Journal of Optimization Theory and Applications (2024) 203:126–145
https://doi.org/10.1007/s10957-024-02478-3

On Necessary Optimality Conditions for Sets of Points in Multiobjective Optimization

Andrea Cristofari[1] · Marianna De Santis[2] · Stefano Lucidi[2]

Received: 6 October 2023 / Accepted: 7 June 2024 / Published online: 3 July 2024

Abstract

Taking inspiration from what is commonly done in single-objective optimization, most local algorithms proposed for multiobjective optimization extend the classical iterative scalar methods and produce sequences of points able to converge to single efficient points. Recently, a growing number of local algorithms that build sequences of sets has been devised, following the real nature of multiobjective optimization, where the aim is that of approximating the efficient set. This calls for a new analysis of the necessary optimality conditions for multiobjective optimization. We explore conditions for sets of points that share the same features of the necessary optimality conditions for single-objective optimization. On the one hand, from a theoretical point of view, these conditions define properties that are necessarily satisfied by the (weakly) efficient set. On the other hand, from an algorithmic point of view, any set that does not satisfy the proposed conditions can be easily improved by using first-order information on some objective functions. We analyse both the unconstrained and the constrained case, giving some examples.

Keywords Multiobjective optimization · Many-objective optimization · Necessary optimality conditions

Mathematics Subject Classification 90C29 · 90C46 · 90C30

Communicated by Ana Luisa Custodio.

✉ Marianna De Santis
marianna.desantis@uniroma1.it

Andrea Cristofari
andrea.cristofari@uniroma2.it

Stefano Lucidi
lucidi@diag.uniroma1.it

1 University of Rome “Tor Vergata”, Rome, Italy

2 Sapienza University of Rome, Rome, Italy

1 Introduction

We consider multiobjective optimization problems of the following form:

$$\min_{x\in\mathcal{F}} f(x) = (f_1(x), \ldots, f_m(x))^T, \tag{1}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, are continuous functions and $\mathcal{F} \subseteq \mathbb{R}^n$. The set $\mathcal{F}$ is commonly known as the *decision space* of the problem, while the image of points in $\mathcal{F}$ through the function f is called the *image space* (or *criterion space*). To characterize the solutions of Problem (1), we use the standard optimality notion based on the dominance in the image space. In particular, given two points $x', x'' \in \mathcal{F}$, we say that x' dominates x'' and $f(x')$ dominates $f(x'')$ if $f_i(x') \leq f_i(x''),\ i = 1, \ldots, m$ and $f(x') \neq f(x'')$. A subset N of $\mathbb{R}^m$ is *stable* with respect to the partial ordering $\leq$, or simply, stable, if $z \nleq z'$ for all $z, z' \in N$. We also say that a subset N' of $\mathbb{R}^m$ is *weakly stable* if $z \nless z'$ for all $z, z' \in N'$. Given a vector $z \in \mathbb{R}^m$ and a subset of indices $I \subseteq \{1, \ldots, m\}$, we denote by z_I the subvector of z with components in I.

Definition 1.1 ((*weakly*) *efficient and* (*weakly*) *nondominated point*) A feasible point $x^* \in \mathcal{F}$ is called an *efficient point* for Problem (1) if there is no point $x \in \mathcal{F}$ such that

$$f_i(x) \leq f_i(x^*) \text{ for } i = 1, \ldots, m \text{ and } f_k(x) < f_k(x^*) \text{ for some } k \in \{1, \ldots, m\}.$$

The image $f(x^*)$ is called *nondominated point*. Moreover, a feasible point $\hat{x} \in \mathcal{F}$ is called a *weakly efficient point* for Problem (1) if there is no point $x \in \mathcal{F}$ such that

$$f_i(x) < f_i(\hat{x}) \quad \text{for } i = 1, \ldots, m.$$

The image $f(\hat{x})$ is called *weakly nondominated point*.

Algorithms for multiobjective optimization aim at approximating the set of all nondominated points, called nondominated set and the efficient set:

Definition 1.2 *(efficient and nondominated set)* The set of (weakly) efficient points of Problem (1) is called (weakly) efficient set and is denoted by $\mathcal{E}$ ($\mathcal{E}_w$). The image set of all (weakly) efficient points is the (weakly) nondominated set and is denoted by $\mathcal{N}$ ($\mathcal{N}_w$) (also known, specifically for $m = 2$, as Pareto front).

Now, we report the so called *weak domination property* (see [20, Chapter 4, Definition 4.9]), which will be used in the next sections.

Definition 1.3 We say that the weak domination property holds for Problem (1) if, for every $u \in \mathcal{F}$, either $u \in \mathcal{E}_w$ or there is some $y \in \mathcal{E}_w$ such that $f_i(y) < f_i(u)$ for all $i = 1, \ldots, m$.

A sufficient condition for the weak domination property to hold is the existence of compact level sets for (at least) one objective function, as shown in the following proposition.

Proposition 1.1 *Let us consider Problem (1) and assume that there exists $j \in \{1, \ldots, m\}$ such that the level sets $\mathcal{L}_j(\alpha) := \{x \in \mathcal{F} \colon f_j(x) \leq \alpha\}$ are compact for all $\alpha \in \mathbb{R}$. Then, the weak domination property holds.*

Proof Take any $u \in \mathcal{F} \setminus \mathcal{E}_w$. We want to show that there exists $y \in \mathcal{E}_w$ such that $f_i(y) < f_i(u)$ for all $i = 1, \ldots, m$. To this aim, consider the following problem:

$$\begin{aligned} &\min_{x \in \mathcal{F}} \varphi(x) := \sum_{i=1}^{m} f_i(x) \\ &\text{s.t.} \quad f_i(x) \leq f_i(u), \; i = 1, \ldots, m. \end{aligned} \tag{2}$$

Since $u \in \mathcal{F} \setminus \mathcal{E}_w$, then there exists $v \in \mathcal{F}$ such that $f_i(v) < f_i(u)$ for all $i = 1, \ldots, m$, implying that the feasible set of Problem (2) is nonempty. Moreover, the feasible set of Problem (2) is compact since $\mathcal{L}_j(f_j(u))$ is compact by hypothesis. Hence, since $\varphi(x)$ is continuous, then an optimal solution y of Problem (2) exists by Weierstrass Theorem. Now, assume by contradiction that $y \notin \mathcal{E}_w$. Since $y \in \mathcal{F}$, then $z \in \mathcal{F}$ exists such that $f_i(z) < f_i(y)$ for all $i = 1, \ldots, m$, implying that z is feasible for Problem (2) and $\varphi(z) < \varphi(y)$. This contradicts the optimality of y and leads to the desired result. □

As in single-objective optimization we can divide algorithms for multiobjective optimization into global and local algorithms. Global algorithms are based on the use of sequences of sets, in order to get information on the global behavior of the problem. It is a debated topic what can be considered as a proper approximation of the nondominated set (see [24] for an overview). We recall, for example, the concept of enclosure defined in [8–11], essentially being a well-structured set in the image space such as a union of boxes, which contains the nondominated set as a subset.

On the other hand, taking inspiration from what is commonly done in single-objective optimization, local algorithms for multiobjective optimization use the fact that if a point does not satisfy *suitable necessary optimality conditions*, then it can be easily improved, in the sense that a "better" new point can be easily defined, with respect to a specific quality measure, which is usually the objective function value. The majority of algorithms proposed in this respect, both for unconstrained and constrained multiobjective problems, extend the classical iterative scalar optimization algorithms, such as the steepest descent [13], Newton [12, 16], external penalty [14], interior point [15], just to name a few. The mentioned approaches produce sequences of points able to converge to single efficient points. In particular, at every iteration, these algorithms look for a new "improved" point, that is, a point that dominates the previous one.

Recently, local algorithms that build sequences of sets have been proposed [3–6, 17–19], trying to go back to the real aim of multiobjective optimization, which is approximating a set (and not a single point). Producing sequences of sets instead of single points necessarily leads to explore new algorithmic approaches. Indeed, at every iteration, such algorithms look for a new "improved" set, namely either a larger set or a set containing points that dominate at least one point from the previous set.

The aim of this work is that of exploring the necessary optimality conditions in multiobjective optimization from a new perspective, namely, the definition of conditions associated to a set of points, instead of a single point. In particular, this can be of interest from both a theoretical and an algorithmic point of view. As in single-objective optimization, these conditions play a twofold role:

- from a theoretical point of view, they define properties that are necessarily satisfied by a global solution, that is, by the (weakly) efficient set $\mathcal{E}_w$;
- from an algorithmic point of view, they characterize sets that cannot be "easily" improved, in the sense that, for example, the standard use of first-order information may not be enough to get a new better set.

Therefore, in the multiobjective optimization context, we look for conditions on sets of points that approximate the efficient set and share the above two features. Such conditions, as stated before, should help to define algorithms that build sequences of stable sets. Furthermore, in presence of many objective functions, such conditions could give a stronger characterization of Pareto stationarity, so that they can be particularly useful in the context of many-objective optimization.

In order to introduce our analysis, we start with the following definition.

Definition 1.4 (*Efficiency-describing collection*) Let $S \subseteq \mathcal{F}$ be a set whose image points with respect to f form a weakly stable set. Given $x \in S$, we say that the collection of sets $\mathcal{T}_S(x) = \{I^1, \dots, I^p\} \neq \emptyset$ is an *efficiency-describing collection* of x if it satisfies the following two conditions:

(i) $I^j \subseteq \{1, \dots, m\}$, $j = 1, \dots, p$, are subsets of objective indices such that there is no point $y \in S$ for which

$$f_i(y) < f_i(x) \quad \forall i \in I^j;$$

(ii) $\bigcup_{i \neq j} I^i \nsubseteq I^j$ for all $j \in \{1, \dots, p\}$, if $p \geq 2$.

Remark 1.1 Note that $\mathcal{T}_S(x) = \{I^1\}$ with $I^1 = \{1, \dots, m\}$ is always an efficiency-describing collection of x, for all $x \in S$, meaning that an efficiency-describing collection always exists for any point $x \in S$. However, it may not be unique. Indeed, in case we are able to detect a proper subset of $\{1, \dots, m\}$ satisfying condition (i) of Definition 1.4, we can also choose a different $\mathcal{T}_S(x)$, not containing the whole set of objective functions, as condition (ii) requires. This means that either $\mathcal{T}_S(x)$ is made of the whole set $\{1, \dots, m\}$ or it is made of proper subsets of $\{1, \dots, m\}$.

Example 1.1 In Fig. 1.1, the image space of a bi-objective instance is depicted. The image of the feasible points through $f(x) = (f_1(x), f_2(x))$ is represented by $f(\mathcal{F})$ and the nondominated set $\mathcal{N}$ is highlighted with a bold line. Let $S = \{x^1, x^2, x^3\}$, the images $f(x^1)$, $f(x^2)$, $f(x^3)$ belong to $\mathcal{N}$, so that in particular they form a stable set. The following sets $\mathcal{T}_S(x^i)$ are efficiency-describing collections of x_i for $i = 1, 2, 3$:

$$\mathcal{T}_S(x^1) = \left\{I^1 = \{1\}\right\}, \quad \mathcal{T}_S(x^2) = \left\{I^1 = \{1, 2\}\right\}, \quad \mathcal{T}_S(x^3) = \left\{I^1 = \{2\}\right\}.$$

We can notice that since $f_1(x^1)$ is the minimum with respect to $f_1(x)$ for $x \in \mathcal{F}$, we have that $f(x^1)$ is a nondominated point in $f(S)$ "thanks to" f_1, while f_2 does

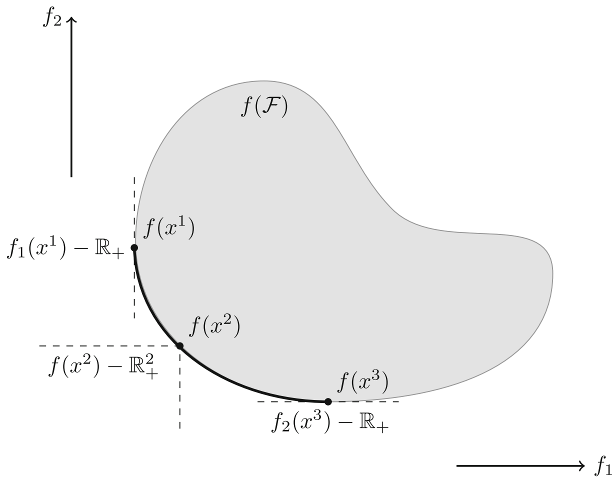

Fig. 1 Image set of a bi-objective instance

not play any role. Equivalently, since $f_2(x^3)$ is the minimum with respect to $f_2(x)$, we have that $f(x^3)$ is nondominated with respect to the images of the other points in S. On the other hand, $f(x^2)$ is nondominated with respect to the images of the other two points as $f_1(x^2) < f_1(x^3)$ and $f_2(x^2) < f_2(x^1)$, meaning that $\mathcal{T}_S(x^2)$ needs to include both indices.

Example 1.1 highlights that, in order to characterize a point whose image is (weakly) nondominated with respect to the images of other points in a set, we do not necessarily have to consider all the objective functions. A trivial case is when we have a set including the minimizer of one objective function. This calls for new necessary optimality conditions when dealing with a set of points. Such conditions should reflect the fact that specific subsets of objective functions might be used to claim that the image of a point is (weakly) nondominated with respect to the images of other points in a set.

The difference with respect to classical conditions is more relevant when considering more than two objective functions, as the following example shows.

Example 1.2 Let us consider a multiobjective problem with five objective functions. Let $S = \{x^1, x^2, x^3, x^4\} \subset \mathcal{F}$ with

$$f(x^1) = \begin{pmatrix} 3 \\ 1 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \quad f(x^2) = \begin{pmatrix} 2 \\ 4 \\ 1 \\ 3 \\ 2 \end{pmatrix}, \quad f(x^3) = \begin{pmatrix} 3 \\ 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}, \quad f(x^4) = \begin{pmatrix} 5 \\ 2 \\ 1 \\ 3 \\ 5 \end{pmatrix}.$$

The set $N = \{f(x^1), f(x^2), f(x^3), f(x^4)\}$ is a stable set. The following sets $\mathcal{T}_S(x^i)$, are efficiency-describing collections of x_i, $i = 1, \ldots, 4$:

$$\begin{aligned}
\mathcal{T}_S(x^1) &= \{I^1 = \{2\},\ I^2 = \{5\},\ I^3 = \{4\}\},\\
\mathcal{T}_S(x^2) &= \{I^1 = \{1\},\ I^2 = \{3\}\},\\
\mathcal{T}_S(x^3) &= \{I^1 = \{4\},\ I^2 = \{2, 3\},\ I^3 = \{3, 5\}\},\\
\mathcal{T}_S(x^4) &= \{I^1 = \{3\}\}.
\end{aligned}$$

Note that the fact that $f(x^1)$ is (weakly) nondominated with respect to the images of the other points in S can be deduced from f_2 or f_5, which attain their least value. Furthermore, $f_4(x^1) \leq f_4(x^i)$ for $i \neq 1$. This suggests that the characterization of the fact that $f(x^1)$ is (weakly) nondominated with respect to the images of the other points in S can be described by the local behavior of these three functions only (i.e., f_2, f_5, f_4). Indeed, the behavior of f_1 can be ignored for x^1 since $f_1(x^1)$ is worse than values attained in other points, like $f_1(x^2)$. In order to describe that $f(x^3)$ is (weakly) nondominated with respect to the images of the other points in S, we can consider either f_4, for which x^3 attains its minimum value, or the subsets made of f_2, f_3 and f_3, f_5. Indeed the index 2 needs to belong to I^2 in order to have x^3 nondominated with respect to x^2, the index 3 is needed in order to have x^3 nondominated with respect to x^1 and the index 5 is needed in combination with 3 in order to have x^3 nondominated with respect to x^4. Note that, in the definition of $\mathcal{T}_S(x^i)$, $i = 1, \ldots, 4$, we are not considering the natural choice of the whole set of objective function indices $\{1, \ldots, 5\}$.

2 The Unconstrained Case

We start by analyzing multiobjective unconstrained problems (i.e., $\mathcal{F} = \mathbb{R}^n$) of the form

$$\min_{x \in \mathbb{R}^n} (f_1(x), \ldots, f_m(x))^T. \tag{3}$$

In the following, we assume that the objective functions $f_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \ldots, m$, are continuously differentiable. In the literature, to characterize (weakly) efficient points, the following condition has been proposed [23].

Proposition 2.1 *If $\mathcal{E}_w \subseteq \mathbb{R}^n$ is the weakly efficient set for Problem (3), then, for all $x^* \in \mathcal{E}_w$ we have that, for all $d \in \mathbb{R}^n$, there exists an index $j \in \{1, \ldots, m\}$ such that $\nabla f_j(x^*)^\top d \geq 0$.*

According to Proposition 2.1, we can introduce the following definition [12]:

Definition 2.1 (Standard Pareto stationary set) Let $S \subseteq \mathbb{R}^n$ be a non-empty set such that $f(S)$ is a weakly stable set. We say that S is a *standard Pareto stationary set* for Problem (3) if, for all $x \in S$ and all $d \in \mathbb{R}^n$, there exists an index $i \in \{1, \ldots, m\}$ such that $\nabla f_i(x)^\top d \geq 0$.

As already mentioned in the introduction, the fact that the image of a point is nondominated with respect to the image of the other points within a set can be characterized by a subset of objective functions. This suggests that we can characterize the (weakly) efficient set of Problem (3) using stronger conditions than those in Proposition 2.1. In particular, there is no need of considering the whole set of objective functions. This can be important when the number of objective functions is large. In those cases, Definition 2.1 poorly characterize sets of efficient points. Indeed, it equivalently states that a set of points is standard Pareto stationary if, for every point in the set, there is no direction that is a descent direction for every objective function, implying that the higher the number of objective functions is, the easier the definition can be satisfied. This is highlighted in the following example.

Example 2.1 Let us consider an unconstrained multiobjective problem with the following objective functions $f_i : \mathbb{R}^2 \to \mathbb{R}$. $i = 1, \dots, 5$,

$$
\begin{aligned}
&f_1(x) = (x_1 - 2)^2 + (x_2 - 3)^2 - 7, \quad f_2(x) = x_1^2 + x_1, \quad f_3(x) = x_2^2 + 3x_2,\\
&f_4(x) = (x_1 - 1)^2 + (x_2 - 1)^2 + 1, \quad f_5(x) = (x_1 + 1)^2 + (x_2 - 1)^2.
\end{aligned}
$$

Their gradients are

$$
\nabla f_1(x) = \begin{pmatrix} 2(x_1 - 2) \\ 2(x_2 - 3) \end{pmatrix}, \quad \nabla f_2(x) = \begin{pmatrix} 2x_1 + 1 \\ 0 \end{pmatrix}, \quad \nabla f_3(x) = \begin{pmatrix} 0 \\ 2x_2 + 3 \end{pmatrix},
$$

$$
\nabla f_4(x) = \begin{pmatrix} 2(x_1 - 1) \\ 2(x_2 - 1) \end{pmatrix}, \quad \nabla f_5(x) = \begin{pmatrix} 2(x_1 + 1) \\ 2(x_2 - 1) \end{pmatrix}.
$$

Let $S = \{x^1, x^2, x^3, x^4\} \subset \mathcal{F}$, with

$$
x^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad x^2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad x^3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad x^4 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
$$

Note that S satisfies Definition 2.1. Indeed, x^1 and x^3 are stationary with respect to functions f_4 and f_5 respectively, so that for any direction $d \in \mathbb{R}^n$ we have that $\nabla f_4(x^1)^T d = 0$ and $\nabla f_5(x^3)^T d = 0$. For $i = 2, 4$, we have that $\nabla f_1(x^i)$, $\nabla f_2(x^i)$ and $\nabla f_3(x^i)$ can be combined with positive coefficients to obtain the zero vector. From Gordan's Theorem of alternative (see Table 2.4.1 [22] or Theorem A.1 in the appendix) this implies that, for any $d \in \mathbb{R}^n$, there is one index $j \in \{1, 2, 3\}$ such that $\nabla f_j(x^i)^T d \geq 0$, $i = 2, 4$.

However, looking at the objective function values at the points in S, we can see that the set S can be easily "improved", for example by moving point x^2 along a descent direction with respect to f_5, as it will be clarified in Example 2.2.

As shown in Example 2.1, there is room to improve the definition of standard Pareto stationarity. One weakness in Definition 2.1 is that it does not fully exploit that the considered set is a weakly stable set. In particular, it ignores the fact that not all the objective functions should necessarily be taken into account. Indeed, there exist objective functions that can be neglected in order to characterize the fact that a point in a weakly stable set is nondominated with respect to the others (as highlighted also in Example 1.2). In this respect, Definition 1.4 comes into play and allows us to state the following result, giving stronger optimality conditions for Problem (3) that include Proposition 2.1.

Proposition 2.2 *Let $\mathcal{E}_w \subseteq \mathbb{R}^n$ be the weakly efficient set of Problem (3), $x^* \in \mathcal{E}_w$ and $\mathcal{T}_{\mathcal{E}_w}(x^*)$ be any efficiency-describing collection of x^*. If the weak domination property holds for Problem (3), then, for all $d \in \mathbb{R}^n$ and all $I \in \mathcal{T}_{\mathcal{E}_w}(x^*)$, there exists an index $i \in I$ such that $\nabla f_i(x^*)^\top d \geq 0$.*

Proof Assume by contradiction that there exists $x^* \in \mathcal{E}_w$, $d \in \mathbb{R}^n$ and $I \in \mathcal{T}_{\mathcal{E}_w}(x^*)$ such that

$$\nabla f_i(x^*)^\top d < 0 \quad \forall i \in I.$$

Then, there exists $\bar{\alpha} > 0$ such that, for any $\alpha \in (0, \bar{\alpha}]$, we have

$$f_i(x^* + \alpha d) < f_i(x^*) \quad \forall i \in I. \tag{4}$$

From the definition of I it follows that, for all $y \in \mathcal{E}_w$, an index $\hat{\imath} \in I$ exists such that

$$f_{\hat{\imath}}(y) \geq f_{\hat{\imath}}(x^*)$$

and, using (4), we also have

$$f_{\hat{\imath}}(x^* + \alpha d) < f_{\hat{\imath}}(x^*).$$

Hence, for all $y \in \mathcal{E}_w$, there exists $\hat{\imath}$ such that

$$f_{\hat{\imath}}(x^* + \alpha d) < f_{\hat{\imath}}(y), \tag{5}$$

implying that $(x^* + \alpha d) \in \mathbb{R}^n \setminus \mathcal{E}_w$. Then, from the weak domination property, we get that $\hat{y} \in \mathcal{E}_w$ exists such that for all i we have

$$f_i(\hat{y}) < f_i(x^* + \alpha d),$$

leading to a contradiction with (5). □

Remark 2.1 Proposition 2.2 allows us to define different necessary optimality conditions according to the choice of $\mathcal{T}_{\mathcal{E}_w}(x^*)$. If we consider $\mathcal{T}_{\mathcal{E}_w}(x^*) = \{I^1\}$ with $I^1 = \{1, \ldots, m\}$ for all $x^* \in \mathcal{E}_w$, then Proposition 2.2 boils down to Proposition 2.1. On the contrary, the larger the number of the subsets of indices within $\mathcal{T}_{\mathcal{E}_w}(x^*)$ and the smaller their cardinality, the stronger the conditions. This highlights how Proposition 2.2 can be particularly useful in the context of many-objective optimization.

Remark 2.2 Condition (ii) in Definition 1.4 is not necessary within the proof of Proposition 2.2. However, it avoids redundant information in the definition of the efficiency-describing collection $\mathcal{T}_{\mathcal{E}_w}(x)$.

From Proposition 2.2, we introduce the following new definition of Pareto stationary set.

Definition 2.2 (*Pareto stationary set*) Let $S \subseteq \mathbb{R}^n$ be a non-empty set such that $f(S)$ is a weakly stable set. We say that S is a *Pareto stationary set* for Problem (3) if, for all $x \in S$, all $d \in \mathbb{R}^n$ and all $I \in \mathcal{T}_S(x)$, there exists an index $i \in I$ such that $\nabla f_i(x)^\top d \geq 0$.

Local algorithms for multiobjective optimization that aim at detecting one single efficient point are generally based on the optimality conditions reported in Proposition 2.1 (see e.g. [13]). In particular, if a point does not satisfy Pareto stationarity conditions, then it is possible to define a direction that is a descent direction for *each* objective function, thus able to produce a new point that dominates the previous one.

When the aim is producing a sequence of sets of points, local algorithms can be based on the conditions introduced in Proposition 2.2. In particular, these conditions on the one hand allow us to certify whether a set S is Pareto stationary. On the other hand, in case a set S does not satisfy these conditions, we can *easily* either increase the cardinality of S or replace some points in S by new points dominating them, as shown in the next result.

Proposition 2.3 *Let $S \subseteq \mathbb{R}^n$ and let $f(S)$ be a weakly stable set. If S is a non-Pareto stationary set for Problem (3), then a point $x \in S$, a direction $d \in \mathbb{R}^n$ and a stepsize $\alpha > 0$ exist such that $f(x + \alpha d)$ is non-dominated by any $f(y)$, $y \in S$. Namely,*

$$\nexists\, y \in S \text{ such that } f_i(y) \leq f_i(x + \alpha d),\ i = 1, \ldots, m, \quad f(y) \neq f(x).$$

Furthermore, $f(x + \alpha d)$ is non-dominated by any $z \in \mathbb{R}^m$ that is dominated by some $f(y)$, $y \in S$.

Proof Since $S \subseteq \mathbb{R}^n$ is a non-Pareto stationary set, there exist $x \in S$, $d \in \mathbb{R}^n$ and $I \in \mathcal{T}_S(x)$ such that

$$\nabla f_i(x)^\top d < 0 \quad \forall i \in I.$$

Then, there exists $\bar{\alpha} > 0$ such that, for any $\alpha \in (0, \bar{\alpha}]$, we have

$$f_i(x + \alpha d) < f_i(x) \quad \forall i \in I. \tag{6}$$

From the definition of I it follows that, for all $y \in S$, there exists an index $\hat{\imath} \in I$ such that

$$f_{\hat{\imath}}(y) \geq f_{\hat{\imath}}(x).$$

Furthermore, from (6) we also have

$$f_{\hat{\imath}}(x+\alpha d) < f_{\hat{\imath}}(x),$$

leading to

$$f_{\hat{\imath}}(x+\alpha d) < f_{\hat{\imath}}(y). \tag{7}$$

Since this holds for all $y \in S$, it follows that $f(x+\alpha d)$ is non-dominated by any $f(y)$, $y \in S$.

Now, consider any $z \in \mathbb{R}^m \setminus f(S)$ dominated by some $f(y)$, $y \in S$. Namely, $f_i(y) \le z_i$ for all $i = 1, \dots, m$ and $z \neq f(y)$. Using (7), we have

$$f_{\hat{\imath}}(x+\alpha d) < f_{\hat{\imath}}(y) \le z_{\hat{\imath}}.$$

It follows that $f(x+\alpha d)$ is non-dominated by z. □

Example 2.2 Let us consider the multiobjective problem proposed in Example 2.1, where the set S is a standard Pareto stationary set, that is, it satisfies Definition 2.1. Assume that this set S is produced by an algorithm that builds sequences of points. The use of the standard Pareto stationarity definition would make the algorithm stop. On the other hand, this would not be the case if using our new stationarity characterization given in Definition 2.2, with $\mathcal{T}_S(x)$ specifically chosen.

Indeed, $f(S) \subset \mathbb{R}^5$ is the following stable set:

$$f(x^1) = \begin{pmatrix} -2 \\ 2 \\ 4 \\ 1 \\ 4 \end{pmatrix}, \quad f(x^2) = \begin{pmatrix} -2 \\ 0 \\ 10 \\ 3 \\ 2 \end{pmatrix}, \quad f(x^3) = \begin{pmatrix} 6 \\ 0 \\ 4 \\ 5 \\ 0 \end{pmatrix}, \quad f(x^4) = \begin{pmatrix} 10 \\ 2 \\ -2 \\ 5 \\ 8 \end{pmatrix}.$$

The following are possible choices of $\mathcal{T}_S(x^i)$, $i = 1, \dots, 4$, satisfying Definition 1.4:

$$\begin{aligned} \mathcal{T}_S(x^1) &= \left\{I^1 = \{1\},\ I^2 = \{4\}\right\}, \\ \mathcal{T}_S(x^2) &= \left\{I^1 = \{1\},\ I^2 = \{2\}\right\}, \\ \mathcal{T}_S(x^3) &= \left\{I^1 = \{2\},\ I^2 = \{5\}\right\}, \\ \mathcal{T}_S(x^4) &= \left\{I^1 = \{3\}\right\}. \end{aligned}$$

Note that even if x^1 is stationary with respect to f_4, since $I^1 = \{1\}$ belongs to $\mathcal{T}_S(x^1)$ and x^1 is non stationary with respect to f_1, then a direction $d \in \mathbb{R}^n$ such that $\nabla f_1(x^1)^T d < 0$ exists. Hence, Definition 2.2 is not satisfied and S is a non-Pareto stationary set for Problem (1). From Proposition 2.3, the set S can be easily "improved" as a direction $d \in \mathbb{R}^n$ and a stepsize $\alpha > 0$ exist such that $f(x^1 + \alpha d)$ is non-dominated by the image of any point in S. Similar arguments apply to the other points x^i, $i = 2, 3, 4$. For example, starting from x^2, we can improve S in different ways, using descent directions for f_1 or f_2, starting from x^3 we can move along a descent direction for f_2 and, starting from x^4, we can move along a descent direction for f_3.

3 The Constrained Case

In this section, we extend the above analysis to the constrained setting. We consider the following constrained multiobjective optimization problem, where the feasible set $\mathcal{F}$ is explicitly defined by inequality and equality constraints:

$$\min_{x \in \mathcal{F}} (f_1(x), \dots, f_m(x))^T \quad \text{with } \mathcal{F} = \{x \in \mathbb{R}^n : g(x) \leq 0, \ h(x) = 0\}. \tag{8}$$

We assume that the functions $f : \mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^p$ and $h : \mathbb{R}^n \to \mathbb{R}^q$ are continuously differentiable. Given a point $x \in \mathcal{F}$, we denote by $A(x)$ the active set, defined as follows:

$$A(x) := \{i \in \{1, \dots, p\} : g_i(x) = 0\}.$$

In the literature, to characterize (weakly) efficient points, the following Fritz-John necessary conditions have been proposed [7, 23].

Proposition 3.1 *If $\mathcal{E}_w \subseteq \mathbb{R}^n$ is the weakly efficient set for Problem (8), then, for all $x \in \mathcal{E}_w$, there exist multipliers $\sigma^* \in \mathbb{R}^m$, $\lambda^* \in \mathbb{R}^p$, $\mu^* \in \mathbb{R}^q$, with $(\sigma^*, \lambda^*, \mu^*) \neq (0, 0, 0)$, such that*

$$\sum_{i=1}^{m} \sigma_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \lambda_j^* \nabla g_j(x^*) + \sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) = 0, \tag{9a}$$

$$\sigma_i^* \geq 0, \qquad i = 1, \dots, m, \tag{9b}$$

$$\lambda_j^* \geq 0, \qquad j = 1, \dots, p, \tag{9c}$$

$$\lambda_j^* g_j(x^*) = 0, \qquad j = 1, \dots, p. \tag{9d}$$

As for the unconstrained case, we can define the standard Pareto stationary set as the set containing points satisfying first-order necessary conditions [7, 23].

Definition 3.1 (*Standard Pareto stationary set*) Let $S \subseteq \mathbb{R}^n$ be a non-empty set such that $f(S)$ is a weakly stable set. We say that S is a *standard Pareto stationary set* for Problem (8) if all $x \in S$ satisfy Fritz-John conditions, that is, for each $x \in S$, there exist multipliers $(\sigma, \lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q$, with $(\sigma, \lambda, \mu) \neq (0, 0, 0)$, such that (9a)–(9d) holds.

As before, Definition 1.4 comes into play and allows us to state stronger conditions for Problem (8) than those in Proposition 3.1, without the need of considering the whole set of objective functions. To this aim we need an intermediate result, based on sets $G(\cdot)$ and $H(\cdot)$ introduced below.

Definition 3.2 Given $\bar{x} \in \mathcal{F}$, we define the sets $G(\bar{x})$ and $H(\bar{x})$ as follows

$$G(\bar{x}) = \{d \in \mathbb{R}^n : \nabla g_i(\bar{x})^T d < 0, \text{ for all } i \in A(\bar{x})\},$$
$$H(\bar{x}) = \{d \in \mathbb{R}^n : \nabla h_j(\bar{x})^T d = 0, \text{ for all } j = 1, \dots, q\}.$$

We now extend the classical result for single objective constrained optimization [1, Theorem 4.3.1] to the multiobjective case, using the sets introduced in Definition 1.4 and in Definition 3.2.

Lemma 3.1 *Let $\mathcal{E}_w \subseteq \mathcal{F}$ be the weakly efficient set for Problem (8) and let $x^* \in \mathcal{E}_w$. If the weak domination property holds for Problem (8) and if $\nabla h_j(x^*)$, $j = 1, \dots, q$ are linearly independent, then, for all $d \in G(x^*) \cap H(x^*)$ and all $I \in \mathcal{T}_{\mathcal{E}_w}(x^*)$, there exists $i \in I$ such that $\nabla f_i(x^*)^\top d \geq 0$.*

Proof Given $I \in \mathcal{T}_{\mathcal{E}_w}(x^*)$, let $F^I(x^*)$ be the set defined as follows

$$F^I(x^*) = \{d \in \mathbb{R}^n : \nabla f_i(x^*)^\top d < 0, \ i \in I\}.$$

Proving the lemma is equivalent to showing that for all $I \in \mathcal{T}_{\mathcal{E}_w}(x^*)$

$$F^I(x^*) \cap G(x^*) \cap H(x^*) = \emptyset.$$

This follows from the proof of [1, Theorem 4.3.1] with minor changes using the weak domination property. □

Thanks to Lemma 3.1 we are able to state a new characterization for a weakly efficient point of Problem (8).

Proposition 3.2 *Let $\mathcal{E}_w$ be the weakly efficient set of Problem (8), $x^* \in \mathcal{E}_w$ and $\mathcal{T}_{\mathcal{E}_w}(x^*)$ be any efficiency-describing collection of x^*. If the weak domination property holds for Problem (8), then, for all $I \in \mathcal{T}_{\mathcal{E}_w}(x^*)$, there exist multipliers $\sigma^* \in \mathbb{R}^m$, $\lambda^* \in \mathbb{R}^p$, $\mu^* \in \mathbb{R}^q$, with $(\sigma^*, \lambda^*, \mu^*) \neq (0, 0, 0)$, such that*

$$\sum_{i \in I} \sigma_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \lambda_j^* \nabla g_j(x^*) + \sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) = 0, \tag{10a}$$
$$\sigma_i^* \geq 0, \quad i \in I, \tag{10b}$$
$$\sigma_i^* = 0, \quad i \notin I, \tag{10c}$$
$$\lambda_j^* \geq 0, \quad j = 1, \dots, p, \tag{10d}$$
$$\lambda_j^* g_j(x^*) = 0, \quad j = 1, \dots, p. \tag{10e}$$

Proof If $\nabla h_j(x^*)$ $j = 1, \dots, q$ are linearly dependent, we can find scalars $\mu_1^*, \dots, \mu_q^*$ not all zeros such that $\sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) = 0$. Letting $\sigma_i^* = 0$, $i \in I$ and $\lambda_j^* = 0$, $j = 1, \dots, p$ the result holds trivially. Now, assume that $\nabla h_j(x^*)$ $j = 1, \dots, q$ are linearly independent. Let C be the matrix whose rows are $\nabla h_l^T(x^*)$, $l = 1, \dots, q$ and D be the matrix whose rows are $\nabla f_i^T(x^*)$, $i \in I$ and $\nabla g_j^T(x^*)$, $j \in A(x^*)$. Then, from Lemma 3.1, we have that the following system

$$Cy = 0, \qquad Dy < 0$$

has no solution. From the Slater's theorem of the alternative (see Table 2.4.1 [22] or Theorem A.2 in the appendix) we have that non-negative σ_i^*, $i \in I$ and λ_j^*, $j \in A(x^*)$ not all zero and $\mu_1^*, \dots, \mu_q^* \in \mathbb{R}$ exist such that

$$\sum_{i \in I} \sigma_i^* \nabla f_i(x^*) + \sum_{j \in A(x^*)} \lambda_j^* \nabla g_j(x^*) + \sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) = 0.$$

Then, letting $\lambda_j^* = 0$, $j \notin A(x^*)$ the result holds. □

Taking inspiration from [21], we further extend our characterization of the weakly efficient points of Problem (8) given in Proposition 3.2, under constraint qualification. This allows us to state the definition of Pareto stationary set and propose how to compute a descent direction when S is not a Pareto stationary set. We first recall the classical Mangasarian–Fromovitz conditions (CMF).

Definition 3.3 The CMF holds at $x \in \mathcal{F}$ if there is no $\alpha_j \geq 0$, $j \in A(x)$, β_l, $l = 1, \dots, q$, such that

$$\sum_{j \in A(x)} \alpha_j \nabla g_j(x) + \sum_{l=1}^{q} \beta_l \nabla h_l(x) = 0, \qquad (\alpha, \beta) \neq (0, 0).$$

Proposition 3.3 *Let $\mathcal{E}_w$ be the weakly efficient set of Problem (8), $x^* \in \mathcal{E}_w$ and $\mathcal{T}_{\mathcal{E}_w}(x^*)$ be any efficiency-describing collection of x^*. If the weak domination property holds for Problem (8) and the CMF holds at x^*, then, for all $I \in \mathcal{T}_{\mathcal{E}_w}(x^*)$, there exist multipliers $\sigma^* \in \mathbb{R}^m$, $\lambda^* \in \mathbb{R}^p$, $\mu^* \in \mathbb{R}^q$, with $\sigma_I^* \neq 0$, such that*

$$\begin{aligned}
\sum_{i \in I} \sigma_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \lambda_j^* \nabla g_j(x^*) + \sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) &= 0, \\
\sigma_i^* \geq 0, \qquad i &\in I, \\
\sigma_i^* = 0, \qquad i &\notin I, \\
\lambda_j^* \geq 0, \qquad j &= 1, \dots, p, \\
\lambda_j^* g_j(x^*) = 0, \qquad j &= 1, \dots, p.
\end{aligned}$$

Proof From Proposition 3.2 we have that x^* satisfies (10a)–(10e). Assume by contradiction that $\sigma_I^* = 0$. Then, (10a) would become

$$\sum_{j=1}^{p} \lambda_j^* \nabla g_j(x^*) + \sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) = 0$$

or equivalently, since $\lambda_j = 0,\ j \notin A(x^*)$,

$$\sum_{j \in A(x^*)} \lambda_j^* \nabla g_j(x^*) + \sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) = 0,$$

so that we get a contradiction to the CMF conditions. □

We are now able to state the definition of Pareto stationary set for multiobjective constrained problems.

Definition 3.4 (*Pareto stationary set*) Let $S \subseteq \mathcal{F}$ be a non-empty set such that $f(S)$ is a weakly stable set. We say that S is a *Pareto stationary set* for Problem (8) if, for all $x \in S$ and all $I \in \mathcal{T}_S(x)$, there exist multipliers $\sigma \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^p$, $\mu \in \mathbb{R}^q$, with $\sigma_I \neq 0$, such that

$$\sum_{i \in I} \sigma_i \nabla f_i(x) + \sum_{j=1}^{p} \lambda_j \nabla g_j(x) + \sum_{l=1}^{q} \mu_l \nabla h_l(x) = 0,$$

$$\begin{aligned} \sigma_i &\geq 0,\ i \in I, \\ \sigma_i &= 0,\ i \notin I, \\ \lambda_j &\geq 0,\ j = 1, \ldots, p, \\ \lambda_j g_j(x) &= 0,\ j = 1, \ldots, p. \end{aligned}$$

Now we characterize how to compute a feasible descent direction when S is not a Pareto stationary set. For sake of simplicity, we assume that the feasible set $\mathcal{F}$ is defined according to inequality constraints only. Note that, in order to prove the result, we need to assume the CMF conditions.

Proposition 3.4 *Let $S \subseteq \mathcal{F}$ and let $f(S)$ be a weakly stable set. If S is a non-Pareto stationary set for Problem (8) with inequality constraints only and if the CMF holds at each point in S, then a point $x \in S$, a direction $d \in \mathbb{R}^n$ and a stepsize $\alpha > 0$ exist such that $(x + \alpha d) \in \mathcal{F}$ and is non-dominated by any $f(y),\ y \in S$. Namely,*

$$\nexists\, y \in S \text{ such that } f_i(y) \leq f_i(x + \alpha d)\ i = 1, \ldots, m, \quad f(y) \neq f(x),$$

Furthermore, $x + \alpha d$ is non-dominated by any $z \in \mathcal{F}$ that is dominated by some $y \in S$.

Proof We start by showing that if $S \subseteq \mathcal{F}$ is a non-Pareto stationary set for Problem (8), we can detect a point $x \in S$, $I \in \mathcal{T}_S(x)$ and a direction $d \in \mathbb{R}^n$ such that

$$\nabla f_j(x)^\top d < 0, \quad \forall j \in I \qquad \text{and} \qquad \nabla g_i(x)^\top d \leq 0, \quad i = 1, \dots, p, \tag{12}$$

with

$$\nabla g_i(x)^\top d < 0, \quad \text{for all } \; i \in A(x). \tag{13}$$

Since $S \subseteq \mathcal{F}$ is a non-Pareto stationary set for Problem (8), we have that a point $x \in S$ exists such that $I \in \mathcal{T}_S(x)$ and no multipliers $\sigma \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^p$ exist for which

$$\begin{array}{ll} \sigma_I \neq 0, & \\ \sigma_i \geq 0, & i \in I, \\ \sigma_i = 0, & i \notin I, \\ \lambda_j \geq 0, & j = 1, \dots, p, \\ \lambda_j g_j(x) = 0, & j = 1, \dots, p, \end{array}$$

and

$$\sum_{i \in I} \sigma_i \nabla f_i(x) + \sum_{j=1}^{p} \lambda_j \nabla g_j(x) = 0.$$

By Motzkin Theorem of alternative (see Table 2.4.1 [22] or Theorem A.3 in the appendix), we have that $\tilde{d} \in \mathbb{R}^n$ exists such that

$$\nabla f_j(x)^\top \tilde{d} < 0, \quad \forall j \in I \qquad \text{and} \qquad \nabla g_i(x)^\top \tilde{d} \leq 0, \quad i = 1, \dots, p.$$

We now show that necessarily $\nabla g_i(x)^\top \tilde{d} < 0$ for all $i \in A(x)$ and we proceed by contradiction. Assume that there is no $d \in \mathbb{R}^n$ such that

$$\nabla f_j(x)^\top d < 0, \quad \forall j \in I \qquad \text{and} \qquad \nabla g_i(x)^\top d < 0, \quad i \in A(x). \tag{14}$$

Then, from Gordan Theorem of alternative (see Table 2.4.1 [22] or Theorem A.1 in the appendix), we have that if (14) does not hold, then there exist $\rho \in \mathbb{R}^m$ and $\tilde{\rho} \in \mathbb{R}^p$, both with non-negative entries, $(\rho, \tilde{\rho}) \neq (0, 0)$ and with

$$\rho_j = 0, \;\; j \notin I \qquad \tilde{\rho}_i = 0, \;\; i \notin A(x),$$

such that

$$\sum_{j \in I} \rho_j \nabla f_j(x) + \sum_{i \in A(x)} \tilde{\rho}_i \nabla g_i(x) = 0.$$

In particular, by multiplying the above expression with the direction $\tilde{d} \in \mathbb{R}^n$, we obtain

$$\sum_{j \in I} \rho_j \nabla f_j(x)^\top \tilde{d} + \sum_{i \in A(x)} \tilde{\rho}_i \nabla g_i(x)^\top \tilde{d} = 0.$$

Since $\nabla f_j(x)^\top \tilde{d} < 0$ for all $j \in I$, we necessarily have $\rho_j = 0$ for all $j \in I$. Therefore

$$\sum_{i \in A(x)} \tilde{\rho}_i \nabla g_i(x) = 0,$$

getting a contradiction with the fact that CMF holds at x.

So, we have proved that $x \in S$, $d \in \mathbb{R}^n$ and $I \in \mathcal{T}_S(x)$ exist such that (12) and (13) hold. In particular, for all $i \in I$, there exists $\alpha_i > 0$ such that

$$f_i(x + \alpha d) < f_i(x) \quad \forall \alpha \in (0, \alpha_i]$$

and

$$g_j(x + \alpha d) \leq 0, \quad \forall j = 1, \dots, p, \ \forall \alpha \in (0, \alpha_i],$$

thanks to the fact that $g_j(x) < 0$ for those $j \notin A(x)$ and $\nabla g_j(x)^\top d < 0$ for $j \in A(x)$. Therefore, for any $\alpha \in (0, \min_{i \in I} \alpha_i]$, we have that $x + \alpha d \in \mathcal{F}$ and we can write

$$f_i(x + \alpha d) < f_i(x) \quad \forall i \in I. \tag{15}$$

From the definition of I, it follows that for every $y \in S$ there exists an index $\hat{\imath} \in I$ such that

$$f_{\hat{\imath}}(y) \geq f_{\hat{\imath}}(x).$$

Furthermore, from (15) we also have

$$f_{\hat{\imath}}(x + \alpha d) < f_{\hat{\imath}}(x),$$

leading to

$$f_{\hat{\imath}}(x + \alpha d) < f_{\hat{\imath}}(y). \tag{16}$$

Since this holds for all $y \in S$, it follows that $(x + \alpha d)$ is non-dominated by any $y \in S$.

Now, consider any $u \in \mathcal{F} \setminus S$ dominated by some $y \in S$. Namely, $f_i(y) \leq f_i(u)$ for all $i = 1, \dots, m$ and $f_i(y) \neq f_i(u)$. Using (16), we have

$$f_{\hat{\imath}}(x + \alpha d) < f_{\hat{\imath}}(y) \leq f_{\hat{\imath}}(u).$$

It follows that $(x + \alpha d)$ is non-dominated by u.

We finally mention that other constraint qualification conditions have been presented and analyzed in the multiobjective literature (see, e.g., [26] and the references therein). In some cases, we can use them in combination with the efficiency-describing collection to give new characterizations of a weakly efficient point. For example, we can define the extended Mangasarian–Fromovitz (EMF) conditions, presented in [2], in light of Definition 1.4.

Definition 3.5 Given $I \subset \{1, \dots, m\}$, the EMF holds at $x \in \mathcal{F}$ if, for all $s \in I$, there is no $\gamma_i \geq 0,\ i \in I, i \neq s, \alpha_j \geq 0,\ j \in A(x), \beta_l,\ l = 1, \dots, q$, such that

$$\sum_{\substack{i \in I \\ i \neq s}} \gamma_i \nabla f_i(x) + \sum_{j \in A(x)} \alpha_j \nabla g_j(x) + \sum_{l=1}^{q} \beta_l \nabla h_l(x) = 0, \qquad (\gamma, \alpha, \beta) \neq (0, 0, 0).$$

The extended Mangasarian–Fromovitz conditions allow us to prove the following further characterization of a weakly efficient point of Problem (8).

Proposition 3.5 *Let $\mathcal{E}_w$ be the weakly efficient set of Problem (8), $x^* \in \mathcal{E}_w$ and $\mathcal{T}_{\mathcal{E}_w}(x^*)$ be any efficiency-describing collection of x^*. If the weak domination property holds for Problem (8) and the EMF holds at x^*, then, for all $I \in \mathcal{T}_{\mathcal{E}}(x^*)$, there exist multipliers $\sigma^* \in \mathbb{R}^m$, $\lambda^* \in \mathbb{R}^p$, $\mu^* \in \mathbb{R}^q$ such that*

$$\begin{aligned} \sum_{i \in I} \sigma_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \lambda_j^* \nabla g_j(x^*) + \sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) = 0,& \\ \sigma_i^* > 0, \qquad i \in I,& \\ \sigma_i^* = 0, \qquad i \notin I,& \\ \lambda_j^* \geq 0, \qquad j = 1, \dots, p,& \\ \lambda_j^* g_j(x^*) = 0, \qquad j = 1, \dots, p.& \end{aligned}$$

Proof From Proposition 3.2 we have that x^* satisfies (10a)–(10e). Assume by contradiction that $s \in I$ exists such that $\sigma_s = 0$. Then, (10a) would become

$$\sum_{\substack{i \in I \\ i \neq s}} \sigma_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \lambda_j^* \nabla g_j(x^*) + \sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) = 0$$

or equivalently, since $\lambda_j = 0,\ j \notin A(x^*)$,

$$\sum_{\substack{i \in I \\ i \neq s}} \sigma_i^* \nabla f_i(x^*) + \sum_{j \in A(x^*)} \lambda_j^* \nabla g_j(x^*) + \sum_{l=1}^{q} \mu_l^* \nabla h_l(x^*) = 0,$$

so that we get a contradiction to the EMF conditions. □

4 Conclusions

In this paper, we have analyzed necessary optimality conditions for continuous multiobjective problems. In particular, our goal is to extend the concept of Pareto stationarity, usually referred to a single point, to characterize a set of points. For both unconstrained and constrained problems, we tried to give a new, flexible and stronger characterization of the (weakly) efficient set, using first-order information and specific subsets of objective function indices. We introduce the definition of *efficiency-describing collection*, allowing us to define the concept of Pareto stationary set and show that a weakly stable set can be algorithmically improved if it is not a Pareto stationary set. In particular, this comes from a methodological interest recently emerged in the literature, see, e.g., [3, 5, 17]. The algorithms proposed in these works rely on the idea of improving, at every iteration, a set of points satisfying specific requirements that can be seen as a violation of our optimality conditions. For what concerns constrained problems, the analysis is driven taking into account different conditions on constraint qualification. Future work might be devoted to the use of the proposed characterization to devise local algorithms for continuous multiobjective optimization converging to Pareto stationary sets. As a final remark, we want to underline that finding a descent direction for all the objective functions becomes a harder and harder task as the number of objectives increases. As a consequence, in presence of many objective functions, the standard Pareto stationary conditions are easier to satisfy. Therefore, our results look particularly appealing and useful in the context of many-objective optimization.

Appendix

For reader's convenience, we report the theorems of the alternative used in the text.

Theorem A.1 (Gordan's Theorem [22]) *Let $A \in \mathbb{R}^{s\times n}$ be a matrix. One and only one of the following systems has solution:*

$$\begin{aligned} Az &< 0 \\ z &\in \mathbb{R}^n \end{aligned} \qquad\qquad \begin{aligned} A^T y &= 0 \\ y &\geq 0 \\ y &\neq 0 \\ y &\in \mathbb{R}^s \end{aligned}$$

Theorem A.2 (Slater's Theorem [22, 25]) *Let $C \in \mathbb{R}^{s_1\times n}$, $D \in \mathbb{R}^{s_2\times n}$ be two matrices. One and only one of the following systems has solution:*

$$\begin{aligned} Cy &= 0 \\ Dy &> 0 \\ y &\in \mathbb{R}^n \end{aligned} \qquad\qquad \begin{aligned} C^T u + D^T v &= 0 \\ v &\geq 0 \\ v &\neq 0 \\ u &\in \mathbb{R}^{s_1} \\ v &\in \mathbb{R}^{s_2} \end{aligned}$$

Theorem A.3 (Motzkin's Theorem [22]) *Let $A \in \mathbb{R}^{s_1 \times n}$, $C \in \mathbb{R}^{s_2 \times n}$, $D \in \mathbb{R}^{s_3 \times n}$ be three matrices, with A non-vacuous. One and only one of the following systems has solution:*

$$\begin{aligned} Ay &> 0 \\ Cy &\geq 0 \\ Dy &= 0 \\ y &\in \mathbb{R}^n \end{aligned} \qquad\qquad \begin{aligned} A^T u + C^T v + D^T z &= 0 \\ u &\geq 0 \\ u &\neq 0 \\ v &\geq 0 \\ u &\in \mathbb{R}^{s_1} \\ v &\in \mathbb{R}^{s_2} \\ z &\in \mathbb{R}^{s_3} \end{aligned}$$

Acknowledgements The authors are very grateful to the anonymous referees for the careful reading of the manuscript and their valuable comments.

Funding Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement.

Data Availibility Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms. Wiley, London (2006)
2. Bigi, G.: Optimality and Lagrangian regularity in vector optimization. Ph. D. Thesis. Dipartimento di Matematica, Università di Pisa (1999)
3. Cocchi, G., Lapucci, M.: An augmented Lagrangian algorithm for multi-objective optimization. Comput. Optim. Appl. **77**(1), 29–56 (2020). https://doi.org/10.1007/s10589-020-00204-z
4. Cocchi, G., Lapucci, M., Mansueto, P.: Pareto front approximation through a multi-objective augmented Lagrangian method. EURO J. Comput. Optim. **9**, 100,008 (2021). https://doi.org/10.1016/j.ejco.2021.100008
5. Cocchi, G., Liuzzi, G., Lucidi, S., Sciandrone, M.: On the convergence of steepest descent methods for multiobjective optimization. Comput. Optim. Appl. **77**, 1–27 (2020). https://doi.org/10.1007/s10589-020-00192-0
6. Custódio, A.L., Madeira, J.A., Vaz, A.I.F., Vicente, L.N.: Direct multisearch for multiobjective optimization. SIAM J. Optim. **21**(3), 1109–1140 (2011). https://doi.org/10.1137/10079731X
7. Da Cunha, N., Polak, E.: Constrained minimization under vector-valued criteria in finite dimensional spaces. J. Math. Anal. Appl. **19**(1), 103–124 (1967). https://doi.org/10.1016/0022-247X(67)90025-X
8. Eichfelder, G., Kirst, P., Meng, L., Stein, O.: A general branch-and-bound framework for continuous global multiobjective optimization. J. Global Optim. **80**, 195–227 (2021). https://doi.org/10.1007/s10898-020-00984-y

9. Eichfelder, G., Stein, O.: Limit sets in global multiobjective optimization. Optimization (2022). https://doi.org/10.1080/02331934.2022.2092479
10. Eichfelder, G., Stein, O., Warnow, L.: A solver for multiobjective mixed-integer convex and nonconvex optimization. J. Optim. Theory Appl. (2023). https://doi.org/10.1007/s10957-023-02285-2
11. Eichfelder, G., Warnow, L.: An approximation algorithm for multi-objective optimization problems using a box-coverage. J. Global Optim. **83**(2), 329–357 (2022). https://doi.org/10.1007/s10898-021-01109-9
12. Fliege, J., Drummond, L.G., Svaiter, B.F.: Newton's method for multiobjective optimization. SIAM J. Optim. **20**(2), 602–626 (2009). https://doi.org/10.1137/08071692X
13. Fliege, J., Svaiter, B.F.: Steepest descent methods for multicriteria optimization. Math. Methods Oper. Res. **51**, 479–494 (2000). https://doi.org/10.1007/s001860000043
14. Fukuda, E.H., Drummond, L.G., Raupp, F.M.: An external penalty-type method for multicriteria. TOP **24**, 493–513 (2016). https://doi.org/10.1007/s11750-015-0406-8
15. Fukuda, E.H., Drummond, L.G., Raupp, F.M.: A barrier-type method for multiobjective optimization. Optimization **69**(11), 2471–2487 (2020). https://doi.org/10.1080/02331934.2019.1576667
16. Gonçalves, M., Lima, F., Prudente, L.: Globally convergent Newton-type methods for multiobjective optimization. Comput. Optim. Appl. **83**(2), 403–434 (2022). https://doi.org/10.1007/s10589-022-00414-7
17. Lapucci, M., Mansueto, P.: Improved front steepest descent for multi-objective optimization. Oper. Res. Lett. **51**(3), 242–247 (2023). https://doi.org/10.1016/j.orl.2023.03.001
18. Lapucci, M., Mansueto, P.: A limited memory quasi-Newton approach for multi-objective optimization. Comput. Optim. Appl. **85**(1), 33–73 (2023). https://doi.org/10.1007/s10589-023-00454-7
19. Liuzzi, G., Lucidi, S., Rinaldi, F.: A derivative-free approach to constrained multiobjective nonsmooth optimization. SIAM J. Optim. **26**(4), 2744–2774 (2016). https://doi.org/10.1137/15M1037810
20. Luc, D.T.: Theory of vector optimization. Springer, Berlin, Heidelberg (1989). https://doi.org/10.1007/978-3-642-50280-4
21. Maciel, M.C., Santos, S.A., Sottosanto, G.N.: Regularity conditions in differentiable vector optimization revisited. J. Optim. Theory Appl. **142**(2), 385–398 (2009). https://doi.org/10.1007/s10957-009-9519-2
22. Mangasarian, O.L.: Nonlinear Programming. SIAM, New Delhi (1994)
23. Miettinen, K.: Nonlinear Multiobjective Optimization, vol. 12. Springer, Berlin (1999)
24. Ruzika, S., Wiecek, M.M.: Approximation methods in multiobjective programming. J. Optim. Theory Appl. **126**(3), 473–501 (2005). https://doi.org/10.1007/s10957-005-5494-4
25. Slater, M.L.: A note on Motzkin's transposition theorem. Econometrica (1951). https://doi.org/10.2307/1905734
26. Stein, O., Volk, M.: Generalized polarity and weakest constraint qualifications in multiobjective optimization. J. Optim. Theory Appl. **198**, 1156–1190 (2023). https://doi.org/10.1007/s10957-023-02256-7