MODIFIED PROXIMAL ALGORITHMS FOR COMMON SOLUTIONS OF MULTI-OBJECTIVE OPTIMIZATION AND FIXED POINT PROBLEMS

FOUZIA AMIR

A Thesis Submitted to the Graduate School of Naresuan University
in Partial Fulfillment of the Requirements
for the Doctor of Philosophy Degree in Mathematics
July 2020

This thesis entitled "Modified proximal algorithms for common solutions of multi-objective optimization and fixed point problems"

by Fouzia Amir

has been approved by the Graduate School as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Mathematics of Naresuan University

Oral Defense Committee

.... Chair
(Professor Suthep Suantai, Ph.D.)

.... Advisor
(Associate Professor Narin Petrot, Ph.D.)

.... Co-Advisor
(Professor Ali Farajzadeh, Ph.D.)

.... Internal Examiner
(Associate Professor Anchalee Kaewcharoen, Ph.D.)

.... Internal Examiner
(Associate Professor Rabian Wangkeeree, Ph.D.)

.... External Examiner
(Assistant Professor Nimit Nimana, Ph.D.)

Approved

....
(Professor Paisarn Muneesawang, Ph.D.)
Dean of the Graduate School
15 JUL 2020

ACKNOWLEDGEMENT

All the glories are to Almighty ALLAH, the Most Merciful and Compassionate, the Most Gracious and Beneficent, Who through his Kindness, Graciousness and countless blessings enabled me to pursue higher ideals in life. His founding lesson passed by His messenger Muhammad (peace be upon him), "Seek knowledge from the candle to the grave", always gave me pleasure to try my level best to seek knowledge and apply it usefully.

My heartiest profound gratitude is to my respected supervisor Associate Prof. Dr. Narin Petrot for his constant and scholarly guidance, enriching and enlightening suggestions, precious and insightful advice, which not only enabled me to carry out this research work and bring my thesis into final shape.

I also humbly acknowledge insightful and inspiring guidance of Respected Prof. Dr. Ali Farajzadeh for his scholarly, perceptive and insightful guidance.

I also acknowledge the open and warm hearted support of Pornpimon Boriwan, Porntip and Warut who always spared precious hours for me inspite of their own hectic schedule.

I also appreciate mentioning that my graduate study was financial supported by the Naresuan University, Phitsanulok, Thailand.

At the end, this task would have been really difficult without the prayers and wishes of my husband (Mutti Ur Rehman) and my parents.

Fouzia Amir

Title	MODIFIED PROXIMAL ALGORITHMS FOR COMMON SOLUTIONS OF MULTI-OBJECTIVE OPTIMIZATION AND FIXED POINT PROBLEMS
Author	Fouzia Amir
Advisor	Associate Professor Dr. Narin Petrot, Ph.D.
Academic Paper	Thesis Ph.D. in Mathematics, Naresuan University, 2019.
Keywords	Proximal point algorithm, convex function, quasi convex function, fixed point, nonexpansive mapping, Mann iteration, locally Lipschitz, multiobjective optimization

ABSTRACT

In this dissertation, there are two main folds of the considered problems. That are the unconstrained multi-objective and constrained multi-objective optimization problems. For the first fold, by using the scalarizing techniques, we suggest the algorithm for finding the solution concept as Pareto critical point for smooth case of the considered problem. In the second fold, we suggest of finding ϵ-quasi weak Pareto optimal point for non-smooth case of the considered problem. Also, inspired by the Mann iteration technique, we mainly focus to the situation when the constraint set is a fixed point set of a class of nonlinear mappings with a special structure as nonexpansive property. The discussions on the obtained methods and the relations to the appeared methods which had appeared in the literatures are provided and pointed out.

LIST OF CONTENTS

CHAPTER I

INTRODUCTION

Life is about decisions. Decisions, no matter if made by a group or an individual, usually involve several conflicting objectives. The observations that real world problems have to be solved optimally according to criteria, which prohibit an "ideal" solution, optimal for each decision-maker under each of the criteria considered, has led to the development of multiobjective optimization. Optimization is central to any problem involving decision making, whether in engineering or in economics. The task of decision making entails choosing between various alternatives. This choice is governed by our desire to make the best decision. The measure of goodness of the alternatives is described by an objective function or performance index. Optimization theory and methods deals with selecting of the best alternative in the sense of the given objective function.

Rigorous mathematical analysis of the optimization problem was carried out during the 20th century, the roots can be traced back to about 300 B.C; when the Greek mathematician Euclid evaluated the minimum distance between a point and a line. Optimization may be regarded as the cornerstone of many areas of applied mathematics, computer science, engineering, and a number of other scientific disciplines. Among other things, optimization plays a key role in finding feasible solutions to real-life problems, from mathematical programming to operations research, economics, management science, business, medicine, life science, and artificial intelligence.

Optimization deals with the study of those kinds of problems in which one has to minimize or maximize one or more objectives that are functions of some real or integer variables. This is executed in a systematic way by choosing the proper values of real or integer variables within an allowed set. Given a defined domain, the main goal of optimization is to study the means of obtaining the best value of some objective function. Single objective optimization deals with the task of finding the "best" solution, which corresponds to the minimum or maximum value of a single objective function that lumps all different objectives into one. This type of optimization is useful as a tool which should provide decision makers with insights into the nature of the problem, but usually cannot provide a set of alternative solutions that trade different objectives against each other. On the contrary, multiobjective optimization (MOP) (multicriteria or multiattribute optimization) deals with the task of simultaneously optimizing two or more conflicting objectives with respect to a set of certain constraints. If the optimization of one objective leads to the automatic optimization of the other, it should not be considered as MOP problem.

However, in many real-life situations we come across problems where an attempt to improve one objective leads to degradation of the other. Such problems belong to the class of MOP problems and appear in several fields including product and process design, network analysis, finance, aircraft design, bioinformatics, the oil and gas industry, automobile design, etc.

Compared to single objective problems, MOP problems are more difficult to solve, because there is no unique solution; rather, there is a set of acceptable trade-off optimal solutions. This set is called Pareto front. MOP is in fact considered as the analytical phase of the multi-objective decision making process, and consists of determining all solutions to the MOP problem that are optimal in the Pareto sense. The preferred solution, the one most desirable to the designer or decision maker, is selected from the Pareto set.

Due to increasing interest in solving real-world MOP problems using proximal point algorithms (PPA), researchers have developed a number of proximal point multi-objective algorithms. The presence of multiple objectives in a problem, in principle, gives rise to a set of optimal solutions (largely known as Pareto-optimal solutions), instead of a single optimal solution. In the absence of any further information, one of these Pareto-optimal solutions cannot be said to be better than the other. This demands a user to find as many Pareto-optimal solutions as possible. Classical optimization methods (including the multi-criterion decision-making methods) suggest converting the multiobjective optimization problem to a single objective optimization problem by emphasizing one particular Pareto-optimal solution at a time. When such a method is to be used for finding multiple solutions, it has to be applied many times, hopefully finding a different solution at each simulation run. There are many algorithms for solving multiobjective optimization problems.

The proximal point algorithm is a widely used tool for solving a variety of convex optimization problems. The algorithm works by applying successively so-called "resolvent" mappings associated to the original object that one aims to optimize. The first instance of what came later to be known as the proximal point algorithm can be found in a short communication from 1970 of Martinet [1]. Starting with the pioneering paper of Rockafellar [2], which clearly fix some existing ideas in the previous literature and gives much more insights on the potential of the algorithm when applies to optimization problems, an important literature has grown on possible extensions and generalizations of this algorithm (see, for example the survey paper [3] and the references therein). Some attention was focused also on the case of multi-objective optimization, see [4–8].

On the other hand, Fixed point theory is a fascinating subject, with an enormous number of applications in various fields of mathematics. Fixed point theory concerns itself with a very simple and basic mathematical setting. It is one of the most powerful and fruitful tools of modern mathematics and may be considered as a core subject of nonlinear analysis. The presence or absence of fixed point is an intrinsic property of a function. However, many necessary and/or sufficient conditions for the existence of such points involve a mixture of algebraic order theoretic or topological properties of mapping or its domain. In a wide range of mathematical problems the existence of a solution is equivalent to the existence of a fixed point for a suitable map. The existence of a fixed point is therefore of paramount importance in several areas of mathematics and other sciences. Fixed point results provide conditions under which maps have solutions. The theory itself is a beautiful mixture of analysis (pure and applied), topology, and geometry. Over the last 50 years or so the theory of fixed points has been revealed as a very powerful and important tool in the study of nonlinear phenomena. In particular, fixed point techniques have been applied in such diverse fields as biology, chemistry, economics, engineering, game theory, and physics. A physical example of a fixed point of a mapping is the center of a whirlpool in a cup of tea when it is stirred. (The fact that the center of the whirlpool moves over time is just due to the fact that the mapping is changing over time.)

It was an integral part of topology at the very birth of the subject in the work of Poincare in the 1880. In the last 50 years, fixed point theory has been flourishing area of research for many mathematicians. The origins of the theory, which dated back to the later part of the nineteenth century, rest in the use of successive approximations used for proving existence of solutions of differential equations introduced independently by Joseph Liouville [9] in 1837 and Charles Emile Picard [10] in 1890. But formally it was started in the beginning of twentieth century as an important part of analysis. The abstraction of this classical theory is the pioneering work of the great Polish mathematician Stefan Banach [11] published in 1922, which provides a constructive method to find the fixed points of a map.

Fixed point theory has played central role in the problems of nonlinear functional analysis and fixed point theorems have provided powerful tools in demonstrating the existence of solution to a large variety of problems in applied mathematics. Fixed point theorems are mainly useful in existence theory for the solutions of differential equations, integral equations, partial differential equations, random differential equations. Also the theory has numerous applications in other related areas like control theory, game theory, economics etc. Besides this, fixed point theory has very fruitful applications in eigenvalue value problems, boundary value problems and best approximation problems. Many existence theorem of analysis can be treated as special cases of suitable fixed point theorems.

This thesis is organized as follows.

In Chapter 2, definitions and auxiliary results are presented as well as nonsmooth analysis.

In Chapter 3, proximal point algorithm for obtaining the solution of a multi-objective optimization problem introduced by Bento et al. [5] is studied, under some weaker conditions, the well-definedness of the generated sequence of algorithm is established when the objective function is locally Lipschitz, and the convergence to a Pareto critical point of the objective function is proved.

In Chapter 4, inexact proximal point algorithm for obtaining the solution of a multi-objective optimization problem is introduced, the well-definedness of the generated sequence of algorithm is established when the objective function is locally Lipschitz, and convergence theorems of the introduced algorithm is presented.

In Chapter 5, modified proximal point algorithm involving fixed point iterates and multi-objective optimization is introduced. The convergent behaviour of the introduced algorithm is studied and discussed. Finally, in Chapter 6, we present proposals for future research and conclusions.

CHAPTER II

PRELIMINARIES CONCEPTS

In this chapter, we give a brief introduction of elementary concepts and some results, which will be used throughout the next chapters.

In this thesis, $\mathbb{R}$ and $\mathbb{N}$ denote the set of all real numbers and the set of all natural numbers, respectively. The notation $\|.\|$ denotes the Euclidean norm and both $\langle a, b\rangle$ or $a^T b$ stand for inner product of a, b in Euclidean space $\mathbb{R}^n$.

The main goal of optimization is to study the means of obtaining the best value of some objective functions. An optimization problem can be defined as follows:
Consider a function $h : \mathbb{R}^n \to \mathbb{R}$, the aim is to determine an element $x^1 \in \mathbb{R}^n$ such that

$$h(x^1) \leq h(x), \quad \forall x \in \mathbb{R}^n,$$

is called the minimization and

$$h(x^1) \geq h(x), \quad \forall x \in \mathbb{R}^n,$$

is known as maximization. Here the domain of h, is called the search space, and the elements of $\mathbb{R}^n$ are called candidate or feasible solutions. The function h is called an objective function/cost function/energy function. A feasible solution that optimizes the objective function is called an optimal solution.

2.1 Multiobjective optimization and optimal points

Now, we remind some basic concepts and properties of multiobjective optimization (MOP), which can be found in [14].
Let $I := \{1, ..., m\}$, then

$$\mathbb{R}^m_+ = \{x \in \mathbb{R}^m : x_j \geq 0, j \in I\},$$

$$\mathbb{R}^m_{++} = \{x \in \mathbb{R}^m : x_j > 0, j \in I\}.$$

Throughout the thesis, the notations $\preceq$ and $\prec$ stand for the following orders on $\mathbb{R}^m$ for $y, z \in \mathbb{R}^m$,

$$z \succeq y (\text{or } y \preceq z) \text{ means that } z - y \in \mathbb{R}^m_+,$$

and

$$z \succ y (\text{or } y \prec z) \text{ means that } z - y \in \mathbb{R}^m_{++}.$$

In MOP, one has to optimize several objective functions simultaneously. Having several objective functions, the notation of "optimum" changes in MOP, because in MOP, the goal is to find compromises (or "tradeoff") rather than a single solution as in global optimization. The notion of "optimum" most commonly adopted is that of originally proposed by Francis Ysidro Edgeworth and later generalized by Vilfredo Pareto.

We consider the following general MOP:

$$\min_{x \in C} H(x),$$

where $H : C \subset \mathbb{R}^n \to \mathbb{R}^m$ is a vector function with $H(x) := (h_1(x), h_2(x), ..., h_m(x))$, $h_i : C \to \mathbb{R}$ and C is nonempty closed set.

The solution notion for MOP is defined with respect to an ordering cone which is used for ordering the criterion space $\mathbb{R}^m$.

Definition 2.1.1. (Pareto optimal point) A point $x^* \in C$ is called Pareto optimal point of H, if there exists no other $x \in C$ such that

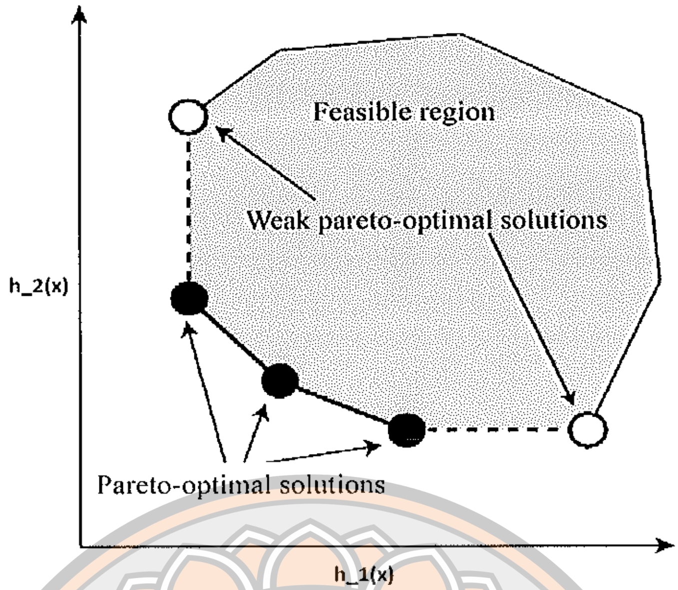
$$H(x) \preceq H(x^*),$$

and $H(x) \neq H(x^*)$,

Definition 2.1.2. (weak Pareto optimal point) A point $x^* \in C$ is said to be weak Pareto optimal point of H, if there exists no other $x \in C$ such that

$$H(x) \prec H(x^*).$$

If we have two objectives $h_1(x)$ and $h_2(x)$, then we can see the Pareto and weak Pareto solutions in the below figure:



Let $\epsilon := (\epsilon_1, ..., \epsilon_m) \in \mathbb{R}^m_+$.

Definition 2.1.3. (ϵ-weak Pareto optimal point) A point $x^* \in C$ is called ϵ-weak Pareto optimal point of H, if there exists no other $x \in C$ such that

$$H(x) + \epsilon \prec H(x^*).$$

Definition 2.1.4. (ϵ-quasi weak Pareto optimal point) A point $x^* \in C$ is known as an ϵ-quasi weak Pareto point of H if there is no $x \in C$ such that

$$H(x) + \epsilon\|x - x^*\| \prec H(x^*).$$

We denote the set of Pareto, weak Pareto, ϵ-weak Pareto and ϵ-quasi weak Pareto solutions of H by $\arg\min\{H(x)|x \in C\}$, $\arg\min_w\{H(x)|x \in C\}$, $\arg\min_{\epsilon w}\{H(x)|x \in C\}$ and $\text{argmin}_{\epsilon q-w}\{H(x)|x \in C\}$, respectively. For more details, we refer to [12] and [13].

Remark 2.1.5. It is noted that Pareto optimal point is also a weak Pareto optimal point but the converse is not true. Also, it is apparent that, if $\epsilon = 0$, then the notions of an ϵ-weakly Pareto solution and an ϵ-weakly quasi Pareto solution defined above coincide with the usual one of a weak Pareto solution. Also, for the case, $\epsilon \neq 0$ we can see that, $\arg\min_w\{H(x)|x \in C\} \subset \arg\min_{\epsilon w}\{H(x)|x \in C\}$ and $\arg\min_w\{H(x)|x \in C\} \subset \arg\min_{\epsilon q-w}\{H(x)|x \in C\}$. While, the sets

$\arg\min_{\epsilon w}\{H(x)|x \in C\}$ and $\arg\min_{\epsilon q-w}\{H(x)|x \in C\}$ might be two different sets. For more details see [13].

For a multi-objective mapping $H : \mathbb{R}^n \to \mathbb{R}^m$, we say that H has a directional derivative at $x \in \mathbb{R}^n$ in the direction of $v \in \mathbb{R}^n$ if

$$D_v H(x) = \lim_{t \to 0} \frac{H(x+tv) - H(x)}{t}.$$

If vector function H is differentiable at $x \in \mathbb{R}^n$, then all of its directional derivatives at x exist. For a differentiable function H, we denote the Jacobian of H at $x \in \mathbb{R}^n$ by

$$JH(x) := (\nabla h_1(x), ..., \nabla h_m(x)),$$

and the image of the Jacobian of H at a point $x \in \mathbb{R}^n$ by

$$Im(JH(x)) := \{JH(x)v = (\langle \nabla h_1(x), v\rangle, ..., \langle \nabla h_m(x), v\rangle), \quad v \in \mathbb{R}^n\}. \quad (2.1.1)$$

A vector function H is differentiable at x, iff all of its components are differentiable at x. For a differentiable function H, $D_v H(x) = JH(x)v$.
The first order optimality condition for problem (2.2) is given by

$$x \in \mathbb{R}^n, \quad Im(JH(x)) \cap (-\mathbb{R}^m_{++}) = \emptyset. \quad (2.1.2)$$

In general, (2.1.2) is necessary, but not sufficient condition, for optimality. A point $x \in \mathbb{R}^n$ satisfying (2.1.2) is called a Pareto critical point (see for instance, [14]).

2.2 Fixed point iteration scheme

Let a mapping $T : \mathbb{R}^n \to \mathbb{R}^n$. Then

$$F(T) = \{x \in \mathbb{R}^n : Tx = x\},$$

is the set of fixed points of mapping T or we can say that a point whose position is not changed by transformation is called a fixed point.
The following definitions will be in the next chapters.

Definition 2: Suppose a mapping $T : \mathbb{R}^n \to \mathbb{R}^n$. The mapping T is called a nonexpansive mapping if it satisfies the following

$$\|Tx - Ty\| \leq \|x - y\|, \ \ \forall x, y \in \mathbb{R}^n. \tag{2.2.1}$$

The following fixed point iteration has been studied by many authors for approximating either fixed points of nonlinear mappings (when theses mapping already have fixed points) or solutions of nonlinear operator equations:

The Mann iteration process

In 1953, Mann invented iterative method, see [15], and used to obtain convergence to a fixed point for many functions. Let a mapping $T : \mathbb{R}^n \to \mathbb{R}^n$, then the Mann iteration process is defined as follows.

$$\begin{cases} x^1 \in \mathbb{R}^n, \text{chosen arbitrarily} \\ x^{k+1} = (1-\alpha_k)x^k + \alpha_k T x^k = M(x^k, \alpha_k, T), \ \ k \in \mathbb{N} \end{cases}$$

where $\{\alpha_k\}$ is real sequence in $(0, 1)$.

In fact, we would like to point out that the modified proximal point algorithm involving single objective optimization and fixed point iteration in the framework of Hilbert spaces and Banach spaces have been intensively studied by many authors, for instance, (see [16, 17]) and the references therein.

2.3 Proximal point algorithm

The proximal point method is a conceptually simple algorithm for minimizing a function $h : \mathbb{R}^n \to \mathbb{R}$ on $\mathbb{R}^n$. At first glance, each proximal subproblem seems no easier than minimizing $h : \mathbb{R}^n \to \mathbb{R}$ in the first place. On the contrary, the addition of the quadratic penalty term often regularizes the proximal subproblems and makes them well conditioned. Despite the improved conditioning, each proximal subproblem still requires invoking an iterative solver. For this reason, the proximal point method has predominantly been thought of as a theoretical/conceptual algorithm, only guiding algorithm design and analysis rather than being implemented directly. The PPA of a closed convex function $h : \mathbb{R}^n \to \mathbb{R}$ is

$$\text{prox}^h_\lambda(x) = \arg\min_{y \in \mathbb{R}^n} \left\{ h(y) + \frac{\lambda}{2}\|y - x\|^2 \right\}. \tag{2.3.1}$$

The proximal point algorithm solves a single optimization problem by solving a sequence of optimization problems (2.3.1) which starts at a point $x^1 \in \mathbb{R}^n$ and generates recursively a sequence of points $\{x^k\}_{k=1}^{\infty}$, where

$$\text{prox}_{\lambda_k}^h(x^k) = x^{k+1} = \arg\min_{y\in\mathbb{R}^n}\left\{h(y) + \frac{\lambda_k}{2}\|y - x^k\|^2\right\}, \tag{2.3.2}$$

and $\{\lambda_k\}$ is a sequence of positive numbers.

The proximal point method has long been ingrained in the foundations of optimization. Recent progress in large scale computing has shown that the proximal point method is not only conceptual, but can guide methodology. Proximal algorithms can be effective and often lead to more easily interpretable numerical methods as compared to the direct methods.

In 2005, Bonnel et al. [6] generalized the famous Rockafellar results from scalar case to vector case, in which they investigate convex vector optimization problem in Hilbert spaces. After that, Ceng and Yao [30] developed both an absolute and a relative version of approximate proximal point algorithm. They considered the approximate proximal method via the subproblems of finding weakly efficient points for suitable regularizations of the original mapping.

Later, in 2015, Papa Quiroz et al. [31] proposed an inexact proximal point method of constrained multiobjective problems involving locally Lipschitz quasiconvex objective functions. They used proximal distances and assumed that the function is also bounded from below, lower semicontinuous for convergence analysis of the method. They proved that the sequence generated by the proposed method converges to a stationary point of the problem. After that, in 2018, João Carlos de O. Souza [34] studied the convergence of exact and inexact versions of the proximal point method with a generalized regularization function in Hadamard manifolds for solving scalar and vectorial optimization problems involving Lipschitz functions. In 2018, Bento et al. [21] considered the exact proximal point method of the constrained nonsmooth multiobjective optimization problem. They used nonscalarization approach for convergence analysis of the method, where the first order optimality condition of the problem is replaced by a necessary condition for weak Pareto points of a multiobjective problem. For more information on the related works in this direction, ones may see [7,8,21,27,32–35] and the references therein.

2.4 Auxiliary concepts

The domain of h, denoted by $\operatorname{dom} h$, is the subset of $\mathbb{R}^n$ on which h is finite valued. A function h is said to be proper when $\operatorname{dom}(h) \neq \emptyset$. We say that a function $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous function at a point $\hat{x} \in \mathbb{R}^n$ if for all sequence $\{x^k\} \subset \mathbb{R}^n$ such that $\lim_{k\to+\infty} x^k = \hat{x}$, we obtain that

$$h(\hat{x}) \leq \liminf_{k\to+\infty} h(x^k).$$

For a closed set $C \subset \mathbb{R}^n$, it is well known that the indicator function of C, $\delta_C : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous function.

We say that a scalar valued function $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is locally Lipschitz continuous at $x \in \mathbb{R}^n$ if there exist a neighborhood U of x and a positive real number L such that

$$|h(z) - h(y)| \leq L\|z - y\|, \quad \forall z, y \in U.$$

Next, we remind Fréchet and Mordukovich subdifferentials concepts.

Definition 2.4.1. Let h be a lower semicontinuous function. The Fréchet subdifferential of h at $x \in \mathbb{R}^n$ is defined by

$$\hat{\partial} h(x) = \begin{cases} \left\{ x^* \in \mathbb{R}^n : \liminf\limits_{y\to x, y\neq x} \dfrac{h(y) - h(x) - \langle x^*, y - x\rangle}{\|y - x\|} \geq 0 \right\}, & if x \in \operatorname{dom} h, \\ \emptyset, & if x \notin \operatorname{dom} h. \end{cases}$$

As noted by Bolte et al. [18], the Fréchet subdifferential is not completely satisfactory in optimization, since $\hat{\partial} h(x)$ might be empty-valued at points of particular interest. This justifies the choice of the following subdifferential:

Definition 2.4.2. Let h be a lower semicontinuous function. The Mordukovich-subdifferential of h at $x \in \mathbb{R}^n$ is defined by

$$\partial h(x) := \left\{ v \in \mathbb{R}^n : \exists (x^k, v^k) \in \operatorname{Graph}(\hat{\partial} h) \text{ with } (x^k, v^k) \to (x, v), h(x^k) \to h(x) \right\},$$

where $\operatorname{Graph}(\hat{\partial} h) := \left\{ (y, u) \in \mathbb{R}^n \times \mathbb{R}^n : u \in \hat{\partial} h(y) \right\}.$

We can see that $\hat{\partial} h(x) \subset \partial h(x)$. In the particular case when h is differentiable at $x \in \mathbb{R}^n$, then $\hat{\partial} h(x) = \partial h(x) = \{\nabla h(x)\}$. If h is convex, then both subdifferentials $\hat{\partial} h(x)$ and $\partial h(x)$ coincide with the usual subdifferential for each $x \in \operatorname{dom} h$.

Definition 2.4.3. Let $C \subset \mathbb{R}^n$ be a nonempty convex set. Then for each $x \in C$, the normal cone is defined by

$$N_C(x) := \{v \in \mathbb{R}^n : \langle v, y - x \rangle \leq 0, \ y \in C\}. \tag{2.4.1}$$

Remark 2.4.4. For nonempty closed and convex set C, $\partial \delta_C(x) = N_C(x)$.

A necessary (but not sufficient) condition for $x \in \operatorname{int}(\operatorname{dom}) h$ to be a minimizer of h is

$$0 \in \partial h(x). \tag{2.4.2}$$

A point $x \in \mathbb{R}^n$ satisfying the above inclusion is called limiting-critical or simply critical. Given a lower semicontinuous function $g : \mathbb{R}^n \to \mathbb{R}$ and $C \neq \emptyset$ a closed and convex set, for the case where $h = g + \delta_C$, we have h is a proper lower semicontinuous function with $\operatorname{dom} h = C$. Then the first order optimality condition takes the following form:

$$0 \in \partial g(x) + N_C(x), \tag{2.4.3}$$

see ([19], Theorem 8.5).
The following propositions are important in subsequent chapters.

Proposition 2.4.5. *Let $h_i : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous at $x \in \mathbb{R}^n$ for all $i \in \{1, ..., m\}$ and $g : \mathbb{R}^n \to \mathbb{R}$ defined by:*

$$g(x) = \max_{1 \leq i \leq m} h_i(x).$$

Then, g is locally Lipschitz continuous at x and

$$\partial g(x) \subset conv\{\partial h_i(x) : i \in I(x)\}, \tag{2.4.4}$$

where "conv" denotes the convex hull of a set and

$$I(x) := \{i \in I : h_i(x) = g(x)\}.$$

Proof. See ([20], Theorem 3.46(ii)). ☐

Proposition 2.4.6. *Let $h_1, h_2 : \mathbb{R}^n \to \mathbb{R}$ be functions such that h_1 is locally Lipschitz continuous at $\bar{x} \in \mathbb{R}^n$ while h_2 is proper lower semicontinuous with $h_2(\bar{x})$ finite. Then*

$$\partial(h_1 + h_2)(\bar{x}) \subset \partial h_1(\bar{x}) + \partial h_2(\bar{x}).$$

Proof. See([19], page 431). ☐

Proposition 2.4.7. *Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper locally Lipschitz function and $\{y^k\} \subset \operatorname{dom} h$ a bounded sequence. If $\{z^k\}$ is a sequence such that $z^k \in \partial h(y^k)$, then $\{z^k\}$ is bounded.*

Proof. The proof follows by combining ([19], Theorem 9.13 and Proposition 5.15) for $f = h, S = \partial h$ and $B = \{y^k\}$. ☐

Note: If in Proposition 2.4.7, we take $\{y^k = \hat{x}\}$ and $\{z^k \in \partial h(y^k = \hat{x})\}$, then $\{z^k\} \subset \mathbb{R}^n$ is bounded. So it has a convergent subsequence and consequently $\partial h(\hat{x})$ is relatively compact, that is, $\partial h(\hat{x})$ is compact.

Proposition 2.4.8. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a $\mathbb{R}^m_+$-quasi convex locally Lipschitz function on $\mathbb{R}^n$. If $g \in \partial h(x)$, such that $\langle h, \bar{x} - x \rangle > 0$, then $h(x) \leq h(\bar{x})$.*

The next result ensures that the set of minimizers of a function, under some assumptions, is nonempty.

Proposition 2.4.9. *[19] Suppose that $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is proper, lower semi-continuous and coercive, then the optimal value h is finite and the set $\{\arg\min h(x) : x \in \mathbb{R}^n\}$ is nonempty and compact.*

Next, we recall some concepts of Clarke directional derivative.
The Clarke directional derivative of a proper locally Lipschitz function $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ at $x \in \mathbb{R}^n$ in the direction of $d \in \mathbb{R}^n$ is denoted by $h^\diamond(x, d)$, and is defined as

$$h^\diamond(x, d) = \lim_{t \to 0} \sup_{y \to x} \frac{h(y + td) - h(y)}{t}.$$

Now, we recall some concepts involving locally Lipschitz functions and nonconvex constrained sets.

Let $C \subset \mathbb{R}^n$ be a nonempty and closed set. We denote the distance function $d : \mathbb{R}^n \to \mathbb{R}$ of a point $x \in \mathbb{R}^n$ to a set $C \subset \mathbb{R}^n$ as

$$d_C(x) := \inf\{\|x - c\| : c \in C\}. \tag{2.4.5}$$

We say that a point $x \in C$ is a Pareto-Clarke critical point of H in C if, for any $T_C(x)$, there exists $i = 1, ..., m$ such that

$$h_i^\diamond(x^*, v) \geq 0, \tag{2.4.6}$$

where $T_C(x) := \{v \in \mathbb{R}^n : d_C^\diamond(x, v) = 0\}$ denotes the set of all tangent vectors to C at x. As mentioned in ([22], page 11), a vector v belongs to $T_C(x)$ if and only if it satisfies the following property: for every sequence $\{x^k\}$ in C converges to x and every sequence t_k in $(0, \infty)$ converging to 0, there is a sequence v^k converging to v such that $x^k + t_k v^k$ belongs to C for all k. The normal cone is the one obtained from tangent cone $T_C(x)$ by polarity.

Therefore, the normal cone $N_C(x)$ to C at x is as follows:

$$\hat{N}_C(x) := \{\varsigma \in \mathbb{R}^n : \langle \varsigma, v \rangle \leq 0, \forall v \in T_C(x)\},$$

see [21]. If C is convex, $\hat{N}_C(x)$ coincides with the cone of normals in the sense of convex analysis; see ([22], Proposition 2.4.4).

Now, we remind Clarke subdifferential concept of scalar and vector functions. The Clarke subdifferential of scalar valued function $h : \mathbb{R}^n \to \mathbb{R}$ at x, denoted by $\partial^\diamond h(x)$, is defined as

$$\partial^\diamond h(x) := \{w \in \mathbb{R}^n : \langle w, d \rangle \leq h^\diamond(x, d), \;\; \forall d \in \mathbb{R}^n\},$$

see Clarke [23].

The Clarke subdifferential of $H : \mathbb{R}^n \to \mathbb{R}^m$ at $x \in \mathbb{R}^n$, denoted by $\partial^\diamond H(x)$, is defined as

$$\partial^\diamond H(x) := \{U \in \mathbb{R}^{m \times n} : U^T d \preceq H^\diamond(x, d), \;\; \forall d \in \mathbb{R}^n\},$$

where $H^\diamond(x, d) := \{h_1^\diamond(x, d), ..., h_m^\diamond(x, d)\}$.

Proposition 2.4.10. *([23], Proposition 1.4)* $h^\circ(x;v) = \max\{\xi \cdot v : \xi \in \partial^\circ h(x)\}$.

Remark 2.4.11. It is noted in [21] that, combining (2.4.6) with Proposition 2.4.10, we have the following alternative definition: a point $x \in \mathbb{R}^n$ is a Pareto-Clarke critical point of H in C if, for any $v \in T_C(x)$, there exist $i \in \{1, ..., m\}$ and $\xi \in \partial h_i(x)$ such that $\langle \xi, v \rangle \geq 0$. Thus, if x is not a Pareto-Clarke critical point of H in C, there exists $v \in T_C(x)$ such that $Uv \prec 0$, $\forall U \in \partial^\circ H(x)$.

The necessary condition for a point to be a Pareto-Clarke critical point of a vector-valued function can be found in Bento et al. ([21] Lemma 1), and is given below.

Proposition 2.4.12. *[21] Let $w \in \mathbb{R}^m_+ \setminus \{0\}$ and assume that C is closed and nonempty set. If $-U^T w \in \hat{N}_C(x)$ for some $U \in \partial^\circ H(x)$, then x is a Pareto Clarke-critical point of H.*

For the nonconvex case, a formula for the Clarke sudifferential of the distance function (2.4.5) defined in Burke, Ferris and Qian [24] is as follows:

Proposition 2.4.13. *[24] Let $C \subset \mathbb{R}^m$ be a nonempty and closed set. If $x \in C$, then*

$$\partial^\circ d_C(x) \subset \mathbb{B}[0,1] \cap \hat{N}_C(x), \tag{2.4.7}$$

where $\mathbb{B}[0,1]$ denotes the closed unit ball in $\mathbb{R}^m$.

Next propositions will be useful in next chapters.

Proposition 2.4.14. *([26], Theorem 3.2.1) Let $C \subset \mathbb{R}^n$ be a non empty set and $h : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function on $\mathbb{R}^n$ with constant L. If $\bar{x}$ is a minimizer for the constrained minimization problem,*

$$\min h(x), \quad x \in C, \tag{2.4.8}$$

and $\tau \geq L$, then $\bar{x}$ is also a minimizer for the unconstrained minimization problem

$$\min\{h(x) + \tau d_C(x)\}, \quad x \in \mathbb{R}^n. \tag{2.4.9}$$

If $\tau > L$ and C is a closed set, then the converse assertion is also true: Any minimizer $\bar{x}$ for the unconstrained problem (2.4.9) is also a minimizer for the constrained problem (2.4.8).

Proposition 2.4.15. *([27], Proposition 2.6.1) Let $\hat{x}$ be a Pareto-Clarke critical point of a locally Lipschitz function $H : \mathbb{R}^n \to \mathbb{R}^m$. If H is $\mathbb{R}^m_+$-convex, then $\hat{x}$ is a weak Pareto solution of the problem (4.0.1).*

Proposition 2.4.16. *([28], Proposition 5.3(ii)) For a function $h : \mathbb{R}^n \to \mathbb{R}$ locally Lipschitz at $\bar{x} \in \mathbb{R}^n$ with modulus $l > 0$, it holds that*

$$\|x^*\| \leq l, \quad \forall x^* \in \partial^\circ h(\bar{x}). \tag{2.4.10}$$

Proposition 2.4.17. *([28], Theorem 5.10) Let $h_1, h_2 : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz functions at $\bar{x} \in \mathbb{R}^n$. Then,*

$$\partial^\circ (h_1 + h_2)(\bar{x}) \subset \partial^\circ h_1(\bar{x}) + \partial^\circ h_2(\bar{x}). \tag{2.4.11}$$

Proposition 2.4.18. *[22] Let $h_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, i = 1, 2, ..., m$, be locally Lipschitz functions at x for all $i = \{1, ..., m\}$. Then the function $h(x) = \max\{h_i(x) | i \in \{1, ..., m\}\}$ is also locally Lipschitz at x and*

$$\partial^\circ h(x) \subset \operatorname{conv}\left\{\partial^\circ h_i(x) | \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1, \lambda_i[h_i(\bar{x}) - h(\bar{x})] = 0\right\}.$$

Proposition 2.4.19. *([29], Theorem 2.1) Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper quasiconvex locally Lipschitz function on $\mathbb{R}^n$. If $x^* \in \partial^\circ h(x)$, such that $\langle x^*, \hat{x} - x\rangle > 0$ then, $h(x) \leq h(\hat{x})$.*

Next, we will see that Lipschitz continuity of a function is the weaker condition than continuously differentiable condition.

Example 2.4.20. *Let $h : \mathbb{R} \to \mathbb{R}$ defined by*

$$h(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{if} \quad x \neq 0 \\ 0 & \text{if} \quad x = 0 \end{cases}$$

It is easy to see that the above example is Lipschitz continuous and differentiable everywhere and

$$h'(x) = \begin{cases} 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x} & \text{if} \quad x \neq 0 \\ 0 & \text{if} \quad x = 0 \end{cases}$$

exists but has an essential discontinuity at $x = 0$.

Next, we recall some basic concepts related to multi-objective optimization.

A sequence $\{x^k\} \subset \mathbb{R}^m$ is called a decreasing sequence if $x^p \prec x^k$ for $k < p$. A point $\bar{x}$ is said to be an infimum of x^k, if there is no x such that $x \preceq \bar{x}$ and $x \preceq x^k$ satisfying $\bar{x} \preceq x^k$ for all $k \in \mathbb{N}$. A vector function $H : \mathbb{R}^n \to \mathbb{R}^m$ is said to be positively semi-continuous if, for every $z \in \mathbb{R}^m_+$, the extended-valued scalar function $x \mapsto \langle H(x), z \rangle$ is lower semicontinuous.

The next definitions and results which can be found in [14] will be useful in the next chapters.

A vector valued function $H : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz if all components of H are locally Lipschitz.

We need the following in the sequel chapters:

Definition 2.4.21. [14] Let $H : \mathbb{R}^n \to \mathbb{R}^m$ be a vectorial function.

- H is called $\mathbb{R}^m_+$-convex iff for every $x, y \in \mathbb{R}^n$, the following holds:

$$H((1-t)x + ty) \preceq (1-t)H(x) + tH(y), \quad t \in [0, 1].$$

- H is called $\mathbb{R}^m_+$-quasi convex iff for every $x, y \in \mathbb{R}^n$, the following holds:

$$H((1-t)x + ty) \preceq \sup\{H(x), H(y)\}, \quad t \in [0, 1],$$

  where the supremum is considered coordinate by coordinate.

- H is called pseudo-convex iff H is differentiable and, for every $x, y \in \mathbb{R}^n$, the following holds:

$$H(y) \prec H(x) \implies JH(x)(y - x) \prec 0.$$

Note that H is convex (resp. quasi-convex) iff H is componentwise convex (resp. quasi-convex), see Definition 6.2 and Corollary 6.6 of ([14], pages 29 and 31), respectively. It is noted in [14] that if H is convex, in particular, it is also quasi-convex (the reciprocal is clearly false). If H is componentwise pseudo-convex, then H is pseudo convex, although the reciprocal is false.

We need the following propositions in our proofs.

Proposition 2.4.22. *[25] Assume that $H : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable. Then H is convex function if, only if, for every $x, y \in \mathbb{R}^n$,*

$$JH(x)(y - x) \preceq H(y) - H(x). \tag{2.4.12}$$

Proposition 2.4.23. *[25] Let $H : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable. Then H is quasi-convex function if, only if, for every $x, y \in \mathbb{R}^n$,*

$$H(y) \prec H(x) \implies JH(x)(y - x) \preceq 0. \tag{2.4.13}$$

Remark 2.4.24. If H is differentiable, then from the characterization (2.4.12), it follows that convexity is a sufficient condition for pseudo-convexity. On the other hand, from the characterization (2.4.13), we obtain that pseudo-convex functions are quasi-convex. Note that the reciprocal, in both the cases, is false.

The next proposition shows that under pseudo-convexity, criticality is equivalent to weak optimality.

Proposition 2.4.25. *[25] Let $H : \mathbb{R}^n \to \mathbb{R}^m$ be a pseudo-convex function and $x \in \mathbb{R}^n$. Then x is a weak Pareto optimal point of H if, only if,*

$$Im(JH(x)) \cap (-\mathbb{R}^m_{++}) = \emptyset.$$

Definition 2.4.26. [14] A subset A of $\mathbb{R}^m$ is said to be $\mathbb{R}^m_+$-complete, if any decreasing sequence of A is bounded by an element of A, i.e., whenever $\{x^k\} \subset A$ is a decreasing sequence, then there exists $x \in A$ such that $x \preceq x^k$ for all $k \geq 0$.

Proposition 2.4.27. *([14], Lemma 3.5) If $A \subset \mathbb{R}^m$ is closed, has a lower bound (i.e., $\exists$ some $a \in A$ such that for all $x \in A$, $a \preceq x$), then A is $\mathbb{R}^m_+$-complete.*

Proposition 2.4.28. *([14], Theorem 3.3) Consider the multi-objective problem (4.0.1). Then $\arg\min\{H(x) | x \in C\}$ in nonempty iff $H(C)$ has a $\mathbb{R}^m_+$-complete section.*

We end this section by recalling some concepts of proximal point mapping.

It was shown in [36] that the fixed point set $\mathrm{F}(\mathrm{prox}^h_\lambda)$ coincides with the set of minimizers of proper convex and lower semi-continuous function $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$.

Some other relevant characteristics of $\operatorname{prox}_\lambda^h$ of function $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ are incorporated in the following couple of lemmas:

Lemma 2.4.29. *[37] Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be proper convex and lower semi-continuous function. For any $\lambda > 0$, the proximal point mapping $\operatorname{prox}_\lambda^h$ is nonexpansive.*

Lemma 2.4.30. *[38] Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be proper convex and lower semi-continuous. Then for all $x, y \in \mathbb{R}^n$ and $\lambda > 0$, the following inequality holds*

$$\frac{1}{2\lambda}\|\operatorname{prox}_\lambda^h x - y\|^2 - \frac{1}{2\lambda}\|x-y\|^2 + \frac{1}{2\lambda}\|x - \operatorname{prox}_\lambda^h x\|^2 \le h(y) - h(\operatorname{prox}_\lambda^h x). \quad (2.4.14)$$

Lemma 2.4.31. *[37] Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper convex and lower semicontinuous function. Then the following identity holds:*

$$\operatorname{prox}_\lambda^h x = \operatorname{prox}_\mu^h\left(\frac{\lambda-\mu}{\lambda}\operatorname{prox}_\lambda^h x + \frac{\mu}{\lambda}x\right), \quad \forall x \in \mathbb{R}^n \text{ and } \lambda > \mu > 0.$$

We will end this section by recalling some auxiliary facts which will be useful for providing the convergence results of the proposed iterative algorithm in next Chapters.

Definition 2.4.32. A sequence $\{x^k\} \subset \mathbb{R}^n$ is said to be Fejér monotone to a nonempty set U iff, for all $x \in U$

$$\|x^{k+1} - x\| \le \|x^k - x\|, \quad k = 0, 1, \ldots$$

Definition 2.4.33. A sequence $\{x^k\} \subset \mathbb{R}^n$ is said to be quasi Fejér monotone to a nonempty set U iff, for all $x \in U$,

$$\|x^{k+1} - x\|^2 \le \|x^k - x\|^2 + \vartheta_k, \quad k = 0, 1, \ldots$$

where $\vartheta_k > 0$ and $\sum_{k=1}^{\infty} \vartheta_k < +\infty$.

Lemma 2.4.34. *[40] Let $U \subset \mathbb{R}^n$ be a nonempty set and $\{x^k\} \subset \mathbb{R}^n$ be a Fejér monotone sequence to U. Then, $\{x^k\}$ is bounded. Moreover, if an cluster point $\overline{x}$ of $\{x^k\}$ belongs to U, the whole sequence $\{x^k\}$ converges to $\overline{x}$ as k goes to $+\infty$.*

Lemma 2.4.35. *[41] Let $U \subset \mathbb{R}^n$ be a nonempty set and $\{x^k\} \subset \mathbb{R}^n$ be a quasi Fejér monotone sequence to U. Then, $\{x^k\}$ is bounded. Moreover, if an cluster point $\overline{x}$ of $\{x^k\}$ belongs to U, the whole sequence $\{x^k\}$ converges to $\overline{x}$ as k goes to $+\infty$.*

Lemma 2.4.36. *[39] Let $x, y \in \mathbb{R}^n$. Let $\alpha \in \mathbb{R}$ and $\mathbb{R}$ denote the set of real numbers. Then*

$$\|\alpha x + (1-\alpha)y\|^2 + \alpha(1-\alpha)\|x-y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2.$$

Proposition 2.4.37. *[6] Let $C \subset \mathbb{R}^n$ be convex set and $H : \mathbb{R}^n \to \mathbb{R}^m$ be a proper $\mathbb{R}^m_+$-convex map. It holds that*

$$\arg\min_w\{H(y)|y \in C\} = \bigcup_{z \in \mathbb{R}^m_+ \setminus \{0\}} \arg\min\{\langle H(y), z\rangle | y \in C\}, \tag{2.4.15}$$

Proposition 2.4.38. *([42], Corollary 1) Let $H : \mathbb{R}^n \to \mathbb{R}^m$ and $C \subset \mathbb{R}^n$ be a nonempty feasible set. It holds*

$$\arg\min_w\{H(x)|x \in \mathbb{R}^n\} \cap C \subseteq \arg\min_w\{H(x)|x \in C\}.$$

CHAPTER III

PROXIMAL POINT METHOD FOR PARETO OPTIMAL POINT OF SMOOTH QUASI-CONVEX FUNCTIONS IN MULTI-OBJECTIVE OPTIMIZATION

In this chapter, we continue to study the Algorithm proposed by Bento et al. [5]. In 2014, Bento et al. [5] presented a scalarized proximal point algorithm for multiobjective optimization by assuming an iterative process. By scalarization methods, one formulates a single objective optimization problem corresponding to a given multi-objective optimization problem. With respect to the convergence analysis, they showed that, for any sequence generated from this algorithm, each accumulation point is a Pareto critical point for the multi-objective function. The Algorithm is given as below:

3.0.1 Algorithm

INITIALIZATION. Choose $x^1 \in \mathbb{R}^n$.
STOPPING CRITERION. If x^k is a Pareto critical point STOP. Otherwise.
ITERATIVE STEP. Take a bounded sequence of positive real numbers $\{\lambda_k\}$, $e := (1, ..., 1) \in \mathbb{R}^m$ and choose $x^{k+1} \in \mathbb{R}^n$ such that

$$x^{k+1} \in \arg\min_{x\in\mathbb{R}^n} \Phi\left(H(x) + \delta_{\Omega_k}(x)e + \frac{\lambda_k}{2}\|x - x^k\|^2 e\right),$$

where

$$\Omega_k := \{x \in \mathbb{R}^n : H(x) \preceq H(x^k)\}, \tag{3.0.16}$$

and δ_{Ω_k} denote the indicator function of Ω_k, and $\Phi : \mathbb{R}^m \to \mathbb{R}$ is given by:

$$\Phi(y) := \max_{i\in I}\langle y, e_i\rangle,$$

where $\{e_i\} \subset \mathbb{R}^m$ is the canonical base of the space $\mathbb{R}^m$. This nonlinear scalarization function can be rewritten as follows:

$$\Phi(y) = \inf\{t \in \mathbb{R} : te \in y + \mathbb{R}^m_+\}, \; e := (1, ..., 1) \in \mathbb{R}^m.$$

The function Φ satisfies the following properties:

$$\begin{gathered}\Phi(x + \alpha e) = \Phi(x) + \alpha, \;\; \Phi(tx) = t\Phi(x), \;\; x \in \mathbb{R}^m, \; \alpha \in \mathbb{R}, \; t \geq 0. \\ x \preceq y \implies \Phi(x) \leq \Phi(y), \;\; x, y \in \mathbb{R}^m.\end{gathered} \tag{3.0.17}$$

In this chapter, we use the weaker assumptions of differentiable and locally Lipschitz properties on the above algorithm of the considered objective function instead of continuously differentiable assumptions. We show that under these assumptions, the method is still well defined and that the accumulation points of any generated sequence, if any, are Pareto critical point for the multi-objective function. Full convergence of the sequence generated by the Algorithm 3.0.1 is also considered.

3.1 Main results

In order to provide the convergence of the Algorithm defined by iterative step 3.0.1, we consider the following assumptions.

Assmuption 3.1.1. *There exists $i_0 \in I$ such that $h_{i_0} : \mathbb{R}^n \to \mathbb{R}$ is bounded below.*

Assmuption 3.1.2. *For all $i \in I$, h_i is locally Lipschitz.*

Assmuption 3.1.3. *H is differentiable.*

Assmuption 3.1.4. $\liminf_{k \to +\infty} \; \lambda_k > 0$.

Proposition 3.1.5. *The Algorithm 3.0.1 is well-defined.*

Proof. Define $\phi_k : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$\phi_k(x) := \Phi\left(H(x) + \delta_{\Omega_k}(x)e + \frac{\lambda_k}{2}\|x - x^k\|^2 e\right), \quad \textit{for each} \quad x \in \mathbb{R}^n. \tag{3.1.1}$$

From Assumption 3.1.1, it follows that ϕ_k is bounded from below. Hence, we have that ϕ_k is coercive. Now, as H is a continuous function, then Ω_k is a closed set and, in particular, δ_{Ω_k} is lower semicontinuous. As ϕ_k is lower semicontinuous and coercive, then using Proposition 2.4.9, we obtain that there exists $x^{k+1} \in \mathbb{R}^n$, which is a global minimizer of ϕ_k. □

Remark 3.1.6. The main improvement of the presented work in [5] is that we consider the differentiability and locally Lipschitz properties of the objective function H instead of a continuously differentiable property (see Assumptions (3.1.2) and (3.1.3)).

Lemma 3.1.7. *Assume that Assumptions 3.1.1 and 3.1.4 hold. Then the sequence $\{x^k\}$ generated by Algorithm 3.0.1 is bounded.*

Proof. For each $k \in \mathbb{N}$, define $\phi_k : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$\phi_k(x) := \Phi\left(H(x) + \delta_{\Omega_k}(x)e + \frac{\lambda_k}{2}\|x - x^k\|^2 e\right), \quad \textit{for each} \quad x \in \mathbb{R}^n. \quad (3.1.2)$$

It follows that $x^{k+1} \in \arg\min_{x\in\mathbb{R}^n} \phi_k(x)$. Then, from the definition of ϕ_k, we get

$$\Phi(H(x^{k+1}))+\delta_{\Omega_k}(x^{k+1})+\frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 \leq \Phi(H(x^k))+\delta_{\Omega_k}(x^k)+\frac{\lambda_k}{2}\|x^k - x^k\|^2.$$

Thus, by the definition of $\delta_{\Omega_k}(.)$, it follows that

$$\frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 \leq \Phi\big(H(x^k)\big) - \Phi\big(H(x^{k+1})\big), \quad k \in \mathbb{N}. \quad (3.1.3)$$

Hence, by $x^k \neq x^{k+1}$, $k \in \mathbb{N}$, we deduce $\frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 > 0$ and so

$$\Phi\big(H(x^{k+1})\big) < \Phi\big(H(x^k)\big),$$

thus by Assumption 3.1.1, we can assert that $\{\Phi(H(x^k))\}$ is a convergent sequence.

Also, by taking the sum of inequality (3.1.3), we obtain

$$\sum_{k=0}^{l} \frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 \leq \sum_{k=0}^{l} \Big(\Phi\big(H(x^k)\big) - \Phi\big(H(x^{k+1})\big) \Big),$$
$$= \Phi\big(H(x^0)\big) - \Phi\big(H(x^{l+1})\big).$$

This implies that the series $\sum \frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2$ is convergent. Using this one, in view of Assumption 3.1.4, we have $\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty$ and

$$\|(x^n - x^0)\| \leq \|\sum_{k=0}^{n-1}(x^{k+1} - x^k)\| \leq \sum_{k=0}^{n-1} \|x^{k+1} - x^k\| \leq \sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty.$$

Subsequently, it follows that $\{x^k\}$ is a bounded sequence. This completes the proof. □

Lemma 3.1.8. *Assume that the Assumptions 3.1.1, 3.1.2, 3.1.3 and 3.1.4 are true. If $\bar{x}$ is a cluster point of $\{x^k\}$ then $\bar{x}$ is a Pareto critical point, provided that Ω_k is a convex set for each $k \in \mathbb{N}$.*

Proof. Observe that, by Assumption 3.1.2, we have Ω_k is a closed subset of $\mathbb{R}^n$, for each $k \in \mathbb{N}$. Next, since $\Omega_{k+1} \subset \Omega_k$, for each $k \in \mathbb{N}$ and $\bar{x}$ is a cluster point of $\{x^k\}$, then there is a subsequence $\{x^{k_j}\}$ converging to $\bar{x}$ and

$$\Omega_0 \supset \Omega_1 \supset ...\Omega_k.$$

thus for $i \in \mathbb{N}$, there exist $j_0 \in \mathbb{N}$ such that $k_{j_0} \geq i$ which implies that $\forall n \geq n_0$, $k_n \geq k_{n_0} \geq i$. So, we obtain

$$\Omega_{k_n} \subset \Omega_{k_{n_0}} \subset \Omega_i.$$

and $x^{k_j} \in \Omega_{k_j}$ for all $i < n$. Since Ω_i is closed and subsequence $x^{k_j} \in \Omega_i$ thus $\bar{x} \in \Omega_i$.

$$\bar{x} \in \cap_{k=0}^{+\infty} \Omega_k =: \Omega. \tag{3.1.4}$$

This shows that Ω is a nonempty, closed and convex subset of $\mathbb{R}^n$.

Now, we will conclude the lemma by contradiction. That is we will assume that $\bar{x}$ is not a Pareto critical point. It would follow that there exists $v \in \mathbb{R}^n$ such that

$$JH(\bar{x})v \prec 0. \tag{3.1.5}$$

Since H is differentiable, then $D_v H(\bar{x}) = JH(\bar{x})v$. Subsequently, from the definition of directional derivative, we have that there is $\delta > 0$ such that

$$H(\bar{x} + sv) \prec H(\bar{x}), \quad \text{for each} \quad s \in (0, \delta).$$

This implies that $\bar{x} + sv \in \Omega_k$, for $k \in \mathbb{N}$ and $s \in (0, \delta)$.

On the other hand, since $x^{k+1} \in \arg\min_{x\in\mathbb{R}^n} \phi_k(x)$, then by (3.0.17) and (2.4.2), we obtain

$$0 \in \partial\Big(\Phi(H(\cdot)) + \delta_{\Omega_k}(\cdot) + \frac{\lambda_k}{2}\|\cdot - x^k\|^2\Big)(x^{k+1}), \quad \text{for each} \quad k \in \mathbb{N}.$$

Thus, from (2.4.3), it follows that

$$0 \in \partial\Big(\Phi(H(\cdot)) + \frac{\lambda_k}{2}\|\cdot - x^k\|^2\Big)(x^{k+1}) + N_{\Omega_k}(x^{k+1}), \quad \text{for each} \quad k \in \mathbb{N}.$$

Note that, from Assumption 3.1.2 together with Proposition 2.4.5, we know that the function $\Phi \circ H$ is locally Lipschitz. Thus, by applying Proposition 2.4.6 with $h_1(\cdot) = \Phi(H(\cdot))$ and $h_2(\cdot) = \frac{\lambda_k}{2}\|\cdot - x^k\|^2$, the last inlusion becomes:

$$0 \in \partial(\Phi \circ H)(x^{k+1}) + \lambda_k(x^{k+1} - x^k) + N_{\Omega_k}(x^{k+1}), \quad \text{for each} \quad k \in \mathbb{N}.$$

Subsequently, there exist sequences $\{w^k\}$, $\{v^k\}$, with $w^{k+1} \in \partial(\Phi \circ H)(x^{k+1})$ and $v^{k+1} \in N_{\Omega_k}(x^{k+1})$ such that

$$0 = w^{k+1} + \lambda_k(x^{k+1} - x^k) + v^{k+1}, \quad \text{for each} \quad k \in \mathbb{N}. \tag{3.1.6}$$

As $\Phi \circ H$ is locally Lipschitz, then by applying Proposition 2.4.7 with $y^k = x^k$, $h = \Phi \circ H$ and $z^k = w^k$, for each $k \in \mathbb{N}$, obey the fact that $\{x^k\}$ is a bounded sequence we obtain that the sequence $\{w^k\}$ is bounded. Subsequently, by (3.1.6), $\{v^k\}$ is also a bounded sequence.
Next, let $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$, which converges to $\bar{x}$. Moreover, let $\bar{w}$ (resp. $\bar{v}$) be a cluster point of $\{w^k\}$ (resp. of $\{v^k\}$). We can assume without loss of generality that the subsequences $\{w^{k_j}\}$ of $\{w^k\}$ and $\{v^{k_j}\}$ of $\{v^k\}$ converge, respectively to $\bar{w}$ and $\bar{v}$ as j goes to infinity. By replacing k by k_j in (3.1.6), letting j goes to infinity and taking into account that $\lim_{j\to+\infty} \lambda_{k_j}(x^{k_{j+1}} - x^{k_j}) = 0$ we obtain:

$$\bar{w} = -\bar{v}. \tag{3.1.7}$$

Now, since $v^{k_j+1} \in N_{\Omega_{k_j}}(x^{k_j+1})$ and $\Omega \subset \Omega_{k_j}$, for each $j \in \mathbb{N}$, from Remark 2.4.4, we have

$$\langle v^{k_j}, x - x^{k_j}\rangle \leq 0, \quad \text{for each} \quad x \in \Omega. \tag{3.1.8}$$

Since v^{k_j} converges to $\bar{v}$, thus by letting j goes to infinity in the last inequality, one can conclude that

$$\langle \bar{v}, x - \bar{x} \rangle \leq 0, \quad \text{for each} \quad x \in \Omega. \tag{3.1.9}$$

So, in view of equality (3.1.7), this leads to

$$\langle \bar{w}, x - \bar{x} \rangle \geq 0, \quad \text{for each} \quad x \in \Omega. \tag{3.1.10}$$

On the other hand, since $\{\Phi(H(x^{k_j}))\}$ converges to $\Phi(H(\bar{x}))$ as j goes to infinity, it follows from the definition of $\partial(\Phi \circ H)$ that $\bar{w} \in \partial(\Phi \circ H)$. Moreover, since H is differentiable, we have $\partial h_i(x) = \{\nabla h_i(x)\}$. Using the characterization (2.4.4) with $g = \Phi \circ H$ and $x = \bar{x}$, there exists a vector $\alpha = (\alpha_1, ..., \alpha_m) \in \mathbb{R}^m_+$, with $\sum_{i=1}^m \alpha_i = 1$ such that

$$\bar{w} = \sum_{i \in I(\bar{x})} \alpha_i \nabla h_i(\bar{x}) = JH(\bar{x})^t \alpha. \tag{3.1.11}$$

For given $s > 0$, combining (3.1.10) and (3.1.11) one has

$$\langle JH(\bar{x})^t \alpha, \bar{x} + sv - \bar{x} \rangle = s\langle JH(\bar{x})^t \alpha, v \rangle = s\langle \alpha, JH(\bar{x})v \rangle \geq 0. \tag{3.1.12}$$

In contrast, by combining the definition of α with (3.1.5), we infer

$$\langle \alpha, JH(\bar{x})v \rangle < 0,$$

which contradicts (3.1.12). This completes the proof. □

Next, we will show the full convergence of the considered algorithm provided that the objective mapping is quasiconvex.

Theorem 3.1.9. *Assume that Assumptions 3.1.1, 3.1.2, 3.1.3 and 3.1.4 hold. If H is a $\mathbb{R}^m_+$-quasi convex function, then the sequence $\{x^k\}$ converges to a Pareto critical point of H.*

Proof. Remind that from (3.1.4) in the proof of Lemma 3.1.8, we know that $\Omega := \cap_{k=0}^{+\infty} \Omega_k$ is a nonempty closed and convex set. Let $x^* \in \Omega$ be given, consider

$$\left\|x^k - x^*\right\|^2 = \left\|x^{k+1} - x^*\right\|^2 + \left\|x^k - x^{k+1}\right\|^2 + 2\langle x^k - x^{k+1}, x^{k+1} - x^* \rangle, \tag{3.1.13}$$

for each $k \in \mathbb{N}$. By following the line of the proof of Lemma 3.1.8, we know that there exist sequences $\{w^k\}$ and $\{v^k\}$ with $w^{k+1} \in \partial(\Phi \circ H)(x^{k+1})$ and $v^{k+1} \in N_{\Omega_k}(x^{k+1})$ such that satisfying (3.1.6), this implies

$$x^k - x^{k+1} = \frac{1}{\lambda_k}\left(w^{k+1} + v^{k+1}\right), \text{ for each } \quad k \in \mathbb{N}.$$

Using this equality together with (4.2.12), we have

$$\left\|x^k - x^*\right\|^2 = \left\|x^{k+1} - x^*\right\|^2 + \left\|x^k - x^{k+1}\right\|^2 - \frac{2}{\lambda_k}\langle w^{k+1} + v^{k+1}, x^* - x^{k+1}\rangle, \tag{3.1.14}$$

for each $k \in \mathbb{N}$.

Next, taking into account that $w^{k+1} \in \partial(\Phi \circ H)(x^{k+1})$ and using the Proposition 2.4.5 with $g = \Phi \circ H$ and $x = \bar{x}$, we see that there exists a vector $\alpha^{k+1} = \left(\alpha_1^{k+1}, ..., \alpha_m^{k+1}\right) \in \mathbb{R}_+^m$, with $\sum_{i=1}^m \alpha_i^{k+1} = 1$ such that

$$w^{k+1} = \sum_{i \in I(x^{k+1})} \alpha_i^{k+1} \nabla h_i(x^{k+1}), \text{ for each } \quad k \in \mathbb{N}. \tag{3.1.15}$$

On the other hand, since $x^* \in \Omega$, we see that $H(x^*) \preceq H(x^{k+1})$. Subsequently, it follows from the quasi-convexity of H and Proposition 2.4.8, that for each $k = 1, 2, ...,$

$$\langle \nabla h_i(x^{k+1}), x^* - x^{k+1}\rangle \leq 0, \text{ for each } i \in \{1, ..., m\}, \text{ and } k \in \mathbb{N}.$$

Thus, by using (5.1.3), we get

$$\langle w^{k+1}, x^* - x^{k+1}\rangle \leq 0, \quad \text{for each} \quad k \in \mathbb{N}. \tag{3.1.16}$$

From another stand point, since $v^{k+1} \in N_{\Omega_k}(x^{k+1})$ and Ω_k is a convex set (because H is quasi convex), for each $k \in \mathbb{N}$, it follows that, we have

$$\langle v^{k+1}, x^* - x^{k+1}\rangle \leq 0, \quad \text{for each} \quad k \in \mathbb{N}. \tag{3.1.17}$$

As $\|x^k - x^{k+1}\|^2 \geq 0$, for each $k \in \mathbb{N}$, the inequality (5.1.5) becomes

$$\left\|x^k - x^*\right\|^2 \geq \left\|x^{k+1} - x^*\right\|^2 - \frac{2}{\lambda_k}\langle w^{k+1}, x^* - x^{k+1}\rangle - \frac{2}{\lambda_k}\langle v^{k+1}, x^* - x^{k+1}\rangle, \tag{3.1.18}$$

for each $k \in \mathbb{N}$. Combining last inequality with (5.1.4) and (3.1.17), we conclude that

$$\|x^{k+1} - x^*\| \leq \|x^k - x^*\|, \quad \text{for each} \quad k \in \mathbb{N}. \tag{3.1.19}$$

This means $\{x^k\}$ is a Fejér monotone to Ω. Thus, in view of Lemma 2.4.34, we conclude that the sequence $\{x^k\}$ converges to $\bar{x}$ as k goes to $+\infty$. Finally, by Lemma 3.1.8, we have $\bar{x}$ is a Pareto critical point of problem (2.2). This completes the proof. □

Corollary 3.1.10. *Under Assumption 3.1.3, if H is pseudo-convex or convex, then the sequence $\{x^k\}$ converges to a weak Pareto optimal point of H.*

Proof. If H is pseudo-convex or convex, in particular, H is $\mathbb{R}^m_+$-quasi convex (see Remark 2.4.24) and the corollary is a consequence of the previous theorem, then the sequence $\{x^k\}$ converges to $\bar{x}$ as k goes to $+\infty$ and $\bar{x}$ is a Pareto critical point. By Proposition 2.4.25, under pseudo-convexity criticality is equivalent to weak optimality which implies that $\bar{x}$ is a weak Pareto optimal point. □

3.2 Conclusion

This chapter presented the proximal point method for solving multiobjective optimization problem under the differentiability, locally Lipschitz and quasi-convex conditions of the cost function. The control conditions to guarantee that the accumulation points of any generated sequence, are Pareto critical points are provided.

CHAPTER IV

PROXIMAL POINT ALGORITHM FOR ϵ-QUASI-WEAK PARETO SOLUTION OF NONSMOOTH LOCALLY LIPSCHITZ MULTIOBJECTIVE OPTIMIZATION

In this chapter, our interest is to consider proximal point method for solving nonsmooth multiobjective optimization problem (4.0.1). Using the same technique as in Bento et al. [21], we propose proximal point algorithm for finding the solution concepts as ϵ-quasi weak Pareto optimal points for constrained nonsmooth multiobjective optimization problem. In terms of Clarke subdifferential, we introduce Fritz-John optimality condition of an ϵ-quasi weak Pareto solution, which we use for convergence analysis of our method. We also show that our proposed algorithm is well defined and the sequence achieved by the proposed algorithm converges to a Pareto-Clarke critical point. For a convex objective function H, we obtain the convergence to a weak Pareto solution of the problem. Throughout this chapter, We consider the following (constrained) multiobjective minimization problem

$$\min_{x \in C} H(x), \tag{4.0.1}$$

where $C \subset \mathbb{R}^n$ a nonempty and closed set and $H : C \to \mathbb{R}^m$ is locally Lipschitz function.

4.1 Necessary optimality condition

In this section, we consider multiobjective optimization problem (4.0.1) of finding the ϵ-quasi weak Pareto point of a vector valued function H subject to the following constrained set

$$C := \{x \in D | g_j(x) \leq 0, \ j = 1, ..., p\},$$

where $D \subset \mathbb{R}^n$ is a nonempty and closed set and $g_j : \mathbb{R}^n \to \mathbb{R}$ is a locally Lipschitz function. We provide necessary conditions for a point $x^* \in C$ to be an ϵ-quasi weak Pareto solution associated to the problem (4.0.1).

Proposition 4.1.1. *Let $x^* \in \arg\min_{\epsilon q-w}\{H(x)|x \in C\}$. Then, there exist $t_i \geq 0$ and $\mu_j \geq 0$ for $i \in \{1, ..., m\}$ and $j \in \{1, ..., p\}$ with $\sum_{i=1}^m t_i + \sum_{j=1}^p \mu_j = 1$ and $\tau > 0$ such that*

$$0 \in \sum_{i=1}^{m} t_i \partial^\diamond h_i(x^*) + \sum_{j=1}^{p} \mu_j \partial^\diamond g_j(x^*) + \sum_{i=1}^{m} t_i \epsilon_i \mathbb{B}_{x^*} + \tau \partial^\diamond d_D(x^*),$$

where $h_i : \mathbb{R}^n \to \mathbb{R}$, $\epsilon_i \in \mathbb{R}^m_+$ for $i \in \{1, ..., m\}$ and $\mathbb{B}_{x^*}$ denotes the closed unit ball of x^*.

Proof. For each $x \in C$, put $\Psi(x) = \max_{\substack{i\in\{1,...,m\} \\ j\in\{1,...,p\}}}\{h_i(x) - h_i(x^*) + \epsilon_i\|x - x^*\|, g_j(x)\}$. Observe that $\Psi(x^*) = \max\{h_i(x^*) - h_i(x^*) + \epsilon_i\|x^* - x^*\|, g_j(x^*)\}$. Since $g_j(x^*) \leq 0$, then we have that $\Psi(x^*) = 0$.
Next, since x^* is an ϵ-quasi weak Pareto optimal point, then there is no $x \in C$ such that

$$h_i(x) + \epsilon_i\|x - x^*\| < h_i(x^*), \quad \forall i \in \{1, ..., m\}. \tag{4.1.1}$$

It can be easily verified that $0 \leq \Psi(x)$, which infers that for all $x \in C$, we have

$$\Psi(x^*) = \inf_{x \in C} \Psi(x).$$

It follows that x^* is also a minimizer to the constrained optimization problem

$$\min_{x \in C} \Psi(x).$$

Proposition 2.4.18 and locally Lipschitz properties of functions h_i and g_j imply that the function Ψ is also locally Lipschitz around x^*. Let L be a locally Lipschitz constant of Ψ at x^* and $\tau \geq L$, then applying the Proposition 2.4.14 to the last problem, we obtain

$$0 \in \partial^\diamond(\Psi(x^*) + \tau d_D(x^*)). \tag{4.1.2}$$

Also, the sum rule (2.4.11) implies that

$$0 \in \partial^\diamond \Psi(x^*) + \tau \partial^\diamond d_D(x^*). \tag{4.1.3}$$

Now, by Proposition 2.4.18 and invoking the sum rule (2.4.11) applied to the Ψ, there exist non-negative real numbers $t_i \geq 0$ and $\mu_j \geq 0$ such that $\sum_{i=1}^m t_i + \sum_{j=1}^p \mu_j = 1$ and

$$\partial^\circ \Psi(x^*) \subset \left\{ \sum_{i=1}^m t_i \partial^\circ h_i(x^*) + \sum_{i=1}^m t_i \epsilon_i B_{x^*} + \sum_{j=1}^p \mu_j \partial^\circ g_j(x^*) \right\}. \tag{4.1.4}$$

and the desired result follows by combining (4.1.3) with (4.1.4). □

4.2 Proximal point algorithm

In this section, we consider $C \subset \mathbb{R}^n$ a nonempty and closed set and $H : C \to \mathbb{R}^m$ is locally Lipschitz function.

Next, we consider the inexact proximal point algorithm for obtaining a Pareto-Clarke critical point of H in C. Take a bounded sequence of positive real numbers $\{\lambda_k\}$ and a sequence $\{e^k\} \subset \mathbb{R}^m_{++}$ such that $\|e^k\| = 1$, for all $k \in \mathbb{N}$. The method generates the sequence $\{x^k\} \in C$ as follows:

4.2.1 Algorithm

INITIALIZATION: Choose an arbitrary initial point

$$x^1 \in C. \tag{4.2.1}$$

STOPPING CRITERION: Given x^k, if x^k is a Pareto-Clarke critical point, then stop. Otherwise go to the iterative step.

ITERATIVE STEP: Take the next iterate $x^{k+1} \in C$ as y such that there exists $\epsilon^k \in \mathbb{R}^m_+$ satisfying

$$y \in \arg\min_{\epsilon^k q - w}\{H(x) + \frac{\lambda_k}{2}\|x - x^k\|^2 e^k | x \in \Omega_k\}, \tag{4.2.2}$$

$$\epsilon^k \preceq \sigma_k \frac{\lambda_k}{2}\|y - x^k\| e^k, \tag{4.2.3}$$

where $\Omega_k := \{x \in C | H(x) \preceq H(x^k)\}$ and $\{\sigma_k\} \subset [0, 1)$.

From now on, we will assume that $0 \prec H$.

4.2.2 Existence of iterates

Proposition 4.2.1. *Let $H : C \to \mathbb{R}^m$ be a continuous function. Then, the sequence $\{x^k\}$ generated by Algorithm 4.2.1, is well defined.*

Proof. We proceed by induction: It holds for $k = 1$ due to (4.2.1). Assume that x^k exists and define

$$H_k(x) := H(x) + \frac{\lambda_k}{2}\|x - x^k\|^2 e^k.$$

Since $x^k \in \Omega_k$, we have $H_k(\Omega_k) \neq \emptyset$. By assumption on H that is $0 \prec H$, we get, $0 \prec H_k(x)$. Now, let $\{y^p\} \subset H_k(\Omega_k)$ such that $y^p \to y$. Since $y^p \in H_k(\Omega_k)$ there exists $z^p \in \Omega_k$ satisfying $y^p = H_k(z^p)$, for any p. We claim that $\{z^p\}$ is bounded, if not, then there is $\{p_j\} \subset \{p\}$ such that $z^{p_j} \to \infty$ as $j \to \infty$, then coercivity of H_k infers that $\|H_k(z^{p_j})\| \to +\infty$ as $j \to \infty$. On the other hand, $\|H_k(z^p)\| \to \|y\|$ because $y^p = H_k(z^p)$ and $y^p \to y$, which is a contradiction. Hence, we proved that $\{z^p\}$ is a bounded sequence. Subsequently, there are $\{z^{p_j}\} \subset \{z^p\}$ and $z \in \mathbb{R}^n$ such that $z^{p_j} \to z$ as $j \to \infty$. Moreover, by the continuity of H, we know that Ω_k is a closed set. Hence, $z \in \Omega_k$. Applying continuity of H_k and using uniqueness of limit, we can assert that $y \in H_k(\Omega_k)$. This proves $H_k(\Omega_k)$ is closed.
Subsequently, by Proposition 2.4.27 and property of $\mathbb{R}^m_+$ that all decreasing sequences having lower bound converges to its infimum, we know that $H_k(\Omega_k)$ is $\mathbb{R}^m_+$-complete. Thus, Proposition 2.4.28 infers that

$$\arg\min_w\{H_k(x)|x \in \Omega_k\}$$

is not empty. Therefore, by Remark 2.1.5, it follows that $\arg\min_{\epsilon^k q - w}\{H_k(x)|x \in \Omega_k\} \neq \emptyset$. □

Remark 4.2.2. Note that if Algorithm 4.2.1 terminates after finite number of iterations, then it terminates at a Pareto-Clarke critical point.

4.2.3 Convergence Analysis

In this section, first we present some results which play an important role in our subsequent considerations. Then, we show that the sequence generated by our algorithm converges to a Pareto-Clarke critical point.

Proposition 4.2.3. *For all $k \in \mathbb{N}$, there exists $A_k \in \mathbb{R}^{m\times n}$, $\alpha^k, \beta^k \in \mathbb{R}^m_+$, $\tau_k > 0$ and $w^k \in \mathbb{R}^n$ such that*

$$A_k^T(\alpha^k + \beta^k) + \lambda_{k-1}\langle e^{k-1}, \alpha^k\rangle(x^k - x^{k-1}) + \langle \epsilon^{k-1}, \alpha^k\rangle v^k + \tau_k w^k = 0, \quad (4.2.4)$$

where $v^k \in \mathbb{B}_{x^k}$, $w^k \in \mathbb{B}[0,1] \cap \hat{N}_C(x^k)$ and $\sum_{i=1}^m(\alpha_i^k + \beta_i^k) = 1$, $\forall k \in \mathbb{N}$.

Proof. For every k, consider the functions

$$W_k(x) := H(x) - H(x^k), \text{ and } H_k(x) := H(x) + \frac{\lambda_k}{2}\|x - x^k\|^2 e^k.$$

As, H and $\|x - x^k\|^2$ are locally Lipschitz, the coordinate functions $(W_k)_i(.) := H(.) - H(x^k)$ and $(H_k)_i(.) := H(.) + \frac{\lambda_k}{2}\|. - x^k\|^2 e^k$, $i \in \{1, ..., m\}$, of $W_k(x)$ and $H_k(x)$, respectively, are also locally Lipschitz.

Since x^k is an ϵ-quasi weak Pareto solution for

$$\min H_{k-1}(x) \text{ such that } W_{k-1}(x) \preceq 0,$$

hence the desired result follows by applying Proposition 4.1.1, for each $k \in \mathbb{N}$ fixed with h_i and g_j by H_{k-1} and W_{k-1}, respectively, and taking into account that from Proposition 2.4.7, we have

$$\partial^\circ d_C(x^k) \subset \mathbb{B}[0,1] \cap \hat{N}_C(x^k), \quad \forall k \in \mathbb{N}.$$

In this case, $A_k^T = [u_1^k....u_m^k]$, where $u_i^k \in \partial^\circ h_i(x^k)$ with $i \in \{1, ..., m\}$, $\alpha^k = (\alpha_1^k, ..., \alpha_m^k)^T$ and $\beta^k = (\beta_1^k, ..., \beta_m^k)^T$. □

Proposition 4.2.4. *If there exists $k \in \mathbb{N}$ such that $x^{k+1} = x^k$, then x^k is a Pareto-Clarke critical point of H.*

Proof. Suppose that for any $k \in \mathbb{N}$, $x^{k+1} = x^k$ which implies that $\epsilon^k = 0$. Then by Proposition 4.2.3, we obtain

$$A_{k+1}^T(\alpha^{k+1} + \beta^{k+1}) + \tau_k w^{k+1} = 0, \quad (4.2.5)$$

which infers that

$$-A_{k+1}^T(\alpha^{k+1} + \beta^{k+1}) \in \hat{N}_C(x^{k+1}). \quad (4.2.6)$$

Since $\sum_{i=1}^m(\alpha_i^{k+1} + \beta_i^{k+1}) = 1$, we can say that $(\alpha^{k+1} + \beta^{k+1}) \in \mathbb{R}^m_+ \backslash \{0\}$. Moreover, $A_{k+1} \in \partial^\circ H(x^{k+1})$, then using Proposition 2.4.12, we obtain the desired result. □

Proposition 4.2.5. *Let $k_0 \in \mathbb{N}$ be such that $\alpha^{k_0} = 0$. Then x^{k_0} is a Pareto-Clarke critical point of H.*

Proof. If there exists $k_0 \in \mathbb{N}$ such that $\alpha_{k_0} = 0$ then, from (4.2.4), we have

$$A_{k_0}^T \beta^{k_0} + \tau_{k_0} w^{k_0} = 0, \tag{4.2.7}$$

where $\tau_{k_0} > 0$, $w^{k_0} \in \hat{N}_C(x^{k_0})$. Since $A_{k_0} \in \partial^\circ H(x^{k_0})$ and $\beta^{k_0} \in \mathbb{R}^m_+ \setminus \{0\}$, the desired result follows by using Proposition 2.4.12. $\square$

From now on, we will assume that the sequences $\{\lambda_k\}$, $\{\epsilon^k\}$ and $\{x^k\}$ are infinite sequences generated by Algorithm 4.2.1, then $\alpha^k \neq 0$ and $x^{k+1} \neq x^k$, in view of Proposition 4.2.4 and 4.2.5, respectively.
Next we prove that every cluster point of x^k, if any, is Parcto-Clarke critical point.

Theorem 4.2.6. *Assume that there exist scalars $a, b, c, d \in \mathbb{R}_{++}$ such that $a \leq \lambda_k \leq b$, $c \leq e_i^k \leq d$, $\sigma_k \leq d < 1$, for all $k \in \mathbb{N}$ and $i \in \{1, ..., m\}$. Then, every cluster point of $\{x^k\}$, if any, is a Pareto-Clarke critical point of H.*

Proof. Since

$$x^{k+1} \in \arg\min_{\epsilon^k q-w}\{H(x) + \frac{\lambda_k}{2}\|x - x^k\|^2 e^k | x \in \Omega_k\},$$

we have

$$\max_{1\leq i\leq m}\{h_i(x^k) - h_i(x^{k+1}) + \epsilon_i^k\|x^k - x^{k+1}\| - \frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 e_i^k\} \geq 0.$$

Hence for any k, there exists some index $i_0 := i_0(k) \in \{1, ..., m\}$, where the maximum in the last inequality is attained. Thus,

$$h_{i_0}(x^k) - h_{i_0}(x^{k+1}) + \epsilon_{i_0}^k\|x^k - x^{k+1}\| - \frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 e_{i_0}^k \geq 0,$$

which provides us

$$\frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 e_{i_0}^k - \epsilon_{i_0}^k\|x^k - x^{k+1}\| \leq h_{i_0}(x^k) - h_{i_0}(x^{k+1}).$$

By (4.2.3) and boundedness assumption of $\{\lambda_k\}$ and $\{e^k\}$, we obtain

$$h_{i_0}(x^k) - h_{i_0}(x^{k+1}) \geq \frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 e_{i_0}^k - \epsilon_{i_0}^k\|x^{k+1} - x^k\|$$

$$\geq \frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 e^k_{i_0} - \sigma_k \frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 e^k_{i_0}$$
$$\geq (1-\sigma_k)\frac{\lambda_k}{2}\|x^{k+1} - x^k\|^2 e^k_{i_0}.$$

Then, from the boundedness of $\{\lambda_k\}$, $\{\epsilon^k\}$ and $\{\sigma_k\}$, we obtain

$$(1-d)\frac{ac}{2}\|x^{k+1} - x^k\|^2 \leq h_{i_0}(x^k) - h_{i_0}(x^{k+1}). \tag{4.2.8}$$

Combining (4.2.2) with the definition of Ω_k, it follows that $\{H(x^k)\}$ is nonincreasing sequence and by assumption on H, i.e, $0 \prec H$, we have that $\{H(x^k)\}$ is a convergent sequence. Hence, by taking $k \to +\infty$ on (4.2.8), we get

$$\lim_{k\to+\infty}(x^{k+1} - x^k) = 0. \tag{4.2.9}$$

Take $\bar{x}$ as a cluster point of $\{x^k\}$, then there exists subsequence $\{x^{k_j}\}$ of $\{x^k\}$ converging to $\bar{x}$. Therefore, by applying Proposition 4.2.3 for the sequence $\{x^{k_j}\}$, we have that there exist sequences $A_{k_j+1} \in \partial^\circ H(x^{k_j+1})$, $\alpha^{k_j+1}, \beta^{k_j+1} \in \mathbb{R}^m_+$ and $v^{k_j+1} \in B_{x^{k_j+1}}$ such that

$$A^T_{k_j+1}(\alpha^{k_j+1}+\beta^{k_j+1})+\lambda_{k_j}\langle e^{k_j},\alpha^{k_j+1}\rangle(x^{k_j+1}-x^{k_j})+\langle\epsilon^{k_j},\alpha^{k_j+1}\rangle v^{k_j+1}+\tau_{k_j+1}w^{k_j+1} = 0, \tag{4.2.10}$$

where $\sum_{i=1}^m(\alpha_i^{k_j+1} + \beta_i^{k_j+1}) = 1$ and $w^{k_j+1} \in \hat{N}_C(x^{k_j+1})$.

From the convergence of $\{x^{k_j}\}$, we obtain that $\{x^{k_j}\}$ is bounded. By locally Lipschitz property of H, it follows by Proposition 2.4.7 that their subgradients are bounded. So from the above conditions the sequences A_{k_j}, v^{k_j}, α^{k_j}, β^{k_j}, w^{k_j} are bounded. Thus, equality (4.2.10) implies that τ_{k_j} is also bounded. Now, without loss of generality, we may assume that the sequences A_{k_j}, v^{k_j}, α^{k_j}, β^{k_j}, w^{k_j} and τ_{k_j} converge to $\bar{A}$, $\bar{v}$, $\bar{\alpha}$, $\bar{\beta}$, $\bar{w}$ and $\bar{\tau}$ respectively. Also, since $\lambda_{k_j}\langle e^{k_j}, \alpha^{k_j+1}\rangle$ is bounded, then by letting k_j goes to infinity in (4.2.10), we obtain

$$\bar{A}^T(\bar{\alpha} + \bar{\beta}) + \bar{\tau}\bar{w} = 0. \tag{4.2.11}$$

Since $\bar{w} \in \hat{N}_C(\bar{x})$, $(\bar{\alpha} + \bar{\beta}) \in \mathbb{R}^m_+ \setminus \{0\}$, $\bar{A} \in \partial^\circ H(\bar{x})$, it follows from (4.2.11) that

$$-\bar{A}^T(\bar{\alpha} + \bar{\beta}) \in \hat{N}_C(\bar{x}),$$

and this together with Proposition 2.4.12, enables us to say that $\bar{x}$ is a Pareto-Clarke critical point of H. This completes the proof. □

Next, we present a technical result that will be useful in the convergence analysis.

Lemma 4.2.7. *Let* $\{z_k\} \subset [0,1)$ *be a sequence such that* $\sum_{k=0}^{\infty} z_k < +\infty$. *Then*

$$\sum_{k=0}^{+\infty} \frac{z_k}{2 - z_k} < +\infty \quad and \quad \prod_{k=0}^{+\infty} \left(1 + \frac{z_k}{2 - z_k}\right) < +\infty.$$

Proof. The first statement follows from the following inequality

$$\frac{z_k}{2 - z_k} < z_k,$$

since $z_k \in [0,1)$ and $\sum_{k=0}^{\infty} z_k < +\infty$. For the second statement, setting $\eta_k = \frac{z_k}{2 - z_k}$, then we have

$$\prod_{k=0}^{l} (1 + \eta_k) \leq \exp(\sum_{k=0}^{l} \eta_k),$$

for all $l \geq 0$. In fact, as the function $x \mapsto \ln(1 + x) - x$ decreases on $[0, +\infty)$ and reaches its maximum value zero at $x = 0$,

$$\ln(1 + \eta_k) \leq \eta_k,$$

for all $k \geq 0$. Summing the last inequality over $k = 0, 1, ..., l$, we have

$$\sum_{k=0}^{l} \ln(1 + \eta_k) \leq \sum_{k=0}^{l} \eta_k,$$

for all $l \geq 0$. The result follows from the product rule for logarithms and from basics properties of the exponential function. □

Next, we will present full convergence theorem of proposed Algorithm 4.2.1. We will consider that $H : C \to \mathbb{R}^m$ is $\mathbb{R}^m_+$-quasi convex, C is convex set, and the following well-known assumption.

$\mathcal{H}1$: The set $(H(x^0) - \mathbb{R}^m_+) \cap H(C)$ is $\mathbb{R}^m_+$-complete.

Theorem 4.2.8. *Assume that* $\mathcal{H}1$ *holds true and* $\sum_{k=0}^{+\infty} \sigma_k < +\infty$. *Then, the sequence* $\{x^k\}$ *generated by the Algorithm 4.2.1, converges to a Pareto-Clarke critical point of* H.

Proof. Define

$$E := \bigcap_{k=0}^{+\infty} \Omega_k.$$

Assumption $\mathcal{H}\mathbf{1}$ implies that E is nonempty. Take $x^* \in E$, which infers that $x^* \in \Omega_k$ for $k \in \mathbb{N}$. It is easy to see that:

$$\|x^k - x^*\|^2 = \|x^{k+1} - x^*\|^2 + \|x^k - x^{k+1}\|^2 + 2\langle x^k - x^{k+1}, x^{k+1} - x^*\rangle, \ \forall k \in \mathbb{N}. \tag{4.2.12}$$

Following the steps of the proof of Theorem 4.2.6,

$$\lambda_k\langle e^k, \alpha^{k+1}\rangle(x^k - x^{k+1}) = A_{k+1}^T(\alpha^{k+1} + \beta^{k+1}) + \langle\epsilon^k, \alpha^{k+1}\rangle v^{k+1} + \tau_{k+1}w^{k+1}, \ \forall k \in \mathbb{N}. \tag{4.2.13}$$

Now, combining (4.2.12) with (4.2.13), we get

$$\begin{aligned} &\frac{\lambda_k b_k}{2}\left(\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 - \|x^k - x^{k+1}\|^2\right) \\ &= \left\langle A_{k+1}^T(\alpha^{k+1} + \beta^{k+1}) + \langle\epsilon^k, \alpha^{k+1}\rangle v^{k+1} + \tau_{k+1}w^{k+1}, x^{k+1} - x^*\right\rangle \\ &= \sum_{i=1}^{m}(\alpha_i^{k+1} + \beta_i^{k+1})\langle u_i^{k+1}, x^{k+1} - x^*\rangle + \sum_{i=1}^{m}\alpha_i^{k+1}\epsilon_i^k\langle v^{k+1}, x^{k+1} - x^*\rangle \\ &+ \tau_{k+1}\langle w^{k+1}, x^{k+1} - x^*\rangle, \end{aligned} \tag{4.2.14}$$

where $b_k = \langle e^k, \alpha^{k+1}\rangle$, $u_i^{k+1} \in \partial^\diamond h_i(x^{k+1}), \forall k \in \mathbb{N}$ and $i \in \{1, ..., m\}$. Since H is $\mathbb{R}_+^m$-quasi convex function, in particular, h_i is quasi convex for each $i \in \{1, ..., m\}$. As $x^* \in \Omega_k$ and $u_i^{k+1} \in \partial^\diamond h_i(x^{k+1})$, it follows by Proposition 2.4.19 that

$$\sum_{i=1}^{m}(\alpha_i^{k+1} + \beta_i^{k+1})\langle u_i^{k+1}, x^{k+1} - x^*\rangle \geq 0, \ \forall k \in \mathbb{N}. \tag{4.2.15}$$

As C is a convex set, $w^{k+1} \in N_C(x^{k+1})$ together with $\tau_{k+1} > 0$ and characterization of convex normal cone imply that

$$\tau_{k+1}\langle w^{k+1}, x^{k+1} - x^*\rangle \geq 0, \ \forall k \in \mathbb{N}. \tag{4.2.16}$$

By combining the inequalities (4.2.15), (4.2.16) with (4.2.14), we obtain

$$\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 - \|x^k - x^{k+1}\|^2 \ \geq -\frac{2}{\lambda_k b_k}\textstyle\sum_{i=1}^{m}\alpha_i^{k+1}\epsilon_i^k\langle v^{k+1}, x^{k+1} - x^*\rangle$$

$$\geq -\sigma_k \|x^{k+1} - x^k\| \|x^* - x^{k+1}\|,$$

for all $k \in \mathbb{N}$. As, $r + s \geq 2\sqrt{rs}$ holds for $r, s \geq 0$, taking $s := \|x^{k+1} - x^k\|$ and $r := \|x^* - x^{k+1}\|$, we obtain

$$\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 - \|x^k - x^{k+1}\|^2 \geq -\frac{\sigma_k}{2} \left[\|x^{k+1} - x^k\|^2 + \|x^* - x^{k+1}\|^2\right],$$

for all $k \in \mathbb{N}$. Thus, we get

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &\leq \left(\frac{1}{1-\sigma_k}\right) \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2 \\ &\leq \left(1 + \frac{\sigma_k}{1-\sigma_k}\right) \|x^k - x^*\|^2, \ \forall k \in \mathbb{N}. \end{aligned} \tag{4.2.17}$$

Since $\sum_{k=0}^{\infty} \sigma_k^2 < +\infty$, it follows that

$$K_0 := \sum_{k=k_0}^{+\infty} \frac{2\sigma_k^2}{1 - 2\sigma_k^2} < +\infty \ \text{ and } \ K_1 := \prod_{j=k_0}^{+\infty} \left(1 + \frac{2\sigma_j^2}{1 - 2\sigma_j^2}\right) < +\infty.$$

By (4.2.17), observe that for all $k \geq k_0$

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &\leq \left(1 + \frac{2\sigma_k^2}{1 - 2\sigma_k^2}\right) \|x^k - x^*\|^2 \\ &\leq \left(1 + \frac{2\sigma_{k-1}^2}{1 - 2\sigma_{k-1}^2}\right)\left(1 + \frac{2\sigma_k^2}{1 - 2\sigma_k^2}\right) \|x^{k-1} - x^*\|^2 \\ &\ \ \vdots \\ &\leq \prod_{j=k_0}^{k} \left(1 + \frac{2\sigma_j^2}{1 - 2\sigma_j^2}\right) \|x^{k_0} - x^*\|^2 \\ &\leq \prod_{j=k_0}^{\infty} \left(1 + \frac{2\sigma_j^2}{1 - 2\sigma_j^2}\right) \|x^{k_0} - x^*\|^2 \\ &= K_1 \|x^{k_0} - x^*\|^2. \end{aligned}$$

This shows that $\{x^k\}$ is bounded. Then (4.2.17) becomes

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 + \frac{2\sigma_k^2}{1 - 2\sigma_k^2} K^2, \ \forall k \in \mathbb{N}. \tag{4.2.18}$$

where $K = \sup_k \|x^k - x^*\|$. Take $\eta_k = \frac{2\sigma_k^2}{1-2\sigma_k^2} K^2$. Since $\eta_k > 0$ and $\sum_{k=1}^{\infty} \eta_k < +\infty$, we obtain that $\{x^k\}$ is quasi-Fejér convergent to E and boundedness of $\{x^k\}$ implies that the sequence $\{x^k\}$ has a cluster point $\bar{x}$. Since Theorem 4.2.6 implies that $\bar{x} \in E$. Therefore using Lemma 2.4.34 with $U = E$, we conclude that the whole sequence $\{x^k\}$ converges to $\bar{x}$ as k goes to $+\infty$, where $\bar{x}$ is a Pareto-Clarke critical point of H. □

Corollary 4.2.9. *If $C = \mathbb{R}^n$, $H : \mathbb{R}^n \to \mathbb{R}^m$ is $\mathbb{R}^m_+$-convex and locally Lipschitz function, then the sequence $\{x^k\}$ converges to a weak Pareto optimal point of H.*

Proof. It is immediate from Proposition 2.4.15. □

4.3 Conclusion

In this chapter, we presented an inexact proximal point algorithm for constrained multiobjective optimization problems under the locally Lipschitz condition of the cost function. Convergence analysis of the considered method, Fritz-John necessary optimality condition of ϵ-quasi weakly Pareto solution in terms of Clarke subdifferential is derived. The suitable conditions to guarantee that the accumulation points of the generated sequences are Pareto-Clarke critical points are provided.

CHAPTER V

HYBRID PROXIMAL POINT ALGORITHM FOR SOLUTION OF CONVEX MULTIOBJECTIVE OPTIMIZATION PROBLEM OVER FIXED POINT CONSTRAINT

In this chapter, we consider the convex constraint multiobjective optimization problem as the fixed point set of nonexpansive mapping. By owing the concepts of proximal method and Mann algorithm, we introduce the algorithm and aim to establish the convergence results of the such proposed iterative algorithm to compute a solution point of the considered constraint convex multiobjective optimization problem. We consider the constraint multiobjective minimization problem:

$$\min_{y \in Fix(T)} H(y), \tag{5.0.1}$$

where $H : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^n \to \mathbb{R}^n$. We will show that under some suitable conditions the following modified proximal point algorithm for multiobjective optimization involving Mann iterate in $\mathbb{R}^n$, $z \in \mathbb{R}^m_+ \setminus \{0\}$:

$$\begin{cases} \tilde{x}^k = \arg\min_{y \in \mathbb{R}^n} \{ \langle H(y) + \frac{\lambda_k}{2} \|y - x^k\|^2 e^k, z \rangle \} \\ x^{k+1} = (1 - \alpha_k)\tilde{x}^k + \alpha_k T\tilde{x}^k \end{cases} \tag{5.0.2}$$

converges to a weak Pareto optimal point of the constraint multiobjective optimization problem (5.0.1).

5.1 Multiobjective Optimization Problem Over Fixed Point Constraint

In this section, we prove the main convergence theorem of proposed iterative scheme. We will work under the following assumptions:

(A1) H is $\mathbb{R}^m_+$-convex function.

(A2) H is positively lower semicontinuous function.

(A3) T is a nonexpansive mapping.

(A4) There is $z \in \mathbb{R}^m_+ \setminus \{0\}$ such that $\Upsilon_z := \arg\min_{y\in\mathbb{R}^n}\{\langle H(y), z\rangle\} \bigcap \mathrm{Fix}(T)$ is a nonempty set.

We proceed with the following main tools.

Lemma 5.1.1. *Assume the the assumptions (A1)- (A4) are satisfied. Then, under the following control conditions:*

(i) $\{\lambda_k\}$ *is a bounded sequence of positive real numbers,*

(ii) $\{\alpha_k\}$ *is a sequence such that* $0 < a \leq \alpha_k \leq b < 1$, $\forall k \geq 1$ *and for some constant* a, b *in* $(0, 1)$,

the sequence $\{x^k\}$ *which is generated by the algorithm (5.0.2), with respect to* z, *satisfies the following items*

(i) $\lim_{k\to\infty} \|x^k - x^*\|$ *exists for all* $x^* \in \Upsilon_z$;

(ii) $\lim_{k\to\infty} \|x^k - \tilde{x}^k\| = 0$;

(iii) $\lim_{k\to\infty} \|Tx^k - x^k\| = 0$.

Proof. Let $x^* \in \Upsilon_z$. So, we have $x^* = Tx^*$ and $x^* \in \arg\min_{y\in\mathbb{R}^n}\{\langle H(y), z\rangle\}$. Then for all $y \in \mathbb{R}^n$, we acquire

$$\langle H(x^*), z\rangle + \frac{\lambda_k}{2}\|x^* - x^*\|^2\langle e^k, z\rangle \leq \langle H(y), z\rangle + \frac{\lambda_k}{2}\|y - x^*\|^2\langle e^k, z\rangle.$$

Let us note that, since H is positively lower semicontinuous, then we have the scalar valued function $\langle H(y), z\rangle$ is lower semicontinuous and convexity of H implies the convexity of $\langle H(y), z\rangle$, hence $\phi_z(y) = \langle H(y), z\rangle$ is proper, convex and lower semicontinuous function.

Now define $\beta_k = \lambda_k \langle e^k, z\rangle$. Note that, $\beta_k > 0$ because $\lambda_k > 0$, $z \in \mathbb{R}^m_+ \setminus \{0\}$, $\{e^k\} \subset \mathbb{R}^m_{++}$. So we get that $x^* = \operatorname{prox}^{\phi_z}_{\beta_k} x^*$.

(i) Now, we first show that $\lim_{k\to\infty} \|x^k - x^*\|$ exists for all $x^* \in \Upsilon_z$. Noting that $\tilde{x}^k = \operatorname{prox}^{\phi_z}_{\beta_k} x^k$ for all $k \geq 1$. So, Lemma 2.4.29 provide us

$$\|\tilde{x}^k - x^*\| = \|\operatorname{prox}^{\phi_z}_{\beta_k} x^k - \operatorname{prox}^{\phi_z}_{\beta_k} x^*\| \leq \|x^k - x^*\|. \tag{5.1.1}$$

It follows from the algorithm (5.0.2) and nonexpansiveness of T that

$$\begin{aligned} \|x^{k+1} - x^*\| &= \|(1-\alpha_k)\tilde{x}^k + \alpha_k T\tilde{x}^k - x^*\| \\ &\leq (1-\alpha_k)\|\tilde{x}^k - x^*\| + \alpha_k\|T\tilde{x}^k - x^*\| \\ &\leq (1-\alpha_k)\|\tilde{x}^k - x^*\| + \alpha_k\|\tilde{x}^k - x^*\| \qquad (5.1.2) \\ &= \|\tilde{x}^k - x^*\| \\ &\leq \|x^k - x^*\|, \ \forall k \geq 1. \end{aligned}$$

This shows that $\{\|x^k - x^*\|\}$ is decreasing and bounded below. Hence $\lim_{k\to\infty} \|x^k - x^*\|$ exists for all $x^* \in \Upsilon_z$.

(ii) In order to proceed for part (ii), we assume without loss of any generality that

$$\lim_{k\to\infty} \|x^k - x^*\| = c \geq 0. \tag{5.1.3}$$

Indeed by (3.1.7), we have

$$\frac{1}{2\lambda_k}\left(\|\tilde{x}^k - x^*\|^2 - \|x^k - x^*\|^2 + \|x^k - \tilde{x}^k\|^2\right) \leq \phi_z(x^*) - \phi_z(\tilde{x}^k).$$

Since $\phi_z(x^*) \leq \phi_z(\tilde{x}^k)$ for all $k \geq 1$, it follows that

$$\|x^k - \tilde{x}^k\|^2 \leq \|x^k - x^*\|^2 - \|\tilde{x}^k - x^*\|^2. \tag{5.1.4}$$

Therefore in order to prove $\lim_{k\to\infty} \|x^k - \tilde{x}^k\| = 0$, it suffices to prove $\|\tilde{x}^k - x^*\| \to c$, because $\|x^k - x^*\| \to c$.

Taking lim inf on both sides of the estimate (5.1.2), we have

$$c \leq \liminf_{k\to\infty} \|\tilde{x}^k - x^*\|. \tag{5.1.5}$$

On the other hand, by taking lim sup on both sides of the estimate (5.1.1), we get

$$\limsup_{k\to\infty} \|\tilde{x}^k - x^*\| \leq \limsup_{k\to\infty} \|x^k - x^*\| = c.$$

Hence, the above estimate together with (5.1.5) implies that

$$\lim_{k\to\infty} \|\tilde{x}^k - x^*\| = c. \tag{5.1.6}$$

Therefore, from (5.1.4), we obtain

$$\lim_{k\to\infty} \|x^k - \tilde{x}^k\| = 0. \tag{5.1.7}$$

(iii) Next, we prove that $\lim_{k\to\infty} \|Tx^k - x^k\| = 0$. As, we observe from Lemma 2.4.36 that

$$\begin{aligned}\|x^{k+1} - x^*\|^2 &= \|(1-\alpha_k)\tilde{x}^k + \alpha_k T\tilde{x}^k - x^*\|^2 \\ &= \|(1-\alpha_k)(\tilde{x}^k - x^*) + \alpha_k(T\tilde{x}^k - x^*)\|^2 \\ &= (1-\alpha_k)\|\tilde{x}^k - x^*\|^2 + \alpha_k\|T\tilde{x}^k - x^*\|^2 \\ &- \alpha_k(1-\alpha_k)\|\tilde{x}^k - T\tilde{x}^k\|^2.\end{aligned}$$

Since T is nonexpansive, it follows that

$$\begin{aligned}\|x^{k+1} - x^*\|^2 &\le (1-\alpha_k)\|\tilde{x}^k - x^*\|^2 + \alpha_k\|\tilde{x}^k - x^*\|^2 - \alpha_k(1-\alpha_k)\|\tilde{x}^k - T\tilde{x}^k\|^2 \\ &\le \|\tilde{x}^k - x^*\|^2 - \alpha_k(1-\alpha_k)\|\tilde{x}^k - T\tilde{x}^k\|^2 \\ &\le \|x^k - x^*\|^2 - a(1-b)\|\tilde{x}^k - T\tilde{x}^k\|^2.\end{aligned}$$

This implies that

$$\|\tilde{x}^k - T\tilde{x}^k\|^2 \le \frac{1}{a(1-b)}(\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2) \to 0 \text{ (as } k \to \infty), \tag{5.1.8}$$

Since T nonexpansive and from (5.1.7), (5.1.8) we obtain that

$$\begin{aligned}\|\tilde{x}^k - Tx^k\| &\le \|\tilde{x}^k - T\tilde{x}^k\| + \|T\tilde{x}^k - Tx^k\| \\ &\le \|\tilde{x}^k - T\tilde{x}^k\| + \|\tilde{x}^k - x^k\| \to 0 \ (as \ \ k \to \infty).\end{aligned} \tag{5.1.9}$$

Now, we can prove that $\lim_{k\to\infty} \|x^k - Tx^k\| = 0$. From (5.1.7) and (5.1.9), we obtain

$$\|x^k - Tx^k\| \le \|x^k - \tilde{x}^k\| + \|\tilde{x}^k - Tx^k\| \to 0 \ (as \ \ k \to \infty). \tag{5.1.10}$$

□

Now we are in position to present our main theorem.

Theorem 5.1.2. *Assume the the assumptions (A1)- (A4) are satisfied. Then, under the following control conditions:*

(i) $\{\lambda_k\}$ *is a bounded sequence of positive real numbers such that* $\lambda_k \geq \lambda > 0$, *for some positive real number* λ;

(ii) $\{\alpha_k\}$ *is a sequence such that* $0 < a \leq \alpha_k \leq b < 1$, $\forall k \geq 1$ *and for some constant* a, b *in* $(0, 1)$,

the sequence $\{x^k\}$, *which is generated by the algorithm (5.0.2), with respect to* z, *converges to a weak Pareto optimal point of the constraint multiobjective optimization problem (5.0.1).*

Proof. In fact, it follows from (5.1.7) and Lemma 2.4.31 that

$$\begin{aligned}\|\operatorname{prox}_\lambda^{\phi_z} x^k - x^k\| &\leq \|\operatorname{prox}_\lambda^{\phi_z} x^k - \tilde{x}^k\| + \|\tilde{x}^k - x^k\| \\ &= \|\operatorname{prox}_\lambda^{\phi_z} x^k - \operatorname{prox}_{\beta_k}^{\phi_z} x^k\| + \|\tilde{x}^k - x^k\| \\ &= \|\operatorname{prox}_\lambda^{\phi_z} x^k - \operatorname{prox}_\lambda^{\phi_z}\left(\frac{\beta_k - \lambda}{\beta_k}\operatorname{prox}_{\beta_k}^{\phi_z} x^k + \frac{\lambda}{\beta_k}x^k\right)\| \\ &+ \|\tilde{x}^k - x^k\| \\ &\leq \|x^k - (1 - \frac{\lambda}{\beta_k})\operatorname{prox}_{\beta_k}^{\phi_z} x^k - \frac{\lambda}{\beta_k}x^k\| + \|\tilde{x}^k - x^k\| \\ &\leq (1 - \frac{\lambda}{\beta_k})\|x^k - \tilde{x}^k\| + \|\tilde{x}^k - x^k\| \to 0 \ (as \ \ k \to \infty)\end{aligned} \tag{5.1.11}$$

Moreover, by (5.1.2), we have that $\{x^k\}$ is Fejér convergent to Υ_z. So, it guarantees that $\{x^k\}$ is bounded. Then there exists a subsequence $\{x^{k_i}\} \subset \{x^k\}$ such that $x^{k_i} \to p^*$. By (5.1.10),

$$\|x^{k_i} - Tx^{k_i}\| \to 0.$$

It follows that $p^* \in \operatorname{Fix}(T)$. Also, from (5.1.11), we have

$$\|x^{k_i} - \operatorname{prox}_\lambda^{\phi_z} x^{k_i}\| \to 0.$$

Since $\operatorname{prox}_\lambda^{\phi_z}$ is a nonexpansive mapping. Then, we get that $p^* \in \operatorname{Fix}(\operatorname{prox}_\lambda^{\phi_z}) = \arg\min_{y\in\mathbb{R}^n}\phi_z(y)$. This shows that $p^* \in \Upsilon_z$. Therefore, using Lemma 2.4.35 with

$U = \Upsilon_z$, we conclude that the whole sequence $\{x^k\}$ converges to p^* as k goes to ∞.

Further, we can see by equality (2.4.15) that $p^* \in \arg\min_w\{H(y)|y \in \mathbb{R}^n\}$. Hence $p^* \in \arg\min_w\{H(y)|y \in \mathbb{R}^n\} \cap \mathrm{Fix}(T)$. Finally, by Proposition 2.4.38, we obtain that the sequence $\{x^k\}$ converges to a weak Pareto optimal point of the constraint multiobjective optimization problem (5.0.1). This completes the proof. □

Remark 5.1.3. It is remarked that if we take constraint set $\mathrm{Fix}(T) = \mathbb{R}^n$, then we get multiobjective optimization problem (2.2), which is done by many authors by different methods, see for instance [6, 7, 9].

5.2 Conclusion

The purpose of this chapter is to consider the convex constraint multiobjective optimization problem, as the fixed point set of nonexpansive mapping. By owing the concepts of proximal method and Mann algorithm, we introduce the algorithm and aim to establish the convergence results of the such proposed iterative algorithm to compute a solution point the considered constraint convex multiobjective optimization problem.

CHAPTER VI

CONCLUSION AND FUTURE WORK

In this thesis, we studied and introduced some proximal point algorithm for multiobjective optimization. This thesis is composed of 5 chapters. We give here necessary and useful information about these chapters.

In chapter 2, we summarized the various well-known definitions and results, which provided and presented a necessary and essential background for the subsequent chapters. All contents of chapter 2 are known and properly referred.

In chapter 3, we considered the proximal point algorithm for multiobjective optimization which was introduced by Bento et al. [5]. The main is to relax the conditions on the considered objective function. Indeed, the work presented in chapter 3 extends the class of functions from continuously differentiable to differentiable and locally Lipschitz.

In chapter 4, we developed an inexact version of proximal point method of Bento et al. [21]. In terms of Clarke subdifferential, we introduced Fritz-John necessary optimality condition of ϵ-quasi weakly Pareto solution which we apply for convergence analysis of proposed method. We have presented that the proposed method is well defined and under some suitable conditions the sequence attained by proposed method converges to a Pareto-Clarke critical point. The newly proposed inexact proximal point algorithm is important because of its practical point of view. The proximal point method is a conceptual algorithm, and its computational performance strongly depends on the method used to solve the subproblems. Hence, in practice computations introduce numerical errors in order to solve the auxiliary minimization problems and these methods usually provide only approximate solutions of the subproblems. Clearly, it is very important, from the view of practice, to study the asymptotic behavior of iterations of the algorithm in the presence of computational errors.

In chapter 5, we have considered the convex constraint multiobjective optimization problem when the constrained set is a fixed point set of nonexpansive mapping. By combining the concepts of proximal method and Mann algorithm, we introduced the algorithm and provided the convergence results of such proposed iterative algorithm to compute a solution point of the considered constraint convex multiobjective optimization problem.

In the future, we do think that the work in chapter 3 is extendable for the problems involving non-differentiable functions by defining the Pareto critical points using directional derivatives. We also intend to introduce the proximal point algorithm for quasi-convex multiobjective optimization constrained to non-expansive fixed point mapping for finding weak Pareto solution, which can be seen as the extension of work done in chapter 5.

The results in chapter 5 also suggest us the algorithm for finding the weak Pareto optimal point of the considered problem (5.0.1), of course, the convergence analysis of this suggested algorithm and also the (new) updated algorithms for finding the Pareto optimal point of this problem (5.0.1) should be considered in the future works. We also intend to propose a subgradient method for essentially quasi-convex multiobjective optimization constrained to non-expansive fixed point mapping for finding Pareto solution.

REFERENCES

REFERENCES

1. Martinet, B.: Regularisation, d' indquations variationelles par approximations succesives, Rev. Franaise d'Inform. Recherche Oper. 4, 154-159, 1970.
2. Rockafellar, R. T.: Monotone operators and the proximal point algorithm, SIAM J. Control Optim., 14, 877-898, 1976.
3. Iusem, A. N.: Augmented Lagrangian methods and proximal point methods for convex optimization, Investigación Operativa 8, 11-49, 1999.
4. Apolinário, H.C.F., Papa Quiroz, E.A., Oliveira, P.R., A scalarization proximal point method for quasiconvex multiobjective minimization. J. of Global Optim,. 64(1), 79-96, 2016.
5. Bento, G.C., Cruz Neto, J.X., Soubeyran, A.: A proximal point-type method for multicriteria optimization. Set-Valued Vari. Anal., 22(3), 557-573, 2014.
6. Bonnel, H., Iusem, A.N., Svaiter, B.F.: Proximal methods in vector optimization. SIAM Journal of Optimization. 15(4), 953-970, 2005.
7. Burachik, R. S., Kaya, C. Y., Rizvi, M.M.: A new scalarization technique to approximate Pareto fronts of problems with disconnected feasible sets, J. Optim. Theory Appl., 162, 428-446, 2014.
8. Durea M., Strugariu, R.: Some remarks on proximal point algorithm in scalar and vectorial cases, Nonlinear Funct. Anal. Appl., 15, 307-319, 2010.
9. Grana Drummond, L. M., Iusem, A. N.: A projected gradient method for vector optimization problems, Comput. Optim. Appl., 28, 5–30, 2004.
10. Picard, E.: Memoire sur la theorie des equations aux derivees partielles et la methode des approximations successives. J. Math. Pures et Appl., vol. 6, 1890.
11. Banach, S.: Sur les oprations dans les ensembles abstraits et leur application aux equations integrales. Fund. Math. vol. 3, 133-181, 1922.
12. Miettinen, K.M.: Nonlinear Multiobjective Optimization. Kluwer Academic Publishers, Boston, 1999.
13. Chuong, T.D., Kim, D.S.: Approximate solutions of multiobjective optimization problems. Positivity 20, 187–207, 2016

14. Luc, D.T.: Theory of Vector Optimization. Lecture Notes in Economics and Mathematical Systems. Springer, Berlin, 1989.

15. Mann, W. R.: Mean value methods in iteration. Proc. Amer. Math. Soc. 4, 506-510, 1953.

16. Chang, S.S., Wu, D.P., Wang, L., Wang, G.: Proximal point algorithms involving fixed point of nonspreading-type multivalued mappings in Hilbert spaces. J. Nonlinear Sci. Appl. 9, 5561–5569, 2016.

17. Phuengrattana, W., Tiammee. J.: Proximal point algorithms for finding common fixed points of a finite family of quasi-nonexpansive multi-valued mappings in real Hilbert spaces. J. Fixed Point Theory Appl. 2018.

18. Bolte, J., Daniilidis, A., Lewis, A., Shiota, M.: Clarke subgradients of stratifiable functions. SIAM J. Optim. 18, 556–572, 2007.

19. Rockafellar, R.T., Wets, R.: Variational Analysis. Grundlehren der Mathematischen Wissenschafte. Springer Verlag, New York, 1998.

20. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation. Vol. I: Basic Theory, Vol. II: Applications. Springer, Berlin, 2006.

21. Bento, G.D.C., Cruz Neto, J.X., López, G., Soubeyran, A., Souza, J.C.O.,: The proximal point method for locally Lipschitz functions in multiobjective optimization with application to the compromise problem. SIAM Journal on Optimization, 28(2), 1104-1120, 2018

22. Clarke, H.F.: Optimization and Nonsmooth Analysis, Classics in Applied Mathematics. SIAM, New York, 1983.

23. Clarke, F. H.: Generalized gradients and applications, Trans. Amer. Math. Soc., 205, 247–262, 1975.

24. Burke, J. V., Ferris, M. C., Qian M.: On the Clarke subdifferential of the distance function of a closed set, J. Math. Anal. Appl., 166, 199–213, 1992.

25. Bento, G.C., Cruz Neto, J.X., Oliveira, P.R., Soubeyran, A.: The self regulation problem as an inexact steepest descent method for multicriteria optimization. Eur. J. Oper. Res., 235, 494-502, 2014.

26. Vinter, R.B.: Optimal Control. Birkhauser, Basel, 2000.

27. Apolinario, H.C., Papa Quiroz, E.A., Oliveria, P.R.: A scalarization proximal point method for quasiconvex multiobjective minimization. J. Glob. Optim. 64, 79–96, 2016.

28. Penot J. P. Circa-Subdifferentials, Clarke Subdifferentials. In: Calculus Without Derivatives. Graduate Texts in Mathematics, vol 266. Springer, New York, NY, 2013.

29. Aussel, D.: Subdifferential properties of quasiconvex and pseudoconvex functions: unified approach. J. Optim. Theory Appl. 97(1), 29–45, 1998.

30. Ceng, Lu-Chuan, and Jen-Chih Yao. "Approximate proximal methods in vector optimization." European Journal of Operational Research 183(1), 1-19, 2007.

31. Quiroz, E.P., Ramirez, L.M. and Oliveira, P.R.: An inexact proximal method for quasiconvex minimization. European Journal of Operational Research, 246(3), 721-729, 2015

32. Branke, J., Deb, K., Miettinen, K., Slowinski, R. eds.: Practical approaches to multiobjective optimization, Dagstuhl Seminar, Dagstuhl, Wadern, Germany, 06501, 2007

33. Bento, G.D.C., da Cruz Neto, J.X. and de Meireles, L.V.: Proximal Point Method for Locally Lipschitz Functions in Multiobjective Optimization of Hadamard Manifolds. Journal of Optimization Theory and Applications, 179(1), 37-52, 2018.

34. Souza, J.C.D.O.,: Proximal Point Methods for Lipschitz Functions on Hadamard Manifolds: Scalar and Vectorial Cases. Journal of Optimization Theory and Applications, 179(3), 745-760, 2018.

35. Eichfelder, G.: An adaptive scalarization method in multiobjective optimization, SIAM J. Optim., 19, 1694-1718, 2009.

36. Ariza-Ruiz, D., Leustean, L. Lóez, G.: Firmly nonexpansive mappings in classes of geodesic spaces. Trans. Amer. Math. Soc., 366, 4299–4322, 2014.

37. Jost, J.: Convex functionals and generalized harmonic maps into spaces of nonpositive curvature, Comment. Math. Helv., 70, 659–673, 1995.

38. Ambrosio, L., Gigli, N., Savaré, G.: Gradient flows in metric spaces and in the space of probability measures, Second edition, Lectures in Mathematics ETH Zurich, Birkhauser Verlag, Basel, 2008.

39. Bauschke H. H., Combettes, P. L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, London, UK, 2017. Set-Valued Variational Analysis. 22(3), 557-573, 2014.

40. Schott, D.: Basic properties of Fejér monotone sequences. Rostocker Math. Kolloqu. 49, 57–74, 1995.

41. Yu. M. Ermolev and A. D. Tuniev, Random Fejer and quasi-Fejer sequences, Theory of Optimal Solutions– Akademiya Nauk Ukrainskoı SSR Kiev, vol. 2, pp. 76–83, 1968; translated in: American Mathematical Society Selected Translations in Mathematical Statistics and Probability, vol. 13, pp. 143–148, 1973.

42. Günther, C., Tammer, C.: Relationships between constrained and unconstrained multi-objective optimization and application in location theory. Mathematical Methods of Operations Research, 84(2), 359-387, 2016.