The Hong Kong Polytechnic University
Department of Applied Mathematics

Multiobjective Optimization Problems, Vector Variational Inequalities and Proximal-type Methods

by

Zhe Chen

A thesis submitted in partial fulfilment of
the requirements for the degree of Doctor of Philosophy

April 2010

CERTIFICATE OF ORIGINALITY

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of a university or other institute of higher learning, except where due acknowledgment is made in the text.

ZHE CHEN

Abstract

The main purpose of this thesis is to study the asymptotical properties of multiobjective optimization (also known as vector optimization) and vector variational inequalities. Based on these asymptotical properties, we construct some proximal-type methods for solving convex multiobjective optimization problems and weak vector variational inequality problems.

We consider a convex vector optimization problem of finding weak Pareto optimal solutions for an extended vector-valued map from a uniformly convex and uniformly smooth Banach space to a real Banach space, with the latter being ordered by a closed, convex and pointed cone with nonempty interior. For this problem, we develop an extension of the classical proximal point method for the scalar-valued convex optimization. In this extension, the subproblems involve the finding of weak Pareto optimal solutions for some suitable regularizations of the original map by virtue of a Lyapunov functional. We present both exact and inexact versions. In the latter case, the subproblems are solved only approximately within an exogenous relative tolerance. In both cases, we prove weak convergence of the sequences generated by the subproblems to a weak Pareto optimal solution of the vector optimization problem.

We also construct a generalized proximal point algorithm to find a weak Pareto optimal solution of minimizing an extended vector-valued map with respect to the positive orthant in finite dimensional spaces. In this extension, the subproblems involve finding weak Pareto optimal solutions for the regularized map by employing a vector-valued Bregman distance function. We prove that the sequence generated by this method converges to a weak Pareto optimal solution of the multiobjective optimization problem by assuming that the original multiobjective optimization problem has a nonempty and compact weak Pareto optimal solution set.

We formulate a matrix-valued proximal-type method to solve a weak vector variational inequality problem with respect to the positive orthant in finite dimensional spaces through normal mappings. We also carry out convergence analysis on the method and prove the convergence of the sequences generated by the matrix-valued proximal-type method to a solution of the original problem under some mild conditions.

Finally, we investigate the nonemptiness and compactness of the weak Pareto optimal solution set of a multiobjective optimization problem with functional constraints via asymptotic analysis. We then employ the obtained results to derive the necessary and sufficient conditions of the weak Pareto optimal solution set of a parametric multiobjective optimization problem.

The study of this thesis has used tools from nonlinear functional analysis, multiobjective programming theory, vector variational inequality theory, asymptotical analysis and numerical linear algebra.

Acknowledgments

First and foremost, I wish to express my deep gratitude to my supervisor Professor Xiaoqi Yang for his constant encouragement and helpful guidance. It has been a precious experience for me to study for my PhD degree under his supervision in the Department of Applied Mathematics at The Hong Kong Polytechnic University. Professor Yang provides me endless new and brilliant ideas on my research (especially when my ideas are all exhausted) and he also gives me many opportunities to explore various fields of mathematics which I have never encountered before. At any time of the day, he is always willing to discuss and give valuable insights to my research. I would like to thank Professor Xuexiang Huang (Chongqing University, China) for his comments and suggestions during the course of this research. I have many fruitful and stimulating discussions with him. I would also like to thank Professor Guang-ya Chen (Chinese Academy of Sciences, China) for his careful reading, useful comments and suggestions on this thesis, which have improved it greatly.

I wish to thank Professor Liqun Qi for his encouragement and help. To Professor Xinmin Yang (Chongqing Normal University, China), Professor Nanjing Huang (Sichuan University, China), Professor Shengjie Li (Chongqing University, China), Professor Chen Ling (Zhejiang University of Finance and Economics, China), Professor Xiyin Zheng (Yunnan University, China), Professor Yiju Wang (Qufu Normal University, China), Dr. Kai Zhang (Shenzhen University, China), Dr. Chunrong Chen (Chongqing University, China) and Mr. Kequan Zhao (Chongqing Normal University, China), I am grateful for their suggestions and help. I also wish to thank Mr. Dingqun Bai, Mr. Zhangyou Chen, Mr. Donghui Fang, Mr. Haiqiao Huang, Mr. Kaiwen Meng and Mr. Jie Xin for their friendship and help. They have made my time at The Hong Kong Polytechnic University more enjoyable. In addition, I would like to acknowledge a number of editors and anonymous referees from several journals, who have given me valuable comments and suggestions which have contributed to a better presentation of several papers, on which this thesis is based.

I would like to thank my parents for encouraging me to keep on studying. Many thanks to Winnie Xueyin Wang, my wife for her love, encouragement, support, sacrifices and tolerance. I thank them for always being by my side, even during those moments when I was astray and down.

Last but not least, thanks are also due to the Research Committee of The Hong Kong Polytechnic University for offering me financial support over the last three years. I sincerely thank Fanny Cheung and Cynthia Hau for their kindness and support. I wish to thank the staff of the Department of Applied Mathematics of The Hong Kong Polytechnic University for their kindness and generosity. I shall always treasure the stimulating research environment and the excellent research facilities of the Department of Applied Mathematics at The Hong Kong Polytechnic University.

Contents

Chapter 1

Introduction

1.1 Proximal-type Methods

Let $f : R^n \to R \cup \{+\infty\}$ be a proper, closed and convex function. Consider the following optimization problem:

$$(\mathrm{P}) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in R^n. \end{array}$$

As f is allowed to be extended-valued, any constrained optimization problem,

$$(\mathrm{CP}) \qquad \begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & x \in K, \end{array}$$

where $f_0 : K \to R$ is a convex function, and K is a closed, nonempty and convex subset of R^n, can easily be transformed into a form of (P) by defining

$$f(x) = \begin{cases} f_0(x), & \text{if } x \in K; \\ +\infty, & \text{else.} \end{cases}$$

Hence, theoretically, there is no loss of generality in considering the form of problem (P). Many theoretical results can be obtained in this (unifying) method for both unconstrained and constrained optimization problems.

Despite the fact that extended-valued functions allow such a unified treatment of both unconstrained and constrained optimization problems, they are typically not tractable from a numerical computational point of view. Therefore, numerical algorithms for the solution of a problem such as problem (P) have to account for the constraints implicitly or, at least, some of these constraints. This can be done quite elegantly by the proximal point methods.

Although our review is concentrated on the application of proximal point methods to optimization problems, these methods may also be applied to several other problem classes such as nonlinear systems of equations, complementarity problems, variational inequalities and generalized equations.

1.1.1 Classical Proximal-Point Methods

The classical proximal-point method was introduced by Martinet [158] in an attempt to alleviate the difficulties in a Tikhonov regularization method [195] when solving an increasingly progressively ill-conditioned problem. Martinet was motivated by a similar approach used in [21] in the case of convex quadratic minimization problems. In the finite-dimensional case, the main difference between the Tikhonov regularization and the proximal point method is that the point to which the sequence produced by the latter method converges cannot be predicted. In the infinite-dimensional setting, which has been considered by many authors, a more important difference exists: for the Tikhonov regularization methods, strong convergence to a solution of the original problem can be proven, whereas for the proximal point method, only weak convergence can be obtained unless further strong assumptions are made.

Rockafellar's paper [173] is an important step toward a wider appreciation of the importance of the proximal point method. In [173], the proximal point algorithm was analyzed for the problem of finding a zero in a maximal monotone map, which includes that of solving a variational inequality problem and a convex optimization problem. Actually, this general framework persists in most subsequent papers dealing with proximal point algorithms. In the mentioned paper, Rockafellar presented several significant results. First the coefficient λ_k [see (1.1.1)] was allowed to vary from iteration to iteration (whereas it was fixed in [158]). Second, and more importantly, the inexact solution of the perturbed subproblems was allowed. Some convergence rate results were given.

Applying it to the minimization problem (P), it generates a sequence $\{x_k\} \subset R^n$ such that x_{k+1} is a solution of the following optimization problem:

$$\min\{f(x) + \frac{\lambda_k}{2}\|x - x_k\|^2\} \tag{1.1.1}$$

for $k = 0, 1, 2, ...,$ $\lambda_k \in (0, \lambda]$ and $\lambda > 0$. The objective function of this subproblem is strictly convex as it is the sum of the original (convex) objective function f and a strictly convex quadratic term. This term is usually called the regularization term. This strictly convex regularization term guarantees that the subproblem (1.1.1) has a unique solution for each $k \in N$. Hence the classical proximal point method is well defined. Furthermore, it has the following global convergence properties.

Theorem 1.1.1 *[101] Let x_k and λ_k be defined in the classical proximal point method (1.1.1). Let $\sigma_k = \sum_{i=1}^{k} \lambda_i$. Let*

$$\bar{f} := \inf\{f(x) |\ x \in R^n\}$$

be the optimal value and

$$\bar{X} := \{\bar{x} \in R^n |\ f(\bar{x}) = \bar{f}\}$$

be the solution set of problem (P). Assume that $\sigma_k \to +\infty$ as $k \to +\infty$. Then, we have the following statements:

(1). The sequence of function values $\{f(x_k)\}$ converges to the optimal value $\bar{f}$;

(2). If $\bar{X}$ is nonempty, the whole sequence $\{x_k\}$ converges to an element of $\bar{X}$.

Theorem 1.1.1 shows some very strong convergence properties under rather weak conditions. In particular, it guarantees the convergence of the whole sequence $\{x_k\}$, even if the solution set $\bar{X}$ contains more than one element. We note that many variations and generalizations of Theorem 1.1.1 have been investigated in the literature. For example, Burachik, Iusem and Svaiter [40] introduced an enlargement of monotone operators and used it to define a family of inexact proximal point algorithms for a variational inequality problem. Ferris [86] demonstrated the finite termination by assuming the "weak sharpness" of the solution. For an extensive discussion on the finite termination property of the proximal point method in the case of both optimization and variational inequality problems, see [166].

Rockafellar also gave another important contribution by showing that the Hestenes-Powell method of multipliers for nonlinear programming is nothing else but the proximal point algorithm applied to the dual optimization problem [174] and that new methods of multipliers can be developed based on this observation. The idea of using under and overrelaxations was put forward by Golshtein and Tretyakov [100], who proposed a scheme similar to the generalized proximal point algorithm, where, λ_k is not allowed to vary from iteration to iteration.

Several other aspects of the proximal point algorithm have been intensively investigated. The first is that of the convergence rate already mentioned in the discussion of [173]. Luque [155] extended Rockafellar's analysis. By assuming various conditions on the growth properties of T^{-1} in a neighborhood of 0 and by considering broad inexactness rules, Luque studied in detail when the sequence x_k generated by the generalized proximal point algorithm converges sublinearly, linearly, or superlinearly. Another contribution of Luque's study is that, in contrast to the analysis in [173], the convergence rate results are obtained without assuming uniqueness of the solution of the unperturbed problem.

The convergence rates alluded to deserve an elaboration. These rates refer to the outer sequence x_k produced by the algorithm. However, to generate these iterates, an iterative inner method has to be used to calculate the resolvent approximately;; the latter typically requires multiple iterations. Therefore, calculating x_k can in general require a possibly huge computational effort. This issue has been recently addressed by Yamashita and Fukushima [204], where the proximal point algorithm is considered to solve a nonlinear complementary problem. Using the Newton method applied to the Fischer-Burmeister equation reformulation of the subproblems, the authors gave conditions ensuring that only one step of the inner Newton method is eventually sufficient to generate a suitable outer x_k, thus ensuring a genuine superlinear convergence rate. A key assumption in this reference is that the point to which the method converges is nondegenerate.

From [173] it is not necessary to compute the exact minimizer of the subproblems (1.1.1) at each step Rockafellar [173] provided some criteria under which inexact solutions still provide similar global convergence properties. However, the criteria of inexactness in Rockafellar [173] are not implementable in general because they assume some knowledge regarding the exact solution of (1.1.1). Solodov and Svaiter have been active in constructing of new variants of the proximal point algorithms [182, 183, 185], where their main concern is the use of some new and more practical and/or effective tolerance requirements in the solution of the subproblems.

The classical proximal-point method can be extended to infinite-dimensional Hilbert spaces for which weak convergence of the iterates $\{x_k\}$ can be shown, (see, e.g. [101, 173] and the references therein). It is worth noting that Güler provided a counterexample showing that a strong convergence does not hold without any further modifications, giving a negative answer to an open question in [173]. The classical proximal point method has been constructed in Banach spaces (see, e.g. [46, 47]). It is interesting that Solodov and Svaiter [184] were able to modify the proximal point algorithm to enforce strong convergence of the method in Hilbert spaces using a technique similar to the one used in the hyperplane projection algorithm. A similar result was achieved in [146] by combining the proximal point algorithm with Tikhonov regularization. Recently, Kamimura and Takahashi [129] have extended some of Solodov and Svaiter's results to more general Banach spaces.

1.1.2 Proximal-type Methods Using Bregman Functions

The simple idea behind each proximal-like method for the solution of convex minimization problems is to replace the strictly convex quadratic term by a nonquadratic term in the regularized subproblem (see, e.g., [23]).

There are several different possibilities in replacing the term $\|x - x_k\|^2$ by another strictly convex distance-like function. The one discussed in this subsection is defined by

$$B_g(x, y) := g(x) - g(y) - \nabla g(y)^\top (x - y) \quad x, y \in K \times intK, \tag{1.1.2}$$

where $g : K \subset R^n \to R$ is a real valued function. According to Facchinei and Pang [85], a Bregman function is defined as follows.

Definition 1.1.1 *Let K be a solid closed convex set in R^n. A function $g : K \to R$ is a Bregman function with zone K if*

(1) g is strictly convex and continuous on K;

(2) g is continuously differentiable on int$K$;

(3) for all $x \in K$ and all constants η, the set

$$\{y \in intK |\ B_g(x, y) \leq \eta\}$$

is bounded;

(4) if x_k is a sequence of points in int$K$ converging to x, then

$$\lim_{k \to +\infty} B_g(x, x_k) = 0.$$

If g *is a Bregman function, then B_g is a Bregman distance. If* g *also satisfies the following condition:*

(5) $\nabla g(intK) = R^n$;

then g is a full range Bregman function.

The standard and formal definition of the Bregman functions was given by Censor and Lent [53] and is based on the work of Bregman [42] on the generalization of the cyclic projection method for finding a point in the intersection of a finite number of closed convex sets. Clearly, $\frac{1}{2}\|x\|^2$ is a full range Bregman function with zone R^n. In what follows, we give two nontrivial examples of the Bregman functions.

Example 1.1.1 *Let $K = R^n_+$. Consider the function $g : K \to R$ defined by*

$$g(x) = \sum_{i=1}^{n} x_i \log x_i,$$

with the convention that $0 \log 0 = 0$.

Direct calculations yield

$$B_g(x, y) = \sum_{i=1}^{n} (x_i \log x_i - y_i \log y_i - (\log y_i + 1)(x_i - y_i)). \tag{1.1.3}$$

Clearly, g is a full range Bregman function. Formula (1.1.3) is known as the Kullback-Leibler relative entropy function, which is widely used in statistics.

Example 1.1.2 *Suppose that*

$$K = \{x \in R^n |\ a_i \leq x_i \leq b_i, i = 1, ..., n\}$$

where for each i*,* $-\infty < a_i < b_i < +\infty$ *and consider the function* $g : K \to R$

$$g(x) = \sum_{i=1}^{n} [(x_i - a_i) \log(x_i - a_i) + (b_i - x_i) \log(b_i - x_i)]$$

where, again, we make the convention that $0 \log 0 = 0$*. An easy calculation shows that*

$$B_g(x, y) = \sum_{i=1}^{n} [(x_i - a_i) \log \frac{(x_i - a_i)}{(y_i - a_i)} + (b_i - x_i) \log \frac{(b_i - x_i)}{(b_i - y_i)}].$$

In this case there is no difficulty in verifying that g is a full range Bregman function.

As observed in [55], there is no universal agreement on the definition of a Bregman function, and different authors usually give different, if clearly related, definitions to achieve their particular goals (see De Pierro and Iusem [72]; Censor and Zenios [54]; Eckstein [77] ; Güler [102]; and Censor et al. [52]). An in-depth study on the properties of the Bregman functions is given in [28].

The Bregman distance function may be used to solve the optimization problem (1.1.1) by generating a sequence x_k in such a way that x_{k+1} is a solution of the following subproblem

$$\min f(x) + \frac{\lambda_k}{2} B_g(x, x_k) \tag{1.1.4}$$

for $k = 0, 1, 2, ...$, where x_0 is a strictly feasible starting point. A convergence result completely identical to Theorem 1.1.1 can then be shown for this method, see (Chen and Teboulle [58]). Other references on these generalized proximal point methods and on the applications of non-quadratic (e.g., logarithmic) proximal point algorithms to decomposition include [15, 16, 14, 122, 193].

1.2 Multiobjective Optimization

Multiobjective optimization (also referred to as vector optimization) is part of mathematical programming dealing with decision problems characterized by multiple and conflicting objective functions to be optimized over a feasible set of decisions.

Multiobjective optimization has two main sources: economic equilibrium and welfare theories of Edgeworth [77] and Pareto [164], and the mathematical backgrounds of the ordered spaces of Cantor [48] and Hausdorff [110]. The game theory of Borel [39] and von Neumann [163] and the production theory of Koopmans [134] also contributed to this area. After the publication of Kuhn and Tucker's paper [139] on the necessary and sufficient conditions for optimality, and of Deubreu's paper [73] on valuation equilibrium and Pareto optimum, multiobjective optimization has been recognized as a useful mathematical discipline.

Let R^n and R^m be Euclidean vector spaces referred to as the decision space and the objective space respectively. Let F be a vector-valued objective function $F : R^n \to R^m$ composed of m real-valued objective functions, $F := (f_1, ..., f_m)$, where $f_i : R^n \to R$ for $i = 1, ..., m$ and $X \subseteq R^n$ a closed subset. A multiobjective optimization problem is given by

$$\min\{F(x)|\ x \in X \subseteq R^n\}. \tag{1.2.1}$$

Decisions are rarely made based on only one criterion; often, decisions are based on several conflicting criteria. A multiobjective optimization model provides the mathematical framework to deal with these situations. In the last fifty years, many theoretical, methodological, and applied studies have been undertaken on multiobjective programming. Several monographs have been published (e.g., Yu [216], Sawaragi, Nakayama and Tanino [177], Jahn [124], Luc [153], Chen, Huang and Yang [65], White [200] and Miettinen [160]). Studies on multiobjective optimization can be divided into three parts: theoretical study, numerical methods, and applications. In this thesis, we focus on the first two parts of multiobjective optimization studies.

1.2.1 Theoretical Studies

An ordinary mathematical programming (or optimization) problem includes only one objective function, and our aim is to find an element that minimizes this function. In other words, the objective space has one dimension; therefore, the ordering in the objective space is trivial in ordinary mathematical programming. In a multiobjective optimization problem, an element that minimizes an objective function does not generally minimize another objective functions. To characterize the optimality of multiobjective optimization problem, we need to define a partial order in the objective space R^m.

Let $C = R^m_+ \subset R^m$ and $C_1 = \{x \in R^m_+ |\ \|x\| = 1\}$. We define, for any $y_1, y_2 \in R^m$,

$$\begin{array}{ll} y_1 \leq_C y_2 & \text{if and only if } y_2 - y_1 \in C; \\ y_1 \nleq_{intC} y_2 & \text{if and only if } y_2 - y_1 \notin intC. \end{array}$$

The extended space of R^m is $\bar{R}^m = R^m \cup \{-\infty_C, +\infty_C\}$, where $-\infty_C$ is an imaginary point; each of the coordinates is $-\infty$ and the imaginary point $+\infty_C$ is analogously understood (with the conventions $\infty_C + \infty_C = \infty_C$, $\mu(+\infty_C) = +\infty_C$ for each positive number μ). The point $y \in R^m$ is a column vector and its transpose is denoted by $y^\top$. The inner product in R^m is denoted by $\langle \cdot, \cdot \rangle$.

Let $K \subset R^n$ be convex and $F : K \to R^m \cup \{+\infty_C\}$ be a vector-valued function. $x^* \in K$ is said to be a *Pareto optimal solution* of F on K if

$$(F(K) - F(x^*)) \cap (-C \backslash \{0\}) = \emptyset,$$

$x^* \in K$ is said to be a *weak Pareto optimal solution* of F on K if

$$(F(K) - F(x^*)) \cap (-intC) = \emptyset.$$

For a set $S^* \subset R^m$, we often use the following notations:

$$Min_C S^* = \{y \in S^* |\ (y - C \backslash \{0\}) \cap S^* = \emptyset\},$$

$$WMin_C S^* = \{y \in S^* |\ (y - intC) \cap S^* = \emptyset\}.$$

This definition provides globally Pareto optimal solutions. We can, of course, define locally Pareto optimal solutions as in the case of ordinary optimization. Any globally Pareto optimal solution is local, but the reverse is true only under appropriate convexity assumptions [51, 154].

A characterization of Pareto optimality and weak Pareto optimality using level sets is given in [81], along with a new solution concept of strict Pareto optimality. In a convex multiobjective programming problem, the set of weak Pareto optimal solutions is the union of the sets of Pareto optimal solutions of problems with parts of the original objective functions [80].

Moreover, approximate solutions to the multiobjective optimization problem are considered. For example, an ε-Pareto optimal solution is defined in [151] as a feasible solution $\bar{x}$ for which there exists no $x \in R^n$ such that

$$f_i(x) \leq f_i(\bar{x}) - \varepsilon_i, \quad \forall i \in [1, ..., m],$$

$$f_i(x) < f_i(\bar{x}) - \varepsilon_i, \quad for\ some\ i \in [1, ..., m],$$

where ε is a vector in R^m_+. Studies on approximate solutions can be found in [75, 147, 149, 201].

In any mathematical problem, the existence of a solution is the first question that should be answered. The existence of Pareto optimal solutions had been discussed in [33, 63, 109, 112, 123]. The existence of Pareto optimal or weak Pareto optimal solutions was also discussed in connection with scalar optimization in [63, 73]. The existence of Pareto optimal solutions with respect to general binary relations was stated in [187]. Research on the domination property was presented in [27, 155]. The concept of a dominator for a multiobjective maximization problem with quasiconvex functions was introduced in [49].

The most fundamental theorem in ordinary nonlinear programming is the Karush-Kuhn-Tucker theorem [58]. Optimality conditions (usually under some appropriate constraint qualifications) for multiobjective optimization problem have been investigated by a number of researchers [57, 59, 108, 148, 180, 198]. Second-order optimality conditions were also studied in [13, 28, 131, 197]. They require second-order approximation sets in the feasible region and new constraint qualifications. Optimality conditions for non-differentiable problems were also discussed by several authors [32, 120, 130, 169, 188]. [121] provided optimality conditions in terms of directional derivatives. If all the functions are locally Lipschitz, we obtained the optimality conditions in which the gradients are replaced with generalized gradients in the KKT-type conditions [67]. In [32] upper Dini derivatives were used to derive optimality conditions. [133] deals with a partially differentiable and partially convex case.

The nonemptiness and compactness of the solution set of a multiobjective optimization problem are important in both theory and methodology. They are an important condition to guarantee the convergence of some algorithms (e.g., proximal-type algorithms, Tikhonov-type regularization algorithms and so on). Recently, Deng [74, 76] obtained some necessary and sufficient conditions for the nonemptiness and compactness of solution sets of a convex vector optimization problem. Huang and Yang [116] also gave characterizations for the nonemptiness and compactness of the set of weak Pareto optimal solutions of a convex vector optimization problem with extended vector-valued functions in terms of the 0-coercivity of some scalar functions. Flores-Bazan [91] established the existence results for finite dimensional vector optimization problems based one the asymptotic description of the functions and sets. Several results have also been obtained on the characterizations of solution sets of convex multiobjective optimization problems [125, 150], and of nonconvex cases [92].

Stability and sensitivity analysis aims to analyze the qualitative and quantitative behavior of the Pareto optimal solutions according to the changes of the parameter values included in the original optimization problem. Stability analysis for set-valued mapping multiobjective optimization problems was investigated in [87, 162, 168, 190]. [190] also studied the stability with respect to the change in domination structure of the decision maker. [127] examined the stability of the compromise solution, not the entire Pareto optimal set. [189] investigated the continuous dependence of solutions on a parameter in a scalarization method. [8] dealt with the stability of not only the Pareto optimal solutions but also of the approximate Pareto optimal solutions. Well-posedness in vector optimization was discussed in [26, 114, 117]. Sensitivity in multiobjective optimization was analyzed by considering the derivatives of the perturbation mapping in [191]. A relationship between the derivative and the Lagrange multiplier vector was also established in [191]. Kuk et al. [140] provided sensitivity analysis on those perturbation mappings. [18] conducted a sensitivity analysis on the duality theory in convex multiobjective programming with right-hand side perturbation. [68] also dealt with sensitivity analysis in multiobjective optimization. [38] discussed differential sensitivity analysis along with second-order efficiency conditions.

Convexity plays a very important role in optimization theory. Various generalizations of convexity for multiobjective optimization problems have been made in the literature. Under pseudo/quasiconvexity, the duality theory for multiobjective programming problems has been studied in [22, 30, 199]. Interesting results in investigating the duality theory and the optimality conditions for multiobjective optimization problems under some generalized convexity assumptions were obtained by Yang et al.

[206, 207, 208, 209].

1.2.2 Numerical Methods

There are two general approaches to generate solution sets of multiobjective optimization problems, scalarization methods and nonscalarizing methods, which convert the multiobjective optimization problems into a scalar-valued objective optimization problem, a sequence of scalar-valued objective optimization problems, or another multiobjective optimization problem, respectively. Under some assumptions, the solution sets of these new programs yield solutions of the original problem. Scalarization methods explicitly employ a scalarizing function to accomplish the conversion, whereas nonscalarizing methods use other means.

A. Scalarization Methods

The traditional approach to solving multiobjective optimization problems is by scalarization. It involves formulating a multiobjective optimization problem through a real-valued scalarizing function t, which is typically a function of the objective functions of the multiobjective optimization problem, auxiliary scalar or vector variables, and/or scalar or vector parameters. Sometimes, the feasible set of the multiobjective optimization problem is additionally restricted by new constraint functions related to the objective functions of the multiobjective optimization problem and/or the new variables introduced. The most well-known scalarization techniques and list-related results in the generation of various classes of solutions of the multiobjective optimization problems can be found in [80, 220].

Many efficient methods have been proposed in investigating the numerical approach of multiobjective optimization problems by scalarization. However, there is one problem: the course of scalarization itself is very complicated in some cases. There is no doubt that it will increase the complexity of the algorithms from the computation viewpoint. Furthermore, many problems cannot be scalarized. Although we can scalarize such problems, the solution to these scalarized problems is another big problem. Thus, developing nonscalarizing approaches to investigate multiobjective optimizations is necessary in both theory and methodology.

B. Nonscalarizing Methods

In contrast to scalarizing approaches discussed in Part A, nonscalarizing methods do not explicitly use a scalarizing function but rather rely on other optimality conditions. The main idea of nonscalarizing methods is the extension of several well-known methods from the scalar case to the vector case. Although the advantages and drawbacks of these and other methods (i.e., scalar case) have been widely discussed in most nonlinear programming books, these extensions are not trivial works because of the different structure between multiobjective optimization and scalar-valued optimizations.

The steepest descent method for multiobjective optimization was dealt with in [90] and [104]. An extension of the projective gradient method to the case of convex constrained vector optimization can be found in [106]. Fliege [88] proposed an efficient interior-point method for approximating the solution set of convex multiobjective optimization problems. A combined homotopy interior-point method for a general multiobjective programming problem was dealt with in [186]. Recenlty, Miglierina et al. [161] constructed a gradient-like method based on suitable directions for box-constrained multi-objective optimizations without "a priori" scalarization. The Newton's Method for unconstrained multiobjective optimization was conducted by Fliege et al. in this working paper [89].

It is worth noting that Bonnel et al. [40] proposed the following vector-valued proximal point algorithm to investigate a convex vector optimization problem in Hilbert space:

(1). Choose $x_0 \in dom(F)$;

(2). Given x_k, if $x_k \in C - ARGMIN_w\{F(x) \mid x \in X\}$, then the algorithm stops, otherwise go to (3);

(3). If $x_k \notin C - ARGMIN_w\{F(x) \mid x \in X\}$, then compute x_{k+1} such that

$$x_{k+1} \in C - ARGMIN_w\{F(x) + \frac{\alpha_k}{2} \parallel x - x_k \parallel^2 e_k \mid x \in \theta_k\}$$

where $\theta_k := \{x \in X \mid F(x) \leq_C F(x_k)\}$, $e_k \in intC$ and $\parallel e_k \parallel= 1$, $C-ARGMIN_w\{F(x) \mid x \in X\}$ is the weak Pareto optimal solution set of F on X.

Convergence analysis for the vector-valued proximal point algorithm was conducted in [40]. They proved that the sequence generated by the vector-valued proximal point algorithm weakly converges to a weak Pareto optimal solution of the vector optimization problem under some mild conditions. The main results of [40] generalized the classical results of Rockafellar's [173] from the scalar case to the vector case.

In [50], Ceng and Yao proposed the following absolute approximate proximal point algorithm to investigate a convex vector optimization problem in Hilbert space:

(1). Choose $x_0 \in dom(F)$;

(2). Given x_k, if $x_k \in C - ARGMIN_w\{F(x) \mid x \in X\}$, then the algorithm stops, otherwise go to (3);

(3). If $x_k \notin C - ARGMIN_w\{F(x) \mid x \in X\}$, take any vector y_k such that

$$y_k \in C - ARGMIN_w\{F(x) + \frac{\alpha_k}{2} \parallel x - x_k - \omega_k \parallel^2 e_k \mid x \in \theta_k\}$$

and then compute the next iterate

$$x_{k+1} = (1 - \beta_k)y_k + \beta_k x_k$$

and go to step (2), where $\theta_k := \{x \in X \mid F(x) \leq_C F(x_k)\}$, $e_k \in intC$ and $\parallel e_k \parallel = 1$, $\beta_k \in (0, 1)$ and $\lim\limits_{k \to +\infty} \beta_k = 0$, ω_k is an error sequence.

Under some suitable conditions, they also proved that any sequence generated by the algorithm weakly converges to a weak Pareto optimal solution of the convex vector optimization problem.

1.3 Vector Variational Inequality

The concept of vector variational inequality in a finite dimensional Euclidean space was first introduced by Giannessi [97]. Let X_0 be a nonempty subset of R^n, and let $f_i : K \to R^n, i \in [1, ..., m]$ be vector-valued functions. Let

$$F := (f_1, ..., f_m), \quad F(x)(v) = (\langle f_1(x), (v)\rangle, ..., \langle f_m(x), (v)\rangle) \tag{1.3.1}$$

for every $x \in X_0$ and $v \in R^n$. The scalar product in a Euclidean space is denoted by $\langle ., . \rangle$.

In [97], the vector variational inequality, presented by the function F and the set X_0, is the following problem: finding a $\bar{x} \in X_0$ such that

$$F(\bar{x})(x - \bar{x}) \notin -R^m_+ \backslash \{0\} \tag{1.3.2}$$

for any $x \in X_0$.

The vector variational inequality problems have many important applications in multiobjective decision-making problems, network equilibrium problems, traffic equilibrium problems, and so on. Due to these significant applications, the study of vector variational inequalities has attracted wide attention. In the last twenty years of development, existence results of solutions, duality theorems and topological properties of solution sets of several kinds of vector variational inequalities have been derived. Chen and Yang [60] investigated general vector variational inequality problems and vector complementary problems in infinite dimensional spaces. Chen [61] considered the vector variational inequality problems with a variable ordering structure. A complete review of the main results of the vector variational inequalities can be found in the monograph [65].

The concept of a gap function is well-known both in the context of convex optimization and variational inequality. The minimization of gap functions is a powerful tool for solving variational inequality. Chen, Huang and Yang [65] generalized the gap functions for variational inequalities to vector variational inequalities. The convexity and differentiability of gap functions are also investigated in the monograph [65].

Vector variational inequalities and their generalizations have been also used as a tool to solve vector (multiobjective) optimization problems. Several authors have discussed relations between vector variational inequalities and vector optimization problems under some convexity or generalized convexity assumptions. Lee [143] showed that a necessary condition for a point to be a weak Pareto optimal solution of a vector optimization problem for differentiable functions is that the point be a solution of a vector variational inequality. Giannessi [98] considered another type vector variational inequality, which is called the Minty type vector variational inequality for gradients: finding an $\bar{x} \in X_0$ such that for any $x \in X_0$, we have

$$(\nabla f_1(x)^\top (x - \bar{x}), ..., \nabla f_m(x)^\top (x - \bar{x})) \notin -R^m_+ \backslash \{0\}. \tag{1.3.3}$$

Giannessi [98] provided the equivalence between the Pareto optimal solutions of a differentiable convex vector optimization problem and the solutions of a Minty type vector variational inequality for gradients which is a vector version of the classical Minty variational inequality for gradients. Moreover, Giannessi [98] also proved the equivalence between the solutions of weak Minty- and Stampacchia- type vector variational inequalities for gradients and weak Pareto optimal solutions of a differentiable convex vector optimization problem. Lee [142] studied the equivalence between nondifferentiable convex vector optimization problems and Minty type vector variational inequality and Stampacchia type vector variational inequality. Yang [212] established the equivalence between a vector variational-like inequality with a multiobjective programming problem for generalized invex functions. The vector variational-like inequality approach was used in [144, 145] to prove some existence theorems for the generalized Pareto optimal solutions of nondifferentiable invex vector optimization problems. The results in [144, 145] are generalizations of the existence results established in [63, 62] for differentiable and convex vector optimization problems and in [132] for differentiable preinvex vector optimization problems. Ansari and Yao [5] proved the equivalence among the Minty vector variational-like inequality, Stampacchia vector variational-like inequality, and a nondifferentiable and nonconvex vector optimization problem. [5] also established an existence theorem for generalized weakly efficient solutions of nondifferentiable nonconvex vector optimization problems by using a fixed point theorem. Yang [211] studied the inverse vector variational inequality problems and their relations with some vector optimization problems. Yang [213] also gave the equivalence between the solutions of a Stampacchia vector variational inequality for gradients and the efficient solutions of a linear fractional vector optimization problem of which the numerators of the objective functions are linear and the denominators of the objective functions are the same linear functions. Several existence results of solutions for vector equilibrium problems can be found in [31, 56, 107, 136] and the references cited therein.

However to the best of our knowledge, there is no numerical method designed that has yet to solve vector variational inequality problems, even no conceptual ones. Motivated by this situations, in this thesis, we attempt to construct a matrix-valued proximal point algorithm to solve a weak vector variational inequality problem (which is an extension of the classical proximal point algorithm proposed by Rockafellar in [173]).

We end this chapter by mentioning that the thesis is based on the following papers written by the author and his colleagues during the period of stay in the Department of Applied Mathematics, The Hong Kong Polytechnic University as a graduate student

[1]. Z. Chen and K. Q. Zhao, A proximal-type method for convex vector optimization problem in Banach spaces, *Numerical Functional Analysis and Optimization*, 30 (2009) 70-81.

[2]. Z. Chen, H. Q. Huang and K. Q. Zhao, Approximate generalized proximal-type method for convex vector optimization problem in Banach spaces, *Computers and Mathematics with Applications*, 57 (2009) 1196-1203.

[3]. Z. Chen, C. H. Xiang, K. Q. Zhao and X. W. Liu, Convergence analysis of Tikhonov-type regularization algorithms for multiobjective optimization problems, *Applied Mathematics and Computation* 211 (2009) 167-172.

[4]. Z. Chen, X. X. Huang and X. Q. Yang, Generalized proximal point algorithms for multiobjective optimization problems, to appear in: *Applicable Analysis*.

[5]. Z. Chen, A Proximal-type Method in Vector Variational Inequalities, in preparation.

[6]. Z. Chen, Asymptotic analysis for parametric multiobjective optimization problems, in preparation.

Chapter 2

Preliminaries

In this chapter, we introduce some definitions, notations and basic results that will be used later in this thesis.

Let $C \subset R^m$ be a closed and convex cone with $intC \neq \emptyset$ and $C_1 = \{x \in C| \, \|x\| = 1\}$. We define, for any $y_1, y_2 \in R^m$,

$$\begin{aligned} y_1 \leq_C y_2 \qquad & \text{if and only if } y_2 - y_1 \in C; \\ y_1 \nleq_{intC} y_2 \qquad & \text{if and only if } y_2 - y_1 \notin intC. \end{aligned}$$

The extended space of R^m is $\bar{R}^m = R^m \cup \{-\infty_C, +\infty_C\}$, where $-\infty_C$ is an imaginary point, each of the coordinates is $-\infty$ and the imaginary point $+\infty_C$ is analogously understood (with the conventions $\infty_C + \infty_C = \infty_C$, $\mu(+\infty_C) = +\infty_C$ for each positive number μ). The point $y \in R^m$ is a column vector and its transpose is denote by $y^\top$. The inner product in R^m is denoted by $\langle ., . \rangle$

2.1 Multiobjective Optimization

In this thesis, we consider the following multiobjective optimization problem:

$$C - MIN\{F(x)| \; x \in R^n\} \quad (MOP)$$

where $F : R^n \to R^m \bigcup \{+\infty_C\}$ and denote by $dom(F) = \{x \in R^n | F(x) \neq +\infty_C\}$ the effective domain of F.

It is known that the following constrained multiobjective optimization problem

$$C - MIN\{F_0(x)|\ x \in X_0\} \quad (CMOP)$$

where X_0 is a nonempty closed convex subset of R^n and $F_0 : X_0 \to R^m$ is a vector-valued function, is equivalent to (MOP), where

$$F(x) = \begin{cases} F_0(x), & x \in X_0, \\ +\infty_C, & \mathrm{x} \notin \mathrm{X}_0, \end{cases}$$

in the sense that they have the same sets of Pareto optimal solutions and the same sets of weak Pareto optimal solutions.

Definition 2.1.1 *[65] Let $K \subset R^n$ be convex. A map $F : K \to R^m \cup \{+\infty_C\}$ is said to be* C-convex *if*

$$F((1-\lambda)x + \lambda y) \leq_C (1-\lambda)F(x) + \lambda F(y)$$

for any $x, y \in K$ and $\lambda \in [0,1]$. F is said to be strictly C-convex *if*

$$F((1-\lambda)x + \lambda y) \leq_{intC} (1-\lambda)F(x) + \lambda F(y)$$

for any $x, y \in K$ with $x \neq y$ and $\lambda \in (0,1)$.

Definition 2.1.2 *[221] A map $F : K \subset X \to R^m \cup \{+\infty_C\}$ is said to be* C-lsc *at $x_0 \in K$ if, for any neighborhood V of $F(x_0)$ in R^m, there exists a neighborhood U of x_0 in R^n such that $F(U \cap K) \subseteq V + C$. The map $F : K \subset X \to R^m \cup \{+\infty_C\}$ is said to be* C-*lsc on K if it is* C-*lsc at every point $x_0 \in K$.*

Remark 2.1.1 *[91] The R^m_+-lower semicontinuity of $F = (F_1, ..., F_m)$ is equivalent to the (usual) lower semicontinuity of each F_i, $i = 1, ..., m$.*

Definition 2.1.3 *[65] Let $K \subset R^n$ be convex and $F : K \to R^m \cup \{+\infty_C\}$ be a vector-valued function. $x^* \in K$ is said to be a* Pareto optimal solution *of F on* K *if*

$$(F(K) - F(x^*)) \cap (-C \backslash \{0\}) = \emptyset,$$

$x^* \in K$ is said to be a weak Pareto optimal solution *of F on K if*

$$(F(K) - F(x^*)) \cap (-intC) = \emptyset.$$

Lemma 2.1.1 *[91] Let $K \subset R^n$ be a closed set, and suppose that $W \subset R^m$ is a closed set such that $W + C \subseteq W$. Assume that $F : K \to R^m \cup \{+\infty_C\}$ is* C-*lsc. The set $P = \{x \in K |\ F(x) - \lambda \in -W\}$ is then closed for all $\lambda \in R^m$.*

Definition 2.1.4 *[153] A cone $C \subseteq R^m$ is called Daniell if any decreasing sequence of R^m having a lower bound converges to its infimum. For example, the cone $C = R^m_+$ has the Daniell property.*

Definition 2.1.5 *[177] A set $K \subset R^m$ is said to have the* domination property *with respect to C, if there exists $k \in R^m$ such that $K \subseteq k + C$.*

Lemma 2.1.2 *[40] If $K \subset R^n$ is a convex set and $F : K \to R^m \cup \{+\infty_C\}$ is a proper C-convex mapping, then*

$$C - ARGMIN_w\{F(x) \mid x \in K\} = \bigcup_{z \in C_1} argmin\{\langle F(x), z\rangle \mid x \in K\}$$

where $C - ARGMIN_w\{F(x) \mid x \in K\}$ is the weak Pareto optimal solution set of F on K.

2.2 Vector Variational Inequality

Let X_0 be a nonempty subset of R^n and let $T_i : X_0 \to R^n, i \in [1, ..., m]$ be vector-valued functions. Let

$$T := (T_1, ..., T_m), \quad T(x) = (T_1(x), ..., T_m(x)), \quad T(x)(v) = (\langle T_1(x), v\rangle, ..., \langle T_m(x), v\rangle)^\top$$

for every $x \in X_0$ and $v \in R^n$. For any $\lambda \in C_1$, a mapping $\lambda(T) : X_0 \to R^n$ is defined by

$$\lambda(T)(x) = T(x)\lambda, \quad x \in X_0. \tag{2.2.1}$$

Definition 2.2.1 *[84] Let $F : X_0 \to R^n$ be a mapping. A variational inequality (VI in short) is a problem of finding $x^* \in X_0$ such that*

$$\text{(VI)} \qquad \langle F(x^*), (x - x^*)\rangle \geq 0, \quad \forall\ x \in X_0,$$

where x^* is the solution of problem (VI).

Definition 2.2.2 *[101] A vector variational inequality (VVI in short) is a problem of finding* $x^* \in X_0$ *such that*

$$\text{(VVI)} \qquad T(x^*)^\top (x - x^*) \not\leq_{C\setminus\{0\}} 0, \quad \forall\, x \in X_0,$$

where x^* *is the solution of problem (VVI).*

Definition 2.2.3 *[65] A weak variational inequality (WVVI in short) is a problem of finding* $x^* \in X_0$ *such that*

$$\text{(WVVI)} \qquad T(x^*)^\top (x - x^*) \not\leq_{intC} 0, \quad \forall x \in X_0,$$

where x^* *is called a solution of problem (WVVI). Denote by* X^* *the solution set of problem (WVVI). Let* $\lambda \in C_1$*, consider the corresponding scalar-valued variational inequality problem of finding* $x^* \in X_0$ *such that:*

$$(VIP_\lambda) \qquad \langle \lambda(T)(x^*), x - x^* \rangle \geq 0, \quad \forall x \in X_0.$$

Denote by X^*_λ *be the solution set of* (VIP_λ)*.*

It is worth noticing that the binary relation $\not\leq_{intC}$ is closed in the sense that if $x_k \to x^*$ as $k \to \infty$, $x_k \not\leq_{intC} 0$, then we have $x^* \not\leq_{intC} 0$ because of the closeness of the set $W =: R^m \setminus (-intC)$.

Definition 2.2.4 *[84] Let* $X_0 \subset R^n$ *be nonempty, closed and convex, and* $F : X_0 \to R^n$ *be a single-valued mapping.*

(i) F *is said to be* monotone *on* X_0 *if, for any* $x_1, x_2 \in X_0$*, there holds*

$$\langle F(x_1) - F(x_2), x_1 - x_2 \rangle \geq 0.$$

(ii) F *is said to be* pseudomonotone X_0 *if, for any* $x_1, x_2 \in X_0$ *, there holds*

$$\langle F(x_2), x_1 - x_2 \rangle \geq 0 \Rightarrow \langle F(x_1), x_1 - x_2 \rangle \geq 0.$$

Clearly, a monotone map is pseudomonotone.

We now give the definitions of C-monotonicity of a matrix-valued map.

Definition 2.2.5 *[65] Let $X_0 \subset R^n$ be nonempty, closed and convex. $T : X_0 \to R^{n\times m}$ is a mapping, which is said to be C-*monotone *on X_0 if, for any $x_1, x_2 \in X_0$, there holds*

$$(T(x_1) - T(x_2))^\top (x_1 - x_2) \geq_C 0.$$

Proposition 2.2.1 *[118] Let X_0 and T be defined as in Definition 2.2.4., we have the following statements:*

(i) T is C-monotone if and only if, for any $\lambda \in C_1$, the mapping $\lambda(T) : X_0 \to R^n$ defined by (2.2.1) is monotone.

(ii) if T is C-monotone, then for any $\lambda \in C_1$, $\lambda(T) : X_0 \to R^n$ is pseudomonotone.

Definition 2.2.6 *[85] Let a set-valued map $G : X_0 \subset R^n \rightrightarrows R^n$ be given, it is said to be* monotone *if*

$$\langle z - \bar{z}, w - \bar{w}\rangle \geq 0$$

for all z *and $\bar{z}$ in X_0, all* w *in $G(z)$ and $\bar{w}$ in $G(\bar{z})$. It is said to be* maximal monotone *if, in addition, the graph*

$$gph(G) = \{(z, w) \in R^n \times R^n \mid w \in G(z)\}$$

is not properly contained in the graph of any other monotone operator from R^n to R^n.

2.3 Asymptotical Analysis

Definition 2.3.1 *[172] Given $b \in R^m$ and an $m \times n$ real matrix* T*, the set*

$$M = \{x \in R^n | Tx = b\}$$

is an affine set *in R^n. Therefore, for any $S \subset R^n$ there exists a unique smallest affine set containing S, which is called the* affine hull *of S and is denoted by $affS$.*

Definition 2.3.2 *[172] The* relative interior *of a convex set* S *in R^n, which we denoted by riS, is defined by*

$$riS = \{x \in affS | \exists \varepsilon > 0, (x + \varepsilon B) \cap (affS) \subset S\},$$

where B is the Euclidean unit ball *in R^n.*

Definition 2.3.3 *[172] Let $S \subset R^n$ be any set. The* closure *clS and interior $intS$ of S can be defined by the formula*

$$clS = \cap\{S + \varepsilon B | \varepsilon > 0\},$$

$$intC = \{x | \exists \varepsilon > 0, x + \varepsilon B \subset S\}.$$

Definition 2.3.4 *[17] Let* K *be a nonempty set in R^n. Then the* asymptotic cone *of the set K, denoted by K^∞, is the set of all vectors $d \in R^n$ that are limits in the direction of the sequence $\{x_k\} \subset K$, namely*

$$K^\infty = \{d \in R^n | \ \exists t_k \to +\infty, \ and \ x_k \in K, \lim_{k \to +\infty} \frac{x_k}{t_k} = d\}. \tag{2.3.1}$$

In the case where K is convex and closed, then, for any $x_0 \in K$,

$$K^\infty = \{d \in R^n | \ x_0 + td \in K, \forall \ t > 0\}. \tag{2.3.2}$$

Lemma 2.3.1 *[175] A set $K \subset R^n$ is bounded if and only if its asymptotic cone is just the zero cone: $K^\infty = \{0\}$.*

Definition 2.3.5 *[17] For any given function $f : R^n \to R \cup \{+\infty\}$, the* asymptotic function *of f is defined as the function f^∞ such that $epif^\infty = (epif)^\infty$, where $epif = \{(x, t) \in R^n \times R | f(x) \leq t\}$ is the epigraph of f. Consequently, we can give the analytic representation of the asymptotic function f^∞:*

$$f^\infty(d) = \inf\{ \ \liminf_{k \to +\infty} \frac{f(t_k d_k)}{t_k} : t_k \to +\infty, d_k \to d\}. \tag{2.3.3}$$

When f is a proper convex and lower semi-continuous (lsc in short) function, we have

$$f^\infty(d) = \sup\{f(x + d) - f(x) | \ x \in dom \ f\} \tag{2.3.4}$$

or equivalently

$$f^\infty(d) = \lim_{t \to +\infty} \frac{f(x + td) - f(x)}{t} = \sup_{t > 0} \frac{f(x + td) - f(x)}{t}, \quad \forall \ d \in dom \ f \tag{2.3.5}$$

and

$$f^\infty(d) = \lim_{t \to 0^+} t f(t^{-1} d), \quad \forall \ d \in dom \ f. \tag{2.3.6}$$

Let $K \subset R^n$. Define the scalar-valued indicator function δ_K as

$$\delta_K(x) = \begin{cases} 0, & x \in K, \\ +\infty, & \text{otherwise}, \end{cases}$$

We have that $\delta_K^\infty = \delta_{K^\infty}$, where $K \subset R^n$ is a nonempty set.

Definition 2.3.6 *[175] The function $f : R^n \to R \cup \{+\infty\}$ is said to be* coercive *if its asymptotic function $f^\infty(d) > 0$, for all $d \neq 0 \in R^n$ and it is said to be* counter-coercive *if its asymptotic function $f^\infty(d) = -\infty$, for some $d \neq 0 \in R^n$.*

Proposition 2.3.1 *[175] Let f_1 and f_2 be lsc and proper on R^n, and suppose that neither is counter-coercive. Then*

$$(f_1 + f_2)^\infty \geq f_1^\infty + f_2^\infty,$$

where the inequality becomes an equality when both functions are convex and $dom f_1 \cap dom f_2 \neq \emptyset$.

Proposition 2.3.2 *[17] Let $f : R^n \to R \cup \{+\infty\}$ be proper convex and lsc. Then the following three statements are equivalent:*

(a) *f is coercive;*

(b) *The optimal set $\{x \in R^n | \ f(x) = \inf f\}$ is nonempty and compact;*

(c) $\liminf\limits_{\|x\| \to +\infty} \frac{f(x)}{\|x\|} > 0.$

Definition 2.3.7 *[175] A set-valued mapping $S : R^n \rightrightarrows R^m$ is said to be* outer semi-continuous *(osc in short) at $\bar{x}$ if*

$$\limsup_{x \to \bar{x}} S(x) \subset S(\bar{x}),$$

where

$$\limsup_{x \to \bar{x}} S(x) := \bigcup_{\{x_k\} \in A} \limsup_{k \to +\infty} S(x_k)$$

$$= \bigcup_{x_k \in A} \{u | \ \exists \ x_k \to \bar{x}, \exists \ u_k \to \bar{u}, with \ u_k \in S(x_k)\},$$

where $A = \{\{x_k\} \subset R^n : x_k \to \bar{x} \ as \ k \to +\infty\}$.

Lemma 2.3.2 *[175] Let $S : R^n \rightrightarrows R^m$ be a set-valued mapping. Then S is outer semicontinuous if and only if gph$S$ is closed in $R^n \times R^n$; moreover, S is outer semicontinuous if and only if S^{-1} is outer semicontinuous.*

Proposition 2.3.3 *[175] For any extended-real-valued functions f_1 and f_2, one has that*

$$\liminf_{x \to \bar{x}} [f_1(x) + f_2(x)] \geq \liminf_{x \to \bar{x}} f_1(x) + \liminf_{x \to \bar{x}} f_2(x) \tag{2.3.7}$$

if the sum on the right hand side is not $\infty - \infty$. On the other hand, one always has

$$\liminf_{x \to \bar{x}} \lambda f(x) = \lambda \liminf_{x \to \bar{x}} f(x), \tag{2.3.8}$$

when $\lambda \geq 0$.

Lemma 2.3.3 *[74] If $K \subset R^n$ is a nonempty, closed and convex set and each component F_i of F is convex, then, the following statements are equivalent:*

(a) $C - ARGMIN_w\{F(x) \mid x \in K\}$ is nonempty and compact;

(b) $argmin\{F_i(x) \mid x \in K\}$ is nonempty and compact for every $i \in [1, ..., m]$;

(c) $K^\infty \cap (\cup_{i=1}^m \{d \in R^n \mid F_i^\infty(d) \leq 0\}) = \{0\}$.

Chapter 3

Proximal-type Methods for A Convex Vector Optimization Problem in Banach Spaces

3.1 Introduction

An important motivation for making analysis about the convergence properties of various proximal point algorithms is related to *The Mesh Independence Principle* [3, 111]. The mesh independence principle replies on infinite dimensional convergence results for predicting the convergence properties of the discrete finite dimensional method. It also provides a theoretical foundation to justify the refinement strategies and help design the refinement process. Many practice problems in economics and engineering are modeled in infinite dimensional spaces, such as optimal control problems, shape optimization problems, and problems of minimal area surface with obstacles, among many others. In many shape optimization problems, the function space is only a Banach and not a Hilbert space. Thus, analyzing the convergence properties of the algorithms in Banach spaces is important and necessary.

A vector optimization problem (VOP in short) is a variant of multiobjective optimization, which is constructed in infinite dimensional spaces.

In this chapter, we consider a convex vector optimization problem of finding weak Pareto optimal solutions for an extended vector-valued map from a uniformly convex and uniformly smooth Banach space to a real Banach space, with the latter being ordered by a closed, convex, and pointed cone with a nonempty interior. We propose both exact and inexact vector-valued proximal point algorithms based on a Lyapunov function. In both cases, we prove that any sequence generated by the algorithms weakly converges to a weak Pareto optimal solution of the vector optimization problem.

This chapter is outlined as follows.

In section 3.2.2, we present the basic definitions, notations and some preliminary results. In section 3.2.3, we propose an exact vector-valued proximal point algorithm based on a Lyapunov functional, carry out convergence analysis on this algorithm, and prove that any sequence generated by the algorithms weakly converges to a weak Pareto optimal solution of the vector optimization problem. In section 3.2.4, we propose an inexact vector-valued proximal point algorithm based on a Lyapunov functional and carry out convergence analysis on this algorithm, in which the subproblems are solved only approximately within a given tolerance. In section 3.2.5, we draw some conclusions and provide some remarks.

3.2 Preliminaries

In this section, we present some basic definitions and propositions for the proof of our main results. A Banach space X is said to be strictly convex if $\| \frac{(x+y)}{2} \| < 1$ for all $x, y \in X$ with $\| x \| = \| y \| = 1$ and $x \neq y$. It is said to be uniformly convex if

$$\lim_{n\to\infty} \| x_n - y_n \| = 0$$

for any two sequences $\{x_n\}$ and $\{y_n\}$ in X such that $\| x_n \| = \| y_n \| = 1$ and

$$\lim_{n\to\infty} \| \frac{x_n + y_n}{2} \| = 1.$$

A uniformly convex Banach space is reflexive and strictly convex. A Banach space X is said to be smooth if the limit

$$\lim_{t\to 0} \frac{\| x + ty \| - \| x \|}{t} \tag{3.2.1}$$

exists for all $x, y \in U$, where $U = \{x \in X : \| x \| = 1\}$. It is said to be uniformly smooth if the limit (3.2.1) is attained uniformly for $x, y \in U$. We know that the space $L^p (1 < p < \infty)$, l^p and Sobolev space W^p_m are uniformly convex and uniformly smooth Banach spaces.

In this chapter, let X be a uniformly convex and uniformly smooth Banach space with norm $\| \,.\, \|_X$, denote by X^* with norm $\| \,.\, \|_{X^*}$ the dual space of X and Y be a real Banach space ordered by a pointed, closed and convex cone C with nonempty interior $intC$, which defines a partial order $\leq_C$ in Y, *i. e.*, for any $y_1, y_2 \in Y$,

$$y_1 \leq_C y_2 \qquad \text{if and only if } y_2 - y_1 \in C;$$

and a binary relation

$$y_1 \not\leq_{intC} y_2 \qquad \text{if and only if } y_2 - y_1 \notin intC.$$

The extended space $\bar{Y} = Y \cup \{-\infty_C, +\infty_C\}$, where $+\infty_C$ and $-\infty_C$ are two distinct elements not belonging to Y. We denote Y^* as the dual space of Y and C^* as the positive polar cone of C, *i.e.*,

$$C^* = \{z \in Y^* : \langle z, y \rangle \geq 0, \forall y \in C\}.$$

Denote by $C^{*+} = \{z \in Y^* : \langle y, z \rangle > 0\}$ for all $y \in C \backslash \{0\}$.

Our analysis holds without requiring reflexivity of the Banach space Y. This detail is not negligible, because we do need the cone C to have a nonempty interior. The prototypical infinite dimensional Banach spaces are the L^p space $(1 \leq p < \infty)$, and their most relevant cones are the so-called positive cones consisting of all p-integrable functions, which are nonnegative almost everywhere. As well known, these cones have an empty interior, except for the case of L^∞, which happens to be nonreflexive. Thus, our analysis covers at least one meaningful example, where the order is induced by a cone in an infinite dimensional space.

We state ([119], Lemma 1.1) as the following lemma.

Lemma 3.2.1 *Let $e \in intC$ be fixed and $C^{*0} = \{z \in C^* | \langle z, e \rangle = 1\}$, then C^{*0} is a weak*-compact subset of C^*.*

The normalized duality mapping $J : X \to 2^{X^*}$ is defined by

$$Jx = \{v \in X^* : \langle x, v\rangle = \|x\|^2 = \|v\|^2\}$$

where $\langle ., .\rangle$ stands for the usual dual product in X. We know that the normalized duality mapping is equal to the identity operator of Hilbert space and has the following important properties [129]:

A. $\| x \|^2 - \| y \|^2 \geq 2\langle x - y, j\rangle$, for all $x, y \in X$ and $j \in Jy$;

B. if X is smooth, then J is single valued;

C. if X is smooth, then J is norm-to-weak* continuous;

D. if X is uniformly smooth, then J is uniformly norm-to-norm continuous on each bounded subset of X.

Definition 3.2.1 *[159] Let X be a normed space. Define a function $\delta_X(\epsilon) : [0, 2] \to [0, 1]$ by the formula*

$$\delta_X(\epsilon) = \inf\{1 - \|\frac{x + y}{2}\| : x, y \in U, \|x - y\| \geq \epsilon\}$$

if $X \neq \{0\}$, and if $X = \{0\}$ by the formula

$$\delta_X(\epsilon) = \begin{cases} 0, & \textit{if } \epsilon = 0; \\ 1, & \textit{if } 0 < \epsilon \leq 2. \end{cases}$$

Then δ_X is the modulus of convexity *of X.*

Definition 3.2.2 *[159] Let X be a normed space. Define a function $\rho_X(\tau) : (0, \infty) \to [0, \infty)$ by the formula*

$$\rho_X(\tau) = \sup\{\frac{1}{2}(\|x + \tau y\| + \|x - \tau y\|) - 1 : x, y \in U\}$$

if $X \neq \{0\}$, and if $X = \{0\}$ by the formula

$$\rho_X(\tau) = \begin{cases} 0, & \textit{if } 0 < \tau < 1; \\ \tau - 1, & \textit{if } \tau \geq 1. \end{cases}$$

Then ρ_X *is the* modulus of smoothness *of* X*. Let* $h_X(\tau) = \rho_X(\tau)/\tau$*, if the space* X *is uniformly smooth. Then,*

$$\lim_{\tau \to 0^+} h_X(\tau) = 0.$$

Lemma 3.2.2 *[8] Let* X *be a real Banach space. Let* $\delta_X(\epsilon)$ *be the modulus of convexity and* $\rho_X(\tau)$ *be the modulus of smoothness of the Banach space* X*. If* $x, y \in X$ *are such that* $\| x \| \leq M$ *and* $\| y \| \leq M$*, then we have*

$$\langle Jx - Jy, x - y \rangle \geq (2L)^{-1} M^2 \delta_X(\frac{\| x - y \|}{2M}) \tag{3.2.2}$$

$$\langle Jx - Jy, x - y \rangle \geq (2L)^{-1} M^2 \delta_{X^*}(\frac{\| Jx - Jy \|}{2M}) \tag{3.2.3}$$

and

$$\| Jx - Jy \| \leq 8M h_X(16L \| x - y \| /M) \tag{3.2.4}$$

where $L > 0$ *is the constant in* $Figiel's$ *inequalities [91].*

Definition 3.2.3 *[40] The set* $(F(x_0) - C) \cap F(X)$ *is said to be* $C-$*complete, if for every sequence* $\{\alpha_n\} \subset X$*, with* $\alpha_0 = x_0$ *such that*

$$F(\alpha_{n+1}) \leq_C F(\alpha_n)$$

for all $n \in N$*, there exists* $\alpha \in X$ *such that* $F(\alpha) \leq_C F(\alpha_n)$ *for all* $n \in N$*.*

Definition 3.2.4 *[40] A map* $F : X \to Y \cup \{+\infty_C\}$ *is said to be positively lower semicontinuous if for every* $z \in C^*$*, the scalar-valued function* $x \to \langle F(x), z \rangle$ *is lower semicontinuous.*

Let's consider following vector optimization problem:

$$C - Min_w\{F(x) | x \in X\} \quad (VOP)$$

where X is a uniformly convex and uniformly smooth Banach space and Y is a real Banach space, $F : X \to Y \cup \{+\infty_C\}$ is *C-convex* and denote by

$$dom(F) = \{x \in X | F(x) \neq +\infty_C\}$$

the effective domain of F.

Notice that any constrained vector optimization problem

$$C - Min_w\{F_0(x)|x \in X_0\} \quad (CVOP)$$

where $F_0 : X_0 \to Y$ is $C - convex$ and X_0 is a nonempty closed convex subset of X, is equivalent to the extended-valued (VOP) with

$$F(x) = \begin{cases} F_0(x), & \text{if } x \in X_0; \\ +\infty, & \text{if } x \notin X_0. \end{cases}$$

in the sense that they have the same set of weak Pareto optimal solutions.

Lemma 3.2.3 *[8] Let $s > 0$ and X be a Banach space. Then X is uniformly convex if and only if there exists a continuous, strictly increasing and convex function $g : [0, \infty) \to [0, \infty)$, $g(0) = 0$ such that*

$$\| x + y \|^2 \geq \| x \|^2 + 2\langle y, j\rangle + g(\| y \|)$$

for all $x, y \in \{z \in X : \| z \| \leq s\}$ and $j \in Jx$.

We now introduce a real-valued function in the Banach space, which has similar properties with $\| x - y \|^2$, is a Hilbert space.

Definition 3.2.5 *Let* X *be a uniformly convex and uniformly smooth Banach space. The* Lyapunov functional $L : X \times X \to R^+$ *is defined by*

$$L(x, y) = \| x \|^2 - 2\langle x, Jy\rangle + \| y \|^2$$

for $x, y \in X$. It is easy to see that

$$(\| x \| + \| y \|)^2 \geq L(x, y) \geq (\| x \| - \| y \|)^2 \tag{3.2.5}$$

for any $x, y \in X$.

Proposition 3.2.1 *[8] Let* X *be a uniformly convex and uniformly smooth Banach space, $\{y_n\}$ and $\{z_n\}$ be two sequences of X. If $L(y_n, z_n) \to 0$ and either $\{y_n\}$ or $\{z_n\}$ is bounded, then we have that $\|y_n - z_n\| \to 0$.*

Proposition 3.2.2 *[8] Let* X *be a uniformly convex and uniformly smooth Banach space and X_0 be a nonempty, closed and convex subset of X. Then there exists a unique element $x^* \in X_0$ such that*

$$L(x^*, x) = \inf\{L(z, x) : z \in X_0\}. \tag{3.2.6}$$

For every nonempty, closed and convex subset X_0 of a uniformly convex and uniformly smooth Banach space X, we can define a map P_{X_0} of X onto X_0 by $P_{X_0}x = x^*$, where x^* is defined by (3.2.6). The map P_{X_0} is coincident with the metric projection when X is a Hilbert space.

Lemma 3.2.4 *[171] Let X be a reflexive Banach space with X^* being strictly convex. f is a proper lower semicontinuous convex function and J is a normalized duality mapping of X into X^*. Then, for all $\lambda > 0$ the mapping $T = \partial f + \lambda J$ maps X onto X^*.*

3.3 Exact Vector-valued Proximal-type Method in Banach Spaces

Lemma 3.3.1 *If $S \subset X$ is a convex set and $F : S \to Y \cup \{+\infty_C\}$ is a proper and $C-$convex map, then*

$$C - ARGMIN_w\{F(x) \mid x \in S\} = \bigcup_{z \in C^{*0}} argmin\{\langle F(x), z\rangle \mid x \in S\}$$

where $C - ARMIN_w\{F(x) \mid x \in S\}$ is the weak Pareto optimal solution set of F.

This follows immediately from Theorem 2.1 in [40].

Let $\{e_k\} \subset intC$ be such that $\| e_k \| = 1$ and $\varepsilon > 0$. For any given x_k, let $\theta_k = \{x \in X \mid F(x) \leq_C F(x_k)\}$. Here we propose the following proximal-type method based on Lyapunov functional (PML, in short):

Step (1) : Take $x_0 \in domF$;

Step (2) : Given any x_k, if $x_k \in C - ARGMIN_w\{F(x) \mid x \in X\}$, then the algorithm stops, otherwise goes to step (3);

Step (3): If $x_k \notin C - ARGMIN_w\{F(x) \mid x \in X\}$, then compute x_{k+1} such that

$$x_{k+1} \in C - ARGMIN_w\{F(x) + \frac{\varepsilon_k}{2} L(x, x_k)e_k \mid x \in \theta_k\} \tag{3.3.1}$$

where $\varepsilon_k \in (0, \varepsilon]$. Go to step (2).

We now present the main results of this section.

Theorem 3.3.1 *Let $F : X \to Y \cup \{+\infty_C\}$ be a proper $C-$convex and positively lower semicontinuous mapping. Then any sequence $\{x_k\}$ generated by the method (PML) is well-defined.*

Proof. Let $x_0 \in domF$ be an initial point and assume that the algorithm has reached step k. We show that the next iterative x_{k+1} does exist. Take any $z \in C^{*0}$ and define a new function $\phi_k(x) : X \to R \cup \{+\infty\}$ as follows:

$$\phi_k(x) = \langle F(x), z\rangle + \delta_{\theta_k}(x) + \frac{\varepsilon_k}{2} L(x, x_k)\langle e_k, z\rangle \tag{3.3.2}$$

where

$$\delta_{\theta_k}(x) = \begin{cases} 0, & \text{if } x \in \theta_k; \\ +\infty, & \text{if } x \notin \theta_k. \end{cases}$$

As F is proper C-convex and positively lower semicontinuous, it is clear that θ_k is a convex and closed set by its definition. Since $x_k \in \theta_k$, θ_k is nonempty. It follows that $\langle F(x), z\rangle + I_{\theta_k}(x)$ is proper convex and lower semicontinuous with respect to x. By the definition of $L(x, x_k)$, we know that

$$L(x, x_k) = \|x\|^2 - 2\langle x, Jx_k\rangle + \|Jx_k\|^2 = \|Jx\|^2 - 2\langle x, Jx_k\rangle + \|Jx_k\|^2.$$

Hence

$$\nabla_x L(x, x_k) = 2Jx - 2Jx_k.$$

As $\{e_k\} \subset intC$ and the definition of C^{*0}, we have that $\langle e_k, z\rangle > 0$. Now we can define $\omega_k = \frac{\varepsilon_k}{2}\langle e_k, z\rangle$, it is easy to check that $\omega_k > 0$ for all $k \in N$. From Lemma 3.2.3, we know that

$$rge\{\partial_x(\langle F(x), z\rangle + I_{\theta_k}(x)) + \omega_k Jx\} = X^*.$$

It follows that

$$rge\{\partial_x(\langle F(x), z\rangle + I_{\theta_k}(x)) + \omega_k Jx - \omega_k Jx_k\} = X^*.$$

That is

$$0 \in rge\{\partial_x(\langle F(x), z\rangle + I_{\theta_k}(x)) + \omega_k Jx - \omega_k Jx_k\}.$$

Thus the subdifferential of $\phi_k(x)$ has some zeros, which are minimizers of $\phi_k(x)$. From Lemma 3.3.1, we have that

$$x_{k+1} \in \arg\min \phi_k(x) \subset C - ARGMIN_w\{F(x) + \frac{\varepsilon_k}{2}L(x, x_k)e_k + I_{\theta_k}(x)e_k\}.$$

x_{k+1} is a solution of (3.3.1). The proof is complete. □

Theorem 3.3.2 *Let the assumptions in Theorem 3.3.1 hold and suppose further that the set* $(F(x_0) - C) \cap F(X)$ *is* C-complete *and* $\bar{X}$ *is nonempty and compact, where* $\bar{X}$ *is the weak Pareto optimal solution set of* F *over* X*. The sequence* $\{x_k\}$ *generated by the method (PML) is then bounded.*

Proof. Based on the method (PML), we know that if the sequence stops at some iteration, it will be a constant thereafter. We now assume the sequence $\{x_k\}$ will not stop after a finite step k. Define $E \subset X$ as follows

$$E = \{x \in X \mid F(x) \leq_C F(x_k), \quad \forall\ k \in N\}. \tag{3.3.3}$$

E is nonempty based on the fact that the set $(F(x_0) - C) \cap F(X)$ is *C-complete*. As x_{k+1} is a weak Pareto optimal solution of problem (3.3.1), there exists $z_k \in C^{*0}$ such that x_{k+1} is a solution of the following problem:

$$(P_k^*) \qquad \min \quad \{\phi_k(x) |\ x \in X\}$$

with $z = z_k$. Thus x_{k+1} satisfies the first-order necessary optimality condition of problem (P_k^*). By the definition of θ_k, we have that $\theta_k \subset domF$, $\emptyset \neq dom(I_{\theta_k}) \subset domF$. Thus, we derive that there exists $\mu_k \in \partial\{\langle F(.) + \frac{\varepsilon_k}{2}L(., x_k)e_k, z_k\rangle(x_{k+1})\}$. Through Theorem 3.23 of [170], one has that

$$\langle x - x_{k+1}, \mu_k\rangle \geq 0, \quad \forall x \in \theta_k. \tag{3.3.4}$$

We define another function as follows:

$$\varphi_k(x) = \langle F(x), z_k\rangle.$$

From (3.3.2) we know that there exists some $\gamma_k \in \partial\varphi_k(x_{k+1})$ such that

$$\mu_k = \gamma_k + \frac{\varepsilon_k}{2}\langle e_k, z_k\rangle(2Jx_{k+1} - 2Jx_k). \tag{3.3.5}$$

Now taking any $x^* \in E$, it is clear that $x^* \in \theta_k$ and we derive

$$\langle x^* - x_{k+1}, \gamma_k + \frac{\varepsilon_k}{2}\langle e_k, z_k\rangle(2Jx_{k+1} - 2Jx_k)\rangle \geq 0. \tag{3.3.6}$$

Based on the definition of the subgradient of φ_k,

$$\langle F(x^*) - F(x_{k+1}), z_k\rangle \geq \langle x^* - x_{k+1}, \gamma_k\rangle.$$

As $x^* \in E$ and $z_k \in C^{*0}$, and it follows that $\langle F(x^*) - F(x_{k+1}), z_k\rangle \leq 0$ and it follows that $\langle x^* - x_{k+1}, \gamma_k\rangle \leq 0$. From (3.3.6), we have

$$\frac{\varepsilon_k}{2}\langle e_k, z_k\rangle\langle x^* - x_{k+1}, X2Jx_{k+1} - 2Jx_k\rangle \geq 0.$$

By defining $\eta_k = \frac{\varepsilon_k}{2}\langle e_k, z_k\rangle$, it is easy to check that $\eta_k > 0$. That is

$$2\langle x^* - x_{k+1}, Jx_{k+1}\rangle - 2\langle x^* - x_{k+1}, Jx_k\rangle$$

$$\begin{aligned}
&= 2\langle x^*, Jx_{k+1}\rangle - 2\langle x_{k+1}, Jx_{k+1}\rangle - 2\langle x^*, Jx_k\rangle + 2\langle x_{k+1}, Jx_k\rangle \\
&= \|x^*\|^2 - 2\langle x^*, Jx_k\rangle + \|x_k\|^2 - \|x^*\|^2 + 2\langle x^*, Jx_{k+1}\rangle - \|x_{k+1}\|^2 \\
&\quad -\|x_{k+1}\|^2 + 2\langle x_{k+1}, Jx_k\rangle - \|x_k\|^2 \geq 0.
\end{aligned} \tag{3.3.7}$$

By the definition of Lyapunov functional $L(x, y)$ and (3.3.7), we obtain the following inequality:

$$L(x_{k+1}, x_k) \leq L(x^*, x_k) - L(x^*, x_{k+1}). \tag{3.3.8}$$

Summing up inequality (3.3.8), we have $\sum_{k=0}^{\infty} L(x_{k+1}, x_k) \leq L(x^*, x_0) < +\infty$. It follows that

$$\lim_{k\to+\infty} L(x_{k+1}, x_k) = 0. \tag{3.3.9}$$

From the property (3.2.5) of Lyapunov functional, we have that

$$(\| x_{k+1} \| - \| x^* \|)^2 \leq L(x^*, x_{k+1}) \leq L(x^*, x_k) \leq L(x^*, x_0) \leq (\| x_0 \| + \| x^* \|)^2$$

which implies

$$\| x_{k+1} \| \leq \| x_0 \| + 2\|x^*\| \quad \forall\ k. \tag{3.3.10}$$

We conclude that $\{x_k\}$ is bounded. The proof is complete. □

Theorem 3.3.3 *Let the assumptions in Theorem 3.3.2 hold. Then any weak cluster point of $\{x_k\}$ belongs to $\bar{X}$.*

Proof. As $\{x_k\}$ is bounded, it has some weak cluster points. Next we will show that all of weak cluster points are weak Pareto optimal solutions of the problem (VOP). Let $\hat{x}$ be one of the weak cluster points of $\{x_k\}$ and $\{x_{k_j}\}$ be a subsequence of $\{x_k\}$, which weakly converges to $\hat{x}$. For any $z \in C_0^*$, we define the function $\psi_z : X \to R \cup \{+\infty\}$ as $\psi_z(x) = \langle F(x), z \rangle$. Since F is positively lower semicontinuous and C-convex, ψ_z is also lower semicontinuous and convex, it follows that $\psi_z(\hat{x}) \leq \liminf_{j \to +\infty} \psi_z(x_{k_j})$. By the fact that $x_{k+1} \in \theta_k$, we have that $F(x_{k+1}) \leq_C F(x_k)$ for every $k \in N$. Thus, $\psi_z(x_{k+1}) \leq \psi_z(x_k)$. Therefore,

$$\psi_z(\hat{x}) \leq \liminf_{j \to +\infty} \psi_z(x_{k_j}) = \inf\{\psi_z(x_k)\}.$$

Hence for any $z \in C_0^*$, we have $\psi_z(\hat{x}) \leq \psi_z(x_k)$, which implies

$$F(\hat{x}) \leq_C F(x_k). \tag{3.3.11}$$

Assume that $\hat{x}$ is not a weak Pareto optimal solution of the problem (VOP), then there exists $\bar{x} \in X$ such that $F(\bar{x}) \leq_{intC} F(\hat{x})$. Taking $z_k \in C^{*0}$ as the same as in the problem (P_k^*), from Lemma 3.2.1 we know that C^{*0} is weak*-compact. By virtue of the Banach-Alaoglu Theorem, there exists $\bar{z} \in C^{*0}$ such that $\bar{z}$ is a weak*-cluster point of $\{z_{k_j}\}$. Without loss of generality, we assume that

$$w^* - \lim_{j \to +\infty} z_{k_j} = \bar{z}.$$

It follows that

$$\langle F(\bar{x}) - F(\hat{x}), z_{k_j} \rangle \geq \langle F(\bar{x}) - F(x_{k_j+1}), z_{k_j} \rangle = \varphi_{k_j}(\bar{x}) - \varphi_{k_j}(x_{k_j+1}), \tag{3.3.12}$$

where φ_{k_j} is defined as the proof of Theorem 3.3.2. From (3.3.5), there exists some $\gamma_{k_j} \in \partial \varphi_{k_j}(x_{k_j+1})$ such that

$$\varphi_{k_j}(\bar{x}) - \varphi_{k_j}(x_{k_j+1}) \geq \langle \bar{x} - x_{k_j+1}, \gamma_{k_j} \rangle$$

$$= \langle \bar{x} - x_{k_j+1}, \mu_{k_j} \rangle - \eta_{k_j} \langle \bar{x} - x_{k_j+1}, 2Jx_{k_j+1} - 2Jx_{k_j} \rangle.$$

Since x_{k_j+1} is a solution of problem $(P_{k_j}^*)$, we have $\langle \bar{x} - x_{k_{j+1}}, \mu_{k_j} \rangle \geq 0$, and we can see that

$$\varphi_{k_j}(\bar{x}) - \varphi_{k_j}(x_{k_j+1}) \geq -\eta_{k_j} \langle \bar{x} - x_{k_j+1}, 2Jx_{k_j+1} - 2Jx_{k_j} \rangle$$

$$-\eta_{k_j}\langle \bar{x} - x_{k_j+1}, Jx_{k_j+1} - Jx_{k_j}\rangle \geq -\eta_{k_j} \parallel Jx_{k_j+1} - Jx_{k_j} \parallel \parallel \bar{x} - x_{k_j+1} \parallel . \tag{3.3.13}$$

By Proposition 3.2.1, (3.3.9) and the fact that $\{x_k\}$ is bounded, it is easy to check that

$$\lim_{j\to+\infty} \parallel x_{k_j+1} - x_{k_j} \parallel = 0. \tag{3.3.14}$$

From inequality (3.2.4), we have

$$\parallel Jx_{k_j+1} - Jx_{k_j} \parallel \parallel \bar{x} - x_{k_j+1} \parallel \leq 8Mh_X(16L \parallel x_{k_j+1} - x_{k_j} \parallel /M) \parallel \bar{x} - x_{k_j+1} \parallel$$

where we use the fact that $\{x_k\}$ is bounded by M. Meanwhile, by (3.3.14) and Definition 3.2.2, we obtain that

$$\lim_{j\to+\infty} h_X(16L\|x_{k_j+1} - x_{k_j}\|/M) = 0. \tag{3.3.15}$$

Thus, based on (3.3.14), (3.3.15) and the boundedness of $\{x_k\}$, we draw the conclusion that the limit of the right expression in (3.3.13) vanishes as $j \to \infty$. It is clear that

$$\langle F(\bar{x}) - F(\hat{x}), \bar{z}\rangle \geq 0. \tag{3.3.16}$$

Then we obtain that (3.3.16) contradicts with the facts that $\bar{z} \in C^{*0}$ and the assumption $F(\bar{x}) \leq_{intC} F(\hat{x})$, thus we can conclude that $\hat{x}$ is a weak Pareto optimal solution of the problem (VOP). The proof is complete. □

Theorem 3.3.4 *Consider the same assumptions as those in Theorem 3.3.2 and further suppose that the normalized dual mapping J is* weak-to-weak continuous. *Then the whole sequence $\{x_k\}$ weakly converges to a weak Pareto optimal solution of the problem (VOP).*

Proof. Let's consider the contrary, assume that there are two $\hat{x}$ and $\tilde{x}$ are weak cluster points of $\{x_k\}$, and that $\{x_{k_j}\}$ and $\{x_{k_i}\}$, which are two subsequence of $\{x_k\}$, which satisfy

$$w - \lim_{j\to+\infty} x_{k_j} = \hat{x}, \quad w - \lim_{i\to+\infty} x_{k_i} = \tilde{x}.$$

Thus, it is clear from Theorem 3.3.3 that $\hat{x}$ and $\tilde{x}$ are weak Pareto optimal solutions of the problem (VOP). Furthermore, from (3.3.11), we obtain that $\hat{x}$ and $\tilde{x} \in E$, that is $L(\hat{x}, x_k)$ and $L(\tilde{x}, x_k)$ are bounded and convergent. Let $\bar{l}_1$ and $\bar{l}_2$ be their limits respectively. Then we have

$$\lim_{k\to+\infty}(L(\hat{x}, x_k) - L(\tilde{x}, x_k)) = \bar{l}_1 - \bar{l}_2. \tag{3.3.17}$$

On the other hand, from the definition of the Lyaponov functional, we obtain that

$$L(\hat{x}, x_k) - L(\tilde{x}, x_k) = \| \hat{x} \|^2 - 2\langle \hat{x} - \tilde{x}, Jx_k\rangle - \| \tilde{x} \|^2 . \tag{3.3.18}$$

Let W be the limit of (3.3.18), taking $k = k_j$ in (3.3.18) and using the weak-to-weak continuity of normalized duality map J, we can derive that $W = -L(\hat{x}, \tilde{x})$. Repeating with $k = k_i$, we have $W = L(\tilde{x}, \hat{x})$. It follows that

$$L(\hat{x}, \tilde{x}) + L(\tilde{x}, \hat{x}) = 0. \tag{3.3.19}$$

That is

$$\langle J(\hat{x}) - J(\tilde{x}), \hat{x} - \tilde{x}\rangle = 0.$$

By (3.2.2), we have

$$0 \geq (2L)^{-1}M^2\delta_X(\frac{\| \hat{x} - \tilde{x} \|}{2M}).$$

From Definition 3.2.1, we derive that

$$\| \hat{x} - \tilde{x} \| = 0.$$

Thus, we obtain that $\hat{x} = \tilde{x}$, which proves the uniqueness of the weak cluster point of $\{x_k\}$. The proof is complete. □

3.4 Inexact Vector-valued Proximal-type Method in Banach Spaces

As in Eckstein [78], the ideal form of the proximal point algorithm is often impractical since, in most cases, iteratively updating x_{k+1} exactly is either impossible or is the same as solving the original problem $0 \in Tx$. Moreover, there seems to be little justification in the effort required to solve the problem accurately when the iterate is far from the solution point. Rockafellar [173] also gave an inexact variant of the proximal point algorithm, that is finding $z^{k+1} \in H$ such that

$$z^{k+1} + c_k v^{k+1} - z^k = \beta^k$$

where β^k is regarded as the error sequence. It was shown that, if $\| \beta^k \| \to 0$ quickly enough such that

$$\sum_{k=0}^{\infty} \| \beta^k \| < \infty,$$

then z^k is weakly convergent to z with $0 \in T(z)$. Because of its relaxed accuracy requirement, the inexact proximal point algorithm is more practical than the exact one. Thus the study of inexact proximal point algorithm has received extensive attention and various forms of the algorithm have been developed for scalar-valued optimization problems and variational inequality problems.

For any given x_k, let $\theta_k = \{x \in X \mid F(x) \leq_C F(x_k)\}$, $\varepsilon_k \in (0, \varepsilon]$ and $\varepsilon > 0$. The sequence $\{e_k\} \subset intC$ such that $\| e_k \| = 1$. We now present an inexact version of vector-valued generalized proximal point algorithm based on Lyapunov functional (GPPAL, in short):

Step (1) : Take $x_0 \in domF$;

Step (2) : Given x_k, if $x_k \in C - ARGMIN_w\{F(x) \mid x \in X\}$, then the algorithm stops; otherwise, go to step (3); and

Step (3): If $x_k \notin C - ARGMIN_w\{F(x) \mid x \in X\}$, then compute x_{k+1} such that

$$x_{k+1} \in C - ARGMIN_w\{F(x) + \frac{\varepsilon_k}{2}(L(x, x_k) - \langle x, \beta_{k+1} \rangle)e_k \mid x \in \theta_k\}, \tag{3.4.1}$$

go to step (2), where $\{\beta_k\} \subset X^*$ is regarded as the error sequence, which is satisfied with

$$\sum_{k=0}^{\infty} \| \beta_k \|_{X^*} < +\infty \tag{3.4.2}$$

and

$$\sum_{k=0}^{\infty} \langle x_k, \beta_k \rangle \ \ exists \ \ and \ \ is \ \ finite. \tag{3.4.3}$$

Before we show the main results of this section, we present the following assumptions of the initial point x_0 and the solution sets:

(A) $\bar{X}$ is nonempty and compact;

(B) the set $(F(x_0)-C)\cap F(X)$ is $C-complete$, which means that for all sequences $\{a_n\}\subset X$, with $a_0=x_0$, such that $F(a_{n+1})\leq_C F(a_n)$ for all $n\in N$, there exists $a\in X$ such that $F(a)\leq_C F(a_n)$ for all $n\in N$.

Theorem 3.4.1 *Let $F: X\to Y\cup\{+\infty_C\}$ be a proper, C- convex and positively lower semicontinuous mapping. Then, any sequence $\{x_k\}$ generated by the algorithm (GPPAL) is well-defined.*

Proof Let $x_0\in domF$ be the initial point and we assume that the algorithm has reached step k, then we will show that the next iterative x_{k+1} does exist. Take any $z\in C^{*0}$ and define a function $\phi_k(x): X\to R\cup\{+\infty\}$ as follows:

$$\phi_k(x)=\langle F(x),z\rangle+I_{\theta_k}(x)+\frac{\varepsilon_k}{2}(L(x,x_k)-\langle x,\beta_{k+1}\rangle)\langle e_k,z\rangle \tag{3.4.4}$$

where

$$I_{\theta_k}(x)=\begin{cases}0, & \text{if } x\in\theta_k;\\ +\infty, & \text{if } x\notin\theta_k.\end{cases}$$

From the assumptions that F is C-convex and positively lower semicontinuous, it is clear that θ_k is a convex set by its definition. Since $x_k\in\theta_k$, θ_k is nonempty. It follows that $\langle F(x),z\rangle+I_{\theta_k}(x)$ is convex and lower semicontinuous with respect to x. By the definition of $L(x,x_k)$, we know that

$$L(x,x_k)=\|x\|^2-2\langle x,Jx_k\rangle+\|Jx_k\|^2=\|Jx\|^2-2\langle x,Jx_k\rangle+\|Jx_k\|^2.$$

Hence

$$\nabla_x L(x,x_k)=2Jx-2Jx_k.$$

By the fact of $\{e_k\}\subset intC$ and the definition of C^{*0}, we have that $\langle e_k,z\rangle>0$. Now we can define $\omega_k=\frac{\varepsilon_k}{2}\langle e_k,z\rangle$, it is easy to check that $\omega_k>0$ for all $k\in N$. From Lemma 3.2.3, we know that

$$rge\{\partial_x(\langle F(x),z\rangle+I_{\theta_k}(x))+\omega_k Jx\}=X^*.$$

It follows that

$$rge\{\partial_x(\langle F(x),z\rangle+I_{\theta_k}(x))+\omega_k Jx-\omega_k Jx_k-\omega_k\beta_{k+1}\}=X^*.$$

That is

$$0 \in rge\{\partial_x(\langle F(x), z\rangle + I_{\theta_k}(x)) + \omega_k Jx - \omega_k Jx_k - \omega_k\beta_{k+1}\}.$$

Thus the subdifferential of $\phi_k(x)$ has some zeros, which are minimizers of $\phi_k(x)$. Thus, the subdifferential of $\phi_k(x)$ has some zeroes, which are minimizers of $\phi_k(x)$. By Lemma 3.3.1, we conclude that such a minimizer satisfies

$$x_{k+1} \in \arg\min \phi_k(x) \subset C - ARGMIN_w\{F(x) + \frac{\varepsilon_k}{2}(L(x, x_k) - \langle x, \beta_{k+1}\rangle)e_k + I_{\theta_k}(x)e_k\}.$$

Hence, it satisfies (3.4.1) and can be taken as x_{k+1}. The proof is complete. □

Theorem 3.4.2 *Let the assumptions in Theorem 3.4.1 hold. Furthermore suppose that the assumption (A) and (B) hold. The sequence $\{x_k\}$ generated by the algorithm (GPPAL) is then bounded.*

Proof. From the algorithm (GPPAL), we know that if the sequence stops at some iterations, it will be a constant thereafter. Now we assume the sequence $\{x_k\}$ will not stop after a finite step k. Define $E \subset X$ as follows

$$E = \{x \in X \mid F(x) \leq_C F(x_k) \quad \forall k \in N\}.$$

It is easy to check E is nonempty by the fact that the set $(F(x_0) - C) \cap F(X)$ is *C-complete*. Since x_{k+1} is a weak Pareto optimal solution of problem (3.4.1) and there exists $z_k \in C^{*0}$ such that x_{k+1} is a solution of following problem

$$(\bar{P}_k) \qquad \min \quad \{\phi_k(x) | \ x \in X\}$$

with $z = z_k$. Thus, x_{k+1} satisfies first-order necessary optimality condition of problem $(\bar{P}_k)$. By the definition of θ_k, we have that $\theta_k \subset domF$, $\emptyset \neq dom(I_{\theta_k}) \subset domF$, then we derive that there exits $\mu_k \in \partial\{\langle F(.) + \frac{\varepsilon_k}{2}(L(., x^k) - \langle ., \beta_{k+1}\rangle)e_k, z_k\rangle(x_{k+1})\}$. By the virtue of Theorem 3.23 of [170], one has that

$$\langle x - x_{k+1}, \mu_k\rangle \geq 0 \tag{3.4.5}$$

for any $x \in \theta_k$. We can now define another function as follows:

$$\varphi_k(x) = \langle F(x), z_k\rangle.$$

From (3.4.4) we know that there exists some $\gamma_k \in \partial\varphi_k(x_{k+1})$ such that

$$\mu_k = \gamma_k + \frac{\varepsilon_k}{2}\langle e_k, z_k\rangle(2Jx_{k+1} - 2Jx_k - \beta_{k+1}). \tag{3.4.6}$$

Taking $x^* \in E$, clearly $x^* \in \theta_k$ and we can derive the following:

$$\langle x^* - x_{k+1}, \gamma_k + \frac{\varepsilon_k}{2}\langle e_k, z_k\rangle(2Jx_{k+1} - 2Jx_k - \beta_{k+1})\rangle \geq 0. \tag{3.4.7}$$

From the definition of the subgradient of φ_k, it follows that

$$\langle F(x^*) - F(x_{k+1}), z_k\rangle \geq \langle x^* - x_{k+1}, \gamma_k\rangle.$$

By the fact $x^* \in E$ and $z_k \in C^{*0}$, it is easy to check that

$$\langle F(x^*) - F(x_{k+1}), z_k\rangle \leq 0$$

and it follows that $\langle x^* - x_{k+1}, \gamma_k\rangle \leq 0$. From (3.4.7) we obtain that

$$\frac{\varepsilon_k}{2}\langle e_k, z_k\rangle\langle x^* - x_{k+1}, 2Jx_{k+1} - 2Jx_k - \beta_{k+1}\rangle \geq 0. \tag{3.4.8}$$

Now we can define $\eta_k = \frac{\varepsilon_k}{2}\langle e_k, z_k\rangle$, it is easy to check that $\eta_k > 0$. By the definition of Lyapunov functional $L(x, y)$ and (3.4.8), we can derive the following inequalities:

$$L(x^*, x_k) - L(x^*, x_{k+1}) - L(x_{k+1}, x_k) - \langle x^* - x_{k+1}, \beta_{k+1}\rangle \geq 0,$$

$$L(x_{k+1}, x_k) \leq L(x^*, x_k) - L(x^*, x_{k+1}) + \langle x_{k+1} - x^*, \beta_{k+1}\rangle. \tag{3.4.9}$$

Thus, we obtain that for all $l \geq 1$

$$L(x^*, x_l) \leq L(x^*, x_0) - \sum_{k=0}^{l-1} L(x_{k+1}, x_k) + \sum_{k=0}^{l-1}\langle x_{k+1} - x^*, \beta_{k+1}\rangle.$$

From the property (3.4.2) of the error sequence, we know that

$$0 \leq \sum_{k=0}^{\infty}\langle x^*, \beta_{k+1}\rangle \leq \sum_{k=0}^{\infty} \| \beta_{k+1} \|\| x^* \| < +\infty,$$

and $\sum\limits_{k=0}^{\infty}\langle x^*, \beta_{k+1}\rangle$ is convergent. Combining with (3.4.3), one has that $\sum\limits_{k=0}^{\infty}\langle x_{k+1} - x^*, \beta_{k+1}\rangle$ exists and is finite. It follows that

$$Q = \sup_{l\geq 1}\{\sum_{k=0}^{l-1}\langle x_{k+1} - x^*, \beta_{k+1}\rangle\} \tag{3.4.10}$$

From the property (3.2.5) of Lyapunov functional $L(x, y)$ and (3.4.10), we obtain that

$$(\| x_l \| - \| x^* \|)^2 \leq L(x^*, x_l) \leq L(x^*, x_0) + Q \leq (\| x_0 \| + \| x^* \|)^2 + Q$$

which implies that

$$\| x_l \| \leq \sqrt{(\| x_0 \| + \| x^* \|)^2 + Q} + \| x^* \| \leq M$$

where M is any real number larger than $\sqrt{(\| x_0 \| + \| x^* \|)^2 + Q} + \| x^* \|$ for all $l \geq 1$. Since $\bar{X}$ is nonempty compact, we conclude that $\{x_k\}$ is bounded. Summing up inequality (3.4.9) and from (3.4.10), we have

$$\sum_{k=0}^{\infty} L(x_{k+1}, x_k) \leq L(x^*, x_0) + \sum_{k=0}^{\infty} \langle x_{k+1} - x^*, \beta_{k+1} \rangle < +\infty.$$

It follows that

$$\lim_{k \to +\infty} L(x_{k+1}, x_k) = 0. \tag{3.4.11}$$

Theorem 3.4.3 *Let assumptions in Theorem 3.4.2 hold. Then any weak cluster points of $\{x_n\}$ belong to $\bar{X}$.*

Proof. Since $\{x_k\}$ is bounded, it has some weak cluster points. Next we will show that all of weak cluster points are weak Pareto optimal solution of problem problem (VOP). Let $\hat{x}$ be one of the weak cluster points of $\{x_k\}$ and $\{x_{k_j}\}$ be a subsequence of $\{x_k\}$, which weakly converges to $\hat{x}$. We define, for each $z \in C^{*0}$, the function $\psi_z : X \to R \cup \{+\infty\}$ as $\psi_z(x) = \langle F(x), z \rangle$. Since F is positively lower semicontinuous and C-convex, ψ_z is also lower semicontinuous and convex, it follows that $\psi_z(\hat{x}) \leq \liminf_{j \to +\infty} \psi_z(x_{k_j})$. By the fact that $x_{k+1} \in \theta_k$, we can see that $F(x_{k+1}) \leq_C F(x_k)$ for $k \in N$. Thus, $\psi_z(x_{k+1}) \leq \psi_z(x_k)$. Therefore,

$$\psi_z(\hat{x}) \leq \liminf_{j \to +\infty} \psi_z(x_{k_j}) = \inf\{\psi_z(x_k)\}.$$

Hence, we have that for all $z \in C^{*0}$

$$\psi_z(\hat{x}) \leq \psi_z(x_k),$$

which implies

$$F(\hat{x}) \leq_C F(x_k).$$

Assuming that $\hat{x}$ is not the weak Pareto optimal solution of problem (VOP), then there exists $\bar{x} \in X$ such that $F(\bar{x}) \leq_{intC} F(\hat{x})$. Taking $z_k \in C^{*0}$ to be the same as that in problem $(\bar{P}_k)$, from Lemma 3.2.1 we know that C^{*0} be weak*-compact. By virtue of Banach-Alaoglu Theorem, there exists $\bar{z} \in C^{*0}$ such that $\bar{z}$ is a weak*-cluster point of $\{z_{k_j}\}$. Without loss of generality, we can assume that

$$w^* - \lim_{j \to +\infty} z_{k_j} = \bar{z} \tag{3.4.12}$$

thus we have that

$$\langle F(\bar{x}) - F(\hat{x}), z_{k_j} \rangle \geq \langle F(\bar{x}) - F(x_{k_j+1}), z_{k_j} \rangle = \varphi_{k_j}(\bar{x}) - \varphi_{k_j}(x_{k_j+1}), \tag{3.4.13}$$

where φ_{k_j} is defined in the proof of Theorem 3.4.2. From (3.4.6), there exist some $\gamma_{k_j} \in \partial \varphi_{k_j}(x_{k_j+1})$ such that

$$\varphi_{k_j}(\bar{x}) - \varphi_{k_j}(x_{k_j+1}) \geq \langle \bar{x} - x_{k_j+1}, \gamma_{k_j} \rangle$$

$$= \langle \bar{x} - x_{k_j+1}, \mu_{k_j} \rangle - \eta_{k_j} \langle \bar{x} - x_{k_j+1}, Jx_{k_j+1} - Jx_{k_j} - \beta_{k_j+1} \rangle.$$

As x_{k_j+1} is the solution of problem $(\bar{P}_{k_j})$, we have $\langle \bar{x} - x_{k_j+1}, \mu_{k_j} \rangle \geq 0$ and we can see that

$$\varphi_{k_j}(\bar{x}) - \varphi_{k_j}(x_{k_j+1}) \geq -\eta_{k_j} \langle \bar{x} - x_{k_j+1}, Jx_{k_j+1} - Jx_{k_j} \rangle - \eta_{k_j} \langle x_{k_j+1} - \bar{x}, \beta_{k_j+1} \rangle$$

$$\geq -\eta_{k_j} \parallel Jx_{k_j+1} - Jx_{k_j} \parallel \parallel \bar{x} - x_{k_j+1} \parallel - \eta_{k_j} \parallel \beta_{k_j+1} \parallel \parallel x_{k_j+1} - \bar{x} \parallel . \tag{3.4.14}$$

By Proposition 3.2.1 and (3.4.11) and the fact of the boundedness of $\{x_k\}$, we obtain that

$$\lim_{j \to +\infty} \parallel x_{k_j+1} - x_{k_j} \parallel = 0. \tag{3.4.15}$$

From inequality (3.2.4), we have

$$\parallel Jx_{k_j+1} - Jx_{k_j} \parallel \parallel \bar{x} - x_{k_j+1} \parallel \leq 8Mh_X(16L \parallel x_{k_j+1} - x_{k_j} \parallel /M) \parallel \bar{x} - x_{k_j+1} \parallel$$

where we use the fact that $\{x_k\}$ is bounded by M. Meanwhile, by (3.4.15) and Definition 3.2.2, we obtain the following:

$$\lim_{j \to +\infty} h_X(16L \| x_{k_j+1} - x_{k_j} \| / M) = 0. \tag{3.4.16}$$

Thus, through (3.4.2), (3.4.15), (3.4.16) and by the fact of the boundedness of $\{x_k\}$, we draw the conclusion that the limit of the right-hand expression in (3.4.14) disappears as $j \to \infty$. Clearly,

$$\langle F(\bar{x}) - F(\hat{x}), \bar{z} \rangle \geq 0. \tag{3.4.16}$$

Then we can see that (3.4.16) contradicts with the facts that $\bar{z} \in C^{*0}$ and the assumption $F(\bar{x}) \leq_{intC} F(\hat{x})$, thus we can conclude that $\hat{x}$ is a weak Pareto optimal solution of problem (VOP). The proof is complete. $\square$

Theorem 3.4.4 *Consider the same assumptions in Theorem 3.4.3 and assume that the normalized dual mapping J is weak-to-weak continuous. Then the whole sequence $\{x_k\}$ weakly converges to a weak Pareto optimal solution of problem (VOP).*

Proof This is similar to the proof of Theorem 3.3.4, and thus we do not include it anymore. The proof is complete. $\square$

3.5 Remarks and Conclusions

In this chapter, we considered a convex vector optimization problem of finding weak Pareto optimal solutions for an extended vector-valued map from a uniformly convex and uniformly smooth Banach space to a real Banach space with respect to the partial order induced by a closed, convex, and pointed cone with a nonempty interior. For this problem, we developed an extension of the proximal point method for scalar-valued convex optimization. In this extension, the subproblems involve the finding of weak Pareto optimal points for the suitable regularization of the original map through the Lyapunov functional. We presented both exact and inexact versions, where the subproblems are solved only approximately within a relative tolerance. In both cases, we proved the weak convergence of the generated sequence to a weak Pareto optimal solution of the vector optimization problem.

Remark 3.5.1 *The difference between the algorithm (GPPAL) and the algorithm (PML) is that there is an error sequence $\{\beta_k\}$ in the step (3) of the algorithm (GPPAL) such that the subproblems are solved only approximately. When the sequence $\{\beta_k\} = 0$, the two algorithm are coincident with each other. Thus, the algorithm (PML) is a special case of the algorithm (GPPAL) and the algorithm (GPPAL) is more practical than the algorithm (PML), due to its relaxed accuracy requirement.*

Chapter 4

Generalized Proximal Point Algorithms for Multiobjective Optimization Problems

4.1 Introduction

The simple idea behind each proximal-like method to solve convex minimization problems is to replace the strictly quadratic term in the regularized subproblem by non-quadratic ones.

The notion of a *Bregman distance function* originated in [42] and the name was firstly used in optimization problems and their related topics by Censor and Lent [53] as a natural extension of the quadratic term in the classical proximal point algorithm. Since then various proximal point algorithms based on Bregman distance functions have been extensively used for scalar-valued optimization problems and other problems (e.g., [28, 54, 58, 72, 134] in finite dimensional spaces, [184] Hilbert spaces and [46, 47] Banach spaces.)

In this chapter, we consider a convex multiobjective optimization problem (MOP, in short) of finding weak Pareto optimal solutions for a map from R^n to $R^m \cup \{+\infty_C\}$, with the latter being ordered by the positive orthant. We construct a generalized proximal point algorithm based on a vector-valued Bregman distance function for solving the problem (MOP), carry out convergence analysis on the algorithm using asymptotic cones and asymptotic functions, and prove the sequence generated by the algorithm converges to a weak Pareto optimal solution of the problem (MOP). The main purpose of this chapter is to extend some important and interesting results in ([40], [47], [60], [53], [58]) to more general cases.

The outline of this chapter is as follows.

In section 4.2, we present the basic definitions, notations, and preliminary results. In section 4.3, we propose a vector-valued generalized proximal point algorithm and carry out convergence analysis. In section 4.4, we draw some conclusions.

In this chapter, let $C = R^m_+$.

4.2 Vector-valued Bregman Distance Functions

Definition 4.2.1 *[65] Let $F : K \subset R^n \to R^m$ be a vector-valued function. F is said to be Gâteaux differentiable at $x_0 \in K$ if there exists an $n \times m$ matrix $DF(x_0)$ such that, for any $v \in R^n$,*

$$DF(x_0)v = \lim_{t \searrow 0} \frac{F(x_0 + tv) - F(x_0)}{t}.$$

$DF(x_0)$ is called the Gâteaux derivative of F at x_0. F is said to be Gâteaux differentiable on K, if F is Gâteaux differentiable at every interior point of K.

Definition 4.2.2 ***Vector-valued Bregman Distance***

Let $G : R^n \to R^m \cup \{+\infty_C\}$ be a proper, strictly C- *convex and* C-*lsc map with closed convex domain $K_0 := dom(G)$ and $intK_0 \neq \emptyset$. Suppose that G is Gâteaux differentiable on K_0 with Gâteaux derivative $DG(.)$. The* **vector-valued Bregman distance** *with respect to G is the map $B_G : K_0 \times intK_0 \to R^m$ defined by*

$$B_G(z, x) := G(z) - G(x) - DG(x)^\top (z - x). \tag{4.2.1}$$

Clearly, $B_G(z, x) \in C$, $\forall\ (z, x) \in K_0 \times intK_0$.

Let us consider the following assumptions on G.

A_1: For any x, y, $z \in intK_0$, if $(DG(x) - DG(y))^\top (z - x) \notin -intC$, then

$$(DG(x) - DG(y))^\top (z - x) \in C.$$

A_2: For any $x \in K_0$, $\lambda \in C_1$, bounded sequences $\{x_k\} \subset intK_0$ and $\{y_k\} \subset intK_0$ such that $\lim_{k\to+\infty} \|x_k - y_k\| = 0$ then

$$\lim_{k\to+\infty} \langle B_G(x, x_k) - B_G(x, y_k), \lambda\rangle = 0.$$

A_3: For any bounded sequences $\{x_k\} \subset K_0$, $\{y_k\} \subset intK_0$ satisfying $\lim_{k\to+\infty} y_k = y$, and for any $\lambda \in C_1$, $\lim_{k\to+\infty} \langle B_G(x_k, y_k), \lambda\rangle = 0$, we have

$$\lim_{k\to+\infty} x_k = y.$$

A_4: For every $y \in R^n$ and $\lambda \in C_1$, there exists $x \in intK_0$ such that $\langle DG(x), \lambda\rangle = y$.

Definition 4.2.3 *The function* G *is said to be a vector-valued Bregman distance function if it satisfies* $A_1 - A_3$.

Definition 4.2.4 *The function* G *is said to be a strengthened vector-valued Bregman distance function if it satisfies* $A_1 - A_4$.

Next, we present two examples of the strengthened vector-valued Bregman distance function in different spaces.

Example 4.2.1 *When* $G(\cdot) = \frac{1}{2} \parallel . \parallel^2 e_m$*, where* $e_m \in intC$ *such that* $\parallel e_n \parallel = 1$*, the Bregman distance with respect to* $\frac{1}{2} \parallel . \parallel^2 e_m$ *can be derived as follows:*

$$B_G(z, x) = \frac{1}{2} \parallel z \parallel^2 e_n - \frac{1}{2} \parallel x \parallel^2 e_m - \langle x, z - x\rangle e_m = \frac{1}{2} \parallel z - x \parallel^2 e_m.$$

It is elementary to check that G is a strengthened vector-valued Bregman distance function, which was investigated by Bonnel et.al. [40] and Ceng and Yao [60].

Example 4.2.2 *Let $X = R^n$, $Y = R^2$ and $G : R^n \to R^2$ be defined by*

$$G(.) = \begin{pmatrix} \sum\limits_{i=1}^{n} a_i x_i \log a_i x_i \\ \sum\limits_{i=1}^{n} b_i x_i \log b_i x_i \end{pmatrix},$$

where $a_i, b_i \in R$ and $\frac{a_i}{b_i} = \beta$ is a constant for $i = 1, ..., n$. The Bregman distance with respect to G can be derived as follows:

$$B_G(x,y) = \begin{pmatrix} \sum\limits_{i=1}^{n} a_i x_i \log \frac{x_i}{y_i} - \sum\limits_{i=1}^{n} a_i (x_i - y_i) \\ \sum\limits_{i=1}^{n} b_i x_i \log \frac{x_i}{y_i} - \sum\limits_{i=1}^{n} b_i (x_i - y_i) \end{pmatrix}.$$

Clearly, G is also a strengthened vector-valued Bregman distance function.

Proposition 4.2.1 *(Three points property). Let $x \in K_0$ and $y, z \in intK_0$. Then, the following relation is true:*

$$(DG(y) - DG(z))^\top (z - x) = B_G(x,y) - B_G(x,z) - B_G(z,y). \tag{4.2.2}$$

Proof. By the definition of B_G, we have

$$DG(z)^\top (x - z) = G(x) - G(z) - B_G(x,z) \tag{4.2.3}$$

$$DG(y)^\top (z - y) = G(z) - G(y) - B_G(z,y) \tag{4.2.4}$$

$$DG(y)^\top (x - y) = G(x) - G(y) - B_G(x,y) \tag{4.2.5}$$

Subtracting (4.2.3) and (4.2.4) from (4.2.5), we obtain the desired result.

4.3 Generalized Proximal Point Algorithm and Convergence Analysis

Denote by $\bar{X}$ the weak Pareto optimal solution set of problem (MOP). We make the following assumptions:

(A) the set $\bar{X}$ is nonempty and compact;

(B) there exists $x_0 \in R^n$ such that $\{x \in R^n | \ F(x) \leq_C F(x_0)\} \subseteq dom(F) \cap intK_0$.

We propose the following vector-valued generalized proximal point algorithm based on a strengthened Bregman distance function (VPPAB, in short):

Step (1) : Take $x_0 \in domF \cap intK_0$ such that $\{x \in R^n | \ F(x) \leq_C F(x_0)\} \subseteq domF \cap intK_0$.

Step (2) : Given x_k, if $x_k \in \bar{X}$, then $x_{k+p} = x_k$ for all $p \geq 1$ and the algorithm stops, otherwise go to step (3).

Step (3): If $x_k \notin \bar{X}$, then compute x_{k+1} satisfying

$$x_{k+1} \in C - ARGMIN_w\{F(x) + \frac{\varepsilon_k}{2} B_G(x, x_k) \mid x \in \theta_k\}, \tag{4.3.1}$$

where $\theta_k := \{x \in R^n \mid F(x) \leq_C F(x_k)\}$, $\varepsilon_k \in (0, \varepsilon], \varepsilon > 0$ and go to step (2).

Next, we establish the main results of this paper.

Theorem 4.3.1 *Let $F : R^n \to R^m \cup \{+\infty_C\}$ be a proper C-convex and C-lower semi-continuous map. Under assumptions (A) and (B), any sequence $\{x_k\}$ generated by the algorithm (VPPAB) is well-defined and bounded.*

Proof. Let $x_0 \in domF \cap intK_0$ be an initial point and we assume the algorithm has reached step k. We show that the next iterative x_{k+1} does exist. Defining a function $W_k(x) : R^n \to R^m \cup \{+\infty_C\}$ as follows:

$$W_k(x) = F(x) + \delta_{\theta_k}(x)e + \frac{\varepsilon_k}{2} B_G(x, x_k) \tag{4.3.2}$$

where $e = (1, \cdots, 1)^\top \in R^m_+$ and $\delta_{\theta_k}(x)$ is the indicator function of set θ_k.

Denote by $\bar{X}_k$ the weak Pareto optimal solution set of the following multiobjective optimization problem:

$$C - MIN\{W_k(x) | \ x \in R^n\} \quad (MOP_k).$$

It is clear that θ_k is a nonempty and convex set by its definition. Since F is C-convex and C-lsc, it follows that each of the coordinates of $\{F(x) + \delta_{\theta_k}(x)e\}$ is convex and lsc with respect to x. By the assumption (A), we know that $\bar{X}$ is nonempty and compact. By virtue of Lemma 2.3.3, for any $i \in [1, ..., m]$ we have

$$F_i^\infty(d) > 0 \quad \forall\ d \neq 0 \in R^n. \tag{4.3.3}$$

From the definition of an indicator function, we know that

$$\delta_{\theta_k}{}^\infty(d) = \delta_{\theta_k{}^\infty}(d) = \begin{cases} 0, & \text{if } d \in \theta_k{}^\infty; \\ +\infty, & \text{if } d \notin \theta_k{}^\infty. \end{cases}$$

Combining this formula with (4.3.3), for any $i \in [1, ..., m]$ we have

$$F_i^\infty(d) + \delta_{\theta_k{}^\infty}(d) > 0 \quad \forall d \neq 0 \in R^n. \tag{4.3.4}$$

Meanwhile, from the assumption A_4 of G, we know that

$$rge(\langle DG(x), \lambda\rangle) = R^n, \quad \forall\ \lambda \in C_1$$

which means that

$$rge(DG_i(x)) = R^n.$$

Thus, there exists $x^* \in domF \cap intK_0$ such that

$$(DG(x^*) - DG(x_k))\lambda^\top = 0.$$

Thus, the following strictly convex optimization problem:

$$\min\{(B_G)_i(x, x_k)|\ x \in \theta_k\},$$

has a unique optimal solution $x = x_k$. Furthermore, from Theorem 3.4.1 in [17], we have

$$(B_G)_i{}^\infty(d, x_k) \geq 0 \quad \forall\ d \neq 0 \in R^n. \tag{4.3.5}$$

Thus, by Proposition 2.3.1, we have that

$$(W_k)_i^\infty(d) = f_i^\infty(d) + \delta_{\theta_k{}^\infty}(d) + \frac{\varepsilon_k}{2}(B_G)_i{}^\infty(d, x_k).$$

Combining (4.3.4) and (4.3.5) with this equation,

$$(W_k)_i^\infty(d) > 0 \quad \forall\ d \neq 0 \in R^n \ and\ i \in [1, ..., m].$$

That is for every $i \in [1, ..., m]$, we have

$$\{d \in R^n |\ F_i^\infty(d) + (B_G)_i{}^\infty(d, x_k) \le 0\} \cap \theta_k^\infty = \{0\}.$$

Thus

$$\cup_{i=1}^m \{d \in \theta_k^\infty |\ F_i^\infty(d) + (B_G)_i{}^\infty(d, x_k) \le 0\} = \{0\}.$$

Denote by $\bar{X}_k$ the weak Pareto optimal solution set of problem (4.3.1). By Lemma 2.3.3, $\bar{X}_k$ is nonempty. Thus, any element of $\bar{X}_k$ can be taken as x_{k+1}.

We claim that the sequence $\{\|x_k\|\}$ is bounded from above. If not, without loss of generality, assume $\|x_k\| \to +\infty$ as $k \to \infty$. From (4.3.3), we know that $\langle F(x), \lambda\rangle$ is coercive for any $\lambda \in C_1$. By Proposition 2.3.2, it follows that

$$\liminf_{\|x_k\| \to +\infty} \frac{\langle F(x_k), \lambda\rangle}{\|x_k\|} > 0. \tag{4.3.6}$$

However by the definition of the algorithm (VPPAB), we have

$$\liminf_{\|x_k\| \to +\infty} \frac{\langle F(x_k), \lambda\rangle}{\|x_k\|} \le \liminf_{\|x_k\| \to +\infty} \frac{\langle F(x_0), \lambda\rangle}{\|x_k\|} = 0, \tag{4.3.7}$$

a contradiction with (4.3.6). Thus, $\|x_k\|$ is bounded. The proof is complete. □

Lemma 4.3.1 *Let the assumptions in Theorem 4.3.1 hold. Then, for any $\lambda \in C_1$, we have*

$$\lim_{k \to +\infty} \langle B_G(x_{k+1}, x_k), \lambda\rangle = 0.$$

Proof. From algorithm (VPPAB), we know that if the sequence stops at some iteration, x_k will be a constant vector thereafter. We now assume that the sequence $\{x_k\}$ will not stop finitely. Define $E \subset R^n$ as follows

$$E = \{x \in R^n \mid F(x) \le_C F(x_k) \quad \forall k \in N\}.$$

Assuming that $F(R^n)$ has the domination property, it follows from the Daniell property of R_+^m that we have E is nonempty. Since x_{k+1} is a weak Pareto optimal solution of problem (4.3.1), there exists a $\lambda_k \in C_1$ such that x_{k+1} is the solution of the following problem (MOP_{λ_k}):

$$\min\{w_{\lambda_k}(x) |\ x \in R^n\},$$

where $w_{\lambda_k}(x) = \langle F(x), \lambda_k\rangle + \frac{\varepsilon_k}{2}\langle B_G(x, x_k), \lambda_k\rangle + \delta_{\theta_k}(x)$. Thus x_{k+1} satisfies the first-order necessary optimality condition of problem (MOP_{λ_k}). It is clear that $\theta_k \subset domF \cap intK_0$ and $\emptyset \neq dom(\delta_{\theta_k}) \subset domF \cap intK_0$. From Theorem 3.23 of [170] and the definition of a strengthened vector-valued Bregman distance function, there exist $\mu_k \in \partial\langle F(.), \lambda_k\rangle(x_{k+1})$ and $\nu_k \in \partial\delta_{\theta_k}(x_{k+1})$ such that

$$\mu_k + \frac{\varepsilon_k}{2}(DG(x_{k+1}) - DG(x_k))\lambda_k + \nu_k = 0.$$

By the fact that $\langle \nu_k, x - x_{k+1}\rangle \leq 0$ for any $x \in \theta_k$, we have that

$$\langle \mu_k + \frac{\varepsilon_k}{2}(DG(x_{k+1}) - DG(x_k))\lambda_k, x - x_{k+1}\rangle \geq 0 \quad \forall x \in \theta_k.$$

Let $x^* \in E$. It is obvious that $x^* \in \theta_k$ for all $k \in N$ and we deduce that

$$\langle \mu_k + \frac{\varepsilon_k}{2}(DG(x_{k+1}) - DG(x_k))\lambda_k, x^* - x_{k+1}\rangle \geq 0.$$

By the definition of subgradient of $\langle F(x_{k+1}), \lambda_k\rangle$, we have that

$$\langle F(x^*) - F(x_{k+1}), \lambda_k\rangle \geq \langle \mu_k, x^* - x_{k+1}\rangle.$$

By the fact that $x^* \in \theta_k$ for all $k \in N$, we have $\langle F(x^*) - F(x_{k+1}), \lambda_k\rangle \leq 0$. It follows that

$$\langle \mu_k, x^* - x_{k+1}\rangle \leq 0$$

and

$$\frac{\varepsilon_k}{2}\langle (DG(x_{k+1}) - DG(x_k)\lambda, x^* - x_{k+1}\rangle \geq 0.$$

That is

$$\frac{\varepsilon_k}{2}((DG(x_{k+1}) - DG(x_k))(x^* - x_{k+1})) \notin -intC. \tag{4.3.8}$$

Furthermore, by the assumption A_1 on G, we obtain

$$((DG(x_{k+1}) - DG(x_k))(x^* - x_{k+1}) \in C.$$

By Proposition 4.2.1 and taking x^* as x, x_k as y , x_{k+1} as z, we have

$$B_G(x^*, x_k) - B_G(x^*, x_{k+1}) - B_G(x_{k+1}, x_k) \geq_C 0.$$

It follows that

$$B_G(x^*, x_{k+1}) \leq_C B_G(x^*, x_k) - B_G(x_{k+1}, x_k). \tag{4.3.9}$$

And from (4.3.9), we have that for each $\lambda \in C_1$

$$\langle B_G(x_{k+1}, x_k), \lambda\rangle \leq \langle B_G(x^*, x_k), \lambda\rangle - \langle B_G(x^*, x_{k+1}), \lambda\rangle, \quad k = 0, 1, ...$$

For all $x^* \in E$ and $\lambda \in C_1$, the sequence $\langle B_G(x^*, x_k), \lambda\rangle$ is a nonnegative and nonincreasing sequence, and is hence convergent. Thus

$$\begin{aligned}\sum_{k=0}^{\infty}\langle B_G(x_{k+1}, x_k), \lambda\rangle &\leq \sum_{k=0}^{\infty}\{\langle B_G(x^*, x_k), \lambda\rangle - \langle B_G(x^*, x_{k+1}), \lambda\rangle\} \\ &\leq \langle B_G(x^*, x_0), \lambda\rangle < \infty\end{aligned}$$

hence

$$\lim_{k\to+\infty}\langle B_G(x_{k+1}, x_k), \lambda\rangle = 0.$$

The proof is complete. □

Theorem 4.3.2 *Let the assumptions in Theorem 4.3.1 and Lemma 4.3.1 hold. Any cluster point of $\{x_k\}$ belongs to $\bar{X}$.*

Proof. If there exists $k_0 \geq 1$ such that $x_{k_0+p} = x_{k_0}, \forall p \geq 1$. Thus, clearly, x_{k_0} is a cluster point of $\{x_k\}$ and it is also a weak Pareto optimal solution of problem (MOP). Now suppose that the algorithm does not terminate finitely. Then, by Theorem 4.3.1, we have that $\{x_k\}$ is bounded and it has some cluster points. Next we will show that all of cluster points are weak Pareto optimal solutions of problem (MOP). Let $\hat{x}$ be one of the cluster points of $\{x_k\}$ and $\{x_{k_j}\}$ be a subsequence of $\{x_k\}$, which converges to $\hat{x}$. Let $\lambda \in C_1$. We define a function $\psi_\lambda : R^n \to R \cup \{+\infty\}$ as $\psi_\lambda(x) = \langle F(x), \lambda\rangle$, $x \in R^n$. Since F is C-lower semicontinuous and C-convex, ψ_λ is also lower semicontinuous and convex. It follows that $\psi_\lambda(\hat{x}) \leq \liminf_{j\to+\infty} \psi_\lambda(x_{k_j})$. By the fact that $x_{k+1} \in \theta_k$, we can see that $F(x_{k+1}) \leq_C F(x_k)$ for $k \in N$. Thus, $\psi_\lambda(x_{k+1}) \leq \psi_\lambda(x_k)$. Therefore,

$$\psi_\lambda(\hat{x}) \leq \liminf_{j\to+\infty} \psi_\lambda(x_{k_j}) = \inf_k\{\psi_\lambda(x_k)\}.$$

Hence, we have that

$$\psi_\lambda(\hat{x}) \leq \psi_\lambda(x_k), \quad \forall\ k.$$

The arbitrariness of λ guarantees the following:

$$F(\hat{x}) \leq_C F(x_k), \quad \forall\ k. \tag{4.3.10}$$

Suppose that, in contrast, there exists $\bar{x} \in R^n$ such that $F(\bar{x}) \leq_{intC} F(\hat{x})$. Taking $\lambda_{k_j} \in C_1$, as C_1 is compact, there exists $\bar{\lambda} \in C_1$ such that $\bar{\lambda}$ is a cluster point of $\{\lambda_{k_j}\}$. Without loss of generality, we assume the following:

$$\lim_{j \to +\infty} \lambda_{k_j} = \bar{\lambda}.$$

Thus we have that

$$\langle F(\bar{x}) - F(\hat{x}), \lambda_{k_j} \rangle \geq \langle F(\bar{x}) - F(x_{k_j+1}), \lambda_{k_j} \rangle = \psi_{\lambda_{k_j}}(\bar{x}) - \psi_{\lambda_{k_j}}(x_{k_j+1}).$$

Let w_λ be the function defined in the proof of Lemma 4.3.1. There exist some $\varphi_{k_j} \in \partial\psi_{\lambda_{k_j}}(x_{k_j+1})$, $\rho_{k_j} \in \partial w_{\lambda_{k_j}}(x_{k_j+1})$ and $\nu_{k_j} \in \partial\delta_{\theta_{k_j}}(x_{k_j+1})$ such that

$$\rho_{k_j} = \varphi_{k_j} + \frac{\varepsilon_{k_j}}{2}(DG(x_{k_j+1}) - DG(x_{k_j}))\lambda_{k_j} + \nu_{k_j}.$$

It follows that

$$\psi_{\lambda_{k_j}}(\bar{x}) - \psi_{\lambda_{k_j}}(x_{k_j+1}) \geq \langle \varphi_{k_j}, \bar{x} - x_{k_j+1} \rangle$$

$$= \langle \rho_{k_j}, \bar{x} - x_{k_j+1} \rangle - \frac{\varepsilon_{k_j}}{2}\langle (DG(x_{k_j+1}) - DG(x_{k_j}))\lambda_{k_j}, \bar{x} - x_{k_j+1} \rangle - \langle \nu_{k_j}, \bar{x} - x_{k_j+1} \rangle. \quad (4.3.11)$$

From the definition of algorithm (VPPAB) and the subdifferential of $\delta_{\theta_{k_j}}$, we have

$$\langle \rho_{k_j}, \bar{x} - x_{k_j+1} \rangle \geq 0, \quad \langle \nu_{k_j}, \bar{x} - x_{k_j+1} \rangle \leq 0. \quad (4.3.12)$$

Combining (4.3.11) with (4.3.12), that is

$$\psi_{\lambda_{k_j}}(\bar{x}) - \psi_{\lambda_{k_j}}(x_{k_j+1}) \geq -\frac{\varepsilon_{k_j}}{2}\langle (DG(x_{k_j+1}) - DG(x_{k_j}))\lambda_{k_j}, \bar{x} - x_{k_j+1} \rangle. \quad (4.3.13)$$

By Proposition 4.2.1, we have

$$\langle (DG(x_{k_j+1}) - DG(x_{k_j}))\lambda_{k_j}, \bar{x} - x_{k_j+1} \rangle$$

$$= \langle B_G(\bar{x}, x_{k_j}) - B_G(\bar{x}, x_{k_j+1}), \lambda_{k_j} \rangle - \langle B_G(x_{k_j+1}, x_{k_j}), \lambda_{k_j} \rangle. \quad (4.3.14)$$

By the fact that $\langle \lambda_{k_j}, B_G(x_{k_j+1}, x_{k_j}) \rangle \geq 0$, we have

$$\langle (DG(x_{k_j+1}) - DG(x_{k_j}))\lambda_{k_j}, \bar{x} - x_{k_j+1} \rangle \leq \langle B_G(\bar{x}, x_{k_j}) - B_G(\bar{x}, x_{k_j+1}), \lambda_{k_j} \rangle.$$

From Lemma 4.3.1, clearly, for any $\lambda \in C_1$,

$$\lim_{j \to +\infty} \langle B_G(x_{k_j+1}, x_{k_j}), \lambda \rangle = 0. \quad (4.3.15)$$

Thus, we have $\lim_{j\to+\infty} x_{k_j+1} = \hat{x}$ by the assumption A_3 on G. All the conditions of assumption A_2 are satisfied. Thus,

$$\lim_{j\to+\infty} \langle B_G(\bar{x}, x_{k_j}) - B_G(\bar{x}, x_{k_j+1}), \lambda_{k_j}\rangle = 0.$$

Thus we have that

$$\langle F(\bar{x}) - F(\hat{x}), \bar{\lambda}\rangle \geq 0 \tag{4.3.16}$$

where $\bar{\lambda}$ is the cluster point of $\{\lambda_{k_j}\}$. We can then conclude that (4.3.16) contradicts with the facts that $\bar{\lambda} \in C_1$ and the assumption $F(\bar{x}) \leq_{intC} F(\hat{x})$, thus we can claim that $\hat{x}$ is a weak Pareto optimal solution of problem (MOP). The proof is complete. □

Theorem 4.3.3 *Assume the same assumptions as in Theorem 4.3.2 and further suppose that $DG(.)$ is norm-to-norm continuous. Then the whole sequence $\{x_k\}$ converges to a weak Pareto optimal solution of problem (MOP).*

Proof Suppose that, in contrast, both $\hat{x}$ and $\tilde{x}$ are two distinct cluster points of $\{x_k\}$ and

$$\lim_{j\to+\infty} x_{k_j} = \hat{x}, \quad \lim_{i\to+\infty} x_{k_i} = \tilde{x},$$

By Theorem 4.3.2, $\hat{x}$ and $\tilde{x}$ are both weak Pareto optimal solutions of problem (MOP). From the proof of Lemma 4.3.1, we have for any $\lambda \in C_1$, the sequence $\langle B_G(\hat{x}, x_k), \lambda\rangle$ and $\langle B_G(\tilde{x}, x_k), \lambda\rangle$ are convergent, let $\bar{l}_1$ and $\bar{l}_2$ be their limits respectively. Then

$$\lim_{k\to+\infty} \{\langle B_G(\hat{x}, x_k), \lambda\rangle - \langle B_G(\tilde{x}, x_k), \lambda\rangle\} = \bar{l}_1 - \bar{l}_2. \tag{4.3.17}$$

Taking $k = k_j$ in (4.3.17) and using the norm-to-norm continuity of $\langle DG(.), \lambda\rangle$, we can derive

$$\bar{l}_1 - \bar{l}_2 = -\langle B_G(\tilde{x}, \hat{x}), \lambda\rangle).$$

Repeating with $k = k_i$, we have

$$\bar{l}_1 - \bar{l}_2 = \langle B_G(\hat{x}, \tilde{x}), \lambda\rangle.$$

It follows that

$$\langle B_G(\hat{x}, \tilde{x}) + B_G(\tilde{x}, \hat{x}), \lambda\rangle = 0.$$

Using (4.2.1), we deduce that

$$\langle (DG(\hat{x}) - DG(\tilde{x}))^\top \lambda, \hat{x} - \tilde{x} \rangle = 0.$$

By C-strict convexity of G, it is easy to check that $\langle \lambda, DG(.) \rangle$ be a strictly monotone operator for any $\lambda \in C_1$, thus we conclude that $\hat{x} = \tilde{x}$, which proves the uniqueness of the cluster point of $\{x_k\}$. The proof is complete. $\square$

4.4 Conclusions

In this chapter, we considered a convex multiobjective optimization problem of a weak Pareto optimal solution for minimizing an extended vector-value map in finite dimensional spaces. We constructed a vector-valued generalized proximal point algorithm based on a strengthened vector-valued Bregman distance function and proved that the sequence generated by this algorithm converges to a weak Pareto optimal solution of the multiobjective optimization problem under the condition that the original optimization problem has a nonempty and compact solution set. Our results are not only some extensions of generalized proximal point algorithms from the scalar case to the vector case but are also some generalizations of vector-valued proximal point algorithm without any exogenously selected vectors in the objective space.

Chapter 5

A Proximal-type Method in Vector Variational Inequalities

5.1 Introduction

The concept of a vector variational inequality problem was firstly introduced by Giannessi [101] in finite dimensional spaces. vector variational inequalities have found many important applications in multiobjective decision making problems, network equilibrium problems, traffic equilibrium problems and so on. These significant applications have made the study of vector variational inequality problems highly attractive Through the last twenty years of development, existence results of solutions, duality theorems and topological properties of solution sets of several kinds of vector variational inequalities have been derived. A complete review of the main results of vector variational inequalities can be found in the monograph [65].

However to the best of our knowledge, no numerical method has yet been designed to solve vector variational inequality problems. Motivated by this situations, in this chapter we firstly attempt to construct a matrix-valued proximal point algorithm for solving a weak vector variational inequality problem (which is an extension of the classical proximal point algorithm proposed by Rockafellar [173]), carry out convergence analysis on the method and prove that the sequence generated by the algorithm converges to a solution of the weak vector variational inequality problem under some mild conditions.

The chapter is organized as follows.

In section 5.2, we present some basic concepts, assumptions and preliminary results. In section 5.3, we introduce a matrix-valued proximal point algorithm and carry out convergence analysis on the algorithm. In section 5.4, we draw the conclusion.

5.2 Subgradients and Normal Mappings

In this section, we present the basic definitions and propositions used in this chapter.

Definition 5.2.1 *[65] Let $F : K \subset R^n \to R^m \cup \{+\infty_C\}$ be a vector-valued mapping. A $n \times m$ matrix V is said to be a weak subgradient of F at $\bar{x} \in K$ if*

$$F(x) - F(\bar{x}) - V^\top(x - \bar{x}) \not\leq_{intC} 0 \quad \forall\, x \in K.$$

Denote by $\partial_C^w F(\bar{x})$ the set of weak subgradients of F on K at $\bar{x}$.

Definition 5.2.2 *[65] Let $F : K \subset R^n \to R^m \cup \{+\infty_C\}$ be a vector-valued mapping. A $n \times m$ matrix V is said to be a strong subgradient of F at $\bar{x} \in K$ if*

$$F(x) - F(\bar{x}) - V^\top(x - \bar{x}) \geq_C 0 \quad \forall\, x \in K.$$

Let $K \subset R^n$ be nonempty, closed and convex. A vector-valued indicator function $\delta(x \mid K)$ of K at x is defined by

$$\delta(x \mid K) = \begin{cases} 0, & \text{if } x \in K; \\ +\infty_C, & \text{if } x \notin K. \end{cases}$$

An important and special case in the theory of weak subgradient is that when $F(x) = \delta(x \mid K)$ becomes a vector-valued indicator function of K. By the Definition 5.2.1, we obtain $V \in \partial_C^w \delta(x^* \mid K)$ if and only if

$$V^\top(x - x^*) \not\geq_{intC} 0 \quad \forall x \in K. \tag{5.2.1}$$

Definition 5.2.3 *A set $VN_K(x^*) \subset R^{n\times m} = \partial_C^w \delta(x^* \mid K)$ is said to be a weak normality operator set to K at x^* if, for every $V \in VN_K(x^*)$, the inequality (5.2.1) holds.*

Clearly, $VN_K(x^*) = \partial_C^w \delta(x^* \mid K)$. As for the scalar case, from [172] we know that $v^* \in \partial \delta_K(x^*) = N_K(x^*)$ if and only if

$$\langle v^*, x - x^* \rangle \leq 0 \quad \forall x \in K \tag{5.2.2}$$

where $\delta_K(x)$ is the scalar-valued indicator function of K. The inequality (5.2.2) means that v^* is normal to K at x^*.

Definition 5.2.4 *Let $VN_K(.) : R^n \rightrightarrows R^{n\times m}$ be a set-valued mapping, which is said to be a weak normal mapping for K, if for any $y \in K$, $V \in VN_K(y)$ such that*

$$V^\top (x - y) \not\geq_{intC} 0, \quad \forall x \in K. \tag{5.2.3}$$

$VN_K(.)$ is said to be strong normal mapping for K, if for any $y \in K$, $V \in VN_K(y)$ such that

$$V^\top (x - y) \leq_C 0, \quad \forall x \in K. \tag{5.2.4}$$

As in [171], the normal mapping for K is a set-valued mapping, which is defined as follows: if for any $y \in K$, $v \in N_K(y)$ such that

$$\langle v, x - y \rangle \leq 0, \quad \forall x \in K.$$

Let $\|A\|_M$ be a matrix norm of the matrix $A \in R^{n\times m}$ (see [156]). In this chapter, we always assume that the matrix norm $\|A\|_M$ is compatible with $\|.\|$, i.e.,

$$\|Ax\| \leq \|A\|_M \|x\|$$

for all $A \in R^{m\times n}$ and $x \in R^n$.

We now introduce a new notion.

Definition 5.2.5 *Let $T : X_0 \to R^{n\times m}$ be a mapping, which is said to be norm sequentially bounded if for any bounded sequence $\{x_k\} \subset X_0$, it holds that the sequence $\{\|T(x_k)\|_M\}$ is bounded.*

Lemma 5.2.1 *[171] Let* K *be a nonempty closed and convex subset of R^n. Let $T_1 : R^n \rightrightarrows R^n$ be the normality mapping to K and $T_2 : R^n \to R^n$ be any single-valued monotone operator such that $K \cap dom(T_2) \neq \emptyset$ and T_2 is continuous on K. Then, we have $T_1 + T_2$ is a maximal monotone operator.*

Lemma 5.2.2 *[175] (*Minty's theorem*) Let $\lambda > 0$ and $T : R^n \rightrightarrows R^n$ be monotone. Then $(I + \lambda T)^{-1}$ is monotone and nonexpansive. Moreover,* T *is maximal monotone if and only if $rge(I + \lambda T) = R^n$. In that case $(I + \lambda T)^{-1}$ is maximal monotone too, and it is a single-valued mapping from all of R^n into itself.*

5.3 Convergence Analysis for a Matrix-valued Proximal-type Method

A weak vector variational inequality (WVVI in short) is a problem of finding $x^* \in X_0$ such that

$$\text{(WVVI)} \qquad T(x^*)^\top (x - x^*) \not\leq_{intC} 0, \quad \forall x \in X_0,$$

where x^* is called a solution of problem (WVVI). Denote by X^* the solution set of problem (WVVI). We also denote by $\hat{X}$ the ideal solution set of problem (WVVI), i.e. for any $x^* \in \hat{X}$, it holds that

$$T(x^*)^\top (x - x^*) \geq_C 0, \quad \forall x \in X_0.$$

Let $\lambda \in C_1$, consider the corresponding scalar-valued variational inequality problem of finding $x^* \in X_0$ such that:

$$(\text{VIP}_\lambda) \qquad \langle \lambda(T)(x^*), x - x^* \rangle \geq 0, \quad \forall x \in X_0.$$

Denote by X^*_λ the solution set of (VIP_λ).

Proposition 5.3.1 *For any $x^* \in X_0$, let $VN_{X_0}(x^*)$ be a weak normality operator set to X_0 at x^* and $V \in VN_{X_0}(x^*)$. Then, there exists a $\lambda \in C_1$ such that $\lambda(V) \in N_{X_0}(x^*)$.*

Proof. By Definition 5.2.3, we know that

$$V^\top(x - x^*) \not\geq_{intC} 0 \quad \forall x \in X_0$$

and

$$V^\top(x - x^*) \in R^m \backslash (intC) \quad \forall x \in X_0.$$

That is

$$V^\top(X_0 - x^*) \subset R^m \backslash (intC)$$

and we have

$$V^\top(X_0 - x^*) \cap intC = \emptyset.$$

By the convexity of X_0, we have that there exists a $\bar{\lambda} \in C \backslash \{0\}$ such that

$$\langle \bar{\lambda}(V), x - x^* \rangle \leq 0, \quad x \in X_0.$$

Since $\|\bar{\lambda}\| > 0$, it follows that

$$\langle \frac{\bar{\lambda}}{\|\bar{\lambda}\|}(V), x - x^* \rangle \leq 0, \quad x \in X_0.$$

Clearly, we have $\frac{\bar{\lambda}}{\|\bar{\lambda}\|} \in C_1$. Without loss of generality, let $\lambda = \frac{\bar{\lambda}}{\|\bar{\lambda}\|}$, we have

$$\langle \lambda(V), x - x^* \rangle \leq 0 \quad \forall x \in X_0.$$

That is $\lambda(V) \in N_{X_0}(x^*)$. The proof is complete. □

Proposition 5.3.2 *[175] The normal mapping $N_{X_0}(x)$ is outer semicontinuous at $\bar{x}$. In other words, if $x_k \to \bar{x}$, $v_k \in N_{X_0}(x_k)$ and $v_k \to \bar{v}$, then $\bar{v} \in N_{X_0}(\bar{x})$*

We denote by $\Phi \subset R^{n\times m}$ the set of the matrices. This satisfies that for any $V \in \Phi$, there exist some $\lambda \in C_1$ such that

$$0 = V\lambda.$$

We propose the following matrix-valued proximal point algorithm (MPPA, in short) for solving problem $(WVVI)$:

Step (1) : Take any $x_0 \in X_0$;

Step (2) : Given any $x_k \in X_0$. If $x_k \in X^*$, then the algorithm stops and let $x_{k+p} = x_k$ for any $p \geq 1$, otherwise goes to step (3);

Step (3) : If $x_k \notin X^*$, then we define x_{k+1} by the following conclusion:

$$T(x_{k+1}) + \varepsilon_k(x_{k+1} - x_k)e_k^\top + VN_{X_0}(x_{k+1}) \subset \Phi \tag{5.3.1}$$

where the sequence $\{e_k\} \subset R^m_{++}$ and $\|e_k\| = 1$, $\varepsilon_k \in (0, \varepsilon]$, $\varepsilon > 0$ and $VN_{X_0}(.)$ is the strong normal mapping to X_0. Go to step (2).

Next we will show the main results of this chapter.

Theorem 5.3.1 *Let $T : X_0 \to R^{n\times m}$ be continuous and C-monotone on X_0, if $domT \cap intX_0 \neq \emptyset$. The sequence $\{x_k\}$ generated by the algorithm (MPPA) is well-defined.*

Proof Let $x_0 \in X_0$ be an initial point and suppose that the algorithm (MPPA) reaches step k. We then show that the next iterate x_{k+1} does exist. Under the assumptions, $T(.)$ is continuous and C-monotone on X_0. By virtue of Proposition 2.2.1, we have $\lambda(T)$ is monotone and continuous on X_0 for any $\lambda \in C_1$. By the definition of strong normality mapping, for any $\lambda \in C_1$, the mapping $VN_{X_0}(x)\lambda$ is a normality mapping to X_0. Thus, by the assumption $\mathrm{dom}T \cap \mathrm{int}X_0 \neq \emptyset$ and Lemma 5.2.1, we obtain that for any $x \in X_0$, the mapping $VN_{X_0}(x)\lambda + T(x)\lambda$ is maximal monotone. Let $\alpha_k = e_k^\top\lambda$. Since $e_k \in R^m_{++}$, obviously the sequence $\{\alpha_k\}$ is positive for every $\lambda \in C_1$. Thus, by Lemma 5.2.2, we conclude that

$$rge\{VN_{X_0}(.)\lambda + T(.)\lambda + \varepsilon_k\alpha_k I(.)\} = R^n.$$

Hence, for any given $\varepsilon_k\alpha_k x_k \in R^n$, there exists a $x_{k+1} \in X_0$ such that

$$\varepsilon_k\alpha_k x_k \in (T + VN_{X_0})(x_{k+1})\lambda + \varepsilon_k\alpha_k x_{k+1} \tag{5.3.2}$$

and

$$0 \in (T + VN_{X_0})(x_{k+1})\lambda + \varepsilon_k\alpha_k x_{k+1} - \varepsilon_k\alpha_k x_k.$$

That is the inclusion (5.3.1) holds. The proof is complete. □

Theorem 5.3.2 *Let the same assumptions as in Theorem 5.3.1 hold. Suppose further that $\hat{X}$ is nonempty. Then, the sequence $\{x_k\}$ generated by the algorithm (MPPA) is bounded.*

Proof. From the algorithm (MPPA), we know that if the algorithm stops at some iteration, the point x_k will be a constant thereafter. Now we assume that the sequence $\{x_k\}$ will not stop after a finite number of iteratives. From (5.3.1), we know that there exist some $\lambda_k \in C_1$ such that

$$0 \in \lambda_k(T)(x_{k+1}) + \varepsilon_k(x_{k+1} - x_k)e_k^\top \lambda_k + VN_{X_0}(x_{k+1})\lambda_k.$$

Thus, there exist some $\gamma_{k+1} \in VN_{X_0}(x_{k+1})$ such that

$$0 = \lambda_k(T)(x_{k+1}) + \varepsilon_k(x_{k+1} - x_k)e_k^\top \lambda_k + \lambda_k(\gamma_{k+1}) \tag{5.3.3}$$

By definition of a strong normal mapping, it holds that $\lambda_k(\gamma_{k+1}) \in N_{X_0}(x_{k+1})$ for any $\lambda_k \in C_1$, that is

$$\langle \lambda_k(\gamma_{k+1}), x - x_{k+1} \rangle \leq 0, \quad \forall\, x \in X_0. \tag{5.3.4}$$

Combining (5.3.3) with (5.3.4), it follows that

$$\langle \lambda_k(T)(x_{k+1}) + \varepsilon_k(x_{k+1} - x_k)e_k^\top \lambda_k, x - x_{k+1} \rangle \geq 0 \quad \forall x \in X_0. \tag{5.3.5}$$

On the other hand, by virtue of the ideal solution set $\hat{X}$ is nonempty, we know that for any given $\lambda_k \in C_1$, the following scalar-valued variational inequality problem (VIP_{λ_k}) has a nonempty solution set, where

$$(VIP_{\lambda_k}) \quad \langle \lambda_k(T)(x^*), x - x^* \rangle \geq 0, \quad \forall\, x \in X_0.$$

Hence, there exists a $x^* \in \hat{X}$ such that x^* is also a solution of problem (VIP_{λ_k}). Hence, we have

$$\langle \lambda_k(T)(x^*), x^* - x_{k+1} \rangle \leq 0.$$

By the C-monotonicity of T, we have that

$$\langle \lambda_k(T)(x_{k+1}), x^* - x_{k+1} \rangle \leq 0. \tag{5.3.6}$$

Combining (5.3.5) with (5.3.6), we obtain that

$$\langle \varepsilon_k {\lambda_k}^\top e_k(x_{k+1} - x_k)^\top, x^* - x_{k+1} \rangle \geq 0.$$

From the proof of Theorem 5.3.1, we know that $\varepsilon_k\langle\lambda_k, e_k\rangle > 0$. It follows that

$$\langle x_{k+1} - x_k, x^* - x_{k+1}\rangle \geq 0$$

$$2\langle x_{k+1} - x_k, x^*\rangle + 2\langle x_k - x_{k+1}, x_{k+1}\rangle \geq 0$$

$$\|x_k\|^2 - 2\langle x_k, x^*\rangle + \|x^*\|^2 - \|x_k\|^2 + 2\langle x_k, x_{k+1}\rangle - \|x_{k+1}\|^2 - \|x_{k+1}\|^2 + 2\langle x_{k+1}, x^*\rangle - \|x^*\|^2 \geq 0.$$

That is

$$\| x_{k+1} - x^* \|^2 \leq \| x_k - x^* \|^2 - \| x_k - x_{k+1} \|^2 . \tag{5.3.7}$$

Clearly, the sequence $\{\| x_k - x^* \|^2\}$ is nonnegative and nonincreasing. Furthermore $\{\| x_k - x^* \|^2\}$ is also bounded below, as denoted by l^* the lower bound of the sequence. By the fact (5.3.7), we have

$$\sum_{k=0}^{\infty} \| x_k - x_{k+1} \|^2 \leq \| x_0 - x^* \|^2 - l^* \leq \| x_0 - x^* \|^2 < \infty$$

and

$$\lim_{k\to+\infty} \| x_k - x_{k+1} \| = 0. \tag{5.3.8}$$

Again from (5.3.7), we obtain that

$$\|x_k - x^*\| \leq \|x_0 - x^*\|$$

for all $x^* \in \hat{X}$. By the nonemptiness of $\hat{X}$, we conclude that $\{x_k\}$ is bounded. The proof is complete. □

Theorem 5.3.3 *Let the same assumptions as in Theorem 5.3.2 hold. We also assume that T is norm sequentially bounded. Then any accumulation point of $\{x_k\}$ is a solution of problem $(WVVI)$.*

Proof. If there exists $k_0 \geq 1$ such that $x_{k_0+p} = x_{k_0}, \forall p \geq 1$. Then, it is clear that x_{k_0} is the unique cluster point of $\{x_k\}$ and it is also a solution of problem $(WVVI)$. Suppose that the algorithm does not terminate finitely. Then, by Theorem 5.3.2, we have that $\{x_k\}$ is bounded and it has some cluster points. Next we show that all of cluster points are solutions of problem $(WVVI)$. Let $\hat{x}$ be a cluster points of $\{x_k\}$ and $\{x_{k_j}\}$ be a subsequence of $\{x_k\}$, which converges to $\hat{x}$. From the limit (5.3.8), we know that $\lim_{j\to+\infty} \| x_{k_j+1} - x_{k_j} \|=0$. That is $x_{k_j+1} \to \hat{x}$ as $j \to +\infty$. By formula (5.3.1), we derive that there exist $\gamma_{k_j+1} \in VN_{X_0}(x_{k_j+1})$ and $\lambda_{k_j} \in C_1$, such that

$$\lambda_{k_j}(T)(x_{k_j+1}) + \lambda_{k_j}(\gamma_{k_j+1}) + \varepsilon_{k_j}(x_{k_j+1} - x_{k_j})e_{k_j}^{\top}\lambda_{k_j} = 0.$$

That is

$$\|\lambda_{k_j}(T)(x_{k_j+1}) + \lambda_{k_j}(\gamma_{k_j+1}) + \varepsilon_{k_j}(x_{k_j+1} - x_{k_j})e_{k_j}^{\top}\lambda_{k_j}\| = 0.$$

It follows that

$$\begin{aligned} 0 \geq &\| \lambda_{k_j}(T)(x_{k_j+1}) + \lambda_{k_j}(\gamma_{k_j+1}) \| - \| \varepsilon_{k_j}(x_{k_j} - x_{k_j+1})e_{k_j}^{\top}\lambda_{k_j} \| \\ =&\| \lambda_{k_j}(T)(x_{k_j+1}) + \lambda_{k_j}(\gamma_{k_j+1}) \| - \varepsilon_{k_j}\alpha_{k_j} \| x_{k_j} - x_{k_j+1} \|, \end{aligned} \tag{5.3.9}$$

where $\alpha_{k_j} = e_{k_j}^{\top}\lambda_{k_j}$. From (5.3.8), we know that $\lim_{j\to+\infty} \| x_{k_j} - x_{k_j+1} \|= 0$. Since $\lambda_{k_j} \in C_1$, by the compactness of C_1, we know that the sequence $\{\lambda_{k_j}\}$ has a convergent subsequence. Without loss of generality, we assume that $\lambda_{k_j} \to \bar{\lambda}$. Furthermore we have $\bar{\lambda} \in C_1$ and $\bar{\lambda} \neq 0$. Thus, taking the limit in (5.3.9), we deduce the following:

$$\lim_{j\to+\infty} \|\lambda_{k_j}(T)(x_{k_j+1}) + \lambda_{k_j}(\gamma_{k_j+1})\| = 0. \tag{5.3.10}$$

We claim that the sequence $\{\lambda_{k_j}(\gamma_{k_j+1})\}$ is bounded. Suppose that, in contrast, without loss of generality, we assume that $\|\lambda_{k_j}(\gamma_{k_j+1})\| \to +\infty$ and $\frac{\lambda_{k_j}(\gamma_{k_j+1})}{\|\lambda_{k_j}(\gamma_{k_j+1})\|} \to \bar{\omega} \in R^n$ and $\bar{\omega} \neq 0$. From (5.3.10), we know that

$$0 = \lim_{j\to+\infty} \frac{\|\lambda_{k_j}(T)(x_{k_j+1}) + \lambda_{k_j}(\gamma_{k_j+1})\|}{\|\lambda_{k_j}(\gamma_{k_j+1})\|} = \lim_{j\to+\infty} \| \frac{\lambda_{k_j}(T)(x_{k_j+1})}{\|\lambda_{k_j}(\gamma_{k_j+1})\|} + \frac{\lambda_{k_j}(\gamma_{k_j+1})}{\|\lambda_{k_j}(\gamma_{k_j+1})\|} \|= \|0+\bar{\omega}\|, \tag{5.3.11}$$

since T is norm sequentially bounded, which yields that

$$\|\lambda_{k_j}(T)(x_{k_j+1})\| \leq \|T(x_{k_j+1})\|_M \|\lambda_{k_j}\| = \|T(x_{k_j+1})\|_M \leq \mu < +\infty$$

for some $\mu > 0$. Obviously, the equality (5.3.11) contradicts with the assumption $\bar{\omega} \neq 0$. Thus, the sequence $\{\lambda_{k_j}(\gamma_{k_j+1})\}$ is bounded. Without loss of generality, we assume that $\lambda_{k_j}(\gamma_{k_j+1}) \to \hat{\omega} \in R^n$. Furthermore, from (5.3.10) and the continuity of T, we derive that

$$\|\bar{\lambda}(T)(\hat{x}) + \hat{\omega}\| = 0.$$

Hence, we have

$$\bar{\lambda}(T)(\hat{x}) + \hat{\omega} = 0.$$

Meanwhile, from the definition of strong normal mapping and Proposition 5.3.2, we have $\hat{\omega} \in N_{X_0}(\hat{x})$. By the definition of $N_{X_0}(\hat{x})$, we know that

$$\langle \hat{\omega}, x - \hat{x} \rangle \leq 0 \quad \forall x \in X_0.$$

That is

$$\langle \bar{\lambda}(T)(\hat{x}), x - \hat{x} \rangle \geq 0 \quad \forall x \in X_0. \tag{5.3.12}$$

Thus

$$T(\hat{x})^{\top}(x - \hat{x}) \notin -intC \quad \forall x \in X_0. \tag{5.3.13}$$

We conclude that $\hat{x}$ is a solution of problem $(WVVI)$. The proof is complete. □

Theorem 5.3.4 *Let the same assumptions as those in Theorem 5.3.3 hold. Suppose further that $X^* = \hat{X}$, the whole sequence $\{x_k\}$ converges to a solution of problem $(WVVI)$.*

Proof Suppose that, in contrast, both $\hat{x}$ and $\tilde{x}$ are two distinct cluster points of $\{x_k\}$ and

$$\lim_{j \to +\infty} x_{k_j} = \hat{x}, \quad \lim_{i \to +\infty} x_{k_i} = \tilde{x}.$$

From Theorem 5.3.3, we know that $\hat{x}$ and $\tilde{x}$ are solutions of problem $(WVVI)$. By virtue of Theorem 5.3.1 and the proof of Theorem 5.3.2, we know that there exist $\hat{\lambda}$ and $\tilde{\lambda} \in C_1$ such that

$$\langle \hat{\lambda}(T)(\hat{x}), \hat{x} - x_{k+1} \rangle \leq 0, \quad \langle \tilde{\lambda}(T)(\tilde{x}), \tilde{x} - x_{k+1} \rangle \leq 0. \tag{5.3.14}$$

By the C-monotonicity of T, we have

$$\langle \hat{\lambda}(T)(x_{k+1}), \hat{x} - x_{k+1} \rangle \leq 0, \quad \langle \tilde{\lambda}(T)(x_{k+1}), \tilde{x} - x_{k+1} \rangle \leq 0. \tag{5.3.15}$$

From (5.3.6), we obtain that

$$\langle x_{k+1} - x_k, \hat{x} - x_{k+1} \rangle \geq 0, \quad \langle x_{k+1} - x_k, \tilde{x} - x_{k+1} \rangle \geq 0. \tag{5.3.16}$$

Similar to (5.3.7), we obtain that

$$\| x_{k+1} - \hat{x} \|^2 \leq \| x_k - \hat{x} \|^2 - \| x_k - x_{k+1} \|^2, \tag{5.3.17}$$

and

$$\| x_{k+1} - \tilde{x} \|^2 \leq \| x_k - \tilde{x} \|^2 - \| x_k - x_{k+1} \|^2 . \tag{5.3.18}$$

Combining (5.3.17) with (5.3.18), we obtain that both sequences $\{\| x_k - \hat{x} \|^2\}$ and $\{\| x_k - \tilde{x} \|^2\}$ are nonnegative and nonincreasing, hence they are convergent. So there exist $\hat{\beta}$, $\tilde{\beta} \in R$ such that

$$\lim_{k \to +\infty} \| x_k - \tilde{x} \| = \tilde{\beta}, \quad \lim_{k \to +\infty} \| x_k - \hat{x} \| = \hat{\beta}. \tag{5.3.19}$$

Clearly, we have

$$\| x_k - \hat{x} \|^2 = \| x_k - \tilde{x} \|^2 + 2\langle x_k - \tilde{x}, \tilde{x} - \hat{x} \rangle + \| \tilde{x} - \hat{x} \|^2 . \tag{5.3.20}$$

Combining (5.3.19) with (5.3.20), we deduce the following

$$\lim_{k \to +\infty} \langle x_k - \tilde{x}, \tilde{x} - \hat{x} \rangle = \frac{1}{2}(\hat{\beta}^2 - \tilde{\beta}^2 - \| \tilde{x} - \hat{x} \|^2). \tag{5.3.21}$$

Taking $k = k_i$ in (5.3.21), we obtain that

$$\hat{\beta}^2 - \tilde{\beta}^2 = \| \tilde{x} - \hat{x} \|^2 .$$

Changing the places of $\hat{x}$ and $\tilde{x}$ in (5.3.20) and repeating $k = k_j$ in (5.3.21), we have that

$$\| \tilde{x} - \hat{x} \|^2 = \tilde{\beta}^2 - \hat{\beta}^2.$$

Thus, we conclude that

$$\| \tilde{x} - \hat{x} \| = 0,$$

which establishes the uniqueness of the cluster points of $\{x_k\}$. The proof is complete. □

5.4 Conclusions

In this chapter, we formulated a matrix-valued proximal point algorithm to solve the weak vector variational inequality problem with respect to the positive orthant in finite dimensional spaces by virtue of normal mappings, carried out convergence analysis on the method and proved the convergence of the generated sequence to a solution of the weak vector variational inequality problem under some mild conditions.

Chapter 6

Asymptotic Analysis for Parametric Multiobjective Optimization Problems

6.1 Introduction

The main idea of some existing algorithms for scalar-valued optimization problems (e.g., proximal-type methods [173], Tikhonov-type regularization algorithms [193], viscosity approximate methods [157] and so on) is to solve a sequence of subproblems instead of solving the origin problem. Thus, the nonemptiness and compactness of solution set of the subproblems is significant in both theory and methodology. What is worth noting is that it is important to guarantee the boundedness of the sequences of solutions obtained by solving subproblems for scalar-valued optimization problems and variational inequality problems (e.g., [9, 10, 11, 12, 17, 16]). Finding sufficient conditions, particularly the necessary and sufficient conditions, which are easy to verify, for the nonemptiness and compactness of the solution set of optimization problems is significant.

In recent years, the study of numerical methods for solving multiobjective optimization problems has received extensive attention. Motivated by the goal of designing more efficient algorithms to solve more complicated multiobjective optimization models, investigating the nonemptiness and compactness of solution sets of multiobjective optimization problems is important and necessary. Deng [73, 76] obtained the necessary and sufficient conditions for the nonemptiness and compactness of solution sets of convex vector optimization problems. Huang and Yang [117] gave characterizations for the nonemptiness and compactness of the set of weak Pareto optimal solutions of an unconstrained/constrained convex vector optimization problem with extended vector-valued functions in terms of the 0-coercivity of some scalar functions. Flores-Bazan [91] established existence results for finite dimensional vector optimization problems based on the asymptotic description of the functions and sets.

In this chapter, we are concerned with the extended-valued parametric multiobjective optimization problem, and we obtain the necessary and sufficient conditions for the nonemptiness and compactness of the weak Pareto optimal solution set of the problem by virtue of asymptotical analysis.

This chapter is organized as follows.

In section 6.2, we present the concepts, basic assumptions and preliminary results. In section 6.3, we propose various necessary and sufficient conditions for the nonemptiness and compactness of the weak Pareto optimal solution set of the parametric multiobjective optimization problem. In section 6.4, we draw the conclusion.

6.2 Parametric Multiobjective Optimization

In this section, we present the basic definitions and propositions used later in this chapter.

Firstly we consider the following extended-valued multiobjective optimization problem with functional constraints,

$$(MOP_{FC}) \qquad \mathrm{Min}_C \quad F_0(x)$$
$$\text{s.t.} \quad x \in S_0 = \{x \in R^n |\ G(x) \leq_D 0\},$$

where $F_0 : R^n \to R^m \cup \{+\infty_C\}$ is a vector-valued function, $(F_0)_i$ is the ith component of F_0 and $G : R^n \to R^l$ is also a vector-valued function, G_j is the jth component of G. Let $D = R^l_+ \subset R^l$ be the positive orthant that defines a partial order in R^l as follows:

for any $y_1, y_2 \in R^l$,

$$y_1 \leq_D y_2 \qquad \text{if and only if } y_2 - y_1 \in D;$$
$$y_1 \nleq_{intD} y_2 \qquad \text{if and only if } y_2 - y_1 \notin intD.$$

Denote by $\bar{X}_{FC}$ the weak Pareto optimal solution set of problem (MOP_{FC}). Let $F : R^n \times R^l \to R^m \cup \{+\infty_C\}$ be a vector-valued perturbed function such that

$$F(x,u) = \begin{cases} F_0(x), & \text{if } G(x) \leq_D u; \\ +\infty_C, & \text{else.} \end{cases}$$

Defining a family of perturbed problems with $F(x,u)$:

$$(MOP_u) \qquad \mathrm{Inf}_C\{F(x,u)|\ x \in R^n\}.$$

Let $A \subset R^l$ be a nonempty set. By $z^* \in \mathrm{Inf}_C A$, we mean that

(1) $z^* \in R^m \cup \{+\infty_C\}$;

(2) $z \nleq_{C\setminus\{0\}} z^*, \forall z \in A$;

(3) $\exists z_k \in A$ such that $z_k \to z^*$.

Clearly, the primal problem (MOP_{FC}) is identical to problem (MOP_u) with $u = 0$. Define the optimal value function by

$$P(u) = \mathrm{Inf}_C\{F(x,u)|\ x \in R^n\}, \quad u \in R^l.$$

Denote by $domP$ the domain of P and $\bar{X}_u$ the weak Pareto optimal solution set of problem (MOP_u) for any $u \in domP$.

Definition 6.2.1 *[175] A function $f : R^n \times R^m \to R \cup \{+\infty\}$ with value $f(x,u)$ is said to be level-bounded in x locally uniformly in u if for each $\bar{u} \in R^m$ and $a \in R$ there is a neighborhood $V \in N(\bar{u})$ along with a bounded set $B \subset R^n$ such that $\{x|\ f(x,u) \leq a\} \subset B$ for all $u \in V$.*

Proposition 6.2.1 *[175] For any collection of sets $K_i \subset R^n$ for $i \in I$, and an arbitrary index set, one has*

$$[\bigcap_{i\in I} K_i]^\infty \subset \bigcap_{i\in I} K_i^\infty, \quad [\bigcup_{i\in I} K_i]^\infty \supset \bigcup_{i\in I} K_i^\infty.$$

The first inclusion holds as an equation for a closed and convex, K_i with nonempty intersection. The second holds as an equation when I is finite.

Proposition 6.2.2 *[175] For any function $f : R^n \to \bar{R}$ and any $\alpha \in R$, one has*

$$\{x|\ f(x) \leq \alpha\}^\infty \subset \{x|\ f^\infty(x) \leq 0\}.$$

This is an equation when f is convex, lsc and proper, and $\{x|\ f(x) \leq \alpha\} \neq \emptyset$.

We state ([175], Theorem 1.17) as the following lemma.

Lemma 6.2.1 *Let's consider*

$$p(u) := \inf_x f(x, u), \quad x(u) := \arg\min_x f(x, u),$$

in the case of a proper and lsc $f : R^n \times R^m \to R \cup \{+\infty\}$ such that $f(x, u)$ is level-bounded in x locally uniformly in u. Then, it holds that the function p is proper and lsc on R^m, and for any $u \in domp$ the set $x(u)$ is nonempty and compact.

Lemma 6.2.2 *(Theorem 3.31 [175]) Let $f : R^n \times R^l \to R \cup \{+\infty\}$ be a proper, convex and lsc function. Then, $f(x, u)$ is level-bounded in x uniformly in u if and only if*

$$f^\infty(d, 0) > 0, \quad \forall\ d \neq 0. \tag{6.2.1}$$

If this is fulfilled, the function $p(u) := \inf_x f(x, u)$ has

$$p^\infty(u) = \inf_x f^\infty(x, u), \tag{6.2.2}$$

attained when finite.

6.3 Asymptotical Analysis for Multiobjective Optimization Problems

Theorem 6.3.1 *Let's consider the problem (MOP_{FC}). Suppose that F_0 is proper, C-convex and C-lsc, and G is proper, D-convex and D-lsc. Then, $\bar{X}_{FC}$ is nonempty and compact if and only if*

$$\cap_{i=1}^l \{d \in R^n|\ G_i^\infty(d) \leq 0\} \bigcap \cup_{j=1}^m \{d \in R^n|\ (F_0)_j^\infty(d) \leq 0\} = \{0\} \tag{6.3.1}$$

Proof. From the assumptions, we know that G is proper, D-convex and D-lsc. That is, G_i is convex and lsc. Through the problem (MOP_{FC}), we obtain that

$$S_0 = \{x \in R^n |\ G(x) \leq_D 0\} = \{x \in R^n |\ G_i(x) \leq 0, \forall\ i \in [1, ..., l]\}.$$

By Lemma 2.1.1, we have S_0 is nonempty, closed and convex. Let $(S_0)_i = \{x \in R^n |\ G_i(x) \leq 0\}$, hence we have

$$\cap_{i=1}^{l} (S_0)_i = S_0$$

and $(S_0)_i$ is also closed and convex for every $i \in [1, ..., l]$. By Proposition 6.2.1, we know that

$$S_0^\infty = (\cap_{i=1}^{l} (S_0)_i)^\infty = \cap_{i=1}^{l} (S_0)_i^\infty.$$

From Proposition 6.2.2, we obtain

$$(S_0)_i^\infty = \{d \in R^n |\ G_i^\infty(d) \leq 0\}.$$

It follows that

$$S_0^\infty = \cap_{i=1}^{l} \{d \in R^n |\ G_i^\infty(d) \leq 0\}. \tag{6.3.2}$$

By virtue of Lemma 2.3.3, we have $\bar{X}_{FC}$ is nonempty and compact if and only if

$$S_0^\infty \bigcap \cup_{j=1}^{m} \{d \in R^n |\ (F_0)_j^\infty(d) \leq 0\} = \{0\}. \tag{6.3.3}$$

Combining (6.3.2) with (6.3.3), we obtain that

$$\cap_{i=1}^{l} \{d \in R^n |\ G_i^\infty(d) \leq 0\} \bigcap \cup_{j=1}^{m} \{d \in R^n |\ (F_0)_j^\infty(d) \leq 0\} = \{0\}.$$

The proof is complete. □

Proposition 6.3.1 *Let the assumptions in Theorem 6.3.1 hold. Then, $F(x, u)$ is C-convex on $R^n \times R^l$.*

Proof. From the definition, we know that

$$F(x, u) = \begin{cases} F_0(x), & \text{if } G(x) \leq_D u; \\ +\infty, & \text{else.} \end{cases}$$

For any $(x_1, u_1), (x_2, u_2) \in R^n \times R^l$, we have that

$$\alpha F(x_1, u_1) = \begin{cases} \alpha F_0(x_1), & \text{if } G(x_1) \leq_D u_1; \\ +\infty, & \text{else.} \end{cases}$$

and

$$(1-\alpha)F(x_2,u_2)=\begin{cases}(1-\alpha)F_0(x_2), & \text{if } G(x_2)\leq_D u_2;\\ +\infty, & \text{else.}\end{cases}$$

That is

$$\alpha F(x_1,u_1)+(1-\alpha)F(x_2,u_2)=\begin{cases}\alpha F_0(x_1)+(1-\alpha)F_0(x_2), & \text{if } G(x_1)\leq_D u_1,\\ & \quad G(x_2)\leq_D u_2;\\ +\infty, & \text{else.}\end{cases}$$

From the D-convexity of G, we obtain that

$$G(\alpha x_1+(1-\alpha)x_2)\leq_D \alpha G(x_1)+(1-\alpha)G(x_2). \tag{6.3.4}$$

If $G(x_1)\leq_D u_1$ and $G(x_2)\leq_D u_2$, the inequality (6.3.4) can be rewritten as follows

$$G(\alpha x_1+(1-\alpha)x_2)\leq_D \alpha G(x_1)+(1-\alpha)G(x_2)\leq_D \alpha u_1+(1-\alpha)u_2. \tag{6.3.5}$$

From definition of $F(x,u)$, we have

$$F(\alpha x_1+(1-\alpha)x_2,\alpha u_1+(1-\alpha)u_2)=\begin{cases}F_0(\alpha x_1+(1-\alpha)x_2), & \text{if } G(\alpha x_1+(1-\alpha)x_2)\\ & \quad \leq_D \alpha u_1+(1-\alpha)u_2;\\ +\infty, & \text{else.}\end{cases}$$

On the other hand, by the C-convexity of F_0, we obtain that

$$F_0(\alpha x_1+(1-\alpha)x_2)\leq_C \alpha F_0(x_1)+(1-\alpha)F_0(x_2). \tag{6.3.6}$$

Thus, we obtain that

$$F(\alpha x_1+(1-\alpha)x_2,\alpha u_1+(1-\alpha)u_2)=F_0(\alpha x_1+(1-\alpha)x_2)$$

$$\leq_C \alpha F_0(x_1)+(1-\alpha)F_0(x_2)=\alpha F(x_1,u_1)+(1-\alpha)F(x_2,u_2)$$

for any $(x_1,u_1),(x_2,u_2)\in R^n\times R^l$, if $G(x_1)\leq_D u_1$ and $G(x_2)\leq_D u_2$. That is, $F(x,u)$ is C-convex on $R^n\times R^l$. The proof is complete. □

Proposition 6.3.2 *Let the assumptions in Proposition 6.3.1 hold and $S: R^l \rightrightarrows R^n$ be a set-valued mapping, where*

$$S(u)=\{x\in R^n|\ G(x)\leq_D u\}.$$

Then, S is outer semicontinuous.

Proof. Let's consider the set

$$gph(S) = \{(u,x) \in R^l \times R^n |\ x \in S(u)\}$$

$$= \{(u,x) \in R^l \times R^n |\ G_i(x) \le u_i, \forall\ i \in [1,...,l]\}.$$

We claim that $gph(S)$ is closed on $R^l \times R^n$. Without loss of generality, we assume that $\{u^k\} \subset R^l$, $\{x^k\} \subset R^n$ such that $(u^k, x^k) \in gph(S)$ and

$$u^k \to \bar{u}, \quad x^k \to \bar{x}. \tag{6.3.7}$$

Denote by u_i^k the *ith* components of u^k. That is for any $i \in [1,...,l]$, we have $u_i^k \to \bar{u}_i$. For any $\epsilon > 0$ and $i \in [1,...,l]$

$$G_i(x^k) \le u_i^k + \epsilon \quad \forall\ k = 1, 2, ...$$

and

$$\liminf_{x^k \to \bar{x}} G_i(x^k) \le G_i(x^k) \le u_i^k + \epsilon, \quad \forall\ k = 1, 2, ... \tag{6.3.8}$$

As $G_i(x)$ is lsc, we obtain

$$\liminf_{x^k \to \bar{x}} G_i(x^k) \ge G_i(\bar{x}). \tag{6.3.9}$$

Combining (6.3.7) and (6.3.8) with (6.3.9), we obtain that

$$\bar{u}_i + \epsilon \ge G_i(\bar{x}), \quad \forall\ i \in [1,...,l].$$

The arbitrariness of ϵ guarantees that

$$\bar{u}_i \ge G_i(\bar{x}), \quad \forall\ i \in [1,...,l].$$

That is

$$G(\bar{x}) \le_D \bar{u},$$

which means that $\bar{x} \in S(\bar{u})$. Thus, $gph(S)$ is closed. By virtue of Lemma 2.3.2, we conclude that the set-valued mapping S is outer semicontinuous. The proof is complete. □

Proposition 6.3.3 *Let $F : R^n \times R^l \to R^m \cup \{+\infty_C\}$ be a perturbed function such that*

$$F(x,u) = \begin{cases} F_0(x), & \textit{if } G(x) \le_D u; \\ +\infty, & \textit{else.} \end{cases}$$

If the assumptions in Proposition 6.3.2 hold, then $F(x,u)$ is C-lsc on $R^n \times R^l$.

Proof From Remark 2.1.1, we know that $F(x,u)$ is C-lsc if and only if all components of $F(x,u)$ are lsc on $R^n \times R^l$. Thus, we only need to check that for any $j \in [1,...,m]$, the following scalar-valued perturbed function is lsc,

$$f_j(x,u) = \begin{cases} (F_0)_j(x), & \text{if } G(x) \leq_D u; \\ +\infty, & \text{else.} \end{cases}$$

We have

$$f_j(x,u) = (F_0)_j(x) + \delta_{S(u)}(x), \tag{6.3.10}$$

where $\delta_{S(u)}(x)$ is an indicator function of set $S(u)$, $i.e.$

$$\delta_{S(u)}(x) = \begin{cases} 0, & \text{if } G(x) \leq_D u; \\ +\infty, & \text{else.} \end{cases}$$

Let $\{x_k\} \subset R^n$ and $\{u_k\} \subset R^l$ be two sequences such that

$$x_k \longrightarrow \bar{x}, \quad u_k \longrightarrow \bar{u}$$

as $k \longrightarrow +\infty$. From (6.3.10), we have

$$\liminf_{k\to+\infty} f_j(x_k,u_k) = \liminf_{k\to+\infty}\{(F_0)_j(x_k) + \delta_{S(u_k)}(x_k)\}, \quad \forall\, j \in [1,...,m].$$

By Proposition 2.3.2, we obtain that

$$\liminf_{k\to+\infty} f_j(x_k,u_k) \geq \liminf_{k\to+\infty}(F_0)_j(x_k) + \liminf_{k\to+\infty} \delta_{S(u_k)}(x_k), \quad \forall\, j \in [1,...,m]. \tag{6.3.11}$$

From Proposition 6.3.2, we know that $S(u_k)$ is outer semicontinuous as $k \longrightarrow +\infty$, that is

$$\limsup_{u_k\to\bar{u}} S(u_k) \subset S(\bar{u})$$

and

$$\bar{x} \in S(\bar{u}).$$

It follows that

$$\liminf_{k\to+\infty} \delta_{S(u_k)}(x_k) = \delta_{S(\bar{u})}(\bar{x}) = 0. \tag{6.3.12}$$

By the C-lower semiconinuouity of F_0, we know that $(F_0)_j(x)$ is lsc for any $j \in [1,...,m]$. That is

$$\liminf_{k\to+\infty}(F_0)_j(x_k) \geq (F_0)_j(\bar{x}), \quad \forall\, j \in [1,...,m]. \tag{6.3.13}$$

Combining (6.3.11) and (6.3.12) with (6.3.13), we have

$$\liminf_{k\to+\infty} f_j(x_k, u_k) \geq (F_0)_j(\bar{x}), \quad \forall\, j \in [1, ..., m].$$

Thus, we obtain that $f_j(x_k, u_k)$ is lsc on $R^n \times R^l$ for any $j \in [1, ..., m]$. Clearly, $F(x, u)$ is C-lsc on $R^n \times R^l$.

Theorem 6.3.2 *Let the assumptions in Proposition 6.3.3 hold. Suppose further that* $\bar{X}_{FC}$ *is nonempty and compact. Then for any* $u \in domP$, $\bar{X}_u$ *is nonempty and compact.*

Proof. From the assumption, we know that $\bar{X}_{FC}$ is nonempty and compact, that is the following problem (MOP_0),

$$(MOP_0) \qquad \mathrm{Inf}_C\{F(x, 0)|\ x \in R^n\}.$$

has a nonempty and compact weak Pareto optimal solution set. By virtue of Lemma 2.3.3, we have

$$\{d \in R^n|\ f_i^\infty(d, 0) \leq 0\} = \{0\}, \quad \forall\, i \in [1, ..., m]. \tag{6.3.14}$$

Denote by $(\bar{X}_u)_i$ the solution set of the following scalar-valued optimization problem

$$\inf\{f_i(x, u)|\ x \in R^n\}, \quad u \in R^l$$

where $f_i(x, u)$ is the *ith* component of $F(x, u)$. From Proposition 6.3.1 and Proposition 6.3.3, we know that $F(x, u)$ is proper, C-convex and C-lsc on $R^n \times R^l$, which means that all components of $F(x, u)$ are proper, convex and lsc. By virtue of Lemma 6.2.2 and (6.3.14), we obtain that $f_i(x, u)$ is level-bounded in x and locally uniformly in u. We also see that all functions $f_i(\cdot, u)$ on R^n have $f_i^\infty(\cdot, u)$ as their asymptotic functions. By Lemma 6.2.1, we conclude that $(\bar{X}_u)_i$ is nonempty and compact. Thus, $\bar{X}_u$ is nonempty and compact for any $u \in domP$.

6.4 Conclusions

In this chapter, we investigated the asymptotical properties of parametric multiobjective optimization. Various necessary and sufficient conditions were given for the nonemptiness and compactness of the weak Pareto optimal solution set of a convex parametric multiobjective optimization problem.

Chapter 7

Conclusions and Suggestions for Future Studies

In this thesis, we studied the asymptotical properties of multiobjective optimization and vector variational inequalities. Based on these asymptotical properties, we constructed some proximal-type methods for solving convex multiobjective optimization problems and weak vector variational inequality problems.

In chapter 3, we considered a convex vector optimization problem of finding weak Pareto optimal solutions for an extended vector-valued map from a uniformly convex and uniformly smooth Banach space to a real Banach space with the latter being ordered by a closed, convex and pointed cone with a nonempty interior. For this problem, we developed an extension of the proximal point method for scalar-valued convex optimization as well as some proximal point algorithms in vector optimization. In this extension, the subproblems involve the finding of weak Pareto optimal solutions for suitable regularizations of the original map. We presented both exact and inexact versions, where the subproblems are solved only approximately within a relative tolerance. In both cases, we proved weak convergence of the sequence generated to a weak Pareto optimal solution, by assuming convexity of the map with respect to C and C-completeness of the initial section.

In chapter 4, we also proposed a generalized proximal point algorithm for finding a weak Pareto optimal solution for minimizing an extended vector-valued map with respect to the positive orthant in finite dimensional spaces through the vector-valued Bregman distance function. We proved that the sequence generated by this algorithm converges to a weak Pareto optimal solution of the multiobjective optimization problem under the condition that the original optimization problem has a nonempty and compact solution set.

In chapter 5, we constructed a matrix-valued proximal-type method to solve a monotone-type weak vector variational inequality, carried out convergence analysis on the method, and proved that the sequence generated by our method converges to a solution of the weak vector variational inequality problem under some mild conditions.

In chapter 6, we investigated the nonemptiness and compactness of the weak Pareto optimal solution set of a convex multiobjective optimization problem with functional constraints via asymptotic analysis. We also employed the obtained results to derive the necessary and sufficient conditions of the weak Pareto optimal solution set of a convex parametric multiobjective optimization problem.

Overall, we obtained some new results and methods for the theory of multiobjective optimization problems and vector variational inequality problems. Some of our results (e.g., Chapter 3 and Chapter 6) include the corresponding results studied by others as special cases; and some of our results (e.g., Chapter 4 and Chapter 5) are original.

However, some of our results are quite abstract, and they are difficult to make numerical tests, due to the fact that the proximal-type method is a conceptual scheme rather than an implementable algorithm. It merely transforms a given problem into a sequence of better behaved subproblems. Thus, the performance of the method depends on the specific algorithm used to solve the subproblems. In this situation, it makes little sense to compare the proximal-type method with other methods in terms of computational efficiency unless a specific procedure is chosen to solve the subproblems. In this thesis, we refrained from discussing the algorithms to solve the subproblems; hence, we did not discuss the implementation issues or comparisons with alternative approaches.

The following is a list of some interesting problems to be dealt with in future research:

[1]. Although the proximal point method is difficult to apply to some practical problems, we should find ways to overcome these difficulties in future research and select some suitable examples for numerical test by choosing a specific procedure.

[2]. In chapter 6, we have investigated the convex multiobjective optimization problem with functional constraints and derived the necessary and sufficient conditions for nonemptiness and compactness of weak Pareto optimal solution set. Thus, we will apply these results to construct some vector-valued primal-dual proximal point methods to solve the multiobjective optimization problems with functional constraints.

[3]. We will attempt to propose some proximal-type methods to solve the set-valued vector variational inequality problem.

[4]. We will attempt to characterize convex composite multiobjective optimization problems, and derive the necessary and sufficient conditions of nonemptiness and compactness of weak Pareto optimal solution set of the problems. We will then apply the obtained results to propose some proximal-type methods to solve the convex composite multiobjective optimization problems.

In studying those problems, we will obtain some new results by using the methods introduced by others, or by introducing new methods to deal with these problems. We will focus more on these and other related problems. We will also intend to obtain more useful results.

Bibliography

[1] B. Aghezzaf and M. Hachimi. Second-order optimality conditions in multiobjective optimization problems, Journal of Optimization Theory and Applications, 102 (1999) 37-50.

[2] Y. Alber and I. Ryazantseva, Nonlinear ill-posed problems of monotone type, Springer-Verlag, Dordrecht, 2006.

[3] E. L. Allgower, K. Böhmer, F. A. Potra and W. C. Rheinboldt, A mesh independence principle for operator equations and their discretizations, SIAM J. Numer. Anal., 23 (1986) 160-169.

[4] E. L. Allgower and K. Böhmer, Application of the mesh independence principle to mesh refinement strategies, SIAM J. Numer. Anal., 24 (1987) 1335-1351.

[5] Q. H. Ansari and J. C. Yao, An existence result for the generalized vector equilibrium problem, Applied Mathematics Letters, 12 (1999) 53-56.

[6] H. Attouch, Convergence of vicosity type approximation methods to particular solutions of variational problems, In Proceeding of Variational Methods Nonlinear Analysis and Differential Equations, 13-43, 1994.

[7] H. Attouch, Variational Convergence for Functions and Operators, Pitman Publishing Inc, 1984.

[8] H. Attouch and H. Riahi, Stability results for Ekeland epsilonvariational principle and cone extremal solutions, Mathematics of Operations Research, 18 (1993) 173-201.

[9] A. Auslander, Noncoercive optimization problem, Mathematics of Operations Research, 21 (1996) 769-781.

[10] A. Auslander, Asymptotic analysis for penalty and barrier methods in variational inequalities, SIAM Journal on Control and Optimization, 37 (1999) 653-671.

[11] A. Auslander, Variational inequalities over the cone of semidefinite positive symmetric matrices and over the Lorentz cone, the Second Japanese-Sino Optimization Meeting, Part II, Kyoto, 2002, Optimization Methods and Software, 18 (2003) 359-376.

[12] A. Auslander and R. Correa, Primal and dual stability results for variational inequalities, Computational Optimization and Applications, 17 (2000) 117-130.

[13] A. Auslander and M. Haddou, An interior proximal point method for convex linearly constrained problems and its extension to variational inequalities, Mathematical Programming, 71 (1995) 77-100.

[14] A. Auslender, M. Teboulle, and S. Ben Tiba, Interior proximal and multiplier methods based on second order homogeneous kernels, Mathematics of Operations Research, 24 (1999) 645-668.

[15] A. Auslender and M. Teboulle, Lagrangian duality and related methods for variational inequalities, SIAM Journal of Optimization, 10 (2000) 1097-1115.

[16] A. Auslender and M. Teboulle, Entropic proximal decomposition methods for convex programs and variational inequalities, Mathematical Programming, 91 (2001) 33-47.

[17] A. Auslander and M. Teboulle, Asymptotic cones and functions in optimization and variational inequalities, Springer 2003.

[18] A. Balbas and P. Jimenez Guerra. Sensitivity analysis for convex multiobjective programming in abstract spaces, Journal of Mathematical Analysis and Applications, 202 (1996) 645-658.

[19] H. H. Bauschke and J. M. Borwein, Lengendre function and the method of random Bregman functions, Journal of Convex Analysis, 4 (1997) 27-67.

[20] E. M. Bednarczuk, An approach to well-posedness in vector optimization: consequences to stability. Parametric optimization, Control and Cybernetics, 23 (1994) 107-122.

[21] R. E. Bellman, R. E. Kalaba and J. Lockett, Numerical inversion of the Laplace transform, Elsevier, New York, 1966.

[22] C. R. Bector, D. Bhatia and P. Jain, Generalized convexity and duality in multiobjective nonsmooth programming, Utilitas Mathematica, 43 (1993) 71-78.

[23] D. P. Bertsekas, Constrained optimization and Lagrange multiplier methods, Academic Press, New York, 1982.

[24] K. Bergstresser, A. Charnes and P. L. Yu, Generation of domination structures and nondominated solutions in multicriteria decision making, Journal of Optimization Theory and Applications, 18 (1976) 3-13.

[25] H. P. Benson, Existence of efficient solutions for vector maximization problems. Journal of Optimization Theory and Applications, 26 (1978) 569-580.

[26] H.P. Benson, An improved definition of proper efficiency for vector maximization with respect to cones, Journal of Mathematical Analysis and Applications, 71 (1979) 232-241.

[27] H. P. Benson, On a domination property for vector maximization with respect to cones, Journal of Optimization Theory and Applications, 39 (1983) 125-132.

[28] H. P. Benson and T. L. Morin, The vector maximization problem: Proper efficiency and stability, SIAM Journal on Applied Mathematics, 32 (1977) 64-72.

[29] H. Bernau, Use of exact penalty functions to determine efficient decisions. European Journal of Operational Research, 49 (1990) 348-355.

[30] D. Bhatia and S. Aggarwal, Optimality and duality for multiobjective nonsmooth programming, European Journal of Operational Research, 57 (1992) 360-367.

[31] M. Bianchi, N. Hadjisavvas and S. Schaible, Vector equilibrium problems with generalized monotone bifunctions, Journal of Optimization Theory and Applications, 92 (1997) 527-542.

[32] G. Bigi and M. Pappalardo, Regularity conditions in vector optimization, Journal of Optimization Theory and Applications, 102 (1999) 83-96.

[33] J. M. Borwein, On the existence of Pareto efficient points, Mathematics of Operations Research, 8 (1983) 64-73.

[34] J. M. Borwein, Proper efficient points for maximizations with respect to cones, SIAM Journal on Control and Optimization, 15 (1977) 57-63.

[35] J. M. Borwein, The geometry of Pareto efficiency over cones, Optimization, 11 (1980) 235-248.

[36] J. M. Borwein and D. Zhuang, Super efficiency in vector optimization, Transactions of the American Mathematical Society, 338 (1993) 105-122.

[37] S. Bolintineanu, Approximate efficiency and scalar stationarity in unbounded nonsmooth convex vector optimization problems, Journal of Optimization Theory and Applications, 106 (2000) 265-296.

[38] S. Bolintineanu and M. El Maghri, Second-order efficiency conditions and sensitivity of efficient points, Journal of Optimization Theory and Applications, 98 (1998) 569-592.

[39] E. Borel, The theory of play and integral equations with skew symmetric kernels, C. R. Acad. Sci. Paris, 173 (1921), 1304- 1308.

[40] H. Bonnel, A. N. Iusem and B. F. Svaiter, Proximal methods in vector optimization, SIAM J. Optimization, 15 (2005) 953-970.

[41] V. J. Bowman, On the relationship of the Tchebycheff norm and the efficient frontier of multiple-criteria objectives. In H. Thieriez, editor, Multiple Criteria Decision Making, volume 130 of Lecture Notes in Economics and Mathematical Systems. Springer Verlag, Berlin, 1976.

[42] L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, USSR Comput. Math. Math.Phys., 7 (1967) 165-186.

[43] F. E. Brower, Existence and perturbation theorems for nonlinear maximal monotone operators in Banach spaces, Bull. Amer. Math. Soc., 73 (1967) 322-327.

[44] R. S. Burachik, D. Butnariu and A. N. Iusem, Iterative methods for solving stochastic convex feasibility problems and applications, Computational Optimization and Applications, 15 (2000) 269-307.

[45] R. S. Burachik and A. N. Iusem, A generalized proximal point algorithm for the variational inequality problem in a Hilbert space, SIAM J. Optimization, 8 (1998) 197-216.

[46] R. S. Burachik and S. Scheimberg, A proximal point algorithm for the variational inequality problem in Banach space, SIAM J. Control and Optimization, 39 (2001) 1633-1649.

[47] D. Butnariu and A. N. Iusem, On a proximal point method for convex optimization in Banach space, Numerical Functional Analysis and Optimization, 18 (1997) 723-744.

[48] G. Cantor, Contributions to the fundation of transfinite set theory, Math. Annalen, 49 (1897) 207-246.

[49] E. Carrizosa and F. Plastria, Dominators for multiple-objective quasiconvex maximization problems, Journal of Global Optimization, 18 (2000) 35-58.

[50] L. C. Ceng and J. C. Yao, Approximate proximal methods in vector optimization, European Journal of Operational Research, 183 (2007) 1-19.

[51] Y. Censor. Pareto optimality in multiobjective problems, Applied Mathematics and Optimization, 4 (1977) 41-59.

[52] Y. Censor, A .N. Iusem and A. Zenios, An interior point method with Bregman functions for the variational inequality problem with paramonotone functions, Mathematical Programming, 81 (1998) 373-400.

[53] Y. Censor and A. Lent, An iterative row-action method for interval convex programming, Jounal of Optimization Theory and Applications, 34 (1981) 321-353.

[54] Y. Censor and S. A. Zenios, The proximal minimization algorithm with D-function, Jounal of Optimization Theory and Applications, 73 (1992) 451-464.

[55] Y. Censor and S.A. Zenios. Parallel Optimization: Theory, Algorithms and Applications, Oxford University Press (New York 1997).

[56] O. Chadli and H. Riahi. On generalized vector equilibrium problems, Journal of Global Optimization, 16 (2000) 33-41.

[57] V. Chankong and Y. Y. Haimes, On the characterization of noninferior solutions of the vector optimization problem, Automatica, 18 (1983) 697-707.

[58] G. Chen and M. Teboulle, Convergence analysis of a proximal- like minimization algorithm using Bregman function, SIAM Journal on Optimization, 3 (1993) 538-543.

[59] G.Y. Chen, Necessary conditions of nondominated solutions in multicriteria decision making, Journal of Mathematical Analysis and Applications, 104 (1984) 38-46.

[60] G. Y. Chen and X. Q. Yang, The vector complementary problem and its equivalences with vector mnimial element in ordered spaces. Journal of Mathematical Analysis and Applications, 153 (1990), 136-158.

[61] G. Y. Chen, Existence of solutions for a vector variational inequality: an extension of the Hartman- Stampacchia theorem, Journal of Optimization Theory and Applications, 74 (1992) 445-456.

[62] G.Y. Chen and B .D. Craven, A vector variational inequality and optimization over an efficient set, Zeitschrift Operations Research, 34 (1990) 1-12.

[63] G.Y. Chen and B. D. Craven, Existence and continuity of solutions for vector optimization, Journal of Optimization Theory and Applications, 81 (1994) 459-468.

[64] G. Y. Chen and W. D. Rong, Characterizations of the Benson proper efficiency for nonconvex vector optimization, Journal of Optimization Theory and Applications, 98 (1998) 365-384.

[65] G.Y. Chen, X. X. Huang and X. Q. Yang, Vector Optimization: Set-Valued and Variational Analysis, Lecture Notes in Economics and Mathematical Systems, 541, Springer-Verlag, Berlin, 2005.

[66] K. L. Chew and E. U. Choo, Pseudolinearity and efficiency, Mathematical Programming, 28 (1984) 226-239.

[67] F. H. Clarke, Optimization and Nonsmooth Analysis, John Wiley X, New York, 1983.

[68] B. D. Craven, On sensitivity analysis for multicriteria optimization, Optimization, 19 (1988) 513-523.

[69] Crouzeix, J. P., Pseudomonotone variational inequality problems: existence of solutions, Mathmatical Programming, 78 (1997) 305-314.

[70] I. Das and J. E. Dennis, Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. SIAM Journal of Optimization, 8 (1998) 631-657.

[71] J. P. Dauer and R. J. Gallagher, Positive proper efficient points and related cone results in vector optimization theory, SIAM Journal on Control and Optimization, 28 (1990) 158-172.

[72] A. De Pierro and A. N. Iusem, A relaxed version of Bregman's method for convex programming, Jounal of Optimization Theory and Applications, 51 (1986) 421-440.

[73] G. Debreu, Valuation equilibrium and Pareto optimum, Proc.Nat.Acad. Sci. USA, 40 (1954) 588-592.

[74] S. Deng, Characterzations of nonemptiness and compactness of solutions sets in convex optimzation, Jounal of Optimization Theory and Applications, 96 (1998) 123-131.

[75] S. Deng, On approximate solutions in convex vector optimization, SIAM Journal on Control and Optimization, 35 (1997) 2128-2136.

[76] S. Deng, Characterizations of the nonemptiness and boundedness of weakly efficient solution sets of convex vector optimization problems in real reflexive Banach spaces, Jounal of Optimization Theory and Applications, 140 (2009) 1-7.

[77] J. Eckstein, Nonlinear proximal point algorithms using Bregman function, with applications to convex programming, Mathematics of Operations Research, 18 (1993) 202-226.

[78] J. Eckstein, Approximate iterations in Bregman-function-based proximal algorithms, Mathmatical Programming, 83 (1998) 113-123.

[79] F. Y. Edgeworth, Mathematical Psychis, C.Kegan Paul Co. London, England, 1881.

[80] M. Ehrgott, Multicriteria Optimization, Springer-Verlag, Berlin, 2000.

[81] M. Ehrgott, H. W. Hamacher, K. Klamroth and S. Nickel, A. Schöel and M.M. Wiecek. A note on the equivalence of balance points and Pareto solutions in multiple-objective programming, Journal of Optimization Theory and Applications, 92 (1997) 209-212.

[82] M. Ehrgott and D.M. Ryan, Constructing robust crew schedules with bicriteria optimization. Journal of Multi-Criteria Decision Analysis, 11 (2002) 139-150.

[83] J. Ester and F. Trötzsch, On generalized notions of efficiency in multicriteria decision making. System Analysis Modelling Simulation, 3 (1986) 153-61.

[84] F. Facchinei and J. S. Pang, Finite-dimensional variational inequalities and complementarity problems, Volume I, Springer Series in Operations Research, Springer, Berlin, 2004.

[85] F. Facchinei and J. S. Pang, Finite-dimensional variational inequalities and complementarity problems, Volume II, Springer Series in Operations Research, Springer, Berlin, 2004.

[86] M.C. Ferris, Finite termination of the proximal-point algorithm, Mathematical Programming, 50 (1991) 359-366.

[87] F. Ferro, An optimization result for set-valued mappings and a stability property in vector problems with constraints, Journal of Optimization Theory and Applications, 90 (1996) 63-77.

[88] J. Fliege, An efficient interior-point method for convex multicriteria optimization problems, Mathematics of Operations Researh, 31 (2006) 825-845.

[89] J. Fliege L.M. Graña Drummond and B.F. Svaiter, Newton's Method for unconstrained multiobjective optimization, SIAM J. optimization, 20 (2009) 602-626.

[90] J. Fliege and B. F. Svaiter, Steepest descent methods for multicriteria optimization, Mathematical Mehtod of Operations Researh, 51 (2000) 479-494.

[91] F. Flores-Bazan, Ideal, weakly efficient solutions for vector problems, Mathematical Programming, 93 (2002) 453-475.

[92] F. Flores-Bazanand and C. Vera, Characterization of the nonemptiness and compactness of solution sets in convex and nonconvex vector optimization, Journal of Optimization Theory and Applications, 130 (2006) 185-208.

[93] W. B. Gearhart, Compromise solutions and estimation of the noninferior set. Journal of Optimization Theory and Applications, 28 (1979) 29-47.

[94] F. W. Gembicki and Y. Y. Haimes, Approach to performance and multiobjective sensitivity optimization: The goal attainment method. IEEE Transactions on Automatic Control, 20 (1975) 769-771.

[95] A. M. Geoffrion, Proper efficiency and the theory of vector maximization, Journal of Mathematical Analysis and Applications, 22 (1968) 618-630.

[96] C. Gert and P. Weidner, Nonconvex separation theorems and some applications in vector optimization. Journal of Optimization Theory and Applications, 67 (1990) 297-320.

[97] F. Giannessi, Theorems of alternative, quadratic programs and complementarity problems. In Cottle, R. W. and Giannessi, F. and Lions, J. L. (eds.), Variational Inequality and Complementarity Problems, Wiley, New York, 1980.

[98] F. Giannessi. On Minty variational principle, in: F. Giannessi, S. Komlosi and T. Rapcsak (eds.), New Trends in Mathematical Programming, 93-99, Kluwer Academic Publishers, Dordrecht, 1998.

[99] C. J. Goh and X. Q. Yang, Scalarization methods for vector variational inequality, Nooconvex Optimization and Its Application, Vol.38, Vector Variational Ineqalities and Vector Equilibria, Franco Giannessi (Ed.), Kluwer Academic Publisher, 2000 217-231.

[100] E. G. Golshtein and N. V. Tretyakov, Modified Lagrangians in convex programming and their generalizations. Mathematical Programming Study, 10 (1979) 86-97.

[101] O. Güler, On the convergence of the proximal point algorithm for convex minimization, SIAM J. Control and Optimization, 29 (1991) 403-419.

[102] O. Güler, Ergodic convergence in proximal point algorithms with Bregman functions, in Du, D.-Z. and Sun, J., Advances in Optimization and Approximation, Kluwer Academic Publishers, (1994) 155-165.

[103] L. M. Graña Drummond and A. N. Iusem, A projected gradient method for vector optimization problems, Computational Optimization and Applications, 28 (2004) 5-30.

[104] L. M. Graña Drummond and B. F. Svaiter, A steepest descent method for vector optimization, Journal of Computational and Applied Mathematics, 175 (2005) 395-414

[105] J. Guddat, F. Guerra Vasquez, K. Tammer and K. Wendler, Multiobjective and Stochastic Optimization Based on Parametric Optimization. Akademie-Verlag, Berlin, 1985.

[106] T. R. Gulati and M. A. Islam, Efficiency and proper efficiency in nonlinear vector maximum problems, European Journal of Operational Research, 44 (1990) 373-382.

[107] N. Hadjisavvas and S. Schaible, From scalar to vector equilibrium problems in the quasimonotone case, Journal of Optimization Theory and Applications, 96 (1998) 297-309.

[108] M. A. Hanson, On sufficiency of the Kuhn-Tucker conditions, Journal of Mathematical Analysis and Applications, 80 (1981) 544-550.

[109] R. Hartley, On cone-efficiency, cone-convexity and cone-compactness, SIAM Journal on Applied Mathematics, 34 (1978) 211-222.

[110] F. Hausdorff, In vestigation concerning order types, Math.- Physische Klasse, 58 (1906) 106-169.

[111] M. Heinkenschloss, Mesh independence for nonlinear least squares problems with norm constraints, SIAM Journal on Optimization, 3 (1993) 81-117.

[112] M. I. Henig, Proper efficiency with respect to cones, Journal of Optimization Theory and Applications, 36 (1982) 387-407.

[113] M. I. Henig, Existence and characterization of efficient decisions with respect to cones, Mathematical Programming, 23 (1982) 111-116.

[114] X. X. Huang, Extended well-posedness properties of vector optimization problems, Journal of Optimization Theory and Applications, 106 (2000) 165-182.

[115] X. X. Huang and X. Q. Yang, Duality and Exact penalization for Vector Optimization via Augmented Lagrangian, Journal of Optimization Theory and Applications, 111 (2001) 615-640.

[116] X. X. Huang and X. Q. Yang, Characterizations of Nonemptiness and Compactness of the Set of Weakly Efficient Solutions for Convex Vector Optimization and Applications, Journal of Mathematical Analysis and Applications, 264 (2001) 270-287

[117] X. X. Huang and X. Q. Yang , Levitin-Polyak well-posedness of constrained vector optimization problems, Journal of Global Optimization, 37 (2007) 287-304.

[118] X. X. Huang and X. Q. Yang, Scalarization of nonemptiness and compactness of solution set of a vector variational vnequality with applications, in preparation.

[119] X. X. Huang, K. L. Teo and X. Q. Yang, Calmness and exact penalization in vector optimization with cone constraints, Computational Optimization and Applications, 35 (2006) 47-67.

[120] Y. Ishizuka, Optimality conditions for directionally differentiable multi-objective programming problems, Journal of Optimization Theory and Applications, 72 (1992) 91-111.

[121] Y. Ishizuka and H. D. Tuan, Directionally differentiable multiobjective optimization involving discrete inclusions, Journal of Optimization Theory and Applications, 88 (1996) 585-616.

[122] A. N. Iusem, B. F. Svaiter and M. Teboulle, Entropy-like proximal methods in convex programming. Mathematics of Operations Research, 19 (1994) 790-814.

[123] J. Jahn, Existence theorems in vector optimization, Journal of Optimization Theory and Applications, 50 (1986) 397-406.

[124] J. Jahn, Vector Optimization: Theory, Applications and Extensions, Springer-Verlag, Berlin, 2004;

[125] V. Jeyakumar, G. M. Lee and N. Dinh, Characterizations of solution sets of convex vector minimization problems, European Journal of Operation Research, 174 (2006) 1380-1395.

[126] V. Jeyakumar and X. Q. Yang, Convex composite multi-objective nonsmooth programming. Mathematical Programming, 59 (1993) 325-343.

[127] E. Jurkiewicz, Stability of compromise solution in multicriteria decision making problems, Journal of Optimization Theory and Applications, 40 (1983) 77-83.

[128] I. Kaliszewski, A modified weighted Tchebycheff metric for multiple objective programming. Computers and Operations Research, 14 (1987) 315-323.

[129] S. Kamimura and W. Takahashi, Strong convergence of a proximal-type algorithm in a Banach space, SIAM J. Optimization, 13 (2003) 938-945.

[130] P. Kanniappan, Necessary conditions for optimality of nondifferentiable convex multiobjective programming, Journal of Optimization Theory and Applications, 40 (1983) 167-174.

[131] H. Kawasaki, Second-order necessary conditions of the Kuhn- Tucker type under new constraint qualification, Journal of Optimization Theory and Applications, 57 (1994) 253-264.

[132] K. R. Kazmi, Existence of solutions for vector optimization, Applied Mathematics Letters, 9 (1996) 19-22.

[133] P. Q. Khanh, Proper solutions of vector optimization problems, Journal of Optimization Theory and Applications, 74 (1992) 105-130.

[134] K. C. Kiwiel, Proximal minimization methods with generalized Bregman functions, SIAM J. Control and Optimization, 35 (1997) 326-349.

[135] K. Klamroth, J. Tind and M. M. Wiecek, Unbiased approximation in multicriteria optimization. Mathematical Methods of Operations Research, 36 (2002) 413-437.

[136] I. V. Konnov and J. C. Yao, Existence of solutions for generalized vector equilibrium problems, Journal of Mathematical Analysis and Applications, 233 (1999) 328-335.

[137] T. C. Koopmans, Analysis of produnction as an efficient combination of activities, in "Activity analysis of production and allocation", Cowles Commission Monograph n.13, T. C., Koopmans ed., John Wiley and Sons, New York, 1951, 33-97.

[138] P. Korhonen and G. Y. Yu, A reference direction approach to multiple objective quadraticlinear programming, European Journal of Operational Research, 102 (1997) 601-610.

[139] H. W. Kuhn and A. W. Tucker, Nonlinear programming, in "Proceeding of the Second Berkeley Symposium on Mathematical Statistics and Probability ", Berkeley, California, 1951 481-492.

[140] H. Kuk, T. Tanino and M. Tanaka, Sensitivity analysis in vector optimization, Journal of Optimization Theory and Applications, 89 (1996) 713-730.

[141] M. Laumen, Newton's mesh independence principle for a class of optimal shape design problems, SIAM J. Control Optim. 37 (1999) 1142-1168.

[142] G. M. Lee, On relations between vector variational inequality and vector optimization problem, in: X.Q. Yang, A.I. Mees, M. Fisher and L. Jennings (eds.), Progress in Optimization, 167-179, Kluwer Academic Publishers, Dordrecht, 2000.

[143] G. M. Lee, D. S. Kim, B. S. Lee and N. D. Yen, Vector variational inequality as a tool for studying vector optimization problems, Nonlinear Analysis: Theory, Methods and Applications, 34 (1998) 745-765.

[144] G. M. Lee, D. S. Kim and H. Kuk, Existence of solutions for vector optimization problems, Journal of Mathematical Analysis and Applications, 220 (1998) 90-98.

[145] G. M. Lee, B. S. Lee and S. S. Chang, On vector quasivariational inequalities, Journal of Mathematical Analysis and Applications, 203 (1996) 626-638.

[146] N. Lehdili and A. Moudafi, Combining the proximal algorithm and Tikhonov regularization, Optimization, 37 (1996) 239-252.

[147] Z. F. Li and S.Y. Wang, ε- efficient solutions in multiobjective optimization, Optimization, 44 (1998) 161-174.

[148] J. G. Lin, Maximal vectors and multiobjective optimization, Journal of Optimization Theory and Applications, 18 (1976) 41-64, .

[149] J. C. Liu, ε-Pareto optimality for nondifferentiable multiobjective programming via penalty function, Journal of Mathematical Analysis and Applications, 198 (1996) 248-261.

[150] R. Lopez and C. Vera, On the set of weakly efficient minimizers for convex multiobjective programming, Operations Research Letters, 36 (2008) 651-655.

[151] P. Loridan, $\varepsilon-solution$ in vector minimization problems, Journal of Optimization Theory and Applications, 43 (1984) 265-276.

[152] D. T. Luc, On the domination property in vector optimization, Journal of Optimization Theory and Applications, 43 (1984) 327-330.

[153] D. T. Luc, Theory of Vector Optimization, Lecture Notes in Economics and Mathematical Systems, 319, Springer-Verlag, Berlin, 1989.

[154] D. T. Luc and S. Schaible, Efficiency and generalized concavity, Journal of Optimization Theory and Applications, 94 (1997) 147-153.

[155] F. J. Luque, Asymptotic convergence analysis of the proximal point algorithm. SIAM Journal on Control and Optimization, 22 (1984) 277-293.

[156] H. L$\ddot{u}$tkepohl, Handbook of Matrices. John Wiley , Chichester, 1996.

[157] P. E. Mainge, A hybrid extragradient-viscosity method for montone opertors and fixed point problems, SIAM J. Control and Optimization, 47 (2008) 1499-1515.

[158] B. Matinet, Regularisation d'in$\acute{e}$quations variationelles par approximations successives, Rev. Francaise Inf. Rech, Oper., 2 (1970) 154-159.

[159] R. E. Megginson, An Introdunction to Banach Space Theory, Graduate Texts in Mathematics 183, Springer-Verlag, Berlin, 1998.

[160] K. M. Miettinen, Nonlinear Multiobjective Optimization, Kluwer Academic Publishers, Dordrecht, 1999.

[161] E. Miglierina, E. Molho and M. C. Recchioni, Box-constrained multi-objective optimization: A gradient-like method without a priori scalarization, European Journal of Operational Research, 188 (2008) 662-682.

[162] P. H. Naccache, Stability in multicriteria optimization, Journal of Mathematical Analysis and Applications, 68 (1979) 441-453.

[163] J. von Neumann, Theory of games and economic behavior, Princeton Univ. Press, Princeton, New Jersey, 1947.

[164] V. Pareto, Cours d'economie politique, F. Rouge, Lausanne, 1896.

[165] A. Pascoletti and P. Serafini, Scalarizing vector optimization problems. Journal of Optimization Theory and Applications, 42 (1984) 499-524.

[166] M. Patriksson, Nonlinear Programming and Variational Inequality Problems. A Unified Approach. Kluwer Academic Publishers (Dordrecht 1999).

[167] T. Pennanen, Local convergence of the proximal point algorithm and multiplier methods without monotonicity, Mathematics of Operations Research, 27 (2002) 170-191.

[168] J. P. Penot and A. Sterna-Karwat, Parametrized multicriteria optimization - continuity and closedness of optimal multifunctions, Journal of Mathematical Analysis and Applications, 120 (1986) 150-168.

[169] F. Plastria and E. Carrizosa, Geometrical characterization of weakly efficient points, Journal of Optimization Theory and Applications, 90 (1996) 217-223.

[170] R. R. Phelps, Convex functions, Monotone Operators and Differentiability, Lecture Notes in Math. 1364 Springer, Berlin, 1988.

[171] R. T. Rockafellar, On the Maximality of Sums of Nonlinear Monotone Operators, Transactions of the American Mathematical Society, 149 (1970) 75-88.

[172] R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1970.

[173] R. T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control and Optimization, 14 (1976) 877-898.

[174] R. T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Mathematics of Operations Research, 1 (1976) 97-116.

[175] R. T. Rockafellar and J. B. Wets, Variational analysis, Springer, 1998.

[176] B. Roy, Problems and methods with multiple objective functions. Mathematical Programming, 1 (1971) 239-266.

[177] Y. Sawaragi, H. Nakayama and T. Tanino, The theory of multiobjective optimization, Academic Press, New York, 1985.

[178] B. Schandl, K. Klamroth, and M.M. Wiecek, Norm-based approximation in multicriteria programming. Computers and Mathematics with Applications, 44 (2002) 925-942.

[179] B. Schandl, K. Klamroth, and M.M. Wiecek, Introducing oblique norms into multiple criteria programming. Journal of Global Optimization, 23 (2002) 81-97.

[180] C. Singh, Optimality conditions in multiobjective differentiable programming, Journal of Optimization Theory and Applications, 53 (1987) 115-123.

[181] R. M. Soland, Multicriteria optimization: A general characterization of efficient solutions. Decision Sciences, 10 (1979) 26-38.

[182] M. V. Solodov and B. F. Svaiter, A hybrid projection-proximal point algorithm, Journal of Convex Analysis, 6 (1999) 59-70.

[183] M. V. Solodov and B.F. Svaiter, Error bounds for proximal point subproblems and associated inexact proximal point algorithms, Mathematical Programming, 88 (2000) 371-389.

[184] M. V. Solodov and B. F. Svaiter, Forcing strong convergence of proximal point iterations in a Hilbert space, Mathematical Programming, 87 (2000) 189-202.

[185] M. V. Solodov and B.F. Svaiter, A unified framework for some inexact proximal point algorithms, Numerical Functional Analysis and Optimization, 22 (2001) 1013-1035.

[186] W. Song and G. M. Yao, Homotopy Method for a General Multiobjective Programming Problem, Journal of Optimization Theory and Applications, 138 (2008) 139-153.

[187] Y. Sonntag and C. Zalinescu, Comparison of existence results for efficient points, Journal of Optimization Theory and Applications, 105 (2000) 161-188.

[188] T. Staib, On necessary and sufficient optimality conditions for multicriterial optimization problems, ZOR - Methods and Models of Operations Research, 35 (1991) 231-248.

[189] A. Sterna-Karwat, Continuous dependence of solutions on a parameter in a scalarization method, Journal of Optimization Theory and Applications, 55 (1987) 417-434.

[190] R. Steuer, Multiple Criteria Optimization: Theory, Computation and Application, John Wiley, New York, 1986.

[191] T. Tanino, Sensitivity analysis in multiobjective optimization, Journal of Optimization Theory and Applications, 56 (1988) 479-499.

[192] T. Tanino and Y. Sawaragi, Stability of nondominated solutions in multicriteria decision making, Journal of Optimization Theory and Applications, 30 (1980) 229-253.

[193] M, Teboulle, Convergence of proximal-like algorithms, SIAM J. Optimization, 7 (1997) 1069-1083.

[194] M. L. TenHuisen and M. M. Wiecek. Efficiency and solution approaches to bicriteria nonconvex programs. Journal of Global Optimization, 11 (1997) 225-251.

[195] A. N. Tikhonov, Solution of in correctly formulated problems and regularization method, Soviet Mathematics Doklady, 4 (1963) 1035-1038.

[196] J. Tind and M. M. Wiecek, Augmented Lagrangian and Tchebycheff approaches in multiple objective programming. Journal of Global Optimization, 14 (1999) 251-266.

[197] S. Y. Wang, Second-order necessary and sufficient conditions in multiobjective programming, Numerical Functional Analysis and Applications, 12 (1991) 237-252.

[198] R. E. Wendell and D. N. Lee, Efficiency in multiple objective optimization problems, Mathematical Programming, 12 (1977) 406-414.

[199] T. Weir and M. Mond, Generalized convexity and duality in multiple objective programming, Bulletin of the Australian Mathematical Society, 39 (1989) 287-299.

[200] D. J. White, Optimality and Efficiency, Wiley, New York, 1982.

[201] D.J. White, Epsilon efficiency, Journal of Optimization Theory and Applications, 49 (1986) 319-337.

[202] D.J. White, Weighting factor extensions for finite multiple objective vector minimization problems. European Journal of Operational Research, 36 (1988) 256-265.

[203] H. K. Xu, Viscosity approximation methods for nonexpansive mappings, Journal of mathematical Analysis and Applications, 298 (2004) 279-291.

[204] N. Yamashita and M. Fukushima. The proximal point algorithm with genuine superlinear convergence for the monotone complementarity problem. SIAM Journal on Optimization 11 (2001) 364-379.

[205] J. C. Yao, Variational inequalities with generalized monotone operators, Mathematics of Operations Reaserch, 19 (1994) 691-705.

[206] X. M. Yang and D. Li, On properties of preinvex function, Journal of mathematical Analysis and Applications, 256 (2001) 229-241.

[207] X. M. Yang, X. Q. Yang and K. L. Teo, Characterizations and Applications of Prequasi-invex Functions, Journal of Optimization Theory and Applications, 110 (2001) 645-668.

[208] X. M. Yang, X. Q. Yang and K. L. Teo, Generalized invexity and generalized invariant monotonicity. Journal of Optimization Theory and Applications, 117 (2003) 607-625.

[209] X. M. Yang, X. Q. Yang and K. L. Teo, Criteria for generalized invex monotonicities. European Journal of Operational Research, 164 (2005) 115-119.

[210] X. Q. Yang, A Hahn-Banach theorem in ordered linear spaces and its applications, Optimization, 2 (1992) 1-9.

[211] X. Q. Yang, Vector variational inequalities and its duality, Nonlinear Analysis, TMA, 21 (1993) 867-877.

[212] X.Q. Yang, Generalized convex functions and vector variational inequalities, Journal of Optimization Theory and Applications, 79 (1993) 563-580.

[213] X.Q. Yang, Vector variational inequality and vector pseudolinear optimization, Journal of Optimization Theory and Applications, 95 (1997) 729-734.

[214] P. L. Yu, A class of solutions for group decision making, Management Science, 19 (1973) 936-946.

[215] P. L. Yu, Cone convexity, cone extreme points, and nondominated solutions in decision problems with multiple objectives, Journal of Optimization Theory and Applications, 14 (1974) 319-377.

[216] P. L. Yu, Multiple-Criteria Decision Making: Concepts, Techniques and Extensions, Plenum Press, New York, 1985

[217] P. L. Yu and G. Leitmann, Compromise solutions, domination structures, and Salukvadzes solution, Journal of Optimization Theory and Applications, 13 (1974) 362-378.

[218] P. L. Yu and G. Leitmann, Nondominated decisions and cone convexity in dynamic multicriteria decision problems, Journal of Optimization Theory and Applications, 14 (1974) 573-584.

[219] D. M. Zhuang, Bases of convex cones and Borweins proper efficiency, Journal of Optimization Theory and Applications, 71 (1991) 613-620.

[220] G. Eichfelder, Adaptive scalarization methods in in multiobjective optimization, Springer, 2008.

[221] J. P. Penot and M. Thera, Semi-continuous in general topology, Arch-Math, 38 (1982), 158-166.