J. Nonlinear Var. Anal. 8 (2024), No. 4, pp. 517-531
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.8.2024.4.03

AN INEXACT NONMONOTONE PROJECTED GRADIENT METHOD FOR CONSTRAINED MULTIOBJECTIVE OPTIMIZATION

XIAOPENG ZHAO[1], HUIJIE ZHANG[1], YONGHONG YAO[1,2,*]

[1]*School of Mathematical Sciences, Tiangong University, Tianjin 300387, China*
[2]*Center for Advanced Information Technology, Kyung Hee University, Seoul 02447, Korea*

Abstract. In this paper, we consider an inexact projected gradient method equipped with a nonmonotone line search rule for smooth constrained multiobjective optimization. In this method, a new nonmonotone line search technique proposed here is employed and the relative errors on the search direction is admitted. We demonstrate that this method is well-defined. Then, we prove that each accumulation point of the sequence generated by this method is Pareto stationary and analyze the convergence rate of the algorithm. When the objective function is convex, the convergence of the sequence to a weak Pareto optimal point of the problem is established.

Keywords. Gradient method; Multiobjective optimization; Nonmonotone line search; Pareto optimality.

1. Introduction

In this paper, we consider the constrained multiobjective optimization problem of the form:

$$\min_{x\in C} F(x), \tag{1.1}$$

where $C \subseteq \mathbb{R}^n$ is a closed and convex set and $F = (f_1, \cdots, f_m)^\top : \mathbb{R}^n \to \mathbb{R}^m$ is a vector-valued function and continuously differentiable on an open superset of C. Such optimization problem has a wide range of applications in many fields, such as economy, finance, engineering, management science, location theory, game theory and so on; see, e.g., [2, 9, 24, 27, 28].

In terms of numerical optimization algorithms, first-order methods are the mainstream algorithms for solving large-scale problems due to their efficiency and computational simplicity. In the present work, we focus on one first-order method, namely the projected gradient method. In scalar optimization, the projected gradient method and its modified versions are often used to solve nonsmooth optimization problems, feasibility problems, variational inequalities, and so on [1, 3, 31, 33]. For the vector optimization, Graña Drummond and Iusem [16] proposed two projected gradient methods for vector optimization problems. For one method, they showed the stationarity of the cluster points without convexity assumptions. For the other method, they proved the convergence to a weakly efficient solution when the objective function is convex. Bello Cruz et al. [4] considered the projected gradient method for quasiconvex multiobjective

*Corresponding author.
E-mail address: zhaoxiaopeng.2007@163.com (X. Zhao), zhanghuijie0826@163.com (H. Zhang), yyhtgu@hotmail.com (Y. Yao).

Received 21 June 2023; Accepted 30 December 2023; Published 1 May 2024.

optimization. They showed the convergence to a stationary point and further to a weakly efficient solution when the multiobjective function is pseudoconvex. In [13], Fliege et al. analyzed the convergence rate of the gradient method for smooth unconstrained multiobjective optimization. We note that these methods are all descent methods, that is, the Armijo-type line search rule is adopted to calculate the stepsize and the sequence of the objective function values is monotone decreasing. While, this monotonicity can considerably slow the convergence rate of the algorithms. In [7], by employing a scalarization process, Brito et al. proposed a projected subgradient method for constrained nondifferentiable convex multiobjective optimization problems and obtained the convergence to a Pareto optimal point of the problem. Recently, based on the Plastria subdifferential, we considered a projected subgradient method for solving constrained nondifferentiable quasiconvex multiobjective optimization problems in [30] and established the convergence to a Pareto optimal point. However, in scalarization approaches, the choice of the parameters is not known in advance, which leaves the modeler and the decision-maker with the burden of choosing them. Moreover, for some problems, unproper choices of the parameters may give rise to unbounded scalar problems [11, 17]. We see that using nonmonotone line search techniques is a good alternative, which does not require parameter information and allows some increase of objective function values in some iterations to improve the convergence speed of the algorithms. In [25], by extending the max-type nonmonotone line search given by Grippo et al. [19] in the scalar context to vector optimization case, Qu et al. proposed nonmonotone gradient methods for convex vector optimization, and established the global convergence and local linear convergence results for the methods. Fazzio and Schuverdt [10] considered a nonmonotone projected gradient method for multiobjective optimization based on the average-type nonmonotone line search technique of [29], and proved the convergence to a weakly Pareto optimal solution when the multiobjective function is convex. Recently, based on the average-type nonmonotone line search, we considered a projected gradient method with exogenously given square summable sequence in the computation of the search direction for the convex multiobjective optimization [32].

Nonetheless, the research on nonmonotone algorithms for solving multiobjective optimization problems is still insufficient. Note that, for scalar optimization, Huang et al. [21] proposed a nonmonotone line search rule, which was verified to be an improved version of the average-type nonmonotone line search technique. Inspired by this, in the present paper, we extend it to the case of multiobjective optimization and propose a new nonmonotone line search strategy.

On the other hand, most of the existing algorithms for vector optimization are exact algorithms. That is, in each iteration, the search direction is obtained by solving the auxiliary subproblem exactly, which increases the computational cost of the algorithm. In view of this, the inexact gradient methods for vector optimization have been considered in some work [6, 12, 15, 18], in which an approximation of the exact search direction was computed at each iteration. Whereas, these methods are descent methods. In this paper, based on the nonmonotone line search rule proposed here, we consider an inexact nonmonotone projected gradient method for multiobjective optimization problem (1.1) and analyze the convergence of the method.

The outline of this paper is as follows. In Section 2, we present some notations and preliminary results which are needed in this work. In Section 3, we propose the method and demonstrate its well-definedness. Some properties of the method are also investigated. Section

4 concerns the convergence analysis of the proposed method without convexity assumptions. It is proved that each accumulation point of the sequence generated by the method is Pareto stationary and the method has a convergence rate of $O(\frac{1}{\sqrt{k}})$. In Section 5, we show the convergence of the method to a weak Pareto solution in the convex case. Finally, some conclusions are given in Section 6.

2. Notations and Preliminary Results

The notations used in the present paper are standard (cf. [22,26]). We denote by $\langle\cdot,\cdot\rangle$ the inner product and by $\|\cdot\|$ the corresponding norm of $\mathbb{R}^n$. $\mathbb{N}$ denotes the set of all positive integers, and $\mathbb{N}^* := \mathbb{N}\cup\{0\}$. The transpose sign is denoted by $\top$.

For a function $f:\mathbb{R}^n\to\overline{\mathbb{R}} := \mathbb{R}\cup\{+\infty\}$, the effective domain of f is denoted by $\mathrm{dom} f := \{x\in\mathbb{R}^n : f(x) < +\infty\}$, and f is called proper if $\mathrm{dom} f \neq \emptyset$. A proper function f is called convex if, for all $x,y\in\mathrm{dom} f$ and for all $t\in[0,1]$, $f(tx+(1-t)y)\le tf(x)+(1-t)f(y)$, while f is called strictly convex if the above inequality is strict for any $x,y\in\mathrm{dom} f$ with $x\neq y$ and $t\in(0,1)$. We call a proper function f is strongly convex with modulus $\tau>0$ (or τ-strongly convex) if, for any $x,y\in\mathrm{dom} f$ and $t\in[0,1]$, $f(tx+(1-t)y)\le tf(x)+(1-t)f(y)-\frac{\tau}{2}t(1-t)\|x-y\|^2$. A proper function f is said to be differentiable at $\bar{x}\in\mathrm{dom} f$ if there exists a vector $v\in\mathbb{R}^n$ with the property that

$$\lim_{x\to\bar{x}}\frac{f(x)-f(\bar{x})-\langle v,x-\bar{x}\rangle}{\|x-\bar{x}\|}=0.$$

Such a v, if it exists, is called the gradient of f at $\bar{x}$ and is denoted by $\nabla f(\bar{x})$. The subdifferential of a proper convex function f at $\bar{x}\in\mathrm{dom} f$ is defined by

$$\partial f(\bar{x}) := \{x^*\in\mathbb{R}^n : f(\bar{x})+\langle x^*,x-\bar{x}\rangle\le f(x),\quad \forall\, x\in\mathrm{dom} f\},$$

and we say that f is subdifferentiable at $x\in\mathrm{dom} f$ if $\partial f(x)\neq\emptyset$. If f is differentiable at $\bar{x}$, then $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$.

It is known that if a τ-strongly convex function f is subdifferentiable at $x\in\mathrm{dom} f$, then, for all $x,y\in\mathrm{dom} f$ and $x^*\in\partial f(x)$,

$$f(y)\ge f(x)+\langle x^*,y-x\rangle+\frac{\tau}{2}\|y-x\|^2. \tag{2.1}$$

The function f is said to be Lipschitz continuous with Lipschitz constant $L>0$ if $|f(x)-f(y)|\le L\|x-y\|$ for all $x,y\in\mathrm{dom} f$.

Let $\mathbb{R}^m_+$ and $\mathbb{R}^m_{++}$ denote the nonnegative orthant and positive orthant of $\mathbb{R}^m$, respectively. We consider the partial order $\preceq$ ($\prec$) induced by $\mathbb{R}^m_+$ ($\mathbb{R}^m_{++}$): for two vectors $x,y\in\mathbb{R}^m$, $x\preceq y$ ($x\prec y$) if and only if $y-x\in\mathbb{R}^m_+$ ($y-x\in\mathbb{R}^m_{++}$). The multiobjective function F is said to be continuously differentiable (convex, respectively, strictly convex), if each component function f_i with $i=1,\cdots,m$ is continuously differentiable (convex, respectively, strictly convex). Let $J_F(x)$ denote the Jacobian matrix of the vector-valued function F in (1.1) at x, that is, $J_F(x) = (\nabla f_1(x),\cdots,\nabla f_m(x))^\top$. The definition of the Pareto optimality for multiobjective optimization problem (1.1) is recalled as follows.

Definition 2.1. A point $x^*\in C$ is said to be

(a) a Pareto optimal point of (1.1) if there does not exist $x\in C$ such that $F(x)\preceq F(x^*)$ and $F(x)\neq F(x^*)$;

(b) a weak Pareto optimal point of (1.1) if there does not exist $x\in C$ such that $F(x)\prec F(x^*)$;

(c) a Pareto stationary point of (1.1) if $J_F(x^*)(C-x^*)\cap(-\mathbb{R}^m_{++})=\emptyset$.

It is known that every Pareto optimal point is also a weak Pareto optimal point, and each weak Pareto optimal point is also a Pareto stationary point. Conversely, if F is convex, then Pareto stationarity implies weak Pareto optimality. Also, it can be verified by definition that if F is strictly convex, then every weak Pareto optimal point is also a Pareto optimal point. Indeed, if we suppose that a weak Pareto optimal point x^* is not a Pareto optimal point, then there exists $x\in\mathbb{R}^n$ such that $F(x)\preceq F(x^*)$ and $F(x)\neq F(x^*)$. Consequently, for all $i=1,\cdots,m$ and $\lambda\in(0,1)$, one can obtain from the strict convexity of f_i that

$$f_i(x^*+\lambda(x-x^*))<\lambda f_i(x)+(1-\lambda)f_i(x^*).$$

It follows that

$$\frac{f_i(x^*+\lambda(x-x^*))-f_i(x^*)}{\lambda}<f_i(x)-f_i(x^*)\leq 0,$$

which further yields that $F(x^*+\lambda(x-x^*))\prec F(x^*)$. This is a contradiction to the weak Pareto optimality of x^*.

Finally, we recall the following quasi-Fejér convergence theorem, which will be used in the convergence analysis of the method.

Definition 2.2. A sequence $\{u_k\}\subseteq\mathbb{R}^n$ is said to be quasi-Fejér convergent to a nonempty set $U\subseteq\mathbb{R}^n$ if, for each $u\in U$, there exists a sequence $\{\varepsilon_k\}\subseteq\mathbb{R}_+$ with $\sum_{k=0}^{\infty}\varepsilon_k<\infty$ such that

$$\|u_{k+1}-u\|^2\leq\|u_k-u\|^2+\varepsilon_k.$$

Proposition 2.1. [8, Theorem 1] *If $\{u_k\}\subseteq\mathbb{R}^n$ is quasi-Fejér convergent to a nonempty set $U\subseteq\mathbb{R}^n$, then $\{u_k\}$ is bounded. Furthermore, if a cluster point $\bar{u}$ of $\{u_k\}$ belongs to U, then* $\lim_{k\to\infty}u_k=\bar{u}$.

3. The Algorithm and Properties

In this section, we propose the inexact projected gradient method equipped with a nonmonotone line search rule introduced for multiobjective optimization problem (1.1). Some properties of the algorithm are also presented.

Given a parameter $\beta>0$, for $x\in C$, define the function $\varphi_x:\mathbb{R}^n\to\overline{\mathbb{R}}$ by

$$\varphi_x(v)=\beta\max_{i\in\{1,\cdots,m\}}\langle\nabla f_i(x),v\rangle+\frac{\|v\|^2}{2},\quad\forall v\in\mathbb{R}^n.$$

In view of the strong convexity of φ_x, it has a unique minimum point on the closed and convex set $C-x$, denoted by $v(x)$, i.e.,

$$v(x):=\text{argmin}_{v\in C-x}\varphi_x(v).\tag{3.1}$$

We denote the optimal value of (3.1) as

$$\theta(x):=\beta\max_{i\in\{1,\cdots,m\}}\langle\nabla f_i(x),v(x)\rangle+\frac{\|v(x)\|^2}{2}.$$

Since $0\in C-x$, we have $\theta(x)\leq 0$. The following characterization of Pareto stationary points of the problem (1.1) in terms of $v(\cdot)$ and $\theta(\cdot)$ follows from [16, Proposition 3].

Lemma 3.1. *For $x \in C$, the following conditions are equivalent:*
(i) *The point x is not a Pareto stationary point of* (1.1)*;*
(ii) $\theta(x) < 0$*;*
(iii) $v(x) \neq 0$.
In particular, $x \in C$ is a Pareto stationary point of (1.1) *if and only if $\theta(x) = 0$.*

For the following continuity of the function $\theta(\cdot)$, one can refer to [14, Proposition 3.4].

Lemma 3.2. *The function $\theta : C \to \mathbb{R}$ is continuous.*

In exact projected gradient methods, the minimizer $v(x)$ of problem $\min_{v \in C-x} \varphi_x(v)$ needs to be calculated in each iteration as the search direction. From a numerical perspective, it is interesting to work with the inexact solution of this problem. In this paper, we consider the approximate search directions as considered in [12, 15, 18].

Definition 3.1. Let $x \in C$ and $\sigma \in [0,1)$. We say that the vector $v \in C-x$ is a σ-approximate direction at x if $\varphi_x(v) \leq (1-\sigma)\theta(x)$.

Note that the exact direction $v(x)$ is always a σ-approximate direction at x for any $\sigma \in [0,1)$. Since φ_x is strongly convex with modulus 1 and $0 \in \partial \varphi_x(v(x))$, we have by (2.1) that

$$\varphi_x(v) - \theta(x) \geq \frac{1}{2}\|v - v(x)\|^2, \quad \text{for any } v \in \mathbb{R}^n. \tag{3.2}$$

Particularly, for a σ-approximate direction v at x, it follows from (3.2) that $\|v - v(x)\|^2 \leq 2\sigma|\theta(x)|$, which establishes the degree of proximity between a σ-approximate direction v and the exact direction $v(x)$ in terms of the optimal value $\theta(x)$.

From Definition 3.1 and Lemma 3.1, the following characterization of Pareto stationary points of the problem (1.1) given by approximate directions can be acquired, which is the theoretical basis for the stopping criterion of Algorithm 3.1.

Proposition 3.1. *Given $x \in C$ and $\sigma \in [0,1)$, x is a Pareto stationary point of* (1.1) *if and only if $v = 0$ is a σ-approximate direction at x.*

Proof. If x is a Pareto stationary point of (1.1), then $\theta(x) = 0$ thanks to Lemma 3.1. Therefore, it follows directly from Definition 3.1 that $v = 0$ is a σ-approximate direction at x. Conversely, if $v = 0$ is a σ-approximate direction at x, then $0 \leq (1-\sigma)\theta(x)$, so $\theta(x) = 0$, which is due to $\theta(x) \leq 0$. Then, applying Lemma 3.1 again yields that x is a Pareto stationary point of (1.1). □

We describe the method considered here as follows.

Algorithm 3.1

Step 1 Choose parameters $\beta > 0$, $\sigma \in [0,1)$, $0 \leq \eta_{\min} \leq \eta_{\max} < 1$, $\delta_{\max} < 1$, and $0 < \delta_{\min} < (1-\eta_{\max})\delta_{\max}$. Let $x_0 \in C$ be an arbitrary initial point. Set $C_0 = F(x_0), Q_0 = 1$ and $k = 0$.

Step 2 If $\nabla f_i(x_k) = 0$ for some $i \in \{1, \cdots, m\}$, then **stop**. Otherwise, compute the inexact search direction $v_k \in C - x_k$ such that $\varphi_{x_k}(v_k) \leq (1-\sigma)\theta(x_k)$.

Step 3 If $v_k = 0$, then **stop**. Otherwise, proceed to Step 4.

Step 4 Choose $\eta_k \in [\eta_{\min}, \eta_{\max}]$ and define

$$Q_{k+1} = \eta_k Q_k + 1. \tag{3.3}$$

Choose $\delta_k \in \left[\delta_{\min}, \frac{\delta_{\max}}{Q_{k+1}}\right]$. Compute a stepsize $\alpha_k \in (0,1]$ as the maximum of $\{\frac{1}{2^j} : j \in \mathbb{N}^*\}$ such that

$$\frac{\eta_k Q_k C_k + F(x_k + \alpha_k v_k)}{Q_{k+1}} \preceq C_k + \delta_k \alpha_k JF(x_k) v_k. \tag{3.4}$$

Step 5 Define

$$x_{k+1} = x_k + \alpha_k v_k \tag{3.5}$$

and

$$C_{k+1} = \frac{\eta_k Q_k C_k + F(x_{k+1})}{Q_{k+1}}. \tag{3.6}$$

Set $k := k+1$ and go back to Step 2.

Observe that, for each k, one can equivalently rewrite C_{k+1} as

$$C_{k+1} = \frac{(\eta_k Q_k + 1)C_k + F(x_{k+1}) - C_k}{Q_{k+1}} = C_k + \frac{F(x_{k+1}) - C_k}{Q_{k+1}}. \tag{3.7}$$

By the definition of Q_{k+1} in (3.3), one can compute that

$$Q_{k+1} = 1 + \sum_{j=0}^{k} \prod_{l=0}^{j} \eta_{k-l} \le 1 + \sum_{j=0}^{k} \eta_{\max}^{j+1} \le \sum_{j=0}^{\infty} \eta_{\max}^{j} \le \frac{1}{1-\eta_{\max}}. \tag{3.8}$$

Then, one has $\delta_{\min} < (1-\eta_{\max})\delta_{\max} \le \frac{\delta_{\max}}{Q_{k+1}}$, so the parameter δ_k can be selected. Moreover, by the definition of Q_{k+1} in (3.3), it can be seen that (3.4) is equivalent to

$$F(x_k + \alpha_k v_k) \preceq C_k + Q_{k+1}\delta_k \alpha_k JF(x_k) v_k. \tag{3.9}$$

If $\eta_k = 0$ for each k, then $C_k = F(x_k)$ and $Q_k = 1$, and the line search (3.9) (if further δ_k is a constant) reduces to the Armijo line search that was adopted in the descent methods [4, 6, 15, 16, 18]. In addition, the average-type nonmonotone line search that was considered in [10, 32] can be regarded as a special form of (3.9) with $\delta_k = \frac{\delta}{Q_{k+1}}$, $\delta \in [\delta_{\min} Q_{k+1}, \delta_{\max}]$.

Note that, if Algorithm 3.1 stops at iteration k, then x_k is a Pareto stationary point of (1.1). Indeed, if $\nabla f_{i_0}(x_k) = 0$ for some $i_0 \in \{1, \cdots, m\}$, then $\langle \nabla f_{i_0}(x_k), x - x_k\rangle = 0$ for any $x \in C$. Thus $J_F(x_k)(C - x_k) \cap (-\mathbb{R}^m_{++}) = \emptyset$, i.e., x_k is a Pareto stationary point. In the case that $v_k = 0$, the Pareto stationarity of x_k follows immediately from Proposition 3.1. Therefore, for any nonstationary point x_k, we have $v_k \ne 0$ and $\nabla f_i(x_k) \ne 0$ for each $i \in \{1, \cdots, m\}$. Then, it can be verified from

$$\varphi_{x_k}(v_k) = \beta \max_{i \in \{1,\cdots,m\}} \langle \nabla f_i(x_k), v_k\rangle + \frac{\|v_k\|^2}{2} \le (1-\sigma)\theta(x_k) \le 0$$

that

$$J_F(x_k) v_k \prec 0 \quad \text{and} \quad \|v_k\| \le 2\beta \|\nabla f_i(x_k)\| \ \text{ for each } i \in \{1, \cdots, m\}. \tag{3.10}$$

Next, we show the well-definedness of the line search in Algorithm 3.1.

Proposition 3.2. *The line search (3.4) in Algorithm 3.1 is well-defined and $F(x_k) \preceq C_k$ for each $k \in \mathbb{N}^*$.*

Proof. We prove this result by mathematical induction. In the case that $k=0$, one has $C_0 = F(x_0)$ and $Q_0 = 1$, and the line search rule (3.4) needs to calculate α_0 such that

$$F(x_0+\alpha_0 v_0) \preceq F(x_0) + Q_1\delta_0\alpha_0 JF(x_0)v_0. \tag{3.11}$$

From $Q_1\delta_0 \le \delta_{\max} < 1$ and $JF(x_0)v_0 \prec 0$, we can see that (3.11) is the standard Armijo line search. Thus (3.11) is true, i.e., such a α_0 exists. Fix any $K\in\mathbb{N}$ and suppose that for all $1\le k\le K$, there exists α_k such that

$$F(x_{k+1}) = F(x_k+\alpha_k v_k) \preceq C_k + Q_{k+1}\delta_k\alpha_k JF(x_k)v_k, \tag{3.12}$$

and $F(x_k)\preceq C_k$.

Now, we consider the case that $k=K+1$. Let us first show that $F(x_{K+1})\preceq C_{K+1}$. To this end, for each $i\in\{1,\cdots,m\}$, we define the function $D^i_{K+1}:\mathbb{R}^+\to\mathbb{R}$ by $D^i_{K+1}(t) = \frac{tC^i_K+f_i(x_{K+1})}{t+1}$, where we use the notation C^i_k to denote the jth component of C_k for any k. Then,

$$D'^i_{K+1}(t) = \frac{C^i_K - f_i(x_{K+1})}{(t+1)^2}. \tag{3.13}$$

Since $JF(x_K)v_K \prec 0$, it follows from (3.12) that $F(x_{K+1})\preceq C_K$, which together with (3.13) implies that $D'^i_{K+1}(t)\ge 0$, i.e., D^i_{K+1} is an increasing function of t. Thus

$$f_i(x_{K+1}) = D^i_{K+1}(0) \le D^i_{K+1}(\eta_K Q_K) = C^i_{K+1},$$

which is equivalent to that $F(x_{K+1})\preceq C_{K+1}$ as $i\in\{1,\cdots,m\}$ is arbitrary.

Next, we prove line search (3.4) holds for $K+1$ by contradiction. Suppose that there does not exist α_{K+1} such that (3.4) holds for $k=K+1$. Then, for any positive integer $l\in\mathbb{N}$, there exists $i(l)\in\{1,\cdots,m\}$ such that

$$\begin{aligned} f_{i(l)}\left(x_{K+1}+\frac{1}{2^l}v_{K+1}\right) &> C^{i(l)}_{K+1} + Q_{K+2}\delta_{K+1}\frac{1}{2^l}\nabla f_{i(l)}(x_{K+1})^\top v_{K+1} \\ &\ge f_{i(l)}(x_{K+1}) + Q_{K+2}\delta_{K+1}\frac{1}{2^l}\nabla f_{i(l)}(x_{K+1})^\top v_{K+1}. \end{aligned}$$

Without loss of generality, we assume that $i(l)=\hat{i}\in\{1,\cdots,m\}$ for all l. Then, by the mean-value theorem, it can be concluded that there exists some $\lambda\in(0,1)$ such that

$$\frac{1}{2^l}\nabla f_{\hat{i}}\left(x_{K+1}+\frac{\lambda}{2^l}v_{K+1}\right)^\top v_{K+1} > Q_{K+2}\delta_{K+1}\frac{1}{2^l}\nabla f_{\hat{i}}(x_{K+1})^\top v_{K+1},$$

so

$$\left(\nabla f_{\hat{i}}\left(x_{K+1}+\frac{\lambda}{2^l}v_{K+1}\right) - \nabla f_{\hat{i}}(x_{K+1})\right)^\top v_{K+1} > (Q_{K+2}\delta_{K+1}-1)\nabla f_{\hat{i}}(x_{K+1})^\top v_{K+1}.$$

Letting $l\to\infty$ in the above inequality and noting the continuous differentiability of $f_{\hat{i}}$, we obtains that $(Q_{K+2}\delta_{K+1}-1)\nabla f_{\hat{i}}(x_{K+1})^\top v_{K+1}\le 0$. Since $Q_{K+2}\delta_{K+1}\le\delta_{\max}<1$, it follows that $\nabla f_{\hat{i}}(x_{K+1})^\top v_{K+1}\ge 0$, which contradicts the fact that $J_F(x_{K+1})^\top v_{K+1}\prec 0$. Consequently, (3.4) is true for $k=K+1$ and the proof is complete. □

Thanks to Proposition 3.2, Algorithm 3.1 is well-defined. From now on, we assume that the sequence $\{x_k\}$ generated by Algorithm 3.1 is infinite. By the definition of the iteration $x_{k+1}=x_k+\alpha_k v_k$ and the convexity of C, the feasibility of $\{x_k\}$, i.e., $\{x_k\}\subseteq C$ can be obtained easily. Moreover, for each k, from the definition of C_{k+1} in (3.6) and $J_F(x_k)v_k \prec 0$, one obtains by (3.4) that $C_{k+1} \preceq C_k$, that is, $\{C_k\}$ is a nonincreasing sequence in $\mathbb{R}^m$.

4. Convergence Analysis: The Nonconvex Case

In this section, we analyze the convergence property of Algorithm 3.1 without the convexity assumption on F. We show that each accumulation point of $\{x_k\}$, if any, is a Pareto stationary point of problem (1.1). Then, we prove that Algorithm 3.1 has a convergence rate of the order of $\frac{1}{\sqrt{k}}$. For this, additional assumption is needed. It is known that the assumption $\{x \in C : F(x) \preceq F(x_k), \forall k \in \mathbb{N}^*\} \neq \emptyset$ is standard for guaranteeing the existence of efficient solutions for vector optimization problems [22], and has been frequently used in the convergence analysis of the algorithms [4, 5, 15, 23, 30]. While, for nonmonotone algorithms, by considering C_k instead of $F(x_k)$, we use the following assumption:

(A1) The set $T := \{x \in C : F(x) \preceq C_k, \forall k \in \mathbb{N}^*\}$ is nonempty.

Recently, this assumption was adopted in [10,32] to prove the convergence of their algorithms proposed therein.

Theorem 4.1. *Assume that (A1) holds. Then, every accumulation point, if any, of the sequence $\{x_k\}$ generated by Algorithm 3.1 is a Pareto stationary point of problem* (1.1).

Proof. Let x^* be an accumulation point of $\{x_k\}$, and let $\{x_{j_k}\}$ be a subsequence of $\{x_k\}$ such that $\lim_{k\to\infty} x_{j_k} = x^*$. Since $\{x_k\} \subseteq C$ and C is closed, it follows that $x^* \in C$. We prove the conclusion of the theorem in the following two cases:

$$\text{(i)} \limsup_{k\to\infty} \alpha_{j_k} > 0, \quad \text{and} \quad \text{(ii)} \limsup_{k\to\infty} \alpha_{j_k} = 0.$$

First, assume that (i) holds. For each $k \in \mathbb{N}^*$, since $J_F(x_k)v_k \prec 0$ and $\delta_k \geq \delta_{\min}$, it follows from (3.4) and (3.6) that

$$C_{k+1} \preceq C_k + \delta_k \alpha_k J_F(x_k)v_k \preceq C_k + \delta_{\min} \alpha_k J_F(x_k)v_k. \tag{4.1}$$

By the assumption (A1), it can be concluded that $\{C_k\}$ is bounded from below, say $R \leq C_k$ for all k. This, together with (4.1), implies that, for any $l \in \mathbb{N}$,

$$\sum_{k=0}^{l} \delta_{\min}(-\alpha_k J_F(x_k)v_k) \preceq \sum_{k=0}^{l} (C_k - C_{k+1}) = C_0 - C_{l+1} \preceq C_0 - R. \tag{4.2}$$

Letting $l \to \infty$ in (4.2), it follows that $\sum_{k=0}^{\infty} \delta_{\min}(-\alpha_k J_F(x_k)v_k) \prec +\infty$. Thus

$$\lim_{k\to\infty} \alpha_k J_F(x_k)v_k = 0. \tag{4.3}$$

By assumption (i), there exists a subsequence $\{\alpha_{j_{l_k}}\}$ of α_{j_k} such that $\lim_{k\to\infty} \alpha_{j_{l_k}} = \alpha > 0$. Combining this with (4.3) yields that $\lim_{k\to\infty} J_F(x_{j_{l_k}})v_{j_{l_k}} = 0$, which can be written componentwisely as $\lim_{k\to\infty} \nabla f_i(x_{j_{l_k}})^\top v_{j_{l_k}} = 0$ for each $i = 1, \cdots, m$. Consequently, we have that, for each $i \in \{1, \cdots, m\}$,

$$\begin{aligned} 0 &\leq \limsup_{k\to\infty} \beta \max_{i\in\{1,\cdots,m\}} \langle \nabla f_i(x_{j_{l_k}}), v_{j_{l_k}} \rangle + \frac{\|v_{j_{l_k}}\|^2}{2} = \limsup_{k\to\infty} \varphi_{x_{j_{l_k}}}(v_{j_{l_k}}) \\ &\leq \limsup_{k\to\infty} (1-\sigma)\theta(x_{j_{l_k}}) = (1-\sigma)\theta(x^*), \end{aligned} \tag{4.4}$$

where the second inequality is due to the fact that $v_{j_{l_k}}$ is a σ-approximate direction and the second equality follows from Lemma 3.2. Then, we obtain from (4.4) that $\theta(x^*) = 0$, so x^* is a Pareto stationary point of problem (1.1) by Lemma 3.1.

Now, we assume that (ii) holds. Since $\{x_{j_k}\}$ is bounded, it follows from the continuous differentiability of F that $J_F(x_{j_k})$ is bounded. Using (3.10), one obtains that $\{v_{j_k}\}$ is also bounded. Considering assumption (ii), one sees that there exists sequence $\{j_{r_k}\}$ of $\{j_k\}$ such that

$$\lim_{k\to\infty} x_{j_{r_k}} = x^*, \quad \lim_{k\to\infty} v_{j_{r_k}} = v^*, \quad \text{and} \quad \lim_{k\to\infty} \alpha_{j_{r_k}} = 0. \tag{4.5}$$

From the σ-approximation of $\{v_{j_{r_k}}\}$ and the characterization of Pareto stationary points in Lemma 3.1, one can obtain that

$$\beta \max_{i\in\{1,\cdots,m\}} \langle \nabla f_i(x_{j_{r_k}}), v_{j_{r_k}}\rangle \le \varphi_{x_{j_{r_k}}}(v_{j_{r_k}}) \le (1-\sigma)\theta(x_{j_{r_k}}) < 0.$$

Then, letting $k \to \infty$ in the above inequality and taking into account (4.5), we have

$$\beta \max_{i\in\{1,\cdots,m\}} \langle \nabla f_i(x^*), v^*\rangle \le (1-\sigma)\theta(x^*) \le 0. \tag{4.6}$$

Now fix a positive integer q. Since $\lim_{k\to\infty} \alpha_{j_{r_k}} = 0$, there exists some $K \in \mathbb{N}$ such that $\alpha_{j_{r_k}} < \frac{1}{2^q}$ for all $k \ge K$, which means that $\frac{1}{2^q}$ does not satisfy the line search (3.4) with $x_{j_{r_k}}$, that is,

$$F\left(x_{j_{r_k}} + \frac{1}{2^q} v_{j_{r_k}}\right) \not\preceq C_{j_{r_k}} + Q_{j_{r_k}+1}\delta_{j_{r_k}} \frac{1}{2^q} J_F(x_{j_{r_k}}) v_{j_{r_k}}, \quad \forall\, k \ge K.$$

Hence, for each $k \ge K$, there exists $i(k) \in \{1,\cdots,m\}$ such that

$$\begin{aligned} f_{i(k)}\left(x_{j_{r_k}} + \frac{1}{2^q} v_{j_{r_k}}\right) &> C_{j_{r_k}}^{i(k)} + Q_{j_{r_k}+1}\delta_{j_{r_k}} \frac{1}{2^q} \nabla f_{i(k)}(x_{j_{r_k}})^\top v_{j_{r_k}} \\ &\ge f_{i(k)}(x_{j_{r_k}}) + \delta_{\max} \frac{1}{2^q} \nabla f_{i(k)}(x_{j_{r_k}})^\top v_{j_{r_k}}, \end{aligned} \tag{4.7}$$

where the second inequality holds due to Proposition 3.2 and that $Q_{j_{r_k}+1}\delta_{j_{r_k}} \le \delta_{\max}$. By considering the subsequence if necessary, we may assume that $i(k) = \tilde{i} \in \{1,\cdots,m\}$ for all $k \ge K$. Then, taking $k \to \infty$ in (4.7) and noting (4.5), one obtains

$$f_{\tilde{i}}\left(x^* + \frac{1}{2^q} v^*\right) \ge f_{\tilde{i}}(x^*) + \delta_{\max} \frac{1}{2^q} \nabla f_{\tilde{i}}(x^*)^\top v^*. \tag{4.8}$$

As the positive integer q is arbitrary, it follows from (4.8) and the continuous differentiability of $f_{\tilde{i}}$ that $\nabla f_{\tilde{i}}(x^*)^\top v^* \ge 0$. Thus $\max_{i\in\{1,\cdots,m\}} \langle \nabla f_i(x^*), v^*\rangle \ge 0$. This, together with (4.6), implies that $\theta(x^*) = 0$, so x^* is a Pareto stationary point of (1.1) by Lemma 3.1. The proof is complete. □

In the remaining part of this section, we analyze the convergence rate of Algorithm 3.1. For this purpose, we consider the following assumption.

(A2) For each $i \in \{1,\cdots,m\}$, ∇f_i is Lipschitz continuous on C with Lipschitz constant L_i, that is, $\|\nabla f_i(x) - \nabla f_i(y)\| \le L_i\|x-y\|$ for all $x, y \in C$.

In the following result, we show the existence of a uniform lower bound on the stepsize $\{\alpha_k\}$, which will be used later in the convergence analysis.

Lemma 4.1. *Assume that (A2) holds. Then, the stepsize* $\{\alpha_k\}$ *satisfies* $\alpha_k \geq \alpha_{\min} := \left\{1, \frac{1-\delta_{\max}}{2\beta L_{\max}}\right\}$ *for all* $k \in \mathbb{N}^*$, *where* $L_{\max} := \max\{L_1, \cdots, L_m\}$.

Proof. Take any $k \in \mathbb{N}^*$. We consider the case that $\alpha_k < 1$. By the definition of α_k, we know that $2\alpha_k$ does not satisfy the line search condition (3.4) of Algorithm 3.1. Hence, there exists an index $i \in \{1, \cdots, m\}$ such that

$$f_i(x_k + 2\alpha_k v_k) > C_k^i + Q_{k+1}\delta_k 2\alpha_k \nabla f_i(x_k)^\top v_k \geq f_i(x_k) + \delta_{\max} 2\alpha_k \nabla f_i(x_k)^\top v_k. \tag{4.9}$$

On the other hand, by the assumption (A2), we have

$$\begin{aligned} f_i(x_k + 2\alpha_k v_k) &\leq f_i(x_k) + 2\alpha_k \nabla f_i(x_k)^\top v_k + \frac{L_i}{2}\|2\alpha_k v_k\|^2 \\ &\leq f_i(x_k) + 2\alpha_k \nabla f_i(x_k)^\top v_k + 2\alpha_k^2 L_{\max}\|v_k\|^2. \end{aligned} \tag{4.10}$$

Combining (4.9) and (4.10), one obtains

$$\alpha_k \geq \frac{(\delta_{\max} - 1)\nabla f_i(x_k)^\top v_k}{L_{\max}\|v_k\|^2}. \tag{4.11}$$

Moreover, it follows from $\varphi_{x_k}(v_k) \leq 0$ that $\nabla f_i(x_k)^\top v_k \leq -\frac{\|v_k\|^2}{2\beta}$. Applying this to (4.11) establishes that $\alpha_k \geq \frac{1-\delta_{\max}}{2\beta L_{\max}}$. Then, the desired result follows by noting that α_k is never larger than 1. □

We conclude this section with the following theorem regarding the convergence rate of Algorithm 3.1.

Theorem 4.2. *Assume that (A2) holds and that there exists some nonempty set* $J \subseteq \{1, \cdots, m\}$ *such that each* f_i *with* $i \in J$ *has a lower bound* $f_i^{\min}$ *on* $\mathbb{R}^n$. *Denote* $F^{\min} := \min_{i \in J} f_i^{\min}$ *and* $F^{\max} := \max_{i \in \{1,\cdots,m\}} f_i(x_0)$. *Let* $\{x_k\}$ *and* $\{v_k\}$ *be the sequences generated by Algorithm 3.1. Then* $\lim_{k\to\infty}\|v_k\| = 0$ *and*

$$\min_{0 \leq l \leq k-1}\|v_l\| \leq \sqrt{\frac{F^{\max} - F^{\min}}{M}}\frac{1}{\sqrt{k}},$$

where $M = \frac{\delta_{\min}\alpha_{\min}}{2\beta}$ *with* $\alpha_{\min}$ *being given in Lemma 4.1.*

Proof. Let $i \in J$. For each $l \in \mathbb{N}^*$, noting (3.6) and $J_F(x_l)v_l \prec 0$, we can obtain from line search (3.4) and Lemma 4.1 that

$$\delta_{\min}\alpha_{\min}|\nabla f_i(x_l)^\top v_l| \leq \delta_l \alpha_l |\nabla f_i(x_l)^\top v_l| \leq C_l^i - C_{l+1}^i.$$

Summing up the above inequality from $l = 0$ until $k-1$ and noting Proposition 3.2 give

$$\delta_{\min}\alpha_{\min}\sum_{l=0}^{k-1}|\nabla f_i(x_l)^\top v_l| \leq \sum_{l=0}^{k-1}\left(C_l^i - C_{l+1}^i\right) = C_0^i - C_k^i \leq F^{\max} - F^{\min}. \tag{4.12}$$

Letting $k \to \infty$ in (4.12) yields

$$\sum_{l=0}^{\infty}|\nabla f_i(x_l)^\top v_l| < \infty. \tag{4.13}$$

On the other hand, for any $l \in \mathbb{N}^*$, it follows from $\varphi_{x_l}(v_l) \leq 0$ that

$$\nabla f_i(x_l)^\top v_l \leq -\frac{\|v_l\|^2}{2\beta}. \tag{4.14}$$

Applying (4.14) to (4.13) yields $\sum_{l=0}^{\infty} \|v_l\|^2 < \infty$, which implies that $\lim_{k\to\infty} \|v_k\| = 0$. Moreover, applying (4.14) to (4.12), one can obtain that

$$k \min_{0\leq l\leq k-1} \|v_l\|^2 \leq \sum_{l=0}^{k-1} \|v_l\|^2 \leq 2\beta \sum_{l=0}^{k-1} |\nabla f_i(x_l)^\top v_l| \leq \frac{2\beta(F^{\max} - F^{\min})}{\delta_{\min}\alpha_{\min}} = \frac{F^{\max} - F^{\min}}{M}.$$

Then the convergence rate is followed. The proof is complete.

□

5. Convergence Analysis: The convex Case

In this section, we investigate the convergence properties of Algorithm 3.1 with the convexity assumption on multiobjective function F. We prove that the sequence $\{x_k\}$ generated by Algorithm 3.1 converges to a weak Pareto optimal point of problem (1.1). Let us start with the following concept. Denote $\Lambda^m := \{\lambda \in \mathbb{R}^m_+ : \sum_{i=1}^m \lambda_i = 1\}$. For a nonempty closed and convex set Ω, let P_Ω denote the projection operator onto Ω.

Definition 5.1. Let $x \in C$. A direction $v \in C - x$ is said to be scalarization compatible (or simply s-compatible) at x if there exists $\lambda \in \Lambda^m$ such that

$$v = P_{C-x}\left(-\beta J_F(x)^\top \lambda\right).$$

The concept of scalarization compatible direction was used in [18] to study the steepest descent method for unconstrained vector optimization, and then was extended in [15] to study the projected gradient method for constrained vector optimization. In this paper, we apply this concept to the case of multiobjective optimization.

Observe that the exact search direction $v(x_k)$ is s-compatible at x_k. Indeed, from the first order optimality condition for $\min_{v\in C-x_k} \varphi_{x_k}(v)$, it can be concluded that there exists $u_k \in \partial\varphi_{x_k}(v(x_k))$ such that

$$\langle u_k, v - v(x_k)\rangle \geq 0, \quad \forall v \in C - x_k. \tag{5.1}$$

Then, by the expression of φ_{x_k} and the formula for the subdifferential of the maximum of convex functions (see, e.g., [20]), we see that there exist $\emptyset \neq J_k \subseteq \{1, \cdots, m\}$ and $\lambda_j^k > 0$ with $j \in J_k$ such that

$$\sum_{j\in J_k} \lambda_j^k = 1, \quad \langle \nabla f_j(x_k), v_k\rangle = \max_{1\leq i\leq m} \langle \nabla f_i(x_k), v_k\rangle, \quad \forall j \in J_k$$

and $u_k = v(x_k) + \beta \sum_{j\in J_k} \lambda_j^k \nabla f_j(x_k)$, which together with (5.1) yields

$$\left\langle v(x_k) + \beta \sum_{j\in J_k} \lambda_j^k \nabla f_j(x_k), v - v(x_k) \right\rangle \geq 0, \quad \forall v \in C - x_k,$$

which is equivalent to

$$v(x_k) = P_{C-x_k}\left(-\beta \sum_{j\in J_k} \lambda_j^k \nabla f_j(x_k)\right).$$

The following result is necessary for obtaining the quasi-Fejér convergence of the sequence generated by Algorithm 3.1.

Lemma 5.1. *Assume that F is convex. Let $\{x_k\}$ be the sequence generated by Algorithm 3.1, where, for all $k \in \mathbb{N}^*$, v_k is a s-compatible direction at x_k given by $v_k = P_{C-x_k}(-\beta J_F(x)^\top \lambda^k)$ with $\lambda^k = \left(\lambda_1^k, \cdots, \lambda_m^k\right)^\top \in \Lambda^m$. Then, for all $x \in C$ and $k \in \mathbb{N}^*$,*

$$\|x_{k+1} - x\|^2 \le \|x_k - x\|^2 + 2\beta\alpha_k \sum_{i=1}^m \lambda_i^k \left(f_i(x) - f_i(x_k)\right) + \frac{2\beta}{\delta_{\min}} \sum_{i=1}^m \left(C_k^i - C_{k+1}^i\right).$$

Proof. Let $x \in C$ and $k \in \mathbb{N}^*$. By (3.5), we have

$$\|x_{k+1} - x\|^2 = \|x_k - x\|^2 - 2\alpha_k \langle v_k, x - x_k\rangle + \alpha_k^2 \|v_k\|^2. \tag{5.2}$$

Since v_k is s-compatible at x_k, one has from the obtuse angle property of projections that

$$\left\langle -\beta \sum_{i=1}^m \lambda_i^k \nabla f_i(x_k) - v_k, v - v_k \right\rangle \le 0, \quad \forall v \in C - x_k.$$

Taking $v = x - x_k$ in the above inequality, we obtain that

$$-\langle v_k, x - x_k\rangle \le \beta \sum_{i=1}^m \lambda_i^k \langle \nabla f_i(x_k), x - x_k\rangle - \beta \sum_{i=1}^m \lambda_i^k \langle \nabla f_i(x_k), v_k\rangle - \|v_k\|^2. \tag{5.3}$$

Since F is convex, it follows that

$$\langle \nabla f_i(x_k), x - x_k\rangle \le f_i(x) - f_i(x_k), \quad \text{for each } i \in \{1, \cdots, m\}. \tag{5.4}$$

Moreover, by using line search (3.4) and noting the definition of C_{k+1} in (3.6), one has that, for each $i \in \{1, \cdots, m\}$,

$$-\langle \nabla f_i(x_k), v_k\rangle \le \frac{C_k^i - C_{k+1}^i}{\delta_k \alpha_k}. \tag{5.5}$$

Applying (5.4) and (5.5) to (5.3), and multiplying both sides by $2\alpha_k$ yield

$$\begin{aligned} -2\alpha_k \langle v_k, x - x_k\rangle &\le 2\beta\alpha_k \sum_{i=1}^m \lambda_i^k (f_i(x) - f_i(x_k)) + \frac{2\beta}{\delta_k} \sum_{i=1}^m \lambda_i^k \left(C_k^i - C_{k+1}^i\right) - 2\alpha_k \|v_k\|^2 \\ &\le 2\beta\alpha_k \sum_{i=1}^m \lambda_i^k (f_i(x) - f_i(x_k)) + \frac{2\beta}{\delta_{\min}} \sum_{i=1}^m \left(C_k^i - C_{k+1}^i\right) - 2\alpha_k \|v_k\|^2, \end{aligned}$$

where the second inequality holds due to the fact that $C_{k+1} \preceq C_k$ and that $\{\lambda_i^k\} \subseteq [0,1]$, $\delta_k \ge \delta_{\min}$. Combining this and (5.2) and noting that $\alpha_k \in (0,1]$, we obtain the desired inequality immediately. □

Now, we establish the convergence of Algorithm 3.1 to a weak Pareto optimal point of the problem (1.1).

Theorem 5.1. *Assume that F is convex and that (A1) holds. Let x_k be the sequence generated by Algorithm 3.1, where, for all $k \in \mathbb{N}^*$, v_k is s-compatible at x_k. Then, $\{x_k\}$ converges to a weak Pareto optimal point of problem (1.1).*

Proof. First, we show that $\{x_k\}$ is quasi-Fejér convergent to T. Take any $x \in T$. Then, by using Lemma 5.1 and taking into account that $F(x_k) \preceq C_k$ for all k (see Proposition 3.2), $\{\lambda_i^k\} \subseteq [0,1]$, and $\alpha_k \in (0,1]$, one sees that, for all $k \in \mathbb{N}^*$,

$$\begin{aligned}\|x_{k+1}-x\|^2 &\leq \|x_k-x\|^2 + 2\beta\alpha_k \sum_{i=1}^m \lambda_i^k \left(C_k^i - f_i(x_k)\right) + \frac{2\beta}{\delta_{\min}} \sum_{i=1}^m \left(C_k^i - C_{k+1}^i\right) \\ &\leq \|x_k-x\|^2 + 2\beta \sum_{i=1}^m \left(C_k^i - f_i(x_k)\right) + \frac{2\beta}{\delta_{\min}} \sum_{i=1}^m \left(C_k^i - C_{k+1}^i\right).\end{aligned}$$

For each k, let

$$\varepsilon_k := 2\beta \sum_{i=1}^m \left(C_k^i - f_i(x_k)\right) + \frac{2\beta}{\delta_{\min}} \sum_{i=1}^m \left(C_k^i - C_{k+1}^i\right).$$

The nonincreasing property of $\{C_k\}$ implies that $\{\varepsilon_k\} \subseteq \mathbb{R}_+$. We now prove that $\sum_{k=0}^{\infty} \varepsilon_k < +\infty$. By (3.7) and (3.8), we can obtain that, for each $i \in \{1,\cdots,m\}$ and $k \in \mathbb{N}$,

$$\begin{aligned}C_k^i - f_i(x_k) &= C_{k-1}^i - f_i(x_k) + \frac{f_i(x_k) - C_{k-1}^i}{Q_k} = \frac{Q_k - 1}{Q_k}\left(C_{k-1}^i - f_i(x_k)\right) \\ &= (Q_k - 1)\left(C_{k-1}^i - C_k^i\right) \leq \frac{\eta_{\max}}{1-\eta_{\max}}\left(C_{k-1}^i - C_k^i\right).\end{aligned} \tag{5.6}$$

Fixing any $N \in \mathbb{N}$ and noting $C_0 = F(x_0)$, we have that

$$\begin{aligned}\sum_{k=0}^N \varepsilon_k &= \sum_{k=1}^N \left(2\beta \sum_{i=1}^m \left(C_k^i - f_i(x_k)\right)\right) + \sum_{k=0}^N \left(\frac{2\beta}{\delta_{\min}} \sum_{i=1}^m \left(C_k^i - C_{k+1}^i\right)\right) \\ &= 2\beta \sum_{i=1}^m \left(\sum_{k=1}^N \left(C_k^i - f_i(x_k)\right)\right) + \frac{2\beta}{\delta_{\min}} \sum_{i=1}^m \left(\sum_{k=0}^N \left(C_k^i - C_{k+1}^i\right)\right) \\ &\leq 2\beta \sum_{i=1}^m \left(\sum_{k=1}^N \frac{\eta_{\max}}{1-\eta_{\max}}\left(C_{k-1}^i - C_k^i\right)\right) + \frac{2\beta}{\delta_{\min}} \sum_{i=1}^m \left(C_0^i - C_{N+1}^i\right) \\ &= \frac{2\beta\eta_{\max}}{1-\eta_{\max}} \sum_{i=1}^m \left(C_0^i - C_N^i\right) + \frac{2\beta}{\delta_{\min}} \sum_{i=1}^m \left(C_0^i - C_{N+1}^i\right) \\ &\leq \left(\frac{2\beta\eta_{\max}}{1-\eta_{\max}} + \frac{2\beta}{\delta_{\min}}\right) \sum_{i=1}^m \left(C_0^i - f_i(x)\right),\end{aligned} \tag{5.7}$$

where the first inequality is due to (5.6) and the second inequality holds because $x \in T$. Letting $N \to \infty$ in (5.7), one has that $\sum_{k=0}^{\infty} \varepsilon_k < +\infty$. Thus $\{x_k\}$ is quasi-Fejér convergent to T as $x \in T$ is arbitrary. Then, it follows from Proposition 2.1 that $\{x_k\}$ is bounded, so it has an accumulation point, denoted by x^*, which is a weak Pareto optimal point of problem (1.1) thanks to Theorem 4.1 and the convexity of F.

Finally, we show that $x^* \in T$. By this and Proposition 2.1, it can be concluded that $\lim_{k\to\infty} x_k = x^*$ as we wanted. Take any $k_0 \in \mathbb{N}$. Let $\{x_{j_k}\}$ be a subsequence of $\{x_k\}$ such that $\lim_{k\to\infty} x_{j_k} = x^*$. By Proposition 3.2 and the nonincreasing property of $\{C_k\}$, we have that $F(x_{j_k}) \preceq C_{j_k} \preceq C_k \preceq C_{k_0}$ for all $k \geq k_0$. Since F is continuous, we then conclude that $F(x^*) = \lim_{k\to\infty} F(x_{j_k}) \preceq C_{k_0}$. Thus $x^* \in T$ as k_0 is arbitrary. The proof is complete. □

6. Conclusion

In this work, we made further studies on the projected gradient method for constrained multiobjective optimization problems. The proposed method is constructed based on a new nonmonotone line search rule proposed here, and in each iteration, the approximate direction is adopted instead solving the subproblem exactly. In the case that the problem is nonconvex, we established the Pareto stationarity of the accumulation points of the sequence generated by this method, and demonstrated that the algorithm has a convergence rate of $O(\frac{1}{\sqrt{k}})$. When the multiobjective function is convex, we proved the convergence of the generated sequence to a weak Pareto optimal point of the problem. In future research, we are interested in further investigating novel nonmonotone techniques as well as the accelerated methods.

References

[1] Y. I. Alber, A. N. Iusem, M. V. Solodov, On the projected subgradient method for nonsmooth convex optimization in a Hilbert space, Math. Program. 81 (1998), 23-35.

[2] U. Bagchi, Simultaneous minimization of mean and variation of flow time and waiting time in single machine systems, Oper. Res. 37 (1989), 118-125.

[3] H. H. Bauschke, J. M. Borwein, On projection algorithms for solving convex feasibility problems, SIAM Rev. 38 (1996), 367-426.

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

[5] G. C. Bento, J. X. Cruz Neto, G. López, A. Soubeyran, J. C. O. Souza, The proximal point method for locally Lipschitz functions in multiobjective optimization with application to the compromise problem, SIAM J. Optim. 28 (2018), 1104-1120.

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

[7] A. S. Brito, J. X. Cruz Neto, P. S. M. Santos, S. S. Souza, A relaxed projection method for solving multiobjective optimization problems, European J. Oper. Res. 256 (2017), 17-23.

[8] R. Burachik, L. M. Graña Drummond, A. N. Iusem, B. F. Svaiter, Full convergence of the steepest descent method with inexact line searches, Optimization 32 (1995), 137-146.

[9] H. Eschenauer, J. Koski, A. Osyczka, Multicriteria Design Optimization, Springer, Berlin, 1990.

[10] N. S. Fazzio, M. L. Schuverdt, Convergence analysis of a nonmonotone projected gradient method for multiobjective optimization problems, Optim. Lett. 13 (2019), 1365-1379.

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

[12] J. Fliege, B. F. Svaiter, Steepest descent methods for multicriteria optimization, Math. Meth. Oper. Res. 51 (2000), 479-494.

[13] J. Fliege, A. I. F. Vaz, L. N. Vicente, Complexity of gradient descent for multiobjective optimization, Optim. Met. Softw. 34 (2019), 949-959.

[14] Ellen H. Fukuda, L. M. Graña Drummond, On the convergence of the projected gradient method for vector optimization, Optimization. 60 (2011), 1009-1021.

[15] Ellen H. Fukuda, L. M. Graña Drummond, Inexact projected gradient method for vector optimization, Comput. Optim. Appl. 54 (2013), 473-493.

[16] L. M. Graña Drummond, A. N. Iusem, A projected gradient method for vector optimization problems, Comput. Optim. Appl. 28 (2004), 5-29.

[17] L. M. Graña Drummond, N. Maculan, B. F. Svaiter, On the choice of parameters for the weighting method in vector optimization, Math. Program. 111 (2008), 201-216.

[18] L. M. Graña Drummond, B. F. Svaiter, A steepest descent method for vector optimization, J. Comput. Appl. Math. 175 (2005), 395-414.

[19] L. Grippo, F. Lampariello, S. Lucidi, A nonmonotone line search technique for Newton's method, SIAM J. Numer. Anal. 23 (1986), 707-716.

[20] J.-B. Hiriart Urruty, C. Lemarèchal, Convex Analysis and Minimization Algorithms, Springer, Berlin, 1993.

[21] S. Huang, Z. Wan, X. Chen, A new nonmonotone line search technique for unconstrained optimization, Numer. Algorithms 68 (2015) 671-689.

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

[23] E. A. Papa Quiroz, S. Cruzado, An inexact scalarization proximal point method for multiobjective quasiconvex minimization, Ann. Oper. Res. 316 (2022), 1445-1470.

[24] M. Pappalardo, Multiobjective optimization: a brief overview. In: A. Chinchuluun, A. Migdalas, P. M. Pardalos, L. Pitsoulis (eds.) Pareto Optimality, Game Theory and Equilibria, pp. 517-528, Springer, 2008.

[25] S. Qu, Y. Ji, J. Jiang, Q. Zhang, Nonmonotone gradient methods for vector optimization with a portfolio optimization application, European J. Oper. Res. 263 (2017), 356-366.

[26] R. T. Rockafellar, Convex Analysis, Priceton University Press, Princeton, 1970.

[27] G. T. Ross, R. M. Soland, A multicriteria approach to the location of public facilities, European J. Oper. Res. 4 (1980), 307-321.

[28] M. Tapia, C. Coello, Applications of multi-objective evolutionary algorithms in economics and finance: A survey, in Proceedings of the IEEE Congress on Evolutionary Computation, IEEE Press, Piscataway, NJ, (2007), 532-539.

[29] H. Zhang, W. W. Hager, A nonmonotone line search technique and its application to unconstrained optimization, SIAM J. Optim. 14 (2004), 1043-1056.

[30] X. Zhao, M. A. Köbis, Y. Yao, J. C. Yao, A projected subgradient method for nondifferentiable quasiconvex multiobjective optimization problems, J. Optim. Theory Appl. 190 (2021), 82-107.

[31] X. Zhao, K. F. Ng, C. Li, J. C. Yao, Linear regularity and linear convergence of projection-based methods for solving convex feasibility problems, Appl. Math. Optim. 78 (2018), 613-641.

[32] X. Zhao, J. C. Yao, Linear convergence of a nonmonotone projected gradient method for multiobjective optimization, J. Global Optim. 82 (2022), 577-594.

[33] X. Zhao, Y. Yao, Modified extragradient algorithms for solving monotone variational inequalities and fixed point problems, Optimization. 69 (2020), 1987-2002.