J. Appl. Numer. Optim. 8 (2026), No. 2, pp. 233-251
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.8.2026.2.05

SEQUENTIAL UPPER APPROXIMATION METHOD FOR SOLVING MULTIOBJECTIVE OPTIMIZATION PROBLEMS

ABDESSAMAD FADIL[1], ABDELOUAFI GHAZI[2,*], AHMED ROUBI[1]

[1]*Hassan First University of Settat, Faculté des Sciences et Techniques, Laboratoire MDSET, 26000, Morocco*
[2]*Laboratory of Applied Mathematics and Scientific Computing, Higher School of Education and Training, Sultan Moulay Slimane University, P.O. Box 523, Beni Mellal, 23000, Morocco*

Abstract. We develop a general algorithm for solving constrained multiobjective optimization problems. Our approach builds on the sequential upper approximation method already used in the framework of scalar optimization, in which a nonlinear optimization problem is solved by solving a sequence of simpler problems where the objective and constraint functions upper estimate those of the original problem. For a broad class of problems, we show that every cluster point of the sequence generated by our algorithm satisfies KKT-type conditions expressed in terms of directional derivatives. Finally we end with some numerical tests to illustrate the behavior of our algorithm.
Keywords. Multiobjective minimization; Nonconvex optimization; Optimality conditions; Successive upper approximation methods.
2020 Mathematics Subject Classification. 90C29, 65K05.

1. INTRODUCTION

Multiobjective programming (MOP) or Pareto optimization encompasses optimization problems with multiple objective functions, each representing distinct goals or criteria. These problems frequently occur in diverse fields such as engineering, economics, logistics, design, industry, agriculture, and environmental management [9, 10, 11, 12, 21, 22, 35]. The goal of solving MOP is to find solutions that effectively balance trade-offs between conflicting objectives, typically seeking a Pareto optimal solution where improving one objective necessitates the compromise of anothers.

Many researchers have developed numerous results concerning necessary/sufficient optimality conditions, as well as resolution algorithms for achieving Pareto optimality. Kanniappan [18] derive the necessary conditions for Pareto optimality of Fritz John and Kuhn-Tucker type, for a convex MOP using classical (i.e., scalarization) optimization method. See also [7, 13, 14, 15, 17, 26, 30, 31, 32, 36]. In [30], Ruíz-Canales et al. investigate the characterization of weakly Pareto solutions for MOP, considering quasiconvex conditions of the objective functions within a convex set of constraints. Necessary optimality conditions for MOP of a difference of convex mappings are established in terms of Lagrange-Fritz-John multipliers by

*Corresponding author.
E-mail address: ghaziab@hotmail.com (A. Ghazi).
Received 1 November 2025; Accepted 23 April 2026; Published online 1 May 2026.

Taa in [32]. Bonnel et al. [5] find weakly Pareto points in MOP using proximal methods. Gao et al. [13] introduce a novel type of properly approximate weak solution for MOP and provides necessary and sufficient conditions for these solutions through nonlinear scalarizations. Using the concepts of strong subdifferential and ε-subdifferential, necessary optimality conditions are established for an ε-weak Pareto minimal point and an ε-proper Pareto minimal point in MOP, where the objective function and constraint set are represented as difference of two vector-valued maps [14]. Ivanov [17] determine the second and first-order optimality conditions of Kuhn–Tucker and Fritz John types for weak Pareto solutions in MOP with inequality constraints.

From the computational point of view, various scalar optimization techniques have been effectively adapted to MOP. These include the projected gradient and subgradient methods [6, 8, 36, 37], steepest descent method [3], Newton method [19], trust-region method [33], and several others.

The famous approach used to solve MOP problems is based on the improvement function [16, 20, 23, 25, 34]. In practice, this approach requires finding a global minimum at each iteration, which is straightforward when the improvement function is convex. However, this is generally not the case if the data defining the problem are nonconvex. For this reason, we will approximate each function with a simpler convex function, and the constraints will also be approximated to define a simpler feasible region, enabling the solution of the new parametric problem. This technique which dates back to [24], was applied to the problem of real-valued optimization and was further developed in [1, 2, 28], allowing for nonconvexity in parts of the constraints or in the objective function.

The paper is organized into the following sections. In Section 2, we present a parametric approach that relies on parametric subproblems. We show that the problem MOP is equivalent to a scalar problem. In Section 3, upper bounding approximation functions and problems are developed, followed by a series of examples and some properties. Section 4 is dedicated to the development of necessary and sufficient optimality conditions for MOP expressed in terms of directional derivatives. The general sequential upper bounding algorithm for solving problem MOP and its convergence are the subject of Section 5. Finally, in Section 6, we perform some numerical tests to illustrate the behavior of the algorithm.

2. Parametric Approach for MOP

Consider a vector programming problem

$$\inf_{x\in X}\left[f(x) := (f_1(x), f_2(x), \dots, f_m(x))\right], \tag{P}$$

where $X = \{x \in C \mid h_j(x) \le 0, \forall j \in J\}$, such that $C \subset \mathbb{R}^n$ a nonempty, closed convex set, and the functions f_i, for $i \in I := \{1,2,\dots,m\}$, and h_j, for $j \in J := \{1,2,\dots,p\}$, are defined on $\mathbb{R}^n$. Let

$$h(x) := \max_{j\in J} h_j(x).$$

Then, the set X can be expressed as

$$X = \{x \in C \mid h(x) \le 0\}.$$

Solving these problems involves finding a solution $\bar{x} \in X$ such that there does not exist any $x \in X$ with $f(x) \neq f(\bar{x})$ and $f_i(x) \le f_i(\bar{x})$ for all $i \in I$. This solution is referred to as a Pareto minimum for (P). In the context of weak Pareto optimization, the objective is to find $\bar{x} \in X$ such that there does not exist any $x \in X$ with $f_i(x) < f_i(\bar{x})$ for all $i \in I$.

To develop our approach, we will use the concept of the improvement function $F_y : X \to \mathbb{R}$, where $y \in X$, defined by

$$F_y(x) := \max_{i \in I} \{f_i(x) - f_i(y)\}.$$

This function was previously employed to deal with vector programming problems. In [20, 23, 34], the authors defined the improvement function in the case of constrained vector programs. For the unconstrained case, the reader can refer to [16, 25].

We relate the following scalar optimization problem (P_y), with $y \in X$, to (P)

$$\inf_{x \in X} F_y(x) \tag{P_y}$$

The next two lemmas explain the link between the problems (P) and (P_y), for $y \in X$.

Lemma 2.1. *A point $\bar{x} \in X$ is a weak Pareto minimum of (P) if and only if it is a global minimum of $(P_{\bar{x}})$.*

Proof. Let $\bar{x} \in X$ be a weak Pareto minimum for (P). Suppose that $\bar{x}$ is not a global minimum for $(P_{\bar{x}})$. Then there exists $z \in X$ for which $F_{\bar{x}}(z) < F_{\bar{x}}(\bar{x})$.

Since $F_{\bar{x}}(\bar{x}) = 0$, it follows that

$$f_i(z) - f_i(\bar{x}) < 0 \quad \text{for all } i \in I,$$

which contradicts the assertion that $\bar{x}$ is a weak Pareto minimum of (P). Then, $\bar{x}$ is a global minimum for $(P_{\bar{x}})$.

Conversely, let's assume that $\bar{x} \in X$ is a global minimum for $(P_{\bar{x}})$. If $\bar{x}$ is not a weak Pareto minimum of (P), then there exists $z \in X$ such that

$$f_i(z) < f_i(\bar{x}) \quad \text{for all } i \in I,$$

which implies that

$$F_{\bar{x}}(z) := \max_{i \in I} \{f_i(z) - f_i(\bar{x})\} < 0 = F_{\bar{x}}(\bar{x}).$$

This contradicts the assumption that $\bar{x}$ is a global minimum for $(P_{\bar{x}})$. □

Lemma 2.2. *Let $\bar{x} \in X$. If the global minimum of $(P_{\bar{x}})$ is unique, Lemma 2.1 holds true if we replace weak Pareto minimum with Pareto minimum.*

Proof. Since a Pareto minimum of (P) is also a weak Pareto minimum of (P), then the first assertion follows from the previous lemma.

Conversely, assume that $\bar{x}$ is a global minimum for $(P_{\bar{x}})$, and $\bar{x}$ is not a Pareto minimum of (P). Then there exists $z \in X$ such that

$$f_i(z) \leq f_i(\bar{x}) \quad \text{for all } i \in I \text{ and } f_j(z) < f_j(\bar{x}) \quad \text{for some } j \in I. \tag{2.1}$$

It follows that

$$F_{\bar{x}}(z) \leq 0. \tag{2.2}$$

On the other hand, $\bar{x}$ is a global minimum for $(P_{\bar{x}})$ entails that

$$F_{\bar{x}}(x) \geq F_{\bar{x}}(\bar{x}) = 0 \quad \text{for all } x \in X.$$

With $x = z$ in that last inequality and the fact that $F_{\bar{x}}(z) \leq 0$, we get $F_{\bar{x}}(z) = 0$.

This means that z is also a global minimum of $(P_{\bar{x}})$. The unicity of such a minimum entails that $z = \bar{x}$. But this cannot hold by the second inequality in (2.1). Hence, $\bar{x}$ is a Pareto minimum of (P). □

In what follows, we will designate by $\tilde{x}_y$ a global minimum of F_y over X. Then we have the following results.

Proposition 2.1. *For every $y \in X$, the following assertions hold true:*

(1) $F_y(y) = 0$ and $F_y(\tilde{x}_y) \leq 0$,
(2) $f_i(\tilde{x}_y) \leq f_i(y)$ for all $i \in I$.

Proof. (1) The definition of $\tilde{x}_y$ indicates that $F_y(\tilde{x}_y) \leq F_y(x)$ for every $x \in X$. Substituting $x = y$ leads to the conclusion that $F_y(\tilde{x}_y) \leq F_y(y) = 0$.

(2) Utilizing the definition of F_y, with $x = \tilde{x}_y$, and Item 1 we derive that $f_i(\tilde{x}_y) - f_i(y) \leq F_y(\tilde{x}_y) \leq 0$ for all $i \in I$. This implies that $f_i(\tilde{x}_y) \leq f_i(y)$ for every $i \in I$.

□

Since solving globally the problem (P_y) is not always possible, we will approximate it with a simpler problem that is easier to solve. This approach will be detailed in the next section.

3. Upper Approximation for MOP

For a complete description of our method, we will be interested to functions that are directionally differentiable. To proceed, we first need to introduce some preliminary concepts and results. Let us begin by recalling that a function φ is said to be directionally differentiable at a point $x \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^n \setminus \{0\}$ if the one-sided directional derivative, defined as follows

$$\varphi'(x;d) = \lim_{t \downarrow 0} \frac{\varphi(x+td) - \varphi(x)}{t}$$

exists.

Bellow, We will list the minimal properties required for our approximating functions. More precisely, we construct approximating functions $\mathscr{U}_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ of f_i at $y \in C$, such that

(A1) $\mathscr{U}_i(y,y) = 0$ for all $i \in I$;
(A2) $\mathscr{U}_i(x,y) \geq f_i(x) - f_i(y)$ for all $i \in I$ and $x \in C$;
(A3) ${\mathscr{U}_i}'((z,y);x-y)|_{z=y} = f_i'(y;x-y)$ for all $i \in I$ and $x \in C$, where we respectively denote by ${\mathscr{U}_i}'((z,y);x-y)|_{z=y}$ and $f_i'(y;x-y)$ the directional derivatives of $\mathscr{U}_i(\cdot,y)$ and $f_i(\cdot)$ at y in the direction $x-y$.

Similarly, we construct approximating functions $\mathscr{H}_j : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ of h_j at $y \in C$, such that

(B1) $\mathscr{H}_j(y,y) = h_j(y)$ for all $j \in J$;
(B2) $\mathscr{H}_j(x,y) \geq h_j(x)$ for all $j \in J$ and $x \in C$;
(B3) $\mathscr{H}_j'((z,y);x-y)\Big|_{z=y} = h_j'(y;x-y)$ for all $j \in J$ and $x \in C$, where we respectively denote by $\mathscr{H}_j'((z,y);x-y)\Big|_{z=y}$ and $h_j'(y;x-y)$ the directional derivatives of $\mathscr{H}_j(\cdot,y)$ and $h_j(\cdot)$ at y in the direction $x-y$.

Using the functions $\mathscr{H}_j(x,y)$, we define the approximating constraint set for the subproblem (P_y) by

$$X_y = \left\{ x \in C \mid \mathscr{H}_j(x,y) \leq 0, \forall j \in J \right\}.$$

Remark 3.1. Note that hypothesis (B2) implies that $X_y \subset X$ for all $y \in C$. Remark on the other hand that, $y \in X_y$ if and only if $y \in X$.

In the following, we will present some examples where it is possible to construct the approximating functions $\mathscr{U}_i$ and $\mathscr{H}_j$ respectively satisfying (A1)-(A3) and (B1)-(B3).

Example 3.1. Assume that the functions f_i and h_j, $i \in I$ and $j \in J$, are directionally differentiable and let $\alpha > 0$. For any $y \in C$ we define

$$\mathscr{U}_i(x,y) := f_i(x) - f_i(y) + \tfrac{\alpha}{2}\|x-y\|^2 \text{ and } \mathscr{H}_j(x,y) := h_j(x) + \tfrac{\alpha}{2}\|x-y\|^2.$$

Example 3.2. Assume that the functions f_i and h_j, $i \in I$ and $j \in J$, are continuously differentiable with gradients respectively L_i and K_j-Lipschitz. From the descent lemma, see e.g. [4, Proposition A.24], for any $x, y \in C$, we have

$$f_i(x) \le f_i(y) + \langle \nabla f_i(y), x-y\rangle + \tfrac{L_i}{2}\|x-y\|^2 \quad \text{for all } i \in I,$$

and

$$h_j(x) \le h_j(y) + \langle \nabla h_j(y), x-y\rangle + \tfrac{K_j}{2}\|x-y\|^2 \quad \text{for all } j \in J.$$

Define the functions

$$\mathscr{U}_i(x,y) := \langle \nabla f_i(y), x-y\rangle + \tfrac{L_i}{2}\|x-y\|^2,$$

and

$$\mathscr{H}_j(x,y) := h_j(y) + \langle \nabla h_j(y), x-y\rangle + \tfrac{K_j}{2}\|x-y\|^2.$$

Example 3.3. Let the functions $f_i = f_i^1 - f_i^2$, $i \in I$ and $h_j = h_j^1 - h_j^2$, $j \in J$, and assume that f_i^ℓ and h_j^ℓ, for $\ell = 1,2$ are convex. Additionally, suppose that the functions f_i^2 and h_j^2 are continuously differentiable for all $i \in I$, $j \in J$. For any $y \in C$ we define the functions

$$f_{i,y}(x) = f_i^1(x) - \left[f_i^2(y) + \langle \nabla f_i^2(y), x-y\rangle\right] \quad \text{for all } i \in I,$$

and

$$h_{j,y}(x) = h_j^1(x) - \left[h_j^2(y) + \langle \nabla h_j^2(y), x-y\rangle\right] \quad \text{for all } j \in J.$$

The subgradient inequalities $f_i^2(y) + \langle \nabla f_i^2(y), x-y\rangle \le f_i^2(x)$ and $h_j^2(y) + \langle \nabla h_j^2(y), x-y\rangle \le h_j^2(x)$, imply that $f_i(x) \le f_{i,y}(x)$ and $h_j(x) \le h_{j,y}(x)$ for all $i \in I$, $j \in J$, and thus

$$f_i(x) - f_i(y) \le f_{i,y}(x) - f_i(y) \quad \text{for all } i \in I.$$

We define the approximating functions by

$$\mathscr{U}_i(x,y) := f_{i,y}(x) - f_i(y) \text{ and } \mathscr{H}_j(x,y) := h_{j,y}(x) \quad \text{for all } i \in I, j \in J \text{ and } y \in C.$$

Example 3.4. Let the index sets I and J be respectively partitioned into three subsets denoted by I_1, I_2, I_3 and J_1, J_2, J_3. Assume that the functions f_i and h_j are defined as in Example 3.1 for $i \in I_1$ and $j \in J_1$, as in Example 3.2 for $i \in I_2$ and $j \in J_2$, and as in Example 3.3 for $i \in I_3$ and $j \in J_3$. Then, the associated functions $\mathscr{U}_i$ and $\mathscr{H}_j$ are defined as follows:

$$\mathscr{U}_i(x,y) = \begin{cases} f_i(x) - f_i(y) + \frac{\alpha}{2}\|x-y\|^2 & \text{if } i \in I_1, \\ \langle \nabla f_i(y), x-y\rangle + \frac{L_i}{2}\|x-y\|^2 & \text{if } i \in I_2, \\ f_{i,y}(x) - f_i(y) & \text{if } i \in I_3, \end{cases}$$

$$\mathscr{H}_j(x,y) = \begin{cases} h_j(x) + \frac{\alpha}{2}\|x-y\|^2 & \text{if } j \in J_1, \\ h_j(y) + \langle \nabla h_j(y), x-y \rangle + \frac{K_j}{2}\|x-y\|^2 & \text{if } j \in J_2, \\ h_{j,y}(x) & \text{if } j \in J_3. \end{cases}$$

Now, to approximate the problem (P_y) for $y \in C$, define the function $\mathscr{U}_y$ by

$$\mathscr{U}_y(x) := \max_{i\in I} \mathscr{U}_i(x,y).$$

Instead of the subproblem (P_y), we associate to (P) its approximating problem

$$\inf_{x\in X_y} \mathscr{U}_y(x), \tag{AP_y}$$

and we denote by x_y the global minimum of $\mathscr{U}_y$ over X_y.

Proposition 3.1. *For all $y \in X$, we have*

(1) $\mathscr{U}_y(y) = 0$, $\mathscr{U}_y(x_y) \leq 0$,
(2) $f_i(x_y) \leq f_i(y)$ for all $i \in I$.

Proof. (1) From the hypothesis (A1), $\mathscr{U}_y(y) = 0$. Since x_y is a global minimum for (AP_y), then $\mathscr{U}_y(x_y) \leq \mathscr{U}_y(x)$ for all $x \in X_y$. Since $y \in X_y$, then with $x = y$ we get $\mathscr{U}_y(x_y) \leq \mathscr{U}_y(y) = 0$.

(2) By (A2), we have that $f_i(x) - f_i(y) \leq \mathscr{U}_y(x)$ for all $i \in I$. With $x = x_y$ in the last inequality and by using Item 1, we deduce that $f_i(x_y) - f_i(y) \leq \mathscr{U}_y(x_y) \leq 0$ for all $i \in I$, which implies in turn that $f_i(x_y) \leq f_i(y)$, for all $i \in I$.

□

The next proposition establishes a link between $(AP_{\bar{x}})$ and (P), where $\bar{x}$ is any weak Pareto minimum for (P).

Proposition 3.2. *Let $\bar{x} \in X$ be a weak Pareto minimum for (P). Then $\bar{x}$ is a global minimum for $(AP_{\bar{x}})$. Conversely, every optimal solution $x_{\bar{x}}$ of $(AP_{\bar{x}})$ is also a weak Pareto minimum for (P).*

Proof. Let $\bar{x} \in X$ be a weak Pareto minimum for (P). Note that $\bar{x} \in X_{\bar{x}}$. Suppose that $\bar{x}$ is not a global minimum for $(AP_{\bar{x}})$. Then there exists $z \in X_{\bar{x}} \subset X$ such that

$$\mathscr{U}_{\bar{x}}(z) < \mathscr{U}_{\bar{x}}(\bar{x}) = 0,$$

where the last equality on the right hand follows from Item 1 of Proposition 3.1 with $y = \bar{x}$.

Since $f_i(z) - f_i(\bar{x}) \leq \mathscr{U}_i(z,\bar{x})$ by (A2), and $X_{\bar{x}}$, it follows that $f_i(z) - f_i(\bar{x}) < 0$ for all $i \in I$, which is a contradiction to the fact that $\bar{x}$ is a weak Pareto minimum for (P). Then $\bar{x}$ is a global minimum for $(AP_{\bar{x}})$.

Conversely, we suppose that $x_{\bar{x}} \in X_{\bar{x}}$ is a global minimum for $(AP_{\bar{x}})$. Note that $x_{\bar{x}} \in X$ since $X_{\bar{x}} \subset X$. If $x_{\bar{x}}$ is not a weak Pareto minimum of (P), then there exists $z \in X$ such that $f_i(z) < f_i(x_{\bar{x}})$ for all $i \in I$. From Proposition 3.1, we have $f_i(x_{\bar{x}}) < f_i(\bar{x})$, and thus, $f_i(z) < f_i(\bar{x})$ for all $i \in I$, which contradicts the fact that $\bar{x}$ is a weak Pareto minimum of (P). □

In the following proposition we show that the same results hold for Pareto minimum.

Proposition 3.3. *Assume that problem (P) has a Pareto minimum $\bar{x} \in X$. Then $\bar{x}$ is a global minimum for $(AP_{\bar{x}})$. Conversely, every optimal solution $x_{\bar{x}}$ of $(AP_{\bar{x}})$ is also a Pareto minimum of problem (P).*

Proof. Since a Pareto minimum of (P) is also a weak Pareto minimum, the first implication follows from Proposition 3.2. Conversely, let $\bar{x} \in X$ be a Pareto minimum of (P). Assume, for contradiction, that there is a global minimum $x_{\bar{x}}$ of $(AP_{\bar{x}})$ which is not a Pareto minimum of (P). Then there exists $z \in X$ such that $f_i(z) \leq f_i(x_{\bar{x}})$ for all $i \in I$, and $f_j(z) < f_j(x_{\bar{x}})$ for some $j \in I$. From Proposition 3.1, Item 2, with $y = \bar{x}$, we have $f_i(x_{\bar{x}}) \leq f_i(\bar{x})$, for all $i \in I$. It follows that $f_i(z) \leq f_i(\bar{x})$ for all $i \in I$, and $f_j(z) < f_j(\bar{x})$ for some $j \in I$, which is absurd since $\bar{x}$ is a Pareto minimum of (P). □

4. Optimality Conditions for MOP

This section is devoted to developing optimality conditions for (P) expressed in terms of directional derivatives. We first recall the expression of the derivative of the function defined by

$$\psi(x) := \max_{1 \leq i \leq N} \psi_i(x),$$

where ψ_i are direcetioanlly differentiable. For this we denote by

$$\Sigma_N = \left\{ \mu = (\mu_1, \cdots, \mu_N)^\top \in \mathbb{R}^N \mid \textstyle\sum_{i=1}^N \mu_i = 1,\, \mu_i \geq 0,\, \forall i = 1, \cdots, N \right\}.$$

Lemma 4.1. *The directional derivative of the function ψ at x in the direction $d \in \mathbb{R}^n$ is given by*

$$\psi'(x;d) = \max_{\mu \in \Sigma_N(x)} \left\{ \sum_{i=1}^N \mu_i \psi_i'(x;d) \right\}, \tag{4.1}$$

where $\Sigma_N(x) := \left\{ \mu \in \Sigma_N \mid \sum_{i=1}^N \mu_i \psi_i(x) = \psi(x) \right\}$.

Proof. Note that the function ψ can be written equivalently as follows

$$\psi(x) = \max_{\mu \in \Sigma_N} \left\{ \sum_{i=1}^N \mu_i \psi_i(x) \right\}.$$

Remark that the maximum in the last expression of $\psi(x)$ is achieved at any $\mu \in \Sigma_N$ such that $\sum_{i=1}^N \mu_i \psi_i(x) = \psi(x)$, i.e., for any $\mu \in \Sigma_N(x)$. Thus, by using [27, Theorem 3.2], we get (4.1). □

In order to develop necessary optimality conditions, we will need the following general assumption.

Assumption 4.1. For any $\mu \in \mathbb{R}$ and $y \in C$, we have

$$\inf_{x \in C} \max_{(\alpha,\beta) \in \Sigma_{m+p}} \left[\sum_{i \in I} \alpha_i \left[\mathscr{U}_i(x,y) - \mu \right] + \sum_{j \in J} \beta_j \mathscr{H}_j(x,y) \right] =$$
$$\max_{(\alpha,\beta) \in \Sigma_{m+p}} \inf_{x \in C} \left[\sum_{i \in I} \alpha_i \left[\mathscr{U}_i(x,y) - \mu \right] + \sum_{j \in J} \beta_j \mathscr{H}_j(x,y) \right].$$

Remark 4.1. Assumption 4.1 is verified, for example, when $\mathscr{U}_i(.,y)$ and $\mathscr{H}_j(.,y)$, for $(i,j) \in I \times J$ and $y \in C$ are convex or quasiconvex, see [29] for further examples.

In the theorem below, we present necessary KKT-type optimality conditions, that any weak Pareto minimum of problem (P) must satisfy, formulated using the directional derivatives of the objective and constraint functions.

Theorem 4.1. *Let $\bar{x} \in X$. Assume that for all $(i,j) \in I \times J$, the functions f_i and h_j are directionally differentiable at $\bar{x}$, that the upper bounding functions $\mathscr{U}_i$ and $\mathscr{H}_j$ are satisfy (A1)-(A3), and (B1)-(B3), respectively, and that Assumption 4.1 is verified with $\mu = 0$. If $\bar{x}$ is a weak Pareto minimum for (P), then there exists $(\bar{\alpha}, \bar{\beta}) \in \Sigma_{m+p}$ such that*

$$\sum_{i\in I} \bar{\alpha}_i f_i'(\bar{x}; x-\bar{x}) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x}; x-\bar{x}) \geq 0 \quad \text{for all } x \in C \tag{4.2}$$

with $\bar{\beta}_j h_j(\bar{x}) = 0$ for all $j \in J$. If in addition there is $x \in C$ such that $h'(\bar{x}; x-\bar{x}) < 0$ then there exists $i \in I$ such that $\bar{\alpha}_i \neq 0$.

Proof. Since $\bar{x}$ is a weak Pareto minimum for (P), it follows from Proposition 3.2 that $\bar{x}$ is also a global minimum for $(AP_{\bar{x}})$. Let us define on C, the function

$$\begin{aligned}\hat{F}_{\bar{x}}(x) &:= \max_{(i,j)\in I\times J} \left\{ \mathscr{U}_i(x,\bar{x}) - \mathscr{U}_{\bar{x}}(\bar{x}),\ \mathscr{H}_j(x,\bar{x}) \right\} \\ &= \max_{(i,j)\in I\times J} \left\{ \mathscr{U}_i(x,\bar{x}),\ \mathscr{H}_j(x,\bar{x}) \right\},\end{aligned}$$

where the last equality follows from Item 1 of Proposition 3.1. Obviously, for all $x \in X_{\bar{x}}$, i.e., $\mathscr{H}_j(x,\bar{x}) \leq 0$ for all $j \in J$, we have $f_i(x) - f_i(\bar{x}) \geq 0$ for all $i \in I$ since $\bar{x}$ is a weak Pareto minimum. By (A2) we get $\mathscr{U}_i(x,\bar{x}) \geq 0$ for all $i \in I$, and the definition of $\mathscr{U}_{\bar{x}}(x)$ entails that $\mathscr{U}_{\bar{x}}(x) \geq 0$. This implies that $\hat{F}_{\bar{x}}(x) \geq 0$,for all $x \in X_{\bar{x}}$. For $x \in C \backslash X_{\bar{x}}$ it holds that $\mathscr{H}_j(x,\bar{x}) > 0$ for some $j \in J$ and this again implies that $\hat{F}_{\bar{x}}(x) \geq 0$. In conclusion $\hat{F}_{\bar{x}}(x) \geq 0$ for all $x \in C$. On the other hand, the function $\hat{F}_{\bar{x}}$ may be expressed by

$$\hat{F}_{\bar{x}}(x) = \max_{(\alpha,\beta)\in\Sigma_{m+p}} \left[\sum_{i\in I} \alpha_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,\bar{x}) \right].$$

Since Assumption 4.1 is fulfilled, with $\mu = 0$, we get

$$\begin{aligned}\inf_{x\in C} \hat{F}_{\bar{x}}(x) &= \inf_{x\in C} \max_{(\alpha,\beta)\in\Sigma_{m+p}} \left[\sum_{i\in I} \alpha_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,\bar{x}) \right] \\ &= \max_{(\alpha,\beta)\in\Sigma_{m+p}} \inf_{x\in C} \left[\sum_{i\in I} \alpha_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,\bar{x}) \right].\end{aligned} \tag{4.3}$$

Note that the function

$$\Sigma_{m+p} \ni (\alpha,\beta) \mapsto \inf_{x\in C} \left[\sum_{i\in I} \alpha_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,\bar{x}) \right]$$

is upper semicontinuous on Σ_{m+p}, since it is the pointwise infinimum of a family of linear functions, a fortiori continuous, and then achieves its maximum on the compact set Σ_{m+p}. Therefore, there exists $(\bar{\alpha}, \bar{\beta}) \in \Sigma_{m+p}$ such that

$$\begin{aligned}&\max_{(\alpha,\beta)\in\Sigma_{m+p}} \inf_{x\in C} \left[\sum_{i\in I} \alpha_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,\bar{x}) \right] \\ &\qquad = \inf_{x\in C} \left[\sum_{i\in I} \bar{\alpha}_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J} \bar{\beta}_j \mathscr{H}_j(x,\bar{x}) \right].\end{aligned} \tag{4.4}$$

Since we showed that $\hat{F}_{\bar{x}}(x) \geq 0$ for all $x \in C$, then from (4.3) and (4.4) we obtain

$$\sum_{i\in I} \bar{\alpha}_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J} \bar{\beta}_j \mathscr{H}_j(x,\bar{x}) \geq 0 \quad \text{for all } x \in C. \tag{4.5}$$

Using the assumptions (A1) and (B1) in (4.5) with $x = \bar{x}$, we get

$$\sum_{i\in I} \bar{\alpha}_i \mathscr{U}_i(\bar{x},\bar{x}) + \sum_{j\in J} \bar{\beta}_j \mathscr{H}_j(\bar{x},\bar{x}) = \sum_{j\in J} \bar{\beta}_j \mathscr{H}_j(\bar{x},\bar{x}) = \sum_{j\in J} \bar{\beta}_j h_j(\bar{x}) \geq 0.$$

But $h_j(\bar{x}) \leq 0$ for all $j \in J$, and thus, $\bar{\beta}_j \mathscr{H}_j(\bar{x},\bar{x}) = \bar{\beta}_j h_j(\bar{x}) = 0$ for all $j \in J$. This together with (4.5) means that $\bar{x}$ minimizes the function

$$x \mapsto \sum_{i\in I} \bar{\alpha}_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J} \bar{\beta}_j \mathscr{H}_j(x,\bar{x})$$

on C. Taking into account that C is convex, it follows that

$$\sum_{i\in I} \bar{\alpha}_i \mathscr{U}_i'((z,\bar{x});x-\bar{x})|_{z=\bar{x}} + \sum_{i\in J} \bar{\beta}_j \mathscr{H}_j'((z,\bar{x});x-\bar{x})|_{z=\bar{x}} \geq 0 \quad \text{for all } x \in C,$$

where $\mathscr{U}_i'((z,\bar{x});x-\bar{x})|_{z=\bar{x}}$ and $\mathscr{H}_j'((z,\bar{x});x-\bar{x})\Big|_{z=\bar{x}}$ are respectively the directional derivatives of $\mathscr{U}_i(\cdot,\bar{x})$ and $\mathscr{H}_i(\cdot,\bar{x})$ at $\bar{x}$ in the direction $x-\bar{x}$. By using (A3) and (B3), we get

$$\sum_{i\in I} \bar{\alpha}_i f_i'(\bar{x};x-\bar{x}) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x};x-\bar{x}) \geq 0 \quad \text{for all } x \in C \tag{4.6}$$

Now, if for all $i \in I$, $\bar{\alpha}_i = 0$, then $\sum_{j\in J} \bar{\beta}_j = 1$ and the fact that $\bar{\beta}_j h_j(\bar{x}) = 0$ for all $j \in J$ implies that $h(\bar{x}) = 0$. On the other hand (4.6) entails that $h'(\bar{x};x-\bar{x}) \geq 0$ for all $x \in C$, contradicting the last assumption of the theorem. □

Remark 4.2. (1) Note that if Assumption 4.1 is fulfilled with the choices $\mathscr{U}_i(x,y) = f_i(x) - f_i(y)$ and $\mathscr{H}_j(x,y) = h_j(x)$ then Theorem 4.1 remains valid since these functions successively satisfy (A1)-(A3) and (B1)-(B3).
(2) Theorem 4.1 also remains valid if the functions f_i and h_j are directionally differentiable and the derivative functions are convex as functions of directions.

In the following we will examine sufficient optimality conditions, that is, conditions under what a point satisfying the necessary optimality conditions stated in Theorem 4.1 is an optimal solution for (P). First, let us recall the notions of quasiconvexity and pseudoconvexity.

Definition 4.1. Let g be a real-valued function directionally differentiable on a set $D \subseteq \mathbb{R}^n$. We say that g is

(1) quasiconvex, if for all $x, y \in D$

$$g(y) \leq g(x) \quad \text{implies} \quad g'(x;y-x) \leq 0;$$

(2) pseudoconvex, if for all $x, y \in D$

$$g(y) < g(x) \quad \text{implies} \quad g'(x;y-x) < 0,$$

or equivaently $g'(x;y-x) \geq 0 \quad \text{implies} \quad g(y) \geq g(x)$.

Now, we are ready to state sufficient optimality conditions for (P).

Theorem 4.2. *Assume that a point $\bar{x} \in X$ satisfies the necessary optimality conditions stated in Theorem 4.1. Assume on the other hand that there exists $\hat{x} \in C$ such that $h'(\bar{x};\hat{x}-\bar{x}) < 0$. If f_i is pseudoconvex for all $i \in I$ and h_j is quasiconvex for all $j \in J$, then $\bar{x}$ is a weak Pareto optimal solution for (P).*

Proof. To show that the converse of Theorem 4.1 is true, assume that (4.2) holds. Then there exists $(\bar{\alpha},\bar{\beta}) \in \Sigma_{m+p}$ such that

$$\sum_{i\in I} \bar{\alpha}_i f_i'(\bar{x};x-\bar{x}) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x};x-\bar{x}) \geq 0 \quad \text{for all } x \in C \tag{4.7}$$

with $\bar{\beta}_j h_j(\bar{x}) = 0$ for all $j \in J$.
Since $\bar{\beta}_j h_j(x) \leq \bar{\beta}_j h_j(\bar{x})$ for all $x \in X$ and $j \in J$, the quasiconvexity of the functions h_j implies that

$$\bar{\beta}_j h_j'(\bar{x};x-\bar{x}) \leq 0 \quad \text{for all } x \in X \text{ and } j \in J.$$

By referring to (4.7) we conclude that

$$\sum_{i\in I} \bar{\alpha}_i f_i'(\bar{x};x-\bar{x}) \geq 0 \quad \text{for all } x \in C. \tag{4.8}$$

Assume that $\bar{x}$ is not a weak Pareto minimum for (P). Then there exists $\hat{x} \in X$ such that $f_i(\hat{x}) < f_i(\bar{x})$ for all $i \in I$. Let $x = \hat{x}$ in (4.8). Then, taking into account that $\bar{\alpha}_i \neq 0$ for some $i \in I$ we conclude that $\sum_{i\in I} \bar{\alpha}_i > 0$ and then $f_{i_0}'(\bar{x};\hat{x}-\bar{x}) \geq 0$ for some $i_0 \in I$. Therefore the pseudconvexity of f_{i_0} implies that $f_{i_0}(\hat{x}) \geq f_{i_0}(\bar{x})$ contradicting our previous assumption. Thereby $\bar{x}$ is a weak Pareto minimum for (P). □

Since the optimal solution $\bar{x}$ is unknown, the previous results suggest us to approximate it by some x^k, at an iteration k, and approximating the function $\mathscr{U}_{\bar{x}}$, iteratively, at each step k by a function $\mathscr{U}_k := \mathscr{U}_{x^k}$, where x^k is a global minimum of the function $\mathscr{U}_{k-1} := \mathscr{U}_{x^{k-1}}$ on the set $X_{k-1} := X_{x^{k-1}}$. The next algorithm describes this sequential upper bounding approximation method.

5. Sequential Upper Bounding Algorithm for MOP

In this section we introduce the general sequential upper bounding algorithm to solve problem (P).

Algorithm 1 Sequential Upper Bounding Algorithm

0. Choose $x^0 \in X$ and let $k = 0$.
1. Find $x^{k+1} \in X_k$ such that $\mathscr{U}_k(x^{k+1}) \leq \mathscr{U}_k(x)$ for all $x \in X_k$.
2. If $\mathscr{U}_k(x^{k+1}) = 0$, stop. Else set $k = k+1$ and return to 1.

We will establish the convergence of the sequences $\{f_i(x^k)\}$ for $i \in I$.

Proposition 5.1. *For all $i \in I$, the sequence $\{f_i(x^k)\}$ is nonincreasing and converges to some $\hat{f}_i$. In addition, if $\hat{f}_i = -\infty$ for all $i \in I$, then the problem (P) has no weak Pareto optimal solution, otherwise the sequences $\{F_k(x^{k+1})\}$ and $\{\mathscr{U}_k(x^{k+1})\}$ converge to 0 as k tends to ∞.*

Proof. Recall that we used the notation $\mathscr{U}_k = \mathscr{U}_{x^k}$ and $X_k = X_{x^k}$ and that x^{k+1} is defined to be a solution of (AP_k). It follows that

$$\mathscr{U}_k(x^{k+1}) \leq \mathscr{U}_k(x) \quad \text{for all } x \in X_k.$$

In particular with $x = x^k$ we get $\mathscr{U}_k(x^{k+1}) \leq \mathscr{U}_k(x^k) = 0$, where the equality $\mathscr{U}_k(x^k) = 0$ follows from Proposition 3.1, Item 1. From (A2), for all $i \in I$ and all $x \in C$, we have

$$\mathscr{U}_k(x) \geq f_i(x) - f_i(x^k).$$

Therefore,

$$0 \geq \mathscr{U}_k(x^{k+1}) \geq f_i(x^{k+1}) - f_i(x^k).$$

This implies that $f_i(x^{k+1}) \leq f_i(x^k)$ for all $i \in I$. It follows that the sequence $\{f_i(x^k)\}$ converges to some $\hat{f}_i \in \mathbb{R} \cup \{-\infty\}$ for all $i \in I$.

Clearly, if $\hat{f}_i = -\infty$ for each $i \in I$, then for all $x \in X$ there exists $\hat{k} \in \mathbb{N}$ such that $f_i(x^k) < f_i(x)$ for all $k \geq \hat{k}$ and $i \in I$, so that the problem (P) cannot have weak Pareto optimal solution.

Assume now that $\hat{f}_i > -\infty$ for some $i \in I$. From the hypothesis (A2), the definition of F_k and the fact that $\mathscr{U}_k(x^{k+1}) \leq \mathscr{U}_k(x^k) = 0$, we have

$$0 \geq \mathscr{U}_k(x^{k+1}) \geq F_k(x^{k+1}) \geq f_i(x^{k+1}) - f_i(x^k). \tag{5.1}$$

On the other hand, $f_i(x^{k+1}) - f_i(x^k) \to 0$, when $k \to \infty$ and (5.1) entails that

$$\mathscr{U}_k(x^{k+1}) \to 0 \text{ and } F_k(x^{k+1}) \to 0 \quad \text{when} \quad k \to \infty. \tag{5.2}$$

This is the desired result. □

Proposition 5.2. *Let Assumption 4.1 be fulfilled. Then, for all $k \in \mathbb{N}$, there exists $(\alpha^k, \beta^k) \in \Sigma_{m+p}$, such that*

$$\sum_{i\in I} \alpha_i^k \left(\mathscr{U}_i(x,x^k) - \mathscr{U}_k(x^{k+1})\right) + \sum_{j\in J} \beta_j^k \mathscr{H}_j(x,x^k) \geq 0 \quad \textit{for all } x \in C. \tag{5.3}$$

Moreover,

$$\sum_{j\in J} \beta_j^k \mathscr{H}_j(x^{k+1},x^k) = 0 \textit{ and } 0 \geq \sum_{j\in J} \beta_j^k \mathscr{H}_j(x^k,x^k) = \sum_{j\in J} \beta_j^k h_j(x^k) \geq \sum_{i\in I} \alpha_i^k \mathscr{U}_i(x^{k+1},x^k).$$

So, the sequence $\left\{\sum_{j\in J} \beta_j^k h_j(x^k)\right\}$ converges to 0 as k tends to ∞.

Proof. Recall that from the definition of x^{k+1} we have

$$\mathscr{U}_k(x) - \mathscr{U}_k(x^{k+1}) \geq 0 \quad \text{for all } x \in X_k, \tag{5.4}$$

where $X_k := \{x \in C \mid \mathscr{H}_j(x,x^k) \leq 0, \forall j \in J\}$. Define the function

$$\mathscr{H}(x,x^k) = \max_{j\in J} \mathscr{H}_j(x,x^k).$$

Obviously, $\mathscr{H}(x,x^k) \leq 0$ if and only if $\mathscr{H}_j(x,x^k) \leq 0$ for all $j \in J$. Define also the function

$$\mathscr{F}_k(x) = \max\left[\mathscr{U}_k(x) - \mathscr{U}_k(x^{k+1}), \mathscr{H}(x,x^k)\right].$$

It is straightforward to show that $\mathscr{F}_k(x) \geq 0$ for all $x \in C$. Indeed, if $x \in X_k$, i.e. $x \in C$ and $\mathscr{H}(x,x^k) \leq 0$, then (5.4) entails that $\mathscr{F}_k(x) \geq 0$, and if $x \in C$ but $x \notin X_k$, i.e. $\mathscr{H}(x,x^k) > 0$ then we have $\mathscr{F}_k(x) > 0$. On the other hand, the function $\mathscr{F}_k$ may be expressed by

$$\mathscr{F}_k(x) = \max_{(\alpha,\beta)\in\Sigma_{m+p}} \left[\sum_{i\in I} \alpha_i \left(\mathscr{U}_i(x,x^k) - \mathscr{U}_k(x^{k+1})\right) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,x^k)\right].$$

Since Assumption 4.1 is fulfilled, with $\mu = \mathscr{U}_k(x^{k+1})$, we get

$$\begin{aligned}\inf_{x\in C} \mathscr{F}_k(x) &= \inf_{x\in C} \max_{(\alpha,\beta)\in\Sigma} \left[\sum_{i\in I} \alpha_i \left(\mathscr{U}_i(x,x^k) - \mathscr{U}_k(x^{k+1})\right) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,x^k)\right] \\ &= \max_{(\alpha,\beta)\in\Sigma} \inf_{x\in C} \left[\sum_{i\in I} \alpha_i \left(\mathscr{U}_i(x,x^k) - \mathscr{U}_k(x^{k+1})\right) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,x^k)\right]. \qquad (5.5)\end{aligned}$$

The function

$$\Sigma_{m+p} \ni (\alpha,\beta) \mapsto \inf_{x\in C} \left[\sum_{i\in I} \alpha_i \left(\mathscr{U}_i(x,x^k) - \mathscr{U}_k(x^{k+1})\right) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,x^k)\right]$$

is upper semicontinuous on Σ_{m+p}, since it is the pointwise infinimum of a family of linear functions, a fortiori continuous, and then achieves its maximum on the compact set Σ_{m+p}. Therefore, for all $k \in \mathbb{N}$ there exists $(\alpha^k, \beta^k) \in \Sigma_{m+p}$ such that

$$\begin{aligned}&\max_{(\alpha,\beta)\in\Sigma_{m+p}} \inf_{x\in C} \left[\sum_{i\in I} \alpha_i \left(\mathscr{U}_i(x,x^k) - \mathscr{U}_k(x^{k+1})\right) + \sum_{j\in J} \beta_j \mathscr{H}_j(x,x^k)\right] \\ &\quad = \inf_{x\in C} \left[\sum_{i\in I} \alpha_i^k \left(\mathscr{U}_i(x,x^k) - \mathscr{U}_k(x^{k+1})\right) + \sum_{j\in J} \beta_j^k \mathscr{H}_j(x,x^k)\right]. \qquad (5.6)\end{aligned}$$

Since we showed that $\mathscr{F}_k(x) \geq 0$ for all $x \in C$, then from (5.6) we obtain

$$\sum_{i\in I} \alpha_i^k \left(\mathscr{U}_i(x,x^k) - \mathscr{U}_k(x^{k+1})\right) + \sum_{j\in J} \beta_j^k \mathscr{H}_j(x,x^k) \geq 0 \quad \text{for all } x \in C. \qquad (5.7)$$

The rest follows by letting in the previous inequality and once $x = x^{k+1}$ once $x = x^k$. □

Before analyzing the convergence of Algorithm 1 in the case it generates an infinite sequence, we will see what happens in the case where the generated sequence is finite.

Theorem 5.1. *If for some $k \in \mathbb{N}$, we have $\mathscr{U}_k(x^{k+1}) = 0$, then x^k verifies the optimality conditions stated in Theorem 4.1.*

Proof. Suppose that Algorithm 1 stops at x^{k+1}, that is, $\mathscr{U}_k(x^{k+1}) = 0$. By invoking Proposition 5.2, there exist $(\alpha^k, \beta^k) \in \Sigma_{m+p}$, such that

$$\sum_{i\in I} \alpha_i^k \left(\mathscr{U}_i(x,x^k) - \mathscr{U}_k(x^{k+1})\right) + \sum_{j\in J} \beta_j^k \mathscr{H}_j(x,x^k) \geq 0 \quad \text{for all } x \in C,$$

and we conclude from the last inequalities therein with $x = x^k$ that

$$\sum_{j\in J} \beta_j^k h_j(x^k) = 0.$$

On the other hand, since $\mathscr{U}_k(x^{k+1}) = 0$, then the last inequality implies that

$$\sum_{i\in I}\alpha_i^k \mathscr{U}_i(x,x^k) + \sum_{j\in J}\beta_j^k \mathscr{H}_j(x,x^k) \geq 0 \quad \text{for all } x\in C.$$

Since the left hand side of the last inequality vanishes at x^k and C is convex, then by passing to the directional derivative, we get

$$\sum_{i\in I}\alpha_i^k \mathscr{U}_i'((z,x^k);x-x^k)|_{z=x^k} + \sum_{j\in J}\beta_j^k \mathscr{H}_j'((z,x^k);x-x^k)|_{z=x^k} \geq 0 \quad \text{for all } x\in C.$$

By using assumptions (A3) and (B3), we conclude that

$$\sum_{i\in I}\alpha_i^k f_i'(x^k;x-x^k) + \sum_{j\in J}\beta_j^k h_j'(x^k;x-x^k) \geq 0 \quad \text{for all } x\in C. \tag{5.8}$$

Noting that $\beta_j^k h_j(x^k) = 0$ for all $j\in J$, we conclude that x^k verifies the necessary optimality conditions stated in Theorem 4.1. □

The next theorem presents the main result regarding the convergence of the sequence $\{x^k\}$ generated by Algorithm 1.

Theorem 5.2. *Assume that the functions f_i and h_j for $i\in I$ and $j\in J$, are directionally differentiable, and that the upper bounding functions $\mathscr{U}_i$ and $\mathscr{H}_j$ for all $i\in I$ and $j\in J$, satisfy assumptions (A1)-(A3) and (B1)-(B3) respectively, and are continuous and that Assumption 4.1 is fulfilled. Then, for every cluster point $\bar{x}$ of the sequence $\{x^k\}$ there exists $(\bar{\alpha},\bar{\beta})\in\Sigma_{m+p}$ such that*

$$\sum_{i\in I}\bar{\alpha}_i f_i'(\bar{x};x-\bar{x}) + \sum_{j\in J}\bar{\beta}_j h_j'(\bar{x};x-\bar{x}) \geq 0 \quad \textit{for all } x\in C$$

with $\bar{\beta}_j h_j(\bar{x}) = 0$ for all $j\in J$. If in addition there is $x\in C$ such that $h'(\bar{x};x-\bar{x})<0$ then there exists $i\in I$ such $\bar{\alpha}_i\neq 0$.

Proof. Let $\{x^k\}$ be the sequence generated by Algorithm 1 and $\{(\alpha^k,\beta^k)\}$ as in (5.3). Let $K\subset\mathbb{N}$, an infinite index set, such that the sequences $\{x^k\}$ and $\{(\alpha^k,\beta^k)\}$ such that converge respectively to limits points $\bar{x}$ and $\{(\bar{\alpha},\bar{\beta})\}$. Then by passing to limit in (5.3) and taking into account that $\mathscr{U}_i$ and $\mathscr{H}_j$ are continuous, and that $\{\mathscr{U}_k(x^{k+1})\}$ converges to 0, we get

$$\sum_{i\in I}\bar{\alpha}_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J}\bar{\beta}_j \mathscr{H}_j(x,\bar{x}) \geq 0 \quad \text{for all } x\in C. \tag{5.9}$$

With $x=\bar{x}$ and by invoking the assumptions (A1) and (B1), we get

$$\sum_{i\in I}\bar{\alpha}_i \mathscr{U}_i(\bar{x},\bar{x}) + \sum_{j\in J}\bar{\beta}_j \mathscr{H}_j(\bar{x},\bar{x}) = \sum_{j\in J}\bar{\beta}_j h_j(\bar{x}) \geq 0.$$

But since $\bar{x}\in X_{\bar{x}}\subset X$ we have $h_j(\bar{x})\leq 0$ for all $j\in J$ which implies, taking into account the last inequality, that $\bar{\beta}_j h_j(\bar{x}) = 0$. Returning to (5.9) we conclude that $\bar{x}$ minimizes the function

$$x\mapsto \sum_{i\in I}\bar{\alpha}_i \mathscr{U}_i(x,\bar{x}) + \sum_{j\in J}\bar{\beta}_j \mathscr{H}_j(x,\bar{x})$$

on C, and taking into account that C is convex, we get

$$\sum_{i\in I}\bar{\alpha}_i \mathscr{U}_i'((z,\bar{x});x-\bar{x})|_{z=\bar{x}} + \sum_{i\in J}\bar{\beta}_j \mathscr{H}_j'((z,\bar{x});x-\bar{x})|_{z=\bar{x}} \geq 0 \quad \text{for all } x\in C.$$

By using (A3) and (B3), we get the desired result. □

Remark 5.1. If for some $i \in I$ the set $\{x \in C \mid f_i(x) \le f_i(x^0)\}$ is bounded (which is the case, e.g., if f_i is inf-compact, then the sequence $\{x^k\}$ is also bounded. Indeed, we established in the previous proof that $f_i(x^{k+1}) \le f_i(x^k)$ for all $i \in I$. Consequently, $f_i(x^{k+1}) \le f_i(x^0)$ for each $i \in I$, which implies that $x^{k+1} \in \{x \in C \mid f_i(x) \le f_i(x^0)\}$.

6. Numerical Tests

This section presents numerical experiments to evaluate the performance of the proposed method.

We recall that our goal is to solve the following multiobjective optimization problem:

$$\inf_{x \in X} [f(x) := (f_1(x), f_2(x), \dots, f_m(x))], \tag{P}$$

where

$$X = \left\{x \in C \mid h_j(x) \le 0, \forall j \in J\right\}$$

and $C \subset \mathbb{R}^n$ a nonempty, closed, convex set.

We introduce the following index sets:

$$I_1 := \{1, 2, \dots, m_1\}, \quad I_2 := \{m_1+1, m_1+2, \dots, m\}, \quad I := I_1 \cup I_2,$$

$$J_1 := \{1, 2, \dots, p_1\}, \quad J_2 := \{p_1+1, p_1+2, \dots, p\}, \quad J := J_1 \cup J_2.$$

We set $m_2 = m - m_1$ and $p_2 = p - p_1$.

We test Algorithm 1 on three types of multiobjective programming problems.

- The first type (Type 1) corresponds to the case where all functions f_i and h_j, for $(i, j) \in I_1 \times J_1$, are continuously differentiable with Lipschitz-continuous gradients, i.e. $I_2 = J_2 = \emptyset$;
- The second type (Type 2) corresponds to the case where the functions f_i, and h_j, for $(i, j) \in I_2 \times J_2$, are DC functions, i.e. $I_1 = J_1 = \emptyset$. In this case, the functions f_i, and h_j, are defined as follows: $f_i(x) = f_i^1(x) - f_i^2(x)$ and $h_j(x) = h_j^1(x) - h_j^2(x)$, for $(i, j) \in I_2 \times J_2$;
- The third type (Type 3) corresponds to the case where the functions f_i, and h_j, for $(i, j) \in I_1 \times J_1$, are continuously differentiable with Lipschitz-continuous gradients, while those for $(i, j) \in I_2 \times J_2$ are DC functions. In this case, the functions f_i, and h_j, for $(i, j) \in I_2 \times J_2$, are given by: $f_i(x) = f_i^1(x) - f_i^2(x)$ and $h_j(x) = h_j^1(x) - h_j^2(x)$, for $(i, j) \in I_2 \times J_2$.

6.1. **Problem statement.** We apply our algorithm to problem (P), where the vector-valued function f is specified through the component functions f_i, for $i = 1, \dots, m$:

$$f_i(x) = \tfrac{1}{2} x^\top P_i x + a_i^\top x + b_i, \text{ for } i \in I_1,$$

and

$$f_i^\ell(x) = \tfrac{1}{2} x^\top P_i^\ell x + a_{i,\ell}^\top x + b_i^\ell, \text{ for } i \in I_2,\ \ell = 1, 2.$$

The feasible set X is given by

$$X = \left\{x \in C \mid h_j(x) \le 0,\ j \in J\right\},$$

where

$$C = \{x \in \mathbb{R}^n \mid 0 \le x_i \le 10,\ i = 1, \dots, n\},$$

and the constraint functions h_j are defined by

$$h_j(x) = \tfrac{1}{2}x^\top C_j x + \alpha_j^\top x + \beta_j, \text{ for } j \in J_1,$$

and

$$h_j^\ell(x) = \tfrac{1}{2}x^\top C_j^\ell x + \alpha_{j,\ell}^\top x + \beta_j^\ell, \text{ for } j \in J_2,\ \ell = 1,2.$$

The data P_i, a_i and b_i for $i \in I_1$, and P_i^ℓ, $a_{i,\ell}$ and b_i^ℓ for $i \in I_2$ with $\ell = 1,2$. C_j, α_j and β_j for $j \in J_1$, and C_j^ℓ, $\alpha_{j,\ell}$ and β_j^ℓ for $j \in J_2$ with $\ell = 1,2$ are constructed as follows:

- the matrices P_i^ℓ for $i \in I_2$ and $\ell = 1,2$ (resp. P_i for $i \in I_1$), are defined by $P_i^\ell := L_{i,\ell}^\top L_{i,\ell}$ (resp. $P_i := L_i^\top L_i$), where $L_{i,\ell}$ (resp. L_i) are $1 \times n$ matrices with components uniformly drawn from $[-1,1]$;
- each element of the vectors $a_{i,\ell}$ for $i \in I_2$ and $\ell = 1,2$ (resp. a_i for $i \in I_1$), is uniformly drawn from $[0,10]$. Similarly, the components of the vectors b_i^ℓ for $i \in I_2$ and $\ell = 1,2$, are uniformly drawn from $[0,1]$;
- the elements b_i for $i \in I_1$, are uniformly drawn from $[80,100]$;
- the matrices C_j^ℓ for $j \in J_2$ and $\ell = 1,2$ (resp. C_j for $j \in J_1$), are defined by $C_j^\ell := N_{j,\ell}^\top N_{j,\ell}$ (resp. $C_j := N_j^\top N_j$), where $N_{j,\ell}$ (resp. N_j) are $1 \times n$ matrices with components uniformly drawn from $[-1,1]$;
- each element of the vectors $\alpha_{j,1}$ for $j \in J_2$ (resp. α_j for $j \in J_1$), is uniformly drawn from $[-1,0]$. The elements $\alpha_{j,2}$ for $j \in J_2$ are uniformly drawn from $[0,1]$;
- the elements β_j^1 for $j \in J_2$ (resp. β_j for $j \in J_1$), are uniformly drawn from $[-1,0]$. The elements β_j^2 for $j \in J_2$ are uniformly drawn from $[0,1]$.

For Type 3, we define the same data, except that the vectors $a_{i,\ell}$ and β_j are constructed as follows:

- the vectors $a_{i,\ell}$ for $i \in I_2$ and $\ell = 1,2$, is uniformly drawn from $[1,10]$. The elements β_j for $j \in J_1$ are uniformly drawn from $[0,1]$.

6.2. **Algorithm implementation and numerical results.** Recall that Algorithm 1 involves solving the following approximating problem associated with (P):

$$\inf_{x\in C} \left\{ \mathscr{U}_k(x) := \max_{i\in I} \mathscr{U}_i(x,x^k) \right\}$$
$$\text{s.t. } \mathscr{H}_j(x,x^k) \le 0, \quad j \in J,$$

where $\mathscr{U}_i(x,x^k)$ and $\mathscr{H}_j(x,x^k)$ are defined in Examples 3.2 and 3.3 respectively for Type 1 and Type 2 problems.

Algorithm 1 was implemented in MATLAB and tested on randomly generated problems, using ten random starting points $x^0 \in [0,1]^n$ for each problem. We report the minimum and maximum values of the objective components of the solution obtained by the algorithm, as well as the value of the function h, the number of iterations, and the total execution time in seconds. These quantities are denoted by $f_{\min}$, $f_{\max}$, h_∞, #Iter, and Time, respectively, in the following tables.

The stopping criterion for our algorithm is $|\mathscr{U}_k(x^{k+1})| \le 1.e-6$.

TABLE 1. Type 1 problems with $n = 50,\ m_1 = 10$ and $p_1 = 20$.

x^0	f_{min}	f_{max}	h_∞	#Iter	Time
1	82.73	96.47	1.18e-12	15	1.00
2	83.98	96.57	1.53e-12	76	2.15
3	82.91	97.23	1.69e-12	27	1.43
4	83.55	96.74	5.38e-13	48	1.90
5	83.25	97.05	-2.29e-14	56	2.02
6	83.68	96.60	8.44e-07	29	1.31
7	84.15	96.47	4.45e-13	29	1.42
8	82.80	96.43	3.89e-13	40	2.18
9	83.31	96.77	4.74e-12	10	1.18
10	84.04	97.02	9.39e-07	21	1.46

TABLE 2. Type 2 problems with $n = 50,\ m_2 = 10$ and $p_2 = 20$.

x^0	f_{min}	f_{max}	h_∞	#Iter	Time
1	-2398.60	-482.19	1.79e-08	6	8.47
2	-2414.40	-403.4	2.52e-08	10	0.39
3	-1890.90	-424.11	2.58e-08	7	0.50
4	-2605.60	-441.22	1.75e-08	19	27.48
5	-2335.30	-392.61	7.60e-08	11	13.22
6	-2458.00	-402.34	-9.21e-09	10	12.58
7	-1810.60	-419.45	1.60e-08	7	0.28
8	-2438.70	-548.62	2.17e-08	13	8.32
9	-2069.40	-477.53	1.30e-08	8	4.16
10	-1791.10	-359.88	1.42e-08	12	0.46

TABLE 3. Type 3 problems with $n = 50,\ m_1 = m_2 = 10$ and $p_1 = p_2 = 20$.

x^0	f_{min}	f_{max}	h_∞	#Iter	Time
1	-50.75	199.06	6.42e-12	68	9.86
2	-43.60	187.24	1.21e-12	44	3.64
3	-42.10	190.21	4.74e-07	41	22.16
4	-51.45	196.91	2.66e-14	84	6.51
5	-37.07	199.6	7.20e-11	118	12.30
6	-31.69	179.52	2.55e-08	90	9.66
7	-36.60	187.14	3.37e-12	31	6.84
8	-41.19	206.55	-1.24e-06	64	3.82
9	-33.43	194.28	-6.38e-12	182	15.12
10	-38.46	197.78	6.03e-09	110	33.58

6.3. **Comments on the numerical results.** With the results of the tests realized for these three types of problems, and reported in Tables 1 to 3, we can note that:

- In Table 1, since our data satisfy the conditions of Theorem 4.2, namely the pseudoconvexity of the objective functions and the quasiconvexity of the constraints, the points obtained are weak Pareto minima of the generated problem.
- Changing the starting point x^0 affects the solution and allows us to obtain different solutions;
- The number of iterations and the total execution time required to reach the solutions are quite reasonable;
- By examining the column of h_∞, we observe that the solutions found lie on the boundary of the feasible set.

7. Conclusion

In this paper, we have developed necessary and sufficient optimality conditions for constrained multiobjective optimization problems (MOP) of the KKT-type. Based on the improvement function, which transforms the original problem into a scalar optimization constrained program, and the approximation of each function defining our problem by a simpler one, we successfully developed an efficient algorithm to solve a wide range of MOPs, including those involving convex function, DC functions and continuously differentiable functions with Lipschitz gradients. The convergence of the algorithm is analyzed, and we show that every cluster point of the sequence generated by the algorithm satisfies KKT-type necessary optimality conditions expressed in terms of directional derivatives. Numerical tests showed good performance of the algorithm for different types of problems.

Acknowledgements

The authors are grateful to the reviewers for their useful suggestions, which improved the content of this paper.

References

[1] A. Auslender, R. Shefi, M. Teboulle, A moving balls approximation method for a class of smooth constrained minimization problems, SIAM J. Optim. 29 (2010), 3232–3259.

[2] A. Beck, A. Ben-Tal, L. Tetruashvili, A sequential parametric convex approximation method with applications to nonconvex truss topology design problems, J. Glob. Optim. 47 (2010), 29–51.

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

[4] D.P. Bertsekas, Nonlinear Programming, 2nd ed., Athena Scientific, Belmont, MA, 1999.

[5] H. Bonnel, A.N. Iusem, B.F. Svaiter, Proximal methods in vector optimization, SIAM J. Optim. 15 (2005), 953–970.

[6] A.S. Brito, J.X.C. Neto, P.S.M. Santos, S.S. Souza, A relaxed projection method for solving multiobjective optimization problems, Eur. J. Oper. Res. 256 (2017), 17–23.

[7] L.C. Ceng, J.C. Yao, Approximate proximal methods in vector optimization, Eur. J. Oper. Res. 183 (2007), 1–19.

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

[9] P. De, J.B. Ghosh, C.E. Wells, On the minimization of completion time variance with a bicriteria extension, Oper. Res. 40 (1992), 1148–1155.

[10] H.A. Eschenauer, J. Koski, A. Osyczka, Multicriteria optimization fundamentals and motivation, In: Multicriteria Design Optimization: Procedures and Applications, pp. 1–32, Springer, Berlin, 1990.
[11] J. Fliege, OLAF—A general modeling system to evaluate and optimize the location of an air polluting facility, OR Spectrum 23 (2001), 117–136.
[12] Y. Fu, U.M. Diwekar, An efficient sampling approach to multiobjective optimization, Ann. Oper. Res 132 (2004), 109–134.
[13] Y. Gao, X. Yang, K.L. Teo, Optimality conditions for approximate solutions of vector optimization problems, J. Ind. Manag. Optim. 7 (2011), 483–500.
[14] X.L. Guo, S.J. Li, Optimality conditions for vector optimization problems with difference of convex maps, J. Optim. Theory Appl. 162 (2014), 821–844.
[15] C. Gutiérrez, B. Jiménez, V. Novo, On second-order Fritz John type optimality conditions in nonsmooth multiobjective programming, Math. Program. 123 (2010), 199–223.
[16] N. Hoseini Monjezi, S. Nobakhtian, An inexact multiple proximal bundle algorithm for nonsmooth nonconvex multiobjective optimization problems, Ann. Oper. Res. 311 (2020), 1123–1154.
[17] V.I. Ivanov, Second- and first-order optimality conditions in vector optimization, Int. J. Inf. Technol. Decis. Mak. 14 (2015), 747–767.
[18] P. Kanniappan, Necessary conditions for optimality of nondifferentiable convex multiobjective programming, J. Optim. Theory Appl. 40 (1983), 167–174.
[19] K. Kumar, D. Ghosh, A. Upadhayay, J.C. Yao, X. Zhao, Quasi-Newton methods for multiobjective optimization problems: A systematic review, Appl. Set-Valued Anal. Optim. 5 (2023), 291-321.
[20] C.G. Liu, K.F. Ng, W.H. Yang, Merit functions in vector optimization, Math. Program. 119 (2008), 215–237.
[21] D.T. Luc, Scalarization and stability, In: Theory of Vector Optimization, pp. 80–108, Springer, Berlin, 1989.
[22] K. Miettinen, Nonlinear Multiobjective Optimization, Springer, New York, 1999.
[23] M.M. Mäkelä, N. Karmitsa, O. Wilppu, Proximal bundle method for nonsmooth and nonconvex multiobjective optimization, In: Mathematical Modeling and Optimization of Complex Structures, pp. 191–204, Springer, 2016.
[24] B.R. Marks, G.P. Wright, A general inner approximation algorithm for nonconvex mathematical programs, Oper. Res. 26 (1978), 681–683.
[25] O. Montonen, N. Karmitsa, M.M. Mäkelä, Multiple subgradient descent bundle method for convex nonsmooth multiobjective optimization, Optimization 67 (2018), 139–158.
[26] N.P. Osmolovskii, V. M. Veliov, Strong metric subregularity of the optimality mapping and second-order sufficient optimality conditions in extremal problems with constraints, Commun. Optim. Theory 2025 (2025), 31.
[27] B.N. Pshenichnyi, Necessary Conditions for an Extremum, Marcel Dekker, New York, 1971.
[28] M. Razaviyayn, M. Hong, Z.Q. Luo, A unified convergence analysis of block successive minimization methods for nonsmooth optimization, SIAM J. Optim. 23 (2013), 1126–1153.
[29] A. Roubi, Theorems of the alternative revisited and some extensions of Yuan's lemma, Set-Valued Var. Anal. 30 (2022), 1213–1232.
[30] P. Ruiz-Canales, A. Rufián-Lizana, A characterization of weakly efficient points, Math. Program. 68 (1995), 205–212.
[31] M. Saadati, O. Morteza, Optimality conditions for robust nonsmooth multiobjective optimization problems in Asplund spaces, Bull. Belg. Math. Soc.-Simon Stevin 28 (2022), 579–601.
[32] A. Taa, Optimality conditions for vector optimization problems of a difference of convex mappings, J. Glob. Optim. 313 (2005), 421–436.
[33] J. Thomann, G. Eichfelder, A trust-region algorithm for heterogeneous multiobjective optimization, SIAM J. Optim. 29 (2019), 1017–1047.
[34] S. Wang, Algorithms for multiobjective and nonsmooth optimization, In: P. Kleinschmidt, F. Radermacher, W. Sweitzer, H. Wildermann (eds.) Methods of Operations Research, vol. 58, pp. 131–142, Athenaum Verlag, Kronberg im Taunus, 1989.
[35] D.J. White, Epsilon-dominating solutions in mean-variance portfolio analysis, Eur. J. Oper. Res. 105 (1998), 457–466.

[36] 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.
[37] X. Zhao, H. Zhang, Y. Yao, An inexact nonmonotone projected gradient method for constrained multiobjective optimization, J. Nonlinear Var. Anal. 8 (2024), 517-531.