International Journal of Pure and Applied Mathematics
Volume 71 No. 4 2011, 643-657

REMARKS ON FRITZ JOHN CONDITIONS FOR PROBLEMS WITH INEQUALITY AND EQUALITY CONSTRAINTS

Giorgio Giorgi
Dipartimento di Economia Politica e Metodi Quantitativi
University of Pavia
5, via San Felice, 27100, Pavia, ITALY

Abstract: This paper contains several remarks and results on the necessary optimality conditions of Fritz John type for a nonlinear programming problem with inequality and equality constraints. The first part is concerned with a simple and brief proof of the Fritz John conditions, proof which benefits from the use of the Bouligand tangent cone. The second part is concerned with the same problem, where a set constraint (which covers the constraints which cannot be expressed by means of neither equalities nor inequalities) is added. Several remarks and results are given to deepen the study of Fritz John type conditions for this case.

AMS Subject Classification: 90C30, 90C46
Key Words: Fritz John conditions, set constraint

1. Introduction

The Fritz John theorem is one of the most important results in mathematical programming. When there are, besides inequality constraints, also equality constraints, the existing proofs are usually quite long and intricated. This is the case, for example, of the paper of Mangasarian and Fromovitz (1967), perhaps the first paper dealing with this topic, of the book of Bazaraa and Shetty (1967) and of Bazaraa, Sherali and Shetty (1993), of the paper of Still and Streng (1996), etc. An interesting paper of McShane (1973) uses the penalty approach and therefore it is useful in those courses on optimization, where also the computational aspects are treated. A similar approach is adopted also by

Received: May 17, 2011

Bertsekas (1999). Pourciau (1980, 1982) presents a proof (that he considers elementary) based on the Lipschitz Fixed Point Theorem, "a result familiar to undergraduates". We do not agree on this point. Recently, Birbil, Frenk and Still (2007) have given a quite elementary proof of the Fritz John (and Karush-Kuhn-Tucker) conditions for a nonlinear programming problem with equality and/or inequality constraints. Their proof relies on some basic results in linear programming and the Bolzano-Weierstrass theorem for compact sets. In our opinion a fruitful approach, also in order to "prepare the ground" for other basic questions of mathematical programming theory (constraint qualifications, Karush-Kuhn-Tucker conditions, etc.), comes from a systematic use of the Bouligand tangent cone and its properties.

This note is organized as follows. In Section 2 we give a quite elementary and brief proof of the classical Fritz John Theorem for a nonlinear programming problem with both inequality and equality constraints and no abstract constraint. In Section 3 we make some considerations for a problem where, besides inequality and equality constraints, there is also an abstract constraint, i. e. a set constraint (which covers the constraints which cannot be expressed by means of neither equalities nor inequalities).

We consider the following two optimization problems

(P) $$\underset{x \in S}{\text{minimize}} f(x)$$

where

$$S = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i \in I = \{1, ...m\}, h_j(x) = 0, \\ j \in J = \{1, ...p\},\ p \leq n\},$$

(P_1) $$\underset{x \in S_1}{\text{minimize}} f(x)$$

where $S_1 = S \cap C$, C being a subset of $\mathbb{R}^n$.

We suppose that f and every $g_i, i \in I$, are differentiable at a feasible point x^0 of (P) or (P_1) and that every $h_j, j \in J$, is continuously differentiable in a neighbourhood of x^0 (we could assume weaker differentiability conditions, but we prefer to adopt a classical approach).

Definition 1. Let $x^0 \in S$; the cone

$$K(x^0) = \left\{y \in \mathbb{R}^n : \nabla g_i(x^0)y \leq 0, \forall i \in I(x^0); \nabla h_j(x^0)y = 0, \forall j \in J\right\},$$

where $I(x^0) = \left\{i \in I : g_i(x^0) = 0\right\}$ is the active index set at x^0, is called the linearizing cone at x^0 for (P).

The cone

$$K^0(x^0) = \left\{y \in \mathbb{R}^n : \nabla g_i(x^0)y < 0, \forall i \in I(x^0); \nabla h_j(x^0)y = 0, \forall j \in J\right\}$$

is called the weak linearizing cone at x^0 or cone of the decreasing directions at x^0 for (P).

A basic tool for treating optimality conditions for (P) or (P_1) is some *local cone approximation* of a set S at a point $x^0 \in S$ (or also $x^0 \in clS$). See, for example, Bazaraa and Shetty (1976), Aubin and Frankowska (1990), Giorgi, Gurraggio and Thierfelder (2004), Giorgi and Guerraggio (1992, 2002).

Definition 2. A sequence $\left\{x^k\right\} \subset R^n \diagdown \left\{x^0\right\}$, with $x^k \longrightarrow x^0$ is said to converge to x^0 in the direction $y \in R^n$ if it holds

$$\lim_{k \to \infty} \frac{x^k - x^0}{\|x^k - x^0\|} = y.$$

Definition 3. Given a set $S \subset R^n$ and a point $x^0 \in S$ (or also $x^0 \in clS$), the cone

$$T(S, x^0) = \left\{\lambda y \in \mathbb{R}^n : \exists \left\{x^k\right\} \subset S, \lim_{k \to \infty} \frac{x^k - x^0}{\|x^k - x^0\|} = y, \lambda \geq 0\right\}$$

is called the Bouligand tangent cone to S at x^0 or also the contingent cone to S at x^0.

It is well known that the contingent cone is a closed cone, but in general not convex. If x^0 is an isolated point of S, then $T(S, x^0) = \{0\}$. The contingent cone $T(S, x^0)$ can be described in various alternative forms, e. g.,

$$T(S, x^0) = \left\{y \in \mathbb{R}^n : \exists \left\{x^k\right\} \subset S, \exists \lambda_k \subset \mathbb{R}_+ \text{ such that } x^k \to x^0, \lambda_k(x^k - x^0) \to y\right\},$$

$$T(S, x^0) = \{y \in \mathbb{R}^n : \forall N(y), \forall \lambda > 0, \exists t \in (0, \lambda), \exists \bar{y} \in N(y) \text{ such that } x + t\bar{y} \in S\}.$$

$T(S, x^0)$ depends only on the structure of S around the point x^0, i. e. if $U(x^0)$ is any neighborhood of x^0, we have $T(S, x^0) = T(S \cap U(x^0), x^0)$. Moreover, if S is a convex set, it holds

$$T(S, x^0) = \operatorname{cl} \operatorname{cone}(S - x^0),$$

where cone S is the convex cone generated by S.

We recall the *Motzkin theorem of the alternative* (see, e. g., Giorgi, Guerraggio and Thierfelder (2004), Mangasarian (1969)).

Theorem 4. *Let be given the matrices A, B and C. The following system*

$$Ay < 0, By \leq 0, Cy = 0$$

is impossible if and only if there exist vectors $u \geq 0, u \neq 0, v \geq 0$ and w (sign-free) such that

$$uA + vB + wC = 0.$$

2. Fritz John Necessary Conditions for Problem (P)

Theorem 5. (Fritz John Theorem) *Let $x^0 \in S$ be a local solution for (P), where f and every $g_i, i \in I$, are differentiable at x^0 and every $h_j, j \in J$, is continuously differentiable in a neighborhood of x^0. Then there exist multipliers $u_0, u_1, ..., u_m \geq 0$ and $v_1, ..., v_p$, not all zero, such that*

$$u_0 \nabla f(x^0) + \sum_{i=1}^{m} u_i \nabla g_i(x^0) + \sum_{j=1}^{p} v_j \nabla h_j(x^0) = 0,$$

$$u_i g_i(x^0) = 0, \quad \forall i \in I.$$

For the proof of this theorem we need some preliminary results, whose utility is also "autonomous" for other questions of nonlinear programming. The next lemma is proved in many books and papers (see, e. g., Hestenes (1975), Gould and Tolle (1971), Varaiya (1967), Guignard (1969), Bazaraa and Shetty (1976)) and is a basic result for optimization problems with only a set constraint. The proof is easy and very brief and will not be repeated here.

Lemma 6. *Let x^0 be a local solution of the problem*

$$\underset{x \in C}{\text{minimize}} f(x)$$

where f is differentiable on an open set containing $C \subset R^n$. Then it holds

$$\nabla f(x^0) y \geq 0, \forall y \in T(C, x^0).$$

We note that the thesis of the above lemma can be equivalently formulated as

$$-\nabla f(x^0) \in \left[T(C, x^0)\right]^*$$

where A^* is the (negative) polar cone of the set $A \subset \mathbb{R}^n$.

Lemma 7. *Let* $x^0 \in S$ *and let* $f, g_i, i \in I$, *and* $h_j, j \in J$, *verify the assumptions of* Theorem 5. *Then we have:*

a) $T(S, x^0) \subset K(x^0)$;

b) If the Jacobian matrix $\nabla h(x^0)$ *has full rank, it holds* $K^0(x^0) \subset T(S, x^0)$;

c) If, moreover, $K(x^0) \neq \varnothing$, *then it holds* $K(x^0) \subset T(S, x^0) = clK^0(x^0)$.

Proof. First of all note that, given relation a), condition c) entails the equality $K(x^0) = T(S, x^0)$.

a) The proof of this part can be found, e. g., in Bazaraa and Shetty (1976) and will not be repeated here. Anyhow, the proof is very brief and easy; moreover, this relation is not essential for the proof of Theorem 5.

b) and c). First we prove c) for a particular case, i. e. we consider the "classical" case of no inequality constraints, i. e. the constraint set is

$$S_0 = \{x \in \mathbb{R}^n : h_j(x) = 0, j = 1, ...p\},$$

where $h_j, j \in J$, is continuously differentiable in a neighbourhood of x^0 and the vectors $\nabla h_j(x^0), j \in J$, are linearly independent. $K^0(x^0)$ is defined accordingly. By the Implicit Function Theorem, we can solve the nonlinear system $h(x) = 0$ in a neighborhood of x^0, expressing p "basic variables" as a function of the remaining $n - p$ "non basic variables": $x_B = H(x_{NB})$. We have also, for $H : \mathbb{R}^{n-p} \longrightarrow \mathbb{R}^p$,

$$\nabla H(x^0_{NB}) = -(\nabla_{NB} h(x^0))(\nabla_B h(x^0))^{-1.}$$

For a vector $y = \begin{pmatrix} y_B \\ y_{NB} \end{pmatrix} \in K^0(x^0)$ we have therefore

$$\nabla h(x^0)y = \nabla_B h(x^0)y_B + \nabla_{NB} h(x^0)y_{NB} = 0,$$

equivalent to $y_B = \nabla H(x^0_{NB})y_{NB}$. We suppose $y_{NB} \neq 0$ (otherwise also $y_B = 0$ and the thesis would be trivial). Without loss of generality we suppose also $\|y_{NB}\| = 1$. There exists therefore a sequence $\left\{x^k_{NB}\right\}$ of non basic variables which converges to x^0_{NB} in the direction y_{NB}. But then it is also convergent the sequence $\{x_B^k\}$ of the corresponding basic variables, with $x_B^k = H(x_{NB}^k)$, to $x_B^0 = H(x_{NB}^0)$. The sequence

$$\{x^k\} = \left\{\begin{pmatrix} x_B^k \\ x_{NB}^k \end{pmatrix}\right\} = \left\{\begin{pmatrix} H(x_{NB}^k \\ x_{NB}^k \end{pmatrix}\right\}$$

is then feasible and convergent to x^0 in the direction $y/\|y\|$, as

$$\frac{x_{NB}^k - x_{NB}^0}{\|x_{NB}^k - x_{NB}^0\|} \longrightarrow y_{NB}$$

and

$$\frac{H(x_{NB}^k) - H(x_{NB}^0)}{\|x_{NB}^k - x_{NB}^0\|} \longrightarrow \nabla H(x_{NB}^0)y_N = y_B.$$

Therefore we have proved that $y \in T(S, x^0)$. Now we prove b). We suppose $I(x^0) \neq \varnothing$, otherwise we would get $K^0(x^0) = K(x^0)$ and the thesis follows from what proved above. Let $y \in K^0(x^0)$, with $\|y\| = 1$, so this vector belongs (being $\nabla h_j(x^0)y = 0, \forall j \in J$ and thanks to what proved above) to the tangent cone of the set $\{x \in \mathbb{R}^n : h_j(x^0) = 0, j \in J\}$, formed of only equality constraints. That is, there exists a sequence$\{x^k\}$ in this set, which converges to x^0 in the direction y. Then, for what regards the active inequality contraints at x^0, the expression

$$\frac{g_i(x^k) - g_i(x^0)}{\|x^k - x^0\|} = \frac{\nabla g_i(x^0)(x^k - x^0) + o(\|x^k - x^0\|)}{\|x^k - x^0\|}$$

converges to $\nabla g_i(x^0)y < 0$, it holds also $g_i(x^k) < 0$ for all $k \in \mathbb{N}$ sufficiently large. For the non active constraints at x^0 the same inequality holds, thanks to continuity. The sequence $\{x^k\}$ is therefore (for large values of $k \in \mathbb{N}$) feasible. So we have proved that $y \in T(S, x^0)$. Now we complete the proof of c), even if this statement is not essential to the proof of Theorem 5. Obviously we have $K^0(x^0) \subset K(x^0)$ and therefore $clK^0(x^0) \subset K(x^0)$. Now we prove that in fact we have equality. Let $y \in K(x^0)$ and let us consider $\bar{y} \in K^0(x^0)$ ($K^0(x^0) \neq \varnothing$ by assumption). By defining $y^k = y + (1/k)\bar{y}$, we construct a sequence $\{y^k\}$ of elements of $K^0(x^0)$ converging to y for $k \longrightarrow \infty$. Therefore $y \in clK^0(x^0)$. We have therefore that $clK^0(x^0) = K(x^0)$. Thanks to a), the equality $T(S, x^0) = K(x^0)$ is proved if $K^0(x^0) \subset T(S, x^0)$. But this is just relation b), proved above. □

Remark 8. The condition

$$K^0(x^0) \neq \varnothing, \nabla h(x^0) \text{ of full rank}$$

is the well-known Mangasarian-Fromovitz constraint qualification, which assures that in Theorem 5 it holds $u_0 > 0$. Moreover, this constraint qualification is both necessary and sufficient for the set $(u_0, u_1, ..., u_m, v_1, ..., v_p)$ of Fritz John multipliers to be bounded and is necessary and sufficient in order that in all elements of the same set, it holds $u_0 > 0$.

Proof of Theorem 5. If the vectors $\left\{\nabla h_1(x^0), ...\nabla h_p(x^0)\right\}$ are linearly dependent, the thesis follows trivially. So, let $\nabla h(x^0)$ be of full rank. Being $x^0 \in S$ a local solution of (P), thanks to Lemmas 6 and 7 (part b)), we have $\nabla f(x^0)y \geq 0, \forall y \in T(S, x^0) \supset K^0(x^0)$. It will be $\nabla f(x^0)y \geq 0$ for all y such that

$$\begin{array}{ll} \nabla g_i(x^0)y < 0, & i \in I(x^0), \\ \nabla h_j(x^0)y = 0, & j \in J. \end{array}$$

That is, the system

$$\left\{ \begin{array}{ll} \nabla f(x^0)y < 0, & \\ \nabla g_i(x^0)y < 0, & i \in I(x^0), \\ \nabla h_j(x^0)y = 0, & j \in J, \end{array} \right.$$

has no solution $y \in \mathbb{R}^n$. Apply Theorem 4 (Motzkin alternative theorem) and put $u_i = 0, \forall i \notin I(x^0)$, to obtain at once the thesis. □

3. Fritz John Necessary Conditions for Problem (P_1)

When, as in problem (P_1), besides inequality and equality constraints, there is also a set constraint $C \subset \mathbb{R}^n$, with C not open, the classical Fritz John necessary conditions of Theorem 5 are no longer valid. Bazaraa and Shetty (1976) provide a counterexample; the conditions of Theorem 5 must be modified, obtaining (with suitable additional assumptions) necessary optimality conditions of the so-called "minimum principle type" (also called from some authors "generalized Lagrange multipliers rule").

Definition 9. Let $x^0 \in S \subset R^n$. The set

$$I(S, x^0) = \{y \in \mathbb{R}^n : \exists N(y), \exists \lambda > 0 \quad \text{such that}$$
$$\sum \forall t \in (0, \lambda), \forall v \in N(y) : x^0 + tv \in S\}$$

is called the cone of interior directions to S at x^0. The set

$$Q(S,x^0) = \{y \in \mathbb{R}^n : \exists N(y) \text{such that} \forall \delta > 0\ \exists t \in (0,\delta) \text{such that} \\ \forall v \in N(y) : x^0 + tv \in S\}$$

is called the cone of quasi-interior directions to S at x^0.

Both these cones are open and it holds $I(S,x^0) \subset Q(S,x^0) \subset T(S,x^0)$. Moreover, $I(S,x^0)$ and $Q(S,x^0)$ are nonempty if $intS \neq \varnothing$. If S is a convex set we have

$$I(S,x^0) = Q(S,x^0) = \text{cone}(intS - x^0).$$

If, moreover, $intS \neq \varnothing$, then

$$I(S,x^0) = intT(S,x^0),$$
$$T(S,x^0) = clI(S,x^0) = \text{cl cone}(S - x^0).$$

See, e.g., Aubin and Frankowska (1990), Giorgi, Guerraggio and Thierfelder (2004), Giorgi and Guerraggio (2002). We recall that, given a set $A \subset \mathbb{R}^n$, A^* is the (negative) polar cone of A.

Bazaraa and Goode (1972) (see also Bazaraa abd Shetty(1976)) prove the following result.

Theorem 10. *Let x^0 be a local solution of* (P_1) *and let the assumptions of* Theorem 2 *be verified. Moreover, let $I(C,x^0)$ be convex. Then there exist scalars $u_0, u_1, ..., u_m \geq 0$ and $v_1, ..., v_p$, not all zero, such that*

$$-\left[u_0 \nabla f(x^0) + \textstyle\sum_{i=1}^m u_i \nabla g_i(x^0) + \sum_{j=1}^p v_j \nabla h_j(x^0)\right] \in I^*(C,x^0)$$
$$u_i g_i(x^0) = 0, \forall i \in I.$$

Remark 11. Giorgi and Guerraggio (1994) give a sharper result, as they prove Theorem 10 with $I(C,x^0)$ substituted by the larger cone $Q(C,x^0)$. However, it is not possible, in general, to replace the open cones $I(C,x^0)$ or $Q(C,x^0)$ by the larger, but closed cone, $T(C,x^0)$. Counterexamples can be given.

Remark 12. We have seen that $I(C,x^0)$ or $Q(C,x^0)$ are convex if C is a convex set. More generally: a set $X \subset$ is locally convex at $x^0 \in X$ if there exists a neighborhood $N(x^0)$ such that $X \cap N(x^0)$ is convex. Then $I(C,x^0)$ or $Q(C,x^0)$ are convex if C is locally convex at x^0 (the same holds for other local cone approximations).

Now, let us consider a nonlinear programming problem with a set constraint but with no equality constraints, i. e.

$$(\mathrm{P}_2) \qquad \underset{x \in S_2}{\text{minimize}} f(x),$$

where $S_2 = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i \in I\} \cap C.$

Giorgi and Guerraggio (1994) prove the following result.

Theorem 13. *Let x^0 be a local solution of* (P_2), *where f and every $g_i, i \in I$, are differentiable at x^0. Then, there exist scalars $\lambda_0 \geq 0, \lambda_i \geq 0, i \in I$, not all zero, such that*

$$\begin{array}{l} -\left[\lambda_0 \nabla f(x^0) + \sum_{i=1}^m \lambda_i \nabla g_i(x^0)\right] \in T_1^*(C, x^0) \\ \lambda_i g_i(x^0) = 0, \forall i \in I, \end{array}$$

where $T_1(C, x^0)$ is a convex subcone of $T(C, x^0)$.

If $T(C, x^0)$ is a convex cone, Theorem 13 gives a sharper result than Theorem 10. This is not surprising, as (P_1) comtains also equality constraints. If $T(C, x^0)$ is not a convex cone, there exist several convex subcones of the same, that can be chosen to represent $T_1 C, x^0)$. It is well known that one of these convex cones is the *Clarke tangent cone* to $x^0 \in C$:

$$\begin{array}{rcl} T_{Cl}(C, x^0) & = & \left\{y \in \mathbb{R}^n : \forall N(y) \exists \lambda > 0, \exists V(x^0), \forall \bar{x} \in V(x^0) \cap C,\right. \\ & & \left. \forall t \in (0, \lambda) \exists \bar{y} \in N(y) : \bar{x} + t\bar{y} \in C\right\}. \end{array}$$

In terms of sequences:

$$\begin{array}{rcl} T_{Cl}(C, x^0) & = & \left\{y \in \mathbb{R}^n : \forall t_n \longrightarrow 0^+, \forall x^n \longrightarrow x^0 \text{ with}\right. \\ & & \left. x^n \in C, \exists y^n \longrightarrow y \text{ such that } x^n + t_n y^n \in C, \forall n \in \mathbb{N}\right\}. \end{array}$$

Obviously, if we can choose the *largest* convex subcone of $T(C, x^0)$, Theorem 13 will be sharper. Theorem 13 can be justified as follows.

If we consider (P_2), with $T(C, x^0)$ convex, e. g. because C is convex, we have obviously

$$x^0 \text{ minimum for } (\mathrm{P}_2) \Longrightarrow \nabla f(x^0) y \geq 0, \forall y \in T(S_2, x^0).$$

Now we can use the simple inclusion

$$T(C, x^0) \cap \left\{y \in \mathbb{R}^n : \nabla g_i(x^0) y < 0, \forall i \in I(x^0)\right\} \subset T(S_2, x^0)$$

which holds true, since the second set is open. So we get

$$\nabla f(x^0) y \geq 0, \forall y \in T(C, x^0) \cap \left\{y \in \mathbb{R}^n : \nabla g_i(x^0) y < 0, i \in I(x^0)\right\}$$

which is equivalent to

$$\{y : \nabla f(x^0)y < 0, \nabla g_i(x^0)y < 0, i \in I(x^0)\} \cap T(C, x^0) = \varnothing.$$

If the first set is empty, we get the classical Fritz John conditions by the Gordan theorem of the alternative (see, e.g., Mangasarian (1969)). If the first set is not empty, then, by the classical separation theorem (the first set is open!), we get the existence of non negative and not all zero multipliers such that

$$-\left[\lambda_0 \nabla f(x^0) + \sum_{i \in I(x^0)} \lambda_i \nabla g_i(x^0) \in T^*(C, x^0)\right].$$

If we take (P_1) into consideration we have in general (even assuming that C is convex)

$$\begin{aligned} & T(C, x^0) \cap \{y \in \mathbb{R}^n : \nabla g_i(x^0)y < 0, \forall i \in I(x^0), \\ \nabla h_j(x^0)y = 0, j \in J\} \not\subset\ & T(S_1, x^0), \end{aligned}$$

as the second set in this last relation is not open.

The results of Bazaraa and Goode (1972) and of Giorgi and Guerraggio (1994) for (P_1) have been subsequently generalized, with regard to a vector optimization problem, by Giorgi, Jimenez and Novo (2004). We report here their Theorem 4.1, referred to (P_1) and to the differentiability assumptions for (P_1).

Theorem 14. *Let x^0 be a local solution of* (P_1) *and let the assumptions of* Theorem 5 *be verified. Moreover, let $P \subset R^n$ be a convex set with $0 \in P$ and let the following regularity condition hold*

(RC) $$\ker \nabla h(x^0) \cap P \subset T(H \cap C, x^0),$$

where $H = \{x \in \mathbb{R}^n : h(x) = 0\}$.

Then there exists (u_0, u, v) such that $(u_0, u) \geq 0$, $(u_0, u, v) \neq 0$ and

$$-\left[u_0 \nabla f(x^0) + \textstyle\sum_{i=1}^m u_i \nabla g_i(x^0) + \sum_{j=1}^p v_j \nabla h_j(x^0)\right] \in P^*$$
$$u_i g_i(x^0) = 0, \forall i \in I.$$

Remark 15. We obtain the same conclusione for $P = I(C, x^0) \cup \{0\}$ or $P = Q(C, x^0) \cup \{0\}$ without assuming (RC). That is, $I(C, x^0)$ or $Q(C, x^0)$ are assumed to be convex, as in Theorem 10 or in its generalization by Giorgi and Guerraggio (1994).

Mangasarian (1969), assuming that C is closed and convex and $intC \neq \varnothing$, obtains for (P_1) the following "minimum principle type conditions":

$$\left[u_0 \nabla f(x^0) + \sum_{i=1}^{m} u_i \nabla g_i(x^0) + \sum_{j=1}^{p} v_j \nabla h_j(x^0)\right](x - x^0) \geq 0, \forall x \in C,$$

in addition to the usual sign conditions on the multipliers and the complementarity conditions. The same relations are obtained by Jahn (1994), always under the assumptions of C convex, $intC \neq \varnothing$, for a more general problem. This is a corollary of Theorem 10, as, if $intC \neq \varnothing$ (C convex), we have seen that $T(C, x^0) = clI(C, x^0) = clQ(C, x^0)$. In this case the polar of the Bouligand tangent cone $T(C, x^0)$ is the so-called *normal cone* to the convex set C at x^0:

$$N(C, x^0) = \left[T(C, x^0)\right]^* = \left[\mathrm{cl\,cone}(C - x^0)\right]^* = (C - x^0)^*.$$

If $intC = \varnothing$, then $I(C, x^0) = Q(C, x^0) = \varnothing$ and Theorem 10 holds trivially. On the other hand, some authors, e. g., Robinson (1976), Giorgi, Jimenez and Novo (2004), obtain the Mangasarian minimum principle type conditions merely under the convexity of C, without imposing $intC \neq \varnothing$. This result, however, seems to be not directly deducible from Theorem 10 or even from Theorem 14, unless $J = \varnothing$ (Theorem 13).

In a quite recent paper Bertsekas and Ozdaglar (2002) consider problem (P_1), with C nonempty closed set of $\mathbb{R}^n$ and with f, every g_i and every h_j continuously differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}$. They obtain "enhanced" Fritz John type conditions, where, instead of looking for a local cone approximation of C whose polar is suitable to get a minimum principle type condition, they use directly a generalization of the normal cone to C at x^0. See also Bertzekas, Ozdaglar and Tseng (2005). Similarly for what concerns the various tangent cones proposed in the literature, there are various definitions of normal cones. Here we refer to what may be called the *general normal cone* to C at $x^0 \in C$ or *Mordukhovich normal cone* to C at $x^0 \in C$ (see Mordukhovich (1976), Rockafellar (1993), Rockafellar and Wets (1998)). This cone, denoted $N_M(C, x^0)$, consists of all vectors $v \in \mathbb{R}^n$ for which there is a sequence of vectors $v^n \longrightarrow v$ and a sequence of points $x^n \longrightarrow x^0$ in C, such that, for each n,

$$v^n(x - x^n) \leq o(\|x - x^n\|), \forall x \in C.$$

Rockafellar and Wets (1998) prove that $v \in N_M(C, x^0)$ if and only if there exist sequences $\{x^n\} \subset C$ and $\{v^n\}$ such that $x^n \longrightarrow x^0, v^n \longrightarrow v$ and $v^n \in [T(C, x^n)]^*$ for all n. $N_M(C, x^0)$ is a closed cone, but in general not convex;

in general we have $[T(C,x^n)]^* \subset N_M(C,x^0)$, however, $N_M(C,x^0)$ may not be equal to $[T(C,x^n)]^*$. When C is convex we have

$$[T(C,x^n)]^* = N_M(C,x^0) = N(C,x^0).$$

Bertsekas and Ozdaglar (2002), besides other properties here not considered, obtain the following version of a generalized Fritz John Theorem.

Theorem 16. *Let x^0 be a local solution of* (P_1), *where f, every g_i and every h_j are continuously differentiable and C is a nonempty closed set of $\mathbb{R}^n$. Then there exist multipliers (u_0,u,v) such that $(u_0,u) \geq 0$, $(u_0,u,v) \neq 0$ and*

$$-\left[u_0\nabla f(x^0) + \textstyle\sum_{i=1}^m u_i\nabla g_i(x^0) + \sum_{j=1}^p v_j\nabla h_j(x^0)\right] \in N_M(C,x^0)$$
$$u_i g_i(x^0) = 0, \forall i \in I.$$

We note that, when C is convex, we obtain, as a corollary, the minimum prin- ciple type condition of Mangasarian, apart from being $intC \neq \varnothing$ or $intC = \varnothing$.

Remark 17. In a pioneering paper Clarke (1976) presented the following generalized Lagrange multiplier rule for (P_1), where f, every g_i and every h_j, are supposed to be locally Lipschitz at x^0 and C is a closed subset of $\mathbb{R}^n$: if x^0 solves (P_1) locally, then there exist numbers (u_0,u_i,v_j) $(i=1,...,m; j=1,...,p)$ not all zero, such that $(u_0,u) \geq 0$ and such that

$$0 \in \left[u_0\partial_{Cl}f(x^0) + \textstyle\sum_{i=1}^m u_i\partial_{Cl}g_i(x^0) + \right.$$
$$\left. + \textstyle\sum_{j=1}^p v_j\partial_{Cl}h_j(x^0) + N_{Cl}(C,x^0)\right],$$
$$u_i g_i(x^0) = 0, \forall i \in I.$$

Here $\partial_{Cl}f(x^0)$ is the Clarke subdifferential of f at x^0 :

$$\partial_{Cl}f(x^0) = \left\{\xi \in \mathbb{R}^n : (\xi,-1) \in N_{Cl}(epif,(x^0,f(x^0)))\right\}$$

and $N_{Cl}(C,x^0)$ is the Clarke normal cone of C at x^0 : $N_{Cl}(C,x^0) = \left[T_{Cl}(C,x^0)\right]^*$. It is known (see, e.g., Clarke (1983)) that if for each point in a neighborhood of x^0, f is continuously differentiable, then $\partial_{Cl}f(x^0) = \left\{\nabla f(x^0)\right\}$. So, in this case, we have another classical minimum type optimality condition for (P_1). Moreover, if C is also convex, again we obtain, as a corollary, the minimum principle type condition of Mangasarian, without imposing $intC \neq \varnothing$, being in this case,

$$T_{Cl}(C,x^0) = T(C,x^0) \text{ and } N(C,x^0) = N_{Cl}(C,x^0) = (C-x^0)^*.$$

References

[1] J.P. Aubin, H. Frankowska, *Set-Valued Analysis* , Birkhäuser, Boston (1990).

[2] M.S. Bazaraa, J.J. Goode, Necessary optimality criteria in mathematical programming in the presence of differentiability, *J. Math. Anal. Appl.*, **40** (1972), 609-621.

[3] M.S. Bazaraa, C.M. Shetty, *Foundations of Optimization*, Springer Verlag, Berlin (1976).

[4] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, *Nonlinear Programming. Theory and Algorithms*, John Wiley & Sons, New York (1993).

[5] D.P. Bertsekas, *Nonlinear Programming*, 2-nd Edition, Athena Scientific, Belmont, Mass (1999).

[6] D.P. Bertsekas, A.E. Ozdaglar, Pseudonormality and Lagrange multiplier theory for constrained optimization, *J. Optim. Theory Appl.*, **114** (2002), 287-343.

[7] D.P. Bertsekas, A.E. Ozdaglar, P. Tseng, Enhanced Fritz John conditions for convex programming, *SIAM J. Optim.*, **16** (2005), 766-797.

[8] S.I. Birbil, J.B. G. Frenk, G.J. Still, An elementary proof of the Fritz-John and Karush-Kuhn-Tucker conditions in nonlinear programming, *European J. of Operational Res.*, **180** (2007), 479-484.

[9] F.H. Clarke, A new approach to Lagrange multipliers, *Math. Oper. Res.*, **1** (1976), 165-174.

[10] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York (1983).

[11] G. Giorgi, A. Guerraggio, On the notion of tangent cone in mathematical programming, *Optimization*, **25** (1992), 11-23.

[12] G. Giorgi, A. Guerraggio, First order generalized optimality conditions for programming problems with a set constraint, In: *Generalized Convexity* (Ed-s: S. Komlosi, T. Rapcsàk and S. Schaible), Proceedings, Pécs, Hungary, 1992, Springer Verlag, Berlin (1994), 171-185.

[13] G. Giorgi, A. Guerraggio, Characterizations, comparisons, algebraic and topological properties of tangent cones, *Journal of Statistics and Management Systems*, **5** (2002), 275-294.

[14] G. Giorgi, A. Guerraggio, J Thierfelder, *Mathematics of Optimization. Smooth and Nonsmooth Case*, Elsevier, Amsterdam (2004).

[15] G. Giorgi, B. Jimenez, V. Novo, Minimum principle-type optimality conditions for Pareto problems, *Int. Journal of Pure and Applied Mathematics*, **10** (2004), 51-68.

[16] F.J. Gould, J. Tolle, A necessary and sufficient condition for constrained optimization, *SIAM Journal on Applied Mathematics*, **20** (1971), 164-172.

[17] M. Guignard, Generalized Kuhn-Tucker conditions for mathematical programming in a Banach space, *SIAM Journal on Control*, **7** (1969), 232-241.

[18] M.R. Hestenes, *Optimization Theory: The Finite-Dimensional Case*, John Wiley, New York (1975).

[19] J. Jahn, *Introduction to the Theory of Nonlinear Optimization*, Springer Verlag, Berlin (1994).

[20] O.L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York (1969).

[21] O.L. Mangasarian, S. Fromovitz, The Fritz John necessary optimality conditions in the presence of equality and inequality constraints, *J. Math. Anal. Appl.*, **17** (1967), 37-47.

[22] E.J. McShane, The Lagrange multiplier rule, *Amer. Math. Monthly*, **80** (1973), 922-925.

[23] B. Mordukhovich, Maximum principle in the problem of time optimal response with nonsmooth constraints, *Journal of Applied Mathematics and Mechanics*, **40** (1976), 960-969.

[24] B.H. Pourciau, Modern multiplier rules, *Amer. Math. Monthly*, **87** (1980), 433-452.

[25] B.H. Pourciau, On the Carathéodory-John multiplier rule, *J. Math. Anal. Appl.*, **87** (1982), 373-381.

[26] S.M. Robinson, First order conditions for general nonlinear optimization, *SIAM J. Appl. Math.*, **30** (1976), 597-607.

[27] R.T. Rockafellar, Lagrange multipliers and optimality, *SIAM Review*, **35** (1993), 183-238.

[28] R.T. Rockafellar, R.-J. Wets, *Variational Analysis*, Springer, Berlin (1998).

[29] G. Still, M. Streng, Optimality conditions in smooth nonlinear programming, *J. Optim. Theory Appl.*, **90** (1996), 483-515.

[30] O. Varaiya, Nonlinear programming in Banach spaces, *SIAM Journal on Applied Mathematics*, **15** (1967), 147-152.