

Characterizing FJ and KKT conditions in nonconvex mathematical programming with applications*

Fabián Flores-Bazán[†]

Giandomenico Mastroeni[‡]

Abstract

In this paper we analyse the Fritz John and Karush-Kuhn-Tucker conditions for a (Gâteaux) differentiable nonconvex optimization problem with inequality constraints and a geometric constraint set. Fritz John conditions are characterized in terms of an alternative theorem which covers beyond standard situations; while characterizations of KKT conditions, without assuming constraints qualifications, are related to strong duality of a suitable linear approximation of the given problem and the properties of its associated image mapping. Such characterizations are suitable for dealing with some problems in structural optimization, where most of the known constraint qualifications fail. In particular, several examples are exhibited showing the usefulness and optimality, in a certain sense, of our results, which even provide much more information than those (including Mordukhovich normal cone or Clarke's one) appearing elsewhere. The case with a single inequality constraint is discussed in details by establishing a hidden convexity in the validity of the KKT conditions. We outline possible applications to a class of mathematical programs with equilibrium constraints (MPEC), as well as to vector equilibrium or quasi-variational inequality problems.

Keywords: FJ conditions, KKT conditions, hidden convexity, strong duality, nonconvex optimization, MPEC.

AMS subject classification: 90C26, 90C46, 90C31, 90C90.

1 Introduction

No doubt that the Karush-Kuhn-Tucker (KKT) optimality conditions remain as a master piece in the development of nonlinear programming, but the Fritz John (FJ) conditions still play a role since no additional requirement is needed as for standard situations.

*This research, for the first author, was supported in part by CONICYT-Chile through FONDECYT 112-0980 and BASAL projects, CMM, Universidad de Chile.

[†]Departamento de Ingeniería Matemática, CI²MA, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Casilla 160-C, Concepción, Chile. E-mail: fflores@ing-mat.udec.cl

[‡]Department of Computer Science, University of Pisa, L.go B. Pontecorvo, 3, 56127 Pisa, Italy. E-mail: gmastroeni@di.unipi.it

KKT conditions were introduced as a generalization of the classical Lagrange multiplier theorem to inequality-constrained problems that nowadays lie on the core of mathematical programming.

Let us consider the following problem with explicit inequality constraints:

$$\min\{f(x) : g_i(x) \leq 0, \quad i = 1, \dots, m, \quad x \in X\}, \quad (1.1)$$

where $f, g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$, are given functions satisfying some differentiability assumptions at $\bar{x} \in X$, and $X \subseteq \mathbb{R}^n$ is usually termed the “geometric constraint set”; whereas

$$K \doteq \{x \in X : g_i(x) \leq 0, \quad i = 1, \dots, m\},$$

is the feasible set to (1.1). For fixed $\bar{x} \in X$, we denote by $I = I(\bar{x}) \doteq \{i : g_i(\bar{x}) = 0\}$ its active index set.

In the simplest situation where X is an open set, the most general optimality conditions for (1.1), due to Fritz John, assert that, if \bar{x} is a local optimal solution to (1.1), then there exist $\lambda_0 \geq 0$, $\lambda_i \geq 0$, $i \in I$, not all zero, satisfying

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) = 0.$$

This is just a consequence of a convex separation theorem. This condition may not be of much interest if the multiplier λ_0 associated with the objective function is zero: the “so called” constraints qualifications (CQ) guarantee that $\lambda_0 > 0$ thus leading to KKT conditions [29, 30]. CQ are conceived in order to prove the fulfillment of KKT conditions whatever the objective function, having \bar{x} as a local minimum on K , is. This property turns out to be equivalent to the Guignard CQ, which has proved to be the more general one including that due to Abadie. Since the pioneering works of Karush, Kuhn and Tucker, many authors have developed such a result in different directions: under standard differentiability assumptions or without such requirements, see [4, 7, 25, 8, 26, 27, 28, 30, 31] and [11, 22, 23].

When $X = \mathbb{R}^n$, another approach uses some penalty functions and their exactness, which means that local solutions to the original problem are also unconstrained local minimizers of the penalty function. A scheme of this kind is presented in [38], where a KKT condition is derived under exactness (being equivalent to a generalized calmness-type condition as shown in [10, 12, 35]) and a certain CQ assumption [38, Theorem 2.1], which involves upper second-order directional derivatives of g_i . In the same paper it is proved that, under exactness, their CQ has no relation with the Guignard CQ, and that exactness alone does not imply the fulfillment of the KKT condition. The same argument, via exactness of the penalty function, is employed in [32] for a minimization problem under complementarity constraints. This is given in terms of some optimality indications sets as introduced in the same paper. Nevertheless, there are situations where our results, as Corollary 5.5 for instance, apply but not [38, Theorem 2.1], see Remark 5.6.

Very recently, enhanced FJ conditions for convex programming and enhanced KKT conditions, under some CQ, have been established in [9] and [39], respectively; whereas further developments of CQ, in view of numerical aspects, are stated in [2, 3].

Precisely, the aim of this paper is to analyze the fulfillment of the FJ and the KKT optimality conditions without any CQ or convexity. In other words, our goal is to characterize the validity of the FJ (Theorem 3.1) and KKT (Theorems 4.1 and 4.3) conditions through the analysis of a suitable linear approximation of the given problem, and the properties of its associated image mapping involving the gradients of the objective function and the active inequality constraints at a (not necessarily local optimal) point $\bar{x} \in X$. Actually, we will see that our Theorem 3.1 provides more information than that involving the Clarke tangent cone, the open cone of interior or quasi-interior directions, and even the normal cone of Mordukhovich, see Example 3.3. Our characterizations will be useful when dealing with problems where most of the constraint qualifications fail as it occurs, for instance, in structural optimization [1].

Unlike the above mentioned results requiring CQ, which involve g_i , $i \in I$, and X , that guarantee KKT conditions for every function f having \bar{x} as a local minimum on K ([26, 27]), our approach allows us to derive assumptions on f , g_i and X , jointly, that ensure the fulfillment of KKT conditions. Thus, we expect to cover situations (including geometric constraints) where the constraints qualification due to Guignard, for instance, fails (see Example 4.5).

We analyze in details the connections between our results and the Guignard along with Mangasarian-Fromovitz CQ [28, 31], providing several examples where our regularity conditions are applicable while the previous ones are not, as expected. Certainly the latter happens with any constraints qualification, recalling that the Guignard CQ is weaker than any other one [26].

The paper is organized as follows. In Section 2, basic definitions and preliminary results on separation and alternative theorems, including a characterization of strong duality, are recalled. Section 3 is devoted to FJ optimality conditions of the alternative type, while Section 4 deals with KKT conditions providing several characterizations in terms of strong duality for a suitable linearization of the given problem and the properties of the associated image. Section 5 is devoted to the analysis of a constrained optimization problem with a single inequality constraint. In particular the convexity of the conical approximation of the geometric constraint set appears as a sufficient and necessary condition (hidden convexity), and most of the results in Section 4 can be stated in a more handfull way. An application of the results of Section 5 is presented in Section 6, where a class of MPEC can be formulated as a minimization problem with a single inequality constraint. In Section 7, we outline further applications of KKT conditions for vector equilibrium problems and quasi-variational inequalities (QVI). Indeed, in [13] an algorithm based on KKT conditions for (QVI) is developed.

2 Basic definitions and preliminary results

Throughout this paper by a cone $P \subseteq \mathbb{R}^n$ we mean a set such that $tP \subseteq P$, for all $t \geq 0$. Given a set $A \subseteq \mathbb{R}^n$, its closure is denoted by \bar{A} ; its convex hull by $\text{co}(A)$ which is the smallest convex set containing A ; its topological interior by $\text{int } A$; $\text{ri } A$ is the relative interior of A , whenever it is convex; $\text{cone}(A)$ stands for the smallest cone containing A , that is,

$$\text{cone}(A) = \bigcup_{t \geq 0} tA,$$

whereas $\overline{\text{cone}}(A)$ denotes the smallest closed cone containing A : obviously $\overline{\text{cone}}(A) = \overline{\text{cone}(A)}$. By $\langle \cdot, \cdot \rangle$ we denote the inner or scalar product in \mathbb{R}^n whose elements are considered column vectors. Thus, $\langle a, b \rangle = a^\top b$ for all $a, b \in \mathbb{R}^n$.

Given a nonempty set $P \subseteq \mathbb{R}^n$, its polar cone, P^* , is defined as

$$P^* = \{\xi \in \mathbb{R}^n : \langle \xi, p \rangle \geq 0 \ \forall p \in P\}.$$

It is well known that, whenever P is a closed convex cone, we have (the bipolar theorem) $P = P^{**} \doteq (P^*)^*$, and in general we have $P^{**} = \overline{\text{co}}(\text{cone } P)$, where $\overline{\text{co}} A$ is the smallest closed and convex set containing A .

We say that a (not necessarily convex) cone, P , is pointed if $\text{co } P \cap (-\text{co } P) = \{0\}$. Notice that since $\text{co}(\text{cone } A) = \text{cone}(\text{co } A)$,

$$\text{cone}(A) \text{ is pointed} \iff \text{cone}(\text{co } A) \text{ is pointed.}$$

Definition 2.1. Let $\emptyset \neq K \subseteq \mathbb{R}^n$ and $\bar{x} \in \overline{K}$, the *Contingent cone* of K at \bar{x} , denoted by $T(K; \bar{x})$, is the set

$$T(K; \bar{x}) \doteq \left\{ v \in \mathbb{R}^n : \exists t_k > 0, \exists x_k \in K, x_k \rightarrow \bar{x}, t_k(x_k - \bar{x}) \rightarrow v \right\}.$$

When K is convex and $\bar{x} \in K$, then it is not hard to prove that

$$T(K; \bar{x}) = \overline{\bigcup_{t \geq 0} t(K - \bar{x})}.$$

Given a convex set $A \subseteq \mathbb{R}^n$, the (outward) normal cone (in the sense of convex analysis) of A at $\bar{x} \in A$, is the set

$$N(A; \bar{x}) \doteq \{\xi \in \mathbb{R}^n : \langle \xi, x - \bar{x} \rangle \leq 0, \ \forall x \in A\}.$$

Thus, $(T(A; \bar{x}))^* = -N(A; \bar{x})$ whenever $\bar{x} \in A$.

Furthermore, we denote $\mathbb{R}_+ \doteq [0, +\infty[$, $\mathbb{R}_{++} \doteq]0, +\infty[$ and $\mathbb{R}_{++}^n = \text{int } \mathbb{R}_+^n$.

The next theorem appears in a more general framework in [19], and it may be considered as a Gordan-type alternative theorem. The proof relies on standard separation arguments.

Theorem 2.2. ([19, Theorem 3.2]) *Let $P \subseteq \mathbb{R}^m$ be a convex closed cone such that $\text{int } P \neq \emptyset$, and $A \subseteq \mathbb{R}^m$ be any nonempty set. Then the following assertions are equivalent:*

- (a) $\exists \lambda^* \in P^* \setminus \{0\}, \langle \lambda^*, a \rangle \geq 0, \ \forall a \in A$;
- (b) $\text{cone}(A + \text{int } P)$ is pointed;
- (c) $\text{co}(A) \cap (-\text{int } P) = \emptyset$.

We can go further when A is the image of a subset $C \subseteq \mathbb{R}^n$ through a linear transformation $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$. The next proposition is taken from [17].

Proposition 2.3. ([17, Proposition 2.7]) *Let F be a real matrix of order $m \times n$, and write $F^\top = (F_1^\top \ \cdots \ F_m^\top)$, where F_i is the i -th row of F . Let $C \subseteq \mathbb{R}^n$ be any nonempty set. Then*

$$F(C) \cap (-\mathbb{R}_{++}^m) = \emptyset \iff \max_{1 \leq i \leq m} \langle F_i^\top, v \rangle \geq 0 \ \forall v \in \overline{C},$$

and the following statements are equivalent:

- (a) $\text{cone}(F(C) + \mathbb{R}_{++}^m)$ is pointed;
- (b) $F(\overline{\text{co}}(C)) \cap (-\mathbb{R}_{++}^m) = \emptyset$;
- (c) $F(\text{co}(C)) \cap (-\mathbb{R}_{++}^m) = \emptyset$;
- (d) $\max_{1 \leq i \leq m} \langle F_i^\top, v \rangle \geq 0 \quad \forall v \in \overline{\text{co}}(C)$;
- (e) $\text{co}(\{F_i^\top : i = 1, \dots, m\}) \cap C^* \neq \emptyset$.

In case $m = 2$, Proposition 2.3 takes a more precise formulation that follows. It will be applied in Section 5 to obtain various characterizations to the validity of the Fritz John conditions under a single inequality constraint. Such a formulation is a consequence of a more general result that appeared in [19].

Proposition 2.4. ([19, Theorems 3.2 and 4.1]) *Let $m = 2$ in the preceding proposition. Then, the following assertions are equivalent:*

- (a) $\exists \lambda^* \geq 0, \lambda^* \neq 0, \langle \lambda^*, a \rangle \geq 0 \quad \forall a \in F(C)$;
- (b) $F(C) \cap (-\mathbb{R}_{++}^2) = \emptyset$ and $\overline{F(C) + \mathbb{R}^2}$ is convex;
- (c) $F(C) \cap (-\mathbb{R}_{++}^2) = \emptyset$ and $F(C) + \mathbb{R}_{++}^2$ is convex;
- (d) $F(C) \cap (-\mathbb{R}_{++}^2) = \emptyset$ and $F(C) + \mathbb{R}^2$ is convex;
- (e) $(F(C) + \mathbb{R}_{++}^2) \cup \{(0, 0)\}$ is pointed.

We will need a further characterization of strong duality. Let Y be a locally convex Hausdorff topological vector space, C be any nonempty set, $\varphi : C \rightarrow \mathbb{R}$, $\psi : C \rightarrow Y$ be any functions, and P be a closed and convex cone in Y . We consider the problem

$$\mu \doteq \inf_{x \in K} \varphi(x), \tag{2.1}$$

where $K \doteq \{x \in C : \psi(x) \in -P\}$.

Let us introduce, as usual, the Lagrangian

$$L(\lambda^*, x) = \varphi(x) + \langle \lambda^*, \psi(x) \rangle.$$

We set $F = (\varphi, \psi)$ so that $F(C) = \{(\varphi(x), \psi(x)) \in \mathbb{R} \times Y : x \in C\}$.

The next result may be found in [20, Theorem 3.2]. It regards a geometric characterization of the strong duality property for (2.1).

Theorem 2.5. *Assume that μ be finite. The following assertions are equivalent:*

- (a) strong duality holds for (2.1), i. e., there exists $\lambda_0^* \in P^*$ such that

$$\inf_{x \in K} \varphi(x) = \inf_{x \in C} L(\lambda_0^*, x);$$

- (b) $\overline{\text{cone}}(\text{co} [F(C) - \mu(1, 0) + (\mathbb{R}_+ \times P)]) \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset$.

The previous theorem, as expected, covers situations where the Slater condition fails. See [20] for details.

3 Fritz John optimality conditions of the alternative-type under inequality constraints

As mentioned in the introduction, the Fritz John optimality conditions will be formulated as an alternative-type result, encompassing existing similar ones for the minimization problem with inequality constraints (1.1).

A natural linear approximation to problem (1.1) arises:

$$\nu_0 \doteq \inf_{v \in G_0(\bar{x})} \nabla f(\bar{x})^\top v, \quad (3.1)$$

where

$$G_0(\bar{x}) \doteq \left\{ v \in \overline{\text{co}} B : \nabla g_i(\bar{x})^\top v < 0, \forall i \in I \right\},$$

and B is a nonempty set in \mathbb{R}^n . As usual, by convention we set $\nu_0 = +\infty$ if $G_0(\bar{x}) = \emptyset$. Obviously this case is the more interesting one. Throughout this paper $f, g_i, i \in I$, are differentiable at $\bar{x} \in X$. Sometimes we will require the continuity of g_i for $i \notin I$. Based on Proposition 2.3 we now establish an alternative-type version of the Fritz John optimality conditions valid at points not necessarily local minima. It has its origin in [16]. Notice that as the set B is bigger, the relation (3.3) is more informative.

Theorem 3.1. (*Fritz John necessary optimality conditions of alternative-type*) *Let us consider problem (1.1), $\bar{x} \in X$, with $X \subseteq \mathbb{R}^n$; $f, g_i, i \in I$, be as above, and let B be a nonempty set in \mathbb{R}^n . Then, exactly one of the following two assertions holds:*

(a) *there exists $v \in \overline{\text{co}} B$ such that*

$$\begin{aligned} \langle \nabla f(\bar{x}), v \rangle &< 0, \\ \langle \nabla g_i(\bar{x}), v \rangle &< 0, \quad i \in I; \end{aligned} \quad (3.2)$$

(b) *there exist $\lambda_0 \geq 0, \lambda_i \geq 0, i \in I$, not all zero, satisfying*

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in B^*, \quad (3.3)$$

or equivalently, $\max_{i \in I} \{ \langle \nabla f(\bar{x}), v \rangle, \langle \nabla g_i(\bar{x}), v \rangle \} \geq 0, \forall v \in \overline{\text{co}} B$.

Furthermore, under the additional assumption that B is a cone, we also have [(a) $\iff \nu_0 = -\infty$], and if $G_0(\bar{x}) \neq \emptyset$, then $\lambda_0 > 0$ and [(b) $\iff \nu_0 = 0$].

Proof. This is a direct application of Proposition 2.3 to

$$F \doteq \begin{pmatrix} \nabla f(\bar{x})^\top \\ \nabla g_I(\bar{x})^\top \end{pmatrix},$$

where $\nabla g_I(\bar{x})^\top$ is the matrix having as rows $\nabla g_i(\bar{x})^\top$ for $i \in I$, and $C = B$.

We prove the second part. The equivalence (a) $\iff \nu_0 = -\infty$ easily follows.

By Proposition 2.3, (b) is equivalent to the impossibility of system (3.2), that is, $\nu_0 = 0$ whenever $G_0(\bar{x}) \neq \emptyset$. Under this latter assumption, we also get $\lambda_0 > 0$, (since otherwise a contradiction is obtained by Proposition 2.3), which is expected because such an assumption is of Slater-type, and implies strong duality for the problem (3.1). This fact will be discussed in more detail at the end of this section. \square

Notice that condition (3.3) can be written as

$$\begin{cases} \lambda_0 \nabla f(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) \in B^*; \\ \lambda_i g_i(\bar{x}) = 0, \quad i = 1, \dots, m, \end{cases} \quad (3.4)$$

provided each g_i is differentiable at \bar{x} .

A condition close to the inconsistency of (3.2) is obtained if \bar{x} is a local solution to (1.1). To be more precise, it is well-known that if g_i is differentiable at \bar{x} , for $i \in I$, and g_i is continuous at \bar{x} for $i \notin I$, local minimality of \bar{x} implies that the system: $v \in T(X; \bar{x})$

$$\begin{cases} \langle \nabla f(\bar{x}), v \rangle < 0, \\ \langle \nabla g_i(\bar{x}), v \rangle < 0, \quad i \in I, \end{cases} \quad (3.5)$$

has no solution, that is, (a) of Theorem 3.1 does not hold, provided that B is a convex subset of $T(X; \bar{x})$, and therefore (b) is true. This is expressed in the following result, which appears in the book by Giorgi, Guerraggio and Thierfelder [25], but its origin goes back to [6]: this result is a direct consequence of our Theorem 3.1. Observe that if $T(X; \bar{x})$ is convex then the next corollary, with $B = T(X; \bar{x})$, provides as much information as possible.

Corollary 3.2. ([25, Theorem 3.6.5]) *Let us consider Problem (1.1) and $\bar{x} \in K$. Let $X \subseteq \mathbb{R}^n$ and B be a nonempty set in \mathbb{R}^n satisfying $\overline{\text{co}} B \subseteq T(X; \bar{x})$. Let f, g_i be as above. If \bar{x} is a local solution to (1.1), then there exist $\lambda_0 \geq 0$, $\lambda_i \geq 0$, $i \in I$, not all zero, satisfying*

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in B^*. \quad (3.6)$$

Note that the assumption $\overline{\text{co}} B \subseteq T(X; \bar{x})$ is equivalent to $\overline{\text{co}}(\text{cone } B) \subseteq T(X; \bar{x})$.

We remark that (3.6) does not include the special case $B^* = -N_M(X; \bar{x})$ where $N_M(X; \bar{x})$ is the Mordukhovich normal cone of X at \bar{x} , since it is, in general, nonconvex, and therefore may not be the polar to any tangential approximation to X at \bar{x} (see [33]). Nevertheless, as it will be seen in Example 3.3, B may be the Clarke tangent cone or the cone of the interior or quasi-interior directions.

The next example, with $T(X; \bar{x})$ being nonconvex, shows an instance where our Theorem 3.1 is applicable providing much more information than Corollary 3.2, even when the Mordukhovich normal cone is considered.

Example 3.3. Consider Problem (1.1) with $f(x_1, x_2) = x_1$, $g(x_1, x_2) = x_2$, and

$$X = \{(x_1, x_2) : x_1 x_2 = 0, x_1 \geq 0, x_2 \geq 0\}, \quad \bar{x} = (0, 0).$$

Then, $T(X; \bar{x}) = X$, which is nonconvex. It follows that $[T(X; \bar{x})]^* = \mathbb{R}_+^2 = \overline{\text{co}}[T(X; \bar{x})]$, and $\bar{x} = (0, 0)$ is a minimum of f on $\{(x_1, x_2) \in X : g(x_1, x_2) \leq 0\}$. Easy computations show that the system (3.2) with $B = T(X; \bar{x})$ (or even for $B = \{x \in \mathbb{R}^2 : p_0^\top x \geq 0\}$, $p_0 \in \mathbb{R}_+^2 \setminus \{0\}$, which contains properly $T(X; \bar{x})$), has no solution, and therefore there exist $\lambda_0, \lambda_1 \geq 0$, not both zero, satisfying

$$\lambda_0 \nabla f(\bar{x}) + \lambda_1 \nabla g(\bar{x}) \in [T(X; \bar{x})]^*. \quad (3.7)$$

Actually any $\lambda_0 \geq 0$ and $\lambda_1 \geq 0$ satisfies (3.7).

$T(X; \bar{x})$ being nonconvex, the only non-trivial convex subcones are

$$B_1 = \{(x_1, 0) : x_1 \geq 0\}, B_2 = \{(0, x_2) : x_2 \geq 0\}.$$

Any of these cones provides, via (3.6), less information than (3.7). Other candidates for a subcone of $T(X, \bar{x})$ are: the Clarke tangent cone of X at \bar{x} , $T_C(X; \bar{x})$, which is always convex, is $\{(0, 0)\}$ in our case; the open cone of interior directions to X at \bar{x} , $I(X; \bar{x})$ ([6, Theorem 3.1]), which in our example is empty; the open cone of quasi-interior directions to X at \bar{x} , $Q(X; \bar{x})$ ([24]), which is also empty here. We point out that $G_0(\bar{x}) = \emptyset$ for any choice of a cone B as above.

Finally, by using Proposition 2.4 in [14], we get

$$N_M(X; \bar{x}) = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 < 0, x_2 < 0\} \cup \{(x_1, x_2) \in \mathbb{R}^2 : x_1 x_2 = 0\}.$$

Hence, the optimality condition ([8])

$$\lambda_0 \nabla f(\bar{x}) + \lambda_1 \nabla g(\bar{x}) \in -N_M(X; \bar{x}),$$

still provides less information than our result (3.7).

Stronger necessary optimality conditions than the Fritz John, are the so-called the KKT conditions. They require that λ_0 appearing in (3.6) be strictly positive. Thus, an easy way to derive KKT conditions from Corollary 3.2, is to impose assumptions yielding a contradiction with (3.6) if $\lambda_0 = 0$ is assumed. For instance the assumption

$$\text{co}(\{\nabla g_i(\bar{x}) : i \in I\}) \cap B^* = \emptyset, \quad (3.8)$$

or, equivalently (by Proposition 2.3) that

$$\text{there exists } v_0 \in \overline{\text{co}} B \text{ such that } \nabla g_i(\bar{x})^\top v_0 < 0, \quad i \in I \quad (3.9)$$

(i. e., $G_0(\bar{x}) \neq \emptyset$) or, equivalently (by Corollary 6.3.2 in [34]) that

$$\text{there exists } v_1 \in \text{ri } \overline{\text{co}} B \text{ such that } \nabla g_i(\bar{x})^\top v_1 < 0, \quad i \in I, \quad (3.10)$$

allow us to conclude that $\lambda_0 > 0$ in (3.6). This is summarized in the next corollary, which is well known in the literature, and the previous assumption is termed as a Mangasarian-Fromovitz type constraints qualification.

Corollary 3.4. *Let us consider Problem (1.1) and $\bar{x} \in K$. Let $X \subseteq \mathbb{R}^n$ and B be a nonempty set in \mathbb{R}^n satisfying $\overline{\text{co}} B \subseteq T(X; \bar{x})$. Let f, g_i , as in the previous corollary and assume that (3.8) (or (3.9) or (3.10)), i.e., $G_0(\bar{x}) \neq \emptyset$, holds. If \bar{x} is a local solution to (1.1), then there exist $\lambda_i \geq 0$, $i \in I$, satisfying*

$$\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in B^*.$$

Observe that the conclusion in the previous corollary remains valid for any function f having \bar{x} as a local minimum on K , which is the nature of the constraints qualification.

In the next section we establish a characterization of the validity of the KKT conditions in a direct way, which will be deduced as a consequence of the strong duality result as presented in Theorem 2.5.

4 Karush-Kuhn-Tucker conditions under inequality constraints

We now derive the KKT conditions without using the Fritz John conditions; instead they will be a consequence of Theorem 2.5.

Consider problem (1.1) with the same notations for K and I as introduced in Section 3.

Let $B \subseteq \mathbb{R}^n$ be a cone. The following approximation of the set K at $\bar{x} \in X$,

$$G'(\bar{x}) \doteq \left\{ v \in \overline{\text{co}} B : \nabla g_i(\bar{x})^\top v \leq 0, i \in I \right\},$$

along with the linear problem:

$$\nu_c \doteq \inf_{v \in G'(\bar{x})} \nabla f(\bar{x})^\top v, \quad (4.1)$$

will play an important role in our characterization of the validity of KKT conditions.

It is not difficult to check that $\nu_c \in \{-\infty, 0\}$. Indeed, obviously $\nu_c \leq 0$; if $\nu_c < 0$ then there exists $v_0 \in G'(\bar{x})$ such that $\nabla f(\bar{x})^\top v_0 < 0$ yielding $\nu_c = -\infty$ since $G'(\bar{x})$ is a cone. We will see below how the value of ν_c is linked to the fulfillment of the KKT optimality conditions under no constraints qualification. Before going on some notations are in order. In what follows $|A|$ denotes the cardinality of the set A .

Let $F : \mathbb{R}^n \rightarrow \mathbb{R} \times \mathbb{R}^{|I|}$ be defined by

$$F(v) \doteq \begin{pmatrix} \nabla f(\bar{x})^\top \\ \nabla g_I(\bar{x})^\top \end{pmatrix} v,$$

where $\nabla g_I(\bar{x})^\top$ denotes the matrix of order $|I| \times n$ whose rows are $\nabla g_i(\bar{x})^\top$ for $i \in I$. Obviously if $I = \emptyset$ then we delete the corresponding submatrix $\nabla g_I(\bar{x})^\top$ from the definition of F .

The next theorem establishes a characterization of the validity of the KKT conditions at points not necessarily local minima. As already observed, if the system $\nabla f(\bar{x})^\top v < 0, v \in G'(\bar{x})$ has a solution, then $\nu_c = -\infty$. In this case, for all $\lambda_i \geq 0, i \in I$, one has

$$\inf_{v \in \overline{\text{co}} B} \langle \nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}), v \rangle \leq \inf_{v \in G'(\bar{x})} \langle \nabla f(\bar{x}), v \rangle = \nu_c = -\infty.$$

Theorem 4.1. *Consider problem (1.1), let $\bar{x} \in X$, and B be a cone in \mathbb{R}^n . Then, the following assertions are equivalent:*

- (a) $\nu_c = 0$ and strong duality for Problem (4.1) holds.
- (b) There exist $\lambda_i \geq 0, i \in I$, satisfying

$$\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in B^*. \quad (4.2)$$

- (c)
$$\overline{[F(\overline{\text{co}} B) + (\mathbb{R}_+ \times \mathbb{R}_+^{|I|})]} \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset. \quad (4.3)$$

- (d)
$$\overline{\text{co}}[F(B) + (\mathbb{R}_+ \times \mathbb{R}_+^{|I|})] \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset. \quad (4.4)$$

(e) $\nu_c = 0$ and for every $i_0 \in I$:

$$\left. \begin{array}{l} v_k \in \overline{\text{co}} B, \|v_k\| \rightarrow +\infty, \nabla g_i(\bar{x})^\top v_k + q_k^i \rightarrow 0, q_k^i \geq 0, \\ i \in I, i \neq i_0, \nabla g_{i_0}(\bar{x})^\top v_k \rightarrow 0, \nabla f(\bar{x})^\top v_k < 0 \end{array} \right\} \implies \limsup_k \nabla f(\bar{x})^\top v_k = 0. \quad (4.5)$$

Proof. Assume that $\nu_c = 0$ and consider Problem (4.1). We observe that strong duality for (4.1) is equivalent to say that there exist $\lambda_i \geq 0$, $i \in I$, satisfying

$$\inf_{v \in G'(\bar{x})} \nabla f(\bar{x})^\top v = \inf_{v \in \overline{\text{co}} B} \langle \nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}), v \rangle. \quad (4.6)$$

Since the left-hand side of (4.6) is $\nu_c = 0$, then (4.6) is equivalent to (4.2). On the other hand, by Theorem 2.5, strong duality for (4.1) holds iff

$$\overline{\text{co}} \text{ cone}[F(\overline{\text{co}} B) - \nu_c(1, 0) + (\mathbb{R}_+ \times \mathbb{R}_+^{|I|})] \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset. \quad (4.7)$$

By noticing that $F(\overline{\text{co}} B) + (\mathbb{R}_+ \times \mathbb{R}_+^{|I|})$ is a convex cone due to linearity of F , we conclude that (4.7) is equivalent to (4.3).

The equivalence between (c) and (d) follows from the equality $\overline{F(\text{co } B)} = \overline{F(\overline{\text{co}} B)}$.

(e) \implies (c): Take $(a, 0) \in F(\overline{\text{co}} B) + (\mathbb{R}_+ \times \mathbb{R}_+^{|I|})$, and suppose that $a < 0$. Then, there exist $v_k \in \overline{\text{co}} B$, $r_k \geq 0$, $q_k^i \geq 0$, $i \in I$, such that $\nabla f(\bar{x})^\top v_k + r_k \rightarrow a$ and $\nabla g_i(\bar{x})^\top v_k + q_k^i \rightarrow 0$ for all $i \in I$. In case $\sup_k \|v_k\| < +\infty$, we obtain, up to a subsequence, $v_k \rightarrow v$. Thus $r_k \rightarrow a - \nabla f(\bar{x})^\top v \geq 0$, which implies $0 > a \geq \nabla f(\bar{x})^\top v$. Then there exists $i_0 \in I$ satisfying $\nabla g_{i_0}(\bar{x})^\top v > 0$ since $\nu_c = 0$. But from the inequality $0 < \nabla g_{i_0}(\bar{x})^\top v_k \leq \nabla g_{i_0}(\bar{x})^\top v_k + q_k^{i_0}$ for all k sufficiently large, we infer that $\nabla g_{i_0}(\bar{x})^\top v = 0$, reaching a contradiction.

We assume now that $\|v_k\| \rightarrow +\infty$. Clearly $\nabla f(\bar{x})^\top v_k < 0$ for all k sufficiently large. Therefore, for such k , there exists $i_k \in I$ such that $\nabla g_{i_k}(\bar{x})^\top v_k > 0$ because of $\nu_c = 0$. Since I is finite, up to a subsequence, we may assume that $i_k = i_0$ for all $k \geq k_0$, and therefore $\nabla g_{i_0}(\bar{x})^\top v_k \geq 0$ for all $k \geq k_0$. It turns out that, $\nabla g_{i_0}(\bar{x})^\top v_k \rightarrow 0$. We are ready to apply (4.5), to conclude (up to a subsequence) that $\nabla f(\bar{x})^\top v_k \rightarrow 0$, implying that $r_k \rightarrow a$, providing a contradiction. This proves (4.3).

(c) \implies (e): By (a) $\nu_c = 0$. Observe that (4.5) does not hold if, and only if there exist $i_0 \in I$ and a sequence v_k such that

$$\left. \begin{array}{l} v_k \in \overline{\text{co}} B, \|v_k\| \rightarrow +\infty, \nabla g_i(\bar{x})^\top v_k + q_k^i \rightarrow 0, q_k^i \geq 0, \\ i \in I, i \neq i_0, \nabla g_{i_0}(\bar{x})^\top v_k \rightarrow 0, \text{ for } i_0 \in I, \nabla f(\bar{x})^\top v_k < 0 \end{array} \right\} \implies \limsup_k \nabla f(\bar{x})^\top v_k < 0.$$

If $\limsup_k \nabla f(\bar{x})^\top v_k$ is finite, then, up to a subsequence, $\lim_k \nabla f(\bar{x})^\top v_k = a < 0$. Set

$$\bar{p}_k \doteq -\frac{\nabla f(\bar{x})^\top v_k}{2} > 0, \quad k \geq \bar{k}, \quad \bar{q}_k^i = q_k^i, \quad i \neq i_0, \quad \bar{q}_k^{i_0} = 0, \quad \forall k.$$

Then,

$$\nabla f(\bar{x})^\top v_k + \bar{p}_k \rightarrow \frac{a}{2} < 0, \quad \nabla g_i(\bar{x})^\top v_k + \bar{q}_k^i \rightarrow 0, \quad i \in I$$

which contradicts (4.3).

If $\limsup_k \nabla f(\bar{x})^\top v_k = -\infty$, set

$$\bar{p}_k \doteq -\nabla f(\bar{x})^\top v_k - 1 > 0, \quad k \geq \bar{k}, \quad \bar{q}_k^i = q_k^i, \quad i \neq i_0, \quad \bar{q}_k^{i_0} = 0, \quad \forall k.$$

Then,

$$\nabla f(\bar{x})^\top v_k + \bar{p}_k \rightarrow -1, \quad \nabla g_i(\bar{x})^\top v_k + \bar{q}_k^i \rightarrow 0, \quad i \in I$$

contradicting (4.3) again. This completes the proof. \square

We now establish another characterization of the fulfillment of KKT conditions on the same line considered in [18]. To that purpose we introduce some notations.

Let $\Delta \doteq \{\alpha \in \mathbb{R}^{|I|} : \sum_{i \in I} \alpha_i = 1, \alpha_i \geq 0, i \in I\}$. Whenever $S_f^-(0) \doteq \{v \in \overline{\text{co}} B : \nabla f(\bar{x})^\top v < 0\} \neq \emptyset$, set

$$s \doteq \min_{\alpha \in \Delta} \sup_{v \in S_f^-(0)} \frac{\sum_{i \in I} \alpha_i \nabla g_i(\bar{x})^\top v}{\nabla f(\bar{x})^\top v} = \sup_{v \in S_f^-(0)} \min_{\alpha \in \Delta} \frac{\sum_{i \in I} \alpha_i \nabla g_i(\bar{x})^\top v}{\nabla f(\bar{x})^\top v},$$

where the second equality follows from the Sion minimax theorem [36].

In case of one single inequality constraint, we will see in the next section that the convex hull in the definition of $S_f^-(0)$ can be avoided, since convexity arises naturally.

Theorem 4.2. *Let B be a cone in \mathbb{R}^n , and assume that $\bar{x} \in X$. The following assertions are equivalent:*

(a) *there exist $\lambda_i \geq 0$, $i \in I$, satisfying*

$$\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in B^*. \quad (4.8)$$

(b) *either $S_f^-(0) = \emptyset$ or $[S_f^-(0)] \neq \emptyset$ with $s < 0$.*

Proof. (a) \implies (b): If $\lambda_i = 0$, for all $i \in I$ then (4.8) collapses to $\nabla f(\bar{x}) \in B^*$, i.e. $S_f^-(0) = \emptyset$. If $\sum_{i \in I} \lambda_i > 0$ and $S_f^-(0) \neq \emptyset$, then from the inequality

$$\frac{1}{\sum_{i \in I} \lambda_i} [\nabla f(\bar{x})^\top v + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x})^\top v] \geq 0, \quad \forall v \in \overline{\text{co}} B,$$

we have

$$\frac{\sum_{i \in I} \lambda_i \nabla g_i(\bar{x})^\top v}{(\sum_{i \in I} \lambda_i) \nabla f(\bar{x})^\top v} \leq -\frac{1}{\sum_{i \in I} \lambda_i} < 0, \quad \forall v \in S_f^-(0),$$

which implies $s < 0$.

(b) \implies (a): If $S_f^-(0) = \emptyset$, as already observed, then (4.8) holds with $\lambda_i = 0$, for all $i \in I$. If $S_f^-(0) \neq \emptyset$ and $s < 0$, then there exists $\bar{\alpha} \in \Delta$ such that

$$\frac{\sum_{i \in I} \bar{\alpha}_i \nabla g_i(\bar{x})^\top v}{\nabla f(\bar{x})^\top v} \leq s, \quad \forall v \in S_f^-(0),$$

and so

$$\nabla f(\bar{x})^\top v - \frac{1}{s} \sum_{i \in I} \bar{\alpha}_i \nabla g_i(\bar{x})^\top v \geq 0, \quad \forall v \in S_f^-(0),$$

i.e.,

$$\nabla f(\bar{x}) - \frac{1}{s} \sum_{i \in I} \bar{\alpha}_i \nabla g_i(\bar{x}) \in [S_f^-(0)]^* = \{-\gamma \nabla f(\bar{x}) + B^* : \gamma \geq 0\},$$

where the last equality is due to Corollaries 6.3.2 and 16.4.2 (taking into account Theorem 6.5) in [34]. From the last relation it is easily seen that (a) is fulfilled with $\lambda_i = -\frac{\bar{\alpha}_i}{(\gamma + 1)s} \geq 0$, $i \in I$, for a suitable $\gamma \geq 0$. Notice that $\lambda_{i_0} > 0$ for some i_0 . \square

We must point out that there are instances satisfying the KKT conditions at points being not necessarily local solutions.

By virtue of Theorem 2.5, the equivalence between (4.3) and (4.4) allows us to infer that we can choose equivalently $\overline{\text{co}} B$ or B in the definition of $G'(\bar{x})$. In other words,

$$\nu_c = \inf_{v \in G'(\bar{x})} \nabla f(\bar{x})^\top v,$$

with

$$G'(\bar{x}) \doteq \left\{ v \in B : \nabla g_i(\bar{x})^\top v \leq 0, i \in I \right\}, \quad (4.9)$$

provided either (4.3) or (4.4) is satisfied.

Before going on, let us put in evidence the main difference of our approach with those existing in the literature. As outlined above most results deriving the KKT conditions require some constraints qualification. A rather general one is that due to Guignard [28], which means, for fixed cone B (the original Guignard CQ was given with $B = T(X; \bar{x})$),

$$[\overline{\text{co}} T(K; \bar{x})]^* = [T(K; \bar{x})]^* = [G'(\bar{x})]^*. \quad (4.10)$$

It is proved in [27, Theorem 1] that Guignard constraints qualification, turns out to be a necessary and sufficient condition for the validity of (4.2) with $B = \mathbb{R}^n$, for every objective function f having a local minimum at \bar{x} . When $B = T(X; \bar{x})$ a similar result is given in [27, Theorem 2].

Set

$$\tilde{G}(\bar{x}) \doteq \left\{ v \in \mathbb{R}^n : \nabla g_i(\bar{x})^\top v \leq 0, i \in I \right\}, \quad (4.11)$$

and

$$\Gamma(\bar{x}) \doteq -[\tilde{G}(\bar{x})]^* = \left\{ y = \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) : \lambda_i \geq 0, i \in I \right\}.$$

In [27, page 13] it is proved that the condition

$$[T(K; \bar{x})]^* = -\Gamma(\bar{x}) + B^* \quad (4.12)$$

is necessary and sufficient for (4.2) to be fulfilled for any objective function f having \bar{x} as a local minimum on K .

We observe that (4.12) implies (4.10). Indeed, because of Corollary 16.4.2 in [34], from (4.12) we have:

$$[\overline{\text{co}} T(K; \bar{x})]^* = [T(K; \bar{x})]^* = -\Gamma(\bar{x}) + B^* = \overline{-\Gamma(\bar{x}) + (\overline{\text{co}} B)^*} = [\tilde{G}(\bar{x}) \cap \overline{\text{co}} B]^* = (G'(\bar{x}))^*.$$

Unlike the above results and many others, which involve conditions on g_i and X that guarantee (4.8) for every function f having \bar{x} as a local minimum on K , our approach allows us to derive conditions on f , g_i and X , jointly, that ensure (4.8). To that end Theorems 4.1 and 4.2 will play an important role. Thus, we expect to cover situations where the constraints qualification (4.10) fails (and so also (4.12) fails), which is illustrated by Example 4.5.

In the remaining part of this paper, unless otherwise specified, we will assume that $B = T(X; \bar{x})$ and $G'(\bar{x})$ is defined as in (4.9), in such a case ν_c will be simply denoted by ν .

Thus, if optimality is imposed at \bar{x} , we obtain the following new theorem which covers situations where $T(X; \bar{x})$ may be non convex, by relaxing condition (4.5) but requiring the

convexity of $F(T(X; \bar{x})) + (\mathbb{R} \times \mathbb{R}_+^{|I|})$. When there is a single inequality constraint such a convexity requirement appears also as a necessary condition as it will be shown in Theorem 5.2.

Theorem 4.3. *Assume in addition that g_i is continuous for $i \notin I$ and that \bar{x} is a local solution to (1.1) and $F(T(X; \bar{x})) + (\mathbb{R} \times \mathbb{R}_+^{|I|})$ is convex. The following assertions are equivalent:*

(a) *for every $i_0 \in I$:*

$$\left. \begin{array}{l} v_k \in T(X; \bar{x}), \nabla g_i(\bar{x})^\top v_k + q_k^i \rightarrow 0, q_k^i \geq 0, \\ i \in I, i \neq i_0, \nabla g_{i_0}(\bar{x})^\top v_k \rightarrow 0, \nabla f(\bar{x})^\top v_k < 0 \end{array} \right\} \implies \limsup_k \nabla f(\bar{x})^\top v_k = 0, \quad (4.13)$$

(b) *there exist $\lambda_i \geq 0, i \in I$, satisfying*

$$\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in [T(X; \bar{x})]^*. \quad (4.14)$$

Proof. (a) \implies (b): By virtue of (d) in Theorem 4.1 (with $B = T(X; \bar{x})$), we need to check that

$$\overline{[F(T(X; \bar{x})) + (\mathbb{R}_+ \times \mathbb{R}_+^{|I|})]} \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset. \quad (4.15)$$

The proof is similar as that in Theorem 4.1. Take $(a, 0) \in \overline{[F(T(X; \bar{x})) + \mathbb{R}_+^{1+|I|}]}$, and suppose that $a < 0$. Then, there exist $v_k \in T(X; \bar{x}), r_k \geq 0, q_k^i \geq 0, i \in I$, such that $\nabla f(\bar{x})^\top v_k + r_k \rightarrow a$ and $\nabla g_i(\bar{x})^\top v_k + q_k^i \rightarrow 0$ for all $i \in I$. Clearly $\nabla f(\bar{x})^\top v_k < 0$ for all k sufficiently large. Therefore, for such $k \in \mathbb{N}$, there exists $i_k \in I$ such that $\nabla g_{i_k}(\bar{x})^\top v_k \geq 0$ since \bar{x} is optimal, see (3.5). Since I is finite, up to a subsequence, we may assume that $i_k = i_0$ for all $k \geq k_0$, and therefore $\nabla g_{i_0}(\bar{x})^\top v_k \geq 0$ for all $k \geq k_0$. It turns out that, $\nabla g_{i_0}(\bar{x})^\top v_k \rightarrow 0$. We are ready to apply (4.13), to conclude (up to a subsequence) that $\nabla f(\bar{x})^\top v_k \rightarrow 0$, implying that $p_k \rightarrow a$, providing a contradiction. This proves (4.15) and the conclusion follows from Theorem 4.1.

(b) \implies (a): It follows from a similar reasoning to that used to prove (c) \implies (e) in Theorem 4.1 with $T(X; \bar{x})$ instead of $\overline{\text{co}} B$. \square

Observe that from the proof of the preceding theorem, it immediately follows that, if \bar{x} is a local solution and (4.13) holds, then $\nu = 0$. Indeed, the latter is equivalent to the condition

$$[F(T(X; \bar{x})) + (\mathbb{R}_+ \times \mathbb{R}_+^{|I|})] \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset,$$

which is implied by (4.15).

The next example shows that the convexity assumption in Theorem 4.3 is not necessary for the validity of the KKT conditions; whereas Example 4.5 shows that it may happen that (4.14) is sharper than (4.8) under (4.10).

Example 4.4. Consider Problem (1.1) with the following data: $f(x_1, x_2) = -x_1 + x_2^2$, $g_1(x_1, x_2) = -x_2$, $g_2(x_1, x_2) = x_1 + x_2$, and

$$X \doteq \left\{ (x_1, x_2) \in \mathbb{R}_+^2 : x_1 x_2 = 0 \right\}.$$

Obviously $\bar{x} \doteq (0, 0)$ is the only feasible point and so the optimal solution. Clearly X is a closed cone, so $T(X; \bar{x}) = X$. We obtain

$$F(v) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -v_1 \\ -v_2 \\ v_1 + v_2 \end{pmatrix}$$

$$F(T(X; \bar{x})) = \left\{ u \in \mathbb{R}^3 : u_1 = -v_1, u_2 = -v_2, u_3 = v_1 + v_2, v_1 v_2 = 0, v_1 \geq 0, v_2 \geq 0 \right\},$$

$$F(T(X; \bar{x})) + \mathbb{R}_+^3 = \left\{ u \in \mathbb{R}^3 : u_1 \geq -v_1, u_2 \geq -v_2, u_3 \geq v_1 + v_2, v_1 v_2 = 0, v_1 \geq 0, v_2 \geq 0 \right\}$$

is a closed non convex set: in fact, $u^1 \doteq (0, -1, 1)$, $u^2 \doteq (-1, 0, 1) \in F(T(X; \bar{x})) + \mathbb{R}_+^3$, but $\frac{1}{2}(u^1 + u^2) = (-\frac{1}{2}, -\frac{1}{2}, 1) \notin F(T(X; \bar{x})) + \mathbb{R}_+^3$ since the set

$$\left\{ (u_1, u_2, u_3) \in \mathbb{R}^3 : -\frac{1}{2} \geq -v_1, -\frac{1}{2} \geq -v_2, 1 \geq v_1 + v_2, v_1 v_2 = 0, v_1 \geq 0, v_2 \geq 0 \right\}$$

is empty. Moreover, note that KKT conditions hold with $\lambda_1^* = \lambda_2^* = 1$.

We further observe that the assumptions of Theorem 4.9 are fulfilled.

Example 4.5. Consider the functions $f(x_1, x_2) \doteq x_2$, $g_1(x_1, x_2) \doteq (x_1 - 1)^2 + (x_2 - 1)^2 - 1$, $g_2(x_1, x_2) \doteq (x_1 - 1)^2 + (x_2 + 1)^2 - 1$, and $X \doteq \mathbb{R}^2$.

Obviously the feasible set is $K = \{\bar{x}\} = \{(1, 0)\}$, and therefore \bar{x} is a minimum point. Then $T(K; \bar{x}) = \{(0, 0)\}$ and $G'(\bar{x}) = \{(x_1, 0) : x_1 \in \mathbb{R}\}$, which implies that $\nu = 0$. On the other hand, since $T(X; \bar{x}) = \mathbb{R}^2$, we obtain $F(T(X; \bar{x})) = F(\mathbb{R}^2) = \{t(1, -2, 2) \in \mathbb{R}^3 : t \in \mathbb{R}\}$. Thus,

$$F(\mathbb{R}^2) + (\mathbb{R}_+ \times \mathbb{R}_+^2) = \{(u, v_1, v_2) \in \mathbb{R}^3 : u \geq t, v_1 \geq -2t, v_2 \geq 2t, t \in \mathbb{R}\},$$

which is convex. Moreover, it easy to check that (4.13) is satisfied. Hence by applying Theorem 4.3, (4.14) holds. The same conclusion is obtained if we apply Theorem 4.2 with $B = \mathbb{R}^n$: here, we get $S_f^-(0) = \{(v_1, v_2) : v_2 < 0\}$ and $s = -2$.

On the other hand, condition (4.10) is satisfied for $B = \{0\}$ and $B = [\tilde{G}(\bar{x})]^*$, and the corresponding KKT conditions (4.8) provide less information than (4.14).

In addition, we point out that $G_0(\bar{x}) = \emptyset$ for any cone B satisfying $\overline{\text{co}} B \subseteq T(X; \bar{x})$, so that Corollary 3.4 is not applicable.

By considering the objective $f(x_1, x_2) = x_1$ instead of the previous one, we get easily that $\nu = -\infty$, that is, the system $\nabla f(\bar{x})^\top v < 0$, $v \in G'(\bar{x})$ has a solution. Consequently, for all $\lambda_i \geq 0$, $i = 1, 2$, one has

$$\inf_{v \in \mathbb{R}^2} \langle \nabla f(\bar{x}) + \sum_{i=1}^2 \lambda_i \nabla g_i(\bar{x}), v \rangle = -\infty.$$

Let us consider a couple of conditions which seem to be more verifiable than (4.13):

- $[v_k \in T(X; \bar{x}), v_k \rightarrow v \neq 0, \nabla f(\bar{x})^\top v_k < 0] \implies \nabla g_j(\bar{x})^\top v > 0, \forall j \in I;$ (4.16)
- for every $i \in I$,

$$[v_k \in T(X; \bar{x}), \nabla g_i(\bar{x})^\top v_k \rightarrow 0, \nabla f(\bar{x})^\top v_k < 0] \implies \limsup_k \nabla f(\bar{x})^\top v_k = 0. \quad (4.17)$$

The following proposition clarifies the relationships between the above two conditions and (4.13).

Proposition 4.6. *Assume that $\bar{x} \in K$. Then, we have*

$$(4.16) \implies (4.17) \implies (4.13).$$

In addition, (4.16) also implies $\nu = 0$.

Proof. Firstly, we easily obtain that (4.16) implies $\nu = 0$ (simply take the constant sequence v to get $\nu = 0$).

(4.16) \implies (4.17): Let $i \in I$, $v_k \in T(X; \bar{x})$ such that $\nabla g_i(\bar{x})^\top v_k \rightarrow 0$, $\nabla f(\bar{x})^\top v_k < 0$. If $v_k \rightarrow 0$, then (4.17) obviously holds. Two possibilities now arise: $\sup_{k \in \mathbb{N}} \|v_k\| < +\infty$ with $v_k \not\rightarrow 0$ and $\sup_{k \in \mathbb{N}} \|v_k\| = +\infty$. In the first case, up to a subsequence, we get $v_k \rightarrow v \neq 0$. In such a case, since $\nabla f(\bar{x})^\top v_k < 0$, by (4.16) we get $\nabla g_j(\bar{x})^\top v > 0$ for all $j \in I$, yielding a contradiction.

In the second case, we may assume, up to a subsequence, that $\|v_k\| \rightarrow +\infty$, $\frac{v_k}{\|v_k\|} \rightarrow v_0 \neq 0$. Since $\nabla f(\bar{x})^\top \frac{v_k}{\|v_k\|} < 0$ by (4.16) we get $\nabla g_j(\bar{x})^\top v_0 > 0$ for all $j \in I$. Moreover, since $\nabla g_i(\bar{x})^\top v_k \rightarrow 0$ then $\nabla g_i(\bar{x})^\top \frac{v_k}{\|v_k\|} \rightarrow 0$, which contradicts that $\nabla g_i(\bar{x})^\top v_0 > 0$. This proves that under (4.16), the conditions in the left-hand side of (4.17) are fulfilled only if $v_k \rightarrow 0$, so that (4.17) holds.

(4.17) \implies (4.13): It is straightforward. \square

The following instance, where $T(X; \bar{x})$ is convex, asserts that condition (4.13) does not imply (4.17).

Example 4.7. Consider the problem

$$\min\{x_2 : x_1 - x_2 \leq 0, -x_1 - x_2 \leq 0, x \in X \doteq \mathbb{R}^2\}.$$

Obviously $\bar{x} \doteq (0, 0)$ is an optimal solution, $T(X; \bar{x}) = X$ and

$$F(v) = \begin{pmatrix} 0 & 1 \\ 1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_2 \\ v_1 - v_2 \\ -v_1 - v_2 \end{pmatrix}.$$

It is a nice exercise to check that condition (4.13) holds. In order to show that (4.17) is not satisfied we reason as follows. Let $i = 1$, $g_1(x) = x_1 - x_2$ and choose $v_k = (v_{1k}, v_{2k})$ with $v_{1k} = v_{2k}$ and $v_{2k} \rightarrow -\infty$ ($v_{2k} < 0$). Then, $\nabla g_1(\bar{x})^\top v_k = 0$ and $\nabla f(\bar{x})^\top v_k = v_{2k} < 0$ for all k , but $v_{2k} \not\rightarrow 0$.

Next example shows that (4.16) (which implies $\nu = 0$) does not guarantee that Fritz John conditions hold.

Example 4.8. Consider the problem

$$\min\{x_1 : -\frac{3}{2}x_1 - x_2 \leq 0, -2x_1 - x_2 \leq 0, x \in X\},$$

where

$$X \doteq \left\{ (x_1, x_2) \in \mathbb{R}^2 : x_2 = 0, x_1 \geq 0 \right\} \cup \left\{ (x_1, x_2) \in \mathbb{R}^2 : x_1 = -x_2, x_1 \leq 0 \right\}.$$

Take $\bar{x} = (0, 0)$. Obviously X is a closed cone so that $T(X; \bar{x}) = X$ and \bar{x} is the optimal solution. Moreover, we obtain

$$F(v) = \begin{pmatrix} 1 & 0 \\ -\frac{3}{2} & -1 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ -\frac{3}{2}v_1 - v_2 \\ -2v_1 - v_2 \end{pmatrix}$$

and so (4.16) is fulfilled. In fact

$$[v^k \in T(X; \bar{x}), v^k \rightarrow v \neq 0, \nabla f(\bar{x})^\top v^k < 0] \implies \nabla g_j(\bar{x})^\top v > 0, \quad \forall j \in I,$$

is equivalent to

$$[(v_1^k, -v_1^k) \rightarrow v \neq 0, v_1^k < 0] \implies [-\frac{3}{2}v_1 - v_2 > 0 \quad \text{and} \quad -2v_1 - v_2 > 0],$$

which is true since $v \doteq (v_1, v_2) = (v_1, -v_1)$ with $v_1 < 0$.

On the other hand, observe that the problem

$$\min\{v_1 : -\frac{3}{2}v_1 - v_2 \leq 0, -2v_1 - v_2 \leq 0, v \in T(X, \bar{x})\},$$

has optimal value $\nu = 0$ (the linearization coincides with the given problem), while $\nu_0 = -\infty$.

Actually,

$$\nu_0 = \min\{v_1 : -\frac{3}{2}v_1 - v_2 < 0, -2v_1 - v_2 < 0, v \in \overline{\text{co}} T(X, \bar{x})\},$$

with $\overline{\text{co}} T(X, \bar{x}) = \{(v_1, v_2) : -v_1 - v_2 \leq 0, v_2 \geq 0\}$. Thus, by Theorem 3.1 the Fritz John conditions are not fulfilled as we can also check directly: we have that $[T(X; \bar{x})]^* = \{(x_1, x_2) : x_1 - x_2 \leq 0, x_1 \geq 0\}$ and FJ conditions are given by the system

$$\lambda_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \lambda_1 \begin{pmatrix} \frac{3}{2} \\ 1 \end{pmatrix} - \lambda_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \lambda \geq 0, \lambda \neq 0, y_1 \geq 0, y_1 - y_2 \leq 0,$$

which is impossible.

Next result establishes a sufficient condition for a KKT point to be a strict local minimum.

Theorem 4.9. *Assume that $\bar{x} \in K$ satisfies (4.14) with $\lambda_i \geq 0$ for $i \in I$. If there exists $i_0 \in I$ such that $\lambda_{i_0} > 0$ and*

$$[v \in T(X; \bar{x}), \nabla f(\bar{x})^\top v = 0 = \nabla g_{i_0}(\bar{x})^\top v] \implies v = 0. \quad (4.18)$$

Then \bar{x} is a strict local solution to Problem (1.1).

Proof. By assumption, we have

$$\nabla f(\bar{x})^\top v \geq -\sum_{i \in I} \lambda_i \nabla g_i(\bar{x})^\top v, \quad \forall v \in T(X; \bar{x}).$$

In particular, since

$$T(K; \bar{x}) \subseteq \{v \in T(X; \bar{x}) : \nabla g_i(\bar{x})^\top v \leq 0, i \in I\} \subseteq T(X; \bar{x}),$$

we get

$$\nabla f(\bar{x})^\top v \geq - \sum_{i \in I} \lambda_i \nabla g_i(\bar{x})^\top v \geq 0, \quad \forall v \in T(K; \bar{x}).$$

Thus, if $\nabla f(\bar{x})^\top v = 0$ then $\lambda_i \nabla g_i(\bar{x})^\top v = 0$ for all $i \in I$, which implies that $\nabla g_{i_0}(\bar{x})^\top v = 0$ since $\lambda_{i_0} > 0$. By (4.18), $v = 0$. Hence

$$\nabla f(\bar{x})^\top v > 0, \quad \forall v \in T(K; \bar{x}) \setminus \{0\},$$

which implies that \bar{x} is a strict local optimal solution to (1.1). \square

For completeness we establish a corollary under the Guignard constraints qualification as expressed in (4.10).

If \bar{x} is a local solution to (1.1) then it is well known that [37, 28, 26]:

$$\nabla f(\bar{x})^\top v \geq 0, \quad \forall v \in T(K; \bar{x}),$$

that is, $\nabla f(\bar{x}) \in [T(K; \bar{x})]^*$. Thus, by imposing the constraints qualification (4.10), we obtain $\nabla f(\bar{x})^\top v \geq 0$, for all $v \in G'(\bar{x})$, or equivalently, the system: $v \in \overline{\text{co}} B$

$$\begin{cases} \langle \nabla f(\bar{x}), v \rangle < 0, \\ \langle \nabla g_i(\bar{x}), v \rangle \leq 0, \quad i \in I, \end{cases} \quad (4.19)$$

has no solution, which is equivalent to have $\nu_c = 0$. Thus the next result is obtained.

Corollary 4.10. *Let $\bar{x} \in K$ be a local solution to (1.1). If (4.10) holds and $F(\overline{\text{co}} B) + \mathbb{R}_+ \times \mathbb{R}_+^{|I|}$ is closed (in particular if $\overline{\text{co}} B$ is polyhedral) with $B \subseteq \mathbb{R}^n$ being a cone, then, there exist $\lambda_i \geq 0$, $i \in I$, satisfying*

$$\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in B^*.$$

Proof. We notice that the impossibility of system (4.19) amounts to writing that

$$F(\overline{\text{co}} B) \cap -(\mathbb{R}_{++} \times \mathbb{R}_+^{|I|}) = \emptyset,$$

or, equivalently

$$(F(\overline{\text{co}} B) + \mathbb{R}_+ \times \mathbb{R}_+^{|I|}) \cap -(\mathbb{R}_{++} \times \mathbb{R}_+^{|I|}) = \emptyset,$$

which, in turn, is equivalent to

$$(F(\overline{\text{co}} B) + \mathbb{R}_+ \times \mathbb{R}_+^{|I|}) \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset. \quad (4.20)$$

Thus, by assumption, (4.3) holds, and therefore, by Theorem 4.1, the conclusion follows. \square

Remark 4.11. The proof of Corollary 4.10 clarifies the connections between Guignard constraints qualification (4.10) and (4.3) of Theorem 4.1. Actually (4.10) allows one to establish directly that, under suitable conditions (as $\overline{\text{co}} B$ polyhedral), (4.20) is equivalent to (4.3). For example, by setting $B \doteq T(X; \bar{x})$, the previous corollary includes the case where X is open under the Guignard constraints qualification.

Let us recall the following condition already mentioned at the end of Section 3, which is related to the Mangasarian-Fromovitz type constraints qualification:

$$\exists \bar{v} \in \text{ri } \overline{\text{co}} B : \langle \nabla g_i(\bar{x}), \bar{v} \rangle < 0, \quad i \in I. \quad (4.21)$$

This assumption implies the validity of the KKT conditions (Corollary 3.4) and is stronger than (4.10) provided $\text{co } B = T(X; \bar{x})$, as the following proposition shows.

Proposition 4.12. *Let $X \neq \emptyset$ and $B \subseteq \mathbb{R}^n$ is a cone such that $\text{co } B \subseteq T(X; \bar{x})$. Assume, in addition that g_i is continuous for $i \notin I$, and that there exists $\bar{v} \in \text{ri } \overline{\text{co}} B$ such that $\langle \nabla g_i(\bar{x}), \bar{v} \rangle < 0$, $i \in I$. Then, $G'(\bar{x}) \subseteq T(K; \bar{x})$, where $G'(\bar{x}) \doteq \{v \in \overline{\text{co}} B : \nabla g_i(\bar{x})^\top v \leq 0, \quad i \in I\}$. If, in addition $\text{co } B = T(X; \bar{x})$, then $G'(\bar{x}) = T(K; \bar{x})$ and, as a consequence, $T(K; \bar{x})$ is convex.*

Proof. Define the set

$$G_0(\bar{x}) \doteq \{v \in \text{ri } \overline{\text{co}} B : \nabla g_i(\bar{x})^\top v < 0, \quad i \in I\}.$$

We first prove that $G_0(\bar{x}) \subseteq T(K; \bar{x})$. Indeed, let $v \in G_0(\bar{x}) \neq \emptyset$. Since $v \in \text{ri } \overline{\text{co}} B \subseteq T(X; \bar{x})$ then there exists $\lambda_k > 0$, $x_k \in X$, $x_k \rightarrow \bar{x}$, $\lambda_k(x_k - \bar{x}) \rightarrow v$. Thus, by the differentiability assumptions

$$g_i(x_k) = g_i(\bar{x}) + \langle \nabla g_i(\bar{x}), x_k - \bar{x} \rangle + \|x_k - \bar{x}\|o(\|x_k - \bar{x}\|), \quad i \in I,$$

and multiplying the previous equalities by λ_k , letting $k \rightarrow +\infty$ and exploiting the choice of v , we have that there exists k_0 such that $g_i(x_k) < 0$ for all $i \in I$ and all $k \geq k_0$. If $i \notin I$, then $g_i(\bar{x}) < 0$ and continuity of g_i implies $g_i(x_k) < 0$, for all $k \geq k_1$, so that there exists k_2 such that the sequence x_k is feasible for $k \geq k_2$: it follows that $v \in T(K; \bar{x})$. This proves $G_0(\bar{x}) \subseteq T(K; \bar{x})$, and therefore

$$G'(\bar{x}) = \overline{G_0(\bar{x})} \subseteq T(K; \bar{x}),$$

where the equality is due to Theorem 6.5 in [34], which proves the first part of the statement. If, furthermore, $\text{co } B = T(X; \bar{x})$, then we have $T(K; \bar{x}) \subseteq G'(\bar{x})$, which along with the previous inclusion complete the proof. \square

Remark 4.13. Proposition 4.12 proves that (4.21) implies

$$[T(K; \bar{x})]^* \subseteq [G'(\bar{x})]^*, \quad (4.22)$$

provided that $\text{co } B \subseteq T(X; \bar{x})$. We note that (4.22) can be considered as a generalized Guignard constraints qualification: actually, Corollary 4.10 is still valid replacing (4.10) with (4.22). Finally, observe that whenever $T(K; \bar{x}) \subseteq \overline{\text{co}} B$, (4.10) and (4.22) are equivalent; in fact, in general, $T(K; \bar{x}) \subseteq \tilde{G}(\bar{x})$ (see (4.11)) so that

$$T(K; \bar{x}) \subseteq [\tilde{G}(\bar{x}) \cap \overline{\text{co}} B] = G'(\bar{x}).$$

Therefore, the opposite inclusion holds in (4.22) and (4.10) follows.

5 The case of a single inequality constraint

We now discuss the Fritz John and the KKT optimality conditions for the minimization problem (1.1) under a single inequality constraint, i.e., with $m = 1$.

Here $g \doteq g_1$, and $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ are differentiable functions at a given $\bar{x} \in X$. This special situation allows us to characterize completely the fulfillment of the Fritz John and the KKT optimality conditions in a more precise way. Both results exploit the special structure of the two-dimensionality of the image space.

We point out that a device on how to reduce a quasiconvex minimization problem with many inequality constraints into one with a single constraint is presented in [18, Theorem 3.6].

We start by establishing several characterization to the validity of the Fritz John optimality conditions. They will be consequences of Proposition 2.4.

In order to apply Proposition 2.4, let us denote as before

$$F \doteq \begin{pmatrix} \nabla f(\bar{x})^\top \\ \nabla g(\bar{x})^\top \end{pmatrix}, \quad \text{i.e., } F(v) \doteq (\nabla f(\bar{x})^\top v, \nabla g(\bar{x})^\top v),$$

and consider

$$C = T(X; \bar{x}), \quad F(C) = \{F(v) \in \mathbb{R}^2 : v \in T(X; \bar{x})\}.$$

We know that (see the proof of Corollary 3.2) if \bar{x} is a local solution to (1.1) with $g(\bar{x}) = 0$, then (3.5) has no solution, i.e., $F(T(X; \bar{x})) \cap (-\mathbb{R}_{++}^2) = \emptyset$. Hence we obtain the following theorem, which is new, and characterizes completely the fulfillment of the Fritz John optimality conditions.

Theorem 5.1. *Let $\bar{x} \in X$ be a local solution to (1.1) with $g(\bar{x}) = 0$. The following assertions are equivalent:*

- (a) $(F(T(X; \bar{x})) + \mathbb{R}_{++}^2) \cup \{(0, 0)\}$ is pointed;
- (b) $F(T(X; \bar{x})) + \mathbb{R}_+^2$ is convex;
- (c) $\overline{F(T(X; \bar{x})) + \mathbb{R}_+^2}$ is convex;
- (d) $F(T(X; \bar{x})) + \mathbb{R}_{++}^2$ is convex;
- (e) there exist $(\lambda_0, \lambda_1) \in \mathbb{R}_+^2 \setminus \{(0, 0)\}$ such that

$$\lambda_0 \nabla f(\bar{x}) + \lambda_1 \nabla g(\bar{x}) \in [T(X; \bar{x})]^*.$$

Proof. It is a direct consequence of Proposition 2.4 with the notations as above. □

Now, let us turn our attention to the KKT optimality conditions. We first notice that for all sets $A \subseteq \mathbb{R}^2$, one has

$$(A + \mathbb{R}_+^2) \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset \implies (A + \mathbb{R}_+^2) \cap (-\mathbb{R}_{++}^2) = \emptyset. \quad (5.1)$$

We associate with problem (1.1) its linearization for a given $\bar{x} \in X$:

$$\nu \doteq \inf_{v \in G'(\bar{x})} \nabla f(\bar{x})^\top v,$$

where

$$G'(\bar{x}) \doteq \left\{ v \in T(X; \bar{x}) : \nabla g(\bar{x})^\top v \leq 0 \right\}, \quad g(\bar{x}) = 0.$$

We already know that $\nu \in \{-\infty, 0\}$, and immediately get

$$\begin{aligned} \nu = 0 &\iff [v \in T(X; \bar{x}), \nabla f(\bar{x})^\top v < 0 \implies \nabla g(\bar{x})^\top v > 0] \\ &\iff F(T(X; \bar{x})) \cap -(\mathbb{R}_{++} \times \mathbb{R}_+) = \emptyset; \\ &\iff [F(T(X; \bar{x})) + \mathbb{R}_+^2] \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset. \end{aligned}$$

The following characterization of the validity of the KKT conditions is obtained for a point not necessarily a local solution and it is the analogue to Theorem 4.1 with $B = T(X; \bar{x})$. Contrary to the case discussed in the preceding section where convexity arises as a sufficient condition, it appears here as a necessary condition as well.

Theorem 5.2. *Assume that $\bar{x} \in X$, $g(\bar{x}) = 0$. The following assertions are equivalent:*

(a) $\exists \lambda^* \geq 0$ such that

$$\nabla f(\bar{x}) + \lambda^* \nabla g(\bar{x}) \in [T(X; \bar{x})]^*. \quad (5.2)$$

(b) $F(T(X; \bar{x})) + \mathbb{R}_+^2$ is convex and

$$\overline{[F(T(X; \bar{x})) + \mathbb{R}_+^2]} \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset. \quad (5.3)$$

(c) $F(T(X; \bar{x})) + \mathbb{R}_+^2$ is convex, $\nu = 0$ and

$$\left. \begin{array}{l} v_k \in T(X; \bar{x}), \|v_k\| \rightarrow +\infty, \\ \nabla g(\bar{x})^\top v_k \rightarrow 0, \nabla f(\bar{x})^\top v_k < 0 \end{array} \right\} \implies \limsup_k \nabla f(\bar{x})^\top v_k = 0. \quad (5.4)$$

Proof. (a) \iff (b): From Theorem 4.1, (a) is equivalent to $\overline{\text{co}[F(T(X; \bar{x})) + \mathbb{R}_+^2]} \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset$. Thus, from (5.1) it follows that $\overline{[F(T(X; \bar{x})) + \mathbb{R}_+^2]} \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset$. On the other hand, by (d) in Proposition 2.4, we infer the convexity of $F(T(X; \bar{x})) + \mathbb{R}_+^2$. This proves one implication; the other is straightforward, again by Theorem 4.1.

(b) \iff (c): We now prove that (5.3) is equivalent to $\nu = 0$ and (5.4). Assume that (5.3) holds. Then, as observed above, $\nu = 0$. Take any $v_k \in T(X; \bar{x})$, $\lim_k \nabla g(\bar{x})^\top v_k = 0$, $\nabla f(\bar{x})^\top v_k < 0$. Thus $\nabla g(\bar{x})^\top v_k > 0$. Suppose, on the contrary, that $\limsup_k \nabla f(\bar{x})^\top v_k = \xi < 0$. First consider $\xi \in \mathbb{R}$. Then, up to a subsequence, $\nabla f(\bar{x})^\top v_k \rightarrow \xi$. Since

$$\begin{pmatrix} \nabla f(\bar{x})^\top \\ \nabla g(\bar{x})^\top \end{pmatrix} v_k \in F(T(X; \bar{x})) + \mathbb{R}_+^2,$$

we get $(\xi, 0) \in \overline{[F(T(X; \bar{x})) + \mathbb{R}_+^2]} \cap (-\mathbb{R}_{++} \times \{0\})$, yielding a contradiction.

Consider now that $\xi = -\infty$. Up to a subsequence, $\nabla f(\bar{x})^\top v_k \rightarrow -\infty$. Due to linearity

$$\begin{pmatrix} \nabla f(\bar{x})^\top \\ \nabla g(\bar{x})^\top \end{pmatrix} v'_k \in F(T(X; \bar{x})) + \mathbb{R}_+^2,$$

where $v'_k = -\frac{v_k}{\nabla f(\bar{x})^\top v_k}$. Thus

$$(-1, 0) \in \overline{[F(T(X; \bar{x})) + \mathbb{R}_+^2]} \cap (-\mathbb{R}_{++} \times \{0\}),$$

yielding a contradiction.

Assume that (5.4) and $\nu = 0$ hold. Take $(a, 0) \in \overline{F(T(X; \bar{x})) + \mathbf{R}_+^2}$, and suppose that $a < 0$. Then, there exist $v_k \in T(X; \bar{x})$, $r_k \geq 0$, $q_k \geq 0$ such that $\nabla f(\bar{x})^\top v_k + r_k \rightarrow a$ and $\nabla g(\bar{x})^\top v_k + q_k \rightarrow 0$. Assume first that $\sup_k \|v_k\| < +\infty$. Up to a subsequence, we get $v_k \rightarrow v \in T(X; \bar{x})$. Thus $r_k \rightarrow a - \nabla f(\bar{x})^\top v \geq 0$, which implies that $0 > a \geq \nabla f(\bar{x})^\top v$. Then $\nabla g(\bar{x})^\top v > 0$ (since $\nu = 0$), and so from the inequality $0 < \nabla g(\bar{x})^\top v_k \leq \nabla g(\bar{x})^\top v_k + q_k$, we obtain $\nabla g(\bar{x})^\top v = 0$, giving a contradiction.

Assume now that $\sup_k \|v_k\| = +\infty$. Up to a subsequence, we have $\|v_k\| \rightarrow +\infty$. Clearly $\nabla f(\bar{x})^\top v_k < 0$ and therefore $\nabla g(\bar{x})^\top v_k > 0$ since $\nu = 0$. This together with the non-negativity of q_k imply $\nabla g(\bar{x})^\top v_k \rightarrow 0$ and $q_k \rightarrow 0$. By (5.4), we obtain, up to a subsequence, that $\nabla f(\bar{x})^\top v_k \rightarrow 0$. This gives $r_k \rightarrow a$, a contradiction, since $a < 0$. \square

In the same spirit of Theorem 4.2, whenever

$$S_f^-(0) \doteq \{v \in T(X; \bar{x}) : \nabla f(\bar{x})^\top v < 0\} \neq \emptyset,$$

set

$$s \doteq \sup_{v \in S_f^-(0)} \frac{\nabla g(\bar{x})^\top v}{\nabla f(\bar{x})^\top v}.$$

Notice that we do not require the convex hull of $T(X; \bar{x})$ in the definition of $S_f^-(0)$ as occurs in Theorem 4.2.

Theorem 5.3. *Assume that $\bar{x} \in X$ is such that $g(\bar{x}) = 0$. The following assertions are equivalent:*

(a) $\exists \lambda^* \geq 0$ such that

$$\nabla f(\bar{x}) + \lambda^* \nabla g(\bar{x}) \in [T(X; \bar{x})]^*;$$

(b) $F(T(X; \bar{x})) + \mathbf{R}_+^2$ is convex and either $S_f^-(0) = \emptyset$ or $[S_f^-(0) \neq \emptyset$ with $s < 0]$.

Proof. (a) \implies (b): The convexity property follows from Theorem 5.2. If $\lambda^* = 0$ then (5.2) collapses to $\nabla f(\bar{x}) \in [T(X; \bar{x})]^*$, i.e., $S_f^-(0) = \emptyset$. If $\lambda^* > 0$ and $S_f^-(0) \neq \emptyset$, then from the inequality

$$\nabla f(\bar{x})^\top v + \lambda^* \nabla g(\bar{x})^\top v \geq 0, \quad \forall v \in T(X; \bar{x}),$$

we have

$$\frac{\nabla g(\bar{x})^\top v}{\nabla f(\bar{x})^\top v} \leq -\frac{1}{\lambda^*} < 0, \quad \forall v \in S_f^-(0),$$

which implies $s < 0$.

(b) \implies (a): First notice that if either $S_f^-(0) = \emptyset$ or $S_f^-(0) \neq \emptyset$ with $s < 0$, we obtain $\nu = 0$. Obviously if $S_f^-(0) = \emptyset$ then (a) holds with $\lambda^* = 0$. Otherwise we easily obtain (5.3) from the fact $s < 0$, and so (a) holds as well. \square

Remark 5.4. A simple application of Theorem 5.1 allows us to assert that:

- $S_f^-(0) = \emptyset$ implies the convexity of $F(B) + \mathbf{R}_+^2$ for all cones $B \subseteq T(X; \bar{x})$;
- $S_f^-(0) \neq \emptyset$ with $s < 0$ implies the convexity of $F(S_f^-(0)) + \mathbf{R}_+^2$. Indeed, If $S_f^-(0) \neq \emptyset$ and $s < 0$, then

$$\frac{\nabla g(\bar{x})^\top v}{\nabla f(\bar{x})^\top v} \leq s, \forall v \in S_f^-(0),$$

and so

$$\nabla f(\bar{x})^\top v - \frac{1}{s} \nabla g(\bar{x})^\top v \geq 0, \quad \forall v \in S_f^-(0).$$

Then, the conclusion follows from Proposition 2.4.

Certainly our previous result applies to situations where \bar{x} needs not to be a local solution.

In case \bar{x} is optimal, we obtain the following new result which includes situations where the Guignard constraints qualification fails.

Corollary 5.5. *Assume that \bar{x} is a local solution to (1.1) such that $g(\bar{x}) = 0$. Then, $F(T(X; \bar{x})) + \mathbb{R}_+^2$ is convex and (5.5) is satisfied if, and only if the KKT conditions (5.2) hold, where*

$$[v_k \in T(X; \bar{x}), \nabla g(\bar{x})^\top v_k \rightarrow 0, \nabla f(\bar{x})^\top v_k < 0] \implies \limsup_k \nabla f(\bar{x})^\top v_k = 0. \quad (5.5)$$

Proof. One implication is as follows. Assuming the KKT conditions, by Theorem 5.2 we obtain $\nu = 0$, the convexity of $F(T(X; \bar{x})) + \mathbb{R}_+^2$, and that (5.5) holds with the additional requirement $\|v_k\| \rightarrow +\infty$. It remains to consider the case $\sup_k \|v_k\| < +\infty$ with $v_k \in T(X; \bar{x})$ satisfying the left hand-side of (5.5). Up to a subsequence, $v_k \rightarrow v \in T(X; \bar{x})$, and so $\nabla g(\bar{x})^\top v = 0$ and $\nabla f(\bar{x})^\top v \leq 0$. Hence $\nabla f(\bar{x})^\top v = 0$ ($\nu = 0$), which is the desired result, completing the proof of (5.5).

The other implication is obtained from Theorem 4.3. \square

By virtue of the preceding corollary, one can ask whether $\nu = 0$ may be substituted by the assumption that \bar{x} be a local solution to (1.1) in Theorem 5.2. The instance

$$\min\{-x_1 + \cos(x_2) : x_1(x_2 + 1) \leq 0, x \in X \doteq \mathbb{R}^2\}, \quad \bar{x} \doteq (0, 0), \quad (5.6)$$

answers this question in the negative form.

Remark 5.6. Corollary 5.5 covers situations where results derived via exactness of penalty functions are not applicable. Indeed, our previous result applies to Example 2.2 in [38]: $\min\{x : x^2 - x \leq 0\} = 0 = \bar{x}$, but [38, Theorem 2.1] is not applicable.

Remark 5.7. The proof of the equivalence between (b) and (c) actually shows that

$$\overline{[F(T(X; \bar{x})) + \mathbb{R}_+^2]} \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset \iff \overline{F(T(X; \bar{x}))} \cap (-\mathbb{R}_{++} \times \mathbb{R}_+) = \emptyset.$$

An instance showing the closure in (5.3) cannot be deleted is exhibited in Example 5.17.

In what follows we exhibit two instances showing the wide applicability of our Theorem 5.2: Example 5.8 allows us to infer the existence of situations where Corollary 3.4 is not applicable (since condition (3.8) is not satisfied for any cone B such that $\overline{\text{co}} B \subseteq T(X; \bar{x})$) but our main theorem applies; whereas Example 5.9 does not satisfy the Guignard constraints qualification (4.10).

Example 5.8. (Corollary 3.4 is not applicable) Consider the problem

$$\min\{-x_1 + x_2 : x_1 - x_2 \leq 0, x \in X\},$$

where $X \doteq \{(x_1, x_2) \in \mathbb{R}_+^2 : x_2 \leq x_1\}$, which has $\bar{x} \doteq (0, 0)$ as an optimal solution. Since X is a closed cone, $T(X; \bar{x}) = X$. Setting

$$F(v) \doteq \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -v_1 + v_2 \\ v_1 - v_2 \end{pmatrix},$$

we get

$$\begin{aligned} F(T(X; \bar{x})) &= \{(u_1, u_2) \in \mathbb{R}^2 : u_1 = -v_1 + v_2, u_2 = v_1 - v_2, v_1 \geq 0, v_2 \geq 0, v_2 \leq v_1\} \\ &= \{(u_1, u_2) \in \mathbb{R}^2 : u_2 = -u_1, u_1 \leq 0\}. \end{aligned}$$

Thus

$$F(T(X; \bar{x})) + \mathbb{R}_+^2 = \{(u_1, u_2) \in \mathbb{R}^2 : u_2 \geq -u_1, u_1 \geq 0\}$$

is a closed convex set. We observe that (5.3) is satisfied, so we can apply Theorem 5.2. It is easily seen that KKT conditions hold with $\lambda^* = 1$.

On the other hand, $G_0(\bar{x}) \doteq \{v \in \overline{\text{co}} B : \nabla g(\bar{x})^\top v < 0\} = \emptyset$ for any cone B satisfying $\overline{\text{co}} B \subseteq T(X; \bar{x})$, so that Corollary 3.4 is not applicable.

Example 5.9. (The Guignard constraints qualification fails) Consider Example 4.5 with the same data but with a slight modification:

$$\min\{x_2 : (x_1 - 1)^2 + (x_2 - 1)^2 - 1 \leq 0, (x_1, x_2) \in X\},$$

where $X \doteq \{(x_1, x_2) \in \mathbb{R}^2 : (x_1 - 1)^2 + (x_2 + 1)^2 - 1 \leq 0\}$. Obviously $K \doteq \{(x_1, x_2) \in X : g_1(x_1, x_2) \leq 0\} = \{(1, 0)\}$. Take $\bar{x} = (1, 0)$, which is a minimum point. Then $T(K; \bar{x}) = \{(0, 0)\}$ and $G'(\bar{x}) = \{(x_1, 0) : x_1 \in \mathbb{R}\}$. This implies that $\nu = 0$. On the other hand $T(X; \bar{x}) = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq 0\}$, and so the Guignard condition (4.10) is not satisfied for $B = T(X; \bar{x})$. Let us check that (b) of the previous theorem is fulfilled. We obtain $F(T(X; \bar{x})) = \{t(1, -2) \in \mathbb{R}^2 : t \leq 0\}$. Thus

$$F(T(X; \bar{x})) + \mathbb{R}_+^2 = \{(u, v) \in \mathbb{R}^2 : u \geq t, v_1 \geq -2t, t \leq 0\}.$$

Due to polyhedrality of $T(X; \bar{x})$ and linearity of F , we get $\overline{F(T(X; \bar{x})) + \mathbb{R}_+^2} = F(T(X; \bar{x})) + \mathbb{R}_+^2$, and therefore (5.3) is fulfilled, proving that (b) of the preceding theorem holds, and so (5.2) is satisfied.

A more verifiable condition implying (5.3), or equivalently $\nu = 0$ and (5.4), is obtained in the next proposition, which is a particular case of Proposition 4.6.

Proposition 5.10. *Assume that $\bar{x} \in X$, $g(\bar{x}) = 0$. If*

$$[v_k \in T(X; \bar{x}), v_k \rightarrow v \neq 0, \nabla f(\bar{x})^\top v_k < 0] \implies \nabla g(\bar{x})^\top v > 0, \quad (5.7)$$

then, $\nu = 0$ and (5.4) holds.

Remark 5.11. The example in (5.6) shows that the reverse implication in the preceding proposition fails. In fact (5.7) becomes

$$[v^k \in T(X; \bar{x}), v^k \rightarrow v \neq 0, -v_1^k < 0] \implies v_1 > 0,$$

which does not hold as we can see by choosing $v^k = (1/k, 1) \in T(X; \bar{x})$, $k \in N$. Hence, (5.7) is not necessary for the validity of the KKT conditions.

The same result is obtained by Example 5.8, that is, (5.7) is not fulfilled. Indeed, let $v_k \doteq (1, m_k)$, with $0 \leq m_k < 1$, $m_k \rightarrow 1$. Then $v_k \in T(X; \bar{x}) = X$ and

$$\nabla f(\bar{x})^\top v_k = -1 + m_k < 0,$$

but since $v_k \rightarrow v = (1, 1)$, we have $\nabla g(\bar{x})^\top v = 0$.

Remark 5.12. By noticing that $\overline{\text{co}} A$ is always polyhedral whenever A is a cone in \mathbb{R}^2 , we conclude that $\overline{[F(T(X; \bar{x})) + \mathbb{R}_+^2]}$ is polyhedral since $F(T(X; \bar{x})) + \mathbb{R}_+^2$ is convex.

In case $F(T(X; \bar{x})) + \mathbb{R}_+^2$ is closed (this holds for instance if $T(X; \bar{x})$ is polyhedral, as in the preceding examples), (b) of Theorem 5.2 reduces to

$$[F(T(X; \bar{x})) + \mathbb{R}_+^2] \cap (-\mathbb{R}_{++} \times \{0\}) = \emptyset,$$

which is equivalent to

$$[v \in T(X; \bar{x}), \nabla f(\bar{x})^\top v < 0] \implies \nabla g(\bar{x})^\top v > 0,$$

or, equivalently $\nu = 0$. In case $T(X; \bar{x})$ is polyhedral, that is, if $T(X; \bar{x}) = \{v \in \mathbb{R}^n : Av \leq 0\}$ for some real $m \times n$ matrix, then the previous implication holds if, and only if (by Farkas lemma) the system

$$\nabla f(\bar{x}) + \lambda^* \nabla g(\bar{x}) + A^\top u = 0, \quad u \in \mathbb{R}_+^m, \lambda^* \geq 0,$$

has solution. This equality expresses actually

$$\nabla f(\bar{x}) + \lambda^* \nabla g(\bar{x}) \in [T(X; \bar{x})]^*,$$

since $[T(X; \bar{x})]^* = \{A^\top v : v \leq 0\}$.

A simpler sufficient condition is expressed in the following result.

Proposition 5.13. *Assume that $\bar{x} \in X$, $g(\bar{x}) = 0$. If either*

(a) $\nu = 0$ and

$$[v \in T(X; \bar{x}), \nabla f(\bar{x})^\top v = 0 = \nabla g(\bar{x})^\top v] \implies v = 0, \quad \text{or} \quad (5.8)$$

(b) \bar{x} is a local solution to (1.1) and

$$[v \in T(X; \bar{x}), \nabla f(\bar{x})^\top v \leq 0 = \nabla g(\bar{x})^\top v] \implies v = 0,$$

then, either $S_f^-(0) = \emptyset$ or $[S_f^-(0) \neq \emptyset \text{ with } s < 0]$ holds.

Proof. In case $S_f^-(0) = \emptyset$ we are done. By assuming that $S_f^-(0) \neq \emptyset$ we shall prove that $s < 0$. If on the contrary, $s = 0$, due to homogeneity, we can find a sequence $v_k \in S_f^-(0)$, $\|v_k\| = 1$, $v_k \rightarrow v$ such that $\frac{\nabla g(\bar{x})^\top v_k}{\nabla f(\bar{x})^\top v_k} \rightarrow 0$. In both cases (a) or (b), we get $\nabla f(\bar{x})^\top v \leq 0$ and $\nabla g(\bar{x})^\top v \geq 0$. The fact that $\nu = 0$ and $s = 0$ allow us to infer that $\nabla f(\bar{x})^\top v = 0 = \nabla g(\bar{x})^\top v$, yielding a contradiction. In case \bar{x} is a local solution to (1.1), $\nabla f(\bar{x})^\top v \leq 0$ and $\nabla g(\bar{x})^\top v > 0$ is impossible since $s = 0$, and so a contradiction is achieved. \square

We observe that Example 3.3 satisfies condition (5.8).

Remark 5.14. The instance in (5.6) and Example 5.8 show that reverse implication in the previous proposition fails in general, since in both cases (5.8) does not hold.

Next remark states that the assumptions of Proposition 5.13 are stronger than that of Proposition 5.10.

Remark 5.15. We claim that

$$[\nu = 0 \text{ and } (5.8)] \implies (5.7).$$

In fact, take $v_k \in T(X; \bar{x})$, $v_k \rightarrow v \neq 0$, such that $\nabla f(\bar{x})^\top v_k < 0$. Since $\nu = 0$, $\nabla g(\bar{x})^\top v_k > 0$ and so $\nabla g(\bar{x})^\top v \geq 0$. We also have $\nabla f(\bar{x})^\top v < 0$. Assume that $\nabla f(\bar{x})^\top v = 0$: if $\nabla g(\bar{x})^\top v = 0$ then $v = 0$ because of (5.8), yielding a contradiction. Hence $\nabla g(\bar{x})^\top v > 0$, the desired result. In case $\nabla f(\bar{x})^\top v < 0$, we get $\nabla g(\bar{x})^\top v > 0$, since $\nu = 0$. This completes the proof of the claim.

It is not hard to see that the reverse implication in the previous remark fails in general.

For completeness we provide a necessary condition involving s .

Corollary 5.16. *Assume that $\bar{x} \in X$ is such that $g(\bar{x}) = 0$. If $S_f^-(0) \neq \emptyset$ with $s < 0$, then*

$$[v \in T(X; \bar{x}), \nabla f(\bar{x})^\top v \leq 0] \implies \nabla g(\bar{x})^\top v \geq 0. \quad (5.9)$$

Proof. First of all observe that, by Theorem 5.3, $S_f^-(0) \neq \emptyset$ with $s < 0$ implies that (5.2) holds with $\lambda^* > 0$ (as it is easily seen by the proof). Then

$$\nabla f(\bar{x})^\top v \geq -\lambda^* \nabla g(\bar{x})^\top v, \quad \forall v \in T(X; \bar{x}),$$

so that (5.9) holds since $\lambda^* > 0$. □

The next example shows that the closure in (b) of Theorem 5.2 cannot be deleted.

Example 5.17. Consider the problem

$$\min\{x_1 + x_2^2 + x_3^2 : x_1^2 + x_2^2 + x_3 \leq 0, x \in X\},$$

where

$$X \doteq \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = \frac{x_1^2}{\sqrt{x_1^2 + x_2^2}}, (x_1, x_2) \neq (0, 0) \right\} \cup \{(0, 0, 0)\}.$$

It is clear that $\bar{x} \doteq (0, 0, 0)$ is the only feasible point and, therefore, the optimal solution. Since X is a closed cone, we get $T(X; \bar{x}) = X$. Moreover, we also obtain $F(v) = (v_1, v_3)$. Thus,

$$\begin{aligned} F(T(X; \bar{x})) = F(X) &= \left\{ (v_1, v_3) \in \mathbb{R}^2 : v_3 = \frac{v_1^2}{\sqrt{v_1^2 + v_2^2}}, (v_1, v_2) \neq (0, 0) \right\} \cup \{(0, 0)\} \\ &= \left\{ (v_1, v_3) \in \mathbb{R}^2 : v_3 > 0, v_1 \neq 0 \right\} \cup \{(0, 0)\}. \end{aligned}$$

It follows that

$$F(T(X; \bar{x})) + \mathbb{R}_+^2 = \left\{ (v_1, v_3) \in \mathbb{R}^2 : v_3 > 0, v_1 < 0 \right\} \cup \mathbb{R}_+^2$$

is a convex set. We observe that

$$[F(T(X; \bar{x})) + \mathbb{R}_+^2] \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset,$$

i.e., $\nu = 0$. But

$$\overline{F(T(X; \bar{x})) + \mathbb{R}_+^2} \cap -(\mathbb{R}_{++} \times \{0\}) \neq \emptyset. \quad (5.10)$$

Note that

$$[T(X; \bar{x})]^* = \{(v_1, v_2, v_3) \in \mathbb{R}^3 : v_1 = v_2 = 0, v_3 \geq 0\},$$

and, according to (5.10), the system

$$\nabla f(\bar{x}) + \lambda \nabla g(\bar{x}) \in [T(X; \bar{x})]^*, \quad \lambda \geq 0,$$

i.e.,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ y \end{pmatrix}, \quad \lambda \geq 0, y \geq 0$$

is clearly impossible.

We also get $s = 0$, and therefore (a) of Theorem 5.3 is not satisfied.

Observe that all the equivalent statements (a)-(e) of Theorem 5.1 are fulfilled.

Next result establishes a sufficient condition for a KKT point to be a strict local minimum, which is a specialization of Theorem 4.9.

Proposition 5.18. *Assume that $\bar{x} \in X$, $g(\bar{x}) = 0$, satisfies (5.2) with $\lambda^* > 0$ (this holds for instance if $S_f^-(0) \neq \emptyset$ with $s < 0$ by the proof of Theorem 5.3) and (5.8). Then \bar{x} is a strict local solution to problem (1.1).*

Notice that Example 3.3 exhibits a situation where $\lambda^* > 0$ is permitted, so that Proposition 5.18 is applicable even with $S_f^-(0) = \emptyset$.

6 A reformulation of a MPEC into one with a single inequality constraint

Given the variational inequality:

$$\text{find } x \in C \text{ such that } \langle A(x), y - x \rangle \geq 0, \quad \forall y \in C, \quad (6.1)$$

where $C \subseteq \mathbb{R}^n$ is a nonempty closed convex set and $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$, let S be the solution set to (6.1) and consider the following problem ($X \subseteq \mathbb{R}^n$)

$$\min\{f(x) : x \in S \cap X\}, \quad (6.2)$$

which is a particular instance of a MPEC. We will see that (6.2) can be reformulated as a standard minimization problem with a single inequality constraint via a differentiable gap function, i.e., a function $p : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ which is non-negative on C and that fulfills the

condition $x \in C$ and $p(x) = 0$ if and only if x is a solution to (6.1). Therefore solving (6.1) is equivalent to the minimization of p on the feasible set C to (6.1).

Unfortunately, the first gap function introduced by Auslender [5] is, in general, not differentiable. This drawback was overcome by Fukushima [21], whose approach considers a continuously differentiable gap function. It was, later on, generalized by Zhu and Marcotte [41]. We briefly describe the results obtained by Fukushima, Zhu and Marcotte. Assume that $G : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is a non-negative, continuously differentiable function, such that $G(x, \cdot) : C \rightarrow \mathbb{R}$ is strongly convex for all $x \in C$, satisfying

$$G(x, x) = 0, \quad \nabla_y G(x, x) = 0, \quad \forall x \in C.$$

Define

$$g(x) \doteq \max_{y \in C} [\langle A(x), x - y \rangle - G(x, y)],$$

then g is a gap function for (6.1); furthermore (see [41]), if $A \in C^1$, then $g \in C^1$ and

$$\nabla g(x) = A(x) - (\nabla A(x))^T (H(x) - x) - \nabla_x G(x, H(x))$$

where $H(x) \doteq \operatorname{argmax}_{y \in C} [\langle A(x), x - y \rangle - G(x, y)]$.

If we consider the particular case

$$G(x, y) \doteq \frac{1}{2} \langle x - y, M(x - y) \rangle,$$

where M is a symmetric and positive definite matrix of order n , we obtain the gap function introduced by Fukushima [21] and, in particular, for $x \in S$, $\nabla g(x) = A(x)$. Indeed, by [21, Proposition 3.1] x is a solution to (6.1) if and only if $H(x) = x$; it is easy to check that $\nabla_x G(x, H(x)) = 0$, which yields $\nabla g(x) = A(x)$.

Thus, problem (6.2) is equivalent to the minimization problem:

$$\min \{ f(x) : g(x) \leq 0, x \in C \cap X \}$$

as one can check it directly.

We now establish a characterization of the KKT optimality conditions as a consequence of our main results of Section 5 by using the Fukushima gap function.

For a given $\bar{x} \in S \cap X$, define

$$F \doteq \begin{pmatrix} \nabla f(\bar{x})^\top \\ A(\bar{x})^\top \end{pmatrix}, \quad \text{i.e., } F(v) \doteq (\nabla f(\bar{x})^\top v, A(\bar{x})^\top v), \quad \nu \doteq \inf_{v \in G'(\bar{x})} \nabla f(\bar{x})^\top v,$$

where

$$G'(\bar{x}) \doteq \left\{ v \in T(C \cap X; \bar{x}) : A(\bar{x})^\top v \leq 0 \right\}.$$

Our Theorem 5.2 yields the required characterizations with $\nabla g(\bar{x}) = A\bar{x}$, and $C \cap X$ instead of X , where the KKT conditions read as follows:

$$\exists \lambda^* \geq 0 \quad \text{such that } \nabla f(\bar{x}) + \lambda^* A(\bar{x}) \in [T(C \cap X; \bar{x})]^*.$$

Thus, we also obtain results derived from Corollary 5.5.

We now provide an example which does not satisfy the Guignard CQ, but our results apply.

Example 6.1. Consider problem (6.1) where $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ and C are defined by

$$A(x_1, x_2) = (x_2^2, x_1 + x_2)^\top, \quad C \doteq \mathbb{R}_+^2.$$

It is not difficult to show that the solution set to (6.1) is given by

$$S = \{x \in \mathbb{R}_+^2 : x_2 = 0\}.$$

Actually, since $C \doteq \mathbb{R}_+^2$, (6.1) is equivalent to the following complementarity problem

$$x_2^2 x_1 = 0, \quad (x_1 + x_2)x_2 = 0, \quad x_1 + x_2 \geq 0, \quad x_1 \geq 0, \quad x_2 \geq 0.$$

Let $g(x) \doteq \max_{y \in C} [A(x), x - y] - \frac{1}{2} \|x - y\|^2$. Then, for any $x \in C$,

$$\begin{aligned} g(x) &= \max_{y \in C} [x_2^2(x_1 - y_1) + (x_1 + x_2)(x_2 - y_2) - \frac{1}{2}(x_1 - y_1)^2 - \frac{1}{2}(x_2 - y_2)^2] \\ &= \begin{cases} x_1 x_2^2 + x_1 x_2 - \frac{1}{2}(x_1^2 - x_2^2) & \text{if } -x_1 + x_2^2 \geq 0; \\ x_1 x_2 + \frac{1}{2}(x_2^2 + x_2^4) & \text{if } -x_1 + x_2^2 \leq 0. \end{cases} \end{aligned}$$

Note that $g(x) = 0$ iff $x \in S$. Let $X \doteq \{x \in \mathbb{R}_+^2 : x_1 x_2 = 0\}$ and consider the problem

$$\min\{x_1^2 + x_2 : g(x) \leq 0, x \in X\},$$

whose feasible set K is given by the solution set to (6.1) (expressed by the constraints $g(x) \leq 0, x \in \mathbb{R}_+^2$) with the additional constraint $x_1 x_2 = 0$.

It is easy to see that $\bar{x} \doteq (0, 0)^\top$ is the optimal solution and X is a closed cone so $T(X; \bar{x}) = X$, moreover $\nabla g(\bar{x}) = A(\bar{x}) = (0, 0)^\top$.

Now, we shall see that our approach is applicable to this reformulation. In particular, Theorem 5.2 will be useful. Indeed, set $B = T(X; \bar{x})$, we observe that the (see (4.10)) Guignard CQ does not hold, as $T(K; \bar{x}) = S$ and $G'(\bar{x}) = \overline{\text{co}} T(X; \bar{x}) = \mathbb{R}_+^2$, while all the equivalent assertions (a), (b) and (c) of Theorem 5.2 are fulfilled. In particular,

$$F(v) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_2 \\ 0 \end{pmatrix}$$

and

$$F(T(X, \bar{x})) = F(X) = \{x \in \mathbb{R}_+^2 : x_2 = 0\},$$

so that (b) of Theorem 5.2 holds. Moreover, $\nabla f(\bar{x}) = (0, 1) \in [T(X; \bar{x})]^*$ and (a) holds with $\lambda^* = 0$.

Finally, note that (c) is fulfilled, since $F(T(X, \bar{x})) + \mathbb{R}_+^2$ is convex and $\nabla f(\bar{x})^\top v \geq 0$, for every $v \in T(X; \bar{x})$.

7 Further applications and concluding remarks

We now present two other simple instances showing the wide applicability of our approach. Let us define the following parametric optimization problem ($P(x)$)

$$\min_{y \in K(x)} \varphi(x, y),$$

where $K(x) \doteq \{y \in X : \psi_i(x, y) \leq 0, i = 1, \dots, m\}$, $\varphi : X \times X \rightarrow \mathbb{R}$, $\psi_i : X \times X \rightarrow \mathbb{R}$, $i = 1, \dots, m$ and X is a nonempty set in \mathbb{R}^n .

We consider the problem of finding $\bar{x} \in K(\bar{x})$ such that \bar{x} is a global (local) solution to $(P(\bar{x}))$. Such a problem arises in many fields as vector optimization, quasi-variational inequalities and vector equilibrium problems. Let us describe a couple of examples.

Consider the following vector equilibrium problem regarding efficient solutions: find

$$\bar{x} \in X : f(\bar{x}, y) \notin -\mathbb{R}_+^m \setminus \{0\}, \quad \forall y \in X, \quad (7.1)$$

where $f : X \times X \rightarrow \mathbb{R}^m$ and $f(x, x) = 0, \forall x \in X$. Set $I \doteq \{1, \dots, m\}$, we assume that $f_i(\bar{x}, \cdot)$ is differentiable at \bar{x} , for all $i \in I$. It is known (see for instance [15]) that \bar{x} is a solution to (7.1) if, and only if \bar{x} is a (global) solution to $(P(\bar{x}))$ with

$$\varphi(x, y) \doteq \sum_{i \in I} f_i(x, y) \quad \text{and} \quad \psi_i(x, y) \doteq f_i(x, y).$$

Thus, the KKT conditions for $(P(\bar{x}))$ become

$$\sum_{i \in I} (1 + \lambda_i) \nabla_y f_i(\bar{x}, \bar{x}) \in [T(X; \bar{x})]^*,$$

which is equivalent (by Theorem 4.3) to: for $i_0 \in I$

$$\left. \begin{aligned} v_k \in T(X; \bar{x}), \quad \nabla_y f_i(\bar{x}, \bar{x})^\top v_k + q_i^k \rightarrow 0, \quad q_i^k \geq 0, i \in I, \\ i \neq i_0, \quad \nabla_y f_{i_0}(\bar{x}, \bar{x})^\top v_k \rightarrow 0, \quad \sum_{i \in I} \nabla_y f_i(\bar{x}, \bar{x})^\top v_k < 0, \end{aligned} \right\}$$

$$\implies \limsup_k \sum_{i \in I} \nabla_y f_i(\bar{x}, \bar{x})^\top v_k = 0,$$

provided X is convex.

The second example is a quasi-variational inequality problem. Given a mapping $H : X \rightarrow \mathbb{R}^n$, consider the quasi-variational inequality problem which consists in finding

$$\bar{x} \in K(\bar{x}) : \langle H(\bar{x}), y - \bar{x} \rangle \geq 0, \quad \forall y \in K(\bar{x}). \quad (7.2)$$

It is easy to see that \bar{x} is a solution to (7.2) if, and only if \bar{x} is a global solution to $(P(\bar{x}))$, with $\varphi(x, y) \doteq \langle H(x), y \rangle$ and K is the set-valued map defined as in $(P(x))$ above, where we further assume that $\psi_i(\bar{x}, \cdot)$ is differentiable at \bar{x} , for $i = \{1, \dots, m\}$.

By Theorem 4.3, it follows that the KKT conditions for (7.2) become: there exist $\lambda_i \geq 0$, $i \in I$, satisfying

$$H(\bar{x}) + \sum_{i \in I} \lambda_i \nabla_y \psi_i(\bar{x}, \bar{x}) \in [T(X; \bar{x})]^*,$$

which is equivalent to: for every $i_0 \in I$,

$$\left. \begin{aligned} v_k \in T(X; \bar{x}), \quad \nabla_y \psi_i(\bar{x}, \bar{x})^\top v_k + q_i^k \rightarrow 0, \quad q_i^k \geq 0, i \in I, \\ i \neq i_0, \quad \nabla_y \psi_{i_0}(\bar{x}, \bar{x})^\top v_k \rightarrow 0, \quad H(\bar{x})^\top v_k < 0, \end{aligned} \right\} \implies \limsup_k H(\bar{x})^\top v_k = 0,$$

provided X is convex.

Conclusions

We have developed several characterizations for the validity of the FJ and KKT optimality conditions and we have obtained results that apply even to situations not satisfying the Guignard CQ. In particular, the case of a single inequality constraint has been analysed in details and then applied to a class of mathematical programs with equilibrium constraints. It should be further developed the question to find suitable subclasses of problems where some of the equivalent assertions to the KKT conditions is easier to verify than the KKT itself.

Acknowledgements

The first author wants to express his gratitude to the members of the Operations Research group of the Department of Computer Science at University of Pisa, where part of this work was carried out, for their very kind hospitality. Both authors want to thank both referees for their constructive criticisms and helpful remarks, which have improved the presentation of the paper.

References

- [1] W. ACHTZIGER, CH. KANZOW, *Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications*, Math. Program. A, 114 (2008), 69–99.
- [2] R. ANDREANI, G. HAESER, M. L. SCHUVERDT AND P. J. S. SILVA, *Two new constraint qualifications and applications*, SIAM J. Optim., 22 (2012), pp. 1109–1135.
- [3] R. ANDREANI, G. HAESER, M. L. SCHUVERDT AND P. J. S. SILVA, *A relaxed constant positive linear dependence constraint qualification and applications*, Math. Program. A, 135 (2012), 255–273.
- [4] S. ARROW, L. HURWICZ AND H. UZAWA, *Constraint qualifications in maximization problems*, Naval Res. Logist. Quart., 8 (1961), 175–191.
- [5] A. AUSLENDER, *Optimisation: méthodes numériques*, Masson, Paris, 1976.
- [6] M. S. BAZARAA AND J. J. GOODE, *Necessary optimality criteria in mathematical programming in the presence of differentiability*, J. Math. Anal. Appl., 40 (1972), pp. 609–621.
- [7] M. S. BAZARAA, H. D. SHERALI AND C. M. SHETTY, *Nonlinear programming: theory and algorithms*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.
- [8] D. P. BERTSEKAS AND A. E. OZDAGLAR, *Pseudonormality and Lagrange multiplier theory for constrained optimization*, J. Optim. Theory Appl., 114 (2002), pp. 287–343.
- [9] D. P. BERTSEKAS, A. E. OZDAGLAR AND P. TSENG, *Enhanced Fritz John conditions for convex programming*, SIAM J. Optim., 16 (2006), pp. 766–797.

- [10] J. V. BURKE, *Calmness and exact penalization*, SIAM J. Control Optim., 29 (1991), pp. 493–497.
- [11] F. H. CLARKE, *A New Approach to Lagrange Multipliers*, Math. Oper. Res., 1 (1976), pp. 165–174.
- [12] F. H. CLARKE, *Optimization and nonsmooth analysis*, John Wiley & Sons, New York, 1983.
- [13] F. FACCHINEI, C. KANZOW AND S. SAGRATELLA, *Solving quasi-variational inequalities via their KKT conditions*, Math. Program. A, 144 (2014), pp. 369–412.
- [14] M. L. FLEGEL AND C. KANZOW, *On M -stationary points for mathematical programs with equilibrium constraints*, J. Math. Anal. Appl., 310 (2005), pp. 286–302.
- [15] F. FLORES-BAZÁN, *Existence theory for finite dimensional pseudomonotone equilibrium problems*, Acta Applicandae Mathematicae, 77 No. 3 (2003), pp. 249–297.
- [16] F. FLORES-BAZÁN, *Fritz John necessary optimality condition of the alternative-type*, J. Optim. Theory Appl., 161 (2014), pp. 807–818.
- [17] F. FLORES-BAZÁN, F. FLORES-BAZÁN AND C. VERA, *Gordan-type alternative theorems and vector optimization revisited*, in Recent developments in vector optimization, Q. H. Ansari and J.-C. Yao, eds., Springer-Verlag, Berlin-Heidelberg, 2012, pp. 29–59.
- [18] F. FLORES-BAZÁN, F. FLORES-BAZÁN AND C. VERA, *A complete characterization of strong duality in nonconvex optimization with a single constraint*, J. Glob. Optim., 53 (2012), pp. 185–201.
- [19] F. FLORES-BAZÁN, N. HADJISAVVAS AND C. VERA, *An optimal alternative theorem and applications to mathematical programming*, J. Glob. Optim., 37 (2007), pp. 229–243.
- [20] F. FLORES-BAZÁN AND G. MASTROENI, *Strong duality in cone constrained nonconvex optimization*, SIAM J. Optim., 23 (2013), pp. 153–169.
- [21] M. FUKUSHIMA, *Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems*, Math. Program. A, 53 (1992), pp. 99–110.
- [22] F. GIANNESI, *Theorems of the alternative and optimality conditions*, J. Optim. Theory Appl., 42 (1984), pp. 331–365.
- [23] F. GIANNESI, *Semidifferentiable functions and necessary optimality conditions*, J. Optim. Theory Appl., 60 (1989), pp. 191–241.
- [24] G. GIORGI AND A. GUERRAGGIO, *First order generalized optimality conditions for programming problems with a set constraint*, in: Generalized Convexity, S. Komlosi, T. Rapcsák and S. Schaible, eds., Proceedings, Pécs, Hungary, 1992, Springer Verlag, Berlin, 1994, pp. 171–185.

- [25] G. GIORGI, A. GUERRAGGIO AND J. THIERFELDER, *Mathematics of optimization: smooth and nonsmooth case*, Elsevier, Amsterdam, 2004.
- [26] F. J. GOULD AND J. W. TOLLE, *A necessary and sufficient qualification for constrained optimization*, SIAM J. Appl. Math., 20 No 2 (1971), pp. 164–172.
- [27] F. J. GOULD AND J. W. TOLLE, *Geometry of optimality conditions and constraint qualifications*, Math. Program., 2 (1972), pp. 1–18.
- [28] M. GUIGNARD, *Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space*, SIAM J. Control, 7 No 2 (1969), pp. 232–241.
- [29] W. KARUSH, *Minima of functions of several variables with inequalities and side conditions*. Master thesis. Department of Mathematics, University of Chicago, 1939.
- [30] H. W. KUHN AND A. W. TUCKER, *Nonlinear programming*, in: J. Neyman, ed., Proceedings of the second Berkeley simposium on mathematical statistics and probability, University of California press, Berkeley, 1950, pp. 481–492.
- [31] O. L. MANGASARIAN AND S. FROMOVITZ, *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, J. Math. Anal. Appl., 17 (1967), pp. 37–47.
- [32] K. W. MENG AND X. Q. YANG, *Optimality conditions via exact penalty functions*, SIAM J. Optim., 20 (2010), pp. 3208–3231.
- [33] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation I. Basic theory*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 330. Springer-Verlag, Berlin, 2006.
- [34] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, 1970.
- [35] A. M. RUBINOV AND X. Q. YANG, *Lagrange-type functions in constrained non-convex optimization*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2003.
- [36] M. SION, *On general minimax theorems*, Pacific J. Math., 8 (1958), pp. 171–176.
- [37] P. P. VARAIYA, *Nonlinear programming in Banach space*, SIAM J. Appl. Math., 15 No 2 (1967), pp. 284–293.
- [38] X. Q. YANG AND Z. Q. MENG, *Lagrange multipliers and calmness conditions of order p* , Math. Oper. Res., 32 (2007), pp. 95–101.
- [39] J. J. YE AND J. ZHANG, *Enhanced Karush-Kuhn-Tucker condition and weaker constraint qualifications*, Math. Program. B, 139 (2013), pp. 353–381.
- [40] J. J. YE, D. L. ZHU AND Q. J. ZHU, *Exact penalization and necessary optimality conditions for generalized bilevel programming problems*, SIAM J. Optim., 7 (1997), pp. 481–507.
- [41] D. L. ZHU AND P. MARCOTTE, *An extended descent framework for variational inequalities*, J. Optim. Theory Appl., 80 (1994), pp. 349–366.