Filomat 34:7 (2020), 2367–2376
https://doi.org/10.2298/FIL2007367A

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Proximal Point Algorithm for Differentiable Quasi-Convex Multiobjective Optimization

Fouzia Amir[a], Ali Farajzadeh[b], Narin Petrot[a,c]

[a] *Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok, 65000, Thailand*
[b] *Department of Mathematics, Faculty of Science, Razi University, BagheAbrisham St. Kermanshah, Iran*
[c] *Centre of Excellence in Nonlinear Analysis and Optimization, Naresuan University, Phitsanulok, 65000, Thailand*

Abstract. The main aim of this paper is to consider the proximal point method for solving multiobjective optimization problem under the differentiability, locally Lipschitz and quasi-convex conditions of the objective function. The control conditions to guarantee that the accumulation points of any generated sequence, are Pareto critical points are provided.

1. Introduction

Let $I := \{1, ..., m\}$, $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m : x_j \geq 0, j \in I\}$, and $\mathbb{R}^m_{++} = \{x \in \mathbb{R}^m : x_j > 0, j \in I\}$. For $y, z \in \mathbb{R}^m$, ($z \geq y$ or $y \leq z$) means that $z - y \in \mathbb{R}^m_+$, and ($z \succ y$ or $y \prec z$) means that $z - y \in \mathbb{R}^m_{++}$. By using these relations, we consider the multiobjective minimization problem as

$$\min_{x \in \mathbb{R}^n} F(x), \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^m$ with,

$$F(x) := (f_1(x), ..., f_m(x)), \quad \textit{for each} \quad x \in \mathbb{R}^n.$$

Multiobjective optimization is the process of simultaneously optimizing two or more real-valued objective functions. It is usually hard to find an optimal solution that satisfies all objectives from the mathematical point of view (i.e., there is no ideal minimizer), and so we consider the following concepts of the solution: a point $x^* \in \mathbb{R}^n$ is called

(i) Pareto optimal point of F, if there exists no $x \in \mathbb{R}^n$ such that $F(x) \leq F(x^*)$ with $F(x) \neq F(x^*)$.

(ii) weak Pareto optimal point of F, if there exists no $x \in \mathbb{R}^n$ such that $F(x) \prec F(x^*)$.

2010 *Mathematics Subject Classification*. 80M50; 58E17; 52A41
Keywords. Multiobjective optimization, Pareto critical point, quasiconvex, pseudoconvex
Received: 08 July 2019; Revised: 04 March 2020; Accepted: 10 March 2020
Communicated by Vasile Berinde
Research supported by Faculty of Science, Naresuan University
Email addresses: fouziaa59@nu.ac.th (Fouzia Amir), faraj1348@yahoo.com (Ali Farajzadeh), narinp@nu.ac.th (Narin Petrot)

It is clear that a Pareto optimal point is also a weak Pareto optimal point but the converse is not true. These types of solution concepts have applications in the economy, industry, agriculture, and other fields, see [15].

Many authors introduced various algorithms for solving the problem (1). Bonnel et al. [7] proved the convergence of the proximal point method of problem (1) for a weak Pareto solution of a multiobjective optimization problem. Bello Cruz et al. [3], considered the projected gradient method for solving the problem of finding a Pareto optimal point of a quasiconvex multiobjective function. Da Cruz Neto et al. [11], extended the classical subgradient method from real-valued minimization to multiobjective optimization for solving quasiconvex nondifferentiable unconstrained multiobjective optimization problems. Chen et al. [10] proposed a new proximal point algorithm by using auxiliary principle technique, based on decomposition method for computing a weakly efficient solution of the constrained multiobjective optimization problem without assuming the nonemptiness of its solution set.

It is well known that multiobjective optimization problems are also solved by using scalarization technique [8, 12–14]. Scalarization means that the problem is converted into a single(scalar) or a family of single objective optimization problems. In this direction, the adaptive problem becomes a real-valued objective function, possibly depending on the chosen parameters. In 2010, Gregorio and Oliveira [16] proved the convergence of the proximal point method by using a logarithmic quadratic proximal scalarization method. Later, Apolinario et al. [1] developed an exact linear scalarization proximal point algorithm to solve the problem (1). In 2014, Bento et al. [4] introduced the following nonlinear scalarized proximal iteration **(BCS)** for solving the multiobjective optimization problem by

$$x_{k+1} \in \arg\min_{x\in\mathbb{R}^n} g\Big(F(x) + I_{\Omega_k}(x)e + \frac{\lambda_k}{2}\|x - x_k\|^2 e\Big), \tag{2}$$

where $\Omega_k := \{x \in \mathbb{R}^n : F(x) \preceq F(x_k)\}$, $\{\lambda_k\}$ is a bounded sequence, I_{Ω_k} denotes the indicator function of Ω_k, $e := (1, ..., 1) \in \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}$ is defined as

$$g(y) := \max_{i\in I}\langle y, e_i\rangle, \quad \textit{for each} \quad y \in \mathbb{R}^m, \tag{3}$$

where $e_i = (0, ..., i, ..., 0)$ is the canonical base of the space $\mathbb{R}^m$. It is well known that the scalar function g equals to the following nonlinear scalarization function

$$g(y) = \inf\{t \in \mathbb{R} : te \in y + \mathbb{R}^m_+\}, \quad \textit{for each} \quad y \in \mathbb{R}^m,$$

(see [19]). It is also worth to point out that the function g fulfills the following properties:

$$\begin{aligned} g(x + \alpha e) =& g(x) + \alpha, \;\; g(tx) = tg(x), \;\; x \in \mathbb{R}^m, \; \alpha \in \mathbb{R}, \; t \geq 0. \\ & x \preceq y \implies g(x) \leq g(y), \;\; x, y \in \mathbb{R}^m. \end{aligned} \tag{4}$$

By using these concepts, Bento et al. [4] proved that under the continuously differentiability of objective functions, the sequence generated by the algorithm **(BCS)** converges to a Pareto critical point. Note that, in the **(BCS)** algorithm, if $x_k = x_{k+1}$, then the algorithm stops to a Pareto critical point.

In the present paper, we will continue the study the **(BCS)** algorithm but under the weaker assumptions of differentiable and locally Lipschitz properties of the considered objective function instead of continuously differentiable assumptions. We show that under these assumptions, the method is still well defined and that the accumulation points of any generated sequence, if any, are Pareto critical point for the multiobjective function. Full convergence of the sequence generated by the algorithm **(BCS)** is also considered.

2. Preliminaries

In this section, we present some basic results and definitions. We say that a real valued function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous function at a point $\hat{x} \in \mathbb{R}^n$ if for all sequence $\{x_k\} \subset \mathbb{R}^n$ such that $\lim_{k\to+\infty} x_k = \hat{x}$, we obtain that

$$f(\hat{x}) \leq \liminf_{k\to+\infty} f(x_k).$$

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

The domain of f, denoted by $dom f$, is the subset of $\mathbb{R}^n$ on which f has a finite values. A function f is said to be proper when $\mathrm{dom} f \neq \emptyset$. For a proper function f, we say that f is locally Lipschitz at $x \in \mathrm{dom}\, f$ if there exist positive real numbers ϵ_x and L_x such that

$$\|f(z) - f(y)\| \leq L_x \|z - y\|, \;\; \forall z, y \in \mathrm{B}(x, \epsilon_x) \cap \mathrm{dom}\, f,$$

where $B(x, \epsilon_x) := \{y \in \mathbb{R}^n : \|x - y\| < \epsilon_x\}$. We say that f is locally Lipschitz if f is locally Lipschitz for all $x \in dom f$.

Next, we remind Fréchet and Mordukovich subdifferential concepts.

Definition 2.1. *Let f be a lower semicontinuous function. The Fréchet subdifferential of f at $x \in \mathbb{R}^n$ is defined by*

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

It was pointed out that the Fréchet subdifferential is not completely satisfactory in optimization, since $\hat{\partial} f(x)$ might be empty-valued at points of particular interest (see [6] for more information), and this justifies the choice of the following subdifferential:

Definition 2.2. *Let f be a lower semicontinuous function. The Mordukovich-subdifferential of f at $x \in \mathbb{R}^n$ is defined by*

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

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

We can see that $\hat{\partial} f(x) \subset \partial f(x)$. In the particular case, when f is differentiable at $x \in \mathbb{R}^n$, we have $\hat{\partial} f(x) = \partial f(x) = \{\nabla f(x)\}$. If f is convex, then both subdifferentials $\hat{\partial} f(x)$ and $\partial f(x)$ coincide with the usual subdifferential for each $x \in \mathrm{dom}\, f$. Note also that the necessary (but not sufficient) condition for $x \in \mathrm{intdom}\, f$ to be a minimizer of f is

$$0 \in \partial f(x). \tag{5}$$

A point $x \in \mathbb{R}^n$ satisfying the above inclusion (5) is called limiting-critical or simply critical point.

Next, we recall the concept of normal cone on a convex set for a function f.

Definition 2.3. *Let $C \subset \mathbb{R}^n$ be a nonempty convex set. Then for each $x \in C$, the normal cone is defined by*

$$N_C(x) := \{v \in \mathbb{R}^n : \langle v, y - x\rangle \leq 0, \;\; y \in C\}. \tag{6}$$

Remark 2.4. *(i) For nonempty closed and convex set $C \subset \mathbb{R}^n$, we have $\partial I_C(x) = N_C(x)$.*

(ii) For a given lower semicontinuous function $h : \mathbb{R}^n \to \mathbb{R}$ and a nonempty, closed and convex subset $C \subset \mathbb{R}^n$, if $f = h + I_C$, it follows that f is a proper lower semicontinuous function with $dom f = C$. Then the first order optimality condition takes the following form:

$$0 \in \partial h(x) + N_C(x), \tag{7}$$

see ([22], Theorem 8.5).

Now, we consider a multiobjective mapping $F : \mathbb{R}^n \to \mathbb{R}^m$. We say that F has a directional derivative at $x \in \mathbb{R}^n$ in the direction of $v \in \mathbb{R}^n$ if

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

For a differentiable function F, we denote the jacobian of F at $x \in \mathbb{R}^n$ by

$$JF(x) := (\nabla f_1(x), ..., \nabla f_m(x)),$$

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

$$Im(JF(x)) := \{JF(x)v = (\langle \nabla f_1(x), v\rangle, ..., \langle \nabla f_m(x), v\rangle), \quad v \in \mathbb{R}^n\}. \tag{8}$$

If F is differentiable then $D_v F(x) = JF(x)v$.

The first order optimality condition for problem (1) is given by

$$x \in \mathbb{R}^n, \;\; Im(JF(x)) \cap (-\mathbb{R}^m_{++}) = \emptyset. \tag{9}$$

In general, (9) is necessary, but not sufficient condition, for optimality. A point $x \in \mathbb{R}^n$ satisfying (9) is called a Pareto critical point (see, for instance, [19]).

The following propositions are important in this paper.

Proposition 2.5. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous at $x \in \mathbb{R}^n$ for all $i \in \{1, ..., m\}$. If $h : \mathbb{R}^n \to \mathbb{R}$ is defined by:*

$$h(x) = \max_{1 \leq i \leq m} f_i(x), \quad \textit{for each} \quad x \in \mathbb{R}^n,$$

then, h is locally Lipschitz continuous function and

$$\partial h(x) = conv\{\partial f_i(x) : i \in I(x)\}, \quad \textit{for each} \quad x \in \mathbb{R}^n \tag{10}$$

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

$$I(x) := \{i \in I : f_i(x) = h(x)\}.$$

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

Proposition 2.6. *Let $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ be functions such that f_1 is locally Lipschitz continuous at $\bar{x} \in \mathbb{R}^n$ while f_2 is proper lower semicontinuous with $f_2(\bar{x})$ finite. Then,*

$$\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x}).$$

Proof. See([22], page 431). □

Proposition 2.7. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper locally Lipschitz function and $\{y^k\} \subset dom f$ a bounded sequence. If $\{z^k\}$ is a sequence such that $z^k \in \partial f(y^k)$, then $\{z^k\}$ is bounded.*

Proof. The proof follows by combining ([22], Theorem 9.13 and Proposition 5.15) for $S = \partial f$ and $B = \{y^k\}$. □

Remark 2.8. *In Proposition 2.7, if we take $\{y^k =: \hat{x} \in \mathbb{R}^n\}$ and $\{z^k \in \partial f(y^k)\}$, then $\{z^k\} \subset \mathbb{R}^n$ is bounded. So it has a convergent subsequence and consequently $\partial f(\hat{x})$ is relatively compact.*

We need the following in the sequel:

Definition 2.9. *[19] Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a multiobjective mapping.*

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

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

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

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

where the supremum is considered coordinatewise.

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

$$F(y) \prec F(x) \implies JF(x)(y-x) \prec 0.$$

Remark 2.10. *Note that F is convex (resp. quasi-convex) iff F is componentwise convex (resp. quasi-convex), see Definition 6.2 and Corollary 6.6 of ([19], pages 29 and 31), respectively. It is easy to see that if F is convex, it is also quasi-convex, and the reciprocal is clearly false. On the other hand, if F is componentwise pseudo-convex, then F is pseudo convex, although the reciprocal is false.*

We need the following propositions in our proofs.

Proposition 2.11. *(see [5]) Assume that $F : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable. Then F is convex function if, only if, for every $x, y \in \mathbb{R}^n$,*

$$JF(x)(y-x) \preceq F(y) - F(x). \tag{11}$$

Proposition 2.12. *(see [5]) Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable. Then F is quasi-convex function if, only if, for every $x, y \in \mathbb{R}^n$,*

$$F(y) \prec F(x) \implies JF(x)(y-x) \preceq 0. \tag{12}$$

Remark 2.13. *If F is differentiable, then from the characterization (11), it follows that convexity is a sufficient condition for pseudo-convexity. On the other hand, from the characterization (12), we obtain that pseudo-convex functions are quasi-convex. Note that the reciprocal, in both the cases, is false, (see, [4], Remark 3.3).*

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

Proposition 2.14. *(see [5]) Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a pseudo-convex function and $x \in \mathbb{R}^n$. Then x is a weak Pareto optimal point of F if and only if*

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

In the next section, we consider the proximal point algorithm for multiobjective optimization, which was introduced by Bento et al. [4]. We show that under some weaker assumptions, which were assumed in [4], the sequence generated by the proximal iteration **(BCS)** terminates at a Pareto critical point.

3. Main Results

In order to provide the convergence of the algorithm **(BCS)**, we consider the following assumptions.

Assumption 1. *There exists $i_0 \in I$ such that $f_{i_0} : \mathbb{R}^n \to \mathbb{R}$ is bounded below.*

Assumption 2. *For all $i \in I$, f_i is locally Lipschitz.*

Assumption 3. *F is differentiable.*

Assumption 4. $\liminf_{k\to+\infty} \lambda_k > 0$.

Remark 3.1. *(i) By following the proof as presented in [4], under Assumption 1, we obtain that the Algorithm defined by the iterative step **(BCS)** is well-defined.*

(ii) The main improvement of the presented work and [4], is that we consider the differentiability and locally Lipschitz properties of the objective function F instead of a continuously differentiable property, (see Assumptions (2) and (3)).

We start with the following lemmas.

Lemma 3.2. *Assume that Assumptions 1 and 4 hold. Then the sequence $\{x_k\}$ generated by Algorithm **(BCS)** is bounded.*

Proof. For each $k \in \mathbb{N} \cup \{0\}$, define $\phi_k : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ by

$$\phi_k(x) := g\Big(F(x) + I_{\Omega_k}(x)e + \frac{\lambda_k}{2}\|x - x_k\|^2 e\Big), \quad \textit{for each} \quad x \in \mathbb{R}^n. \tag{13}$$

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

$$g(F(x_{k+1})) + I_{\Omega_k}(x_{k+1}) + \frac{\lambda_k}{2}\|x_{k+1} - x_k\|^2 \le g(F(x_k)) + I_{\Omega_k}(x_k) + \frac{\lambda_k}{2}\|x_k - x_k\|^2.$$

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

$$\frac{\lambda_k}{2}\|x_{k+1} - x_k\|^2 \le g\Big(F(x_k)\Big) - g\Big(F(x_{k+1})\Big), \quad k \in \mathbb{N} \cup \{0\} \tag{14}$$

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

$$g\Big(F(x_{k+1})\Big) < g\Big(F(x_k)\Big),$$

thus by Assumption 1, we can assert that $\{g(F(x_k))\}$ is a convergent sequence.

Also, by taking the sum of inequality (14), we obtain

$$\begin{aligned}\sum_{k=0}^{l} \frac{\lambda_k}{2}\|x_{k+1} - x_k\|^2 &\le \sum_{k=0}^{l} \Big(g\Big(F(x_k)\Big) - g\Big(F(x_{k+1})\Big)\Big),\\ &= g\Big(F(x_0)\Big) - g\Big(F(x_{l+1})\Big).\end{aligned}$$

This implies that the series $\sum \frac{\lambda_k}{2}\|x_{k+1} - x_k\|^2$ is convergent. Using this one, in view of Assumption 4, we have $\sum_{k=0}^{+\infty} \|x_{k+1} - x_k\|^2 < +\infty$. Subsequently, it follows that $\{x_k\}$ is a bounded sequence. This completes the proof.
□

Lemma 3.3. *Assume that the Assumptions 1, 2, 3 and 4 are true. If $\bar{x}$ is a cluster point of $\{x_k\}$ then $\bar{x}$ is a Pareto critical point, provided that Ω_k is a convex set for each $k \in \mathbb{N} \cup \{0\}$.*

Proof. Observe that, by Assumption 2, we have Ω_k is a closed subset of $\mathbb{R}^n$, for each $k \in \mathbb{N} \cup \{0\}$. Next, since $\Omega_{k+1} \subset \Omega_k$, for each $k \in \mathbb{N} \cup \{0\}$ and $\bar{x}$ is a cluster point of $\{x_k\}$, we can show that

$$\bar{x} \in \cap_{k=0}^{+\infty}\Omega_k =: \Omega. \tag{15}$$

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

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

$$JF(\bar{x})v \prec 0. \tag{16}$$

Subsequently, there is $\delta > 0$ such that

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

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

On the other hand, since $x_{k+1} \in \arg\min_{x \in \mathbb{R}^n} \phi_k(x)$, then by (4) and (5), we obtain

$$0 \in \partial\Big(g(F(\cdot)) + I_{\Omega_k}(\cdot) + \frac{\lambda_k}{2}\|\cdot - x_k\|^2\Big)\Big(x_{k+1}\Big), \quad \text{for each} \quad k \in \mathbb{N} \cup \{0\}.$$

Thus, from (7), it follows that

$$0 \in \partial\Big(g(F(\cdot)) + \frac{\lambda_k}{2}\|\cdot - x_k\|^2)\Big)(x_{k+1}) + N_{\Omega_k}(x_{k+1}), \quad \text{for each} \quad k \in \mathbb{N} \cup \{0\}.$$

Note that, from Assumption 2 together with Proposition 2.5, we know that the function $g \circ F$ is locally Lipschitz. Thus, by applying Proposition 2.6 with $f_1(\cdot) = g(F(\cdot))$ and $f_2(\cdot) = \frac{\lambda_k}{2}\|\cdot - x_k\|^2$, the last inlusion becomes:

$$0 \in \partial(g \circ F)(x_{k+1}) + \lambda_k(x_{k+1} - x_k) + N_{\Omega_k}(x_{k+1}), \quad \text{for each} \quad k \in \mathbb{N} \cup \{0\}.$$

Subsequently, there exist sequences $\{w_k\}$, $\{v_k\}$, with $w_{k+1} \in \partial(g \circ F)(x_{k+1})$ and $v_{k+1} \in N_{\Omega_k}(x_{k+1})$ such that

$$0 = w_{k+1} + \lambda_k\Big(x_{k+1} - x_k\Big) + v_{k+1}, \quad \text{for each} \quad k \in \mathbb{N} \cup \{0\}. \tag{17}$$

As $g \circ F$ is locally Lipschitz, then by applying Proposition 2.7 with $y_k = x_k$, $f = g \circ F$ and $z_k = w_k$, for each $k \in \mathbb{N} \cup \{0\}$, obey the fact that $\{x_k\}$ is a bounded sequence we obtain that the sequence $\{w_k\}$ is bounded. Subsequently, by (17), $\{v_k\}$ is also a bounded sequence.

Next, let $\{x_{k_j}\}$ be a subsequence of $\{x_k\}$, which converges to $\bar{x}$. Moreover, let $\bar{w}$ (resp. $\bar{v}$) be a cluster point of $\{w_k\}$ (resp. of $\{v_k\}$). We can assume without loss of generality that the subsequences $\{w_{k_j}\}$ of $\{w_k\}$ and $\{v_{k_j}\}$ of $\{v_k\}$ converge, respectively, to $\bar{w}$ and $\bar{v}$ as j goes to infinity. By replacing k by k_j in (17), letting j goes to infinity and taking into account that $\lim_{j \to +\infty} \lambda_{k_j}\Big(x_{k_{j+1}} - x_{k_j}\Big) = 0$, we obtain:

$$\bar{w} = -\bar{v}. \tag{18}$$

Now, since $v_{k_j+1} \in N_{\Omega_{k_j}}(x_{k_j+1})$ and $\Omega \subset \Omega_{k_j}$, for each $j \in \mathbb{N} \cup \{0\}$, from Remark 2.4, we have

$$\langle v_{k_j}, x - x_{k_j} \rangle \leq 0, \quad \text{for each} \quad x \in \Omega. \tag{19}$$

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

$$\langle \bar{v}, x - \bar{x} \rangle \leq 0, \quad \text{for each} \quad x \in \Omega. \tag{20}$$

So, in view of equality (18), this leads to

$$\langle \bar{w}, x - \bar{x} \rangle \geq 0, \quad \text{for each} \quad x \in \Omega. \tag{21}$$

On the other hand, since $\{g(F(x_{k_j}))\}$ converges to $g(F(\bar{x}))$ as j goes to infinity, it follows from the definition of $\partial(g \circ F)$ that $\bar{w} \in \partial(g \circ F)$. Moreover, since F is differentiable, we have $\partial f_i(x) = \{\nabla f_i(x)\}$. Using the characterization (10) with $h = g \circ F$ and $x = \bar{x}$, there exists a vector $\alpha = (\alpha_1, ..., \alpha_m) \in \mathbb{R}^m_+$, with $\sum_{i=1}^m \alpha_i = 1$ such that

$$\bar{w} = \sum_{i \in I(\bar{x})} \alpha_i \nabla f_i(\bar{x}) = JF(\bar{x})^t \alpha. \tag{22}$$

For given $s > 0$, combining (21) and (22) one has

$$\langle JF(\bar{x})^t \alpha, \bar{x} + sv - \bar{x} \rangle = s\langle JF(\bar{x})^t \alpha, v \rangle = s\langle \alpha, JF(\bar{x})v \rangle \geq 0. \tag{23}$$

In contrast, by combining the definition of α with (16), we infer

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

which contradicts (23). This completes the proof. □

Next, we will show the full convergence of the considered algorithm, provided that the objective mapping is quasiconvex. To do this, we need the following concept and its properties.

Definition 3.4. *[9] A sequence $\{z_k\} \subset \mathbb{R}^n$ is said to be Fejér monotone with respect to a nonempty set U if, for all $z \in U$,*

$$\|z_{k+1} - z\| \leq \|z_k - z\|, \quad \textit{for each} \quad k \in \mathbb{N} \cup \{0\}.$$

The following result on Fejér monotone is well known.

Lemma 3.5. *Let $U \subset \mathbb{R}^n$ be a nonempty set and $\{z_k\} \subset \mathbb{R}^n$ be a Fejér monotone sequence with respect to a nonempty set U, then:*

(i) The sequence $\{z^k\}$ is bounded.

(ii) If a cluster point $\bar{z}$ of $\{z^k\}$ belongs to U, the whole sequence $\{z^k\}$ converges to $\bar{z}$ as k goes to $+\infty$.

Proof. See Schott [23], (Theorem 2.7). □

Theorem 3.6. *Assume that Assumptions 1, 2, 3 and 4 hold. If F is a quasi-convex function, then the sequence $\{x_k\}$ converges to a Pareto critical point of F.*

Proof. Remind that from (15) in the proof of Lemma 3.3, we know that $\Omega := \cap_{k=0}^{+\infty} \Omega_k$ is a nonempty closed and convex set. Let $x^* \in \Omega$ be given, consider

$$\|x_k - x^*\|^2 = \|x_{k+1} - x^*\|^2 + \|x_k - x_{k+1}\|^2 + 2\langle x_k - x_{k+1}, x_{k+1} - x^* \rangle, \text{ for each} \quad k \in \mathbb{N} \cup \{0\}. \tag{24}$$

By following the line of the proof of Lemma 3.3, we know that there exist sequences $\{w_k\}$ and $\{v_k\}$, with $w_{k+1} \in \partial(g \circ F)(x_{k+1})$ and $v_{k+1} \in N_{\Omega_k}(x_{k+1})$ such that satisfying (17), this implies

$$x_k - x_{k+1} = \frac{1}{\lambda_k}\Big(w_{k+1} + v_{k+1}\Big), \text{ for each} \quad k \in \mathbb{N} \cup \{0\}.$$

Using this equality together with (24), we have

$$\|x_k - x^*\|^2 = \|x_{k+1} - x^*\|^2 + \|x_k - x_{k+1}\|^2 - \frac{2}{\lambda_k}\langle w_{k+1} + v_{k+1}, x^* - x_{k+1} \rangle, \text{ for each} \quad k \in \mathbb{N} \cup \{0\}. \tag{25}$$

Next, taking into account that $w_{k+1} \in \partial(g \circ F)(x_{k+1})$ and using the Proposition 2.5 with $h = g \circ F$ and $x = \bar{x}$, we see that there exists a vector $\alpha^{k+1} = \left(\alpha_1^{k+1}, ..., \alpha_m^{k+1}\right) \in \mathbb{R}^m_+$, with $\sum_{i=1}^m \alpha_i^{k+1} = 1$ such that

$$w_{k+1} = \sum_{i \in I(x_{k+1})} \alpha_i^{k+1} \nabla f_i(x_{k+1}), \quad \text{for each} \quad k \in \mathbb{N} \cup \{0\}. \tag{26}$$

On the other hand, since $x^* \in \Omega$, we see that $F(x^*) \preceq F(x_{k+1})$. Subsequently, it follows from the quasi-convexity of F and Proposition 2.12, that for each $k = 0, 1, ...,$

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

Thus, by using (26), we get:

$$\langle w_{k+1}, x^* - x_{k+1} \rangle \leq 0, \quad \text{for each} \quad k \in \mathbb{N} \cup \{0\}. \tag{27}$$

From another stand point, since $v_{k+1} \in N_{\Omega_k}(x_{k+1})$ and Ω_k is a convex set(because F is quasi convex), for each $k \in \mathbb{N} \cup \{0\}$, it follows that, we have

$$\langle v_{k+1}, x^* - x_{k+1} \rangle \leq 0, \quad \text{for each} \quad k \in \mathbb{N} \cup \{0\}. \tag{28}$$

As $\|x_k - x_{k+1}\|^2 \geq 0$, for each $k \in \mathbb{N} \cup \{0\}$, the inequality (25) becomes

$$\|x_k - x^*\|^2 \geq \|x_{k+1} - x^*\|^2 - \frac{2}{\lambda_k}\langle w_{k+1}, x^* - x_{k+1} \rangle - \frac{2}{\lambda_k}\langle v_{k+1}, x^* - x_{k+1} \rangle, \quad \text{for each} \quad k \in \mathbb{N} \cup \{0\}. \tag{29}$$

Combining last inequality with (27) and (28), we conclude that

$$\|x_{k+1} - x^*\| \leq \|x_k - x^*\|, \quad \text{for each} \quad k \in \mathbb{N} \cup \{0\}. \tag{30}$$

This means $\{x_k\}$ is a Fejér monotone to Ω. Thus, in view of Lemma 3.5, we conclude that the sequence $\{x_k\}$ converges to $\bar{x}$ as k goes to $+\infty$. Finally, by Lemma 3.3, we have $\bar{x}$ is a Pareto critical point of problem (1). This completes the proof. □

Corollary 3.7. *Under Assumption 3, if F is pseudo-convex or convex, then the sequence $\{x_k\}$ converges to a weak Pareto optimal point of F.*

Proof. If F is pseudo-convex or convex, in particular, F is quasi-convex (see Remark 2.13) and the corollary is a consequence of the previous theorem, then the sequence $\{x_k\}$ converges to $\bar{x}$ as k goes to $+\infty$ and $\bar{x}$ is a Pareto critial point. By Proposition 2.14, under pseudo-convexity criticality is equivalent to weak optimality, which implies that $\bar{x}$ is a weak Pareto optimal point. □

4. Conclusion

This paper considered the proximal point algorithm for multiobjective optimization, which was introduced by Bento et al. [4]. The main aim of this paper is to relax the conditions on the considered objective function. Indeed, the work presented in this article extends the class of functions from continuously differentiable to differentiable and locally Lipschitz. We do think that this work is extendable for the problems involving non-differentiable functions by defining the Pareto critical points using directional derivatives.

References

[1] Apolinário, H.C.F., Papa Quiroz, E.A., Oliveira, P.R., A scalarization proximal point method for quasiconvex multiobjective minimization. J. of Global Optim,. 64(1), (2016) 79-96.

[2] Bagirov, A.M., Karmitsa, N., Mäkelä M.: Introduction to Nonsmooth Optimization: Theory, Practice and Software Springer, Cham, Heidelberg (2014)

[3] Bello Cruz, J.Y., Lucambio Perez, L.R. and Melo, J.G.: Convergence of the projected gradient method for quasiconvex multiobjective optimization. Nonlinear Anal., 74, (2011) 5268-5273.

[4] Bento, G.C., Cruz Neto, J.X., Soubeyran, A.: A proximal point-type method for multicriteria optimization. Set-Valued Vari. Anal., 22(3), (2014) 557-573.

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

[6] Bolte, J., Daniilidis, A., Lewis, A., Shiota, M.: Clarke subgradients of stratifiable functions. SIAM J. of Optim., 18(2), (2007) 556-572.

[7] Bonnel, H., Iusem, A.N., Svaiter, B.F.: Proximal methods in vector optimization. SIAM J. of Optim., 15(4), (2005) 953-970.

[8] Burachik, R. S., Kaya, C. Y., Rizvi, M.M.: A new scalarization technique to approximate Pareto fronts of problems with disconnected feasible sets, J. Optim. Theory Appl., 162, (2014) 428-446.

[9] Bauschke H. H., Combettes P. L.: A weak-to-strong convergence principle for Fejer monotone methods in Hilbert spaces, Mathematics of Operations Research 26, (2001), 248–264.

[10] Chen, J., Ansari, Q. H., Liou, Y. C., Yao, J. C.: A proximal point algorithm based on decomposition method for cone constrained multiobjective optimization problems, Comput. Optim. Appl. 65, no. 1 (2016): 289-308.

[11] Cruz Neto, J. X., Silva, G. J. P., Ferreira, O. P., Lopes, J. O.: A subgradient method for multiobjective optimization, Comput. Optim. Appl., 54, (2013) 461-472.

[12] Durea M., Strugariu, R.: Some remarks on proximal point algorithm in scalar and vectorial cases, Nonlinear Funct. Anal. Appl., 15, (2010) 307-319.

[13] Eichfelder, G.: An adaptive scalarization method in multiobjective optimization, SIAM J. Optim., 19, (2009) 1694-1718.

[14] Fliege, J. Grana Drummond, L. M., Svaiter, B. F.: Newton's method for multiobjective optimization, SIAM J. Optim., 20, (2009) 602-626.

[15] Gal, T., Hanne, T.: On the development and future aspects of vector optimization and MCDM. A tutorial. In: Multicriteria Anal., (Coimbra, 1994), Springer, Berlin (1997) 130-145.

[16] Gregorio, R., Oliveira, P.R.: A Logarithmic-quadratic proximal point scalarization method for multiobjective programming. J. of Global Optim., 49, (2010) 361-378.

[17] Ishizuka, Y., Optimality conditions for directionally differentiable multiobjective programming problems. J. of Optim. Theory and Appl., 72, (1992) 91-111.

[18] Ishizuka, Y. and Tuan, H. D.: Directionally differentiable multiobjective optimization involving discrete inclusions. J. of Optim. Theory and Appl., 88, (1996) 585-616.

[19] Luc, D.T.: Theory of Vector Optimization. Lecture Notes in Economics and Mathematical Systems. Springer, Berlin (1989)

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

[21] Nguyen, M. H., Luu, D. V.: On necessary conditions for efficiency in directionally differentiable optimization problems. Cahiers de la Maison des Sciences Economiques (2006)

[22] Rockafellar, R.T., Wets, R.: Variational Analysis. Grundlehren der Mathematischen Wissenschafte. Springer Verlag, New York (1998)

[23] Schott, D.: Basic properties of Fejér monotone sequences. Rostocker Mathematische Kolloquium. 49, (1995) 57-74.

[24] Ward, D.E.: Directional derivative calculus and optimality conditions in nonsmooth mathematical programming. J. of Inform. and Optim. Sciences 10, (1989) 81-96.