leq 0, \quad \forall \xi \in \partial f_i(\bar{x}), \forall i \in \{1, \dots, p\} \\ \xi^T d &\leq 0, \quad \forall \xi \in \partial g_i(\bar{x}), \forall i \in I(\bar{x}) \\ d &\neq 0. \end{aligned}$$

Hence, the following system has no solution $d \in \mathbf{R}^n$:

$$\begin{aligned} \xi^T d &\leq 0, \quad \forall \xi \in \partial f_i(\bar{x}), \forall i \in \{1, \dots, p\} \\ \xi^T d &\leq 0, \quad \forall \xi \in \partial g_i(\bar{x}), \forall i \in I(\bar{x}) \\ d_1 &> 0. \end{aligned}$$

Using the semi-infinite Farkas Theorem 1.9, we have $e_1 \in \text{cl}(I_{\bar{x}})$. Similarly, it can be shown that $\pm e_i \in \text{cl}(I_{\bar{x}})$ for each $i \in \{1, 2, \dots, p\}$. Here, e_i denotes the i th unit vector. Therefore, $\text{cl}(I_{\bar{x}}) = \mathbf{R}^n$. Since $A_{\bar{x}}$ is a convex set whose closure is equal to \mathbf{R}^n , we have $I_{\bar{x}} = \mathbf{R}^n$ and the proof is complete. \square

Corrolary 3.1. *Assume that f_i ($i = 1, \dots, p$) and g_j ($j = 1, \dots, m$) in Problem (3.21) are continuously differentiable. If \bar{x} is a robust efficient solution of Problem (3.21) which satisfies (CQ), then*

$$\text{pos}\{\nabla f_1(\bar{x}), \dots, \nabla f_p(\bar{x})\} + \text{pos}\{\nabla g_i(\bar{x}) : i \in I(\bar{x})\} = \mathbf{R}^n.$$

Theorem 3.5 provides a converse version of Theorem 3.4. Theorems 3.4 and 3.5 extend Theorem 3.4 of Georgiev et al. [32].

Theorem 3.5. *Let f_i ($i = 1, \dots, p$) and g_j ($j = 1, \dots, m$) in Problem (3.21) be*

convex. If \bar{x} is an efficient solution and

$$\text{pos}\left(\bigcup_{i=1}^p \partial f_i(\bar{x})\right) + \text{pos}\left(\bigcup_{i \in A(\bar{x})} \partial g_i(\bar{x})\right) = \mathbf{R}^n,$$

then \bar{x} is a robust efficient solution of Problem (3.21).

Proof. We prove the theorem by contradiction. Suppose that \bar{x} is not robust. Then, according to Theorem 3.3, there exists a nonzero vector $\bar{d} \in T_X(\bar{x}) \cap G(\bar{x})$. From the convexity assumption, we get $\xi^T \bar{d} \leq 0$, for each $\xi \in \text{pos}(\partial g_i(\bar{x}))$ and each $i \in I(\bar{x})$. Also, $\xi^T \bar{d} \leq 0$, for each $\xi \in \text{pos}(\partial f_i(\bar{x}))$ and each $i \in \{1, 2, \dots, p\}$, because of $\bar{d} \in G(\bar{x})$. On the other hand, by the assumption of the theorem, $\bar{d} = \sum_{i=1}^p u_i \xi_i + \sum_{j \in I(\bar{x})} v_j \zeta_j$, for some $u_i, v_j \geq 0$, $\xi_i \in \text{pos}(\partial f_i(\bar{x}))$, and $\zeta_j \in \text{pos}(\partial g_j(\bar{x}))$. Therefore, $\bar{d}^T \bar{d} \leq 0$. Hence, we get $\bar{d} = 0$ which gives a contradiction. \square

Corrolary 3.2. Assume that f_i ($i = 1, \dots, p$) and g_j ($j = 1, \dots, m$) in Problem (3.21) are differentiable and convex. If \bar{x} is an efficient solution and

$$\text{pos}\{\nabla f_1(\bar{x}), \dots, \nabla f_p(\bar{x})\} + \text{pos}\{\nabla g_i(\bar{x}) : i \in I(\bar{x})\} = \mathbf{R}^n,$$

then \bar{x} is a robust efficient solution of Problem (3.21).

Although the compactness assumption is essential in Theorem 3.1 (see Example 3.2), the following result shows that Theorem 3.1 remains valid without compactness of the feasible set for convex programming problems.

Theorem 3.6. Let f_i ($i = 1, 2, \dots, p$) and g_j ($j = 1, 2, \dots, m$) be convex in Problem (3.21). If \bar{x} is a robust efficient solution of Problem (3.21), then \bar{x} is a properly efficient solution of Problem (3.21).

Proof. Suppose that \bar{x} is not a properly efficient solution. Therefore, there exist $\{x_i\} \subseteq X$, increasing sequence $\{M_i\}$ of positive real numbers, and $k \in \{1, \dots, p\}$, such that $M_i \rightarrow +\infty$,

$$f_k(x_i) < f_k(\bar{x}) \quad \forall i, \tag{3.22}$$

and

$$\frac{f_k(\bar{x}) - f_k(x_i)}{f_j(x_i) - f_j(\bar{x})} > M_i, \text{ for each } j \in \{1, \dots, p\} \text{ with } f_j(x_i) > f_j(\bar{x}). \quad (3.23)$$

Define $Q_i = \{j : f_j(x_i) > f_j(\bar{x})\}$. This set is nonempty because \bar{x} is efficient. Without loss of generality, by choosing an appropriate subsequence, Q_i is a constant set for all i indices. So, we denote it by Q . Also, define the feasible set of Problem (3.21) by $X = \{x \in \mathbf{R}^n : g_j(x) \leq 0, j = 1, 2, \dots, m\}$. Without loss of generality, we assume that the sequence $\{\frac{x_i - \bar{x}}{\|x_i - \bar{x}\|}\}$ converges to some nonzero vector d . Setting $t_i = \min\{\frac{1}{i}, \|x_i - \bar{x}\|\}$ and $d_i = \frac{x_i - \bar{x}}{\|x_i - \bar{x}\|}$, we have $t_i \downarrow 0$ and $\bar{x} + t_i d_i \in X$, according to the convexity assumptions. Hence, $d \in T_X(\bar{x})$. Due to the convexity assumption, we get

$$\begin{aligned} f_j(x_i) &\geq f_j(\bar{x}) + \xi^T(x_i - \bar{x}), \quad \forall \xi \in \partial f_j(\bar{x}), \quad \forall j \in \{1, \dots, p\} \setminus Q, \\ \Rightarrow \quad \xi^T(x_i - \bar{x}) &\leq 0, \quad \forall \xi \in \partial f_j(\bar{x}), \quad \forall j \in \{1, \dots, p\} \setminus Q, \\ \Rightarrow \quad \xi^T d &\leq 0, \quad \forall \xi \in \partial f_j(\bar{x}), \quad \forall j \in \{1, \dots, p\} \setminus Q. \end{aligned}$$

Moreover, from (3.23) and the convexity of the objective functions, we have

$$\begin{aligned} \xi^T(x_i - \bar{x}) &\leq f_j(x_i) - f_j(\bar{x}), \quad \forall \xi \in \partial f_j(\bar{x}), \quad \forall j \in Q, \\ &< \frac{f_k(\bar{x}) - f_k(x_i)}{M_i}, \quad \forall \xi \in \partial f_j(\bar{x}), \quad \forall j \in Q, \\ &\leq \frac{1}{M_i} \eta^T(\bar{x} - x_i), \quad \forall \eta \in \partial f_k(\bar{x}). \end{aligned}$$

Thus,

$$\xi^T d \leq 0, \quad \forall \xi \in \partial f_j(\bar{x}), \quad \forall j \in Q.$$

Therefore, $d \in T_X(\bar{x}) \cap G(\bar{x})$. This is a contradiction because of Theorem 3.2, and the proof is complete. \square

Remark 3.1. *The robust solution studied in the present chapter may not exist in some special cases, though these solutions (if exist) have nice properties as compared to non-robust points. An efficient point is robust if it stays efficient under small*

linear perturbations. Let us assume that the f_i and g_j functions are differentiable here. Under some CQs and appropriate assumptions, the KKT/FJ condition

$$\sum_{i=1}^p \lambda_i \nabla f_i(\bar{x}) + \sum_{j \in I(\bar{x})} \mu_j \nabla g_j(\bar{x}) = 0$$

for some nonnegative μ_j 's and some nonnegative λ_i 's (not all zero), is necessary for the efficiency of \bar{x} . If some objective function, say f_1 , is perturbed, then $\nabla f_1(\bar{x})$ is alerted and hence to preserve the KKT/FJ condition (efficiency of \bar{x}), the Lagrange multiplier(s) of some other objective function(s) or some constraint function(s) should be changed. Hence, at least one other objective function or at least one constraint function is required for robustness, i.e., $m + p \geq 2$. Thus, there is no robust solution for unconstrained single objective problems. To show this analytically, let \bar{x} be an arbitrary optimal solution of $\min_{x \in \mathbf{R}^n} h(x)$, where $h : \mathbf{R}^n \rightarrow \mathbf{R}$. Then $\nabla h(\bar{x}) = 0$ which implies $\nabla h(\bar{x}) + C \neq 0$ for each $C \neq 0$. Therefore, \bar{x} is not optimal for $\min_{x \in \mathbf{R}^n} h(x) + Cx$ for each $C \neq 0$. Hence, \bar{x} is not robust for $\min_{x \in \mathbf{R}^n} h(x)$. Thus, this unconstrained problem does not have any robust solution.

Now, consider an unconstrained multi-objective programming problem $\min_{x \in \mathbf{R}^n} f(x)$, with $f : \mathbf{R}^n \rightarrow \mathbf{R}^p$ and $p \geq 2$. Here, $m = 0$. If \bar{x} is a robust solution, then by Corollary 3.1, $\text{pos}\{\nabla f_1(\bar{x}), \dots, \nabla f_p(\bar{x})\} = \mathbf{R}^n$, and hence $p \geq n + 1$.

For constrained problem (3.21) satisfying the assumptions of Corollary 3.1, if \bar{x} is a robust solution, then $p + m \geq n + 1$. This is not restrictive for practical cases, because in practice the problem has at least $2n$ constraints due to the lower and upper bounds on variables.

Remark 3.2. The necessary condition presented in Theorem 3.2 provides a tie-in to the gradient-like descent methods existing in the literature for solving vector optimization problems; see [22] and [28]. Extending these numerical tools to generate robust solution(s) can be worth studying in future.

3.3 Robustness radius

In this short section, we compute a radius of robustness. For a given vector $a \in \mathbf{R}^p$, the vector a^+ is obtained from a by substituting all negative components by zero. It is not difficult to show that $\|a^+\|$ is equal to the distance from a to $-\mathbf{R}_{\geq}^p = \{x \in \mathbf{R}^p : x \leq 0\}$.

Lemma 3.1. *Let X be a closed and convex set and f_i ($i = 1, \dots, p$) be convex. Let $d \in T_X(\bar{x})$ with $\|d\| = 1$. If \bar{x} is a robust solution of Problem (3.3), then $\|(f'(\bar{x}; d))^+\| > 0$ and it is equal to the optimal value of the following problem,*

$$\sup\{t : f'(\bar{x}; d) + tCd \notin -\mathbf{R}_{\geq}^p, \quad \forall \|C\| \leq 1\}.$$

Proof. First, we show that $f'(\bar{x}; d) \notin -\mathbf{R}_{\geq}^p$. If $f'(\bar{x}; d) \leq 0$, then due to the convexity of f , we have $d \in G(\bar{x})$, which gives a contradiction according to Theorem 3.2. The proof of the second part is similar to that of Lemma 4.2 of [32]. \square

Theorem 3.7. *Under the assumptions of Lemma 3.1, the optimal value of the following problem is positive and it is a robustness radius for \bar{x} :*

$$\begin{aligned} & \min \|(f'(\bar{x}; d))^+\| \\ & \text{s.t. } d \in T_X(\bar{x}), \\ & \|d\| = 1. \end{aligned}$$

Proof. Let ρ be the optimal value of the given problem. Thus, by Lemma 3.1, $\rho > 0$. Now, we show that ρ is a robustness radius for \bar{x} . If it is not a robustness radius, then there exist some $x^o \in X$ and some matrix C^o such that $\|C^o\| < \rho$, and

$$f(x^o) + C^o x^o \leq f(\bar{x}) + C^o \bar{x}. \quad (3.24)$$

Setting $d^o = \frac{x^o - \bar{x}}{\|x^o - \bar{x}\|}$, we have $\|d^o\| = 1$ and $d^o \in T_X(\bar{x})$ due to the convexity of X . Furthermore, by convexity of f , we get

$$f'(\bar{x}; d^o) + C^o d^o = \frac{f'(\bar{x}; x^o - \bar{x})}{\|x^o - \bar{x}\|} + C^o d^o \leq \frac{f(x^o) - f(\bar{x})}{\|x^o - \bar{x}\|} + \frac{C^o(x^o - \bar{x})}{\|x^o - \bar{x}\|}.$$

Therefore, according to (3.24), we get

$$f'(\bar{x}; d^o) + C^o d^o \in -\mathbf{R}_{\geq}^p. \quad (3.25)$$

Defining

$$\rho^o = \sup\{t : f'(\bar{x}; d^o) + tC^o d^o \notin -\mathbf{R}_{\geq}^p, \quad \forall \|C\| \leq 1\}, \quad (3.26)$$

we have $\rho \leq \rho^o$. Furthermore, for each $t \in (0, \rho^o)$ and each C with $\|C\| \leq 1$, we have $f'(\bar{x}; d^o) + tC^o d^o \notin -\mathbf{R}_{\geq}^p$. This is a contradiction with (3.25) by setting $t = \|C^o\|$ and $C = \frac{C^o}{\|C^o\|}$, and the proof is complete. \square

It can be seen that the optimal value of the optimization problem considered in the above theorem is equal to the maximum robustness radius, if one furthermore assumes the equality of the Bouligand tangent cone and the cone of feasible directions.

3.4 Comparison with other notions

In the following, we highlight the relationships between the robustness notion considered in this chapter and some notions mentioned in Section 3.1.2. See also [32] for some comparisons.

Proposition 3.4. *Let \bar{x} be a robust solution of Problem (3.3) with radius ϵ . Then, \bar{x} is flimsy and highly robust efficient for the following uncertain multi-objective optimization problem,*

$$\begin{aligned} \min \quad & f(C, x) \\ \text{s.t.} \quad & x \in X, C \in U, \end{aligned}$$

with uncertain set $U = \{C_{p \times n} : \|C\| < \epsilon\}$ and uncertain objective $f(C, x) = f(x) + Cx$.

Proof. The proof is not difficult and is hence omitted. \square

The following proposition provides a connection between the robustness and point-based minmax robust efficiency.

Proposition 3.5. *Let \bar{x} be a robust solution of Problem (3.3) with radius ϵ . Then considering any $\bar{\epsilon} \in (0, \epsilon)$, the vector \bar{x} is a point-based minmax robust efficient solution with $U = \{C_{p \times n} : \|C^i\| \leq \frac{\bar{\epsilon}}{\sqrt{p}}, \forall i = 1, 2, \dots, p\}$ and $f(C, x) = f(x) + Cx$.*

Proof. Let $f_u^{\max}(x) = \max_{C \in U} f(x) + Cx$. To the contrary, assume that there exists some $x^o \in X$ such that $f_u^{\max}(x^o) \leq f_u^{\max}(\bar{x})$. If $\bar{x} = 0$, then $f_u^{\max}(\bar{x}) = f(0)$ and hence by considering a $p \times n$ matrix C with $\|C^i\| \leq \frac{\bar{\epsilon}}{\sqrt{p}}, i = 1, 2, \dots, p$, we get $\|C\| = (\sum_{i=1}^p \|C^i\|^2)^{\frac{1}{2}} < \epsilon$, and

$$f(x^o) + Cx^o \leq f(0).$$

These relations contradict the robustness of \bar{x} (in the sense of Definition 3.6). Now, assume that $\bar{x} \neq 0$. Then $f_u^{\max}(\bar{x}) = f(\bar{x}) + \frac{\bar{\epsilon}\|\bar{x}\|}{\sqrt{p}}e$, where e is a vector with all components being equal to one. Now, we consider a $p \times n$ matrix \bar{C} , with $\bar{C}^i = \frac{\bar{\epsilon}}{\sqrt{p}\|\bar{x}\|}\bar{x}^T$. We get $\bar{C} \in U$ and

$$f(x^o) + \bar{C}x^o \leq f_u^{\max}(x^o) \leq f_u^{\max}(\bar{x}) = f(\bar{x}) + \frac{\bar{\epsilon}\|\bar{x}\|}{\sqrt{p}}e = f(\bar{x}) + \bar{C}\bar{x}.$$

Furthermore, $\|\bar{C}\| = \sqrt{\sum_{i=1}^p \|\bar{C}^i\|^2} < \epsilon$. Hence, $\|\bar{C}\| < \epsilon$, and

$$f(x^o) + \bar{C}x^o \leq f(\bar{x}) + \bar{C}\bar{x}.$$

These relations contradict the robustness of \bar{x} (in the sense of Definition 3.6) and the proof is complete. \square

The next theorem states relationship between robust efficiency and set-based minmax robust efficiency. Moreover, we also investigate the connection with hull-based

minmax robust efficiency.

Theorem 3.8. *Let \bar{x} be a robust solution of Problem (3.3) with radius ϵ . Then \bar{x} is a set-based minmax robust solution of the following uncertain problem:*

$$\begin{aligned} \min \quad & f(C, x) \\ \text{s.t.} \quad & x \in X, C \in U, \end{aligned}$$

with uncertain set $U = \{C_{p \times n} : \|C\| < 0.5\epsilon\}$ and uncertain objective $f(C, x) = f(x) + Cx$. In addition, \bar{x} is also hull-based minmax robust efficient.

Proof. To the contrary assume that

$$f_U(x^o) \subseteq f_U(\bar{x}) - \mathbf{R}_{\geq}^p \tag{3.27}$$

for some $x^o \in X$. This implies

$$\forall C \in U \quad \exists \bar{C} \in U \text{ s.t. } f(x^o) + Cx^o \leq f(\bar{x}) + \bar{C}\bar{x}. \tag{3.28}$$

The two vectors x^o and \bar{x} can not be zero. If $x^o = 0$, then by (3.28),

$$f(0) + \bar{C}(0) \leq f(\bar{x}) + \bar{C}\bar{x},$$

or some $\bar{C} \in U$. This contradicts the robustness assumption. Moreover, if $\bar{x} = 0$, then by considering $C = 0$ in (3.28), there exists some \bar{C} with $\|\bar{C}\| \leq 0.5\epsilon$ such that $f(x^o) \leq f(\bar{x})$. This contradicts the efficiency of \bar{x} . Hence, $x^o \neq 0$ and $\bar{x} \neq 0$.

Now, define

$$M = \{\lambda \in \mathbf{R}_{\geq}^p : \|\lambda\| \leq 1, \sum_{j=1}^p \lambda_j \geq 1\}.$$

It is clear that M is a nonempty compact convex set. Let $F : M \rightrightarrows M$ be a set-valued mapping defined by

$$F(\lambda) = \{\lambda' \in M : f(x^o) + \frac{\epsilon\|x^o\|}{2\|\lambda\|}\lambda \leq f(\bar{x}) + \frac{\epsilon\|\bar{x}\|}{2\|\lambda'\|}\lambda'\}.$$

We show that $F(\lambda)$ is nonempty and convex for each $\lambda \in M$.

Let $\lambda \in M$. Defining the $p \times n$ matrix $C^o := \frac{\epsilon}{2\|\lambda\|\|x^o\|} \lambda x^{oT}$. We have $\|C^o\| \leq 0.5\epsilon$, and hence by (3.28), there exists some $p \times n$ matrix \bar{C} such that $\|\bar{C}\| \leq 0.5\epsilon$, and

$$f(x^o) + \frac{\epsilon\|x^o\|}{2\|\lambda\|} \lambda \leq f(\bar{x}) + \bar{C}\bar{x}. \quad (3.29)$$

Consider $\bar{\lambda}$ with $\bar{\lambda}_i = \|\bar{C}^i\|$. Define $\lambda' := \frac{\bar{\lambda}}{\|\bar{\lambda}\|}$. By considering the Cauchy-Schwarz inequality and $\frac{\epsilon}{2\|\bar{\lambda}\|} \geq 1$, we have

$$\begin{aligned} f(x^o) + \frac{\epsilon\|x^o\|}{2\|\lambda\|} \lambda &\leq f(\bar{x}) + \bar{C}\bar{x} \\ &\leq f(\bar{x}) + \|\bar{x}\|\bar{\lambda} \\ &\leq f(\bar{x}) + \frac{\epsilon\|\bar{x}\|}{2\|\bar{\lambda}\|} \bar{\lambda} \\ &\leq f(\bar{x}) + \frac{\epsilon\|\bar{x}\|}{2\|\lambda'\|} \lambda' \end{aligned}$$

Therefore, due to $\lambda' \in M$, we have $\lambda' \in F(\lambda)$, and hence $F(\lambda)$ is nonempty.

To prove the convexity, let $\lambda_1, \lambda_2 \in F(\lambda)$ and $v \in (0, 1)$. First, we assume that $\|\lambda_1\| = \|\lambda_2\| = 1$. Then, by definition of $F(\lambda)$, we get

$$f(x^o) + \frac{\epsilon\|x^o\|}{2\|\lambda\|} \lambda \leq f(\bar{x}) + \frac{\epsilon\|\bar{x}\|}{2} (v\lambda_1 + (1-v)\lambda_2)$$

Due to $\|v\lambda_1 + (1-v)\lambda_2\| \leq 1$ and $v\lambda_1 + (1-v)\lambda_2 \geq 0$, we can infer

$$f(x^o) + \frac{\epsilon\|x^o\|}{2\|\lambda\|} \lambda \leq f(\bar{x}) + \frac{\epsilon\|\bar{x}\|}{2\|v\lambda_1 + (1-v)\lambda_2\|} (v\lambda_1 + (1-v)\lambda_2).$$

Hence, $v\lambda_1 + (1-v)\lambda_2 \in F(\lambda)$ when $\|\lambda_1\| = \|\lambda_2\| = 1$. Now, considering two arbitrary vectors $\lambda_1, \lambda_2 \in F(\lambda)$ and $v \in (0, 1)$, there are $\gamma > 0$ and $\mu \in (0, 1)$ such that

$$v\lambda_1 + (1-v)\lambda_2 = \gamma \left(\mu \frac{\lambda_1}{\|\lambda_1\|} + (1-\mu) \frac{\lambda_2}{\|\lambda_2\|} \right). \quad (3.30)$$

Notice that $0 < \|\lambda_1\|, \|\lambda_2\| \leq 1$. By definition of $F(\lambda)$, it is clear that $\frac{\lambda_1}{\|\lambda_1\|}, \frac{\lambda_2}{\|\lambda_2\|} \in F(\lambda)$. Furthermore, if $\lambda' \in F(\lambda)$ and $\gamma\lambda' \in M$, for some $\gamma > 0$, then $\gamma\lambda' \in F(\lambda)$.

Therefore, according to (3.30), we have $v\lambda_1 + (1 - v)\lambda_2 \in F(\lambda)$. Hence, F is convex-valued. It is clear that graph of F is closed. Therefore, by the Kakutani fixed-point Theorem 1.4, there exists some $\lambda^* \in M$ such that

$$f(x^o) + \frac{\epsilon \|x^o\|}{2\|\lambda^*\|} \lambda^* \leq f(\bar{x}) + \frac{\epsilon \|\bar{x}\|}{2\|\lambda^*\|} \lambda^*. \quad (3.31)$$

The above inequality does not hold as equality, since otherwise due to (3.28) we have

$$f(\bar{x}) + \frac{\epsilon \|\bar{x}\|}{2\|\lambda^*\|} \lambda^* = f(x^o) + \frac{\epsilon \|x^o\|}{2\|\lambda^*\|} \lambda^* \leq f(\bar{x}) + \tilde{C}\bar{x},$$

for some \tilde{C} with $\|\tilde{C}\| \leq 0.5\epsilon$. Then $\frac{\epsilon \|\bar{x}\|}{2\|\lambda^*\|} \lambda^* \leq \tilde{C}\bar{x}$. By the Cauchy-Schwarz inequality, we get $\frac{\epsilon \|\bar{x}\|}{2\|\lambda^*\|} \lambda^* \leq \|\bar{x}\|d$ in which $d \in \mathbf{R}^p$ with $d_i = \|\tilde{C}^i\|$. Therefore, $\|\tilde{C}\| = \|d\| > 0.5\epsilon$ which is a contradiction. Thus, inequality (3.31) holds and it does not hold as equality.

On the other hand, by Cauchy-Schwarz inequality,

$$f(x^o) + \frac{\epsilon \bar{x}^T x^o}{2\|\lambda^*\| \|\bar{x}\|} \lambda^* \leq f(x^o) + \frac{\epsilon x^{oT} x^o}{2\|\lambda^*\| \|x^o\|} \lambda^*.$$

Hence, according to (3.31),

$$f(x^o) + \frac{\epsilon \bar{x}^T x^o}{2\|\lambda^*\| \|\bar{x}\|} \lambda^* \leq f(\bar{x}) + \frac{\epsilon \bar{x}^T \bar{x}}{2\|\lambda^*\| \|\bar{x}\|} \lambda^*.$$

Therefore, setting $C^o = \frac{\epsilon}{2\|\lambda^*\| \|\bar{x}\|} \lambda^* \bar{x}^T$, we have $\|C^o\| < \epsilon$ and

$$f(x^o) + C^o x^o \leq f(\bar{x}) + C^o \bar{x}.$$

The two last relations contradict the robustness of \bar{x} (in the sense of Definition 3.6) and the proof of the first part is complete.

As $co(f_U(\bar{x})) = f_U(\bar{x})$, \bar{x} is also hull-based minmax robust efficient. \square

It is not difficult to see that Theorem 3.8 will be valid if one replaces 0.5ϵ , in the considered uncertainty set, with any $\bar{\epsilon} \in (0, \epsilon)$. Note that the converse of none of the propositions given in this section holds, in general.

3.5 Modification of the objective function

In this section, we consider two robustness aspects of (weakly/properly) efficient solutions. In the first one, we consider a convex combination of the objective function of Problem (3.21) with a new special function. The second robustness aspect is due to adding a new objective function to the problem. In both cases, we examine preserving the weak/proper/robust efficiency.

Consider the following problem for $\alpha \in [0, 1]$:

$$\begin{aligned} \min \quad & f(x) + (1 - \alpha)h(x)q \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i \in \{1, 2, \dots, m\}, \end{aligned}$$

where $h : \mathbf{R}^n \rightarrow \mathbf{R}$ is a convex function and $q \in \mathbf{R}_{\geq}^p$ is a p -vector with nonnegative components. We denote this program by $(MOP)_{\alpha}$, and this program coincides with (3.21) when $\alpha = 1$.

Note: Throughout this section, we assume that the functions h, f_i ($i = 1, \dots, p$) and g_j ($j = 1, \dots, m$) are convex and there is \hat{x} such that $g_j(\hat{x}) < 0$ ($j = 1, \dots, m$).

Theorem 3.9 presents a sufficient condition for properly efficient solutions of problems (3.21) and $(MOP)_0$ to remain properly efficient for $(MOP)_{\alpha}$.

Theorem 3.9. *If \bar{x} is a properly efficient solution of both problems (3.21) and $(MOP)_0$, then \bar{x} is a properly efficient solution of $(MOP)_{\alpha}$ for each $\alpha \in (0, 1)$.*

Proof. Since \bar{x} is a properly efficient solution of Problem (3.21), then there exist $\lambda \in \mathbf{R}^p$ and $w \in \mathbf{R}^m$ such that

$$0 \in \sum_{i=1}^p \lambda_i \partial f_i(\bar{x}) + \sum_{j=1}^m w_j \partial g_j(\bar{x}), \quad w_j g_j(\bar{x}) = 0, \quad j = 1, \dots, m, \quad \lambda > 0, \quad w \geq 0.$$

Also, since \bar{x} is a properly efficient solution of Problem $(MOP)_0$, there exist $\mu \in \mathbf{R}^p$ and $v \in \mathbf{R}^m$ such that $\mu > 0, v \geq 0$, and

$$0 \in \sum_{i=1}^p \mu_i \partial f_i(\bar{x}) + q^T \mu \partial h(\bar{x}) + \sum_{j=1}^m v_j \partial g_j(\bar{x}), \quad (3.32)$$

$$v_j g_j(\bar{x}) = 0, \quad j = 1, \dots, m. \quad (3.33)$$

Notice that the convexity of h is crucial in obtaining (3.32).

Let $\alpha \in (0, 1)$. We define t and γ as follows:

$$t := \frac{\alpha q^T \mu}{\alpha q^T \mu + (1-\alpha) q^T \lambda},$$

$$\gamma := t\lambda + (1-t)\mu.$$

It is clear that $0 < t < 1$ and $\gamma > 0$. Also,

$$(1-t)q^T \mu = (1-\alpha)q^T \gamma. \quad (3.34)$$

Thus,

$$\sum_{i=1}^p \gamma_i \partial f_i(\bar{x}) + (1-\alpha)q^T \gamma \partial h(\bar{x}) = t \sum_{i=1}^p \lambda_i \partial f_i(\bar{x}) + (1-t) \sum_{i=1}^p \mu_i \partial f_i(\bar{x}) + (1-t)q^T \mu \partial h(\bar{x}).$$

Therefore, setting $z = tw + (1-t)v$, we get

$$\begin{aligned} 0 &\in \sum_{i=1}^p (t\lambda_i + (1-t)\mu_i) \partial f_i(\bar{x}) + (1-t)q^T \mu \partial h(\bar{x}) + \sum_{j=1}^m (tw_j + (1-t)v_j) \partial g_j(\bar{x}) \\ &= \sum_{i=1}^p \gamma_i \partial f_i(\bar{x}) + (1-\alpha)q^T \gamma \partial h(\bar{x}) + \sum_{j=1}^m z_j \partial g_j(\bar{x}), \end{aligned}$$

where $\gamma > 0$ and $z \geq 0$. Therefore, \bar{x} is a global minimizer for

$$\begin{aligned} \min \quad & \sum_{i=1}^p \gamma_i f_i(x) + (1-\alpha)q^T \gamma h(x) \\ \text{s.t.} \quad & g_j(x) \leq 0, \quad j = 1, \dots, m. \end{aligned}$$

This implies that \bar{x} is a properly efficient solution of $(MOP)_\alpha$, according to Theorem 3.11 of [24]. \square

The following two results give sufficient conditions for efficient (respectively weakly) solutions of problems (3.21) and $(MOP)_0$ to remain efficient (resp. weakly) for $(MOP)_\alpha$. These results extend Proposition 2.2 of [32].

Theorem 3.10. *Let \bar{x} be an efficient solution of both Problems (3.21) and $(MOP)_0$.*

Then \bar{x} is an efficient solution of $(MOP)_\alpha$ for each $\alpha \in (0, 1)$.

Proof. Let $\alpha \in (0, 1)$. To the contrary assume that there exists a feasible point, \hat{x} , such that

$$\begin{aligned} f(\hat{x}) + (1 - \alpha)qh(\hat{x}) &\leq f(\bar{x}) + (1 - \alpha)qh(\bar{x}), \\ f(\hat{x}) + (1 - \alpha)qh(\hat{x}) &\neq f(\bar{x}) + (1 - \alpha)qh(\bar{x}). \end{aligned}$$

If $h(\bar{x}) < h(\hat{x})$, then

$$f(\hat{x}) - f(\bar{x}) \leq (1 - \alpha)q(h(\bar{x}) - h(\hat{x})) \leq (\text{and } \neq) 0.$$

This contradicts the efficiency of \bar{x} for (3.21). Hence, we assume $h(\bar{x}) \geq h(\hat{x})$. Due to the convexity assumption, we have

$$\begin{aligned} &f\left(\frac{1}{2}\hat{x} + \frac{1}{2}\bar{x}\right) + qh\left(\frac{1}{2}\hat{x} + \frac{1}{2}\bar{x}\right) \\ &\leq \frac{1}{2}f(\hat{x}) + \frac{1}{2}f(\bar{x}) + \frac{1}{2}qh(\hat{x}) + \frac{1}{2}qh(\bar{x}) \\ &\leq (\text{and } \neq) f(\bar{x}) + \frac{1}{2}(1 - \alpha)q\left(h(\bar{x}) - h(\hat{x})\right) + \frac{1}{2}qh(\hat{x}) + \frac{1}{2}qh(\bar{x}) \\ &= f(\bar{x}) + q\left(h(\bar{x}) + \frac{\alpha}{2}(h(\hat{x}) - h(\bar{x}))\right) \leq f(\bar{x}) + qh(\bar{x}). \end{aligned}$$

Hence, setting $z = \frac{1}{2}\hat{x} + \frac{1}{2}\bar{x}$, the vector z is feasible and

$$f(z) + qh(z) \leq f(\bar{x}) + qh(\bar{x}).$$

This contradicts the efficiency of \bar{x} for $(MOP)_0$, and completes the proof. \square

Theorem 3.11. *Let \bar{x} be a weakly efficient solution of both Problems (3.21) and $(MOP)_0$. Then \bar{x} is a weakly efficient solution of $(MOP)_\alpha$ for each $\alpha \in (0, 1)$.*

Proof. The proof is similar to that of Theorem 3.10 and is hence omitted. \square

Theorem 3.12 gives a sufficient condition for robust efficient solutions of Problems (3.21) and $(MOP)_0$ to remain robust efficient for $(MOP)_\alpha$.

Theorem 3.12. *If \bar{x} is a robust efficient solution for both Problems (3.21) and $(MOP)_0$, then \bar{x} is a robust efficient solution of $(MOP)_\alpha$ for each $\alpha \in [0, 1]$.*

Proof. Let $\alpha \in [0, 1]$. By Theorem 3.10, \bar{x} is efficient for Problem $(MOP)_\alpha$. Now, we show that \bar{x} is robust for $(MOP)_\alpha$. By Theorem 3.4,

$$\text{pos}\left(\bigcup_{i=1}^p \partial f_i(\bar{x})\right) + \text{pos}\left(\bigcup_{i \in A(\bar{x})} \partial g_i(\bar{x})\right) = \mathbf{R}^n$$

and

$$\text{pos}\left(\bigcup_{i=1}^p \partial(f_i + q_i h)(\bar{x})\right) + \text{pos}\left(\bigcup_{i \in I(\bar{x})} \partial g_i(\bar{x})\right) = \mathbf{R}^n.$$

By the above two equalities, and since all the ∂ -sets are convex here, we have $\text{pos}\left(\bigcup_{i=1}^p \partial(f_i + (1 - \alpha)q_i h)(\bar{x})\right) + \text{pos}\left(\bigcup_{i \in I(\bar{x})} \partial g_i(\bar{x})\right) = \mathbf{R}^n$. Therefore, \bar{x} is a robust efficient solution for $(MOP)_\alpha$, because of Theorem 3.5. \square

In the rest of this section, we examine adding a new objective function to Problem (3.21). Consider the following multi-objective optimization problem, denoted by $(MOPh)$:

$$\begin{aligned} & \min \begin{pmatrix} f(x) \\ h(x) \end{pmatrix} \\ & \text{s.t. } g_i(x) \leq 0 \quad i = 1, \dots, m, \end{aligned}$$

where $h : \mathbf{R}^n \rightarrow \mathbf{R}$. The following two theorems address some connections between the properly efficient solutions of the two problems (3.21) and $(MOPh)$. Recall that the functions h , f_i , and g_j are convex.

Theorem 3.13. *Let \bar{x} be a properly efficient solution of Problem $(MOPh)$. If $\partial h(\bar{x}) \subseteq \text{pos}\left(\bigcup_{i=1}^p \partial f_i(\bar{x})\right) + \text{pos}\left(\bigcup_{i \in I(\bar{x})} \partial g_i(\bar{x})\right)$, then \bar{x} is a properly efficient solution of Problem (3.21).*

Proof. Since \bar{x} is a properly efficient solution of Problem $(MOPh)$, then there

exist $\lambda \in \mathbf{R}^p$ and $w \in \mathbf{R}^m$ such that

$$\begin{aligned} 0 &\in \sum_{i=1}^p \lambda_i \partial f_i(\bar{x}) + \partial h(\bar{x}) + \sum_{j=1}^m w_j \partial g_j(\bar{x}), \\ w_j g_j(\bar{x}) &= 0, \quad j = 1, \dots, m \quad \lambda > 0, \quad w \geq 0. \end{aligned}$$

Therefore, by assumption of the theorem, $0 \in \sum_{i=1}^p \bar{\lambda}_i \partial f_i(\bar{x}) + \sum_{j=1}^m \bar{w}_j \partial g_j(\bar{x})$, for some $\bar{\lambda} > 0, \bar{w} \geq 0$. This implies that \bar{x} is a properly efficient solution of Problem (3.21). \square

By a manner similar to the proof Theorem 3.13, it can be shown that the result is valid for weak efficient solutions as well. The following example shows that this result may not be valid for efficient solutions.

Example 3.4. Let $g(x) = f(x) = x$ and $h(x) = x^2$. It is clear that $\bar{x} = 0$ is an efficient solution of (MOPh) and $\{\nabla h(0)\} \subseteq \text{pos}(\nabla f(0))$ but $\bar{x} = 0$ is not an efficient solution of Problem (3.21).

The following result gives even more insights into the connection between the proper efficient solutions of the two problems (3.21) and (MOPh), when the f_i and g_j functions are continuously differentiable.

Theorem 3.14. (i) If \bar{x} is a properly efficient solution of both Problems (3.21) and (MOPh), then there exist vectors $u \in \mathbf{R}^p$ and $v \in \mathbf{R}^m$ such that $u > 0$, and

$$\left(\sum_{i=1}^p u_i \nabla f_i(\bar{x}) + \sum_{j=1}^m v_j \nabla g_j(\bar{x}) \right) \in \partial h(\bar{x}).$$

(ii) Let \bar{x} be a properly efficient solution of Problem (3.21). If there exist vectors $u \in \mathbf{R}^p$ and $v \in \mathbf{R}^{|I(\bar{x})|}$ such that $v \geq 0$ and

$$\left(\sum_{i=1}^p u_i \nabla f_i(\bar{x}) - \sum_{j \in I(\bar{x})} v_j \nabla g_j(\bar{x}) \right) \in \partial h(\bar{x}),$$

then \bar{x} is a properly efficient solution of Problem (MOPh).

Proof. (i) Since \bar{x} is a properly efficient solution of Problem (MOPh), then there

exist $(\lambda, \lambda_{p+1}) \in \mathbf{R}^p \times \mathbf{R}$ and $w \in \mathbf{R}^m$ such that

$$\begin{aligned} 0 &\in \sum_{i=1}^p \lambda_i \nabla f_i(\bar{x}) + \lambda_{p+1} \partial h(\bar{x}) + \sum_{j=1}^m w_j \nabla g_j(\bar{x}), \\ w_j g_j(\bar{x}) &= 0, \quad j = 1, \dots, m, \quad \lambda > 0, \quad w \geq 0. \end{aligned}$$

Therefore, there exists some $d \in \partial h(\bar{x})$ such that

$$d = - \sum_{i=1}^p \frac{\lambda_i}{\lambda_{p+1}} \nabla f_i(\bar{x}) - \sum_{j=1}^m \frac{w_j}{\lambda_{p+1}} \nabla g_j(\bar{x}).$$

On other hand, \bar{x} is a properly efficient solution of Problem (3.21). Therefore, there exist $\lambda' \in \mathbf{R}^p$ and $w' \in \mathbf{R}^m$ such that

$$0 = \sum_{i=1}^p \lambda'_i \nabla f_i(\bar{x}) + \sum_{j=1}^m w'_j \nabla g_j(\bar{x}), \quad (3.35)$$

$$w'_j g_j(\bar{x}) = 0, \quad j = 1, \dots, m \quad \lambda' > 0, \quad w' \geq 0. \quad (3.36)$$

Let $t > \max_{1 \leq i \leq p} \left\{ \frac{\lambda_i}{\lambda'_i \lambda_{p+1}} \right\}$. We have

$$d = \sum_{i=1}^p \left(t \lambda'_i - \frac{\lambda_i}{\lambda_{p+1}} \right) \nabla f_i(\bar{x}) + \sum_{j=1}^m \left(t w'_j - \frac{w_j}{\lambda_{p+1}} \right) \nabla g_j(\bar{x}).$$

Setting $u_i := t \lambda'_i - \frac{\lambda_i}{\lambda_{p+1}}$ and $v_j = t w'_j - \frac{w_j}{\lambda_{p+1}}$, completes the proof of part (i).

(ii) Setting $\mu_j = 0$, for each $j \notin I(\bar{x})$, by the assumption of the theorem, we have $0 = - \sum_{i=1}^p u_i \nabla f_i(\bar{x}) + d + \sum_{j=1}^m \mu_j \nabla g_j(\bar{x})$, for some $d \in \partial h(\bar{x})$. On other hand, since \bar{x} is a properly efficient solution of Problem (3.21), there exist $\lambda' \in \mathbf{R}^p$ and $w' \in \mathbf{R}^m$ satisfying (3.35) and (3.36). For $t > \max_{1 \leq i \leq p} \left\{ \frac{u_i}{\lambda'_i} \right\}$, we have

$$\begin{aligned} 0 &= \sum_{i=1}^p (t \lambda'_i - u_i) \nabla f_i(\bar{x}) + d + \sum_{j=1}^m (t w'_j + \mu_j) \nabla g_j(\bar{x}) \\ &\in \sum_{i=1}^p (t \lambda'_i - u_i) \nabla f_i(\bar{x}) + \partial h(\bar{x}) + \sum_{j=1}^m (t w'_j + \mu_j) \nabla g_j(\bar{x}). \end{aligned}$$

Therefore, \bar{x} is a properly efficient solution of $(MOPh)$, and the proof is complete. \square

The following example shows that part (i) of the above theorem may not hold when some f_i or g_j functions are nonsmooth. A similar example can be constructed

for part (ii).

Example 3.5. Let $f : \mathbf{R} \rightarrow \mathbf{R}$ be defined by

$$f(x) = \begin{cases} x_2 - x_1, & x_2 > 0 \\ -x_1, & x_2 \leq 0. \end{cases}$$

Consider the following optimization problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) = x_1 - x_2 \leq 0. \end{aligned}$$

The functions f and g are convex and

$$\partial f \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \left\{ \begin{pmatrix} -1 \\ \alpha \end{pmatrix} : \alpha \in [0, 1] \right\}, \quad \partial g \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

With $\lambda = \alpha = \mu = 1$, we have

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} -1 \\ \alpha \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Therefore, $\bar{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is an optimal solution of the above problem. Now, consider the function $h(x) = x_1$ and the following problem

$$\begin{aligned} \min \quad & \begin{pmatrix} f(x) \\ h(x) \end{pmatrix} \\ \text{s.t.} \quad & g(x) = x_1 - x_2 \leq 0. \end{aligned} \tag{3.37}$$

We have $\partial h(\bar{x}) = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$. Also, for $\lambda_1 = \lambda_2 = 1$ and $\alpha = \mu = 0$, we get

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \lambda_1 \begin{pmatrix} -1 \\ \alpha \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Therefore, $\bar{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a properly efficient solution of Problem (3.37). Hence, in this example, $\bar{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a properly efficient solution of both problems (3.21) and (MOPh), while there is not any $\lambda > 0$ satisfying

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} -1 \\ \alpha \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

for some $\mu \in \mathbf{R}$ and $\alpha \in [0, 1]$. It shows that part (i) of Theorem 3.14 may not be valid in the presence of nonsmooth f_i or g_j functions.

The last theorem of this section establishes a connection between the robust solutions of two problems (3.21) and (MOPh).

Theorem 3.15. *If \bar{x} be a robust efficient solution of Problem (3.21), then \bar{x} is a robust efficient solution of (MOPh). The converse holds if*

$$\partial h(\bar{x}) \subseteq \text{pos} \left(\bigcup_{i=1}^p \partial f_i(\bar{x}) \right) + \text{pos} \left(\bigcup_{i \in I(\bar{x})} \partial g_i(\bar{x}) \right).$$

Proof. The results follow from Theorems 3.3 and 3.6. □

Chapter 4

Stability and sensitivity analysis

4.1 Introduction

This chapter is concerned with sensitivity analysis and stability in parametrized multi-objective optimization. A fundamental question in optimization is the investigation of behavior of solution set and efficient values under perturbation. Considerable work has been devoted to these topics under the titles of stability and sensitivity analysis. The stability term is used for examination of continuity of optimal value and solution set under perturbation. By sensitivity analysis we mean studying differentiability of the optimal value function and solution set. Moreover, this served as an important motivation for proposing some generalized derivatives including proximal subdifferential.

Here, we only mention some works dedicated to sensitivity analysis and stability in multi-objective optimization. We refer the reader to [12] and references therein for single objective optimization.

Stability in optimization problems is both theoretically and practically important. It has been used as a useful tool in post-optimal analysis. There are many papers addressing stability in vector optimization. Some scholars have considered parametric vector optimization with the objective function and the feasible set depending on a parameter lying in a Banach space. They provide some sufficient conditions for lower and upper semicontinuity of set-valued mappings corresponding to feasible set,

efficient set and so on; see, e.g., [53, 54, 65, 72] and the references therein. Moreover, Sawaragi et al. [65] also investigate the problems with parameter-dependent ordering cones. Bednarczuk [5, 6] introduces the concept of strict efficiency and she establishes Holder like continuity of feasible solution set and optimal value under strict efficiency and some other appropriate assumptions.

Some scholars study stability from a different standpoint. They consider a sequence of vector optimization problems whose feasible set and objective function tend to a given set and a given function in the sense of Painlevé-Kuratowski or other sense, respectively. They give some sufficient conditions for convergence of the sequence of efficient sets to the efficient set of the given problem; see [52, 58] and the references therein.

Sensitivity analysis is also a strong tool in post-optimal analysis. There are numerous works devoted to this concept in vector optimization. Tanino [70, 71] is among the first scholars who have dedicated some studies to this issue. He applies contingent derivative for analyzing efficient solutions and efficient values [70, 71]. Shi [67] introduces the concept of TP cone. Some results on contingent derivative of the efficient value mapping are improved and extended by using this concept [49, 67]. We refer the readers to excellent survey [72].

Lee et al. [51] use Proto derivative for analyzing parametric vector optimization. The authors obtain some sufficient conditions for Proto differentiability of efficient solution set and efficient values. Chuong et al. [15] use the generalized Clarke epiderivative notion to study the problem. They provide some formulas for calculating this derivative. Chuong [14] studies the sensitivity analysis via s-derivative. He also gives some formulas for inner and outer estimation of s-derivative for efficient values.

The concept of coderivative has also been applied to investigating the behavior of efficient solutions and efficient values; see [13, 47] and the references therein.

Here, we study parametric multi-objective optimization problem whose feasible set is given explicitly by some inequalities. Throughout this chapter, all functions are assumed to be locally Lipschitz. We establish some theorems for strict semi-differentiability of feasible set and feasible values. Moreover, we give some formulas

for their derivatives. We also provide some sufficient conditions for their pseudo-Lipschitz continuity. Thanks to the mentioned theorem, we provide sufficient conditions for semi-differentiability and pseudo-Lipschitz continuity of efficient solutions and efficient values. Furthermore, some formulas for their derivatives are given. In addition, some counterexamples are given to clarify the theoretical results.

The rest of the chapter unfolds as follows. Section 2 is devoted to strict semi-differentiability and pseudo-Lipschitz continuity of the set-valued mappings of feasible set and feasible values. We establish that under mild conditions, the set-valued mapping of feasible set and its feasible value are strictly semi-differentiable. Moreover, we investigate the pseudo-Lipschitz continuity of the mentioned set-valued mappings. In Section 3, we provide some sufficient conditions for pseudo-Lipschitz continuity and semi-differentiability of the efficient solutions and efficient values.

4.2 Parametric multi-objective optimization

Consider the following parametric multi-objective optimization problem, denoted as $P(u)$,

$$\begin{aligned} \min \quad & f(u, x) \\ \text{s.t.} \quad & g(u, x) \leq 0, \end{aligned} \tag{4.1}$$

where $f : \mathbf{R}^n \times \mathbf{R}^q \rightarrow \mathbf{R}^p$ and $g : \mathbf{R}^n \times \mathbf{R}^q \rightarrow \mathbf{R}^m$ are locally Lipschitz functions. Here, x is a decision vector and u is a parameter. Associated with this problem, the following set-valued mappings, called feasible solution, feasible value, efficient solution and efficient value mappings, are respectively defined by

$$\begin{aligned} X(u) &:= \{x \in \mathbf{R}^n : g(u, x) \leq 0\} \\ Y(u) &:= \{f(u, x) : x \in X(u)\} \\ E(u) &:= \{x \in X(u) : x \text{ is an efficient solution of } P(u)\} \\ V(u) &:= \{f(u, x) : x \in E(u)\} \end{aligned}$$

In the sequel, we say that constraint qualification CQ holds at $(\bar{u}, \bar{x}) \in \text{gph}X$ if

$$0 \notin \text{co}(\cup_{i \in I(\bar{u}, \bar{x})} \partial_x g_i(\bar{u}, \bar{x}))$$

where $I(\bar{u}, \bar{x})$ denotes the index set of active constraints at (\bar{u}, \bar{x}) and ∂_x stands for generalized gradient with respect to x . It is well-known that this constraint qualification is equivalent to

$$M(\mathbf{R}^n) + \mathbf{R}_{\geq}^{|I(\bar{u}, \bar{x})|} = \mathbf{R}^{|I(\bar{u}, \bar{x})|}, \quad \forall M \in \partial_x g_{I(\bar{u}, \bar{x})},$$

where $g_{I(\bar{u}, \bar{x})}$ is the function defined by the active components of g at (\bar{u}, \bar{x}) and $|I(\bar{u}, \bar{x})|$ is equal to the cardinal number of $I(\bar{u}, \bar{x})$. The next remark provides a sufficient condition for pseudo-Lipschitz continuity of X at a given point [62].

Remark 4.1. *Let $(\bar{u}, \bar{x}) \in \text{gph}X$. If CQ holds at (\bar{u}, \bar{x}) , then X is pseudo-Lipschitz continuous at (\bar{u}, \bar{x}) .*

Proposition 4.1. *Let $(\bar{u}, \bar{x}) \in \text{gph}X$ and let g be regular at this point. If CQ holds at (\bar{u}, \bar{x}) , then X is strictly semi-differentiable at (\bar{u}, \bar{x}) and for each $l \in \mathbf{R}^q$*

$$D_{s\text{-low}}X(\bar{u}, \bar{x})(l) = \{d : g'_i(\bar{u}, \bar{x}; l, d) \leq 0, \forall i \in I(\bar{u}, \bar{x})\}. \quad (4.2)$$

Proof. Since CQ holds at (\bar{u}, \bar{x}) and g is regular at this point, $\text{gph}X$ is regular at (\bar{u}, \bar{x}) ; see Theorem 1.8. Therefore, X is Clarke differentiable at the given point. Strict semi-differentiability follows from Proposition 1.2 and Remark 4.1. Due to the mentioned facts, derivative of X at (\bar{u}, \bar{x}) is given by the above formula. \square

The following examples illustrate that all the given assumptions of Proposition 4.1 are essential.

Example 4.1. *Let $g(u, x) = x - \max\{u, 1\}$. It can be verified easily at $(\bar{u}, \bar{x}) = (1, 1)$ that*

$$DX(\bar{u}, \bar{x})(1) = \{l : -\infty < l \leq 1\}, \quad D_{s\text{-low}}X(\bar{u}, \bar{x})(1) = \{l : -\infty < l \leq 0\}.$$

Thus, X is not strictly semi-differentiable at (\bar{u}, \bar{x}) . Here, g is not regular.

Example 4.2. Let $g(u, x) = x^2 - u^2$. It can be verified easily at $(\bar{u}, \bar{x}) = (0, 0)$ that

$$DX(\bar{u}, \bar{x})(1) = \{l : -1 \leq l \leq 1\}, \quad D_{s\text{-low}}X(\bar{u}, \bar{x})(1) = \emptyset.$$

Thus, X is not strictly semi-differentiable at (\bar{u}, \bar{x}) . Here, CQ does not hold at the given point.

In the rest of the section, we deal with pseudo-Lipschitz continuity and strict semi-differentiability of Y at a given point. To express the result more concisely, we define the set-valued mapping $\hat{X} : \text{gph}Y \rightrightarrows \mathbf{R}^n$ given by $\hat{X}(u, y) = \{x \in X(u) : f(u, x) = y\}$.

Theorem 4.1. Let $(\bar{u}, \bar{y}) \in \text{gph}Y$ and let $\text{gph}Y$ be closed around (\bar{u}, \bar{y}) . Under each of the following assumptions, Y is pseudo-Lipschitz continuous at (\bar{u}, \bar{y}) .

(i) there is $\bar{x} \in \hat{X}(\bar{u}, \bar{y})$ such that for any sequence $\{(u_n, y_n)\} \subseteq \text{gph}Y$ with $(u_n, y_n) \rightarrow (\bar{u}, \bar{y})$, there exists a sequence $\{x_n\}$ with $x_n \in \hat{X}(u_n, y_n)$ admitting \bar{x} as a cluster point and X is pseudo-Lipschitz continuous at (\bar{u}, \bar{x}) ;

(ii) the set-valued mapping \hat{X} is locally compact on $N \cap \text{gph}Y$ for some neighborhood N of (\bar{u}, \bar{y}) . X is pseudo-Lipschitz continuous at each $(\bar{u}, \bar{x}) \in \hat{X}(\bar{u}, \bar{y})$.

Proof. According to Theorem 1.3, Y is pseudo-Lipschitz continuous at (\bar{u}, \bar{y}) if and only if $(\sigma, 0) \in N_{\text{gph}Y}(\bar{u}, \bar{y})$ implying $\sigma = 0$. We argue by contradiction. Let Y be not pseudo-Lipschitz continuous at (\bar{u}, \bar{y}) . So, $(\sigma, 0) \in N_{\text{gph}Y}(\bar{u}, \bar{y})$, for some $\sigma \neq 0$.

Hence, there are sequences $\{u_n, y_n\} \subseteq \text{gph}Y$ and $(\sigma_n, \mu_n) \rightarrow (\sigma, 0)$ with $(\sigma_n, \mu_n) \in \hat{N}_{\text{gph}Y}(u_n, y_n)$ and $(u_n, y_n) \rightarrow (\bar{u}, \bar{y})$. On the account of condition (i), there is a sequence $\{x_n\}$ (without relabelling) with $x_n \in \hat{X}(u_n, y_n)$ and $x_n \rightarrow \bar{x}$. Due to the local Lipschitz continuity of f , for each n , we have

$$\begin{aligned} & \langle \sigma_n, u - u_n \rangle + \langle \mu_n, y - y_n \rangle \leq o(\|u - u_n\| + \|y - y_n\|), \quad \forall (u, y) \in \text{gph}Y, \\ \Rightarrow & \langle \sigma_n, u - u_n \rangle + \langle \mu_n, f(u, x) - y_n \rangle \leq o(\|u - u_n\| + \|f(u, x) - f(u_n, x_n)\|), \quad \forall (u, x) \in \text{gph}X, \\ \Rightarrow & \langle \sigma_n, u - u_n \rangle \leq o(\|u - u_n\| + \|x - x_n\|) + \epsilon_n(\|u - u_n\| + \|x - x_n\|), \quad \forall (u, x) \in \text{gph}X, \end{aligned}$$

with positive sequence $\{\epsilon_n\}$ tending to zero. We can infer from Theorem 1.2 that $(\sigma, 0) \in N_{gphX}(\bar{u}, \bar{x})$. This is a contradiction with pseudo-Lipschitz continuity of X at (\bar{u}, \bar{x}) , and so Y is pseudo-Lipschitz continuous at (\bar{u}, \bar{y}) .

The pseudo-Lipschitz continuity of Y at (\bar{u}, \bar{y}) under (ii) can be proved in a similar way. \square

It is easy to see that if $\cup_{u \in U} X(u)$ is bounded for some neighborhood U of \bar{u} , then the first part of condition (ii) in Theorem 4.1 is fulfilled. In the following theorem, we provide another hypothesis for having Theorem 4.1. Before we state the theorem, we need to recall some notions.

Consider the set-valued mapping X . The outer horizon limit of X at \bar{u} is denoted and defined by $X^\infty(u) := \{\lim_n t_n x_n : t_n \downarrow 0, u_n \rightarrow u, x_n \in X(u_n)\}$. The outer horizon limit of a set-valued mapping at a given point is a closed cone. The reader can see [64] for more information about the outer horizon limit.

The asymptotic function of f with respect to the set-valued mapping X at u is defined and denoted by $f_X^\infty(u; d) := \{\lim_n t_n f(u_n, x_n) : t_n \downarrow 0, u_n \rightarrow u, t_n x_n \rightarrow d, x_n \in X(u_n)\}$ (with some modifications). In general, asymptotic function is a closed-valued set-valued mapping. We refer the reader to [55] for more details.

Theorem 4.2. *Let $(\bar{u}, \bar{y}) \in gphY$ be given and let $gphY$ be closed around (\bar{u}, \bar{y}) . Assume the following conditions hold:*

- (i) X is pseudo-Lipschitz continuous at (\bar{u}, \bar{x}) for every $\bar{x} \in \hat{X}(\bar{u}, \bar{y})$;
- (ii) $\{v : f_X^\infty(u; v) = 0\} \cap X^\infty(\bar{u}) = \{0\}$.

Then Y is pseudo-Lipschitz continuous at (\bar{u}, \bar{y}) .

Proof. For each sequence $\{(u_n, y_n)\} \subseteq gphY$ with $(u_n, y_n) \rightarrow (\bar{u}, \bar{y})$, there exists a sequence $\{x_n\}$ with $x_n \in \hat{X}(u_n, y_n)$. If $\{x_n\}$ has no cluster point, then $\{\|x_n\|\}$ converges to ∞ . By taking $t_n = 1/\|x_n\|$, we may assume $\{t_n x_n\}$ converges to some nonzero vector d . Then

$$0 = \lim_{n \rightarrow \infty} t_n y_n \in f_X^\infty(u; d),$$

which shows that $d \in \{v : f_X^\infty(u; v) = 0\}$. This contradicts the hypothesis because $d \in X^\infty(\bar{u})$ and $d \neq 0$. Hence, $\{x_n\}$ admits some cluster points. Moreover, on account of continuity of f and g , cluster points belong to $\hat{X}(\bar{u}, \bar{y})$. Now, the result can be obtained similar to Theorem 4.1. \square

Example 4.3. Let $X(u) = \{x \in R : 0 \leq x \leq \sqrt{|u|}\} \cup [2, 3]$ and

$$f(u, x) = \begin{cases} x, & x \leq 1 \\ 2 - x, & x \geq 1. \end{cases}$$

It can be verified easily that neither (i) nor (ii) holds at $(\bar{u}, \bar{y}) = (0, 0)$ and is not pseudo-Lipschitz continuous at the point.

In the next theorem, we provide some sufficient conditions for strict semi-differentiability of Y .

Theorem 4.3. Let $y \in Y(\bar{u})$ be given and let $\text{gph}Y$ be closed around (\bar{u}, \bar{y}) . If the following conditions hold, then Y is strictly semi-differentiable at (\bar{u}, \bar{y}) .

- (i) there exists $\bar{x} \in \hat{X}(\bar{u}, \bar{y})$ such that for each sequence $\{(u_n, y_n)\} \subseteq \text{gph}Y$ with $(u_n, y_n) \rightarrow (\bar{u}, \bar{y})$, \bar{x} is a cluster point for some sequence $\{x_n\}$ with $\{(u_n, y_n, x_n)\} \subseteq \text{gph}\hat{X}$;
- (ii) X is pseudo-Lipschitz continuous and strictly semi-differentiable at (\bar{u}, \bar{x}) and f is regular at this point;
- (iii) $DX(\bar{u}, \bar{x})(0) \cap \text{Ker}f'_x(\bar{u}, \bar{x}) = \{0\}$,

where $\text{Ker}f'_x(\bar{u}, \bar{x}) = \{d : f'(\bar{u}, \bar{x}; 0, d) = 0\}$.

In addition,

$$D_{s\text{-low}}Y(\bar{u}, \bar{y})(l) = \{f'(\bar{u}, \bar{x}; l, d) : \forall d \in D_{s\text{-low}}X(\bar{u}, \bar{x})(l)\}, \quad \forall l \in \mathbf{R}^q. \quad (4.3)$$

Proof. First, we show that if $(l, e) \in T_{\text{gph}Y}$, then $e \in \{f'(\bar{u}, \bar{x}; l, d) : \forall d \in D_{s\text{-low}}X(\bar{u}, \bar{x})(l)\}$.

Let $(l, e) \in T_{\text{gph}Y}$. By the definition of Bouligand tangent cone, there exist $\{(u_n, y_n)\} \subseteq$

$\text{gph}Y$ and $t_n \downarrow 0$ with $t_n^{-1}(u_n - \bar{u}, y_n - \bar{y}) \rightarrow (l, e)$. According to the assumptions (without relabelling), there is a sequence $x_n \in X(u_n)$ with $y_n = f(u_n, x_n)$ and $x_n \rightarrow \bar{x}$. Consider the sequence of $(x_n - \bar{x})/t_n$. If it is not bounded, then we may assume $(x_n - \bar{x})/\|x_n - \bar{x}\|$ converges to some nonzero vector d and $t_n/\|x_n - \bar{x}\|$ converges to zero as n tends to ∞ . Then we get

$$0 = \lim_{n \rightarrow \infty} (y_n - \bar{y})/\|x_n - \bar{x}\| = \lim_{n \rightarrow \infty} (f(u_n, x_n) - f(\bar{u}, \bar{x}))/\|x_n - \bar{x}\| = f'(\bar{u}, \bar{x}; 0, d),$$

which contradicts the hypotheses. Hence, $(x_n - \bar{x})/t_n$ is bounded and without loss of generality we assume that $t_n^{-1}(x_n - \bar{x}) \rightarrow d$. As X is strictly semi-differentiable at (\bar{u}, \bar{x}) , $d \in D_{s\text{-low}}X(\bar{u}, \bar{x})(l)$. Moreover, due to the regularity of f at this point $e = f'(\bar{u}, \bar{x}; l, d)$.

Now, we show that if $e = f'(\bar{u}, \bar{x}; l, d)$ for some $d \in D_{s\text{-low}}X(\bar{u}, \bar{x})(l)$, then $e \in D_{s\text{-low}}Y(\bar{u}, \bar{y})(l)$. Consider an arbitrary sequences $\{(u_n, y_n)\} \subseteq \text{gph}Y$ and $\{t_n\}$ with $(u_n, y_n) \rightarrow (\bar{u}, \bar{y})$ and $t_n \downarrow 0$. By (i), there is a subsequence $\{x_{n_k}\}$ with $x_{n_k} \in \hat{X}(u_{n_k}, y_{n_k})$ and $x_{n_k} \rightarrow \bar{x}$. From strictly semi-differentiability of X at (\bar{u}, \bar{x}) , there are sequences $l_{n_k} \rightarrow l$ and $d_{n_k} \rightarrow d$ with $\{(u_{n_k} + t_{n_k}l_{n_k}, x_{n_k} + t_{n_k}d_{n_k})\} \subseteq \text{gph}X$. Consider the sequence $\{y_{n_k}\}$ given by $y_{n_k} = f(u_{n_k}, x_{n_k})$. Let $e_{n_k} = t_{n_k}^{-1}(f(u_{n_k} + t_{n_k}l_{n_k}, x_{n_k} + t_{n_k}d_{n_k}) - f(u_{n_k}, x_{n_k}))$. Thus, $\{(u_{n_k} + t_{n_k}l_{n_k}, y_{n_k} + t_{n_k}e_{n_k})\} \subseteq \text{gph}Y$, and regularity of f at (\bar{u}, \bar{x}) implies $e_{n_k} \rightarrow e$. In view of Lemma 1.1, we have $(l, e) \in T_{\text{gph}Y}^c(\bar{u}, \bar{y})$. From Theorem 4.1 and Proposition 1.2, the set-valued mapping Y is strictly semi-differentiable at the given point. This establishes the result. \square

In the next theorem, we provide another sufficient condition for the strict semi-differentiability of Y at a given point.

Theorem 4.4. *Let $y \in Y(\bar{u})$ and let $\text{gph}Y$ be closed around (\bar{u}, \bar{y}) . Suppose that the following properties hold:*

- (i) $\{v : f_X^\infty(u; v) = 0\} \cap X^\infty(\bar{u}) = \{0\}$;
- (ii) X is pseudo-Lipschitz continuous and strictly semi-differentiable on $\hat{X}(\bar{u}, \bar{y})$ and f is regular on this set;

(iii) $DX(\bar{u}, \bar{x})(0) \cap \text{Ker} f'_x(\bar{u}, \bar{x}) = \{0\}$, for every $\bar{x} \in \hat{X}(\bar{u}, \bar{y})$.

Then Y is strictly semi-differentiable at (\bar{u}, \bar{y}) and for each $l \in \mathbf{R}^q$

$$D_{s\text{-low}}Y(\bar{u}, \bar{y})(l) = \{f'(\bar{u}, \bar{x}; l, d) : \forall \bar{x} \in \hat{X}(\bar{u}, \bar{y}), \forall d \in D_{s\text{-low}}X(\bar{u}, \bar{x})(l)\}.$$

Proof. The proof is similar to that of Theorem 4.3 and Theorem 4.2. □

Remark 4.2. The hypothesis $DX(\bar{u}, \bar{x})(0) \cap \text{Ker} f'_x(\bar{u}, \bar{x}) = \{0\}$ in Theorem 4.3 can be replaced by “ (i) holds in Theorem 4.1 and for some positive constant k

$$\|x_n - \bar{x}\| \leq k\|u_n - \bar{u}\| + k\|y_n - \bar{y}\|.”$$

This statement can be established in a similar way.

The following example shows that the hypotheses assumed in Theorems 4.3, Theorems 4.4 and Remark 4.2 are essential for deriving strict semi-differentiability.

Example 4.4. Let $X(u) = \{(x_1, x_2) : x_1 + x_2 = 1, x_1, x_2 \geq 0\}$ and $f(u, x) = u(x_1 - x_2)$. At $(\bar{u}, \bar{y}) = (0, 0)$, we have

$$\begin{aligned} T_{\text{gph}Y}(\bar{u}, \bar{y}) &= \text{pos}\{(1, 1), (1, -1)\} \cup \text{pos}\{(-1, 1), (-1, -1)\}, \\ D_{s\text{-low}}Y(\bar{u}, \bar{y})(l) &= \{0\}, \end{aligned}$$

for each l . Thus, Y is not strictly semi-differentiable at (\bar{u}, \bar{y}) .

In the next lemma we provide a sufficient condition for having the hypothesis of Remark 4.2.

Lemma 4.1. Let $h := \begin{bmatrix} f(u, x) \\ g_{I(\bar{u}, \bar{x})}(\bar{u}, \bar{x}) \end{bmatrix}$. Assume that for each $M \in \partial_x h(\bar{u}, \bar{x})$, we have $M(\mathbf{R}^n) + \{0\} \times \mathbf{R}_{\geq}^{|I(\bar{u}, \bar{x})|} = \mathbf{R}^{|I(\bar{u}, \bar{x})|+p}$. Then for each sequence $\{(u_n, y_n)\} \subseteq \text{gph}Y$ with $(u_n, y_n) \rightarrow (\bar{u}, \bar{y})$, there exists $x_n \rightarrow \bar{x}$ with $x_n \in X(u_n)$ and $y_n = h(x_n, u_n)$ for

sufficiently large n . Moreover,

$$\|x_n - \bar{x}\| \leq k(\|u_n - \bar{u}\| + \|y_n - \bar{y}\|), \quad (4.4)$$

with k being a positive constant independent of the sequence $\{(u_n, y_n)\}$.

Proof. This follows from Theorem 1.7. \square

We say CQ1 holds at (\bar{u}, \bar{x}) if we have assumptions of Lemma 4.1.

Since efficient points of the two sets $Y(u)$ and $Y(u) + \mathbf{R}_{\geq}^p$ are the same, it would be in some cases beneficial to consider the set-valued mapping $Y + \mathbf{R}_{\geq}^p$, where $Y + \mathbf{R}_{\geq}^p(u) = Y(u) + \mathbf{R}_{\geq}^p$. For instance, when the functions g and f are convex with respect to x , $Y(u)$ may not be convex while $Y(u) + \mathbf{R}_{\geq}^p$ is convex. So, we are equipped with the great tools of convex analysis. Another noticeable case is when $Y(u)$ is not closed while $Y(u) + \mathbf{R}_{\geq}^p$ is closed. In what follows we investigate pseudo-Lipschitz continuity and strict semi-differentiability of this set-valued mapping.

Theorem 4.5. *Let $(\bar{u}, \bar{y}) \in \text{gph}Y$ and let Y be closed at \bar{u} . If the following conditions hold, then $Y + \mathbf{R}_{\geq}^p$ is pseudo-Lipschitz continuous at (\bar{u}, \bar{y}) .*

- (i) *Y is pseudo-Lipschitz continuous at (\bar{u}, \bar{y}) and $\text{gph}Y$ and $\text{gph}(Y + \mathbf{R}_{\geq}^p)$ are closed around it;*
- (ii) *$Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p = \{0\}$.*

Proof. We show that if $(\sigma, 0) \in N_{\text{gph}Y + \mathbf{R}_{\geq}^p}(\bar{u}, \bar{y})$ then $\sigma = 0$. On the contrary, suppose that $\sigma \neq 0$. By the definition of limiting cone, there are sequences $\{(u_n, y_n + r_n)\} \subseteq \text{gph}Y + \mathbf{R}_{\geq}^p$ and $\{(\sigma_n, \mu_n)\}$ such that $(u_n, y_n + r_n) \rightarrow (\bar{u}, \bar{y})$ and $(\sigma_n, \mu_n) \rightarrow (\sigma, 0)$. Note that $\{r_n\}$ is bounded, since If this is not the case, then we would assume without loss of generality that $r_n \rightarrow \infty$ and $\|r_n\|^{-1}r_n \rightarrow r$, where r is a nonzero vector in \mathbf{R}_{\geq}^p . Therefore,

$$\lim_{n \rightarrow \infty} (u_n - \bar{u}, \frac{y_n}{\|r_n\|}) = (0, -r),$$

which contradicts our assumptions. Moreover, $r_n \rightarrow 0$, since if this does not hold, then it leads to a contradiction with $\bar{y} \in V(\bar{u})$. As $\mu_n \rightarrow 0$, there is a sequence $\{\epsilon_n\} \subseteq \mathbf{R}_{>}$ such that

$$\langle \sigma_n, u - u_n \rangle \leq o(\|u - u_n\| + \|z - y_n - r_n\|) + \epsilon_n(\|u - u_n\| + \|z - y_n - r_n\|), \quad \forall (u, z) \in \text{gph}Y + \mathbf{R}_{\geq}^p,$$

and $\epsilon_n \downarrow 0$. Moreover, clearly we have

$$\langle \sigma_n, u - u_n \rangle \leq o(\|u - u_n\| + \|y - y_n\|) + \epsilon_n(\|u - u_n\| + \|y - y_n\|), \quad \forall (u, y) \in \text{gph}Y,$$

which implies $\sigma \in N_{\text{gph}Y}(\bar{u}, \bar{y})$ (Theorem 1.2), contradicting pseudo-Lipschitz continuity of Y at (\bar{u}, \bar{y}) . \square

The contingent differentiability and the Proto differentiability of $Y + \mathbf{R}_{\geq}^p$ have been investigated in [51, 72]. In the following theorem, we present sufficient conditions for strict semi-differentiability of $Y + \mathbf{R}_{\geq}^p$.

Theorem 4.6. *Let $(\bar{u}, \bar{y}) \in \text{gph}V$. Assume that the following conditions hold:*

- (i) *Y is strictly semi-differentiable at (\bar{u}, \bar{y}) and Y is closed at \bar{u} ;*
- (ii) *$Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p = \{0\}$.*

Then, $Y + \mathbf{R}_{\geq}^p$ is strictly semi-differentiable at (\bar{u}, \bar{y}) and

$$D_{s\text{-low}}(Y + \mathbf{R}_{\geq}^p)(\bar{u}, \bar{y})(l) = D_{s\text{-low}}Y(\bar{u}, \bar{y})(l) + \mathbf{R}_{\geq}^p \quad \forall l \in \mathbf{R}^q. \quad (4.5)$$

Proof. The proof includes two parts. First we show that if $(l, e) \in T_{\text{gph}Y + \mathbf{R}_{\geq}^p}(\bar{u}, \bar{y})$, then $e \in D_{s\text{-low}}Y(\bar{u}, \bar{y})(l) + \mathbf{R}_{\geq}^p$. Next, we establish that $e \in D_{s\text{-low}}Y(\bar{u}, \bar{y})(l) + \mathbf{R}_{\geq}^p$ implies $e \in D_{s\text{-low}}(Y + \mathbf{R}_{\geq}^p)(\bar{u}, \bar{y})(l)$, which completes the proof.

let $(l, e) \in T_{\text{gph}(Y + \mathbf{R}_{\geq}^p)}(\bar{u}, \bar{y})$. By definition, there are sequences $\{(u_n, y_n + r_n)\} \subseteq \text{gph}(Y + \mathbf{R}_{\geq}^p)$ and $t_n \downarrow 0$ such that $t_n^{-1}(u_n - \bar{u}, y_n + r_n - \bar{y}) \rightarrow (l, e)$. Similar to the proof of Theorem 4.5, it can be shown that $r_n \rightarrow 0$. Then, if $t_n^{-1}r_n$ has a convergent subsequence, without loss of generality, we may assume that $t_n^{-1}r_n \rightarrow \bar{r}$. Hence,

$t_n^{-1}y_n \rightarrow \bar{e} \in D_{s-low}Y_{(\bar{u}, \bar{y})}(\zeta)$, which implies $e \in D_{s-low}Y(\bar{u}, \bar{y})(\zeta) + \mathbf{R}_{\geq}^p$, for otherwise, $t_n^{-1}r_n$ is unbounded. We may assume that $\|r_n\|^{-1}y_n$ converges to a nonzero vector $\bar{r} \in -\mathbf{R}_{\geq}^p$. Therefore,

$$\lim_{n \rightarrow \infty} (u_n - \bar{u}, \frac{y_n}{\|r_n\|}) = (0, \bar{r}),$$

which is impossible. Hence, $t_n^{-1}d_n$ has a convergent subsequence and the first part is established.

For the second part, let $e \in D_{s-low}Y(\bar{u}, \bar{y})(l)$, $d \in \mathbf{R}_{\geq}^p$, $l_n \rightarrow \bar{l}$ and $\{(u_n, y_n + r_n)\} \subseteq \text{gph}(Y + \mathbf{R}_{\geq}^p)$ such that $(u_n, y_n + r_n) \rightarrow (\bar{u}, \bar{y})$. We can show similar to the proof of Theorem 4.5 that $r_n \rightarrow 0$, and $y_n \rightarrow \bar{y}$. Since Y is strictly semi-differentiable at (\bar{u}, \bar{y}) , there exists $e_n \rightarrow e$ such that $\{(u_n + t_n l_n, y_n + t_n e_n)\} \subseteq \text{gph}Y$. Hence, $\{(u_n + t_n l_n, y_n + r_n + t_n e_n)\} \subseteq \text{gph}(Y + \mathbf{R}_{\geq}^p)$, and the proof is complete. \square

The property $Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p = \{0\}$ plays an important role here. It is readily seen that if Y is bounded from below on a neighborhood of \bar{u} then this property holds. In the following propositions, we provide some sufficient conditions for having this property.

Proposition 4.2. *Let $(\bar{u}, \bar{y}) \in \text{gph}V$ and let Y be closed at \bar{u} . Assume that there is a neighborhood N of \bar{u} such that $Y(u)$ is convex for each $u \in N$. If Y is pseudo-Lipschitz at (\bar{u}, \bar{y}) , then $Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p = \{0\}$.*

Proof. Suppose, contrary to our claim, that there is a nonzero vector $e \in Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p$. By definition, there are sequences $\{(u_n, y_n)\} \subseteq \text{gph}Y$ and $\{t_n\} \subseteq \mathbf{R}_{>}$ such that $(u_n, t_n y_n, t_n) \rightarrow (\bar{u}, e, 0)$. Y being pseudo-Lipschitz at (\bar{u}, \bar{y}) , there is a sequence $\{(u_n, z_n)\} \subseteq \text{gph}Y$ with $z_n \rightarrow \bar{y}$. Thus, for n sufficiently large $t_n y_n + (1 - t_n)z_n \in Y(u_n)$. This implies $\bar{y} - e \in Y(\bar{u})$ which contradicts efficiency of \bar{y} . Thus, $e = 0$ and the proof is complete. \square

Shi [67] and Lee [51] obtained formula (4.5) for contingent derivative and Proto derivative, respectively, under the assumption of $D_s Y(\bar{u}, \bar{y})(0) \cap -\mathbf{R}_{\geq}^p = \{0\}$ instead of hypotheses (ii). In the following proposition, we show that this hypothesis is satisfied when we have $D_s Y(\bar{u}, \bar{y})(0) \cap -\mathbf{R}_{\geq}^p = \{0\}$.

Proposition 4.3. *Let $\bar{y} \in V(\bar{u})$ be a given point. If $D_s Y(\bar{u}, \bar{y})(0) \cap -\mathbf{R}_{\geq}^p = \{0\}$, then*

$$(i) \quad DY_{(\bar{u}, \bar{y})}(0) \cap -\mathbf{R}_{\geq}^p = \{0\};$$

$$(ii) \quad Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p = \{0\}.$$

Proof. Suppose the assertion of the proposition is false. If we have a nonzero vector $e \in DY_{(\bar{u}, \bar{y})}(0) \cap -\mathbf{R}_{\geq}^p$, then $e \in D_s Y(\bar{u}, \bar{y})(0) \cap -\mathbf{R}_{\geq}^p$ which is impossible. Thus, there is a nonzero vector $e \in Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p$. By definition, there exist sequences $\{(u_n, y_n)\} \subseteq \text{gph} Y$ and $t_n \downarrow 0$ with $u_n \rightarrow \bar{u}$ and $t_n y_n \rightarrow e$. This implies $(t_n(u_n - \bar{u}), t_n(y_n - \bar{y})) \rightarrow (0, e)$ which contradicts our assumption. Thus, the proof is complete. \square

Shi [67] gave some conditions under which $D_s Y(\bar{u}, \bar{y})(0) \cap -\mathbf{R}_{\geq}^p = \{0\}$; see Proposition 3.2. We have this property when $\text{gph}(Y + \mathbf{R}_{\geq}^p)$ is convex; see [72]. As a result, the mentioned conditions imply $DY_{(\bar{u}, \bar{y})}(0) \cap -\mathbf{R}_{\geq}^p = \{0\}$ and $Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p = \{0\}$. The following example shows that the hypotheses (ii) and (iii) in Theorems 4.5 and 4.6 are essential for having pseudo-Lipschitz continuity and strict semi-differentiability, respectively.

Example 4.5. *Let*

$$Y(u) = \begin{cases} \{u^2, u^{-1}\}, & u < 0 \\ 0, & x \geq 1. \end{cases}$$

Note that Y is pseudo-Lipschitz and strictly semi-differentiable at $(0, 0)$, but $Y + \mathbf{R}_{\geq}$ is neither pseudo-Lipschitz nor strictly semi-differentiable at this point. Indeed, $D_{s\text{-low}}(Y + \mathbf{R}_{\geq})(0, 0)(-1) = \mathbf{R}_{\geq}$ but $D(Y + \mathbf{R}_{\geq})(0, 0)(-1) = \mathbf{R}$.

We refer the reader to Example 3.4 of [70] to see that proper efficiency and closedness are also necessary in Theorem 4.6.

In general, pseudo-Lipschitz continuity and strict semi-differentiability of Y and $Y + \mathbf{R}_{\geq}^p$ are not related to each other. Y may be pseudo-Lipschitz continuous and strictly semi-differentiable at a given point, while $Y + \mathbf{R}_{\geq}^p$ does not fulfill these properties at the given point, and vice versa. To check pseudo-Lipschitz and strict semi-differentiability of $Y + \mathbf{R}_{\geq}^p$, one can consider the following parametric multi-objective

optimization problem:

$$\begin{aligned} \min \quad & f(u, x) + r \\ \text{s.t.} \quad & g(u, x) \leq 0 \\ & r \geq 0, \end{aligned}$$

where x and r are decision variables and utilize related theorems.

4.3 Pseudo-Lipschitz continuity and semi-differentiability of the efficient set and the Pareto frontier

In this section, we study pseudo-Lipschitz continuity and semi-differentiability of the efficient solution and the efficient value mappings. Since strict semi-differentiability of the efficient solution and the efficient value mappings hold only under strict conditions, we only concentrate on semi-differentiability. For example, consider the optimization problem with $X(u) = Y(u) = \{(x_1, x_2) : 2x_1 + x_2 \geq 0, x_1 + x_2 \geq 0\}$. While X and Y are independent of parameter, neither V nor E is strictly semi-differentiable at $(0, (0, 0))$.

Definition 4.1. *Y is called locally dominated near \bar{u} , if for U neighborhood of \bar{u} we have*

$$Y(u) \subseteq V(u) + \mathbf{R}_{\geq}^p, \quad \forall u \in U.$$

The following lemma gives some conditions to guarantee local domination property for Y near a given point.

Lemma 4.2. *Let $\bar{y} \in V(\bar{u})$. Suppose that Y is closed-valued on a neighborhood of \bar{u} and Y is pseudo-Lipschitz at (\bar{u}, \bar{y}) . If $Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p = \{0\}$, then Y is locally dominated near \bar{u} . As a result, for some neighborhood of \bar{u} , V is nonempty.*

Proof. Suppose the result is false. Due to closed-valuedness and pseudo-Lipschitz continuity of Y , there are sequences $u_n \rightarrow \bar{u}$, $\{y_n\}$ and $\{z_n\}$ such that $y_n, z_n \in Y(u_n)$,

$z_n \leq y_n$ and $\|z_n\| \geq n + \|y_n\|^2$. Let us extract a subsequence $\{\|z_n\|^{-1}z_n\}$ without relabeling that converges to some nonzero vector r . Therefore,

$$\lim_{n \rightarrow \infty} (u_n, \|z_n\|^{-1}z_n) = (0, r).$$

Since $z_n \leq y_n$, $r \in -\mathbf{R}_{\geq}^p$ which contradicts our assumptions. In virtue of pseudo-Lipschitz continuity of Y , Y is nonempty in some neighborhood of \bar{u} . This implies non-emptiness of V in a neighborhood of \bar{u} . \square

Theorem 4.7. *Let Y be pseudo-Lipschitz at $(\bar{u}, \bar{y}) \in \text{gph}V$. If for some neighborhood $U \times O$ of (\bar{u}, \bar{y}) the following conditions hold, then V is pseudo-Lipschitz at this point:*

- (i) Y is closed on U and $\text{gph}V$ is closed around (\bar{u}, \bar{y}) .
- (ii) $DY(u, y)(0) \cap -\mathbf{R}_{\geq}^p = \{0\}$ for all $(u, y) \in (U \times O) \cap \text{gph}V$;
- (iii) $Y^\infty(u) \cap -\mathbf{R}_{\geq}^p = \{0\}$ for each $u \in U$.

Proof. We argue by contradiction. If it is false, then there exists $(\sigma, 0) \in N_{\text{gph}V}(\bar{u}, \bar{y})$ such that $\sigma \neq 0$. We consider two cases: either $(\sigma, 0)$ belongs to Fréchet normal cone or not.

Let $(\sigma, 0) \in \hat{N}_{\text{gph}V}(\bar{u}, \bar{y})$. By definition,

$$\limsup_{(u, y) \xrightarrow{\text{gph}V} (\bar{u}, \bar{y})} \frac{\langle \sigma, u - \bar{u} \rangle}{\|u - \bar{u}\| + \|y - \bar{y}\|} \leq 0. \quad (4.6)$$

Moreover, since Y is pseudo-Lipschitz at this point, $(\sigma, 0) \notin N_{\text{gph}V}(\bar{u}, \bar{y})$. This implies $(\sigma, 0) \notin \hat{N}_{\text{gph}V}(\bar{u}, \bar{y})$. Thus, there are $\epsilon > 0$ and $\{(u_n, y_n)\} \subseteq \text{gph}Y$ such that $(u_n, y_n) \rightarrow (\bar{u}, \bar{y})$ and

$$\lim_{n \rightarrow \infty} \frac{\langle \sigma, u_n - \bar{u} \rangle}{\|u_n - \bar{u}\| + \|y_n - \bar{y}\|} > 2\epsilon. \quad (4.7)$$

Y is locally dominated around \bar{u} (Lemma 4.2), and so we can choose a sequence $\{r_n\} \subseteq \mathbf{R}_{\geq}^p$ such that $y_n - r_n \in V(u_n)$ for n sufficiently large. Note that $\{r_n\}$ tends to zero, since if it were not the case, either $\{r_n\}$ has a convergent subsequence tending

to a nonzero vector or it tends to infinity. In the first case, because $y_n \rightarrow \bar{y}$, we obtain $\hat{y} \in Y(\bar{u})$ with $\hat{y} \leq \bar{y}$. This is a contradiction with $\bar{y} \in V(\bar{u})$. In the second case, reducing to subsequence if necessary,

$$\lim_{n \rightarrow \infty} (u_n, \frac{y_n - r_n}{\|r_n\|}) = (0, -r),$$

where $r \in \mathbf{R}_{\geq}^p$. This is a contradiction with condition (iii).

We can infer from (4.7) that for each $\delta > 0$ and for n sufficiently large, we have

$$\frac{\epsilon(\|u_n - \bar{u}\| + \|y_n - \bar{y}\|)}{\|u_n - \bar{u}\| + \|y_n - r_n - \bar{y}\|} \leq \frac{\langle \sigma, u_n - \bar{u} \rangle}{\|u_n - \bar{u}\| + \|y_n - r_n - \bar{y}\|} \leq \delta.$$

The above inequality and the triangle inequality imply

$$\lim_{n \rightarrow \infty} \frac{(u_n - \bar{u}, y_n - \bar{y})}{\|r_n\|} = (0, 0).$$

Without loss of generality, we assume that $\{\|r_n\|^{-1}r_n\}$ converges to a nonzero vector $r \in \mathbf{R}_{\geq}^p$. Thus, we have

$$\lim_{n \rightarrow \infty} \frac{(u_n - \bar{u}, y_n - r_n - \bar{y})}{\|r_n\|} = (0, -r),$$

which is a contradiction with $D_Y(\bar{u}, \bar{y})(0) \cap -\mathbf{R}_{\geq}^p = \{0\}$. Hence, if $(\theta, 0) \in \hat{N}_{gphV}(\bar{u}, \bar{y})$, then $\theta = 0$.

Now, we consider the other case. By definition, there are sequences $\{(u_n, y_n)\} \subseteq gphV$ and $(\sigma_n, \mu_n) \rightarrow (\sigma, 0)$ such that $(u_n, y_n) \rightarrow (\bar{u}, \bar{y})$ and $(\sigma_n, \mu_n) \in \hat{N}_{gphV}(u_n, y_n)$. Since Y is pseudo-Lipschitz at (\bar{u}, \bar{y}) , there is $\epsilon > 0$ such that for each sufficiently large n , there exists a sequence $\{(u_n^i, y_n^i)\} \subseteq gphY$ with

$$\lim_{i \rightarrow \infty} \frac{\langle \sigma_n, u_n^i - u_n \rangle + \langle \mu_n, y_n^i - y_n \rangle}{\|u_n^i - u_n\| + \|y_n^i - y_n\|} > 2\epsilon, \quad (4.8)$$

and $(u_n^i, y_n^i) \rightarrow (u_n, y_n)$. If it is not the case, then we obtain a contradiction $(\sigma, 0) \in N_{gphY}(\bar{u}, \bar{y})$. Similar to the first case, for sufficiently large n , we obtain a vector

$r \in -\mathbf{R}_{\geq}^p$ such that $r \in DY(u_n, y_n)(0)$. This leads to a contradiction with the assumptions, and the proof is complete. \square

In the following propositions, we provide a lower and upper estimation of $D_{low}V(\bar{u}, \bar{y})$.

Proposition 4.4. *Let $(\bar{u}, \bar{y}) \in gphV$. Then for each $l \in \mathbf{R}^q$,*

$$D_{low}V(\bar{u}, \bar{y})(l) \subseteq \{W \min D_{low}Y(\bar{u}, \bar{y})(l)\} \quad (4.9)$$

Proof. this can be proven similar to Theorem 3.3. of [49]. \square

Theorem 4.8. *Let $(\bar{u}, \bar{y}) \in gphV$. Assume that the following conditions hold:*

- (i) *Y is locally dominated around \bar{u} and Y is closed at \bar{u} ;*
- (ii) *$DY_{(\bar{u}, \bar{y})}(0) \cap -\mathbf{R}_{\geq}^p = \{0\}$;*
- (iii) *$Y^\infty(\bar{u}) \cap -\mathbf{R}_{\leq}^p = \{0\}$.*

Then for each $l \in \mathbf{R}^q$,

$$\min D_{low}Y(\bar{u}, \bar{y})(l) \subseteq D_{low}V(\bar{u}, \bar{y})(l). \quad (4.10)$$

Proof. Let $e \in \{\min D_{low}Y(\bar{u}, \bar{y})(l)\}$. From the assumptions, for $t_n \downarrow 0$ and $l_n \rightarrow l$, there is a sequence $e_n \rightarrow e$ such that $\{(\bar{u} + t_n l_n, \bar{y} + t_n e_n)\} \subseteq gphY$. By the assumptions, for sufficiently large n , we can construct a sequence $\{h_n\}$ such that $\{(\bar{u} + t_n l_n, \bar{y} + t_n h_n)\} \subseteq gphV$ and $\bar{y} + t_n h_n \leq \bar{y} + t_n e_n$. The sequence $\{t_n h_n\}$ tends to zero, since otherwise, either it tends to infinity or it has a convergent subsequence. Suppose that $\{t_n h_n\}$ tends to infinity and $\|h_n\|^{-1} h_n \rightarrow h$. Since $h_n \leq e_n$, we have $h \leq 0$. Thus,

$$\lim_{n \rightarrow \infty} \left(\bar{u} + t_n l_n, \frac{\bar{y} + t_n h_n}{t_n \|h_n\|} \right) = (0, h),$$

which contradicts condition (iii). For the second case, without loss of generality, we assume that $t_n h_n \rightarrow -r$. Similar to the former case it is shown that $r \in \mathbf{R}_{\geq}^p$. Hence, $\bar{y} - r \in Y(\bar{u})$ which contradicts $\bar{y} \in V(\bar{u})$. Thus, $\{t_n h_n\}$ tends to zero. The proof is

complete by showing that $\{h_n\}$ tends to e . Assume that this is not the case. Then, either it tends to infinity or it has a subsequence not converging to e . If $\{h_n\}$ tends to infinity, then we may assume that $\|h_n\|^{-1}h_n \rightarrow h \in -\mathbf{R}_{\geq}^p$. Consequently,

$$\lim_{n \rightarrow \infty} \left(\frac{\bar{u} + t_n l_n - \bar{u}}{t_n \|h_n\|}, \frac{\bar{y} + t_n h_n - \bar{y}}{t_n \|h_n\|} \right) = (0, h),$$

which contradicts condition (ii). Now, without loss of generality, assume that $h_n \rightarrow h$ with $h_n \neq e$. By virtue of the construction of $\{h_n\}$, $h \leq e$. Thus,

$$\lim_{n \rightarrow \infty} \left(\frac{\bar{u} + t_n l_n - \bar{u}}{t_n}, \frac{\bar{y} + t_n h_n - \bar{y}}{t_n} \right) = (l, h)$$

which is impossible. Therefore, $\{h_n\}$ tends to e , which is the desired result. \square

The following example shows that the upper and lower estimations of $D_{low}V(\bar{u}, \bar{y})(l)$ can be strict.

Example 4.6. Let $Y(u) = \{y \in \mathbf{R}^2 : y_1^2 + y_2^2 \leq 1\}$. Y and V are strictly semi-differentiable and semi-differentiable at $(\bar{u}, \bar{y}) = (0, (-1, 0))$, respectively, and $D_{low}V(\bar{u}, \bar{y})(1) = \{e \in \mathbf{R}^2 : e_1 = 0, e_2 \leq 0\}$ with

$$\begin{aligned} W \min DY_{s-low}Y(\bar{u}, \bar{y})(1) &= \{e \in \mathbf{R}_{\geq}^2 : e_1 = 0\}, \\ \min DY_{s-low}Y(\bar{u}, \bar{y})(1) &= \emptyset. \end{aligned}$$

Definition 4.2. We call $\bar{y} \in V(\bar{u})$ a uniformly proper efficient point, if there is a convex, closed and pointed cone K such that $\mathbf{R}_{\geq}^n \subseteq \text{int}(K) \cup \{0\}$. Furthermore, for some neighborhood N of (\bar{u}, \bar{y}) ,

$$Y(u) \cap (y - K) = \{y\}, \quad \forall (u, y) \in \text{gph}V \cap N.$$

The next theorem gives some conditions under which the inclusion given in Theorem 4.8 holds by an equality.

Proposition 4.5. If $\bar{y} \in V(\bar{u})$ is a uniformly proper efficient point, then for each

$l \in \mathbf{R}^q$ we have

$$D_{low}V(\bar{u}, \bar{y})(l) \subseteq \{\min D_{low}Y(\bar{u}, \bar{y})(l)\}. \quad (4.11)$$

Proof. Let l be given and $e \in D_{low}V(\bar{u}, \bar{y})(l)$. Suppose to the contrary that $e \notin \{\min D_{low}Y(\bar{u}, \bar{y})(l)\}$. Then there is some $e' \leq e$ with $e' \in D_{low}Y(\bar{u}, \bar{y})(l)$. Let $t_n \downarrow 0$ and $d_n \rightarrow d$ be given. By definition, there are $e_n \rightarrow e$ and $e'_n \rightarrow e'$ with $\{(\bar{u} + t_n d_n, \bar{y} + t_n e_n)\} \subseteq \text{gph}V$ and $\{(\bar{u} + t_n d_n, \bar{y} + t_n e'_n)\} \subseteq \text{gph}Y$. Thus, we have

$$\lim_{n \rightarrow \infty} \bar{y} + t_n e'_n - (\bar{y} + t_n e_n) = \lim_{n \rightarrow \infty} e'_n - e_n = e' - e.$$

This contradicts the uniform proper efficiency hypothesis. \square

Remark 4.3. *The Propositions 4.4, 4.8 and 4.5 hold when D_{low} is substituted by $D, D_{low}, D_{s-low}, D_{s-up}$ and D_a providing Y is semi-differentiable at (\bar{u}, \bar{y}) .*

The next theorem investigates the semi-differentiability of V at a given point.

Theorem 4.9. *Let the following conditions hold:*

- (i) Y is semi-differentiable at (\bar{u}, \bar{y}) and it is locally dominated around \bar{u} ;
- (ii) $DY_{(\bar{u}, \bar{y})}(0) \cap -\mathbf{R}_{\geq}^p = \{0\}$;
- (iii) $Y^\infty(\bar{u}) \cap -\mathbf{R}_{\geq}^p = \{0\}$.

If $\bar{y} \in V(\bar{u})$ is a uniformly proper efficient point, then V is semi-differentiable at (\bar{u}, \bar{y}) and

$$D_{low}V(\bar{u}, \bar{y})(l) = \{\min D_{low}Y(\bar{u}, \bar{y})(l)\}. \quad (4.12)$$

Proof. The proof involves two parts. First, we show that $e \in DV(\bar{u}, \bar{y})(l)$ implies $e \in \{\min D_{low}Y(\bar{u}, \bar{y})(l)\}$. Second, we show that $e \in \{\min D_{low}Y(\bar{u}, \bar{y})(l)\}$ leads to $e \in D_{low}V(\bar{u}, \bar{y})(l)$.

Let $e \in DV(\bar{u}, \bar{y})(l)$. From the definition of the Bouligand tangent cone, there are

$l_n \rightarrow l$, $e_n \rightarrow e$ and $t_n \downarrow 0$ with $\{(\bar{u} + t_n l_n, \bar{y} + t_n e_n)\} \subseteq gphV$. Due to the semi-differentiability of Y , $e \in D_{low}Y(\bar{u}, \bar{y})(l)$. If e belongs to $\{\min D_{low}Y(\bar{u}, \bar{y})(l)\}$, then this part is complete; otherwise, there is $(l, h) \in D_{s-low}Y(\bar{u}, \bar{y})(l)$ with $h \leq e$. As Y is semi-differentiable, there is a sequence h_n with $\{(\bar{u} + t_n l_n, \bar{y} + t_n h_n)\} \subseteq gphY$. Since \bar{y} is a uniformly proper efficient point of $Y(\bar{u})$, there is a cone K such that

$$Y(\bar{u} + t_n l_n) \cap (\bar{y} + t_n e_n - K) = \{\bar{y} + t_n e_n\}, \quad \forall n \geq n_0, \quad (4.13)$$

where $n_0 \in \mathbf{N}$. Since $h \leq e$, for sufficiently large n , this leads to a contradiction with (4.13).

The rest of proof follows from that of Theorem 4.8. □

In Theorem 4.9 if one substitutes semi-differentiability of Y for prorto-differentiability, then prorto-differentiability of V follows in a similar manner. This is stronger than the conditions given by Lee-Huy [51], only in the convex case. Tanino [71] proposed the notion of normal efficiency for the case that Y is a convex set-valued mapping ($gphY$ is a convex set). He proved the relation (4.12) with respect to contingent derivative under the convexity condition and normal efficiency. Theorem 4.9 can be adapted such that we have Tanino's result without convexity and normal efficiency. In the following lemma, we extend the definition of normal efficiency for the case that Y is convex-valued. Furthermore, we show that normal efficiency implies uniformly proper efficiency even for this extension.

Definition 4.3. *Let Y be convex-valued. $\bar{y} \in V(\bar{u})$ is said to be a normal proper point, if $N_{Y(\bar{u})}(\bar{y}) \subseteq -int(\mathbf{R}_{\geq}^p)$, where $N_{Y(\bar{u})}(\bar{y})$ is the normal cone in the sense of convex analysis.*

Theorem 4.10. *Let Y be convex-valued. Assume that Y is pseudo-Lipschitz at $(\bar{u}, \bar{y}) \in gphV$. If \bar{y} is normal proper, then it is uniformly proper efficient.*

Proof. Suppose the assertion of the lemma is false. Then, for each $\epsilon > 0$ there is a

sequence $\{(u_n, y_n)\} \subseteq \text{gph}V$ such that

$$\{0\} \not\subseteq (Y(u_n) - y_n) \cap -K_\epsilon,$$

where $K_\epsilon = \cap_{i=1}^p \{d : \langle e_i + \epsilon e, d \rangle \geq 0\}$, e_i is the unit vector of the standard basis and e is a vector with all components being to one. It is clear that if for some convex, closed and pointed cone $\mathbf{R}_{\geq}^p \subseteq \text{int}(K) \cup \{0\}$, then there is ϵ such that $K_\epsilon \subseteq K$. Moreover, by using Farkas' lemma, if we have

$$\begin{cases} Y(u_n) \cap (y_n - \mathbf{R}_{\geq}^p) = \{y_n\} \\ \{0\} \not\subseteq (Y(u_n) - y_n) \cap -K_\epsilon, \end{cases}$$

then there is $\lambda_n \geq 0$ such that $-\lambda_n \in N_{Y(u_n)}(y_n)$, $\sum_{i=1}^p \lambda_n^i = 1$ and $\min_{1 \leq i \leq p} \lambda_n^i \leq \epsilon$. We consider a sequence with $\epsilon_n \downarrow 0$. By Cantor's diagonal argument and the aforementioned fact, if necessary reducing to a subsequence, we can construct sequence $\{(u_n, y_n)\} \subseteq \text{gph}V$ and $\{\lambda_n\}$ such that for some $j \in \{1, \dots, p\}$

$$\begin{cases} \{0\} \not\subseteq (Y(u_n) - y_n) \cap -K_{\epsilon_n}, \\ (u_n, y_n) \rightarrow (\bar{u}, \bar{y}), \lambda_n \rightarrow \lambda, \\ \lambda_n \geq 0, \lambda_n^j = \min_{1 \leq i \leq p} \lambda_n^i \leq \epsilon_n, \\ -\lambda_n \in N_{Y(u_n)}(y_n), \sum_{i=1}^p \lambda_n^i = 1. \end{cases}$$

Since Y is pseudo-Lipschitz at (\bar{u}, \bar{y}) , for a neighborhood N of \bar{y} and for each $z \in Y(\bar{u}) \cap N$, there is a sequence $\{(u_n, z_n)\} \subseteq \text{gph}Y$ with $z_n \rightarrow z$. Thus, for sufficiently large n ,

$$\langle \lambda_n, z_n - y_n \rangle \geq 0.$$

This implies $-\lambda \in N_{Y(\bar{u}) \cap N}(\bar{u})$. Since $Y(u)$ is convex, $-\lambda \in N_{Y(\bar{u})}(\bar{u})$ which contradicts normal efficiency of \bar{y} . \square

The following example shows that we can have uniformly proper efficiency at a given point while not having normally proper.

Example 4.7. Let $Y(u) = \mathbf{R}_{\geq}^2$. It is easily seen that zero is K -minimal but $N_{Y(0)}(0) = \mathbf{R}_{\geq}^2$.

Proposition 4.6. Let $(\bar{u}, \bar{x}) \in \text{gph}E$ and let CQ1 hold at this point. If V is pseudo-Lipschitz continuous at $(\bar{u}, f(\bar{u}, \bar{x}))$, then E is pseudo-Lipschitz at the given point.

Proof. Since CQ1 holds at (\bar{u}, \bar{x}) , \hat{X} is pseudo-Lipschitz at this point. Thus, pseudo-Lipschitz continuity of E at (\bar{u}, \bar{x}) follows from pseudo-Lipschitz continuity of V . \square

Theorem 4.11. Let $(\bar{u}, \bar{x}) \in \text{gph}E$. Assume that the following assumptions hold:

- (i) f is regular at (\bar{u}, \bar{x}) ;
- (ii) CQ1 holds at (\bar{u}, \bar{x}) ;
- (iii) V is pseudo-Lipschitz and semi-differentiable at $(\bar{u}, f(\bar{u}, \bar{x}))$ and $\text{gph}V$ is closed around it.

Then, E is semi-differentiable at (\bar{u}, \bar{x}) and

$$D_{\text{low}}E(\bar{u}, \bar{x})(l) = \{d \in D_{s\text{-low}}X(\bar{u}, \bar{x})(l) : C'(\bar{u}, \bar{x}; l, d) \in D_{\text{low}}V_{(\bar{u}, \bar{y})}(\bar{u}, \bar{x})(l)\} \quad (4.14)$$

Proof. $E(u)$ can be shown as $\{x \in \mathbf{R}^n : g(u, x) \in K\}$, where $g(u, x) = \begin{bmatrix} (u, x) \\ (u, C(u, x)) \end{bmatrix}$ and $K = \text{gph}X \times \text{gph}V$. It follows from Remark 4.1 and Proposition 4.1 that X is pseudo-Lipschitz and strictly semi-differentiable at (\bar{u}, \bar{x}) . From the assumptions and the mentioned facts, all the assumptions of Corollary 5.7 in [2] are fulfilled. (here, differentiability can be substituted by regularity). Thus, E is Proto differentiable at (\bar{u}, \bar{x}) and we have (4.14). Since E is pseudo-Lipschitz at the given point, it is also semi-differentiable at (\bar{u}, \bar{x}) . \square

Bibliography

- [1] ABDELAZIZ, F. B. Solution approaches for the multiobjective stochastic programming. *European Journal of Operational Research* 216, 1 (2012), 1–16.
- [2] AMAHROQ, T., AND THIBAUT, L. On proto-differentiability and strict proto-differentiability of multifunctions of feasible points in perturbed optimization problems. *Numerical functional analysis and optimization* 16, 9-10 (1995), 1293–1307.
- [3] AUBIN, J. P. Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions. Tech. rep., DTIC Document, 1980.
- [4] AUBIN, J.-P. Lipschitz behavior of solutions to convex minimization problems. *Mathematics of Operations Research* 9, 1 (1984), 87–111.
- [5] BEDNARCZUK, E. *Stability Analysis for Parametric Vector Optimization Problems*, vol. 442. Institute of Mathematics, Polish Academy of Sciences, 2007.
- [6] BEDNARCZUK, E. M. On lower lipschitz continuity of minimal points. *Discussiones Mathematicae, Differential Inclusions, Control and Optimization* 20, 2 (2000), 245–255.
- [7] BEN-TAL, A., EL GHAOUI, L., AND NEMIROVSKI, A. *Robust Optimization*. Princeton University Press, 2009.
- [8] BENSON, H. P. Existence of efficient solutions for vector maximization problems. *Journal of Optimization Theory and Applications* 26, 4 (1978), 569–580.

- [9] BENSON, H. P. An improved definition of proper efficiency for vector maximization with respect to cones. *Journal of Mathematical Analysis and Applications* 71, 1 (1979), 232–241.
- [10] BERTSIMAS, D., BROWN, D. B., AND CARAMANIS, C. Theory and applications of robust optimization. *SIAM review* 53, 3 (2011), 464–501.
- [11] BOKRANTZ, R., AND FREDRIKSSON, A. On solutions to robust multiobjective optimization problems that are optimal under convex scalarization. *arXiv preprint arXiv:1308.4616* (2013).
- [12] BONNANS, J. F., AND SHAPIRO, A. *Perturbation Analysis of Optimization Problems*. Springer Science & Business Media, 2013.
- [13] CHUONG, T. D. Clarke coderivatives of efficient point multifunctions in parametric vector optimization. *Nonlinear Analysis: Theory, Methods & Applications* 74, 1 (2011), 273–285.
- [14] CHUONG, T. D. Derivatives of the efficient point multifunction in parametric vector optimization problems. *Journal of Optimization Theory and Applications* 156, 2 (2013), 247–265.
- [15] CHUONG, T. D., AND YAO, J.-C. Generalized clarke epiderivatives of parametric vector optimization problems. *Journal of Optimization Theory and Applications* 146, 1 (2010), 77–94.
- [16] CLARKE, F. H. *Optimization and Nonsmooth Analysis*, vol. 5. Siam Publications, 1990.
- [17] CLARKE, F. H. *Functional Analysis, Calculus of Variations and Optimal Control*, vol. 264. Springer Science & Business Media, 2013.
- [18] DEB, K. *Multi-Objective Optimization Using Evolutionary Algorithms*, vol. 16. John Wiley & Sons, 2001.

- [19] DEB, K., AND GUPTA, H. Introducing robustness in multi-objective optimization. *Evolutionary Computation* 14, 4 (2006), 463–494.
- [20] DONTCHEV, A. L., AND ROCKAFELLAR, R. T. *Implicit Functions and Solution Mappings*, vol. 194. Springer, 2009.
- [21] DORSCH, D., JONGEN, H. T., RÜCKMANN, J.-J., AND SHIKHMAN, V. On the local representation of piecewise smooth equations as a lipschitz manifold. *Journal of Mathematical Analysis and Applications* 411, 2 (2014), 916–930.
- [22] DRUMMOND, L. G., AND IUSEM, A. N. A projected gradient method for vector optimization problems. *Computational Optimization and Applications* 28, 1 (2004), 5–29.
- [23] ECKER, J. G., AND KOUADA, I. Finding efficient points for linear multiple objective programs. *Mathematical Programming* 8, 1 (1975), 375–377.
- [24] EHRGOTT, M. *Multicriteria optimization*. Springer Science & Business Media, 2006.
- [25] EHRGOTT, M., IDE, J., AND SCHÖBEL, A. Minmax robustness for multi-objective optimization problems. *European Journal of Operational Research* 239, 1 (2014), 17–31.
- [26] EICHFELDER, G. *Adaptive Scalarization Methods in Multiobjective Optimization*. Springer, 2008.
- [27] EVANGELOS, T. Multi-criteria decision making methods: A comparative study. *Netherland: Kluwer Academic Publication* (2000).
- [28] FLIEGE, J., AND SVAITER, F. B. Steepest descent methods for multicriteria optimization. *Mathematical Methods of Operations Research* 51, 3 (2000), 479–494.
- [29] FRANKLIN, J. N. *Methods of Mathematical Economics: Linear and Nonlinear Programming, Fixed-Point Theorems*, vol. 37. SIAM, 2002.

- [30] GEARHART, W. Compromise solutions and estimation of the noninferior set. *Journal of Optimization Theory and Applications* 28, 1 (1979), 29–47.
- [31] GEOFFRION, A. M. Proper efficiency and the theory of vector maximization. Tech. rep., DTIC Document, 1967.
- [32] GEORGIEV, P. G., LUC, D. T., AND PARDALOS, P. M. Robust aspects of solutions in deterministic multiple objective linear programming. *European Journal of Operational Research* 229, 1 (2013), 29–36.
- [33] GIANNESI, F., MASTROENI, G., AND PELLEGRINI, L. On the theory of vector optimization and variational inequalities. image space analysis and separation. In *Vector variational inequalities and vector equilibria*. Springer, 2000, pp. 153–215.
- [34] GIORGI, G., GUERRAGGIO, A., AND GUERRAGGIO, A. On the notion of tangent cone in mathematical programming. *Optimization* 25, 1 (1992), 11–23.
- [35] GOBERNA, M. A., JEYAKUMAR, V., LI, G., AND VICENTE-PÉREZ, J. Robust solutions to multi-objective linear programs with uncertain data. *European Journal of Operational Research* 242, 3 (2015), 730–743.
- [36] GOBERNA, M. A., JEYAKUMAR, V., LI, G., AND VICENTE-PÉREZ, J. Robust solutions of multiobjective linear semi-infinite programs under constraint data uncertainty. *SIAM Journal on Optimization* 24, 3 (2014), 1402–1419.
- [37] GOBERNA, M. A., AND LÓPEZ, M. A. *Linear Semi-Infinite Optimization*. Wiley, 1998.
- [38] GOERIGK, M., AND SCHÖBEL, A. Algorithm engineering in robust optimization. *arXiv preprint arXiv:1505.04901* (2015).
- [39] GUDDAT, J. *Multiobjective and Stochastic Optimization Based on Parametric Optimization*, vol. 26. Akademie-Verlag, 1985.

- [40] GUERRAGGIO, A., MOLHO, E., AND ZAFFARONI, A. On the notion of proper efficiency in vector optimization. *Journal of Optimization Theory and Applications* 82, 1 (1994), 1–21.
- [41] GUTJAHR, W. J., AND PICHLER, A. Stochastic multi-objective optimization: a survey on non-scalarizing methods. *Annals of Operations Research* (2013), 1–25.
- [42] HENIG, M. Proper efficiency with respect to cones. *Journal of Optimization Theory and Applications* 36, 3 (1982), 387–407.
- [43] IDE, J., AND KÖBIS, E. Concepts of efficiency for uncertain multi-objective optimization problems based on set order relations. *Mathematical Methods of Operations Research* 80, 1 (2014), 99–127.
- [44] IDE, J., AND SCHÖBEL, A. Robustness for uncertain multi-objective optimization: a survey and analysis of different concepts. *OR Spectrum* 38, 1 (2016), 235–271.
- [45] JEYAKUMAR, V., AND LUC, D. T. *Nonsmooth Vector Functions and Continuous Optimization*, vol. 10. Springer Science & Business Media, 2007.
- [46] KALISZEWSKI, I. *Quantitative Pareto Analysis by Cone Separation Technique*. Springer Science & Business Media, 2012.
- [47] KHAN, A. A., TAMMER, C., AND ZĂLINESCU, C. *Set-valued Optimization*. Springer, 2015.
- [48] KUHN, H., AND TUCKER, A. Nonlinear programming. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability* (1951), vol. 1, pp. 481–492.
- [49] KUK, H., TANINO, T., AND TANAKA, M. Sensitivity analysis in vector optimization. *Journal of Optimization Theory and Applications* 89, 3 (1996), 713–730.

- [50] KUROIWA, D., AND LEE, G. M. On robust multiobjective optimization. *Vietnam J. Math* 40, 2&3 (2012), 305–317.
- [51] LEE, G., AND HUY, N. On sensitivity analysis in vector optimization. *Taiwanese Journal of Mathematics* 11, 3 (2007), pp–945.
- [52] LI, X., WANG, Q., AND LIN, Z. Stability results for properly quasi convex vector optimization problems. *Optimization* 64, 5 (2015), 1329–1347.
- [53] LUC, D. T. Theory of vector optimization, 1989.
- [54] LUC, D. T. *Multiobjective Linear Programming*. Springer International Publishing, 2016.
- [55] LUC, D. T., AND PENOT, J.-P. Convergence of asymptotic directions. *Transactions of the American Mathematical Society* 353, 10 (2001), 4095–4121.
- [56] MARLER, R., AND ARORA, J. Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization* 26, 6 (2004), 369–395.
- [57] MIETTINEN, K. *Nonlinear Multiobjective Optimization*, vol. 12. Springer Science & Business Media, 2012.
- [58] MIGLIERINA, E., AND MOLHO, E. Convergence of minimal sets in convex vector optimization. *SIAM Journal on Optimization* 15, 2 (2005), 513–526.
- [59] MORDUKHOVICH, B. S. *Variational Analysis and Generalized Differentiation I: Basic Theory*, vol. 330. Springer Science & Business Media, 2006.
- [60] PENOT, J.-P. Differentiability of relations and differential stability of perturbed optimization problems. *SIAM Journal on Control and Optimization* 22, 4 (1984), 529–551.
- [61] POURKARIMI, L., AND SOLEIMANI-DAMANEH, M. Robustness in deterministic multi-objective linear programming with respect to the relative interior and angle deviation. *Optimization* (2016).

- [62] ROCKAFELLAR, R. T. Lipschitzian properties of multifunctions. *Nonlinear Analysis: Theory, Methods & Applications* 9, 8 (1985), 867–885.
- [63] ROCKAFELLAR, R. T. Proto-differentiability of set-valued mappings and its applications in optimization. In *Annales de l'IHP Analyse non linéaire* (1989), vol. 6, pp. 449–482.
- [64] ROCKAFELLAR, R. T., AND WETS, R. J.-B. *Variational analysis*, vol. 317. Springer Science & Business Media, 2009.
- [65] SAWARAGI, Y NAKAYAMA, H., AND TANINO, T. *Theory of multiobjective optimization*, vol. 176. Elsevier, 1985.
- [66] SHAPIRO, A., DENTCHEVA, D., AND RUSZCZYŃSKI, A. *Lectures on stochastic programming: modeling and theory*, vol. 16. SIAM, 2014.
- [67] SHI, D. S. Contingent derivative of the perturbation map in multiobjective optimization. *Journal of Optimization Theory and Applications* 70, 2 (1991), 385–396.
- [68] SOYSTER, A. L. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research* 21, 5 (1973), 1154–1157.
- [69] TAA, A. Set-valued derivatives of multifunctions and optimality conditions. *Numerical Functional Analysis and Optimization* 19, 1-2 (1998), 121–140.
- [70] TANINO, T. Sensitivity analysis in multiobjective optimization. *Journal of Optimization Theory and Applications* 56, 3 (1988), 479–499.
- [71] TANINO, T. Stability and sensitivity analysis in convex vector optimization. *SIAM Journal on Control and Optimization* 26, 3 (1988), 521–536.
- [72] TANINO, T. Sensitivity analysis in mcdm. In *Multicriteria Decision Making*. Springer, 1999, pp. 173–201.

- [73] THIBAUT, L. Tangent cones and quasi-interiorly tangent cones to multifunctions. *Transactions of the American Mathematical Society* 277, 2 (1983), 601–621.
- [74] WALLENIS, J., AND ZIONTS, S. *Multiple Criteria Decision Making: From Early History to the 21st Century*. World Scientific, 2011.
- [75] WIECEK, M. M., EHRGOTT, M., AND ENGAU, A. Continuous multiobjective programming. In *Multiple Criteria Decision Analysis*. Springer, 2016, pp. 739–815.