# The projected gradient direction in vector optimization: continuity and dual characterization

N. Fazzio[1], L. F. Prudente[2*], M. L. Schuverdt[1]

[1]Department of Mathematics, FCE, University of La Plata, La Plata, Bs. As., Argentina.

[2]Institute of Mathematics and Statistics, Federal University of Goiás, Goiânia, GO, Brazil.

*Corresponding author(s). E-mail(s): lfprudente@ufg.br;
Contributing authors: nfazzio@mate.unlp.edu.ar;
schuverd@mate.unlp.edu.ar;

**Abstract**

This technical note proves that, for a smooth vector optimization problem on a closed convex feasible set ordered by a pointed cone, the projected gradient direction depends continuously on the decision variable. Our argument is based on a simple and direct proof via a fixed-domain reformulation of the subproblem. We then give a necessary and sufficient dual characterization of this direction and show that its associated set-valued dual variable mapping is outer semicontinuous.

**Keywords:** Vector optimization, multiobjective optimization, projected gradient, continuity, dual characterization.

## 1 Introduction

We study the vector optimization problem on a nonempty closed convex set $\Omega \subset \mathbb{R}^n$ endowed with a partial order induced by a closed, convex, pointed cone with non-empty interior $K \subset \mathbb{R}^m$:

$$\min_K \{ F(x) \mid x \in \Omega \}, \tag{1}$$

where $F\colon \mathbb{R}^n \to \mathbb{R}^m$, $F(x) := (f_1(x), \dots, f_m(x))$ is a continuously differentiable function and $u \preceq_K v$ if and only if $v - u \in K$. When $K$ is the nonnegative orthant

$\mathbb{R}^m_+$, the induced order is the usual componentwise order, and this recovers the classical multiobjective optimization problem. This kind of problem arises in a wide range of applications, including engineering, economics, finance, management science, and medical treatment planning such as radiotherapy (see, e.g., [6, 37]).

A fruitful research line over the last two decades extends classical scalar algorithms directly to vector-valued objectives rather than relying on scalarizations [13, 30] or heuristics [28]. For unconstrained problems, Fliege and Svaiter first introduced a steepest-descent method in [16], and subsequently Graña-Drummond and Svaiter generalized that approach to vector optimization with a general ordering cone in [24]. Several other extensions have since been proposed in the unconstrained setting, including Newton-type [11, 18, 22, 25, 38], quasi-Newton [27, 32–34], conjugate gradient [21, 29], subgradient [2], and proximal algorithms [7]. For constrained problems, the projected gradient method, introduced by Graña-Drummond and Iusem in [23], has played a central role. It has become the basis for several important developments, including convergence analyses [3, 19, 40], inexact and nonmonotone variants [14, 20, 41, 42], and Barzilai–Borwein adaptations [9, 10, 31, 39]. Beyond projected gradient schemes, other approaches have also been investigated, such as conditional gradient methods [1], interior-point [15] and augmented Lagrangian schemes [8, 12], and SQP algorithms [17].

At each $x \in \Omega$, the projected gradient direction is defined as the solution of a strongly convex scalar subproblem. While its optimal value is known to be continuous [19, Proposition 3.4] (and serves as a standard merit function in convergence analyses), the continuity of the direction itself had not been formally established. Although this property can be regarded as an instance of the classical *Maximum Theorem* of Berge [4] and, more directly, of Hogan's formulation [26, Corollary 8.1], no explicit verification had been carried out in the projected gradient setting. We close this gap by presenting a simple and direct proof, based on a fixed-domain reformulation of the subproblem, making the result both rigorous and accessible to a broad audience. In addition, we provide a necessary and sufficient dual characterization of the projected gradient direction and prove that the associated set-valued multiplier mapping is outer semicontinuous. Since many advanced methods build on the projected gradient framework, our results not only reinforce the foundational convergence theory, but also provide the rigorous basis needed to extend convergence and sensitivity analyses to these more sophisticated algorithms.

This paper is organized as follows. In Section 2, we recall basic concepts, including the definition of the projected gradient direction. Section 3 reformulates the projected gradient subproblem in a fixed domain and proves the continuity of this direction with respect to the decision variable. In Section 4, we derive a necessary and sufficient dual characterization and establish the outer semicontinuity of the associated multiplier mapping. Finally, Section 5 contains concluding remarks.

**Notation:** $\langle \cdot, \cdot \rangle$ is the usual inner product, and $\| \cdot \|$ is the Euclidean norm. Given $x \in \mathbb{R}^n$, $JF(x)$ denotes the Jacobian of $F$ at $x$. $P_\Omega(x)$ will denote the projection of $x \in \mathbb{R}^n$ onto $\Omega$, i.e., $P_\Omega(x) \coloneqq \operatorname{argmin}_{y \in \Omega} \|x - y\|$. If $S \subset \mathbb{R}^m$, then the conic hull and the convex hull of $S$ are denoted by $\operatorname{cone}(S)$ and $\operatorname{conv}(S)$, respectively. We denote by $\Delta_r$ the $r$-dimensional simplex, i.e., $\Delta_r \coloneqq \{\lambda \in \mathbb{R}^r \mid \lambda_j \geq 0, \sum_j \lambda_j = 1\}$.

## 2 Basic concepts

Let $K$ be a cone as in the introduction and denote its positive polar cone by $K^* := \{w \in \mathbb{R}^m \mid \langle w, y\rangle \geq 0,\ \forall y \in K\}$. Let $C \subset K^* \setminus \{0\}$ be a compact set such that $K^* = \text{cone}(\text{conv}(C))$. We can take, for example, $C = \{w \in K^* \mid \|w\| = 1\}$. Assuming that $K$ is polyhedral, we can choose $C$ as the finite set of extreme rays of $K^*$. In particular, for the multiobjective optimization case where $K = \mathbb{R}^m_+$, $C$ can be taken as the canonical basis of $\mathbb{R}^m$.

Define the functions $\mathcal{D}\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and $\mathcal{A}\colon \mathbb{R}^n \times \mathbb{R}^n \rightrightarrows K^*$ by

$$\mathcal{D}(x, d) := \max\{\, \langle JF(x)d, w\rangle \mid w \in C\,\}, \tag{2}$$

and

$$\mathcal{A}(x, d) := \{\, w \in C \mid \langle JF(x)d, w\rangle = \mathcal{D}(x, d)\,\}, \tag{3}$$

i.e., $\mathcal{A}(x, d)$ is the set of maximizers in the definition of $\mathcal{D}$.

**Lemma 1** *Let $\mathcal{D}\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and $\mathcal{A}\colon \mathbb{R}^n \times \mathbb{R}^n \rightrightarrows K^*$ be defined as in (2) and (3), respectively. Then the following properties hold:*

*(a) The mapping $\mathcal{D}(\cdot, \cdot)$ is continuous.*
*(b) For any fixed $x \in \mathbb{R}^n$, the subdifferential of the function $d \mapsto \mathcal{D}(x, d)$ is given by*

$$\partial\mathcal{D}(x, d) = JF(x)^\top \operatorname{conv} \mathcal{A}(x, d), \quad \forall d \in \mathbb{R}^n.$$

*(c) Let $x, d \in \mathbb{R}^n$ and $w \in \operatorname{conv} \mathcal{A}(x, d)$. Then, $w$ has a finite representation $w = \sum_{j=1}^{r} \lambda_j z_j$ with $z_j \in \mathcal{A}(x, d)$, $\lambda := (\lambda_1, \ldots, \lambda_r) \in \Delta_r$, and $r \leq m + 1$.*

*Proof* Item *(a)* follows from [24]. Indeed, the continuity of $\mathcal{D}(x, d)$ follows from it being the maximum of continuous affine functions over the compact set $C$. Item *(b)* is a direct consequence of Danskin's Theorem, see [5, Proposition 4.5.1]. Item *(c)* is a direct consequence of Carathéodory's Theorem, see [5, Proposition 1.3.1]. □

We now recall the *projected gradient direction* associated with (1), see [23]. Given $x \in \Omega$, consider the following constrained scalar-valued minimization problem:

$$\min \left\{ \mathcal{D}(x, d) + \tfrac{1}{2}\|d\|^2 \;\middle|\; d \in \Omega - x \right\}, \tag{4}$$

where $\Omega - x := \{p - x \mid p \in \Omega\}$. The *projected gradient direction* for $F$ at $x \in \Omega$ is defined by the unique optimal solution of (4), namely

$$v(x) := \operatorname{argmin} \left\{ \mathcal{D}(x, d) + \tfrac{1}{2}\|d\|^2 \;\middle|\; d \in \Omega - x \right\}. \tag{5}$$

Note that, since the objective function of (4) is strongly convex and $\Omega$ is a closed and convex set, $v(x)$ is well defined for every $x \in \Omega$. The following proposition will be useful to provide a necessary and sufficient condition for the optimality of (4).

**Proposition 2** *[5, Proposition 4.7.2] Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a convex function. A vector $x^*$ minimizes $\varphi$ over a convex set $X \subset \mathbb{R}^n$ if and only if there exists a subgradient $q \in \partial\varphi(x^*)$ such that $q^\top(x - x^*) \geq 0$, for all $x \in X$.*

A well known characterization of the Euclidean projector is as follows.

**Lemma 3** *[5, Proposition 2.2.1] For every $x \in \mathbb{R}^n$, a vector $z \in \Omega$ is equal to $P_\Omega(x)$ if and only if $(y - z)^\top(x - z) \leq 0$ for all $y \in \Omega$.*

The following elementary lemma ensures that a bounded sequence with a unique limit point actually converges. Since the result is standard (see, for instance, [36, Chapter 3]), we omit its proof.

**Lemma 4** *Let $\{u^k\} \subset \mathbb{R}^n$ be a bounded sequence. Assume that every convergent subsequence of $\{u^k\}$ has the same limit point $u^* \in \mathbb{R}^n$. Then, the entire sequence $\{u^k\}$ converges to $u^*$.*

We finish this section by recalling a key concept regarding set-valued functions [35, Chapter 5].

**Definition 1** *A set-valued mapping $W : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is* outer semicontinuous *at $x^* \in \operatorname{dom} W \coloneqq \{x \in \mathbb{R}^n \mid W(x) \neq \emptyset\}$ if*

$$\limsup_{x \to x^*} W(x) \subset W(x^*),$$

*where*

$$\limsup_{x \to x^*} W(x) \coloneqq \Big\{\lambda \in \mathbb{R}^m \;\Big|\; \exists\, x^k \to x^* \text{ and } \exists\, \lambda^k \to \lambda \text{ with } \lambda^k \in W(x^k)\Big\}. \tag{6}$$

*If $W$ is outer semicontinuous at every point in a set $\Omega \subset \operatorname{dom} W$, then it is said to be outer semicontinuous on $\Omega$.*

# 3 Continuity of the projected gradient direction

To establish the continuity of the projected gradient direction, it is convenient to reformulate problem (4) in terms of the translated variable $p \coloneqq x + d$. This yields

$$\min \left\{ \mathcal{D}(x, p - x) + \tfrac{1}{2}\|p - x\|^2 \;\middle|\; p \in \Omega \right\}. \tag{7}$$

Since the objective function in (7) is strongly convex and the feasible set $\Omega$ is closed and convex, this problem has a unique minimizer. Accordingly, we define

$$u(x) \coloneqq \operatorname{argmin} \left\{ \mathcal{D}(x, p - x) + \tfrac{1}{2}\|p - x\|^2 \;\middle|\; p \in \Omega \right\}, \tag{8}$$

so the projected gradient direction is given by

$$v(x) = u(x) - x. \tag{9}$$

Before proceeding, we record a boundedness property of $u(\cdot)$.

**Lemma 5** *The mapping $u\colon \Omega \to \Omega$ given by (8) is bounded on compact subsets of $\Omega$.*

*Proof* The result follows directly from the boundedness of $v(x)$ on compact subsets, established in [20, Proposition 2.5], since $u(x) = x + v(x)$ and $x$ ranges over a compact set. □

The next theorem establishes the continuity of $v(\cdot)$.

**Theorem 6** *The projected gradient direction mapping $v\colon \Omega \to \mathbb{R}^n$ defined in (5) is continuous on $\Omega$.*

*Proof* From (9), it suffices to prove that $u(\cdot)$ from (8) is continuous. Assume $\{x^k\} \subset \Omega$ is such that $x^k \to x^* \in \Omega$. We must show $u(x^k) \to u(x^*)$. Since $\{x^k\}$ is bounded, by Lemma 5, the sequence $\{u(x^k)\}$ is also bounded. Therefore, it has convergent subsequences. Pick any convergent subsequence $\{u(x^{k_\ell})\}$ and denote its limit by $\bar{u}$. By definition of $u(x^{k_\ell})$, we know that

$$\mathcal{D}(x^{k_\ell}, u(x^{k_\ell}) - x^{k_\ell}) + \tfrac{1}{2} \left\| u(x^{k_\ell}) - x^{k_\ell} \right\|^2 \leq \mathcal{D}(x^{k_\ell}, p - x^{k_\ell}) + \tfrac{1}{2} \left\| p - x^{k_\ell} \right\|^2, \quad \forall p \in \Omega.$$

Choosing $p = u(x^*)$ and letting $\ell \to \infty$ in the above inequality, by the continuity of $\mathcal{D}$ (Lemma 1 *(a)*), we find

$$\mathcal{D}(x^*, \bar{u} - x^*) + \tfrac{1}{2} \left\| \bar{u} - x^* \right\|^2 \leq \mathcal{D}(x^*, u(x^*) - x^*) + \tfrac{1}{2} \left\| u(x^*) - x^* \right\|^2.$$

On the other hand, by the definition of $u(x^*)$ as the minimizer in (8) implies

$$\mathcal{D}(x^*, u(x^*) - x^*) + \tfrac{1}{2} \left\| u(x^*) - x^* \right\|^2 \leq \mathcal{D}(x^*, p - x^*) + \tfrac{1}{2} \left\| p - x^* \right\|^2, \quad \forall p \in \Omega.$$

Combining the two inequalities, we deduce that $\bar{u}$ is also an optimal solution for the problem (7) with $x = x^*$. By uniqueness, $\bar{u} = u(x^*)$. Hence *every* convergent subsequence of $\{u(x^k)\}$ has the same limit $u(x^*)$. By Lemma 4, this implies that the whole sequence $\{u(x^k)\}$ converges to $u(x^*)$. This shows that $u(\cdot)$ is continuous, and thus $v(\cdot)$ is also continuous in view of (9). □

As an immediate corollary, we obtain the continuity of the projected gradient merit function, providing an alternative proof of [19, Proposition 3.4].

**Corollary 7** *The optimal value mapping $\theta : \Omega \to \mathbb{R}$, defined by*

$$\theta(x) := \mathcal{D}(x, v(x)) + \tfrac{1}{2}\|v(x)\|^2,$$

*is continuous.*

*Proof* This follows directly from Theorem 6, which establishes the continuity of $v(\cdot)$, together with the continuity of $\mathcal{D}(\cdot, \cdot)$ (Lemma 1 *(a)*). □

**Remark 1** *Although the continuity of $v(\cdot)$ could also be deduced from general maximum-theorem results for parametric minimization (see, e.g., Hogan's Corollary 8.1 [26]), the fixed-domain reformulation used above allows for a direct and elementary proof that avoids the machinery of point-to-set mappings and is often easier to adapt to related contexts.*

# 4 Dual characterization and semicontinuity of the dual variables

We now provide a necessary and sufficient dual characterization of the projected gradient direction $v(x)$ in the vector optimization setting. This not only clarifies its structure but also enables the analysis of continuity properties of the associated dual variables.

**Theorem 8** *Let $x \in \Omega$. Then $v \in \mathbb{R}^n$ is the unique solution of (4) (i.e., $v = v(x)$) if and only if there exists $\bar{w} \in \operatorname{conv} \mathcal{A}(x, v)$ such that*

$$v = P_\Omega\big(x - JF(x)^\top \bar{w}\big) - x.$$

*Proof* Fix $x \in \Omega$ and set $\varphi(d) := \mathcal{D}(x, d) + \frac{1}{2}\|d\|^2$. By Proposition 2, a vector $v \in \mathbb{R}^n$ is the unique minimizer of $\varphi$ on $\Omega - x$ if and only if there exists a subgradient $q \in \partial\varphi(v)$ such that

$$q^\top (d - v) \geq 0, \quad \forall d \in \Omega - x. \tag{10}$$

From Lemma 1 *(b)*, we know that every subgradient of $\varphi(v)$ can be written as $q = v + JF(x)^\top \bar{w}$ for some $\bar{w} \in \operatorname{conv} \mathcal{A}(x, v)$. Substituting this expression for $q$ into (10) and letting $p := x + d$, we obtain

$$\big(x - JF(x)^\top \bar{w} - (x + v)\big)^\top \big(p - (x + v)\big) \leq 0, \quad \forall p \in \Omega.$$

Lemma 3 shows that this inequality is equivalent to

$$v = P_\Omega\big(x - JF(x)^\top \bar{w}\big) - x.$$

Since each step is a two-way equivalence, the conclusion holds in both directions, and the proof is complete. □

Theorem 8 refines [20, Proposition 4.1], which proved only the forward implication and produced $\bar{w} \in \operatorname{conv}(C)$. Here we obtain an "if and only if" statement and show that $\bar{w}$ can in fact be chosen in the smaller set $\operatorname{conv} \mathcal{A}(x, v)$. Moreover, the theorem shows that such a vector $\bar{w}$ can be interpreted as an optimal dual multiplier associated with the scalar convex subproblem defining the projected gradient direction. It also extends the result of [42, Section 5], valid for the particular case $K = \mathbb{R}^m_+$, to an arbitrary ordering cone.

To further clarify the meaning of the dual vector $\bar{w}$ in Theorem 8, we next show that in the multiobjective case it admits a natural interpretation as a Lagrange multiplier.

**Remark 2** *In the multiobjective optimization case, the vector $\bar{w}$ in Theorem 8 admits a direct interpretation as a Lagrange multiplier. Taking $C$ as the canonical bases of $\mathbb{R}^m$, we get* $\operatorname{conv} \mathcal{A}(x, v(x)) \subset \Delta_m$ *and hence $\bar{w} \in \Delta_m$. By introducing an auxiliary variable $\tau \in \mathbb{R}$ and rewriting (4) as*

$$\min_{\tau, d} \; \tau + \tfrac{1}{2}\|d\|^2 \quad \text{s.t.} \quad \nabla f_j(x)^\top d \leq \tau \; (j \in \{1, \ldots, m\}), \; d \in \Omega - x,$$

*one sees that the components of $\bar{w}$ are precisely the KKT multipliers associated with the constraints $\nabla f_j(x)^\top d \leq \tau$. Thus $\bar{w}$ is nothing but the Lagrange–multiplier vector for the scalar reformulation of (4).*

Theorem 8 also suggests a connection with the unconstrained steepest descent direction in the multiobjective case, as detailed in the next remark.

**Remark 3** *Consider the multiobjective case ($K = \mathbb{R}^m_+$, $C$ the canonical basis). The multipliers defining the unconstrained steepest descent direction can be obtained from the dual problem*

$$\min\Big\{ \tfrac{1}{2}\|\sum_{j=1}^{m} \lambda_j \nabla f_j(x)\|^2 \ \Big|\ \lambda \in \Delta_m \Big\}, \tag{11}$$

*see [16]. Its solution $\lambda \in \Delta_m$ determines the direction*

$$d_{\mathrm{SD}}(x) = -\sum_{j=1}^{m} \lambda_j \nabla f_j(x),$$

*where the multipliers $\lambda_j$ are positive only for those indices $j$ such that $\nabla f_j(x)^\top d_{\mathrm{SD}}(x) = \mathcal{D}\big(x, d_{\mathrm{SD}}(x)\big)$. Define the trial vector*

$$v_{\mathrm{trial}}(x) \coloneqq P_\Omega\big(x + d_{\mathrm{SD}}(x)\big) - x.$$

*If every index $j$ with $\lambda_j > 0$ corresponds to an element of the active set $\mathcal{A}\big(x, v_{\mathrm{trial}}(x)\big)$, then $\lambda \in \operatorname{conv} \mathcal{A}\big(x, v_{\mathrm{trial}}(x)\big)$ and, by Theorem 8,*

$$v(x) = v_{\mathrm{trial}}(x).$$

*In particular, if $x + d_{\mathrm{SD}}(x) \in \Omega$, then $v_{\mathrm{trial}}(x) = d_{\mathrm{SD}}(x)$ and this condition trivially holds, implying that $v(x) = d_{\mathrm{SD}}(x)$. Thus, when the active-set condition is satisfied, one can recover $v(x)$ from a single projection together with a check of the active constraints. For small values of $m$, this verification is inexpensive: the steepest descent direction is cheap to compute, since it is obtained from a small dual problem (11), which even admits a closed-form solution in the biobjective case, and the set $\mathcal{A}(x, v_{\mathrm{trial}}(x))$ is obtained from a single matrix-vector product. Hence, this shortcut can be applied in practice with negligible additional computational cost. The same reasoning applies whenever the ordering cone $K$ is finitely generated. Its positive polar $K^*$ is then also finitely generated, and taking $C$ equal to its set of extreme rays leads to an active-set test and a shortcut for computing $v(x)$ that are directly analogous to those described above.*

Define the set-valued mapping $W \colon \Omega \rightrightarrows K^*$ by

$$W(x) \coloneqq \Big\{ w \in \operatorname{conv} \mathcal{A}(x, v(x)) \ \Big|\ v(x) = P_\Omega\big(x - JF(x)^\top w\big) - x \Big\}. \tag{12}$$

Given $x \in \Omega$, the set $W(x)$ is exactly the collection of all dual variable vectors associated to $v(x)$. Clearly, $W(x)$ is nonempty and compact. We now prove its outer semicontinuity.

**Theorem 9** *The set-valued mapping $W \colon \Omega \rightrightarrows K^*$ given by (12) is outer semicontinuous on $\Omega$.*

*Proof* Let $x^* \in \Omega$ be arbitrary and take any $w^* \in \limsup_{x\to x^*} W(x)$. By (6), there are sequences $x^k \in \Omega$ with $x^k \to x^*$ and $w^k \in W(x^k)$ with $w^k \to w^*$. By the definition of $W(\cdot)$ in (12), for each $k$, we have $w^k \in \operatorname{conv} \mathcal{A}(x^k, v(x^k))$ and

$$v(x^k) = P_\Omega\big(x^k - JF(x^k)^\top w^k\big) - x^k. \tag{13}$$

By the continuity of $JF(\cdot)$, $v(\cdot)$ (see Theorem 6), and the projection operator $P_\Omega(\cdot)$, taking the limit as $k \to \infty$ in (13) gives

$$v(x^*) = P_\Omega\big(x^* - JF(x^*)^\top w^*\big) - x^*.$$

To finish the proof, we must show that $w^* \in \operatorname{conv} \mathcal{A}(x^*, v(x^*))$. Recalling that $w^k \in \operatorname{conv} \mathcal{A}(x^k, v(x^k))$, by Lemma 1 *(c)*, each $w^k$ admits a representation with at most $m+1$ terms:

$$w^k = \sum_{j=1}^{m+1} \lambda_j^k z_j^k \quad \text{with} \quad \lambda^k \in \Delta_{m+1} \text{ and } z_j^k \in \mathcal{A}(x^k, v(x^k)) \subset C.$$

Hence, for every $j$ and $k$, we obtain

$$\langle z_j^k, JF(x^k)v(x^k)\rangle = \mathcal{D}(x^k, v(x^k)), \tag{14}$$

Because $C$ and $\Delta_{m+1}$ are compact, we may assume (taking a subsequence if necessary) that $z_j^k \to z_j^* \in C$ and $\lambda_j^k \to \lambda_j^*$ with $\lambda_j^* \geq 0$ and $\sum_{j=1}^{m+1} \lambda_j^* = 1$. Taking the limit as $k \to \infty$ in (14) and using the continuity of $JF(\cdot)$, $v(\cdot)$ and $\mathcal{D}(\cdot,\cdot)$, we have

$$\langle z_j^*, JF(x^*)v(x^*)\rangle = \mathcal{D}(x^*, v(x^*)),$$

hence $z_j^* \in \mathcal{A}(x^*, v(x^*))$ for each $j \in \{1, \dots, m+1\}$. Therefore,

$$w^* = \sum_{j=1}^{m+1} \lambda_j^* z_j^* \in \operatorname{conv} \mathcal{A}(x^*, v(x^*)),$$

which shows $w^* \in W(x^*)$ and proves $\limsup_{x\to x^*} W(x) \subset W(x^*)$. Thus, $W$ is outer semicontinuous on $\Omega$. □

This result has a natural algorithmic interpretation. If a projected gradient method generates a primal–dual sequence $\{(x^k, w^k)\}$ with $w^k \in W(x^k)$ and a subsequence $\{x^{k_\ell}\}$ converges to some $x^* \in \Omega$, then every limit point of $\{w^{k_\ell}\}$ necessarily belongs to $W(x^*)$. In other words, the dual variables remain consistent in the limit, so that accumulation points of the dual sequence are valid multipliers associated with the corresponding primal limit point.

# 5 Final remarks

This note establishes three contributions to the theory of projected gradient methods in vector optimization. First, we proved the continuity of the projected gradient direction through a direct and elementary argument. Second, we obtained a necessary and sufficient dual characterization, extending known results from the multiobjective case to arbitrary cones. Finally, we proved the outer semicontinuity of the multiplier mapping, showing that for any convergent subsequence of projected gradient iterates, the associated dual sequence has limit points that are valid multipliers for the primal limit.

Beyond their intrinsic theoretical interest, these properties have direct algorithmic implications. The continuity of $v(\cdot)$ helps prevent instabilities that may arise from numerical perturbations or errors in the computation of the projected gradient direction, which is especially relevant in inexact variants where the direction is computed only approximately. In particular, continuity guarantees that small perturbations in the approximate direction do not compromise the descent condition. Moreover, it is also relevant for the analysis of line search procedures in projected gradient schemes, including nonmonotone variants: small variations in the accepted step size lead to nearby iterates, and continuity ensures that the corresponding projected gradient directions remain close as well. Likewise, the outer semicontinuity of the set-valued mapping $W(\cdot)$, which associates dual variable vectors to the projected gradient direction at each primal iterate, plays an important theoretical role in the convergence analysis of projected gradient-type methods and ensures stability in the behavior of dual sequences. Consequently, accumulation points of the dual variables remain valid multipliers corresponding to the primal limit, preventing erratic dual behavior and supporting convergence guarantees for algorithms that update both primal and dual variables simultaneously. Altogether, these results reinforce the theoretical foundations of projected gradient algorithms and support the development of more advanced schemes in vector optimization.

# Declarations

**Funding**

L. F. Prudente was partially supported by CNPq, Grants 307294/2023-4 and 403197/2025-2. M. L. Schuverdt was partially supported by CONICET.

**Conflicts of interest**

The authors declare that they have no conflict of interest.

# References

[1] P. B. Assunção, O. P. Ferreira, and L. F. Prudente. Conditional gradient method for multiobjective optimization. *Comput. Optim. Appl.*, 78(3):741–768, 2021.

[2] J. Y. Bello Cruz. A subgradient method for vector optimization problems. *SIAM J. Optim.*, 23(4):2169–2182, 2013.

[3] J. Y. Bello Cruz, L. R. Lucambio Pérez, and J. G. Melo. Convergence of the projected gradient method for quasiconvex multiobjective optimization. *Nonlinear Anal.*, 74(16):5268–5273, 2011.

[4] C. Berge. *Topological Spaces: Including a Treatment of Multi-valued Functions, Vector Spaces and Convexity.* Macmillan, New York, 1963.

[5] D. Bertsekas, A. Nedic, and A. Ozdaglar. *Convex Analysis and Optimization.* Athena Scientific optimization and computation series. Athena Scientific, 2003.

[6] V. Bhaskar, S. K. Gupta, and A. K. Ray. Applications of multiobjective optimization in chemical engineering. *Rev. Chem. Eng.*, 16(1):1–54, 2000.

[7] H. Bonnel, A. N. Iusem, and B. F. Svaiter. Proximal methods in vector optimization. *SIAM J. Optim.*, 15(4):953–970, 2005.

[8] G. A. Carrizo, N. S. Fazzio, M. D. Sánchez, and M. L. Schuverdt. Scaled-PAKKT sequential optimality condition for multiobjective problems and its application to an augmented Lagrangian method. *Comput. Optim. Appl.*, 89(3):769–803, 2024.

[9] G. A. Carrizo, N. S. Fazzio, and M. L. Schuverdt. A nonmonotone projected gradient method for multiobjective problems on convex sets. *J. Oper. Res. Soc. China*, 12(2):410–427, 2024.

[10] J. Chen, L. Tang, and X. Yang. A Barzilai-Borwein descent method for multiobjective optimization problems. *Eur. J. Oper. Res.*, 311(1):196–209, 2023.

[11] T. D. Chuong. Newton-like methods for efficient solutions in vector optimization. *Comput. Optim. Appl.*, 54(3):495–516, 2013.

[12] G. Cocchi and M. Lapucci. An augmented Lagrangian algorithm for multi-objective optimization. *Comput. Optim. Appl.*, 77(1):29–56, 2020.

[13] G. Eichfelder. *Adaptive Scalarization Methods in Multiobjective Optimization.* Springer Berlin Heidelberg, 2008.

[14] N. S. Fazzio and M. L. Schuverdt. Convergence analysis of a nonmonotone projected gradient method for multiobjective optimization problems. *Optim. Lett.*, 13(6):1365–1379, 2019.

[15] J. Fliege. An efficient interior-point method for convex multicriteria optimization problems. *Math. Oper. Res.*, 31(4):825–845, 2006.

[16] J. Fliege and B. F. Svaiter. Steepest descent methods for multicriteria optimization. *Math. Methods Oper. Res.*, 51(3):479–494, 2000.

[17] J. Fliege and A. I. F. Vaz. A method for constrained multiobjective optimization based on SQP techniques. *SIAM J. Optim.*, 26(4):2091–2119, 2016.

[18] J. Fliege, L. M. Graña Drummond, and B. F. Svaiter. Newton's method for multiobjective optimization. *SIAM J. Optim.*, 20(2):602–626, 2009.

[19] E. H. Fukuda and L. M. Graña Drummond. On the convergence of the projected gradient method for vector optimization. *Optimization*, 60(8-9):1009–1021, 2011.

[20] E. H. Fukuda and L. M. Graña Drummond. Inexact projected gradient method for vector optimization. *Comput. Optim. Appl.*, 54(3):473–493, 2013.

[21] M. L. N. Gonçalves and L. F. Prudente. On the extension of the Hager–Zhang conjugate gradient method for vector optimization. *Comput. Optim. Appl.*, 76 (3):889–916, 2020.

[22] M. L. N. Gonçalves, F. S. Lima, and L. F. Prudente. Globally convergent Newton-type methods for multiobjective optimization. *Comput. Optim. Appl.*, 83(2): 403–434, 2022.

[23] L. M. Graña Drummond and A. N. Iusem. A projected gradient method for vector optimization problems. *Comput. Optim. Appl.*, 28(1):5–29, 2004.

[24] L. M. Graña Drummond and B. F. Svaiter. A steepest descent method for vector optimization. *J. Comput. Appl. Math.*, 175(2):395–414, 2005.

[25] L. M. Graña Drummond, F. M. P. Raupp, and B. F. Svaiter. A quadratically convergent Newton method for vector optimization. *Optimization*, 63(5):661–677, 2014.

[26] W. W. Hogan. Point-to-set maps in mathematical programming. *SIAM Rev.*, 15 (3):591–603, 1973.

[27] M. Lapucci and P. Mansueto. A limited memory quasi-Newton approach for multi-objective optimization. *Comput. Optim. Appl.*, 85(1):33–73, 2023.

[28] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler. Combining convergence and diversity in evolutionary multiobjective optimization. *Evol. Comput.*, 10(3):263–282, 2002.

[29] L. R. Lucambio Pérez and L. F. Prudente. Nonlinear conjugate gradient methods for vector optimization. *SIAM J. Optim.*, 28(3):2690–2720, 2018.

[30] K. Miettinen. *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media, 1999.

[31] V. Morovati, L. Pourkarimi, and H. Basirzadeh. Barzilai and Borwein's method for multiobjective optimization problems. *Numer. Algorithms*, 72(3):539–604, 2016.

[32] V. Morovati, H. Basirzadeh, and L. Pourkarimi. Quasi-Newton methods for multiobjective optimization problems. *4OR*, 16(3):261–294, 2017.

[33] Z. Povalej. Quasi-Newton method for multiobjective optimization. *J. Comput. Appl. Math.*, 255:765–777, 2014.

[34] L. F. Prudente and D. R. Souza. A quasi-Newton method with wolfe line searches for multiobjective optimization. *J. Optim. Theory Appl.*, 194(3):1107–1140, 2022.

[35] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.

[36] W. Rudin. *Principles of Mathematical Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1976.

[37] T. Stewart, O. Bandte, H. Braun, N. Chakraborti, M. Ehrgott, M. Göbelt, Y. Jin, H. Nakayama, S. Poles, and D. Di Stefano. Real-world applications of multi-objective optimization. In J. Branke, K. Deb, K. Miettinen, and R. Słowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 285–327. Springer, Berlin, Heidelberg, 2008. ISBN 978-3-540-88908-3.

[38] J. Wang, Y. Hu, C. K. Wai Yu, C. Li, and X. Yang. Extended Newton methods for multiobjective optimization: majorizing function technique and convergence analysis. *SIAM J. Optim.*, 29(3):2388–2421, 2019.

[39] J.-J. Wang, L.-P. Tang, and X.-M. Yang. Spectral projected subgradient method with a 1-memory momentum term for constrained multiobjective optimization problem. *J. Glob. Optim.*, 89(2):277–302, 2024.

[40] X. Zhao, Q. Sun, L. Liu, and S. Y. Cho. Convergence analysis of a projected gradient method for multiobjective optimization problems. *J. Nonlinear Var. Anal.*, 5(6):929–938, 2021.

[41] X. Zhao, J. Yao, and Y. Yao. A nonmonotone gradient method for constrained multiobjective optimization problems. *J. Nonlinear Var. Anal*, 6(6):693–706, 2022.

[42] X. Zhao, H. Zhang, and Y. Yao. An inexact nonmonotone projected gradient method for constrained multiobjective optimization. *J. Nonlinear Var. Anal.*, 8 (4):517–531, 2024.