J. Nonlinear Var. Anal. 6 (2022), No. 6, pp. 693-706
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.6.2022.6.07

A NONMONOTONE GRADIENT METHOD FOR CONSTRAINED MULTIOBJECTIVE OPTIMIZATION PROBLEMS

XIAOPENG ZHAO[1], JEN-CHIH YAO[2], YONGHONG YAO[1,*]

[1]*School of Mathematical Sciences, Tiangong University, Tianjin 300387, China*
[2]*Research Center for Interneural Computing,*
China Medical University Hospital, China Medical University, Taichung 40447, Taiwan

Abstract. In this paper, we consider a nonmonotone gradient method for smooth constrained multiobjective optimization problems. Under mild assumptions, we demonstrate the Pareto stationarity of the accumulation point of the sequence generated by this method, while the convergence of the full sequence to a weak Pareto optimal solution of the problem is proven when the function is convex. Further, by imposing some assumptions on the gradients of the objective functions and the search directions, the linear convergence of the function value sequence to the optimal value is provided. The initial point in the convergence results established here can be any one in the constraint set.

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

1. Introduction

Multiobjective optimization refers to the process of optimizing several objective functions simultaneously, which can be formulated in the following form:

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

where $C \subseteq \mathbb{R}^n$ is a constraint set and $F : \mathbb{R}^n \to \mathbb{R}^m$ is a vector-valued function. This kind of problem can be found in various areas, such as engineering, economy, finance, management science, radiotherapy, and so on [1, 2, 3, 4, 5], and has been widely studied by many researchers [6, 7, 8].

In view of its extensive applications, the research on the numerical algorithms for solving multiobjective optimization problems has received a lot of attention and many iterative methods have been proposed, including projected gradient method [9, 10, 11], steepest descent method [12, 13], proximal point method [14, 15, 16], conjugate gradient method [17], Newton method [18, 19], trust-region method [20, 21] and so on.

Among these methods for solving multiobjective optimization, scalarization methods [15, 22, 23] and descent methods [9, 13, 19, 24] are mainly two different kinds of approaches. Based on

*Corresponding author.
E-mail addresses: zhaoxiaopeng.2007@163.com (X. Zhao), yaojc@mail.cmu.edu.tw (J.C. Yao), yyhtgu@hotmail.com (Y. Yao).

Received March 22, 2022; Accepted July 30, 2022.

the scalarization technique, scalarization methods compute the Pareto or weak Pareto optimal solutions by choosing some parameters in advance and reformulating the original vector-valued problems into the parameterized scalar-valued ones. The weighting method is a widely used scalarization technique, which minimizes a linear combination of the objectives with the vector of "weights". However, as pointed out in [18, 25], when using this method for some problems, most choices of the parameters may give rise to unbounded (and thus unsolvable) scalar problems. Usually, descent methods do not require any priori parameter information. In descent methods, Armijo line search rule is a frequently used search strategy for calculating the iteration stepsize t_k (see, for example, [9, 12, 19, 24]): given $\gamma \in (0,1)$, the stepsize t_k is defined as

$$t_k = \max\left\{2^{-j} : j \in \mathbb{N}, F(x_k + 2^{-j}v_k) \preceq F(x_k) + \gamma 2^{-j} JF(x_k)v_k\right\}, \tag{1.2}$$

where v_k is the search direction and $JF(x_k)$ is the Jacobian matrix of F at x_k. Recently, Fliege et al. [24] studied the gradient descent method for unconstrained multiobjective optimization problems with each objective function being convex and its gradient being Lipschitz continuous, in which another line search rule was also considered: given $\gamma \in (0,1)$, the stepsize t_k satisfies

$$t_k = \max\left\{2^{-j} : j \in \mathbb{N}, F(x_k + 2^{-j}v_k) \preceq F(x_k) + 2^{-j} JF(x_k)v_k + 2^{-j}\frac{\gamma}{2}\|v_k\|^2 e\right\}, \tag{1.3}$$

where e is the vector of ones in $\mathbb{R}^m$.

Under the above line search rules, the sequence of the objective values of the vectorial function is strictly monotone decreasing, and this monotonicity plays an important role in the convergence analysis of the algorithm. But as stated by Grippo et al [26] for the scalar optimization, the enforcing monotonicity of the function values can considerably slow the rate of convergence in the intermediate stages of the minimization process.

To improve this situation, some nonmonotone line search techniques have been proposed in scalar optimization [26, 27, 28, 29], which have been verified numerically that can increase the possibility of finding the optimal solution and improve the convergence speed of the algorithms. Note that the nonmonotone line search technique introduced by Zhang and Hager [29], which requires that an average of the successive function values decreases, has been shown to be more efficient than either the monontone or the traditional nonmonotone schemes in terms of taking fewer function and gradient evaluations on average.

Recently, nonmonotone line search techniques have also been applied to multiobjective optimization [11, 30, 31, 32]. Particularly, Fazzio and Schuverdt [33] extended the nonmonotone line search rule in [29] for scalar optimization to the multiobjective optimization and changed the Armijo search rule (1.2) as follows: given $\gamma \in (0,1)$, $C_0 = F(x_0)$, $Q_0 = 1$, $0 \le \eta_{\min} \le \eta_{\max} \le 1$, $\eta_k \in [\eta_{\min}, \eta_{\max}]$, the stepsize t_k satisfies

$$t_k = \max\left\{2^{-j} : j \in \mathbb{N}, F(x_k + 2^{-j}v_k) \preceq C_k + \gamma 2^{-j} JF(x_k)v_k\right\}, \tag{1.4}$$

and update

$$C_{k+1} = \frac{\eta_k Q_k C_k + F(x_{k+1})}{Q_{k+1}} \quad \text{and} \quad Q_{k+1} = \eta_k Q_k + 1. \tag{1.5}$$

Based on the line search rule (1.4), the authors in [33] considered a nonmonotone projected gradient method for the constrained multiobjective optimization problem (1.1) when C is a closed and convex set and F is a continuously differentiable vectorial function. They showed the stationarity of the accumulation points of the sequences generated by the proposed algorithm, and then established the convergence to weak Pareto points when the function F is convex. Very recently in [11], by adopting the nonmonotone line search rule (1.4), we considered a projected gradient method with exogenously given square summable sequence in the computation of the search direction for the same multiobjective optimization problem when F is convex. We proved the convergence of the full sequence generated by the algorithm to a weak Pareto optimal point. Furthermore, under some appropriate Lipschitz continuity assumption of the gradients of objective functions, we established the linear convergence result of the proposed method.

It should be noted that although some achievements have been made in the research of nonmonotone algorithms in multiobjective optimization, the research about this topic is relatively recent and insufficient. Therefore, this problem still deserves our continuous and in-depth study.

In this work, we consider the multiobjective optimization problem (1.1) in the case when C is a closed and convex set, and each objective function f_i is continuously differentiable with the gradient being Lipschitz continuous. By applying the idea in nonmonotone line search technique (1.4) and (1.5) to the sufficient decrease condition (1.3), we propose a nonmonotone algorithm to solve the multiobjective optimization problem that considered herein. Under suitable assumptions, we show that any accumulation point of the sequence generated by the proposed method is Pareto stationary, and the whole sequence converges to a weak Pareto optimal point when the objective function is convex. Further, under some additional assumptions on the gradients of the objective functions, the linear convergence result for this method is also established.

The remainder of this paper is organized as follows. In Section 2, we present the notations and preliminary results that to be used in the present paper. In Section 3, we propose the nonmonotone algorithm, for which the well-definedness is illustrated and some important properties of the algorithm that will be used in the sequal are provided. Section 4 contains the main results, where we establish the convergence analysis of the algorithm in different cases when the objective function is nonconvex, respectively, convex. Finally, some concluding remarks are given in Section 5.

2. NOTATIONS AND PRELIMINARY RESULTS

The notations used in the present paper are standard (cf. [34, 35, 36]). We denote by $\langle\cdot,\cdot\rangle$ the inner product of $\mathbb{R}^n$ and by $\|\cdot\|$ its corresponding norm. The closed unit ball of $\mathbb{R}^n$ is denoted by $\mathbf{B}$ and the transpose sign is denoted by ^{T}. Let $\mathbb{R}^m_+$ and $\mathbb{R}^m_{++}$ denote the nonnegative orthant and positive orthant of $\mathbb{R}^m$, respectively. For two vectors $x,y\in\mathbb{R}^m$, we write $x\preceq y$ (resp., $x\prec y$) if and only if $y-x\in\mathbb{R}^m_+$ (resp., $y-x\in\mathbb{R}^m_{++}$).

In the present paper, we consider the multiobjective optimization problem (1.1) with $C\subseteq\mathbb{R}^n$ being a nonempty closed and convex set and $F:\mathbb{R}^n\to\mathbb{R}^m$ being a continuously differentiable vector-valued function on an open superset of C denoted by

$$F:=(f_1,\cdots,f_m)^T, \tag{2.1}$$

where each f_i, $i=1,\cdots,m$, is a real-valued function and its gradient is Lipschitz continuous with constant $L_i>0$ on C, i.e.,

$$\|\nabla f_i(x)-\nabla f_i(y)\|\le L_i\|x-y\|,\qquad \text{for all } x,y\in C.$$

We denote $L_{\max} := \max\{L_1, \cdots, L_m\}$. The multiobjective function F is said to be continuously differentiable or convex if each component function f_i with $i = 1, \cdots, m$ is continuously differentiable or convex.

We use $J_F(x)$ to denote the Jacobian matrix of F at x, that is,

$$J_F(x) = (\nabla f_1(x), \cdots, \nabla f_m(x))^T.$$

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

(a) a Pareto optimal point (or Pareto efficient solution) of F on C 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 (or weak Pareto efficient solution) of F on C if there does not exist $x \in C$ such that $F(x) \prec F(x^*)$,

(c) a Pareto stationary point (or a Pareto critical point) of F on C if

$$J_F(x^*)(C - x^*) \cap (-\mathbb{R}^m_{++}) = \emptyset.$$

It is well-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, but the converse is not always true. However, if F is convex, then Pareto stationarity implies weak Pareto optimality.

We recall the following quasi-Fejér convergence theorem, which has been widely used to analyze the gradient and subgradient methods; see, for example, [9, 22, 33, 37].

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

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

Proposition 2.1. ([38, 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 $\lim_{k\to\infty} u_k = \bar{u}$.*

The algorithm's quality of convergence will be discussed in terms of linear convergence. Recall that a sequence $\{u_k\} \subseteq \mathbb{R}^n$ is said to converge linearly to its limit u (with rate θ) if $\theta \in [0, 1)$ and there is some $\alpha \ge 0$ such that $\|u_k - u\| \le \alpha\theta^k$ for all k.

3. The Algorithm and Properties

In this section, by combining Zhang and Hager's nonmonotone line search technique [29] with the descent condition (1.3) that was recently considered in [24], we propose the nonmonotone gradient algorithm for multiobjective optimization problem (2.1). Then, we present some basic properties of the algorithm that will be used for convergence analysis in the next section.

The nonmonotone gradient algorithm considered herein is formally stated as follows.

Algorithm 3.1. Step 1 Choose parameters $\gamma \in (0, 1)$ and $0 \le \eta_{\min} \le \eta_{\max} \le 1$. 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 search direction v_k:

$$v_k := \operatorname{argmin}_{v \in C - x_k} \varphi_k(v), \tag{3.1}$$

where $\varphi_k(v) = \max_{1 \le i \le m} \langle \nabla f_i(x_k), v \rangle + \frac{\|v\|^2}{2}$.

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

Step 4 Compute a stepsize $t_k \in (0,1]$ as the maximum of

$$T_k := \left\{ t = \frac{1}{2^j} | j \in \mathbb{N}, F(x_k + tv_k) \preceq C_k + tJF(x_k)v_k + \frac{\gamma t}{2} \|v_k\|^2 e \right\} \tag{3.2}$$

where $e = (1, \cdots, 1)^T \in \mathbb{R}^m$. Then, set

$$x_{k+1} = x_k + t_k v_k. \tag{3.3}$$

Step 5 Choose $\eta_k \in [\eta_{\min}, \eta_{\max}]$ and define

$$Q_{k+1} = \eta_k Q_k + 1, \quad C_{k+1} = \frac{\eta_k Q_k C_k + F(x_{k+1})}{Q_{k+1}}. \tag{3.4}$$

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

Observe that (one can also see [33]) for each k, C_{k+1} can be equivalently rewritten as

$$C_{k+1} = \frac{(\eta_k Q_k + 1)C_k + F(x_{k+1}) - C_k}{Q_{k+1}} = C_k + \frac{F(x_{k+1}) - C_k}{Q_{k+1}}, \tag{3.5}$$

and so

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

Note that if $\eta_k = 0$ for each k, then $C_k = F(x_k)$ and the line search (3.2) reduces to the monotone one that recently considered in [24] for the gradient descent method. While if $\eta_k = 1$ for each k, then $C_k = \frac{1}{k+1}\sum_{i=0}^{k} F(x_i)$ is the average of all the previous function values. Thus, similar to the selection of the parameters involved in other nonmonotone linear search rules ([11, 33]), the choice of η_k controls the degree of the non-monotonicity of the line search (3.2) (the line search closely approximates the monotone one as η_k approaches 0, and the scheme becomes more nonmonotone as η_k approaches 1).

Now, we illustrate the validity of the stopping criteria in Algorithm 3.1. For this, we need to formulate some useful inequalities first. Note that the function φ_k in (3.1) is strongly convex. Then, it follows from the first order optimality condition for $\min_{v \in C - x_k} \varphi_k(v)$ that there exists $u_k \in \partial \varphi_k(v_k)$ such that

$$\langle u_k, v - v_k \rangle \geq 0, \quad \forall v \in C - x_k. \tag{3.7}$$

Then, by the expression of φ_k and the formula for the subdifferential of the maximum of convex functions (see, e.g., [39]), there exist $\emptyset \neq J_k \subseteq \{1, \cdots, m\}$ and $\lambda_j^k > 0$ with $j \in J_k$ such that

$$\sum_{j \in J_k} \lambda_j^k = 1, \quad \langle \nabla f_j(x_k), v_k \rangle = \max_{1 \leq i \leq m} \langle \nabla f_i(x_k), v_k \rangle, \quad \forall j \in J_k \tag{3.8}$$

and

$$u_k = v_k + \sum_{j \in J_k} \lambda_j^k \nabla f_j(x_k). \tag{3.9}$$

Combining (3.7) and (3.9) yields

$$\left\langle v_k + \sum_{j \in J_k} \lambda_j^k \nabla f_j(x_k), v - v_k \right\rangle \geq 0, \quad \forall v \in C - x_k. \tag{3.10}$$

Proposition 3.1. *If the Algorithm 3.1 stops at iteration k, then x_k is a Pareto stationary point, and further a weak Pareto optimal point when F is convex.*

Proof. Take $k \in \mathbb{N}$. If $\nabla f_{i_0}(x_k) = 0$ for some $i_0 \in \{1, \cdots, m\}$, then for any $x \in C$, $\langle \nabla f_{i_0}(x_k), x - x_k \rangle = 0$, and 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 when $v_k = 0$, it follows from (3.10) that

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

This implies that, for any $x \in C$, there exists $j \in J_k \subseteq \{1, \cdots, m\}$ such that $\langle \nabla f_j(x_k), x - x_k \rangle \geq 0$. Consequently, $J_F(x_k)(C - x_k) \cap (-\mathbb{R}^m_{++}) = \emptyset$, and x_k is a Pareto stationary point. The further conclusion of the proposition follows immediately from the equivalence between the Pareto stationarity and the weak Pareto optimality of x_k when the objective function F is convex. The proof is complete. □

Henceforth, we suppose that Algorithm 3.1 does not have a finite termination. The following result shows the well-definedness of Algorithm 3.1 and a useful relation between the generated sequence $\{x_k\}$ and $\{C_k\}$, where the proof of the conclusion $F(x_k) \preceq C_k$ draws on Zhang and Hager's technique for scalar optimization [29].

Lemma 3.1. *The nonmonotone line search rule (3.2) holds, and so Algorithm 3.1 is well-defined. The sequence $\{x_k\}$ generated by Algorithm 3.1 belongs to C and $F(x_k) \preceq C_k$ for each $k \in \mathbb{N}$.*

Proof. We will prove the conclusion of this lemma by induction. First, consider the case when $k = 0$. By using the Lipschitz continuity of ∇f_i with $i \in \{1, \cdots, m\}$, one can obtain that, for all $t \in (0, \frac{\gamma}{L_i}]$,

$$\begin{aligned} f_i(x_0 + tv_0) &\leq f_i(x_0) + t \nabla f_i(x_0)^T v_0 + \frac{L_i}{2} \|tv_0\|^2 \\ &\leq f_i(x_0) + t \nabla f_i(x_0)^T v_0 + \frac{\gamma t}{2} \|v_0\|^2. \end{aligned}$$

Thus we have that, for all $t \in (0, \frac{\gamma}{L_{\max}}]$,

$$F(x_0 + tv_0) \preceq F(x_0) + tJF(x_0)v_0 + \frac{\gamma t}{2} \|v_0\|^2 e. \tag{3.11}$$

Note that $C_0 = F(x_0)$. Then, (3.11) implies that the line search rule (3.2) in Algorithm 3.1 holds for $k = 0$, and so the stepsize t_0 can be computed by the fact that in the backtracking scheme the stepsize starts at one and is halved each time. And it can be verified that $x_1 = x_0 + t_0 v_0 \in C$ by noting that the initial iterate x_0 belongs to C, $v_0 \in C - x_0$ and the convexity of C.

Moreover, by the definition of the search direction and the fact that $0 \in C - x_0$, one has that

$$\varphi_0(v_0) = \max_{1 \leq i \leq m} \langle \nabla f_i(x_0), v_0 \rangle + \frac{\|v_0\|^2}{2} \leq \varphi_0(0) = 0,$$

which means that

$$t_0 JF(x_0)v_0 + \frac{\gamma t_0}{2} \|v_0\|^2 e \preceq t_0 \left(JF(x_0)v_0 + \frac{\|v_0\|^2}{2} e \right) \preceq 0, \tag{3.12}$$

where the first inequality holds because $\gamma \in (0, 1)$. This, together with the line search rule (3.2) (for $k = 0$) and the update iterate (3.3), implies that $F(x_1) \preceq C_0 = F(x_0)$.

Now, let us consider the case when $k \geq 1$. Without loss of generality, we assume that $F(x_k) \preceq C_{k-1}$ and $x_k \in C$. Similarly to obtaining (3.11), one can obtain by using the Lipschitz continuity of ∇f_i that

$$F(x_k + tv_k) \preceq F(x_k) + tJF(x_k)v_k + \frac{\gamma t}{2}\|v_k\|^2 e \quad \text{for all } t \in \left(0, \frac{\gamma}{L_{\max}}\right]. \tag{3.13}$$

For each $i = 1, \cdots, m$, define the function $D_{k,i} : \mathbb{R} \to \mathbb{R}$ by

$$D_{k,i}(t) = \frac{tC_{k-1}^i + f_i(x_k)}{t+1},$$

where C_{k-1}^i is the ith component of C_{k-1}. It is easy to compute that $D'_{k,i}(t) = \frac{C_{k-1}^i - f_i(x_k)}{(t+1)^2} \geq 0$ for all t (due to the assumption $F(x_k) \preceq C_{k-1}$), and thus

$$f_i(x_k) = D_{k,i}(0) \leq D_{k,i}(\eta_{k-1}Q_{k-1}) = \frac{\eta_{k-1}Q_{k-1}C_{k-1}^i + f_i(x_k)}{\eta_{k-1}Q_{k-1} + 1} = C_k^i,$$

which is equivalent to that $F(x_k) \preceq C_k$. Combining this with (3.13) yields that the line search rule (3.2) holds for k. Then, the stepsize t_k can be obtained and so is $x_{k+1} := x_k + t_k v_k$, and $x_{k+1} \in C$ can also be checked easily by noting that $v_k \in C - x_k$. The proof is complete. □

Remark 3.1. Similarly to (3.12) in the proof of Lemma 3.1, for all $k \in \mathbb{N}$, one can obtain $t_k JF(x_k)v_k + \frac{\gamma t_k}{2}\|v_k\|^2 e \preceq 0$ by $\varphi_k(v_k) \leq 0$, and thus $F(x_{k+1}) \preceq C_k$ due to (3.2). Applying this to (3.5), it follows that, for each k, $C_{k+1} \preceq C_k$, that is, $\{C_k\}$ is a nonincreasing sequence in $\mathbb{R}^m$.

Next, we show that the stepsize $\{t_k\}$ generated in Algorithm 3.1 has a lower bound.

Lemma 3.2. *The stepsize in Algorithm 3.1 always satisfies* $t_k \geq t_{\min} \equiv \min\left\{\frac{\gamma}{2L_{\max}}, 1\right\}$.

Proof. Fix any $k \in \mathbb{N}$. By the definition of t_k, we know that $2t_k$ does not satisfy the inequality in the nonmonotone line search rule (3.2) of Algorithm 3.1. Consequently, there exists an index $i \in \{1, \cdots, m\}$ such that

$$\begin{aligned} f_i(x_k + 2t_k v_k) &> C_k^i + 2t_k \nabla f_i(x_k)^T v_k + \frac{\gamma}{2} 2t_k \|v_k\|^2 \\ &\geq f_i(x_k) + 2t_k \nabla f_i(x_k)^T v_k + \gamma t_k \|v_k\|^2, \end{aligned} \tag{3.14}$$

where the second inequality holds thanks to Lemma 3.1. On the other hand, by the Lipschitz continuity of ∇f_i, we have

$$f_i(x_k + 2t_k v_k) \leq f_i(x_k) + 2t_k \nabla f_i(x_k)^T v_k + \frac{L_i}{2}\|2t_k v_k\|^2. \tag{3.15}$$

Combining (3.14) and (3.15), one obtains $t_k \geq \frac{\gamma}{2L_i} \geq \frac{\gamma}{2L_{\max}}$. The conclusion then follows by noting that t_k is never larger than one. □

4. Convergence Analysis

In this section, we concentrate on the convergence property of Algorithm 3.1, and we consider it when the function F is nonconvex, respectively, convex. The following technical result will be used in the convergence analysis.

Proposition 4.1. *Assume that $\eta_{\max} < 1$. Then, for all $x \in C$ and $k \in \mathbb{N}$, there exists $\{\lambda_j^k\}_{j=1}^m \subseteq [0,1]$ satisfying $\sum_{j=1}^m \lambda_j^k = 1$, and*

$$\|x_{k+1} - x\|^2 \le \|x_k - x\|^2 + 2t_k \sum_{j=1}^m \lambda_j^k \langle \nabla f_j(x_k), x - x_k \rangle + \frac{2}{1-\eta_{\max}} \sum_{j=1}^m (C_k^j - C_{k+1}^j).$$

Proof. Let $x \in C$ and $k \in \mathbb{N}$. By (3.3), we have that

$$\begin{aligned} \|x_{k+1} - x\|^2 &= \|x_k - x\|^2 - 2t_k \langle v_k, x - x_k \rangle + t_k^2 \|v_k\|^2 \\ &= \|x_k - x\|^2 - 2t_k \langle v_k, x - x_k - v_k \rangle + t_k(t_k - 2)\|v_k\|^2. \end{aligned} \tag{4.1}$$

Taking $v = x - x_k$ in (3.10) and letting $\lambda_j^k = 0$ for $j \notin J_k$, one can get that

$$2t_k \langle v_k, x - x_k - v_k \rangle \ge 2t_k \sum_{j=1}^m \lambda_j^k \langle \nabla f_j(x_k), v_k \rangle - 2t_k \sum_{j=1}^m \lambda_j^k \langle \nabla f_j(x_k), x - x_k \rangle.$$

By using $F(x_{k+1}) \preceq C_k + t_k JF(x_k) v_k + \frac{\gamma t_k}{2} \|v_k\|^2 e$ in the last inequality, and noting that $F(x_{k+1}) \preceq C_k$ and $\sum_{j=1}^m \lambda_j^k = 1$ with each $0 \le \lambda_j^k \le 1$ (see (3.8)), it follows that

$$\begin{aligned} &2t_k \langle v_k, x - x_k - v_k \rangle \\ \ge & 2 \sum_{j=1}^m \lambda_j^k \left(f_j(x_{k+1}) - C_k^j - \frac{\gamma t_k}{2} \|v_k\|^2 \right) - 2t_k \sum_{j=1}^m \lambda_j^k \langle \nabla f_j(x_k), x - x_k \rangle \\ \ge & 2 \sum_{j=1}^m (f_j(x_{k+1}) - C_k^j) - 2t_k \sum_{j=1}^m \lambda_j^k \langle \nabla f_j(x_k), x - x_k \rangle - \gamma t_k \|v_k\|^2. \end{aligned} \tag{4.2}$$

Applying (4.2) to (4.1), we obtain

$$\begin{aligned} &\|x_{k+1} - x\|^2 \\ \le & \|x_k - x\|^2 + 2 \sum_{j=1}^m (C_k^j - f_j(x_{k+1})) + 2t_k \sum_{j=1}^m \lambda_j^k \langle \nabla f_j(x_k), x - x_k \rangle + t_k(\gamma + t_k - 2)\|v_k\|^2 \\ \le & \|x_k - x\|^2 + 2 \sum_{j=1}^m (C_k^j - f_j(x_{k+1})) + 2t_k \sum_{j=1}^m \lambda_j^k \langle \nabla f_j(x_k), x - x_k \rangle, \end{aligned} \tag{4.3}$$

where the last inequality holds since both γ and t_k are less than or equal to 1. Moreover, by the definition of Q_{k+1} in (3.4), one can compute that

$$Q_{k+1} = 1 + \sum_{j=0}^k \prod_{l=0}^j \eta_{k-l} \le 1 + \sum_{j=0}^k \eta_{\max}^{j+1} \le \sum_{j=0}^\infty \eta_{\max}^j \le \frac{1}{1-\eta_{\max}}. \tag{4.4}$$

By this and (3.6), one then gets that

$$\sum_{j=1}^m (C_k^j - f_j(x_{k+1})) = \sum_{j=1}^m Q_{k+1}(C_k^j - C_{k+1}^j) \le \frac{1}{1-\eta_{\max}} \sum_{j=1}^m (C_k^j - C_{k+1}^j). \tag{4.5}$$

Thus the desired inequality can be acquired by combining (4.3) and (4.5). $\square$

4.1. **The nonconvex case.**

Theorem 4.1. *Assume that $\eta_{\max} < 1$ and that F is bounded from below. Then, every accumulation point, if any, of the sequence $\{x_k\}$ generated by Algorithm 3.1 is a Pareto stationary point.*

Proof. Let x^* be an accumulation point of $\{x_k\}$. Since $\{x_k\} \subseteq C$ (by Lemma 3.1) and C is closed, it follows that $x^* \in C$. Suppose that x^* is not a Pareto stationary point. Then, by definition, there exists $\hat{x} \in C$ such that $J_F(x^*)(\hat{x} - x^*) \in -\mathbb{R}^m_{++}$, i.e.,

$$\max_{1\le i\le m} \langle \nabla f_i(x^*), \hat{x} - x^* \rangle < 0. \tag{4.6}$$

Applying Proposition 4.1 with $x = \hat{x}$, we have that, for all k,

$$\|x_{k+1} - \hat{x}\|^2 \le \|x_k - \hat{x}\|^2 + 2t_k \max_{1\le i\le m} \langle \nabla f_i(x_k), \hat{x} - x_k \rangle + \frac{2}{1-\eta_{\max}} \sum_{j=1}^m (C_k^j - C_{k+1}^j). \tag{4.7}$$

Take $N \in \mathbb{N}$. Summing (4.7) from $k = 0, \cdots, N$, one derives

$$\begin{aligned}
&-2\sum_{k=0}^N t_k \max_{1\le i\le m} \langle \nabla f_i(x_k), \hat{x} - x_k \rangle \\
\le& \sum_{k=0}^N \left(\|x_k - \hat{x}\|^2 - \|x_{k+1} - \hat{x}\|^2\right) + \frac{2}{1-\eta_{\max}} \sum_{k=0}^N \left(\sum_{j=1}^m (C_k^j - C_{k+1}^j) \right) \\
=& \|x_0 - \hat{x}\|^2 - \|x_{N+1} - \hat{x}\|^2 + \frac{2}{1-\eta_{\max}} \sum_{j=1}^m \left(\sum_{k=0}^N (C_k^j - C_{k+1}^j) \right) \\
\le& \|x_0 - \hat{x}\|^2 + \frac{2}{1-\eta_{\max}} \sum_{j=1}^m (C_0^j - C_{N+1}^j).
\end{aligned} \tag{4.8}$$

Since F is bounded from below, we can deduce from $F(x_k) \preceq C_k$ for all k in Lemma 3.1 that C_k is also bounded from below. Then, it follows from (4.8) that

$$-2\sum_{k=0}^N t_k \max_{1\le i\le m} \langle \nabla f_i(x_k), \hat{x} - x_k \rangle < +\infty. \tag{4.9}$$

Moreover, we know from Lemma 3.2 that $t_k \ge t_{\min} > 0$ for all k. This, together with (4.9), implies that

$$\lim_{k\to\infty} \left(-\max_{1\le i\le m} \langle \nabla f_i(x_k), \hat{x} - x_k \rangle \right) = 0. \tag{4.10}$$

However, for any subsequence $\{x_{j_k}\}$ of $\{x_k\}$ such that $\lim_{k\to\infty} x_{j_k} = x^*$, one can obtain by (4.6) that

$$\lim_{k\to\infty} \left(-\max_{1\le i\le m} \langle \nabla f_i(x_{j_k}), \hat{x} - x_{j_k} \rangle \right) = -\max_{1\le i\le m} \langle \nabla f_i(x^*), \hat{x} - x^* \rangle > 0,$$

which is a contradiction to (4.10). Therefore, we conclude that x^* is a Pareto stationary point. $\square$

4.2. **The convex case.** In this subsection, we consider Algorithm 3.1 in the case when imposing the convexity assumption on F. To attain stronger convergence results, we need additional assumption.

It is known that the condition $\{x \in C : F(x) \preceq F(x_k), \forall k \in \mathbb{N}\} \neq \emptyset$ has been frequently used in the convergence analysis of many algorithms for vector optimization, see for example the projected gradient method [9, 10, 37], steepest descent method [12], proximal point method [14, 16, 40]. This assumption is related to the completeness of the image of F, namely, all nonincreasing sequences with respect to $\preceq$ in the image of F have a lower bound. And the completeness is a standard assumption for ensuring existence of efficient points for vector optimization problems [34].

While for nonmonotone algorithm, by considering C_k instead of $F(x_k)$, Fazzio and Schuverdt [33] used the following assumption to prove the convergence of the algorithm proposed therein:

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

Recently, we also used this assumption to show the convergence of another nonmonotone projected gradient method [11]. Next, by applying the assumption (A1) again, we study the convergence of Algorithm 3.1 considered in the present paper.

Theorem 4.2. *Assume that (A1) holds and $\eta_{\max} < 1$. Then, the sequence $\{x_k\}$ generated by Algorithm 3.1 converges to a weak Pareto optimal point.*

Proof. First, we show that $\{x_k\}$ is quasi-Fejér convergent to T. Take any $x \in T$. By the gradient inequality of each f_j, we have that, for all k,

$$\langle \nabla f_j(x_k), x - x_k \rangle \leq f_j(x) - f_j(x_k) \leq C_k^j - f_j(x_k).$$

From this and Proposition 4.1, one can obtain that, for all k,

$$\begin{aligned}
&\|x_{k+1} - x\|^2 \\
\leq &\|x_k - x\|^2 + 2t_k \sum_{j=1}^{m} \lambda_j^k \langle \nabla f_j(x_k), x - x_k \rangle + \frac{2}{1-\eta_{\max}} \sum_{j=1}^{m} (C_k^j - C_{k+1}^j) \\
\leq &\|x_k - x\|^2 + 2t_k \sum_{j=1}^{m} \lambda_j^k (C_k^j - f_j(x_k)) + \frac{2}{1-\eta_{\max}} \sum_{j=1}^{m} (C_k^j - C_{k+1}^j) \\
\leq &\|x_k - x\|^2 + 2 \sum_{j=1}^{m} (C_k^j - f_j(x_k)) + \frac{2}{1-\eta_{\max}} \sum_{j=1}^{m} (C_k^j - C_{k+1}^j),
\end{aligned} \tag{4.11}$$

where the last inequality is established by the conclusion $F(x_k) \preceq C_k$ in Lemma 3.1, the stepsize $t_k \leq 1$ and that $\lambda_j^k \leq 1$. For each k, let

$$\varepsilon_k := 2 \sum_{j=1}^{m} (C_k^j - f_j(x_k)) + \frac{2}{1-\eta_{\max}} \sum_{j=1}^{m} (C_k^j - C_{k+1}^j).$$

Note that $C_{k+1} \preceq C_k$ (see Remark 3.1), so we have $\varepsilon_k \geq 0$, and $\|x_{k+1} - x\|^2 \leq \|x_k - x\|^2 + \varepsilon_k$ by (4.11) for all k.

We are going to prove that $\sum_{k=0}^{\infty}\varepsilon_k<+\infty$. By (4.8) in the proof of Theorem 4.1 and noting that $x\in T$, we have that, for any $N\in\mathbb{N}$,

$$\sum_{k=0}^{N}\left(\frac{2}{1-\eta_{\max}}\sum_{j=1}^{m}(C_k^j-C_{k+1}^j)\right)=\frac{2}{1-\eta_{\max}}\sum_{j=1}^{m}(C_0^j-C_{N+1}^j)\le\frac{2}{1-\eta_{\max}}\sum_{j=1}^{m}(C_0^j-f_j(x)),$$

and so

$$\sum_{k=0}^{\infty}\left(\frac{2}{1-\eta_{\max}}\sum_{j=1}^{m}(C_k^j-C_{k+1}^j)\right)<+\infty \tag{4.12}$$

by letting N goes to $+\infty$. Moreover, by (3.5) and (4.4), one has that, for each $j\in\{1,\cdots,m\}$ and $k\in\mathbb{N}$,

$$\begin{aligned}C_k^j-f_j(x_k)&=C_{k-1}^j-f_j(x_k)+\frac{f_j(x_k)-C_{k-1}^j}{Q_k}=(C_{k-1}^j-f_j(x_k))\frac{Q_k-1}{Q_k}\\&=(C_{k-1}^j-C_k^j)(Q_k-1)\le\frac{\eta_{\max}}{1-\eta_{\max}}(C_{k-1}^j-C_k^j).\end{aligned} \tag{4.13}$$

Then, by defining $C_{-1}=C_0$, it follows from (4.13) that, for any $N\in\mathbb{N}$,

$$\begin{aligned}\sum_{k=0}^{N}\left(2\sum_{j=1}^{m}(C_k^j-f_j(x_k))\right)&\le\sum_{k=0}^{N}\left(\frac{2\eta_{\max}}{1-\eta_{\max}}\sum_{j=1}^{m}(C_{k-1}^j-C_k^j)\right)\\=\frac{2\eta_{\max}}{1-\eta_{\max}}\sum_{j=1}^{m}(C_0^j-C_N^j)&\le\frac{2\eta_{\max}}{1-\eta_{\max}}\sum_{j=1}^{m}(C_0^j-f_j(x)).\end{aligned}$$

Thus

$$\sum_{k=0}^{\infty}\left(2\sum_{j=1}^{m}(C_k^j-f_j(x_k))\right)<+\infty.$$

Combining this with (4.12) yields that $\sum_{k=0}^{\infty}\varepsilon_k<+\infty$ and consequently $\{x_k\}$ is quasi-Fejér convergent to T.

Then, by Proposition 2.1, we have that $\{x_k\}$ is bounded and so has an accumulation point, denoted by x^*, which is a weak Pareto optimal point thanks to Theorem 4.1 (note that the proof of Theorem 4.1 still holds when the assumption "F is bounded from below" is replaced by (A1)) and the convexity of F.

In the remaining proof, we show that $x^*\in T$. By this and Proposition 2.1 again, the convergence of $\{x_k\}$ to x^* is then obtained. Let $\{x_{j_k}\}$ be a subsequence of $\{x_k\}$ such that $\lim_{k\to\infty}x_{j_k}=x^*$. Take any $k_0\in\mathbb{N}$. By Lemma 3.1 and the nonincreasing property of $\{C_k\}$, one has that for all $k\ge k_0$, $F(x_{j_k})\preceq C_{j_k}\preceq C_k\preceq C_{k_0}$. Since F is continuous, one then gets that $F(x^*)=\lim_{k\to\infty}F(x_{j_k})\preceq C_{k_0}$. Thus, $x^*\in T$ as k_0 is arbitrary, which completes the proof. □

At the end of this subsection, we give a linear convergence result of Algorithm 3.1. Recently, we proved the linear convergence of the nonmonotone algorithm studied in [11] by using the following condition on the search direction and the gradients of f_i:

(A2) There exists positive constant c such that, for each $i\in\{1,\cdots,m\}$ and $k\in\mathbb{N}$,

$$\langle\nabla f_i(x_k),v_k\rangle\le -c\|\nabla f_i(x_k)\|.$$

In the present work, by employing the similar techniques that used in [11], we establish the linear convergence of Algorithm 3.1. First of all, we can obtain the following useful lemma under the assumption (A2).

Lemma 4.1. *Assume that (A2) holds. Then, there exists $\alpha > 0$ such that*

$$f_i(x_{k+1}) \leq C_k^i - \alpha \|\nabla f_i(x_k)\| \quad \text{for each } i \in \{1, \cdots, m\} \text{ and } k \in \mathbb{N}.$$

Proof. Take $i \in \{1, \cdots, m\}$ and $k \in \mathbb{N}$. By the definition of the search direction v_k and noting that $0 \in C - x_k$, we can obtain by $\varphi_k(v_k) \leq \varphi_k(0) = 0$ that $\frac{\|v_k\|^2}{2} \leq -\langle \nabla f_i(x_k), v_k \rangle$. Combining this with the nonmonotone line search rule (3.2) yields

$$f_i(x_{k+1}) \leq C_k^i + t_k(1-\gamma)\langle \nabla f_i(x_k), v_k \rangle. \tag{4.14}$$

Then, applying the lower bound of t_k in Lemma 3.2 and assumption (A2) to (4.14), one can obtain

$$f_i(x_{k+1}) \leq C_k^i - c t_{\min}(1-\gamma)\|\nabla f_i(x_k)\|,$$

where $t_{\min}$ is defined as that in Lemma 3.2. Consequently, the conclusion holds with $\alpha := c t_{\min}(1-\gamma)$. □

Below, we state the linear convergence result of Algorithm 3.1. Applying Lemma 3.1, Lemma 4.1 and Theorem 4.2, the following theorem can be obtained by using the similar techniques as that used in [11, Theorem 2] and so we omit the proof here for simplicity.

Theorem 4.3. *Assume that (A1), (A2) hold and $\eta_{\max} < 1$. Let $\{x_k\}$ be the sequence generated by Algorithm 3.1. Then, there exits $\theta \in (0,1)$ such that*

$$F(x_k) - F(x^*) \preceq \theta^k (F(x_0) - F(x^*)) \quad \text{for all } k,$$

where $x^* := \lim_{k\to\infty} x_k$ is a weak Pareto optimal point. Specifically, the convergence rate θ is given by

$$\theta = 1 - \frac{\alpha(1-\eta_{\max})}{\alpha + (\rho + \|x^*\|)(1+2L_{\max})}$$

with α as that in Lemma 4.1, the constant ρ satisfying $\{x_k\} \subseteq \rho\mathbf{B}$ and $L_{\max}$ is the maximal Lipschitz constant L_i of ∇f_i, i.e., $L_{\max} := \max\{L_1, \cdots, L_m\}$.

5. Conclusion

It has been verified numerically in the optimization literature that the nonmonotone strategies often perform better than monotone ones. In this paper, we presented a gradient method equipped with the Zhang and Hager's nonmonotone line search technique (in multiobjective case) for smooth multiobjective optimization. We established the Pareto stationarity of the accumulation points of the sequence generated by the proposed method for nonconvex multiobjective optimization problems. When the multiobjective function was convex, the convergence of the generated sequence to a weak Pareto optimal point and further the linear convergence of the function values to the optimal value were proved.

The future work includes several aspects. For example, it is of interest to investigate the linear convergence rate of the sequence generated by the nonmonotone algorithm (rather than the function value sequence), the accelerated gradient methods, and propose novel nonmonotone techniques to find good Pareto fronts for multiobjective optimization problems.

Acknowledgments
The research of the first author was supported by the National Natural Science Foundation of China (Grant Number 11801411). The research of the second author was supported by the Grant MOST (111-2115-M-039-001-MY2). The research of the third author was partially supported in part by University Innovation Team of Tianjin [Grant Number TD13-5033].

REFERENCES

[1] U. Bagchi, Simultaneous minimization of mean and variation of flow time and waiting time in single machine systems, Oper. Res. 37 (1989), 118-125.
[2] H. Eschenauer, J. Koski, A. Osyczka, Multicriteria Design Optimization, Springer, Berlin, 1990.
[3] E. Schreibmann, M. Lahanas, L. Xing, D. Baltas, Multiobjective evolutionary optimization of the number of beams, their orientations and weights for intensity-modulated radiation therapy, Phys. Med. Biol. 49 (2004), 747-770.
[4] T. Stewart, O. Bandte, H. Braun, N. Chakraborti, M. Ehrgott, M. Gobelt, Y. Jin, H. Nakayama, S. Poles, D. Di Stefano, Real-world applications of multiobjective optimization, in Multiobjective Optimization: Interactive and Evolutionary Approaches, J. Branke, K. Deb, K. Miettinen, and R. S lowinski, eds., Lecture Notes in Comput. Sci. 5252, Springer, Cham, (2008), 285-327.
[5] 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.
[6] J. Branke, K. Deb, K. Miettinen, R. Slowinski, Multiobjective Optimization: Interactive and Evolutionary Approaches, Springer, Berlin, 2008.
[7] G.Y. Chen, X.X. Huang, X.Q. Yang, Vector Optimization: Set-Valued and Variational Analysis, Springer, Berlin, 2005.
[8] J. Jahn, Vector Optimization: Theory, Applications, and Extensions, Springer, New York, 2011.
[9] L.M. Graña Drummond, A.N. Iusem, A projected gradient method for vector optimization problems, Comput. Optim. Appl. 28 (2004), 5-29.
[10] X.P. 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.
[11] X.P. Zhao, J.C. Yao, Linear convergence of a nonmonotone projected gradient method for multiobjective optimization, J. Global Optim. 82 (2022), 577-594.
[12] 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.
[13] J. Fliege, B.F. Svaiter, Steepest descent methods for multicriteria optimization, Math. Methods Oper. Res. 51 (2000), 479-494.
[14] L. Ceng, J.C. Yao, Approximate proximal methods in vector optimization, European J. Oper. Res. 183 (2007), 1-19.
[15] E.A. Papa Quiroz, H.C.F. Apolinário, K.D. Villacorta, P.R. Oliveira, A linear scalarization proximal point method for quasiconvex multiobjective minimization, J. Optim. Theory Appl. 183 (2019), 1028-1052.
[16] E.A. Papa Quiroz, S. Cruzado, An inexact scalarization proximal point method for multiobjective quasiconvex minimization, Ann. Oper. Res. (2020), https://doi.org/10.1007/s10479-020-03622-8.
[17] L.R. Lucambio Pérez, L.F. Prudente, Nonlinear conjugate gradient methods for vector optimization, SIAM J. Optim. 28 (2018), 2690-2720.
[18] J. Fliege, L.M. Graña Drummond, B.F. Svaiter, Newton's method for multiobjective optimization, SIAM J. Optim. 20 (2009), 602-626.
[19] J. Wang, Y. Hu, C.K.W. Yu, C. Li, X. Yang, Extended Newton methods for multiobjective optimization: Majorizing function technique and convergence analysis, SIAM J. Optim. 29 (2019), 2388-2421.
[20] G.A. Carrizo, P.A. Lotito, M.C. Maciel, Trust region globalization strategy for the nonconvex unconstrained multiobjective optimization problem, Math. Program. 159 (2016), 339-369.

[21] J. Thomann, G. Bichfelder, A trust-region algorithm for heteroheneous multiobjective optimization, SIAM J. Optim. 29 (2019), 1017-1047.
[22] 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.
[23] R.S. Burachik, C.Y. Kaya, M.M. Rizvi, A new scalarization technique and new algorithms to generate Pareto fronts, SIAM J. Optim. 27 (2017), 1010-1034.
[24] J. Fliege, A.I. F.Vaz, L.N. Vicente, Complexity of gradient descent for multiobjective optimization, Optim. Methods Softw. 34 (2019), 949-959.
[25] 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.
[26] L. Grippo, F. Lampariello, S. Lucidi, A nonmonotone line search technique for Newton's method, SIAM J. Numer. Anal. 23 (1986), 707-716.
[27] Y.H. Dai, On the nonmonotone line search, J. Optim. Theory Appl. 112 (2002), 315-330.
[28] P.L. Toint, An assessment of non-monotone line search techniques for unconstrained optimization, SIAM J. Sci. Comput. 17 (1996), 725-739.
[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] N. Mahdavi-Amiri, F. Salehi Sadaghiani, A superlinearly convergent nonmonotone quasi-Newton method for unconstrained multiobjective optimization, Optim. Methods Softw. 35 (2020), 1223-1247.
[31] K. Mita, E.H. Fukuda, N. Yamashita, Nonmonotone line searches for unconstrained multiobjective optimization problems, J. Global Optim. 75 (2019), 63-90.
[32] S.J. Qu, Y. Ji, J.L. Jiang, Q.P. Zhang, Nonmonotone gradient methods for vector optimization with a portfolio optimization application, European J. Oper. Res. 263 (2017), 356-366.
[33] N.S. Fazzio, M.L. Schuverdt, Convergence analysis of a nonmonotone projected gradient method for multiobjective optimization problems, Optim. Lett. 13 (2019), 1365-1379.
[34] D.T. Luc, Theory of Vector Optimization, Lecture Notes in Economics and Mathematical Systems, Springer, Berlin, 1989.
[35] R.T. Rockafellar, Convex Analysis, Priceton University Press, Priceton, 1970.
[36] R.T. Rockafellar, R. J.-B. Wets, Variational Analysis, Springer, Berlin, 1998.
[37] J.Y. Bello Cruz, L.R. Lucambio Pérez, J.G. Melo, Convergence of the projected gradient method for quasi-convex multiobjective optimization, Nonlinear Anal. 74 (2011), 5268-5273.
[38] 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.
[39] J.-B. Hiriart Urruty, C. Lemarèchal, Convex Analysis and Minimization Algorithms, Springer, Berlin, 1993.
[40] 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.