J Optim Theory Appl
DOI 10.1007/s10957-013-0456-8

Fritz John Necessary Optimality Conditions of the Alternative-Type

Fabián Flores-Bazán

Received: 22 February 2013 / Accepted: 18 October 2013

Abstract An alternative-type version of the Fritz John optimality conditions is established at points not necessarily optimal, which covers situations where no result appearing elsewhere is applicable. As a by-product, a versatile formulation of these necessary Fritz John optimality conditions along with a simple proof is provided. This encompasses several versions appearing in the literature. A variant of the KKT conditions is also presented.

Keywords Nonlinear programming · Fritz John conditions · KKT conditions

1 Introduction

The Fritz John necessary optimality conditions along with the Karush–Kuhn–Tucker have proved to be a masterpiece in the development of nonlinear optimization. There is no standard way to present these results in most textbooks devoted to mathematics and even engineering students, see Bazaraa et al. [1], Bector et al. [2]. The main goal of this note is to provide an alternative-type version of the Fritz John optimality condition at points not necessarily optimal. Its proof uses simple separation theorems on convex sets and properties of polar cones. Our formulation is versatile and requires the notion of the contingent cone. Its versatility is shown by recovering various versions appearing in the common literature, and it is suitable for expository purposes.

Section 2 starts with the basic definitions of a contingent cone, polar cone, and a new characterization of a disjunction in term of pointedness. In addition, a Gordan-type theorem of the alternative, suitable for our purpose, is established. The alternative-type version of the Fritz John optimality conditions is established in

F. Flores-Bazán (✉)
Departamento de Ingeniería Matemática, Universidad de Concepción, CI2MA, Casilla 160-C, Concepción, Chile
e-mail: fflores@ing-mat.udec.cl

Published online: 05 November 2013

Sect. 3 (see Theorem 3.1). Some particular situations appearing in the literature are discussed in Sect. 4. Finally, Sect. 5 presents a new variant of the KKT optimality conditions under a Mangasarian–Fromovitz type constraint qualification. Some conclusions are outlined at the end.

2 Basic Definitions and Preliminary Results

In what follows, given a set $A \subseteq \mathbb{R}^n$, its closure is denoted by $\mathrm{cl}\, A$; its convex hull by $\mathrm{conv}\, A$ which is the smallest convex set containing A; its topological interior by $\mathrm{int}\, A$. We set $\mathrm{cone}\, A := \bigcup_{t \geq 0} tA$.

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

$$T(K; \bar{x}) := \big\{ v \in \mathbb{R}^n : \exists t_k > 0, \exists x_k \in K, x_k \to \bar{x}, t_k(x_k - \bar{x}) \to v \big\}.$$

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

$$T(K; \bar{x}) = \mathrm{cl}\left(\bigcup_{t \geq 0} t(K - \bar{x}) \right).$$

By $\langle \cdot, \cdot \rangle$ we mean 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

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

It well known that, whenever P is a closed and convex cone, we have (the bipolar theorem; reflexivity) $P = P^{**} := (P^*)^*$, and in general we have $P^{**} = \mathrm{cl}(\mathrm{conv}(\mathrm{cone}\, P))$.

We say that a (not necessarily convex) cone, P, is pointed iff

$$\mathrm{conv}\, P \cap (-\mathrm{conv}\, P) = \{0\}.$$

Notice that

$$\mathrm{cone}\, A \text{ is pointed} \quad \Longleftrightarrow \quad \mathrm{cone}(\mathrm{conv}\, A) \text{ is pointed.} \tag{1}$$

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

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

Part of the next theorem appears, in a more general framework, in Theorem 3.2 of [3].

Theorem 2.1 *Let* $P \subseteq \mathbb{R}^m$ *be a convex and closed cone such that* $\mathrm{int}\, P \neq \emptyset$, *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) $\operatorname{cone}(A + \operatorname{int} P)$ *is pointed*;
(c) $\operatorname{conv}(A) \cap (-\operatorname{int} P) = \emptyset$.

Proof (a) $\Longrightarrow$ (b) Suppose that $0 = \sum_{i=1}^{l} x_i$ with $x_i \in \operatorname{cone}(A + \operatorname{int} P)$. We shall prove that $x_i = 0$ for all i. By choice, $x_i = t_i(a_i + p_i)$ with $t_i \geq 0$, $a_i \in A$, $p_i \in \operatorname{int} P$ for $i = 1, \dots, l$. We may assume that $\sum_{i=1}^{l} t_i > 0$, which implies that $\sum_{i=1}^{l} t_i a_i \in -\operatorname{int} P$. This yields a contradiction under (a), since the inequality in (a) also holds for all $a \in \operatorname{conv} A$, and

$$\operatorname{int} P = \bigl\{p \in P : \langle q, p\rangle > 0, \forall q \in P^*, q \neq 0\bigr\}.$$

(b) $\Longrightarrow$ (c) By (1), $\operatorname{cone}(\operatorname{conv}(A) + \operatorname{int} P)$ is pointed. Assume on the contrary that $\operatorname{conv}(A) \cap (-\operatorname{int} P) \neq \emptyset$. Then, $0 \in \operatorname{conv}(A) + \operatorname{int} P$. This implies that

$$\operatorname{cone}\bigl(\operatorname{conv}(A) + \operatorname{int} P\bigr) = \mathbb{R}^m,$$

contradicting (b).

(c) $\Longrightarrow$ (a) By applying a standard theorem on separation of convex sets, we get the existence of $p \in \mathbb{R}^n \setminus \{0\}$ and $\alpha \in \mathbb{R}$ such that

$$\langle p, z\rangle \geq \alpha, \quad \forall z \in \operatorname{conv} A, \quad \text{and} \quad \langle p, w\rangle \leq \alpha, \quad \forall w \in -\operatorname{cl}(\operatorname{int} P) = -P. \tag{2}$$

From the first inequality of (2), it follows that $\langle p, a\rangle \geq \alpha$ for all $a \in A$, and from the second inequality we get $\alpha \geq 0$. Hence $p \in P^*$, proving the desired result. □

Remark 2.1 It is not difficult to check that for any set $A \subseteq \mathbb{R}^m$,

$$\begin{aligned} A \cap (-\operatorname{int} P) = \emptyset \quad &\Longleftrightarrow \quad \operatorname{cl}(A) \cap (-\operatorname{int} P) = \emptyset \quad \Longleftrightarrow \quad (A + P) \cap (-\operatorname{int} P) = \emptyset \\ &\Longleftrightarrow \quad \operatorname{cl}\bigl(\operatorname{cone}(A + P)\bigr) \cap (-\operatorname{int} P) = \emptyset. \end{aligned}$$

We can go further when A is the image of a subset $C \subseteq \mathbb{R}^n$ through a linear transformation $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^m$. The next proposition can be considered as a Gordan-type theorem of the alternative, and it has its origin in Proposition 2.7 of [4].

Proposition 2.1 *Let $\mathcal{F}$ be a real matrix of order $m \times n$, and write $\mathcal{F}^\top = (\mathcal{F}_1^\top \dots \mathcal{F}_m^\top)$, where $\mathcal{F}_i$ is the ith row of $\mathcal{F}$. Let $C \subseteq \mathbb{R}^n$ be any nonempty set. Then*

$$\mathcal{F}(C) \cap \bigl(-\operatorname{int} \mathbb{R}^m_+\bigr) = \emptyset \quad \Longleftrightarrow \quad \max_{1 \leq i \leq m} \bigl\langle \mathcal{F}_i^\top, v\bigr\rangle \geq 0, \quad \forall v \in \operatorname{cl}(C),$$

and the following statements are equivalent:

(a) $\operatorname{cone}(\mathcal{F}(C) + \operatorname{int} \mathbb{R}^m_+)$ *is pointed*;
(b) $\mathcal{F}(\operatorname{cl}(\operatorname{conv}(C))) \cap (-\operatorname{int} \mathbb{R}^m_+) = \emptyset$;
(c) $\mathcal{F}(\operatorname{conv}(C)) \cap (-\operatorname{int} \mathbb{R}^m_+) = \emptyset$;
(d) $\max_{1 \leq i \leq m} \langle \mathcal{F}_i^\top, v\rangle \geq 0, \forall v \in \operatorname{cl}(\operatorname{conv}(C))$;
(e) $\operatorname{conv}(\{\mathcal{F}_i^\top : i = 1, \dots, m\}) \cap C^* \neq \emptyset$.

Proof The first part is straightforward. By the previous theorem,

$$\operatorname{cone}\big(\mathcal{F}(C)+\operatorname{int}\mathbb{R}^m_+\big) \text{ is pointed} \quad\Longleftrightarrow\quad \operatorname{conv}\big(\mathcal{F}(C)\big)\cap\big(-\operatorname{int}\mathbb{R}^m_+\big)=\emptyset.$$

It is easy to check that $\operatorname{conv}(\mathcal{F}(C))=\mathcal{F}(\operatorname{conv}(C))$ and

$$\begin{aligned}\mathcal{F}\big(\operatorname{conv}(C)\big)\cap\big(-\operatorname{int}\mathbb{R}^m_+\big)=\emptyset \quad&\Longleftrightarrow\quad \mathcal{F}\big(\operatorname{cl}\big(\operatorname{conv}(C)\big)\big)\cap\big(-\operatorname{int}\mathbb{R}^m_+\big)=\emptyset\\ &\Longleftrightarrow\quad \operatorname{cl}\big(\mathcal{F}\big(\operatorname{conv}(C)\big)\big)\cap\big(-\operatorname{int}\mathbb{R}^m_+\big)=\emptyset.\end{aligned}$$

Both relations, along with (a) of Theorem 2.1, amount to writing

$$\operatorname{conv}\big(\big\{\mathcal{F}_i^\top : i=1,\ldots,m\big\}\big)\cap C^*\neq\emptyset,$$

allowing us to conclude all the remaining equivalences. □

3 The Fritz John Optimality Conditions of the Alternative-Type

Let us consider the minimization problem with inequality constraints:

$$\min f(x) \quad \text{s.t.} \quad g_i(x)\le 0, \quad i=1,\ldots,m; \quad x\in X, \tag{3}$$

where $f, g_i : \mathbb{R}^n \to \mathbb{R}$, $i=1,\ldots,m$, are given functions, and $X\subseteq\mathbb{R}^n$ is any nonempty set. Define the feasible set to (3) as

$$K := \big\{x\in X : g_i(x)\le 0, i=1,\ldots,m\big\}.$$

For a fixed $\bar{x}\in K$, we associate its active index set as

$$I = I(\bar{x}) := \big\{i : g_i(\bar{x})=0\big\}. \tag{4}$$

Based on Proposition 2.1, we establish a new version, as an alternative-type result, of the Fritz John optimality conditions at points not necessarily optimal, which is new in the literature.

Theorem 3.1 (Fritz John necessary optimality conditions of alternative-type) *Let us consider problem* (3) *and $\bar{x}\in K$, with $X\subseteq\mathbb{R}^n$. Let f, g_i, $i\in I$, be differentiable at $\bar{x}$. Then, exactly one of the following two assertions holds*:

(a) *There exists $v\in\mathbb{R}^n$ such that*

$$\begin{aligned}&\big\langle\nabla f(\bar{x}), v\big\rangle<0, \qquad v\in\operatorname{cl}\big(\operatorname{conv}\big[T(X;\bar{x})\big]\big);\\ &\big\langle\nabla g_i(\bar{x}), v\big\rangle<0, \qquad i\in I,\end{aligned} \tag{5}$$

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

$$\lambda_0\nabla f(\bar{x})+\sum_{i\in I}\lambda_i\nabla g_i(\bar{x})\in\big[T(X;\bar{x})\big]^*, \tag{6}$$

or equivalently, $\max_{i\in I}\{\langle\nabla f(\bar{x}), v\rangle, \langle\nabla g_i(\bar{x}), v\rangle\}\ge 0, \forall v\in T(X;\bar{x})$.

Furthermore, if each g_i is differentiable at $\bar{x}$, condition (6) *can be written as*

$$\begin{cases} \lambda_0 \nabla f(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) \in [T(X;\bar{x})]^*; \\ \lambda_i g_i(\bar{x}) = 0, \quad i = 1, \ldots, m. \end{cases} \tag{7}$$

Proof This is a direct application of Proposition 2.1 to

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

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

Before continuing, some remarks about the previous result are in order. First of all, it is well-known that, besides differentiability of g_i at $\bar{x}$, for $i \in I$, the continuity of g_i at $\bar{x}$ for $i \notin I$ and the local minimality of $\bar{x}$ imply that the system

$$\begin{aligned} &\langle \nabla f(\bar{x}), v \rangle < 0, \qquad v \in T(X;\bar{x}); \\ &\langle \nabla g_i(\bar{x}), v \rangle < 0, \qquad i \in I \end{aligned} \tag{8}$$

has no solution (it will be proved in Corollary 3.1 by completeness). Thus, if $T(X;\bar{x})$ is convex, we have the impossibility of system (5), and so, by Theorem 3.1, (b) holds. This will be expressed in Corollary 3.1, whose formulation is very versatile as we shall show in Sect. 4, and it encompasses many recent results appearing in the literature. Thus, the interesting case is when $T(X;\bar{x})$ is nonconvex. On the other hand, we cannot expect, in general, that proving the impossibility of (5) is easier than checking the fulfillment of (6). Although there are situations where it happens; see Example 3.1, for instance (with one single inequality constraint). Indeed, there it holds

$$G_0(\bar{x}) := \big\{ v \in \mathrm{cl}\big(\mathrm{conv}\big[T(X;\bar{x})\big]\big) : \langle \nabla g(\bar{x}), v \rangle < 0 \big\} = \emptyset \tag{9}$$

and

$$F_0(\bar{x}) := \big\{ v \in \mathrm{cl}\big(\mathrm{conv}\big[T(X;\bar{x})\big]\big) : \langle \nabla f(\bar{x}), v \rangle < 0 \big\} = \emptyset. \tag{10}$$

Observe that in the mentioned example, it is verified that

$$\big\{ \nabla f(\bar{x}), \nabla g(\bar{x}) \big\} \subseteq T(X;\bar{x}) \quad \text{and} \quad \mathrm{cl}\big(\mathrm{conv}\big[T(X;\bar{x})\big]\big) = \big[T(X;\bar{x})\big]^*,$$

from which the emptiness of $G_0(\bar{x})$ or $F_0(\bar{x})$ easily follows. One can further develop this line of reasoning to show that a class of problems, without a convexity assumption on $T(X;\bar{x})$, under which system (5) has no solution.

We now compare with a result which, in our opinion, is the most general one concerning the validity of the Fritz John optimality conditions. It appears in the book by Giorgi, Guerraggio and Therfelder [5], Theorem 3.6.5, but its origin goes back to [6]. Some recent remarks on Fritz John optimality conditions in the same direction as in [5] are presented in [7], Theorem 13.

Theorem 3.2 ([5], Theorem 3.6.5; [7], Theorem 13) *Let $\bar{x} \in K$ be a local solution to problem (3) with $X \subseteq \mathbb{R}^n$. Let f, g_i, $i \in I$, be differentiable at $\bar{x}$. Then, for every convex subcone T_1 of $T(X;\bar{x})$ there exist $\lambda_0 \geq 0$, $\lambda_i \geq 0$, $i \in I$, not all zero, satisfying*

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in [T_1]^*. \tag{11}$$

The next example shows an instance where our previous theorem is applicable, providing a sharper result, whereas Theorem 3.2 yields less information.

Example 3.1 Let us take $f(x_1, x_2) = x_1$, $g(x_1, x_2) = x_2$, 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

$$[T(X;\bar{x})]^* = \mathbb{R}^2_+ = \mathrm{cl}\big(\mathrm{conv}[T(X;\bar{x})]\big),$$

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 corresponding system (5) has no solution, and therefore there exist λ_0, $\lambda_1 \geq 0$, not both zero, such that

$$\lambda_0 \nabla f(\bar{x}) + \lambda_1 \nabla g(\bar{x}) \in [T(X;\bar{x})]^*. \tag{12}$$

In this case, any $\lambda_0 \geq 0$ and $\lambda_1 \geq 0$ satisfies such conditions.

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

$$T_1 = \{(x_1, 0) : x_1 \geq 0\}, \qquad T_2 = \{(0, x_2) : x_2 \geq 0\}.$$

It is easy to see that for any of these cones, (11) provides less information than (12). Other candidates for T_1 are: the Clarke tangent cone of X at $\bar{x}$, $T_C(X;\bar{x})$, which is always convex and 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 $\bar{x}$, $Q(X;\bar{x})$ [8], which is also empty here.

We now establish a corollary that is a consequence of our Theorem 3.1. The versatility of its formulation is shown in Sect. 4.

Corollary 3.1 (Fritz John necessary optimality conditions) *Let us consider problem (3) and $\bar{x} \in K$. Let $X \subseteq \mathbb{R}^n$ be such that $T(X;\bar{x})$ is convex. Let f, g_i, $i \in I$, be differentiable at $\bar{x}$; g_i, $i \notin I$, be continuous at $\bar{x}$. If $\bar{x}$ is a local solution to (3), then there exist $\lambda_0 \geq 0$, $\lambda_i \geq 0$, $i \in I$, not all zero, satisfying*

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in [T(X;\bar{x})]^*. \tag{13}$$

Proof We claim that (a) of Theorem 3.1 does not hold, proving that the system (8) has no solution. Even though the proof is standard, we shall provide it just for convenience of the reader. Suppose, on the contrary, that v is a solution to (8) and there exist sequences $\lambda_k > 0$, $x_k \in X$, $x_k \to \bar{x}$, satisfying $\lambda_k(x_k - \bar{x}) \to v$. By differentiability at $\bar{x}$,

$$f(x_k) - f(\bar{x}) = \big\langle \nabla f(\bar{x}), x_k - \bar{x} \big\rangle + \|x_k - \bar{x}\| o\big(\|x_k - \bar{x}\|\big)$$

with $o(t) \to 0$ as $t \to 0$. Multiplying this equality by λ_k, letting $k \to +\infty$ and using the first inequality of (8), we get the existence of k_1 such that

$$f(x_k) < f(\bar{x}), \quad \forall k \geq k_1. \tag{14}$$

It only remains to check that x_k is feasible for all k sufficiently large to reach a contradiction.

Let $i \in I$. We get similarly as for f

$$g_i(x_k) - g_i(\bar{x}) = \big\langle \nabla g_i(\bar{x}), x_k - \bar{x} \big\rangle + \|x_k - \bar{x}\| o\big(\|x_k - \bar{x}\|\big).$$

Multiplying this equality by λ_k and letting $k \to +\infty$, we obtain, for some k_2,

$$g_i(x_k) < 0, \quad \forall i \in I, \forall k \geq k_2. \tag{15}$$

Since g_i is continuous for $i \notin I$, there exists k_3 such that

$$g_i(x_k) < 0, \quad \forall i \notin I, \forall k \geq k_3. \tag{16}$$

Combining (15) and (16), we conclude that x_k is feasible for all k sufficiently large. Hence $\bar{x}$ cannot be a local solution to (3), showing that (a) of the previous theorem is not valid, and therefore (b) holds, and the corollary follows. $\square$

4 Some Particular Situations

Let us show some interesting specializations appearing in the literature where $T(X; \bar{x})$ is convex.

4.1 The Set X Is Not Convex with $T(X; \bar{x})$ Being Convex

Let us consider the problem with an additional quadratic equality constraint:

$$\min f(x) \quad \text{s.t.} \quad g_i(x) \leq 0, \quad i = 1, \ldots, m; \quad q(x) = 0; \quad x \in \mathbb{R}^n, \tag{17}$$

where q is a function of the form

$$q(x) := \frac{1}{2} x^\top A x + a^\top x + \alpha,$$

with A being a (real) symmetric matrix, $a \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. Clearly, the set

$$X := \big\{ x \in \mathbb{R}^n : q(x) = 0 \big\}$$

is not necessarily convex even if q is convex. Let $\bar{x}$ be feasible for problem (17). It is not difficult to find that (see, for instance, [9], Theorem 2.1)

$$T(X;\bar{x}) = \{v \in \mathbb{R}^n : (A\bar{x}+a)^\top v = 0\} \quad \text{if } A\bar{x}+a \neq 0;$$

whereas

$$T(X;\bar{x}) = \{v \in \mathbb{R}^n : v^\top A v = 0\} \quad \text{if } A\bar{x}+a = 0.$$

This set, in general, is nonconvex. If, additionally, q is convex, that is, A is positive semidefinite, a more precise formulation may be obtained since $v^\top A v = 0 \Longleftrightarrow Av = 0$:

$$T(X;\bar{x}) = \begin{cases} (A\bar{x}+a)^\perp & \text{if } A\bar{x}+a \neq 0; \\ \ker A & \text{if } A\bar{x}+a = 0. \end{cases} \tag{18}$$

Thus,

$$[T(X;\bar{x})]^* = \begin{cases} \mathbb{R}(A\bar{x}+a) & \text{if } A\bar{x}+a \neq 0; \\ (\ker A)^\perp = A(\mathbb{R}^n) & \text{if } A\bar{x}+a = 0. \end{cases}$$

Hence, the Fritz John conditions (13) reduce to the existence of λ_0, λ_i, $i \in I$, not all zero, $\lambda \in \mathbb{R}$, $y \in \mathbb{R}^n$, such that

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i\in I} \lambda_i \nabla g_i(\bar{x}) = \begin{cases} \lambda(A\bar{x}+a) & \text{if } A\bar{x}+a \neq 0; \\ Ay & \text{if } A\bar{x}+a = 0. \end{cases}$$

4.2 The Set X Is Convex

In this case, $T(X;\bar{x})$ is convex as well, and since $T(X;\bar{x}) = \mathrm{cl}(\bigcup_{t\geq 0} t(X-\bar{x}))$, we get $[T(X;\bar{x})]^* = -N(X;\bar{x})$, and so (13) can be written as

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i\in I} \lambda_i \nabla g_i(\bar{x}) \in -N(X;\bar{x}).$$

Thus, Theorem 3.2.2 in [2] is obtained.

4.3 The Set X Is Open or $\bar{x} \in \operatorname{int} X$

In this situation, $T(X;\bar{x}) = \mathbb{R}^n$, and therefore condition (13) reduces to

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

This is nothing else than Theorem 4.2.8 in [1], and Theorem 3.2.1 in [2] when X is the entire space $\mathbb{R}^n$.

4.4 The Set X Is an Affine Subspace

This case deals with $X = \{x \in \mathbb{R}^n : Hx = d\} = \bar{x} + \ker H$, with H being a real $p \times n$ matrix and $\bar{x} \in X$. Thus, we obtain

$$T(X;\bar{x}) = \ker H \quad \text{and} \quad \big[T(X;\bar{x})\big]^* = (\ker H)^* = (\ker H)^{\perp}.$$

Hence (13) is expressed as

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i\in I} \lambda_i \nabla g_i(\bar{x}) \in (\ker H)^{\perp} = H^{\top}\big(\mathbb{R}^p\big),$$

that is, there exist $y_i \in \mathbb{R}$, $i = 1, \dots, p$, such that

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

4.5 The Set X Is Polyhedral

Let us consider (see, for instance, Birbil et al. [10])

$$\min f(x) \quad \text{s.t.} \quad g_i(x) \le 0, \quad i = 1, \dots, m; \quad h_j^{\top} x \le d_j, \quad j = 1, \dots, p; \quad x \in \mathbb{R}^n. \tag{19}$$

In this case, we can take $X = \{x \in \mathbb{R}^n : Hx \le d\}$ with H being a $p \times n$ matrix, and refine (13). More precisely, by denoting $h_j^{\top}$ to be the rows of H and setting $J := \{j : h_j^{\top}\bar{x} = d\}$, the conclusion of Theorem 3.1 reduces to the existence of $\lambda_0 \ge 0$, $\lambda_i \ge 0$, $i \in I$, not all zero, and $u_j \ge 0$, $j \in J$, such that

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i\in I} \lambda_i \nabla g_i(\bar{x}) + \sum_{j\in J} u_j h_j = 0. \tag{20}$$

Indeed, by setting $C \doteq \{v \in \mathbb{R}^n : h_j^{\top} v \le 0, j \in J\}$, which is a closed convex cone, one can easily check that $T(X;\bar{x}) = C$ (see Lemma 5.1.4 in Bazaraa et al. [1]) By applying Proposition 2.1, we get the existence of $\lambda_0 \ge 0$, $\lambda_i \ge 0$, $i \in I$, not all zero, such that

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

The conclusion follows once we notice that $C^* = \{H_J u : u \le 0\}$, with H_J being the matrix with columns h_j for $j \in J$. Thus we recovered the Fritz John conditions as they appear in [10].

5 The Karush–Kuhn–Tucker Optimality Conditions

For the sake of completeness, we now derive the necessary optimality condition due to Karush, Kuhn and Tucker from the Fritz John conditions, without the local optimality of $\bar{x}$.

Theorem 5.1 (Karush, Kuhn and Tucker necessary optimality conditions) *Let us consider* (3). *Let* $f, g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I$, *be differentiable functions at* $\bar{x} \in K$. *Assume that system* (5) *has no solution and*

$$\operatorname{conv}\big(\{\nabla g_i(\bar{x}) : i \in I\}\big) \cap \big[T(X;\bar{x})\big]^* = \emptyset \tag{21}$$

holds, then there exist real numbers $\lambda_i \geq 0$, $i \in I$, *such that*

$$\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \in \big[T(X;\bar{x})\big]^*. \tag{22}$$

Proof By Theorem 3.1, we get (6). If $\lambda_0 = 0$, we obtain

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

Since not all λ_i are zero, the previous condition contradicts (21). Hence $\lambda_0 > 0$, and so (22) is satisfied. □

Let us show that the constraint qualification (21) encompasses some of the classical constraint qualifications appearing in the literature. It will be discussed in case $\bar{x}$ is a local solution and $T(X;\bar{x})$ is convex. First of all, by Proposition 2.1, (21) is equivalent to the existence of $v \in T(X;\bar{x})$ satisfying $\langle \nabla g_i(\bar{x}), v\rangle < 0$, $i \in I$, which is a Mangasarian–Fromovitz-type constraint qualification.

5.1 Linear Independence when X Is Open ([1], Theorem 4.2.13; [2], Theorem 3.3.1)

In this case, (21) reduces to

$$\sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) \neq 0, \quad \text{for all } \lambda_i \geq 0, i \in I, \sum_{i \in I} \lambda_i = 1,$$

which is implied by the linear independence of $\{\nabla g_i(\bar{x}) : i \in I\}$. We will see that our constraint qualification (21) (when X is open) is weaker than the linear independence of $\{\nabla g_i(\bar{x}) : i \in I\}$. Indeed, simply take $f(x_1, x_2) = x_1^2 + x_2^2$, $g_1(x_1, x_2) = x_1^2 + x_1$, $g_2(x_1, x_2) = x_1^2 + 2x_1$ and $g_3(x_1, x_2) = x_1^2 + x_2^2 - 1$. Here, $X = \mathbb{R}^2$ and $\bar{x} = (0, 0)$ and $I = \{1, 2\}$. Then, $\{\nabla g_1(0), \nabla g_2(0)\} = \{(1, 0), (2, 0)\}$ is not linearly independent but condition (21) is satisfied. Hence the KKT optimality conditions hold.

5.2 Linear Equality Constraints and Full Column Rank ([11], Theorem 1)

We consider $X = \{x \in \mathbb{R}^n : Hx = d\}$, with H being a real matrix of order $p \times n$. Our constraint qualification (21) is implied by the assumption that the matrix $(\nabla g_I(\bar{x}) H^\top)^\top$ has full row rank, where $(\nabla g_I(\bar{x}))$ denotes the matrix whose columns are the gradients $\nabla g_i(\bar{x})$ for $i \in I$. Indeed, under the assumption of full column rank, the system

$$\big\langle \nabla g_i(\bar{x}), v\big\rangle < 0, \quad i \in I, Hv = 0, v \in \mathbb{R}^n,$$

has a solution. Let v be one of them. From this it follows that any ξ belonging to $\mathrm{co}(\{\nabla g_i(\bar{x}) : i \in I\})$ satisfies $\langle \xi, v \rangle < 0$, which implies that (21) holds, since $v \in \ker H$ and

$$\left[T(X;\bar{x})\right]^* = (\ker H)^* = (\ker H)^{\perp}.$$

5.3 X is Polyhedral ([10], Sect. 1)

By considering problem (19) in Sect. 4.5, we re-obtain the result in [10]. Indeed, we notice the constraint qualification in this paper: there exists $v \in C$ such that $\max_{i \in I}\{\langle \nabla g_i(\bar{x}), v \rangle\} < 0$ is equivalent to (21).

6 Conclusions

An alternative-type version of the Fritz John optimality conditions is presented. Such a result may be seen as a characterization of the FJ points beyond optimality and convexity. It means that further investigation should be carried out to understand the nature of FJ points and so of KKT points. Indeed, the problem of characterizing KKT points, without any constraint qualification (CQ), will be the subject of future research. This is motivated by applications where some mathematical programs with equilibrium constraints (for instance, in structural optimization) do not satisfy standard CQ.

From our characterization of FJ points, the standard result under convexity of $T(X;\bar{x})$, provided $\bar{x}$ is a local solution, is recovered. Hence, an open question is to find classes of problems for which proving the impossibility of (5) is easier than checking the fulfillment of (6) whenever $T(X;\bar{x})$ is not convex. In this respect, we exhibit a problem with inequality constraint, which may lead to a further development and finding new conditions implying the validity of the FJ condition.

Acknowledgements This research was supported in part by CONICYT-Chile through FONDECYT 112-0980, BASAL Projects, CMM, Universidad de Chile, and MATH-AmSud 13math-01.

References

1. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms. Wiley, Hoboken (2006)
2. Bector, C.R., Chandra, S., Dutta, J.: Principles of Optimization Theory. Alpha Science International, Oxford (2005)
3. Flores-Bazán, F., Hadjisavvas, N., Vera, C.: An optimal alternative theorem and applications to mathematical programming. J. Glob. Optim. **37**, 229–243 (2007)
4. Flores-Bazán, F., Flores-Bazán, F., Vera, C.: Gordan-type alternative theorems and vector optimization revisited. In: Ansari, Q.H., Yao, J.-C. (eds.) Recent Developments in Vector Optimization, pp. 29–59. Springer, Berlin (2012)
5. Giorgi, G., Guerraggio, A., Therfelder, J.: Mathematics of Optimization: Smooth and Nonsmooth Case. Elsevier, Amsterdam (2004)
6. Bazaraa, M.S., Goode, J.J.: Necessary optimality criteria in mathematical programming in the presence of differentiability. J. Math. Anal. Appl. **40**, 609–621 (1972)

7. Giorgi, G.: Remarks in Fritz John conditions for problems with inequality and equality constraints. Int. J. Pure Appl. Math. **71**, 643–657 (2011)
8. Giorgi, G., Guerraggio, A.: First order generalized optimality conditions for programming problems with a set constraint. In: Komlosi, S., Rapcsàk, T., Schaible, S. (eds.) Generalized Convexity, Proceedings, Pécs, Hungary, 1992, pp. 171–185. Springer, Berlin (1994)
9. Di, S., Poliquin, R.: Contingent cone to a set defined by equality and inequality constraints at a Fréchet differentiable point. J. Optim. Theory Appl. **81**, 469–478 (1994)
10. Birbil, S.I., Frenk, J.B.G., Still, G.J.: An elementary proof of the Fritz John and Karush–Kuhn–Tucker conditions in nonlinear programming. Eur. J. Oper. Res. **180**, 479–484 (2007)
11. Brezhneva, O.A., Tret'yakov, A.A., Wright, S.E.: A simple and elementary proof of the Karush–Kuhn–Tucker theorem for inequality-constrained optimization. Optim. Lett. **3**, 7–10 (2009)