Comput Optim Appl
DOI 10.1007/s10589-012-9494-7

A subgradient method for multiobjective optimization

J.X. Da Cruz Neto · G.J.P. Da Silva · O.P. Ferreira · J.O. Lopes

Received: 18 February 2011

Abstract A method for solving quasiconvex nondifferentiable unconstrained multiobjective optimization problems is proposed in this paper. This method extends to the multiobjective case of the classical subgradient method for real-valued minimization. Assuming the basically componentwise quasiconvexity of the objective components, full convergence (to Pareto optimal points) of all the sequences produced by the method is established.

Keywords Pareto optimality or efficiency · Multiobjective optimization · Subgradient method · Quasi-Féjer convergence

J.X. Da Cruz Neto was partially supported by CNPq GRANT 301625-2008 and PRONEX-Optimization (FAPERJ/CNPq).

G.J.P. Da Silva was partially supported by PADCT-CNPq.

O.P. Ferreira was supported in part by FUNAPE/UFG, CNPq Grants 201112/2009-4, 475647/2006-8 and PRONEX-Optimization (FAPERJ/CNPq).

J.O. Lopes was partially supported by INCTMAT-CNPq.

J.X. Da Cruz Neto (✉) · J.O. Lopes
DM, Universidade Federal do Piauí, Teresina, PI 64049-500, Brazil
e-mail: jxavier@ufpi.br

J.O. Lopes
e-mail: jurandir@ufpi.br

G.J.P. Da Silva · O.P. Ferreira
IME, Universidade Federal de Goiás, Campus II, Caixa Postal 131, Goiânia, GO 74001-970, Brazil

G.J.P. Da Silva
e-mail: geci@mat.ufg.br

O.P. Ferreira
e-mail: orizon@mat.ufg.br

Published online: 21 June 2012

1 Introduction

In multiobjective optimization, several objective functions have to be simultaneously minimized. For nontrivial problems, no single point will minimize all given objective functions at once, and so the concept of optimality is to be replaced by the concept of *Pareto optimality* or *efficiency*. One should recall that a point is called *Pareto optimal* or *efficient*, if there is no different point with the same, or smaller, objective function values, such that there is a decrease in at least one objective function value.

Multiobjective optimization problems are relevant in many economical applications, industry, agriculture, and others fields (see [5, 20]). Various procedures are used to solve multiobjective optimization problems, such as, for example: the weighting method (see [6, 10, 13, 18]), steepest descent method (see [4, 7, 8]) and nonsmooth multiobjective optimization methods (see [14, 15]).

The purpose of this paper is to present a method to find a solution to the unconstrained multiobjective optimization problem, where each objective function is a quasiconvex (non-necessarily differentiable) function. As far as we know, this is the first extension of a classical scalar-valued method to the context of nonsmooth multiobjective. It is worth to point out that, even if the objective functions under consideration are differentiable, our method is different from that proposed by [4, 7, 8] as, in this case, we use a divergent series stepsize rule and a convex combination of the subgradients of the objective functions. Note that this is not, in general, a descent direction method.

This article is organized as follows. In Sect. 2, we state the problem under consideration (see (2.1) below) and the necessary notation. Section 3 presents an algorithm to solve that problem. In Sect. 4, we study the convergence analysis of this algorithm under certain hypotheses.

2 Basic assumptions and properties

The following notation is used throughout our presentation. We denote by $\langle \cdot, \cdot \rangle$ the usual inner product of $\mathbb{R}^n$ and by $\|\cdot\|$ its corresponding norm. Consider the function $F : \mathbb{R}^n \to \mathbb{R}^m$, given by $F(x) = (f_1(x), f_2(x), \dots, f_m(x))$. We are interested in the problem

$$\min_{x \in \mathbb{R}^n} F(x) \tag{2.1}$$

with the following meaning: a point $x^* \in \mathbb{R}^n$ is a *Pareto Optimal* or *efficient* solution (2.1) if, and only if, there exists no $x \in \mathbb{R}^n$ such that $F(x) \le F(x^*)$ and $F(x) \ne F(x^*)$. Here, the inequality sign $\le$ between vectors is to be understood in a componentwise sense. Similarly, in what follows, a strict inequality $F(x) < F(x^*)$ is to be understood componentwise, too. We denote by $S^* \subset \mathbb{R}^n$ the set of all Pareto optimal or efficient points.

Remark 2.1 Note that x^* is a Pareto optimal point if there is no other point x such that $f_i(x) \le f_i(x^*)$ $\forall i = 1, \dots, m$ and $f_j(x) < f_j(x^*)$ for at least one j.

Let us start by introducing the definition of the so-called convex and quasiconvex function.

Definition 2.2 A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be convex if, for every $x, y \in \mathbb{R}^n$ and for every $\alpha \in [0, 1]$ the following inequality holds:

$$f\big((1-\alpha)x + \alpha y\big) \le (1-\alpha) f(x) + \alpha f(y).$$

Definition 2.3 A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be quasiconvex if, for every $x, y \in \mathbb{R}^n$ and for every $\alpha \in [0, 1]$ the following inequality holds:

$$f\big((1-\alpha)x + \alpha y\big) \le \max\big\{f(x), f(y)\big\}.$$

Remark 2.4 If f is a convex function, then f is a quasiconvex function.

For classical multiobjective optimization purposes, these notions have been naturally adapted to functions with values in a finite-dimensional Euclidean space, as follows:

Definition 2.5 A function $F : \mathbb{R}^n \to \mathbb{R}^m$, given by $F(x) = (f_1(x), \dots, f_m(x))$ is said to be componentwise convex (quasiconvex) if all component functions $f_1, \dots, f_m$ are convex (quasiconvex).

Definition 2.6 For a given $\lambda \in \mathbb{R}$ the level (respectively, strict level) set of f, corresponding to λ, is the set:

$$S_f^{\le}(\lambda) = \big\{x \in \mathbb{R}^n \,|\, f(x) \le \lambda\big\},$$

respectively,

$$S_f^{<}(\lambda) = \big\{x \in \mathbb{R}^n \,|\, f(x) < \lambda\big\}.$$

The next proposition characterizes quasiconvex functions.

Proposition 2.7 *A function $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex if, and only if $S_f^{\le}(\lambda)$ is a convex set for all $\lambda \in \mathbb{R}$.*

Proof See [1]. □

Definition 2.8 The normal cone of $C \subset \mathbb{R}^n$ at point $z \in \mathbb{R}^n$ is denoted and defined by

$$N_C(z) = \big\{w \in \mathbb{R}^n \,|\, \langle w, x - z\rangle \le 0 \ \forall x \in C\big\}.$$

Proposition 2.9 *Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. Then f is quasiconvex if, and only if*

$$\nabla f(x) \in N_{S_f^{<}(f(x))} \quad \forall x \in \mathbb{R}^n.$$

Proof See [1]. □

Let us recall that the usual Fenchel-Moreau subdifferential (see [17]):

Definition 2.10 Given a function $f : \mathbb{R}^n \to \mathbb{R}$ and a point $z \in \mathbb{R}^n$, the Fenchel-Moreau subdifferential of f at z is defined and denoted by:

$$\partial f(z) = \big\{ s \in \mathbb{R}^n : \langle s, x - z \rangle \leq f(x) - f(z) \ \forall x \in \mathbb{R}^n \big\}. \tag{2.2}$$

We use the adapted subdifferentials of quasiconvex functions, namely the *Plastria subdifferential* (see [16]) and *infradifferential*, or *Gutiérrez subdifferential* (see [9]), which are defined as follows:

Definition 2.11 Given a function $f : \mathbb{R}^n \to \mathbb{R}$ and a point $z \in \mathbb{R}^n$, the Plastria's lower subdifferential of f at z is defined and denoted by

$$\partial^{<} f(z) = \big\{ s \in \mathbb{R}^n : \langle s, x - z \rangle \leq f(x) - f(z) \ \forall x \in S_f^{<}(z) \big\}.$$

Definition 2.12 Given a function $f : \mathbb{R}^n \to \mathbb{R}$ and a point $z \in \mathbb{R}^n$, the Gutiérrez subdifferential of f at z is defined and denoted by:

$$\partial^{\leq} f(z) = \big\{ s \in \mathbb{R}^n : \langle s, x - z \rangle \leq f(x) - f(z) \ \forall x \in S_f^{\leq}(z) \big\}.$$

We emphasize that the Plastria subdifferential and Gutiérrez subdifferential defined above is an adaptation to the quasiconvex case of the Fenchel-Moreau subdifferential, as defined in (2.2). It is clear that

$$\partial f(z) \subset \partial^{\leq} f(z) \subset \partial^{<} f(z).$$

Moreover, for every $s \in \partial^{\leq} f(z)$ it follows that $\{\theta s | \theta \geq 1\} \subset \partial^{\leq} f(z)$. This shows that the set $\partial^{\leq} f(z)$ is not bounded. Hence, $\partial f(z)$ and $\partial^{\leq} f(z)$ or $\partial^{<} f(z)$ are generally different, even for convex functions.

Proposition 2.13 *Let $z \in \mathbb{R}^n$. If the closure of $S_f^{<}(f(z))$ equals $S_f^{\leq}(f(z))$ then $\partial^{<} f(z) = \partial^{\leq} f(z)$.*

Proof See [19]. □

Definition 2.14 A function $f : \mathbb{R}^n \to \mathbb{R}$ quasiconvex is said to be essentially quasiconvex if each local minimizer is global.

Remark 2.15 If f is a convex function, then f is essentially a quasiconvex function.

Corollary 2.16 *If $f : \mathbb{R}^n \to \mathbb{R}$ is essentially a quasiconvex continuous function, then the closure of $S_f^{<}(f(z))$ equals $S_f^{\leq}(f(z))$ for all non-minimizer z of f. Consequently, $\partial^{<} f(z) = \partial^{\leq} f(z)$ for all non-minimizer z of f.*

Proof Let x belonging to the closure of $S^{<}_{f}(f(z))$, then there exists a sequence $\{x^k\} \subset S^{<}_{f}(f(z))$ such that $\lim_{k\to\infty} x^k = x$. Since $f(x^k) < f(z)$ we have

$$\lim_{k\to\infty} f\left(x^k\right) \leq f(z) \quad \Longrightarrow \quad f(x) \leq f(z) \quad \Longrightarrow \quad x \in S^{\leq}_{f}\big(f(z)\big).$$

Let $x \in S^{\leq}_{f}(f(z))$, we have that

$$f(x) \leq f(z) \quad \Longrightarrow \quad x \in S^{<}_{f}\big(f(z)\big) \quad \text{or} \quad f(x) = f(z).$$

Since z is not a minimizer of f, we obtain that x is not a local minimizer. Hence we can obtain a convergent sequence $\{x^k\}$ such that $\lim_{k\to\infty} x^k = x$ and $f(x^k) < f(z)$ for all k. Thus $x^k \in S^{<}_{f}(f(z))$ for all k, which implies x belonging to the closure of $S^{<}_{f}(f(z))$. So, the closure of $S^{<}_{f}(f(z))$ equals $S^{\leq}_{f}(f(z))$ for all non-minimizer z of f. From Proposition 2.13 we obtain that $\partial^{<} f(z) = \partial^{\leq} f(z)$ for all non-minimizer z of f. □

Definition 2.17 A function $f : \mathbb{R}^n \to \mathbb{R}$ is said Lipschitz if there exists a number $0 \leq L < \infty$ such that

$$\big|f(x) - f(y)\big| \leq L\|x - y\|, \quad \forall x, y \in \mathbb{R}^n.$$

We summarize in the next theorem the main results of the Plastria subdifferential and Gutiérrez subdifferential for the case of quasiconvex and Lipschitz functions:

Theorem 2.18 *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be quasiconvex and Lipschitz continuous function with constant L_i for $i = 1, 2, \ldots, m$. Then the following conditions hold*:

(i) *$\partial^{<} f_i(z) \neq \emptyset$ for every $z \in \mathbb{R}^n$. Moreover, there exists $s_i \in \partial^{<} f_i(z)$ with $\|s_i\| \leq L_i$ for $i = 1, 2, \ldots, m$;*
(ii) *$\partial^{<} f_i(z)$ is a closed and convex set for every $z \in \mathbb{R}^n$ and $i = 1, 2, \ldots, m$;*
(iii) *$0 \in \partial^{<} f_i(z)$ if only if $z \in \mathbb{R}^n$ is global minimizer of f_i, in which case $\partial^{<} f_i(z) = \mathbb{R}^n$, for $i = 1, 2, \ldots, m$.*

Proof See [16]. □

Theorem 2.19 *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be essentially quasiconvex and Lipschitz continuous functions for $i = 1, 2, \ldots, m$. Then the following conditions hold*:

(i) *$\partial^{\leq} f_i(z) \neq \emptyset$ for every $z \in \mathbb{R}^n$ for $i = 1, 2, \ldots, m$.*
(ii) *$\partial^{\leq} f_i(z)$ is a closed and convex set for every $z \in \mathbb{R}^n$ and for $i = 1, 2, \ldots, m$;*
(iii) *$0 \in \partial^{\leq} f_i(z)$ if only if $z \in \mathbb{R}^n$ is global minimizer of f_i, in which case $\partial^{\leq} f_i(z) = \{0\}$, for $i = 1, 2, \ldots, m$.*

Proof See [19]. □

Remark 2.20 Since $\partial^{\leq} f_i(z) \subset \partial^{<} f_i(z)$, by Theorem 2.18 there exists $s_i \in \partial^{\leq} f_i(z)$ with $\|s_i\| \leq L_i$, for $i = 1, 2, \ldots, m$.

We should point that there is no loss of generality in supposing that all f_i are essentially quasiconvex and Lipschitz continuous with the same constant, say L (it suffices to take $L = \max_{1\leq i\leq m} L_i$).

The following simple example illustrates that, in general, $\nabla f(x) \notin \partial^< f(x)$.

Example 1 Let $f : \mathbb{R} \to \mathbb{R}$ be the function given by

$$f(x) = \begin{cases} -9x - 54, & \text{if } x \leq -3, \\ x^3, & \text{if } -3 < x < 3, \\ 9x, & \text{if } x \geq 3. \end{cases}$$

It is easy to compute that

$$\nabla f(-2) = 12 \notin \partial^< f(-2) = [19, +\infty).$$

Moreover, the function in this example is essentially quasiconvex and Lipschitz continuous, but not convex. Nevertheless, in the case of essentially quasiconvex and Lipschitz continuous functions we can assure that a multiple of $\nabla f(x)$ belongs to $\partial^< f(x)$. More precisely, we have the following proposition:

Theorem 2.21 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a quasiconvex and Lipschitz function with constant L. Then, for all $v \in N_{S^<_f(f(z))}$ and $0 < \|v\| < 1$, the vector $\tilde{v} = Lv$ belongs to $\partial^< f(z)$. Additionally, for all $u \in \partial^< f(z)$ and $u \neq 0$, the vector $\tilde{u} = Lu/\|u\|$ belongs to $\partial^< f(z)$.*

Proof See [3]. □

Corollary 2.22 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a quasiconvex, differentiable and Lipschitz function with constant L. If $z \in \mathbb{R}^n$ such that $\nabla f(z) \neq 0$ then $g := L\nabla f(z)/\|\nabla f(z)\| \in \partial^< f(z)$.*

Proof It follows from Proposition 2.9 and Theorem 2.21. □

We present now the preliminary result, in the form of recurrence numerical inequalities, which is needed in our convergence analysis. We recall next the so called quasi-Fejér convergence theorem, used in many papers to analyze the gradient and subgradient methods (see [2, 8, 12]).

Definition 2.23 A sequence $\{w^k\} \subset \mathbb{R}^n$ is called quasi-Fejér convergent to a set $U \subset \mathbb{R}^n$ if for every $u \in U$ there exists a sequence $\{\epsilon_k\} \subset \mathbb{R}_+$ such that

$$\left\|w^{k+1} - u\right\|^2 \leq \left\|w^k - u\right\|^2 + \epsilon_k,$$

with $\sum_{k=0}^{\infty} \epsilon_k < \infty$.

Proposition 2.24 *If $\{w^k\} \subset \mathbb{R}^n$ is a quasi-Fejér convergent sequence to a nonempty set U then $\{w^k\}$ is bounded. Further, if a cluster point $\overline{w}$ of $\{w^k\}$ belongs to U then $\lim_{k\to\infty} w^k = \overline{w}$.*

Proof See [2]. □

3 The algorithm

We describe in this section the algorithm which is introduced in this paper to solve (2.1). In order to guarantee its well-definedness, we will assume the following condition:

Assumption 3.1 *The components $f_i : \mathbb{R}^n \to \mathbb{R}$ of the objective function $F : \mathbb{R}^n \to \mathbb{R}^m$ are essentially quasiconvex and Lipschitz continuous functions with constant L, for $i = 1, 2, \ldots, m$.*

Now we define our algorithm:

Algorithm

1. **Initialization.** Take $\delta \in (0, 1)$, $x^0 \in \mathbb{R}^n$ and set $k = 0$.
2. **Iterative step.** Take ${s_i}^k \in \partial^{\leq} f_i(x^k)$ such that $\|{s_i}^k\| \leq L$ for $i = 1, \ldots, m$. Define

$$s^k \in \arg\max_{1 \leq i \leq m} \|s_i^k\|. \tag{3.1}$$

   (a) **Stopping criterion.** If $\|s^k\| = 0$ or $x^k \in S^*$ (Pareto optimal set) then **STOP**. Otherwise, continue.
   (b) **Definition of x^{k+1}.** Define

$$x^{k+1} = x^k - \frac{\alpha_k}{\|s^k\|} \sum_{i=1}^{m} \lambda_{i,k} s_i^k; \tag{3.2}$$

   where

$$\sum_{k=1}^{\infty} \alpha_k = \infty, \quad \sum_{k=0}^{\infty} \alpha_k^2 < \infty \quad \text{and} \tag{3.3}$$

$$\{\lambda_{i,k}\} \subset [\delta, 1] \quad \text{for } i = 1, \ldots, m, \qquad \sum_{i=1}^{m} \lambda_{i,k} = 1. \tag{3.4}$$

Remark 3.2 It follows immediately that, if $m = 1$, i.e., if $F : \mathbb{R}^n \to \mathbb{R}$, we have that

$$x^{k+1} = x^k - \frac{\alpha_k}{\|s^k\|} s^k,$$

which is the scalar-valued subgradient method under condition (3.3).

Proposition 3.3 *Suppose that Assumption 3.1 holds. Thus, for $i = 1, 2, \ldots, m$ there exists ${s_i}^k \in \partial^{\leq} f_i(x^k)$ for all k.*

Proof It follows from Theorem 2.19. □

Remark 3.4 The proposition above guarantees the well-definedness of our algorithm.

From now on, we assume that sequence $\{x^k\}$ generated by the algorithm is infinite, without explicitly mentioning it in the statement of our results. This assumption is reasonable because: (a) if s^k is equal to 0 then $s_i^k = 0$ for all $i = 1, \dots, m$. Hence x^k is a minimizer of each f_i (see Theorem 2.19) and thus a Pareto optimal solution to the problem (2.1); or (b) If $x^k \in S^*$ for some $k \geq 0$ then the problem (2.1) is solved. Therefore, let us assume that $x^k \notin S^*$ for all k.

4 Convergence analysis

In this section, we prove that the sequence $\{x^k\}$ generated by our algorithm converges to a Pareto optimal. In order to do this, we will also assume the following condition:

Assumption 4.1 *For all sequence* $\{z^k\} \subset \mathbb{R}^n \setminus S^*$, *there exists* $z \in \mathbb{R}^n$ *such that* $F(z) \leq F(z^k)$ *for* $k = 0, 1, 2, \dots$.

Remark 4.2 We should point out some consequences from the above assumption:

(i) S^* must be non empty for the case $m = 1$;
(ii) If the above assumption is true for all $\{z_k\} \subset \mathbb{R}^n$ then there is a unique optimal value;
(iii) If $\{F(z^k)\}$ is a $\mathbb{R}^m_+$-decreasing sequence (see [13] for this definition) then $z^k \notin S^*$ for all $k = 0, 1, 2, \dots$, which is standard in the analysis of extensions of classical methods to vector optimization (see [4, 8]).

For the sequence $\{x^k\}$ generated by our algorithm we define the following set:

$$T := \{y \in \mathbb{R}^n : F(y) \leq F(x^k)\ \forall k\}.$$

Under the Assumption 4.1 we can guarantee this set is nonempty.

Proposition 4.3 *For any* $y \in T$, *it holds that*

$$\|x^{k+1} - y\|^2 \leq \|x^k - y\|^2 + \alpha_k^2 - \frac{2\alpha_k}{\|s_\ell^k\|} \sum_{i=1}^{m} \lambda_{i,k}\big(f_i(x^k) - f_i(y)\big).$$

As a result,

$$\|x^{k+1} - y\|^2 \leq \|x^k - y\|^2 + \alpha_k^2, \quad k = 0, 1, \dots.$$

Proof First, we prove the following inequality:

$$\|x^{k+1} - x^k\| \leq \alpha_k, \quad k = 0, 1, \dots. \tag{4.1}$$

In order to prove this inequality, note that, from (3.2), the triangular inequality and (3.1) we get

$$\|x^{k+1} - x^k\| = \frac{\alpha_k}{\|s^k\|}\left\|\sum_{i=1}^{m}\lambda_{i,k}s_i^k\right\| \le \frac{\alpha_k}{\|s^k\|}\sum_{i=1}^{m}\lambda_{i,k}\|s_i^k\| \le \alpha_k.$$

After simple algebraic manipulations, we obtain, for all $y \in \mathbb{R}^n$ that

$$\|x^{k+1} - y\|^2 = \|x^k - y\|^2 + \|x^{k+1} - x^k\|^2 - 2\langle x^{k+1} - x^k, y - x^k\rangle, \quad k = 0, 1, \ldots.$$

Therefore, using the last equality, (4.1) and (3.2) it easy to conclude that

$$\|x^{k+1} - y\|^2 \le \|x^k - y\|^2 + \alpha_k^2 - \frac{2\alpha_k}{\|s_\ell^k\|}\sum_{i=1}^{m}\lambda_{i,k}\langle s_i^k, x^k - y\rangle.$$

Since $s_i^k \in \partial^{\le} f_i(x^k)$, we have, for any $y \in S_f^{\le}(f(x^k))$ and for $i = 1, 2, \ldots, m$ that

$$\langle s_i^k, x^k - y\rangle \ge f_i(x^k) - f_i(y).$$

Combining this with the last inequality, we obtain the first statement of the proposition.

For proving the last statement of the proposition, use the first one and that for any $y \in T$ we have $f_i(x^k) - f_i(y) \ge 0$ for $i = 1, 2, \ldots, m$. □

Now we show that all sequences produced by the method are convergent.

Theorem 4.4 *The sequence $\{x^k\}$ converges to some $x^* \in T$.*

Proof In view of (3.3) and the second inequality in Proposition 4.3, $\{x^k\}$ is quasi-Fejér convergent to T and so, since $T \ne \emptyset$, by Proposition 2.24, it is a bounded sequence. Since $L \ge \|s^k\|$, then for $y \in T$ we obtain

$$\frac{\alpha_k}{\|s^k\|}\sum_{i=1}^{m}\lambda_{i,k}\big(f_i(x^k) - f_i(y)\big) \ge \frac{\alpha_k}{L}\sum_{i=1}^{m}\lambda_{i,k}\big(f_i(x^k) - f_i(y)\big).$$

For the sake of a simpler notation, let us define the sequence

$$\gamma_k := \frac{2}{L}\sum_{i=1}^{m}\lambda_{i,k}\big(f_i(x^k) - f_i(y)\big) \ge 0, \quad k = 0, 1, \ldots.$$

Hence, by using the first inequality in Proposition 4.3, we conclude that

$$\alpha_k\gamma_k \le \|x^k - y\|^2 - \|x^{k+1} - y\|^2 + \alpha_k^2, \quad k = 0, 1, \ldots.$$

Using the last inequality we have, for $p \in \mathbb{N}$

$$\sum_{k=0}^{p}\alpha_k\gamma_k \le \|x^0 - y\|^2 - \|x^{p+1} - y\|^2 + \sum_{k=0}^{p}\alpha_k^2,$$

which, by (3.3) implies that $\sum_{k=0}^{\infty} \alpha_k \gamma_k < \infty$. We conclude that $\liminf \gamma_k = 0$ because also $\sum_{k=0}^{+\infty} \alpha_k = \infty$. Since $\gamma_k \geq 0$, there exists an subsequence $\{\gamma_{k_j}\}$ such that $\lim_{j\to\infty} \gamma_{k_j} = 0$. Using (3.4) and the definition of sequence γ_k we obtain that

$$\lim_{j\to\infty} f_i\left(x^{k_j}\right) = f_i(y), \quad i = 1, 2, \ldots, m.$$

Let x^* be a cluster point of $\{x^{k_j}\}$ such that $x^* = \lim_{l\to\infty} x^{k_{j_l}}$. Then we have

$$f_i\left(x^*\right) = \lim_{l\to\infty} f_i\left(x^{k_{j_l}}\right) = \lim_{j\to\infty} f_i\left(x^{k_j}\right) = f_i(y), \quad i = 1, \ldots, m.$$

Thus, it follows, from the last equation that $x^* \in T$, and so, by Proposition 2.24 the whole sequence $\{x^k\}$ converges to x^*. □

The following theorem is our main result.

Theorem 4.5 *Sequence $\{x^k\}$ converges to a Pareto optimal solution of the problem* (2.1).

Proof From Theorem 4.4, we have that $\{x^k\}$ converges to a point $x^* \in T$. Suppose that x^* is not a Pareto optimal solution of the problem (2.1). Then, there exists $\bar{x}$ and an index j such that

$$f_i(\bar{x}) \leq f_i\left(x^*\right), \quad i = 1, 2, \ldots, m \quad \text{and} \quad f_j(\bar{x}) < f_j\left(x^*\right). \tag{4.2}$$

Since $x^* \in T$, we have from the definition of T and the above conditions that $\bar{x} \in T$. Using the first inequality in Proposition 4.3 and (3.4), we conclude that

$$\left\|x^{k+1} - \bar{x}\right\|^2 \leq \left\|x^k - \bar{x}\right\|^2 + \alpha_k^2 - \frac{2\delta\alpha_k}{L} \sum_{i=1}^{m} \left(f_i\left(x^k\right) - f_i(\bar{x})\right).$$

Now, from (4.2) and the fact that $x^* \in T$

$$0 < \rho := \frac{\delta}{L} \sum_{i=1}^{m} \left(f_i\left(x^*\right) - f_i(\bar{x})\right) \leq \frac{\delta}{L} \sum_{i=1}^{m} \left(f_i\left(x^k\right) - f_i(\bar{x})\right).$$

Combining the last two inequalities, we obtain

$$\left\|x^{k+1} - \bar{x}\right\|^2 \leq \left\|x^k - \bar{x}\right\|^2 + \alpha_k(\alpha_k - 2\rho). \tag{4.3}$$

As $\lim_{k\to\infty} \alpha_k = 0$ and $\rho > 0$, there exists k_0 such that

$$\alpha_k < \rho \quad \text{for all } k \geq k_0.$$

From (4.3) and the last inequality, we conclude that

$$\rho\alpha_k \leq \left\|x^k - \bar{x}\right\|^2 - \left\|x^{k+1} - \bar{x}\right\|^2, \quad k \geq k_0.$$

Take $l \in \mathbb{N}, l \geq k_0$. Adding up in the above inequality from $j = k_0$ to $j = l$, we get

$$\sum_{j=k_0}^{l} \alpha_j \leq \frac{1}{\rho}\left(\|x^{k_0} - \overline{x}\|^2 - \|x^{l+1} - \overline{x}\|^2\right) \leq \frac{1}{\rho}\|x^{k_0} - \overline{x}\|^2.$$

It follows that $\sum_{j=k_0}^{\infty} \alpha_j < \infty$, in contradiction with $\sum_{k=0}^{\infty} \alpha_j = \infty$. Therefore x^* is a Pareto optimal solution of the multiobjective problem with objective function F. □

5 Final remarks

We proposed a subgradient type method for nonsmooth unconstrained multiobjective optimization where the objective function is componentwise essentially quasiconvex and Lipschitz continuous. Under reasonable assumptions, we proved that there exists an efficient solution and, for any initial point, the sequence produced by this method converges to Pareto optimal solution of the problem.

In this paper, we do not study the subgradient method for multiobjective optimization for its practical value. Indeed, it is not suitable for "real life" multiobjective problems like to the classical subgradient method for scalar-valued optimization problem. However, it is one that comes first to mind when starting from the ideas of the classical subgradient method, in our attempt to deal with more efficient methods (for example, ϵ-subgradient methods, bundle methods and cutting-plane algorithm for multiobjective optimization). On the other hand, we expect that it will be the first step toward a more efficient methods to come in the future, comparable to the classical subgradient method for scalar-valued nonsmooth optimization, which is bad but basic, since it is the departure point for many other more sophisticated and efficient algorithms, for instance, ϵ-subgradient methods, bundle methods and cutting-plane algorithm (see [11] for a very clear exposition). It remains an open question how to extend more efficient procedures as bundle methods and cutting-plane algorithm to multiobjective optimization. We foresee further progress along this path in the future.

References

1. Avriel, M.: Nonlinear Programming: Analysis and Methods. Dover, New York (2003)
2. Burachik, R.S., Graña Drummond, L.M., Iusem, A.N., Svaiter, B.F.: Full convergence of the steepest descent method with inexact line search. Optimization **32**, 137–146 (1995)
3. Censor, Y., Segal, A.: Algorithms for the quasiconvex feasibility problem. J. Comput. Appl. Math. **185**, 34–50 (2006)
4. Flieg, J., Svaiter, B.F.: Steepest descent methods for multicriteria optimization. Math. Methods Oper. Res. **51**, 479–494 (2000)
5. Gal, T., Hanne, T.: On the development and future aspects of vector optimization and MCDM. In: Cláco, J. (ed.) Multicriteria Analysis, pp. 130–145. Springer, Berlin (1997)
6. Geoffrion, A.M.: Proper efficiency and the theory of vector maximization. J. Math. Anal. Appl. **22**, 618–630 (1968)
7. Graña Drummond, L.M., Iusem, A.N.: A projected gradient method for vector optimization problems. Comput. Optim. Appl. **28**, 5–29 (2004)
8. Graña Drummond, L.M., Svaiter, B.F.: A steepest descent method for vector optimization. J. Comput. Appl. Math. **175**, 395–414 (2005)

9. Gutiérrez Díez, J.M.: Infragradients and directions of decrease. Rev. R. Acad. Cienc. Exactas Fís. Nat. Madr. **78**, 523–532 (1984)
10. Haimes, Y.Y., Lasdon, L.S., Wismer, D.A.: On a bicriterion formulation of the problems of integrated system identification and system optimization. IEEE Trans. Syst. Man Cybern. **SMC-1**, 296–297 (1971)
11. Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms II: Advanced Theory and Bundle Methods. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 306. Springer, Berlin (1993)
12. Iusem, A.N., Svaiter, B.F., Teboulle, M.: Entropy-like proximal methods in convex programming. Math. Oper. Res. **19**, 790–814 (1994)
13. Luc, T.D.: Theory of Vector Optimization. Lecture Notes in Economics and Mathematical Systems, vol. 319. Springer, Berlin (1989)
14. Mäkelä, M.M., Ännikkö, T.M.: Numerical solution of nonsmooth optimal control problems with an application to the continuous casting process. Adv. Math. Sci. Appl. **4**, 491–515 (1994)
15. Miettinen, K., Mäkelä, M.M.: An interactive method for nonsmooth multiobjective optimization with an application to optimal control. Optim. Methods Softw. **2**, 31–44 (1993)
16. Plastria, F.: Lower subdifferentiable functions and their minimization by cutting planes. J. Optim. Theory Appl. **46**, 37–53 (1985)
17. Rockafellar, B.F.: Convex Analysis. Princeton University Press, Princeton (1970)
18. Tammer, C., Winkler, K.: A new scalarization approach and applications in multicriteria. J. Nonlinear Convex Anal. **4**, 365–380 (2003)
19. Xu, H., Rubinov, A.M., Glover, B.M.: Strict lower subdifferentiability and applications. J. Aust. Math. Soc. Ser. B, Appl. Math **40**, 379–391 (1999)
20. White, D.J.: A bibliography on the applications of mathematical programming multiple objective methods. J. Oper. Res. Soc. **41**, 669–691 (1990)