An inexact strategy for the projected gradient algorithm in vector optimization problems on variable ordered spaces

J.Y. Bello-Cruz∗ G. Bouza Allende†

November, 2018

Abstract

The variable order structures model situations in which the comparison between two points depends on a point-to-cone application. In this paper, an inexact projected gradient method for solving smooth constrained vector optimization problems on variable ordered spaces is presented. It is shown that every accumulation point of the generated sequence satisfies the first order necessary optimality condition.

The behavior of this scheme is also studied under K-convexity of the objective function where the convergence of all accumulation points is established to a weakly efficient point. Moreover, the convergence results are derived in the particular case in which the problem is unconstrained and if exact directions are taken as descent directions. Furthermore, we investigate the proposed method to optimization models where the domain of the variable order application and the objective function are the same. In this case, similar concepts and convergence results are presented. Finally, some computational experiments designed to illustrate the behavior of the proposed inexact methods versus the exact ones (in terms of CPU time) are performed.

2010 Mathematics Subject Classification: 90C29, 90C52, 65K05, 35E10

Keywords: Gradient method – K-convexity – Variable order – Vector optimization – Weakly efficient points

1 Introduction

The variable order structure [15] is a natural extension of the well-known fixed (partial) order given by a closed, pointed and convex cone. This kind of orderings models situations in which the comparison between two points depends on a set-valued application. The research focused on vector optimization problems on variable ordered spaces and its applications. These problems have recently received much attention from the optimization community due to its broad applications in several different areas. Variable order structures (VOS) given by a point-to-cone valued application were well studied in [15–17], motivated by important applications. VOS appear in

∗Corresponding author. Department of Mathematical Sciences, Northern Illinois University. Watson Hall 366, DeKalb, IL, USA - 60115. E-mail: yunierbello@niu.edu.
†Facultad de Matemática y Computación, Universidad de La Habana, La Habana, Cuba - 10400. E-mail: gema@matcom.uh.cu
medical diagnosis [13], portfolio optimization [37], capability theory of wellbeing [3], psychological modeling [2], consumer preferences [30, 31] and location theory, etc; see, for instance, [1, 17]. The main goal of these models is to find an element of a certain set such that the evaluation of the objective function cannot be improved by the image of any other feasible point with respect to the variable order. So, their mathematical description corresponds to the called Optimization Problem(s) on Variable Ordered Spaces (OPVOS(s)). For the above reasons, it is important to find efficient solution algorithms for solving these kinds of models.

OPVOSs have been treated in [14], in the sense of finding a minimizer of the image of a vector function, with respect to an ordered structure depending on points in the image. It is a particular case of the problem described in [16], where the goal of the model is to find a minimum of a set. Here we will consider a partial (variable) order defined by the cone-valued application which is used to define our problem - OPVOS. We want to point out that OPVOSs generalize the classical vector optimization problems. Indeed, it corresponds to the case in which the order is defined by a constant cone valued application. Many approaches have been proposed to solve the classical constrained vector optimization, such as projected gradient methods, proximal points iterations, weighting techniques schemes, Newton-like and subgradient methods; see, for instance, [4, 7–9, 18, 19, 21–24, 29, 35]. The use of extensions of these iterative algorithms to the variable ordering setting is currently a promising idea. It is worth noting that, as far as we know, only a few of these schemes mentioned above have been proposed and studied in the variable ordering setting; as e.g., the steepest descent algorithm and sub-gradient-like algorithms for unconstrained problems; see, for instance, [5, 6].

In this work, due to its simplicity and the adaptability to the structure of the vectorial problem, we present an inexact projected gradient method for solving constrained variable order vector problems. The properties of the accumulation points of the generated sequence are studied and its convergence is also analyzed under convexity. Moreover, we derive the convergence of the exact projected gradient method and the inexact projected gradient one for the unconstrained problem. Finally, analogous results are obtained if the variable order is given by a point-to-cone application whose domain coincides with the image of the objective function.

This work is organized as follows. The next section provides some notations and preliminary results that will be used in the remainder of this paper. We also recall the concept of $K$-convexity of a function on a variable ordered space and present some properties of this class. Section 3 is devoted to the presentation of the inexact projected gradient algorithm. The convergence of the sequence generated by the projected gradient method is shown in Section 4. Then, under the $K$-convexity of the objective function and the convexity of the set of feasible solutions, we guarantee that the generated sequence is bounded and all its accumulation points are solutions of the variable order problem. Section 5 discusses the properties of this algorithm when the variable order is taken as a cone-value set from the image of the objective function. Section 6 introduces some examples illustrating the behavior of both proposed methods. Finally, some final remarks are given.

2 Preliminaries

In this section we present some preliminary results and definitions. First we introduce some useful notations: Throughout this paper, we write $p := q$ to indicate that $p$ is defined to be equal to $q$ and we write $\mathbb{N}$ for the nonnegative integers $\{0, 1, 2, \ldots \}$. The inner product in $\mathbb{R}^n$ will be denoted by $\langle \cdot, \cdot \rangle$ and the induced norm by $\| \cdot \|$. The closed ball centered at $x$ with radius $r$ is represented by
Let $x^*$ be a weak solution of problem (1). If $F$ is a continuously differentiable function, then $x^*$ is a stationary point.
Proof. Suppose that \( x^* \) is a weak solution of problem (1). Fix \( d \in C - x^* \). By definition there exists \( c \in C \), such that \( d = c - x^* \). Since \( C \) is a convex set, for all \( \alpha \in [0, 1] \), \( x^* + \alpha d \in C \). Since \( x^* \) is a weak solution of problem (1), \( F(x^* + \alpha d) - F(x^*) \not\in \text{int}(K(x^*)) \). Hence,

\[
F(x^* + \alpha d) - F(x^*) \in (-\text{int}(K(x^*)))^c. \tag{3}
\]

The Taylor expansion of \( F \) at \( x^* \) leads us to \( F(x^* + \alpha d) = F(x^*) + \alpha J_F(x^*)d + o(\alpha) \). The last equation together with (3) implies \( \alpha J_F(x^*)d + o(\alpha) \in (-\text{int}(K(x^*)))^c \). Using that \((-\text{int}(K(x^*)))^c \) is a closed cone, and since \( \alpha > 0 \), it follows that

\[
J_F(x^*)d + \frac{o(\alpha)}{\alpha} \in (-\text{int}(K(x^*)))^c. \tag{4}
\]

Taking limit in the above inclusion, when \( \alpha \) goes to 0, and using the closedness of \((-\text{int}(K(x^*)))^c \), we obtain that \( J_F(x^*)d \in (-\text{int}(K(x^*)))^c \), establishing that \( x^* \in S^c \). ■

In classical optimization, stationarity is also a sufficient condition for weak minimality under convexity assumptions. For vector optimization problems on variable ordered spaces, the convexity concept was introduced in Definition 3.1 of [5] as follows:

**Definition 2.2** We say that \( F \) is a K–convex function on \( C \) if for all \( \lambda \in [0, 1] \), \( x, \bar{x} \in C \),

\[
F(\lambda x + (1 - \lambda)\bar{x}) \in \lambda F(x) + (1 - \lambda)F(\bar{x}) - K(\lambda x + (1 - \lambda)\bar{x}).
\]

It is worth noting that in the variable order setting the convexity of \( \text{epi}(F) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid F(x) \leq y - K(x)\} \) is equivalent to the \( K \)-convexity of \( F \) iff \( K(x) \equiv K \) for all \( x \in \mathbb{R}^n \); see Proposition 3.1 of [5]. As already shown in [5], \( K \)-convex functions have directional derivatives under natural assumptions; see Proposition 3.5 of [5]. In particular, if \( Gr(K) \) is closed and \( F \in C^1 \) is \( K \)-convex, then we have the gradient inclusion inequality as follows:

\[
F(x) - F(\bar{x}) \in J_F(\bar{x})(x - \bar{x}) + K(\bar{x}), \quad x \in C, \ \bar{x} \in C.
\]

In the next proposition, we study the relation between stationarity, descent directions and weak solution concept in the constrained sense for problem (1) extending to the variable order setting the results presented in Proposition 1 of [22] and Lemma 5.2 of [21].

**Proposition 2.3** Let \( K \) be a point-to-cone and closed mapping, and \( F \in C^1 \) be a \( K \)-convex function. Then:

(i) The point \( x^* \) is a weak solution of problem (1) iff it is a stationary point.

(ii) If for all \( d \in C - x^* \), \( J_F(x^*)d \not\in -K(x^*) \setminus \{0\} \), then \( x^* \) is a minimizer.

Proof. (i): Let \( x^* \in S^c \), where \( S^c \) is the set of the stationary points. If \( x^* \) is not a weak minimizer then there exists \( x \in C \) such that \( -k_1 := F(x) - F(x^*) \not\in \text{int}(K(x^*)) \). By the convexity of \( F \), for some \( k_2 \in K(x^*) \), we have

\[
-k_1 = F(x) - F(x^*) = J_F(x^*)(x - x^*) + k_2.
\]

It follows from the above equality that

\[
J_F(x^*)(x - x^*) = -(k_1 + k_2). \tag{4}
\]
Moreover, since \( K(x^*) \) is a convex cone, \( k_1 \in \text{int}(K(x^*)) \) and \( k_2 \in K(x^*) \), it holds that \( k_1 + k_2 \in \text{int}(K(x^*)) \). Thus, the last two equalities imply that \( J_F(x^*)(x - x^*) \in -\text{int}(K(x^*)) \), which contradicts the fact that \( x^* \) is a stationary point because \( x \) belongs to \( C \) and hence \( x - x^* \in C - x^* \). The conversely implication was already shown in Proposition 2.1.

**(ii):** By contradiction suppose that there exists \( x \in C \) such that \( F(x) - F(x^*) = -k_1 \), where \( k_1 \in K(x^*) \setminus \{0\} \). Combining the previous condition with (4), it follows that

\[
J_F(x^*)(x - x^*) = -(k_1 + k_2) \in -K(x^*).
\]

Using that \( J_F(x^*)(x - x^*) \notin -K(x^*) \setminus \{0\} \), we get that \( k_1 + k_2 = 0 \), and as \( k_1, k_2 \in K(x^*) \), \( k_1 = -k_2 \). It follows from the pointedness of the cone \( K(x^*) \) that \( k_1 = k_2 = 0 \), contradicting the fact that \( k_1 \neq 0 \).

It is worth mentioning that the concept of \( K \)-convexity for \( F \) depends of the point-to-cone mapping \( K \). Thus, this general approach covers several convexity concepts, from the scalar setting to the vector one and it can be used to model a large number of applications; see, for instance, [2,3,13]. In Section 5 we discuss another variable order when the point-to-cone application depends of the image set of \( F \), such kind of variable orders was introduced and studied in [5,6].

The Inexact Projected Gradient Method to solve problem (1) is presented in the next section.

### 3 The Inexact Projected Gradient Method

This section is devoted to present an inexact projected gradient method for solving constrained smooth problems equipped with a variable order. This method uses an Armijo-type line-search, which is done on inexact descent feasible directions. The proposed scheme here has two main differences with respect to the approach introduced in [5]. (i) it solves constrained problems. (ii) it accepts approximate solutions of the subproblem \((P_\beta)\) with some tolerance.

In the following, several constrained concepts and results will be presented and proved, which will be used in the convergence analysis of the proposed method below.

We start this section by presenting some definitions and basic properties of some auxiliary functions and sets, which will be useful in the convergence analysis of the proposed algorithms. Firstly, we define the set valued mapping \( G : \mathbb{R}^n \rightrightarrows \mathbb{R}^m \), which for each \( x \), defines the set of the normalized generators of \( K^*(x) \), i.e. \( G(x) \subseteq K^*(x) \cap S(0,1) \) is a compact set such that the cone generated by its convex hull is \( K^*(x) \). Although the set of the dual cone \( K^*(x) \cap S(0,1) \) fulfils those properties, in general, it is possible to take smaller sets; see, for instance, [27,28,34]. On the other hand, assuming that \( F \in C^1 \), we consider the support function \( \rho : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \) as

\[
\rho(x,w) := \max_{y \in G(x)} y^T w. \tag{5}
\]

\( \rho(x,w) \) was extensively studied for the vector optimization in Lemma 3.1 of [21] and it is useful to define the useful auxiliary function \( \phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \), as

\[
\phi(x,v) := \max_{y \in G(x)} y^T J_F(x)v. \tag{6}
\]

Then, we are ready to introduce the following auxiliary subproblem, for each \( x \in \mathbb{R}^n \) and \( \beta > 0 \), as

\[
\min_{v \in C-x} \left\{ \frac{\|v\|^2}{2} + \beta \phi(x,v) \right\}. \tag{P_\beta}
\]
**Remark 3.1** Since \( G(x) \) is compact, the function \( \phi(x, \cdot) : \mathbb{R}^n \to \mathbb{R} \) is well defined for each \( x \in \mathbb{R}^n \). Moreover, it is a continuous function.

Next proposition provides a characterization of the stationarity using the auxiliary function \( \phi \), defined in (6). The unconstrained version of the following proposition can be found in Proposition 4.1 of [5]. A version of this proposition was presented in Lemma 2.4 of [20] for vector optimization.

**Proposition 3.2** The following statements hold:

(i) For each \( x \in \mathbb{R}^n \), \( \max_{y \in G(x)} y^T \bar{w} \) if and only if \( \bar{w} \in -\text{int}(K(x)) \).

(ii) The point \( x \) is not stationary iff there exists \( v \in C - x \) such that \( \phi(x, v) < 0 \).

(iii) If \( \phi(x, v) < 0 \) and \( \beta > 0 \), then there exists \( \bar{\lambda} > 0 \) such that \( \parallel \lambda v \parallel^2 + \beta \phi(x, \lambda v) < 0 \) for all \( \lambda \in (0, \bar{\lambda}] \).

(iv) For each \( x \in \mathbb{R}^n \), subproblem \( (P_x) \) has a unique solution, denoted by \( v(x) \).

**Proof.** (i): The result of this item follows as in Proposition 4.1(i) of [5].

(ii): Note that, fixing \( x \), it follows from (6) that \( \phi(x, v) = \rho(x, J_F(x)v) \). Then, by the definition of stationarity and item (i), the statement holds true.

(iii): It follows from the definition of \( \phi(x, v) \) that \( \phi(x, \cdot) \) is a positive homogeneous function. Thus, for all \( \lambda > 0 \),
\[
\frac{\parallel \lambda v \parallel^2}{2} + \beta \phi(x, \lambda v) = \lambda \left( \frac{\parallel v \parallel^2}{2} + \beta \phi(x, v) \right).
\]

(7)

Since \( \phi(x, v) < 0 \), there exists \( \bar{\lambda} > 0 \) small enough such that \( \lambda \parallel v \parallel^2 + \beta \phi(x, v) < 0 \). Hence, (7) together with the above inequality implies that \( \frac{\parallel \lambda v \parallel^2}{2} + \beta \phi(x, \lambda v) < 0 \), for all \( \lambda \in (0, \bar{\lambda}] \), as desired.

(iv): Using the definition of the function \( \phi(x, v) \), given in (7), it is easy to prove that \( \phi \) is a sublinear function as well. Hence, \( \phi(x, \cdot) \) is a convex function, and then, \( \parallel v \parallel^2 + \beta \phi(x, v) \) is a strongly convex function. Since \( C \) is a convex set, \( C - x \) is also convex and therefore, subproblem \( (P_x) \) has a unique minimizer.

Based on Proposition 3.2(iii), we can define \( v(x) \) as the unique solution of problem \( (P_x) \) and \( y(x, v) \) is an element of the compact set \( G(x) \) such that \( y(x, v)^T J_F(x)v = \phi(x, v) \). Next we will discuss about the continuity of the function
\[
\theta_\beta(x) := \frac{\parallel v(x) \parallel^2}{2} + \beta \phi(x, v(x)),
\]
which is related with the one defined in (35) of [22].

The following proposition is the constrained version of Proposition 4.2 in [5]. Items (i)-(ii), (iii) and (iv) can be seen as a version for the variable vector optimization of Proposition 3 of [22], Proposition 2.5 of [20] and Proposition 3.4 of [21], respectively.

**Proposition 3.3** Let \( F \in C^1 \) and fix \( \beta > 0 \). Then, the following hold

(i) \( \theta_\beta(x) \leq 0 \) for all \( x \in C \).
(ii) \( x \) is a stationary point iff \( \theta_\beta(x) = 0 \).

(iii) \( \|v(x)\| \leq 2\beta \|J_F(x)\| \).

(iv) If \( G \) is a closed application, then \( \theta_\beta \) is an upper semi-continuous function on \( C \).

Proof. (i): Note that as \( 0 \in C - x \) for all \( x \in C \) and \( \theta_\beta(x) \leq \frac{\|0\|^2}{2} + \beta \phi(x, 0) = 0 \).

(ii): As shown in Proposition 3.2(ii), \( x \) is a non stationary point iff for some \( v \in C - x, \phi(x, v) < 0 \). Then, by Proposition 3.2(iii), there exists \( \hat{v} \in C - x \) such that \( \frac{\lambda^2}{2} \|v\|^2 + \lambda \beta \phi(x, v) < 0 \) and hence \( \theta_\beta(x) < 0 \).

(iii): By (i), \( 0 \geq \theta_\beta(x) = \frac{\|v(x)\|^2}{2} + \beta y(x, v(x))^T J_F(x)v(x) \). Then, after some algebra, we get
\[
\frac{\|v(x)\|^2}{2} \leq -\beta y(x, v(x))^T J_F(x)v(x) \leq \beta \|y(x, v(x))^T J_F(x)v(x)\|
\]
Using that \( \|y(x, v(x))\| = 1 \), it follows from the above inequality that
\[
\frac{\|v(x)\|^2}{2} \leq \beta \|J_F(x)\| \|v(x)\|
\]
and the result follows after dividing the above inequality by the positive term \( \|v(x)\|^2 / 2 \neq 0 \).

(iv): Now we prove the upper semi-continuity of the function \( \theta_\beta \). Let \( (x^k)_{k \in \mathbb{N}} \) be a sequence converging to \( x \). Take \( \hat{x} \in C \) such that \( v(x) = \hat{x} - x \) and also denote \( \hat{x}^k := v^k + x^k \). It is clear that, for all \( k \in \mathbb{N} \), \( \hat{x} - x^k \in C - x^k \), and so,
\[
\theta_\beta(x^k) = \frac{\|x^k - x^k\|^2}{2} + \beta \phi(x^k, \hat{x}^k - x^k)
\leq \frac{\|\hat{x} - x^k\|^2}{2} + \beta \phi(x^k, \hat{x} - x^k)
= \frac{\|\hat{x} - x^k\|^2}{2} + \beta y_k^T J_F(x^k)(\hat{x} - x^k).
\]
Since each \( y_k := y(x^k, \hat{x} - x^k) \) belongs to the compact set \( G(x^k) \subseteq K^*(x^k) \cap S(0, 1) \subseteq B(0, 1) \) for all \( k \in \mathbb{N} \), then the sequence \( (y_k)_{k \in \mathbb{N}} \) is bounded because is in \( \bigcup_{k \in \mathbb{N}} G(x^k) \subseteq B(0, 1) \). Therefore, there exists a convergent subsequence of \( (y_k)_{k \in \mathbb{N}} \). We can assume without lost of generality that \( \lim_{k \to \infty} y_k = y \), and also since \( G \) is closed, \( y \in G(x) \). Taking limit in (9), we get
\[
\limsup_{k \to \infty} \theta_\beta(x^k) \leq \limsup_{k \to \infty} \frac{\|\hat{x} - x^k\|^2}{2} + \beta y_k^T J_F(x^k)(\hat{x} - x^k)
= \frac{\|\hat{x} - x\|^2}{2} + \beta y^T J_F(x)(\hat{x} - x)
\leq \frac{\|\hat{x} - x\|^2}{2} + \beta \phi(x, \hat{x} - x) = \theta_\beta(x).
\]
Then, the function \( \theta_\beta \), defined in (8), is upper semi-continuous. \( \blacksquare \)
Lemma 3.4 Consider any \( x, \hat{x} \in C \) and \( z \in \mathbb{R}^n \). If \( J_F \) is locally Lipschitz, \( d_H(G(x), G(\hat{x})) \leq L_G \| x - \hat{x} \| \) for some \( L_G > 0 \) and \( C \) is bounded, then

\[
|\phi(x, z) - \phi(\hat{x}, z)| \leq L \| x - \hat{x} \|,
\]

for some \( L > 0 \). Hence \( \phi \) is a continuous function in the first argument.

Proof. By Proposition 4.1(iv) of [5], and using the Lipschitz assumption for \( G \) in \( C, \rho(x, w) \), defined in (5), is also a Lipschitz function for all \( (x, w) \in C \times W \) for any bounded subset \( W \subset \mathbb{R}^n \). That is, if \( \| w_i \| \leq M, i = 1, 2 \) with \( M > 0 \), then

\[
|\rho(x_1, w_1) - \rho(x_2, w_2)| \leq \hat{L} \| x_1 - x_2 \| + \| w_1 - w_2 \|,
\]

where \( \hat{L} := L_G M \). Taking (10) for \( x_1 = x, x_2 = x^k, w_1 = J_F(x)(\hat{x}^k - x^k) \) and \( w_2 = J_F(x^k)(\hat{x}^k - x^k) \), we get

\[
|\rho(x, J_F(x)(\hat{x}^k - x^k)) - \rho(x^k, J_F(x^k)(\hat{x}^k - x^k))| \leq \hat{L} \| x - x^k \| + \| (J_F(x) - J_F(x^k))(\hat{x}^k - x^k) \| \\
\leq \hat{L} \| x - x^k \| + \| J_F(x) - J_F(x^k) \| \| \hat{x}^k - x^k \|,
\]

because of the compactness of \( C \) and the continuity of \( J_F, \| J_F(x)(\hat{x}^k - x^k) \| \leq M \) for all \( k \in \mathbb{N} \) and \( x \in C \). Noting that

\[
\phi(x, \hat{x}^k - x^k) - \phi(x^k, \hat{x}^k - x^k) = \rho \left( x, J_F(x)(\hat{x}^k - x) \right) - \rho \left( x^k, J_F(x^k)(\hat{x}^k - x^k) \right),
\]

and due to \( J_F \) is locally Lipschitz and (10), it follows that

\[
|\phi(x, \hat{x}^k - x^k) - \phi(x^k, \hat{x}^k - x^k)| \leq (\hat{L} + L \hat{M}) \| x - x^k \|,
\]

where \( \hat{L} \) is the Lipschitz constant of \( J_F \) and \( \| \hat{x}^k - x^k \| \leq \hat{M} \) for all \( k \in \mathbb{N} \). This proves the continuity of \( \phi \) in the first argument.

Now we can prove the lower semicontinuity of \( \theta_\beta \) by following similar ideas of the result presented in Proposition 3.4 of [21] for vector optimization.

Proposition 3.5 Let \( F \in C^1 \) and consider any \( x, \hat{x} \in C \) with \( C \) bounded. Then, if \( d_H(G(x), G(\hat{x})) \leq L_G \| x - \hat{x} \| \) for some \( L_G > 0 \) and \( J_F \) is locally Lipschitz, \( \theta_\beta \) is a lower semicontinuous function on \( C \).

Proof. We consider the function \( \theta_\beta(x) \). Note further that

\[
\theta_\beta(x) \leq \beta \phi(x, \hat{x}^k - x) + \frac{\| \hat{x}^k - x \|^2}{2} \\
= \theta_\beta(x^k) + \beta \left[ \phi(x, \hat{x}^k - x) - \phi(x^k, \hat{x}^k - x^k) \right] + \frac{\| \hat{x}^k - x \|^2 - \| \hat{x}^k - x^k \|^2}{2} \\
= \theta_\beta(x^k) + \beta \left[ \phi(x, \hat{x}^k - x) - \phi(x^k, \hat{x}^k - x^k) \right] + \frac{1}{2} \left[ -2 \langle \hat{x}^k, x^k - x \rangle + \| x \|^2 - \| \hat{x}^k \|^2 \right] .
\]

Thus, taking limit in the previous inequality and using Lemma 3.4, we get

\[
\lim_{k \to \infty} \phi(x, \hat{x}^k - x) - \phi(x^k, \hat{x}^k - x^k) = 0.
\]
Also, it is follows that \( \lim_{k \to \infty} \frac{1}{2} [\|x\|^2 - \|x^k\|^2] - \langle \dot{x}^k, x^k - x \rangle = 0 \). Hence,

\[
\theta_{\beta}(x) \leq \liminf_{k \to \infty} \left\{ \theta_{\beta}(x^k) + \beta \left[ \phi(x, x^k - x) - \phi(x, z^k - x^k) \right] - \langle \dot{x}^k, x^k - x \rangle + \frac{\|x\|^2 - \|x^k\|^2}{2} \right\} \\
= \liminf_{k \to \infty} \theta_{\beta}(x^k),
\]

establishing the desired result.

Now we recall the concept of \( \delta \)-approximate direction introduced in Definition 3.1 of [20].

**Definition 3.6** Let \( x \in C \) and \( \beta > 0 \). Given \( \delta \in [0,1) \), we say that \( v \) is a \( \delta \)-approximate solution of problem \((P_x)\) if \( v \in C - x \) and \( \beta \phi(x,v) + \frac{\|v\|^2}{2} \leq (1 - \delta) \theta_{\beta}(x) \). If \( v \neq 0 \) we say that \( v \) is a \( \delta \)-approximate direction.

Hence, from a numerical point of view, it would be interesting to consider algorithms in which the line-search is given over a \( \delta \)-approximate solution of subproblem \((P_x)\) instead of on an exact solution of it.

**Remark 3.7** Note that if the solution of \((P_x)\) is 0, then the only possible \( \delta \)-approximate solution is \( v = 0 \). In other case, since \( \theta_{\beta}(x) < 0 \), there exist feasible directions \( v \) such that

\[
\beta \phi(x,v) + \frac{\|v\|^2}{2} \in [\theta_{\beta}(x), (1 - \delta) \theta_{\beta}(x)].
\]

In particular \( v(x) \), the solution of \((P_x)\), is always a \( \delta \)-approximate solution. As was presented in [12], a variant to this approach consists in to find \( v \) such that \( J_F(x_k)v \in \cup_{x \in C} K(x) \). However, this approach will converge to a stationary point as in the case we have proposed. Non-dominated points are minimizers under well-known conditions. A future research line will be to find, in this case, how we can guarantee that the stationary points, that the algorithm converges, are non-dominated ones.

Next we present an inexact algorithm for solving problem (1). The algorithm requires the following exogenous parameters: \( \delta \in [0,1) \) and \( \sigma, \gamma \in (0,1) \) and \( 0 < \hat{\beta} \leq \beta < +\infty \).

**Inexact Projected Gradient Method (IPG Method).** Assume that \( \beta_k \in [\beta, \hat{\beta}] \) for all \( k \in \mathbb{N} \).

**Initialization:** Take \( x^0 \in \mathbb{R}^n \) and \( \beta_0 \).

**Iterative step:** Given \( x^k \) and \( \beta_k \), compute \( v^k \), \( \delta \)-approximate solution of \((P_{x^k})\). If \( v^k = 0 \), then stop. Otherwise compute

\[
j(k) := \min \left\{ j \in \mathbb{Z}_+ : F(x^k) - F(x^k + \gamma_j v^k) + \sigma \gamma_j J_F(x^k)v^k \in K(x^k) \right\}. \tag{12}
\]

Set \( x^{k+1} = x^k + \gamma_k v^k \in C \), with \( \gamma_k = \gamma j(k) \).

It is worth noting that **IPG Method** extends Algorithm 3.3 of [20] to the variable order setting. Next proposition proves that the stepsize \( \gamma_k \) is well defined for all \( k \in \mathbb{N} \), i.e., there exists a finite positive integer \( j \) that fulfills Armijo-type rule given in (12) at each step of **IPG Method**. The proof of the next result uses a similar idea to the presented in Proposition 2.2 of [20].
Proposition 3.8 Subproblem (12) has a finite solution, i.e., there exists an index \( j(k) < +\infty \) which is solution of (12).

Proof. If \( v^k = 0 \) then \textbf{IPG Method} stops. Otherwise, if \( v^k \neq 0 \) then by Proposition 3.3(ii), \( x^k \) is not a stationary point and \( \theta_{\beta_k}(x^k) < 0 \). Moreover,

\[
\beta_k \phi(x^k, v^k) \leq \beta_k \phi(x^k, v^k) + \frac{\|v^k\|^2}{2} \leq (1 - \delta) \theta_{\beta_k}(x^k) < 0.
\]

Note further that \( \phi(x^k, v^k) = \max_{y \in G(x^k)} y^T J_F(x^k)v^k < 0 \). Thus, it follows from Proposition 3.2(i) that

\[
J_F(x^k)v^k \in -\text{int}(K(x^k)). \tag{13}
\]

Using the Taylor expansion of \( F \) at \( x^k \), we obtain that

\[
F(x^k) + \sigma \gamma J_F(x^k)v^k - F(x^k + \gamma v^k) = (\sigma - 1) \gamma J_F(x^k)v^k + o(\gamma). \tag{14}
\]

Since \( \sigma < 1 \) and \( K(x^k) \) is a cone, it follows from (13) that \((\sigma - 1) \gamma J_F(x^k)v^k \in \text{int}(K(x^k))\). Then, there exists \( \ell \in \mathbb{N} \) such that, for all \( j \geq \ell \), we get \((\sigma - 1) \gamma J_F(x^k)v^k + o(\gamma) \in K(x^k)\). Combining the last inclusion with (14), we obtain \( F(x^k) + \sigma \gamma J_F(x^k)v^k - F(x^k + \gamma v^k) \in K(x^k) \) for all \( j \geq \ell \).

Hence (12) holds for \( j(k) = \ell \).

\[
\blacksquare
\]

Remark 3.9 After this proposition it is clear that given \((x^k, v^k), j(k)\) is well-defined. Furthermore, the sequence generated by \textbf{IPG Method} is always feasible. Indeed, as \( x^k, x^k + v^k \in C, \gamma_k \in (0,1] \) and \( C \) is convex, \( x^{k+1} = x^k + \gamma_k v^k \in C \).

4 Convergence Analysis of \textbf{IPG Method}

In this section we prove the convergence of \textbf{IPG Method} presented in the previous section. First we consider the general case and then the result is refined for \( K \)-convex functions. From now on, \((x^k)_{k \in \mathbb{N}} \) denote the sequence generated by \textbf{IPG Method}. We begin the section with the following lemma.

Lemma 4.1 Let \( F \in C^1 \). Assume that

(a) \( \cup_{x \in C} K(x) \subseteq K \), where \( K \) is a closed, pointed and convex cone.

(b) The application \( G \) is closed.

(c) \( d_H(G(x), G(\bar{x})) \leq L_G \|x - \bar{x}\| \), for all \( x, \bar{x} \in C \).

If \( x^* \) is an accumulation point of \((x^k)_{k \in \mathbb{N}} \), then \( \lim_{k \to \infty} F(x^k) = F(x^*) \).

Proof. Let \( x^* \) be any accumulation point of the sequence \((x^k)_{k \in \mathbb{N}} \) and denote \((x^i)_{i \in \mathbb{N}} \) a subsequence of \((x^k)_{k \in \mathbb{N}} \) such that \( \lim_{k \to \infty} x^i = x^* \). It follows from the definition of Armijo-type line-search in (12) that

\[
F(x^{k+1}) - F(x^k) - \sigma \gamma J_F(x^k)v^k \in -K(x^k). \tag{15}
\]

Since \textbf{IPG Method} does not stop after finitely many steps, \( v_k \neq 0 \), which means that \( \phi(x^k, v^k) < 0 \). By Proposition 3.2(i), this means that \( J_F(x^k)v^k \in -\text{int}(K(x^k)) \). Multiplying the last inclusion by \( \sigma \gamma_k > 0 \) and summing with (15), we get from the convexity of \( K(x^k) \) that

\[
F(x^{k+1}) - F(x^k) - \sigma \gamma_k J_F(x^k)v^k + \sigma \gamma_k J_F(x^k)v^k \in -\text{int}(K(x^k)).
\]
Thus, $F(x^{k+1}) - F(x^k) \in -\text{int}(K(x^k))$. Since $\bigcup_{x \in C} K(x) \subseteq K$ from (a), it holds that $\text{int}(K(x)) \subseteq \text{int}(K)$ for all $x$, and $F(x^{k+1}) - F(x^k) \in -\text{int}(K)$. Hence, $(F(x^k))_{k \in \mathbb{N}}$ is decreasing with respect to cone $K$. The continuity of $F$ implies that $\lim_{k \to \infty} F(x^k) = F(x^*)$. Then, to prove that the whole sequence $(F(x^k))_{k \in \mathbb{N}}$ converges to $F(x^*)$, we use that it is decreasing with respect to cone $K$, which is a closed, pointed and convex cone; see, for instance, [25, 36]. Thus, we get that $\lim_{k \to \infty} F(x^k) = F(x^*)$, as desired.

Next we present an analogous result as was proved in Proposition 3.3(iii) where $v^k$ is a $\delta$-solution of subproblem $(P_{\delta})$, which gives us a upper bound for the norm of $v^k$. Next lemma is a version of Proposition 2.5 of [20] to the variable order setting.

**Lemma 4.2** Let $(x^k)_{k \in \mathbb{N}}$ and $(\beta_k)_{k \in \mathbb{N}}$ be sequences generated by IPG Method. Then, $\|v^k\| \leq 2\beta_k \|J_F(x^k)\|$. 

**Proof.** By the definition of $\delta$-approximate direction $\beta_k \phi(x^k, v^k) + \frac{\|v^k\|^2}{2} \leq (1 - \delta) \theta_k(x^k)$. As was shown in Proposition 3.3(ii), $(1 - \delta) \theta_k(x^k) \leq 0$, since $x^k \in C$. Thus, $\frac{\|v^k\|^2}{2} \leq -\beta_k \phi(x^k, v^k)$ and the result follows as in Proposition 3.3(iii).

Next we prove the stationarity of the accumulation points of the generate sequence. Some arguments used in the proof of the next theorem are similar to those of Theorem 3.5 of [20] and Theorem 5.1 of [5] for fixed and variable vector optimization, respectively.

**Theorem 4.3** Suppose that

(a) $\bigcup_{x \in C} K(x) \subseteq K$, where $K$ is a closed, pointed and convex cone.

(b) The application $G$ is closed.

(c) $d_H(G(x), G(\hat{x})) \leq L_G \|x - \hat{x}\|$, for all $x, \hat{x} \in C$.

(d) $J_F$ is a locally Lipschitz function.

If $(\beta_k)_{k \in \mathbb{N}}$ is a bounded sequence, then all the accumulation points of $(x^k)_{k \in \mathbb{N}}$ are stationary points of problem (1).

**Proof.** Let $x^*$ be an accumulation point of the sequence $(x^k)_{k \in \mathbb{N}}$. Denote $(x^i)_{k \in \mathbb{N}}$ any convergent subsequence to $x^*$. Since $F \in C^1$, Lemma 4.2 implies that the subsequence $(v^i)_{k \in \mathbb{N}}$ is also bounded and hence has a convergent subsequence. Without loss of generality, we assume that $(v^i)_{k \in \mathbb{N}}$ converges to $v^*$, $\beta_i$ and $\gamma_i$ converge to $\beta^*$ and $\gamma^*$, respectively. Recall that $\rho(x, w) = \max_{y \in G(x)} y^T w$.

By definition we have $F(x^{k+1}) - F(x^k) - \sigma \gamma_k J_F(x^k) v^k \in -K$. Using Proposition 3.2(i), implies that $\rho(x^k, F(x^{k+1}) - F(x^k) - \sigma \gamma_k J_F(x^k) v^k) \leq 0$. Since the function $\rho$ is sublinear, as shown in Proposition 3.2 (iv), we get

$$\rho(x^k, F(x^{k+1}) - F(x^k)) \leq \sigma \gamma_k \rho(x^k, J_F(x^k) v^k).$$  \hspace{1cm} (16)

Using the semilinear property for $\rho$ in the second argument, we can rewrite (16) as $\rho(x^k, F(x^k)) - \rho(x^k, F(x^{k+1})) \geq -\sigma \gamma_k \rho(x^k, J_F(x^k) v^k) \geq 0$. Now considering the subsequences $\{x^i\}$ and $\{v^i\}$, where $v^i = v(x^i)$ on this inequality, we have

$$\lim_{k \to \infty} \left[ \rho(x^i, F(x^i)) - \rho(x^i, F(x^{i+1})) \right] \geq -\sigma \lim_{k \to \infty} \gamma_k \rho(x^i, J_F(x^i) v^i) \geq 0.$$
As already was observed in the proof of Lemma 3.4, \( \rho \) is continuous and moreover from Lemma 4.1, we have \( \lim_{k \to \infty} F(x^k) = F(x^*) \). Thus,

\[
\lim_{k \to \infty} \left[ \rho \left( x^{i_k}, F(x^{i_k}) \right) - \rho \left( x^{i_k}, F(x^{i_k+1}) \right) \right] = \rho \left( x^*, F(x^*) \right) - \rho \left( x^*, F(x^*) \right) = 0.
\]

These facts imply that \( \lim_{k \to \infty} \gamma_i \rho \left( x^{i_k}, I_F(x^{i_k})v^{i_k} \right) = 0 \). Hence we can split our analysis in two cases \( \gamma^* > 0 \) and \( \gamma^* = 0 \).

**Case 1:** \( \gamma^* > 0 \). Here

\[
\lim_{k \to \infty} \phi(x^{i_k}, v^{i_k}) = \lim_{k \to \infty} \rho \left( x^{i_k}, I_F(x^{i_k})v^{i_k} \right) = 0. \tag{17}
\]

Suppose that

\[
\theta_{\beta_i}(x^*) = \|v(x^*)\|^2/2 + \beta_i \phi(x^*, v(x^*)) < -\epsilon < 0, \tag{18}
\]

where \( v(x^*) = \hat{x} - x^* \) with \( \hat{x} \in \mathbb{C} \). Due to the continuity of \( \phi(\cdot, \cdot) \) in both arguments, Lemma 3.4 and (17) imply that

\[
\phi(x^{i_k}, v^{i_k}) > -\frac{(1 - \delta)\epsilon}{\max_{k \in \mathbb{N}} \beta_k} = -\frac{(1 - \delta)\epsilon}{\hat{\beta}}
\]

for \( k \) large enough. After note that \( (\beta_k)_{k \in \mathbb{N}} \) is a positive and bounded sequence, then

\[
\|v^{i_k}\|^2/2 + \beta_i \phi(x^{i_k}, v^{i_k}) \geq \beta_i \phi(x^{i_k}, v^{i_k}) > -\beta_i \frac{(1 - \delta)\epsilon}{\hat{\beta}} \geq -(1 - \delta)\epsilon. \tag{19}
\]

By definition of the subsequence \( (v^{i_k})_{k \in \mathbb{N}} \), we have, for all \( v^{i_k} \in \mathbb{C} - x^{i_k} \) and \( v \in \mathbb{C} - x^{i_k} \),

\[
(1 - \delta) \left( \frac{\|v\|^2}{2} + \beta_i \phi(x^{i_k}, v) \right) \geq (1 - \delta) \theta_{\beta_i}(x^{i_k}) \geq \frac{\|v^{i_k}\|^2}{2} + \beta_i \phi(x^{i_k}, v^{i_k}). \tag{20}
\]

Combining (19) and (20), we obtain that \( (1 - \delta) \left( \frac{\|v\|^2}{2} + \beta_i \phi(x^{i_k}, v) \right) > -(1 - \delta)\epsilon \). In particular consider \( \delta^k = \hat{x} - x^{i_k} \). Dividing by \( (1 - \delta) > 0 \), we obtain

\[
\frac{\|\delta^k\|^2}{2} + \beta_i \phi(x^{i_k}, \delta^k) > -\epsilon.
\]

By the continuity of function \( \phi \) with respect to the first argument and taking limit in the previous inequality, lead us to the following inequality \( \|v(x^*)\|^2/2 + \beta^* \phi(x^*, v(x^*)) \geq -\epsilon \). This fact and (18) imply

\[
-\epsilon > \frac{\|v(x^*)\|^2}{2} + \beta^* \phi(x^*, v(x^*)) \geq -\epsilon,
\]

which is a contradiction. Thus, we can conclude that \( \theta_{\beta_i}(x^*) \geq 0 \) and, hence, using Proposition 3.3, \( x^* \) is a stationary point if \( \limsup_{k \to \infty} \gamma_i > 0 \).

**Case 2:** \( \gamma^* = 0 \). We consider the previously defined convergent subsequences \( (x^{i_k})_{k \in \mathbb{N}}, (\beta_i)_{k \in \mathbb{N}}, (v^{i_k})_{k \in \mathbb{N}}, (\gamma_i)_{k \in \mathbb{N}} \) convergent to \( x^*, \beta^*, v^* \) and \( \gamma^* = 0 \), respectively. Since \( \beta^* > 0 \), we get that

\[
\rho \left( x^{i_k}, I_F(x^{i_k})v^{i_k} \right) \leq \rho \left( x^{i_k}, I_F(x^{i_k})v^{i_k} \right) + \frac{\|v^{i_k}\|^2}{2\beta_i}.
\]
Since $v^k$ is a $δ$-approximate solution of $(P_{v^k})$, see Definition 3.6, then
\[
ρ \left( x^k, J_F(x^k)v^k \right) + \frac{\|v^k\|^2}{2β_k} \leq \frac{(1−δ)}{β_k}(x^k) < 0.
\]
It follows from taking limit above that $ρ(x^*, J_F(x^*)v^*) ≤ \frac{\|v^*\|^2}{2β_*} ≤ 0$. Fix $q ∈ \mathbb{N}$. Then, for $k$
large enough $F(x^k + γqv^k) ≠ F(x^k) + σγqJ_F(x^k)v^k - K(x^k)$, as there exists $\tilde{y}_k ∈ G(x^k)$ such that
\[
\langle F(x^k + γqv^k) - F(x^k) - σγqJ_F(x^k)v^k, \tilde{y}_k \rangle > 0,
\]
holds that
\[
ρ \left( x^k, F(x^k + γqv^k) - F(x^k) - σγqJ_F(x^k)v^k \right) ≥ 0.
\]
Taking limit as $k$ tends to $+∞$, and using that $ρ$ is a continuous function, then
\[
ρ(x^*, F(x^* + γqv^*) - F(x^*) - σγqJ_F(x^*)v^*) ≥ 0.
\]
But $ρ(x, ·)$ is a positive homogeneous function, so,
\[
ρ \left( x^*, \frac{F(x^* + γqv^*) - F(x^*)}{γq} - σJ_F(x^*)v^* \right) ≥ 0.
\]
Taking limit as $q$ tends to $+∞$, we obtain $ρ(x^*, (1−σ)J_F(x^*)v^*) ≥ 0$. Finally, since
\[
ρ(x^*, J_F(x^*)v^*) ≤ 0,
\]
holds $ρ(x^*, J_F(x^*)v^*) = 0$ and by Proposition 3.2(ii), this is equivalent to say that $x^* ∈ S^δ$. ■

The above result generalizes Theorem 5.1 of [5], where the exact steepest descent method for unconstrained problems was studied. Recall that at the exact variant of the algorithm the direction $v^k$
is computed as an exact solution of problem $(P_{v^k})$. In order to fill the gap between these two cases, we present two direct consequences of the above result, the inexact method for unconstrained problems and the exact method for the constrained problem.

**Corollary 4.4** Suppose that conditions (a)-(d) of Theorem 4.3 are fulfilled. Then all accumulation points of the sequence $(x^k)_{k∈\mathbb{N}}$ generated by the exact variant of IPG Method are stationary points of problem (1).

**Proof.** Apply Theorem 4.3 to the case $δ = 0$. ■

**Corollary 4.5** Suppose that conditions (a)-(d) of Theorem 4.3 are fulfilled for $C = \mathbb{R}^n$. If $(β_k)_{k∈\mathbb{N}}$ is a bounded sequence, then all accumulation points of $(x^k)_{k∈\mathbb{N}}$ computed by IPG Method are stationary points of problem (1).

**Proof.** Directly by applying Theorem 4.3 for $C = \mathbb{R}^n$. ■

The result presented in Theorem 4.3 assumes the existence of accumulation points. We want to emphasize that this is a fact that takes place even when the projected gradient method is applied to the solution of classical scalar problems, i.e., $m = 1$ and $K(x) = \mathbb{R}_+$. The convergence of the whole sequence generated by the algorithm is only possible under stronger assumptions as convexity. Now, based on quasi-Fejér theory, we will prove the full convergence of the sequence generated by IPG Method when we assume that $F$ is $K$-convex. We start by presenting its definitions and its properties.

**Definition 4.6** Let $S$ be a nonempty subset of $\mathbb{R}^n$. A sequence $(z^k)_{k∈\mathbb{N}}$ is said to be quasi-Fejér convergent to $S$ iff for all $x ∈ S$, there exists $k$ and a summable sequence $(ε_k)_{k∈\mathbb{N}} ⊂ \mathbb{R}_+$ such that $\|z^{k+1} - x\|^2 ≤ \|z^k - x\|^2 + ε_k$ for all $k ≥ k$. 

13
This definition originates in [10] and has been further elaborated in [26]. A useful result on quasi-Fejér sequences is the following.

**Fact 4.7** If \((z^k)_{k \in \mathbb{N}}\) is quasi-Fejér convergent to \(S\) then,

(i) The sequence \((z^k)_{k \in \mathbb{N}}\) is bounded.

(ii) If an accumulation point of \((z^k)_{k \in \mathbb{N}}\) belongs to \(S\), then the whole sequence \((z^k)_{k \in \mathbb{N}}\) converges.

**Proof.** See Theorem 1 of [11].

For guaranteeing the convergence of IPG Method, we introduce the following definition which is related with the one presented in Definition 4.2 of [20].

**Definition 4.8** Let \(x \in C\). A direction \(v \in C - x\) is scalarization compatible (or simply s-compatible) at \(x\) if there exists \(w \in \text{conv}(G(x))\) such that \(v = P_{C-x}(-\beta F(x)w)\).

Now we proceed as in Proposition 4.3 of [20]. In the following we present the relation between inexact and s-compatible directions.

**Proposition 4.9** Let \(x \in C\), \(w \in \text{conv}(G(x))\), \(v = P_{C-x}(-\beta F(x)w)\) and \(\delta \in [0, 1)\). If

\[
\beta \phi(J_F(x)v) \leq (1 - \delta)\beta \langle w, J_F(x)v \rangle - \frac{\delta}{2}\|v\|^2,
\]

then \(v\) is a \(\delta\)-approximate projected gradient direction.

**Proof.** See Proposition 4.3 of [20].

We start the analysis with a technical result which is related with the proof of Lemma 5.3 of [20].

**Lemma 4.10** Suppose that \(F\) is \(K\)-convex. Let \((x^k)_{k \in \mathbb{N}}\) be a sequence generated by IPG Method where \(v^k\) is an s-compatible direction at \(x^k\), given by \(v^k = P_{C-x^k}(-\beta_k F(x)w^k)\), with \(w^k \in \text{conv}(G(x^k))\) for all \(k \in \mathbb{N}\). If for a given \(\hat{x} \in C\) we have \(F(\hat{x}) - F(x^k) \in -K(x^k)\), then

\[
\|x^{k+1} - \hat{x}\|^2 \leq \|x^k - \hat{x}\|^2 + 2\beta_k\gamma_k\|\langle w^k, J_F(x^k)v^k \rangle\|.
\]

**Proof.** Since \(x^{k+1} = x^k + \gamma_k v^k\), we have \(\|x^{k+1} - \hat{x}\|^2 = \|x^k - \hat{x}\|^2 + \gamma_k^2\|v^k\|^2 - 2\gamma_k\langle v^k, \hat{x} - x^k \rangle\). Let us analyze the rightmost term of the above expression. It follows from the definition of \(v^k\) and the obtuse angle property of projections that \((-\beta_k F(x^k)w^k - v^k, v - v^k) \leq 0\), for all \(v \in C - x^k\). Taking \(v = \hat{x} - x^k \in C - x^k\) on the above inequality, we obtain

\[
-\langle v^k, \hat{x} - x^k \rangle \leq \beta_k\langle w^k, F(x^k)(\hat{x} - x^k) \rangle - \beta_k\langle w^k, J_F(x^k)v^k \rangle - \|v^k\|^2.
\]

Now, it follows from the convexity of \(F\) that \(\langle w^k, J_F(x^k)(\hat{x} - x^k) \rangle \leq \langle w^k, F(\hat{x}) - F(x^k) \rangle\). Also the fact \(F(\hat{x}) \leq_K(x^k) F(x^k)\), i.e., \(F(\hat{x}) - F(x^k) \in -K(x^k)\), together with \(w^k \in K^*(x^k)\) imply that \(\langle w^k, F(\hat{x}) - F(x^k) \rangle \leq 0\). Moreover, by using \(J_F(x^k)v^k \in \text{int}(-K(x^k))\) and \(w^k \in G(x^k) = K^*(x^k) \cap S(0, 1)\), we have \(\langle w^k, J_F(x^k)v^k \rangle < 0\). Thus, we get \(-\langle v^k, \hat{x} - x^k \rangle \leq \beta_k\|\langle w^k, J_F(x^k)v^k \rangle\| - \|v^k\|^2\). The result follows because \(\gamma_k \in (0, 1]\).

We still need to make a couple of supplementary assumptions, which are standard in convergence analysis of classical (scalar-valued) methods and in its extensions to the vector optimization setting.

14
Assumption 4.4: Let \((z^k)_{k \in \mathbb{N}} \in F(C)\) be a sequence such that \(z^k - z^{k+1} \in K(z^k)\) for all \(k \in \mathbb{N}\) and \(z \in C, z^k - z \in \mathcal{K}\) for some closed, convex and pointed cone \(\mathcal{K}\). Then there exists \(\hat{z} \in C\) such that \(F(\hat{z}) \preceq_{\mathcal{K}} z^k\) for all \(k \in \mathbb{N}\), i.e., \(F(\hat{z}) - z^k \in -\mathcal{K}\). In this case we said that the sequence \((z^k)_{k \in \mathbb{N}}\) is bounded below with respect to \(\mathcal{K}\).

Recently, it was observed in [32] that this assumption could be replaced by assuming that the restriction of \(F\) on \(C\) has compact sections. This assumption is related to the completeness of the image of \(F\). It is important to mention that completeness is a standard assumption for ensuring existence of efficient points in vector problems [34].

Assumption 4.5: The search direction \(v^k\) is s-compatible at \(x^k\), that is to say, \(v^k = P_{C-x^k}(-\beta J_F(x^k)^T w^k)\), where \(w^k \in \text{conv}(G(x^k))\) for all \(k \in \mathbb{N}\).

This assumption holds automatically in the exact case. Moreover, it has been widely used in the literature in the vector case; see, for instance, [20]. Version of these assumptions are also used in [20], when the order is given by a constant cone.

The next result is an extension to the variable order setting of Theorem 5.2 of [20].

**Theorem 4.11** Assume that \(F\) is \(K\)-convex and that Assumptions 4.4 and 4.5 hold. If \(\text{int}(\cap_{k \in \mathbb{N}} K(x^k)) \neq \emptyset\) and there exists \(\mathcal{K}\), a pointed, closed and convex cone such that \(K(x^k) \subseteq \mathcal{K}\) for all \(k \in \mathbb{N}\), then every sequence generated by the inexact projected gradient method (IPG Method) is bounded and its accumulation points are weakly efficient solutions.

**Proof.** Let us consider the set \(T := \{x \in C : F(x^k) - F(x) \in K(x^k), \forall k\}\), and take \(\hat{x} \in T\), which exists by Assumption 4.4. Since \(F\) is a \(K\)-convex function and Assumption 4.5 holds, it follows from Lemma 4.10 that

\[
\|x^{k+1} - \hat{x}\|^2 \leq \|x^k - \hat{x}\|^2 + 2\beta_k \gamma_k \langle w^k, J_F(x^k) v^k \rangle,
\]

for all \(k \in \mathbb{N}\). By the definition of \(v^k\), it is a descent condition. This means that \(-J_F(x^k) v^k \in K(x^k)\). Hence \(\langle w^k, J_F(x^k) v^k \rangle \leq 0\). Then,

\[
\|x^{k+1} - \hat{x}\|^2 - \|x^k - \hat{x}\|^2 \leq 2\beta_k \gamma_k \langle w^k, J_F(x^k) v^k \rangle \leq -2\beta_k \gamma_k \langle w^k, J_F(x^k) v^k \rangle.
\]

On the other hand as \(\mathcal{K}\) is a closed, convex and pointed cone with nonempty interior, \(\mathcal{K}^*\) is also a closed, convex and pointed cone with nonempty interior. Since \(K(x^k) \subseteq \mathcal{K}\), it holds that \(\mathcal{K}^* \subseteq K^*(x^k)\). Hence \(\mathcal{K}^* \subseteq \cap_{k \in \mathbb{N}} K^*(x^k)\). Let \(\omega_1, \ldots, \omega_m \in \mathcal{K}^*\) be a basis of \(\mathbb{R}^m\) which exists because \(\text{int}(\mathcal{K}^*) \neq \emptyset\).

Then, there exists \(\alpha^k_1, \ldots, \alpha^k_m \in \mathbb{R}\) such that \(w^k = \sum_{i=1}^m \alpha^k_i \omega_i\). Substituting in (22),

\[
\|x^{k+1} - \hat{x}\|^2 - \|x^k - \hat{x}\|^2 \leq -2\beta_k \gamma_k \sum_{i=1}^m \alpha^k_i \langle \omega_i, J_F(x^k) v^k \rangle.
\]

On the other hand, since \(-J_F(x^k) v^k \in K(x^k)\), \(\omega_1, \ldots, \omega_m \in \mathcal{K}^* \subseteq K^*(x^k)\) and \(\beta_k, \gamma_k > 0\) for all \(k \in \mathbb{N}\), it holds \(\langle \omega_i, -2\beta_k \gamma_k J_F(x^k) v^k \rangle \geq 0\). Without loss of generality, we can assume that \(\|\omega_i\| = 1\). Then, \(\alpha^k_i\) is uniformly bounded, i.e. there exists \(M > 0\) such that for all \(k, i\) \(|\alpha^k_i| \leq M\). Hence,

\[
\|x^{k+1} - \hat{x}\|^2 - \|x^k - \hat{x}\|^2 \leq -2M \beta_k \gamma_k \sum_{i=1}^m \langle \omega_i, J_F(x^k) v^k \rangle.
\]
By the Armijo-type line-search in (12), \( F(x^{k+1}) - F(x^k) - \gamma_k \sigma J_F(x^k) v^k \in -K(x^k) \). Recall that \( \omega_i \in \cap_{k \in \mathbb{N}} K^\ast(x^k) \), we obtain \( \langle \omega_i, F(x^k) - F(x^{k+1}) \rangle \geq \langle \omega_i, -\gamma_k J_F(x^k) v^k \rangle \). It follows from (24) that
\[
\|x^{k+1} - \hat{x}\|^2 - \|x^k - \hat{x}\|^2 \leq 2\frac{M}{\sigma} \beta_k \sum_{i=1}^{m} \langle \omega_i, F(x^k) - F(x^{k+1}) \rangle.
\] (25)

For the Fejér convergence of \( (x^k)_{k \in \mathbb{N}} \) to \( T \), it is enough to prove that the term \( \beta_k \sum_{i=1}^{m} \langle \omega_i, F(x^k) - F(x^{k+1}) \rangle \geq 0 \) is summable at all \( k \in \mathbb{N} \). Since \( \beta_k \leq \hat{\beta} \) for all \( k \in \mathbb{N} \),
\[
\sum_{k=0}^{n} \beta_k \sum_{i=1}^{m} \langle \omega_i, F(x^k) - F(x^{k+1}) \rangle \leq \hat{\beta} \sum_{i=1}^{m} \langle \omega_i, F(x^0) - F(x^{n+1}) \rangle.
\] (26)

As consequence of the Armijo-type line-search, we have \( F(x^k) - F(x^{k+1}) \in K(x^k) \subset K \). So, \( (F(x^k))_{k \in \mathbb{N}} \) is a decreasing sequence with respect to \( K \). Furthermore by Assumption 4.4, it is bounded below, also with respect to the order given by \( K \), by \( F(\hat{x}) \), where \( \hat{x} \in T \). Hence, the sequence \( (F(x^k))_{k \in \mathbb{N}} \) converges and using (26) in the inequality below, we get
\[
\sum_{k=0}^{\infty} \beta_k \sum_{i=1}^{m} \langle \omega_i, F(x^k) - F(x^{k+1}) \rangle = \lim_{n \to \infty} \sum_{k=0}^{n} \beta_k \sum_{i=1}^{m} \langle \omega_i, F(x^k) - F(x^{k+1}) \rangle \\
\leq \hat{\beta} \lim_{n \to \infty} \sum_{i=1}^{m} \langle \omega_i, F(x^0) - F(x^{n+1}) \rangle \\
= \hat{\beta} \sum_{i=1}^{m} \langle \omega_i, F(x^0) - F(\hat{x}) \rangle < +\infty.
\]

So, the quasi-Fejér convergence is fulfilled.

Since \( \hat{x} \) is an arbitrary element of \( T \), it is clear that \( (x^k)_{k \in \mathbb{N}} \) converges quasi-Fejér to \( T \). Hence, by Fact 4.7, it follows that \( (x^k)_{k \in \mathbb{N}} \) is bounded. Therefore, \( (x^k)_{k \in \mathbb{N}} \) has at least one accumulation point, which, by Theorem 4.3 is stationary. By Proposition 2.3, this point is also weakly efficient solution, because \( F \) is \( K \)-convex. Moreover, since \( C \) is closed and the whole sequence is feasible, then this accumulation point belongs to \( C \).  

\[ \blacksquare \]

## 5 Another Variable Order

As was noticed in Section 6 of [6] the variable order structure could formulated in two different ways. Moreover, Examples 3.1 and 3.2 in [6] illustrate the differences of considering one order or the other. Thus, the variable order for the optimization problem may also depend of new order by using the cone valued mapping \( \hat{K} : \mathbb{R}^m \Rightarrow \mathbb{R}^m \) where \( \hat{K}(y) \) is a convex, closed and pointed cone for all \( y \in C \subset \mathbb{R}^m \) where \( C \) is the feasible set. It is worth noting that the domain of the new mapping \( \hat{K} \) is in \( \mathbb{R}^m \) and the orderings considered in the previous sections, are defined by applications whose domain is \( \mathbb{R}^n \). As already discussed in [6], convexity can be defined and convex functions satisfy nice properties such as the existence of subgradients.

Given a closed and convex set \( C \), we say that \( x^* \in C \) solves the optimization problem
\[
\hat{K} - \min F(x) \text{ s.t. } x \in C,
\] (27)
if, for all \( x \in C \),
\[
F(x) - F(x^*) \notin -\mathring{K}(F(x^*)) \setminus \{0\}.
\]

Here we can assume that \( \mathring{K} : F(C) \subseteq \mathbb{R}^m \Rightarrow \mathbb{R}^m \). We shall mention that the main difference between the above problem and (1) yields in the definition of the variable order given now by \( \mathring{K} \).

For a more detailed study of the properties of the minimal points and their characterizations and convexity concept on this case; see [6, 15].

In this framework the definitions of weak solution and stationary point are analogous. The main difference is that instead of \( K(x^*) \), the cone \( \mathring{K}(F(x^*)) \) is considered to define the variable partial order. That is, the point \( x^* \) is stationary, iff for all \( d \in C - x^* \), we have \( J_F(x^*)d \notin -\text{int}(\mathring{K}(F(x^*))) \).

Then, similarly as in the case of problem (1), the following holds.

**Proposition 5.1** If \( F \) is a continuously differentiable function and \( C \) is a convex set, weak solutions of (27) are stationary points. Moreover if \( F \) is also convex with respect to \( \mathring{K} \), the converse is true.

**Proof.** It follows the same lines of the proof of Propositions 2.1 and 2.3: the Taylor expansion of \( F \) and the closedness of \( \mathring{K}(F(x^*)) \) imply the result. \( \blacksquare \)

The inexact algorithm is adapted in the following way

---

**F-Inexact Projected Gradient Method (FIPG Method).** Given \( 0 < \hat{\beta} \leq \hat{\beta}_k \leq \hat{\beta} < +\infty, \delta \in (0, 1) \) and \( \sigma, \gamma \in (0, 1) \).

**Initialization:** Take \( x^0 \in \mathbb{R}^n \) and \( \beta_0 \).

**Iterative step:** Given \( x_k \) and \( \beta_k \), compute \( v^k \), \( \delta \)-approximate solution of \( (Q_{x^k}) \). If \( v^k = 0 \), then stop. Otherwise compute
\[
\ell(k) := \min \left\{ \ell \in \mathbb{Z}_+ : F(x^k) + \sigma \gamma^\ell J_F(x^k)v^k - F(x^k + \gamma^\ell v^k) \in \mathring{K}(F(x^k)) \right\}.
\]

Set \( x^{k+1} = x^k + \gamma_k v^k \in C \), with \( \gamma_k = \gamma^{\ell(k)} \).

---

Here the auxiliary problem \( (Q_{x^k}) \) is defined as
\[
\min_{v \in C - x^k} \left\{ \frac{\|v\|^2}{2} + \beta_k \phi(x^k, v) \right\},
\]
where \( \phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \),
\[
\phi(x, v) := \max_{y \in G(F(x))} y^T J_F(x)v,
\]
for \( G : \mathbb{R}^m \Rightarrow \mathbb{R}^m \) generator of \( \mathring{K}^*(F(x)) := [\mathring{K}(F(x))]^* \).

With this ordering, the function \( \phi \) characterizes the stationarity. Furthermore, subproblem \( (Q_{x^k}) \) has a unique solution which is \( v^k = 0 \) if and only if \( x^k \) is a stationary point. Results analogous to those proven in Propositions 3.3 and 3.5 are also true. These facts implies that **FIPG Method** is well defined, i.e., if it stops, then the computed point is a stationary point and in other case there exists \( \ell(k) \) which satisfies the Armijo-type line-search (28). So, only the convergence of a sequence generated by it must be studied.

As in the last section, we analyze the convergence of the functional values sequence \((F(x^k))_{k \in \mathbb{N}}\).
**Lemma 5.2** Suppose that $x^*$ is an accumulation point of $(x^k)_{k \in \mathbb{N}}$ of the sequence generated by **FIPG Method**. If $\bigcup_{x \in C} \hat{K}(F(x)) \subseteq \mathcal{K}$, where $\mathcal{K}$ is a closed, pointed and convex cone, $G(F(x))$ is a closed application such that $d_H(G(F(x)), G(F(\hat{x}))) \leq L_{GF}\|x - \hat{x}\|$, for all $x, \hat{x} \in C$, then $\lim_{k \to \infty} F(x^k) = F(x^*)$.

**Proof.** The result is again proven by the existence of a non-increasing sequence with an accumulation point.

Next, with the help of the last Lemma, we summarize the convergence of the generated sequence with the following result.

**Theorem 5.3** Suppose that

(a) $\bigcup_{x \in C} \hat{K}(F(x)) \subseteq \mathcal{K}$, where $\mathcal{K}$ is a closed, pointed and convex cone.

(b) The application $G \circ F$ is closed.

(c) $d_H(G(F(x)), G(F(\hat{x}))) \leq L_{GF}\|x - \hat{x}\|$, for all $x, \hat{x} \in C$.

(d) $J_F(x)$ is a locally Lipschitz function.

If $(\beta_k)_{k \in \mathbb{N}}$ is a bounded sequence, then all accumulation points of $(x^k)_{k \in \mathbb{N}}$ generated by **FIPG Method** are stationary points of problem (27).

**Proof.** It follows from the same lines of the proof of Theorem 4.3.

We want to point out that in the last theorem the condition

$$d_H(G(F(x)), G(F(\hat{x}))) \leq L_{GF}\|x - \hat{x}\|, \quad \forall x, \hat{x} \in C$$

replaces the similar one given in Theorem 4.3, i.e., $d_H(G(y), G(\hat{y})) \leq L_{G}\|y - \hat{y}\|$, for all $y, \hat{y} \in C$. Moreover, if $F$ is Lipschitz on $C$, then this last condition implies the condition (c) in Theorem 5.3 by taking $y = F(x)$ and $\hat{y} = F(\hat{x})$.

Next result is an extension to the variable order setting of Theorem 5.2 of [20].

**Theorem 5.4** Assume that $F$ is $\hat{K}$-convex and additionally:

(a) If $(z^k)_{k \in \mathbb{N}} \subset F(C)$ is a sequence such that $z^k - z^{k+1} \in \hat{K}(F(z^k))$ for all $k \in \mathbb{N}$ and $z \in C, z^k - z \in \mathcal{K}$ for some closed, convex and pointed cone $\mathcal{K}$, $\bigcup_{k \in \mathbb{N}} \hat{K}(F(z^k)) \subseteq \mathcal{K}$, then there exists $\hat{z} \in C$ such that $F(\hat{z}) \preceq z^k$ for all $k \in \mathbb{N}$.

(b) The search direction $v^k$ is s-compatible at $x^k$, i.e., $v^k = P_{C-x^k}(-\beta_F(x^k)^T w^k)$, where $w^k \in \text{conv}(G(x^k))$, for all $k \in \mathbb{N}$.

(c) $\text{int}(\bigcap_{k \in \mathbb{N}} \hat{K}(F(x^k))) \neq \emptyset$.

(d) There exists $\mathcal{K}$, a pointed, closed and convex cone such that $\hat{K}(F(x^k)) \subseteq \mathcal{K}$ for all $k \in \mathbb{N}$.

Then every sequence generated **FIPG Method** is bounded and its accumulation points are weakly efficient solutions.

**Proof.** It follows from the same lines of the proof of Theorem 4.11 using now the new variable order structure.
6 Illustrative Examples

In this section we present three examples two for problem (1) and one for (27), illustrating how both proposed methods work starting at ten different random initial points. We verify our assumptions in each problem and make some comparisons between the proposed methods by changing the exactness of the approximate gradient direction and its exact versions.

The algorithms were implemented in MatLab R2012 and ran at a Intel(R) Atom(TM) CPU N270 at 1.6GHz. All starting points are not solutions and are randomly generated. Despite the fact that it may not be an easy task to compute the positive dual cone of a given one, the computation of (approximate) directions is, in general, complicated. Indeed, after the definition, the optimal value of problem \((P_x)\) must be known. The use of \(s\)-compatible directions at the iteration \(k\) of the proposed methods, see Definition 4.8, is recommended in the case in which the exact projection onto the set \(C - x_k\) is not too complicated. This is the case of the set of feasible solutions of the next example. Clearly in all examples below the defining order cone-valued mappings are closed applications and Lipschitz with respect to the Hausdorff distance. The stopping criteria were \(\|\sigma^k\| < 10^{-4}\) and also when the 30-iterations is reached. The solutions were displayed with four digits, CPU time was recorded in seconds and moreover, the number of iterations was also shown in each case.

Example 6.1 We consider the vector problem as (1) with

\[
K - \min F(x) = (x + 1, x^2 + 1), \text{ s.t. } x \in [0, 1],
\]

where \(F : \mathbb{R} \rightarrow \mathbb{R}^2\) and the variable order is given by \(K : \mathbb{R} \Rightarrow \mathbb{R}^2\),

\[
K(x) := \{(z_1, z_2) \in \mathbb{R}^2 : z_1 \geq 0, (x^2 + 1)z_1 - (x + 1)z_2 \leq 0\}.
\]

In this model the closed interval \([0, \sqrt{2} - 1] \approx [0, 0.4142]\) is the set of minimizers.

**IPG Method** was run ten times by using ten random initial points, out of the set of minimizers, ended at solution points, which are obtained after the verification of the stopping criterion. The method gives the following data:

<table>
<thead>
<tr>
<th>Instances</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Points</td>
<td>0.6557</td>
<td>0.6948</td>
<td>0.8491</td>
<td>0.9340</td>
<td>0.6787</td>
<td>0.7577</td>
<td>0.7431</td>
<td>0.4387</td>
<td>0.6555</td>
<td>0.9502</td>
</tr>
<tr>
<td>Solutions</td>
<td>0.4115</td>
<td>0.4128</td>
<td>0.4140</td>
<td>0.4135</td>
<td>0.4116</td>
<td>0.4131</td>
<td>0.4127</td>
<td>0.4136</td>
<td>0.4114</td>
<td>0.4130</td>
</tr>
<tr>
<td>CPU Time</td>
<td>0.0001</td>
<td>0.0250</td>
<td>0.0001</td>
<td>0.0156</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.1094</td>
<td>0.0156</td>
<td>0.0781</td>
</tr>
<tr>
<td>N(#) Iterations</td>
<td>16</td>
<td>19</td>
<td>23</td>
<td>26</td>
<td>17</td>
<td>20</td>
<td>20</td>
<td>4</td>
<td>16</td>
<td>28</td>
</tr>
</tbody>
</table>

Note that in all cases optimal solutions were computed and that the solution points are in the set of optimal solutions. Moreover, to illustrate how the inexactness of **IPG Method** affect the performance (average of CPU time and the average of the number of iterations for 10 instances) of the convergence is considering for the above example for different values of \(\delta\) the approximate solution parameter.

<table>
<thead>
<tr>
<th>(\delta) – Approximation Parameters</th>
<th>(\delta = 0)</th>
<th>(\delta = 0.25)</th>
<th>(\delta = 0.5)</th>
<th>(\delta = 0.75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average of CPU Time</td>
<td>0.4819</td>
<td>0.2232</td>
<td>0.0244</td>
<td>0.0251</td>
</tr>
<tr>
<td>Average of N(#) Iterations</td>
<td>9</td>
<td>15</td>
<td>19</td>
<td>21</td>
</tr>
</tbody>
</table>

The above table shows that when the inexactness increases, CPU time decreases but the number of iterations increases.
Next example is a non-convex problem corresponding to the model studied in the previous section.

**Example 6.2** [cf. Example 4.13 of [14]] Consider the following vector problem as (27)

\[
\bar{K} - \min_{x} F(x) = (x_1, x_2), \text{ s.t. } \pi \leq x_1^2 + x_2^2 \leq 2\pi,
\]

where \(F : \mathbb{R}^2 \rightarrow \mathbb{R}^2\) and the variable order is leading by the cone \(\hat{K} : \mathbb{R}^2 \rightharpoonup \mathbb{R}^2\),

\[
\hat{K}(y) := \left\{ z = (z_1, z_2) \in \mathbb{R}^2 : \|z\|_2 \leq \left[ \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} y \right]^T z / \pi \right\}.
\]

The set of solutions (stationary points) of this problem is directly computed by using the definition of \(S^s\) given in (2).

\[
\{ (x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = \pi \text{ or } x_1^2 + x_2^2 = 2\pi \}.
\]

**FIPG Method** computes the following points:

<table>
<thead>
<tr>
<th>Instances</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Points</td>
<td>1.8650</td>
<td>1.7525</td>
<td>2.4190</td>
<td>1.9573</td>
<td>0.7931</td>
<td>1.2683</td>
<td>1.8135</td>
<td>-2.0485</td>
<td>-0.6446</td>
<td>-0.8561</td>
</tr>
<tr>
<td>Solutions</td>
<td>1.1632</td>
<td>0.9850</td>
<td>1.7705</td>
<td>1.7492</td>
<td>0.8634</td>
<td>1.4208</td>
<td>1.7456</td>
<td>-1.7438</td>
<td>-0.8016</td>
<td>-0.9535</td>
</tr>
<tr>
<td>CPU Time</td>
<td>0.2969</td>
<td>0.2344</td>
<td>0.2188</td>
<td>0.1719</td>
<td>0.1563</td>
<td>0.5625</td>
<td>0.2656</td>
<td>0.1875</td>
<td>0.2031</td>
<td>0.3125</td>
</tr>
<tr>
<td>No ({\}) Iterations</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>17</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Note that the solutions computed at all iterations of the proposed method belong to the set of optimal solutions.

The last example is widely studied in Section 9.2 of the book [15].

**Example 6.3** [cf. Example 9.5 of [15]] Consider the following problem

\[
K - \min_{x} F(x) = (x_1, x_2), \text{ s.t. } x \in C
\]

where

\[
C := \left\{ x = (x_1, x_2) \in [0, \pi] \times [0, \pi] : \left| x_1^2 + x_2^2 - 1 - \frac{1}{10} \cos \left( 16 \arctan \left( \frac{x_1}{x_2} \right) \right) \right| \geq 0, \right.
\]

\[
(x_1 - 0.5)^2 + (x_2 - 0.5)^2 \leq 0.5.
\]

For the variable ordering structure is considered the map \(K : \mathbb{R}^2 \rightharpoonup \mathbb{R}^2\), by

\[
K(x) := \left\{ z \in \mathbb{R}^2 : \|z\|_2 \leq \frac{2}{\min_{i=1,2} x_i} x_i^T z \right\}.
\]

**IPG Method** stating at ten random initial points performs as following:

<table>
<thead>
<tr>
<th>Instances</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Point</td>
<td>0.9735</td>
<td>0.7932</td>
<td>0.8403</td>
<td>0.9847</td>
<td>0.7508</td>
<td>0.9786</td>
<td>0.9790</td>
<td>0.9679</td>
<td>0.8082</td>
<td>0.8965</td>
</tr>
<tr>
<td>Solution</td>
<td>0.9011</td>
<td>0.7407</td>
<td>0.7854</td>
<td>0.9096</td>
<td>0.7004</td>
<td>0.9050</td>
<td>0.9054</td>
<td>0.8967</td>
<td>0.7551</td>
<td>0.7754</td>
</tr>
<tr>
<td>CPU Time</td>
<td>0.0938</td>
<td>0.0313</td>
<td>0.0625</td>
<td>0.0313</td>
<td>0.0469</td>
<td>0.0156</td>
<td>0.0313</td>
<td>0.0313</td>
<td>0.0313</td>
<td>0.0156</td>
</tr>
</tbody>
</table>

In this case, the maximum number of iterations is achieved. Nevertheless, good approximations to minimal elements have been computed at each of the ten instances presented in the above table.
For all the above examples, we have also run the respectively exact version by taking the same initial points and the comparison between them in term of the CPU time (in seconds) are shown in the following figures.

![Example 6.1](image1)

![Example 6.2](image2)

![Example 6.3](image3)

This shows that the inexact versions are significantly faster (CPU Time), in almost all instances, than the exact ones. However, for all initial data of the above examples, the exact versions of the proposed methods take fewer iterations than the inexact ones to converge to a solution. It is worth emphasizing that the exact versions have to solve exactly the harder subproblems $P_{x_k}$ and $Q_{x_k}$ to find the descent direction at each iteration $k$. This is a serious drawback for the computational implementation point of view (avoidable for the above examples), making the exact implementation inefficient in general.

7 Final Remarks

The projected gradient method is one of the classical and basic schemes for solving constrained optimization problems. In this paper, we have extended the exact and unconstrained scheme proposed in [20]. The new proposed scheme solves now smooth and constrained vector optimization problems under a variable ordering by taking inexact descent directions. This inexact projected approach promotes future research on other efficient variants for these kind of problems. As it is shown in the examples above it is more effective and implementable than the exact one. Moreover, now constrained variable optimization problems can be solved by using the proposed methods. However, the full convergence of the generated sequence to a weakly efficient solution is still an open problem in variable order setting.

Future work will be addressed to investigate this theoretical issue; also for some particular instances in which the objective function is a non-smooth function, extending the projected subgradient method proposed in [4] to the variable order setting. The numerical behavior of these approaches under $K$-convexity of the non-smooth objective function remains open and is a promising direction to be investigated. Despite its computational shortcomings, it hopefully sets the foundations of future more efficient and general algorithms for this setting.

Acknowledgements: JYBC was partially supported by the National Science Foundation (NSF) Grant DMS - 1816449. The results of this paper were developed during the stay of the second author at the University of Halle-Wittenberg supported by the Humboldt Foundation. The authors would like to extend their gratitude towards anonymous referees whose suggestions helped us to improve the presentation of this manuscript.
References


