SIAM J. CONTROL OPTIM.
Vol. 39, No. 5, pp. 1441–1460

MULTIPLIER RULES UNDER MIXED ASSUMPTIONS OF DIFFERENTIABILITY AND LIPSCHITZ CONTINUITY*

JANE J. YE†

Abstract. In this paper we study nonlinear programming problems with equality, inequality, and abstract constraints where some of the functions are Fréchet differentiable at the optimal solution, some of the functions are Lipschitz near the optimal solution, and the abstract constraint set may be nonconvex. We derive Fritz John type and Karush–Kuhn–Tucker (KKT) type first order necessary optimality conditions for the above problem where Fréchet derivatives are used for the differentiable functions and subdifferentials are used for the Lipschitz continuous functions. Constraint qualifications for the KKT type first order necessary optimality conditions to hold include the generalized Mangasarian–Fromovitz constraint qualification, the no nonzero abnormal multiplier constraint qualification, the metric regularity of the constraint region, and the calmness constraint qualification.

Key words. necessary optimality conditions, Fréchet differentiability, subdifferentials, constraint qualifications, metric regularity, calmness

AMS subject classifications. 49K10, 90C30

PII. S0363012999358476

1. Introduction.

The classical multiplier rule usually requires that the objective function and the inequality constraints be differentiable, the equality constraints be continuously differentiable at the optimal solution, and the abstract constraint set be convex with nonempty interior (e.g., see Bazaraa, Sherali, and Shetty [1] and Mangasarian [14]).

Over the last three decades, the classical multiplier rule was extended under two different assumptions: differentiability and Lipschitz continuity.

On the one hand, the classical multiplier rule was extended in the direction of eliminating the smoothness assumption while keeping the differentiability assumption. In the case where there is no abstract constraint, based on a correction theorem, Halkin [9] proved that the classical multiplier rule holds under the weaker assumption which requires only that the equality constraints be Fréchet differentiable at the optimal solution and continuous in a neighborhood of the optimal solution. Based on a multidimensional intermediate value theorem, Di [7] derived some first order and second order multiplier rules for nonlinear programming problems with equality, inequality, and abstract constraints where all functions are Fréchet differentiable at the optimal solution and continuous in a neighborhood of the optimal solution and the abstract constraint set is convex.

On the other hand, in nonsmooth analysis the classical multiplier rule was generalized in the direction of replacing the differentiability assumption by the Lipschitz continuity assumption. Under the assumption that all functions are Lipschitz near the optimal solution and the abstract constraint set is closed, Clarke [3] derived a generalized multiplier rule involving the Clarke generalized gradient and the Clarke normal

*Received by the editors July 6, 1999; accepted for publication (in revised form) August 23, 2000; published electronically January 19, 2001. The research of this paper was partially supported by NSERC and the University of Victoria.
http://www.siam.org/journals/sicon/39-5/35847.html
†Department of Mathematics and Statistics, University of Victoria, Victoria BC, Canada V8W 3P4 (janeye@Math.UVic.CA).

cone. The Clarke generalized gradient of a function would reduce to the usual derivative only when the function is strictly differentiable (for example, when the function is continuously differentiable). Hence, when all functions involved are continuously differentiable and the abstract constraint set is convex, the generalized multiplier rule of Clarke would recover the classical multiplier rule. However, the Clarke generalized gradient of a Lipschitz continuous function may be strictly larger than the set which consists of the usual derivative when the function is Fréchet differentiable but not strictly differentiable. In the case when the abstract set is convex, Ioffe [11] showed that the Clarke generalized multiplier rule can be sharpened by replacing the Clarke generalized gradient by the Michel–Penot subdifferential which coincides with the usual derivative when the function is Gâteaux differentiable. Other results in this direction also include Mordukhovich's combined multiplier rule [16] and the Treiman's multiplier rule [19].

In this paper we study first order necessary optimality conditions for nonlinear programming problems with equality, inequality, and abstract constraints where some of the functions are Fréchet differentiable at the optimal solution, some of the functions are Lipschitz near the optimal solution, and the abstract constraint set may be nonconvex. For the above nonlinear programming problem with mixed assumptions on differentiability and Lipschitz continuity, since a differentiable function may not be Lipschitz continuous, the only applicable necessary optimality conditions in the literature are fuzzy multiplier rules (see, e.g., Borwein and Zhu [2]). Although in a finite dimensional space the fuzzy multiplier rule reduces to an exact multiplier rule, it involves the singular subdifferential of the non-Lipschitz functions. Our purpose is to derive exact (i.e., nonfuzzy) first order multipler rules which do not involve any singular subdifferentials for the above problem where Fréchet derivatives are used for the differentiable functions and subdifferentials are used for the Lipschitz continuous functions.

To be more precise, we consider the following optimization problem:

$$
\begin{array}{lll}
\text{(P)} & \text{minimize} & f(x) \\
 & \text{subject to} & g_i(x) \le 0, \quad i = 1, 2, \ldots, I, \\
 & & h_j(x) = 0, \quad j = 1, 2, \ldots, J, \\
 & & \phi_k(x) \le 0, \quad k = 1, 2, \ldots, K, \\
 & & \psi_l(x) = 0, \quad l = 1, 2, \ldots, L, \\
 & & x \in \Omega,
\end{array}
$$

where $f, g_i(i = 1, 2, \ldots, I), h_i(j = 1, 2, \ldots, J), \phi_k(k = 1, 2, \ldots, K), \psi_l(l = 1, 2, \ldots, L)$ are the objective function and the constraint functions from a Banach space X to R. Ω is a closed subset of X and I, J, K, L are given integers. Generally one has $I \ge 1, J \ge 1, K \ge 1, L \ge 1$, but we allow I, J, K, or $L = 0$ to signify the case in which there are no explicit constraints of the type.

Let $\bar{x}$ be a local optimal solution to (P). Denote by $I(\bar{x}) := \{i : g_i(\bar{x}) = 0\}$ and $K(\bar{x}) := \{k : \phi_k(\bar{x}) = 0\}$ the index sets of the binding constraints. We always make the following basic assumptions on the constraint functions.

(A) $g_i(i \in I(\bar{x})), h_j(j = 1, 2, \ldots, J)$ are Fréchet differentiable at $\bar{x}$ and $g_i(i \notin I(\bar{x}))$ are continuous at $\bar{x}$. $\phi_k(k \in K(\bar{x})), \psi_l(l = 1, 2, \ldots, L)$ are Lipschitz near $\bar{x}$ and $\phi_k(k \notin K(\bar{x}))$ are continuous at $\bar{x}$.

Our main results include the following multiplier rules.

THEOREM 1.1 (Fritz John necessary optimality conditions for the case $L = 0$).

Let $\bar{x}$ be a local optimal solution of (P) *with $L = 0$. Suppose that f is either Fréchet differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, in addition to assumption* (A), *$h_j(j = 1, 2, \ldots, J)$ are continuous in a neighborhood of $\bar{x}$, and there exists a vector that is hypertangent* (*see Definition* 2.4) *to the abstract constraint set Ω at $\bar{x}$. Then there exist scalars $\lambda \geq 0, \alpha_i \geq 0(i \in I(\bar{x})), \beta_j(j = 1, 2, \ldots, J), \gamma_k \geq 0(k \in K(\bar{x}))$ not all zero such that*

$$0 \in \lambda \partial^{\diamond} f(\bar{x}) + \sum_{i \in I(\bar{x})} \alpha_i \nabla g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) + \sum_{k \in K(\bar{x})} \gamma_k \partial^{\diamond} \phi_k(\bar{x}) + N(\bar{x}, \Omega),$$

where $\partial^{\diamond}$ denotes the Michel–Penot subdifferential, ∇ denotes the Fréchet derivative, and $N(\bar{x}, \Omega)$ denotes the Clarke normal cone to Ω at $\bar{x}$.

Remark 1. Note that in the case where f is Fréchet differentiable at $\bar{x}$, $\partial^{\diamond} f(\bar{x}) = \{\nabla f(\bar{x})\}$ in the above multiplier rule. As it was shown by Fernandez [8], the continuity assumption of the equality constraints h_j in Theorem 1.1 cannot be removed.

THEOREM 1.2 (Fritz John necessary optimality conditions for the case $I = J = 0$). *Let $\bar{x}$ be a local optimal solution of* (P) *with $I = J = 0$. Suppose that the objective function f is Fréchet differentiable at $\bar{x}$, $\phi_k(k \in K(\bar{x})), \psi_l(l = 1, 2, \ldots, L)$ are Lipschitz near $\bar{x}$ and $\phi_k(k \notin K(\bar{x}))$ are continuous at $\bar{x}$. Then there exist scalars $\lambda \geq 0, \gamma_k \geq 0(k \in K(\bar{x})), \eta_l(l = 1, 2, \ldots, L)$ not all zero such that*

$$0 \in \lambda \nabla f(\bar{x}) + \sum_{k \in K(\bar{x})} \gamma_k \partial \phi_k(\bar{x}) + \sum_{l=1}^{L} \eta_l \partial \psi_l(\bar{x}) + N(\bar{x}, \Omega),$$

where ∂ denotes the Clarke generalized gradient and $N(x, \Omega)$ denotes the Clarke normal cone to Ω at $\bar{x}$. Moreover, if X is an Asplund space which is a Banach space whose separable subspaces have separable duals (as is the case for reflexive spaces), under the above assumptions, there exist scalars $\lambda \geq 0, \gamma_k \geq 0(k \in K(\bar{x})), \eta_l(l = 1, 2, \ldots, L)$ not all zero such that

$$0 \in \lambda \nabla f(\bar{x}) + \sum_{k \in K(\bar{x})} \gamma_k \hat{\partial} \phi_k(\bar{x}) + \hat{\partial} \left(\sum_{l=1}^{L} \eta_l \psi_l \right) (\bar{x}) + \hat{N}(\bar{x}, \Omega),$$

where $\hat{\partial}$ denotes the limiting subdifferential and $\hat{N}(\bar{x}, \Omega)$ denotes the limiting normal cone to Ω at $\bar{x}$.

As in smooth and Lipschitz optimization we also give constraint qualifications under which the scalar λ in the above theorems is nonzero such as the generalized Mangasarian–Fromovitz constraint qualification (GMFCQ), the no nonzero abnormal multiplier constraint qualification (NNAMCQ), the metric regularity of the constraint region (metric regularity CQ), and the calmness constraint qualification (calmness CQ).

We organize the paper as follows. In the next section we provide preliminaries that will be used in the paper. In section 3, we prove Theorem 1.1, the Fritz John type necessary optimality condition for the case where there are no Lipschitz continuous equality constraints. In section 4, we introduce constraint qualifications, discuss the relationship between the (GMFCQ) and (NNAMCQ), and prove that under constraint qualifications such as the (NNAMCQ), the metric regularity CQ and the calmness CQ, λ in Theorems 1.1 can be taken as 1. An example is given to show that when the objective function is not Lipschitz but only Fréchet differentiable, the metric regularity CQ may not imply the calmness CQ. Hence the well-known relationships between these constraint qualifications may not hold when some of the functions are not Lipschitz but Fréchet differentiable. However, it turns out that the Karush–Kuhn–Tucker (KKT) conditions can usually be derived directly. We prove that unlike the Fritz John type condition (Theorem 1.1), under the metric regularity CQ and the calmness CQ the KKT condition holds even in the case where $L \neq 0$, and the continuity assumption of the Fréchet differentiable equality constraints is not needed. In section 5, we derive KKT type necessary optimality conditions for the case where all constraint functions are Lipschitz continuous and the objective function is Fréchet differentiable under the constraint qualification (NNAMCQ), the metric regularity CQ, and the calmness CQ. Theorem 1.2, the Fritz John type necessary optimality condition, then follows as an easy consequence.

2. Preliminaries. This section contains some background material on nonsmooth analysis which will be used throughout the paper. We give only concise definitions that will be needed in the paper. For more detailed information on the subject, our references are Clarke [4], Clarke, Ledyaev, Stern, and Wolenski [5], Loewen [13], and Mordukhovich and Shao [17].

We first give the following notations that will be used throughout the paper. For a vector $v \in R^n$, v_i is the ith components of v. For any Banach space X we denote its norm by $\|\cdot\|$ and consider the dual space X^* equipped with the weak-star topology w^*, where $\langle\cdot,\cdot\rangle$ means the canonic pairing. As usual, B and B^* stand for the open unit balls in the space and the dual space in question. Note that intΩ, clΩ, and $co\Omega$ mean, respectively, the interior, the closure, and the convex hull of an arbitrary nonempty set $\Omega \subset X$, while the notation cl^* is used for the weak-star topological closure in X^*.

For a set-valued map $\Phi : X \rightrightarrows X^*$, we denote by

$$\limsup_{x\to\bar{x}} \Phi(x)$$

the sequential Kuratowski–Painlevé upper limit with respect to the norm topology in X and the weak-star topology in X^*, i.e.,

$$\limsup_{x\to\bar{x}} \Phi(x) := \{x^* \in X^* | \exists \text{ sequences } x_k \to \bar{x}, x_k^* \xrightarrow{w^*} x^*, \text{ with } x_k^* \in \Phi(x_k) \forall k = 1, 2, \ldots\}.$$

We now give some concepts for various normal cones.

DEFINITION 2.1. *Let Ω be a nonempty subset of a Banach space X and let $\epsilon \geq 0$.*

(i) *Given $x \in$ clΩ, the set*

$$N_\epsilon^F(x,\Omega) := \left\{ x^* \in X^* | \limsup_{x'\to x, x'\in\Omega} \frac{\langle x^*, x'-x\rangle}{\|x'-x\|} \leq \epsilon \right\} \tag{2.1}$$

is called the set of Fréchet ϵ-normals to Ω at x. When $\epsilon = 0$, the set (2.1) is a cone which is called the Fréchet normal cone to Ω at x and is denoted by $N^F(x,\Omega)$.

(ii) *Let $\bar{x} \in$ clΩ. The nonempty cone*

$$\hat{N}(\bar{x},\Omega) := \limsup_{x\to\bar{x}, \epsilon\downarrow 0} N_\epsilon^F(x,\Omega) \tag{2.2}$$

is called the limiting normal cone to Ω at $\bar{x}$.

Using the definitions for normal cones, we now give definitions for corresponding subdifferentials of a single-valued map.

DEFINITION 2.2. *Let X be a Banach space and $\varphi : X \to R \cup \{+\infty\}$ be l.s.c. (lower semicontinuous) and finite at $x \in X$. The sets*

$$\partial^F_\epsilon \varphi(x) := \{x^* \in X^* | (x^*, -1) \in N^F_\epsilon((x, \varphi(x)), epi\varphi)\}, \tag{2.3}$$
$$\hat{\partial}\varphi(x) := \{x^* \in X^* | (x^*, -1) \in \hat{N}((x, \varphi(x)), epi\varphi)\},$$

where $epi\varphi := \{(x, v) : v \geq \varphi(x)\}$ denotes the epigraph of φ, are called, respectively, the Fréchet ϵ-subdifferential and the limiting subdifferential of φ at x. When $\epsilon = 0$, the set (2.3) *is called the Fréchet subdifferential of φ at x and is denoted by $\partial^F \varphi(x)$.* It is known that the Fréchet subdifferential has the following analytic expression:

$$\partial^F \varphi(x) = \left\{ x^* \in X^* | \liminf_{x' \to x} \frac{\varphi(x') - \varphi(x) - \langle x^*, x' - x\rangle}{\|x' - x\|} \geq 0 \right\}. \tag{2.4}$$

Let X be any Banach space, $\bar{x} \in X$, and $\varphi : X \to R$ be any continuous function. Then the Michel–Penot directional derivative of φ at $\bar{x}$ in the direction $v \in X$ introduced in [15] is given by

$$\varphi^\square(\bar{x}; v) := \sup_{w \in X} \limsup_{t \downarrow 0} \frac{\varphi(\bar{x} + t(v + w)) - \varphi(\bar{x} + tw)}{t}$$

and the Michel–Penot subdifferential of φ at $\bar{x}$ is given by the set

$$\partial^\diamond \varphi(\bar{x}) := \{x^* \in X^* | \langle x^*, v\rangle \leq \varphi^\square(\bar{x}; v) \ \forall v \in X\}.$$

It is known (see [15, Proposition 1.3]) that when a function φ is Gâteaux differentiable at $\bar{x}$, $\partial^\diamond \varphi(\bar{x}) = \{\nabla \varphi(\bar{x})\}$.

The following properties of the Michel–Penot directional derivatives and the Michel–Penot subdifferentials will be useful.

PROPOSITION 2.3 (see [15, 11]). *Let X be a Banach space, $x \in X$, and f be Lipschitz near x with constant L_f. Then*

(i) *The function $v \to f^\square(x; v)$ is finite, positively homogeneous, and subadditive on X.*
(ii) *As a function of v, $f^\square(x; v)$ is Lipschitz continuous with constant L_f on X.*
(iii) *$\partial^\diamond f(x)$ is a nonempty, convex, weak*-compact subset of X^* and $\|x^*\| \leq L_f$ for every $x^* \in \partial^\diamond f(x)$.*

Let X be any Banach space, $\bar{x} \in X$, and $\varphi : X \to R$ be Lipschitz near $\bar{x}$. Then the Clarke generalized derivative of φ at $\bar{x}$ in the direction $v \in X$ is given by

$$\varphi^0(\bar{x}; v) := \limsup_{x \to \bar{x}, t \downarrow 0} \frac{\varphi(x + tv) - \varphi(x)}{t}$$

and the Clarke generalized gradient of φ at $\bar{x}$ is given by the set

$$\partial \varphi(\bar{x}) := \{x^* \in X^* | \langle x^*, v\rangle \leq \varphi^0(\bar{x}; v) \ \forall v \in X\}.$$

Let Ω be a nonempty subset of a Banach space X and consider its distance function, that is, the function $d_\Omega(\cdot) : X \to R$ defined by

$$d_\Omega(x) = \inf\{\|x - c\| : c \in \Omega\}.$$

The Clarke tangent cone to Ω at $\bar{x}$ is defined by

$$T(\bar{x},\Omega) := \{v \in X|\ d^0_\Omega(\bar{x};v) = 0\}$$

and the Clarke normal cone to Ω at $\bar{x}$ is defined by polarity with $T(\bar{x},\Omega)$:

$$N(\bar{x},\Omega) := \{x^* \in X^* | \langle x^*, v\rangle \le 0 \quad \forall v \in T(\bar{x},\Omega)\}.$$

Definition 2.4 (hypertangent). *Let X be a Banach space. A vector v in X is said to be hypertangent to the set $\Omega \subseteq X$ at the point $x \in \Omega$ if for some $\epsilon > 0$,*

$$y + tw \in \Omega \quad \forall y \in (x + \epsilon B) \cap \Omega, w \in v + \epsilon B, t \in (0, \epsilon).$$

It follows easily that any vector v hypertangent to Ω at x belongs to $T(x,\Omega)$. It is possible to have no hypertangents at all. However, it is clear that when Ω is a convex set with nonempty interior, then any vector $x^* - x$ with $x^* \in \mathrm{int}\Omega$ is hypertangent to Ω at x.

It is known that in any Banach space X and for any $\epsilon \ge 0$

$$N^F_\epsilon(\bar{x},\Omega) \subseteq \hat{N}(\bar{x},\Omega) \subseteq N(\bar{x},\Omega),$$
$$\partial^F_\epsilon \varphi(\bar{x}) \subseteq \hat{\partial}\varphi(\bar{x}) \subseteq \partial^\diamond \varphi(\bar{x}) \subseteq \partial\varphi(\bar{x})$$

and in any Asplund space, the following precise relationships hold [17, Theorems 2.9 and 8.11]:

(i) For any closed set $\Omega \subseteq X$ and $\bar{x} \in \Omega$ one has

$$\hat{N}(\bar{x},\Omega) = \limsup_{x\to\bar{x}} N^F(x,\Omega),$$
$$N(\bar{x};\Omega) = \mathrm{cl}^* co \hat{N}(\bar{x},\Omega).$$

(ii) For any function $\varphi : X \to R$ which is Lipschitz near $\bar{x} \in X$, one has

$$\hat{\partial}\varphi(\bar{x}) = \limsup_{x\to\bar{x}} \partial^F\varphi(x),$$
$$\partial\varphi(\bar{x}) = \mathrm{cl}^* co \hat{\partial}\varphi(\bar{x}).$$

We now summarize the sum rules and chain rules for the various subdifferentials in the literature. For convenience, we do not intend to quote the results under the most general assumptions. Instead, we provide the results under the assumptions we need in our paper. For example, since when Y is finite dimensional, a function $\varphi : X \to Y$ is Lipschitz near $\bar{x} \in X$ is strictly Lipschitzian at $\bar{x}$ in the sense of [17]; Propositions 2.5(ii) and 2.6(ii) are special cases of the results in [17].

Proposition 2.5 (sum rules).

(i) (*See, e.g., the proof of* [6, Lemma 2.2].) *Let X be a Banach space and $\bar{x} \in X$. Let $\varphi_1 : X \to R$ be Fréchet differentiable at $\bar{x}$ and $\varphi_2 \to R \cup \{+\infty\}$ be finite and l.s.c. at $\bar{x}$. Then*

$$\partial^F(\varphi_1 + \varphi_2)(\bar{x}) = \nabla\varphi_1(\bar{x}) + \partial^F\varphi_2(\bar{x}).$$

(ii) (*See* [17, Proposition 2.5 and Theorem 4.1].) *Let X be an Asplund space and $\bar{x} \in X$. Let $\varphi_i : X \to R \cup \{+\infty\}, i = 1, 2,$ be l.s.c. at $\bar{x}$ and one of these functions is Lipschitz near $\bar{x}$. Then one has*

$$\hat{\partial}(\varphi_1 + \varphi_2)(\bar{x}) \subseteq \hat{\partial}\varphi_1(\bar{x}) + \hat{\partial}\varphi_2(\bar{x}).$$

(iii) (*See* [4, Proposition 2.3.3].) *Let X be a Banach space and $\bar{x} \in X$. Let $\varphi_i : X \to R, i = 1, 2,$ be Lipschitz near $\bar{x}$. Then one has*

$$\partial(\varphi_1 + \varphi_2)(\bar{x}) \subseteq \partial\varphi_1(\bar{x}) + \partial\varphi_2(\bar{x}).$$

PROPOSITION 2.6 (chain rules).

(i) (*See* [5, Theorem 2.5].) *Let X be a Banach space and $\bar{x} \in X$. Suppose that $\varphi : X \to R^n$ is Lipschitz near $\bar{x}$ and $f : R^n \to R$ is Lipschitz near $\varphi(\bar{x})$. Then*

$$\partial(f \circ \varphi)(\bar{x}) \subseteq \mathrm{cl}^* co \cup_{y^* \in \partial f(\varphi(\bar{x}))} \partial\langle y^*, \varphi\rangle(\bar{x}).$$

(ii) (*See* [17, Proposition 2.5 and Corollary 6.3].) *Moreover, if X is an Asplund space, then*

$$\hat{\partial}(f \circ \varphi)(\bar{x}) \subseteq \cup_{y^* \in \hat{\partial} f(\varphi(\bar{x}))} \hat{\partial}\langle y^*, \varphi\rangle(\bar{x}).$$

The following exact penalty results given by Clarke in [4, Proposition 2.4.3] will often be used in the paper.

PROPOSITION 2.7. *Let C be a closed subset of X. Assume that f attains a minimum over C at $\bar{x} \in C$ and f is Lipschitz near $\bar{x}$ with constant $L_f > 0$. Then for all $K \geq L_f$, the function $g(y) = f(y) + K d_C(y)$ also attains a minimum over X at $\bar{x}$.*

3. Proof of Theorem 1.1. We need only to prove the theorem under the assumption that there do not exist scalars $\alpha_i \geq 0 (i \in I(\bar{x})), \beta_j (j = 1, 2, \dots, J), \gamma_k \geq 0 (k \in K(\bar{x}))$ not all zero such that

$$0 \in \sum_{i \in I(\bar{x})} \alpha_i \nabla g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) + \sum_{k \in K(\bar{x})} \gamma_k \partial^{\diamond} \phi_k(\bar{x}) + N(\bar{x}, \Omega). \tag{3.1}$$

Indeed, if (3.1) is satisfied by some scalars $\alpha_i \geq 0 (i \in I(\bar{x})), \beta_j (j = 1, 2, \dots, J), \gamma_k \geq 0 (k \in K(\bar{x}))$ that are not all zero, then by taking $\lambda = 0$ we obtain the Fritz John condition.

Case 1, $J \neq 0$.

Since $\nabla h_j(\bar{x}) (j = 1, 2, \dots, J)$ are linearly independent by assumption (3.1), by the correction theorem of Halkin [9, Theorem F], there exist a neighborhood U of $\bar{x}$ and a continuous mapping ζ from U into X such that $\zeta(\bar{x}) = 0, \nabla\zeta(\bar{x}) = 0$ and

$$h_j(x + \zeta(x)) = \langle \nabla h_j(\bar{x}), x - \bar{x}\rangle \quad \forall x \in U, \; j = 1, 2, \dots, J. \tag{3.2}$$

We shall now prove that there is no $v \in \mathrm{int}T(\bar{x}, \Omega)$ such that

$$f^{\square}(\bar{x}; v) < 0, \tag{3.3}$$

$$\langle \nabla g_i(\bar{x}), v\rangle < 0, \quad i \in I(\bar{x}), \tag{3.4}$$

$$\langle \nabla h_j(\bar{x}), v\rangle = 0, \quad j = 1, 2, \dots, J, \tag{3.5}$$

$$\phi_k^{\square}(\bar{x}; v) < 0, \quad k \in K(\bar{x}). \tag{3.6}$$

By contradiction, we assume that there exists $v^* \in \mathrm{int}T(\bar{x}, \Omega)$ such that (3.3)–(3.6) hold. Let

$$\theta(t) = \bar{x} + tv^* + \zeta(\bar{x} + tv^*) \;\; \forall t \in [0, 1].$$

Then by virtue of (3.2) and (3.5), for all $\tau \in (0,1]$ small enough, $h_j(\theta(\tau)) = 0$ for all $j = 1, 2, \ldots, J$. Since $\theta(0) = \bar{x}, \nabla\theta(0) = v^*$, by the chain rule,

$$\lim_{t\to 0^+} \frac{g_i(\theta(t)) - g_i(\theta(0))}{t} = \langle \nabla g_i(\bar{x}), v^* \rangle \quad \forall i \in I(\bar{x}).$$

Consequently, by virtue of (3.4),

$$\lim_{t\to 0^+} \frac{g_i(\theta(t)) - g_i(\theta(0))}{t} < 0 \quad \forall i \in I(\bar{x}).$$

That is, for all $\tau \in (0,1]$ small enough,

$$g_i(\theta(\tau)) < 0 \quad \forall i \in I(\bar{x}).$$

Since ϕ_k is Lipschitz near $\bar{x}$,

$$\phi_k^{\square}(\bar{x}; v^*) = \sup_{w\in X} \limsup_{v'\to v^*, t\downarrow 0} \frac{\phi_k(\bar{x} + t(v' + w)) - \phi_k(\bar{x} + tw)}{t} \quad \forall k \in K(\bar{x}).$$

Consequently, by virtue of (3.6), we have for all $\tau \in (0,1]$ small enough,

$$\frac{\phi_k(\bar{x} + \tau v^* + \zeta(\bar{x} + \tau v^*)) - \phi_k(\bar{x})}{\tau} < 0 \quad \forall k \in K(\bar{x}).$$

That is, for all $\tau \in (0,1]$ small enough,

$$\phi_k(\theta(\tau)) < 0 \quad \forall k \in K(\bar{x}).$$

Similarly since $f^{\square}(\bar{x}, v) = \langle \nabla f(\bar{x}), v \rangle$ when f is Fréchet differentiable, for all $\tau \in (0,1]$ small enough, $f(\theta(\tau)) < f(\bar{x})$ by virtue of (3.3). By assumption, there exists a hypertangent to Ω at $\bar{x}$. By Rockafellar (see [4, Theorem 2.4.8]), the set of all hypertangents to Ω at $\bar{x}$ coincides with the interior of the Clarke tangent cone to Ω at $\bar{x}$. So for all $\tau \in (0,1]$ small enough,

$$\bar{x} + tv^* + \zeta(\bar{x} + \tau v^*) = \bar{x} + \tau \left[v^* + \frac{\zeta(\bar{x} + \tau v^*)}{\tau} \right] \in \Omega.$$

By the continuity assumptions at $\bar{x}$ for $g_i (i \notin I(\bar{x})), \phi_k(k \notin K(\bar{x}))$, for all $\tau \in (0,1]$ small enough,

$$g_i(\theta(\tau)) < 0 \quad \forall i \notin I(\bar{x}),$$
$$\phi_k(\theta(\tau)) < 0 \quad \forall k \notin K(\bar{x}).$$

Hence there exists $\tau \in (0,1]$ such that

$$\begin{aligned} & f(\theta(\tau)) < f(\bar{x}), \\ & g_i(\theta(\tau)) < 0, \qquad i = 1, 2, \ldots, I, \\ & h_j(\theta(\tau)) = 0, \qquad j = 1, 2, \ldots, J, \\ & \phi_k(\theta(\tau)) < 0, \qquad k = 1, 2, \ldots, K, \\ & \theta(\tau) \in \Omega, \end{aligned}$$

which contradicts the fact that $\bar{x}$ is a local optimal solution of (P).

Since $T(\bar{x},\Omega)$ is a closed convex cone and $f^{\square}(\bar{x};v), \phi_k^{\square}(\bar{x};v)$ are continuous in v (see Proposition 2.3), by virtue of nonexistence of $v \in \mathrm{int}T(\bar{x},\Omega)$ satisfying (3.3)–(3.6), the nonemptyness of $\mathrm{int}T(\bar{x},\Omega)$, and Proposition 4.4, $v = 0$ is a solution to the following problem:

$$\begin{aligned} \min \quad & f^{\square}(\bar{x};v) \\ \text{s.t.} \quad & \langle \nabla g_i(\bar{x}), v\rangle \le 0, \quad i \in I(\bar{x}), \\ & \langle \nabla h_j(\bar{x}), v\rangle = 0, \quad j = 1,2,\ldots,J, \\ & \phi_k^{\square}(\bar{x};v) \le 0, \quad k \in K(\bar{x}), \\ & v \in T(\bar{x},\Omega). \end{aligned}$$

Applying the generalized multiplier rule of Clarke [4, Theorem 6.1.1], there exist scalars $\lambda \ge 0, \alpha_i \ge 0 (i \in I(\bar{x})), \beta_j (j = 1,2,\ldots,J), \gamma_k \ge 0 (k \in K(\bar{x}))$ not all zero such that

$$\begin{aligned} 0 \in \lambda \partial_v f^{\square}(\bar{x};0) &+ \sum_{i \in I(\bar{x})} \alpha_i \nabla g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) \\ &+ \sum_{k \in K(\bar{x})} \gamma_k \partial_v \phi_k^{\square}(\bar{x};0) + N(0, T(\bar{x},\Omega)), \end{aligned}$$

where ∂_v denotes the generalized gradient with respect to v.

By definition, $\xi \in \partial^{\diamond} f(\bar{x})$ if and only if

$$\langle \xi, v\rangle \le f^{\square}(\bar{x};v) \;\forall v \in X. \tag{3.7}$$

Since $f^{\square}(\bar{x};v)$ is a convex function of v (see Proposition 2.3) and obviously $f^{\square}(\bar{x};0) = 0$, (3.7) holds if and only if $\xi \in \partial_v f^{\square}(\bar{x};0)$. Hence, $\partial_v f^{\square}(\bar{x};0) = \partial^{\diamond} f(\bar{x})$. Similarly, $\partial_v \phi_k^{\square}(\bar{x};0) = \partial^{\diamond} \phi_k(\bar{x})$. Since $\xi \in N(\bar{x},\Omega)$ if and only if $\langle \xi, v\rangle \le 0$ for all $v \in T(\bar{x},\Omega)$,

$$N(0, T(\bar{x},\Omega)) = N(\bar{x},\Omega).$$

Hence the Fritz John condition holds in this case.

Case 2, $J = 0, I \ne 0$.

We shall now prove that there is no $v \in \mathrm{int}T_{\Omega}(\bar{x})$ such that

$$f^{\square}(\bar{x};v) < 0, \tag{3.8}$$

$$\langle \nabla g_i(\bar{x}), v\rangle < 0, \quad i \in I(\bar{x}), \tag{3.9}$$

$$\phi_k^{\square}(\bar{x};v) < 0, \quad k \in K(\bar{x}). \tag{3.10}$$

By contradiction, we assume that there exists $v^* \in \mathrm{int}T(\bar{x},\Omega)$ such that (3.8)–(3.10) hold. Since $g_i, i \in I(\bar{x})$ are differentiable at $\bar{x}$, for $t > 0$ small enough,

$$g_i(\bar{x} + tv^*) = g_i(\bar{x}) + t\langle \nabla g_i(\bar{x}), v^*\rangle + \alpha_i(\bar{x}, tv^*)t\|v^*\| \quad \forall i \in I(\bar{x}),$$

where $\lim_{t\to 0} \alpha_i(\bar{x}, tv^*) = 0$ for $i \in I(\bar{x})$.

By virtue of (3.9), for $\tau > 0$ small enough,

$$\langle \nabla g_i(\bar{x}), v^*\rangle + \alpha_i(\bar{x}, \tau v^*)\|v^*\| < 0$$

and hence for $\tau > 0$ small enough,

$$g_i(\bar{x} + \tau v^*) < 0, \quad i = 1,2,\ldots,I.$$

By virtue of (3.10), we have for all $\tau \in (0,1]$ small enough,

$$\frac{\phi_k(\bar{x}+\tau v^*) - \phi_k(\bar{x})}{\tau} < 0 \quad \forall k \in K(\bar{x}).$$

That is, for all $\tau \in (0,1]$ small enough,

$$\phi_k(\bar{x}+\tau v^*) < 0, \quad k = 1,2,\ldots,K.$$

Similarly, we can prove that for all τ small enough,

$$f(\bar{x}+\tau v^*) < f(\bar{x}).$$

Since v^* is a hypertangent to Ω at $\bar{x}$, $\bar{x}+\tau v^* \in \Omega$ for $\tau > 0$ small enough. Hence there exists $\tau > 0$ such that

$$\begin{aligned}
&f(\bar{x}+\tau v^*) < f(\bar{x}),\\
&g_i(\bar{x}+\tau v^*) < 0, \quad i = 1,2,\ldots,I,\\
&\phi_k(\bar{x}+\tau v^*) < 0, \quad k = 1,2,\ldots,K,\\
&\bar{x}+\tau v^* \in \Omega,
\end{aligned}$$

which contradicts the fact that $\bar{x}$ is a local optimal solution of (P).

The remaining proof is similar to Case 1.

4. Constraint qualifications and the KKT conditions. In this section we introduce four constraint qualifications which ensure the KKT conditions hold and discuss the relationships among them.

The first constraint qualification for the case $L = 0$ follows naturally from the Fritz John necessary optimality condition as in the following proposition.

THEOREM 4.1 (KKT condition for the case $L = 0$ under the (NNAMCQ)). *In addition to the assumptions of Theorem* 1.1*, assume that there is no nonzero abnormal multiplier, i.e.,*

$$\begin{aligned}
(4.1)\qquad &0 \in \sum_{i\in I(\bar{x})} \alpha_i \nabla g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) + \sum_{k\in K(\bar{x})} \gamma_k \partial^\diamond \phi_k(\bar{x}) + N(\bar{x},\Omega),\\
&\alpha_i \geq 0,\ i \in I(\bar{x}),
\end{aligned}$$

implies that $\alpha_i = 0$ for all $i \in I(\bar{x}), \beta_j = 0$ for all $j = 1,2,\ldots,J, \gamma_k = 0$ for all $k \in K(\bar{x})$. Then $\lambda > 0$ in the conclusion of Theorem 1.1.

Proof. By Theorem 1.1, there exist scalars $\lambda \geq 0, \alpha_i \geq 0 (i \in I(\bar{x})), \beta_j (j = 1,2,\ldots,J), \gamma_k \geq 0 (k \in K(\bar{x}))$ not all zero such that

$$(4.2)\quad 0 \in \lambda \partial^\diamond f(\bar{x}) + \sum_{i\in I(\bar{x})} \alpha_i \nabla g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) + \sum_{k\in K(\bar{x})} \gamma_k \partial^\diamond \phi_k(\bar{x}) + N(\bar{x},\Omega).$$

The case $\lambda = 0$ is impossible. Indeed, if $\lambda = 0$ in the above condition, then the inclusion (4.2) coincides with inclusion (4.1). The assumption then rules out this possibility. ◻

Motivated by the above KKT condition we define the following constraint qualification for the general problem (P).

DEFINITION 4.2. *We say that* (P) *satisfies the* (NNAMCQ) *if*

$$0 \in \sum_{i \in I(\bar{x})} \alpha_i \nabla g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) + \sum_{k \in K(\bar{x})} \gamma_k \partial^{\diamond} \phi_k(\bar{x}) + \sum_{l=1}^{L} \eta_l \partial^{\diamond} \psi_l(\bar{x}) + N(\bar{x}, \Omega),$$
$$\alpha_i \geq 0,\ i \in I(\bar{x}), \gamma_k \geq 0,\ k \in K(\bar{x}),$$

implies that $\alpha_i = 0$ *for all* $i \in I(\bar{x}), \beta_j = 0$ *for all* $j = 1, 2, \ldots, J, \gamma_k = 0$ *for all* $k \in K(\bar{x}), \eta_l = 0$ *for all* $l = 1, 2, \ldots, L$.

We now prove that the (NNAMCQ) is closely related to but weaker than the (GMFCQ) defined as follows.

DEFINITION 4.3. *We say that* (P) *satisfies the* (GMFCQ) *at* $\bar{x}$ *if there exists* $d_0 \in \text{int}T(\bar{x}, \Omega)$ *such that*

(i) $\langle \nabla g_i(\bar{x}), d_0 \rangle < 0, \phi_k^{\Box}(\bar{x}; d_0) < 0\ \forall i \in I(\bar{x}), k \in K(\bar{x})$,

(ii) $\langle \nabla h_j(\bar{x}), d_0 \rangle = 0, \psi_l^{\Box}(\bar{x}; d_0) = 0,\ i = 1, 2, \ldots, J, l = 1, 2, \ldots, L$,

(iii) *for any* $\xi_l \in \partial^{\diamond} \psi_l(\bar{x}), l = 1, \ldots, L$, $\{\nabla h_1(\bar{x}), \ldots, \nabla h_J(\bar{x}), \xi_1, \ldots, \xi_L\}$ *are linearly independent.*

PROPOSITION 4.4. *The* (GMFCQ) *implies* (NNAMCQ). *Under the assumption that* $\text{int}T(\bar{x}, \Omega) \neq \emptyset$, *the* (GMFCQ) *and* (NNAMCQ) *are equivalent.*

Proof. Since the proof of (GMFCQ) implying (NNAMCQ) is exactly similar to the proof in the case $I = J = 0$ [12, Proposition 4.3], we omit the proof.

We now prove the reverse statement under the assumption that $\text{int}T(\bar{x}, \Omega) \neq \emptyset$. Suppose that the (NNAMCQ) holds but not the (GMFCQ). If for some $\xi_l \in \partial^{\diamond} \psi_l(\bar{x}), l = 1, 2, \ldots, L$, $\{\nabla h_1(\bar{x}), \ldots, \nabla h_J(\bar{x}), \xi_1, \ldots, \xi_L\}$ are linearly dependent, then there exist scalars $\beta_j (j = 1, 2, \ldots, J), \eta_l (l = 1, 2, \ldots, L)$ not all zero such that

$$\begin{aligned} 0 &\in \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) + \sum_{l=1}^{L} \eta_l \partial^{\diamond} \psi_l(\bar{x}) \\ &\subseteq \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) + \sum_{l=1}^{L} \eta_l \partial^{\diamond} \psi_l(\bar{x}) + N(\bar{x}, \Omega), \end{aligned}$$

which contradicts the fact that there is no nonzero abnormal multiplier for (P). If there is no $d_0 \in \text{int}T(\bar{x}, \Omega)$ satisfying items (i) and (ii), then in the case $I \neq 0$, $d = 0$ must be an optimal solution to the following problem:

$$\begin{aligned} \min \quad & \langle \nabla g_{\bar{i}}(\bar{x}), d \rangle \\ \text{s.t.} \quad & \langle \nabla g_i(\bar{x}), d \rangle \leq 0,\ i \in I(\bar{x}) \backslash \{\bar{i}\}, \\ & \langle \nabla h_j(\bar{x}), d \rangle = 0,\ j = 1, \ldots, J, \\ & \phi_k^{\Box}(\bar{x}; d) \leq 0,\ k \in K(\bar{x}), \\ & \psi_l^{\Box}(\bar{x}; d) = 0,\ l = 1, \ldots, L, \\ & d \in T(\bar{x}, \Omega), \end{aligned}$$

where $\bar{i} \in I(\bar{x})$ and in the case where $I = 0$ but $K \neq 0$, $d = 0$ must be an optimal solution to the following problem:

$$\begin{aligned} \min \quad & \phi_{\bar{k}}^{\Box}(\bar{x}; d) \\ \text{s.t.} \quad & \langle \nabla h_j(\bar{x}), d \rangle = 0,\ j = 1, \ldots, J, \\ & \phi_k^{\Box}(\bar{x}; d) \leq 0,\ k \in K(\bar{x}) \backslash \{\bar{k}\}, \end{aligned}$$

$$\psi_l^\square(\bar{x}; d) = 0,\ l = 1, \dots, L,$$
$$d \in T(\bar{x}, \Omega),$$

where $\bar{k} \in K(\bar{x})$. Applying the generalized multiplier rule of Clarke completes the proof. □

In Lipschitz optimization, it is well known that the calmness condition is the weakest constraint qualification. We now extend the definition of the calmness condition [4] to our setting.

DEFINITION 4.5 (calmness). *Let $\bar{x}$ be a solution of* (P). (P) *is calm at $\bar{x}$ provided that there exist $\epsilon > 0$ and $\mu > 0$ such that for all $(p, q, u, v) \in \epsilon B_{I+J+K+L}$ and all $x \in \bar{x} + \epsilon B$ satisfying*

$$g(x) + p \le 0, h(x) + q = 0, \phi(x) + u \le 0, \psi(x) + v = 0, x \in \Omega \tag{4.3}$$

one has

$$f(\bar{x}) \le f(x) + \mu\|(p, q, u, v)\|,$$

where B_n denotes the open unit ball in R^n, $g(x) := (g_1(x), g_2(x), \dots, g_I(x))^t$ and $h(x), \phi(x), \psi(x)$ are the vector-valued mappings defined similarly.

We now prove that the calmness condition is also a constraint qualification in our setting. It is interesting to note that unlike the Fritz John type condition (Theorem 1.1) the KKT conditions under either the calmness condition (Theorem 4.2) or the one under the metric regularity condition (Theorems 4.8 and 4.10) hold even for problem (P) with $L \neq 0$. Moreover, under either the calmness condition or the metric regularity condition, the Fréchet differentiable equality constraints do not need to be continuous near the optimal solution.

THEOREM 4.6 (KKT condition under calmness CQ). *Let $\bar{x}$ be a solution of* (P). *Suppose that the objective function f is either Fréchet differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, the constraint functions satisfy assumption* (A)*, and there exists a vector that is hypertangent to Ω at $\bar{x}$. If* (P) *is calm at $\bar{x}$, then there exist $\alpha_i \ge 0 (i \in I(\bar{x})), \beta_j (j = 1, 2, \dots, J), \gamma_k \ge 0 (k \in K(\bar{x})), \eta_l (l = 1, 2, \dots, L)$ such that*

$$0 \in \partial^\diamond f(\bar{x}) + \sum_{i \in I(\bar{x})} \alpha_i \nabla g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) + \sum_{k \in K(\bar{x})} \gamma_k \partial^\diamond \phi_k(\bar{x})$$
$$+ \sum_{l=1}^{L} \eta_l \partial^\diamond \psi_l(\bar{x}) + N(\bar{x}, \Omega).$$

Proof. By the definition of calmness, $(x, p, u) = (\bar{x}, 0, 0)$ is a local solution to

$$\begin{aligned} \min \quad & f(x) + \mu(\|(p, u)\| + \sum_{j=1}^{J} |h_j(x)| + \sum_{l=1}^{L} |\psi_l(x)|) \\ \text{s.t.} \quad & g(x) + p \le 0, \\ & \phi(x) + u \le 0, \\ & x \in \Omega. \end{aligned}$$

For a function $g_i(x)$, denote by $g_i^+(x) := \max\{g_i(x), 0\}$. Since $g(x) - g^+(x) \le 0, \phi(x) - \phi^+(x) \le 0$ and $g(\bar{x}) - g^+(\bar{x}) = 0, \phi(\bar{x}) - \phi^+(\bar{x}) = 0$, taking $p = -g^+(x), u = -\phi^+(x)$, by the calmness condition $\bar{x}$ is also a local solution of the following problem:

$$\begin{aligned} \min \quad & f(x) + \mu\Bigg(\sqrt{I+K}\max\{g_1(x),\ldots,g_I(x),\phi_1(x),\ldots,\phi_K(x),0\} \\ & \qquad + \sum_{j=1}^{J}|h_j(x)| + \sum_{l=1}^{L}|\psi_l(x)|\Bigg) \\ \text{s.t.} \quad & x \in \Omega. \end{aligned}$$

That is, $(x,r,s,t) = (\bar{x},0,0,0)$ is a local solution of the following problem:

$$\begin{aligned} \min \quad & f(x) + \mu\left(\sqrt{I+K}r + \sum_{j=1}^{J} s_j + \sum_{l=1}^{L} t_l\right) \\ \text{s.t.} \quad & r \geq g_i(x),\ i = 1,2,\ldots,I, \\ & r \geq \phi_k(x),\ k = 1,2,\ldots,K, \\ & r \geq 0, \\ & s_j \geq h_j(x),\ j = 1,2,\ldots,J, \\ & s_j \geq -h_j(x),\ j = 1,2,\ldots,J, \\ & t_l \geq \psi_l(x),\ l = 1,2,\ldots,L, \\ & t_l \geq -\psi_l(x),\ l = 1,2,\ldots,L, \\ & x \in \Omega. \end{aligned}$$

It is straightforward to verify that the (NNAMCQ) for the above problem is satisfied and the Lagrange multiplier rule with $\lambda = 1$ for the original problem follows from applying Theorem 4.1 to the above problem. □

We also extend the notion of metric regularity in smooth and Lipschitz optimization to our setting.

DEFINITION 4.7. *Let C denote the constraint region of* (P) *and $\bar{x} \in C$. C is said to be metrically regular at $\bar{x}$ if there exist positive constants μ, ϵ such that for all $(p,q,u,v) \in \epsilon B$ and all $x \in \bar{x} + \epsilon B$ satisfying* (4.3)*, one has*

$$d_C(x) \leq \mu\|(p,q,u,v)\|.$$

As in smooth and Lipschitz optimization, the metric regularity is stronger than the calmness condition in our setting when the objective function is Lipschitz continuous.

THEOREM 4.8 (KKT condition under the metric regularity assumption when the objective function is Lipschitz). *Let $\bar{x}$ be a solution of* (P)*. Assume that the objective function f is Lipschitz near $\bar{x}$, the constraint functions satisfy assumption* (A)*, and there exists a vector that is hypertangent to Ω at $\bar{x}$. If the constraint region is metrically regular at $\bar{x}$, then the KKT condition as stated in the conclusion of Theorem* 4.6 *also holds.*

Proof. Since the objective function f is Lipschitz near $\bar{x}$, by virtue of Proposition 2.7, $\bar{x}$ is a local solution to the following problem:

$$\min \quad f(x) + L_f d_C(x),$$

where L_f denotes the Lipschitz constant of f near $\bar{x}$ and C is the constraint region of (P). By the metric regularity, $(x,p,q,u,v) = (\bar{x},0,0,0,0)$ is a local solution to the following problem:

$$\begin{aligned} \min \quad & f(x) + L_f \mu \|(p,q,u,v)\| \\ \text{s.t.} \quad & g(x)+p \le 0, h(x)+q \le 0, \phi(x)+u \le 0, \psi(x)+v = 0, x \in \Omega. \end{aligned}$$

That is, the calmness CQ is satisfied at $\bar{x}$ and hence the conclusion of Theorem 4.6 also holds. ∎

Unlike the case where the objective function is Lipschitz continuous, when the objective function is only differentiable, the metric regularity of a constraint region may not imply the calmness as illustrated by the following example.

Example. Consider the following optimization problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x = 0, \end{aligned}$$

where

$$f(x) := \begin{cases} x^2 \sin \frac{1}{x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

It is clear that f is differentiable everywhere with

$$f'(x) := \begin{cases} 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Hence f is differentiable at the optimal solution $\bar{x} = 0$ but not Lipschitz near $\bar{x} = 0$. The constraint region $\{x : x = 0\}$ is metrically regular since the constraint function is linear. However, the problem is not calm at $\bar{x} = 0$ since $\bar{x} = 0$ is not a solution to the perturbed problem

$$\min \quad f(x) + \mu \|x\|$$

for any $\mu > 0$.

However, although the metric regularity is not stronger than the calmness condition when the objective function is not Lipschitz, it turns out that the metric regularity is still a constraint qualification when the objective function is Fréchet differentiable. In the remainder of this section, we would like to prove the KKT condition under the metric regularity assumption when the objective function is Fréchet differentiable. First we prove the following formula for the Fréchet normal cone to the feasible region C and then we use the result to derive the multiplier rules.

LEMMA 4.9. *Let $\bar{x}$ be a feasible solution of* (P). *Assume that the constraint functions satisfy assumption* (A) *and there exists a vector that is hypertangent to Ω at $\bar{x}$. If C, the feasible region of* (P), *is metrically regular at $\bar{x}$, then*

$$N^F(\bar{x}, C) \subset \left\{ \sum_{i \in I(\bar{x})} \alpha_i \nabla g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \nabla h_j(\bar{x}) + \sum_{k \in K(\bar{x})} \gamma_k \partial^\diamond \phi_k(\bar{x}) \right.$$
$$\left. + \sum_{l=1}^{L} \eta_l \partial^\diamond \psi_l(\bar{x}) + N(\bar{x}, \Omega) : \alpha_i \ge 0, \gamma_k \ge 0, i \in I(\bar{x}), k \in K(\bar{x}) \right\}.$$

Proof. Let ξ be any element in $N_C^F(\bar{x})$. Then for any $\lambda \downarrow 0$ there exists $\delta > 0$ such that

$$\langle \xi, x' - \bar{x} \rangle \le \lambda \|x' - \bar{x}\| \qquad \forall x' \in C \cap (\bar{x} + \delta B).$$

That is, $\bar{x}$ is a local solution to the following problem:

$$\begin{array}{ll} \min & -\langle \xi, x' \rangle + \lambda \|x' - \bar{x}\| \\ \text{s.t.} & x' \in C. \end{array}$$

Since the objective function of the above problem is Lipschitz continuous, by virtue of Proposition 2.7, $\bar{x}$ is a local solution to the following problem:

$$\min \quad -\langle \xi, x' \rangle + \lambda \|x' - \bar{x}\| + L d_C(x'),$$

where $L \geq \|\xi\| + \lambda$ for all $\lambda > 0$. By the metric regularity, $\bar{x}$ is a local solution to the following problem:

$$\begin{array}{lll} (\mathrm{P}') & \min & -\langle \xi, x' \rangle + \lambda \|x' - \bar{x}\| \\ & & +L\mu(\sqrt{I+K} \max\{g_1(x'), \ldots, g_I(x'), \phi_1(x'), \ldots, \phi_K(x'), 0\} \\ & & +\|h(x')\| + \|\psi(x')\|) \\ & \text{s.t.} & x' \in \Omega. \end{array}$$

Or equivalently, $(x', r, s, t) = (\bar{x}, 0, 0, 0)$ is a local solution to the following problem:

$$\begin{array}{ll} \min & -\langle \xi, x' \rangle + \lambda \|x' - \bar{x}\| + M \left(r + \sum_{j=1}^{J} s_j + \sum_{l=1}^{L} t_l \right) \\ \text{s.t.} & r \geq g_i(x'), \quad i = 1, 2, \ldots, I, \\ & r \geq \phi_k(x'), k = 1, 2, \ldots, K, \\ & r \geq 0, \\ & s_j \geq h_j(x'), \quad j = 1, 2, \ldots, J, \\ & s_j \geq -h_j(x'), \quad j = 1, 2, \ldots, J, \\ & t_l \geq \psi_l(x'), \quad l = 1, 2, \ldots, L, \\ & t_l \geq -\psi_l(x'), \quad l = 1, 2, \ldots, L, \\ & x' \in \Omega, \end{array}$$

with $M = L\mu\sqrt{I+K}$. One can easily verify that the (NNAMCQ) for the above problem is satisfied. Applying Theorem 4.1, there exist $\alpha_i^\lambda (i \in I(\bar{x}))$, $\beta_j^\lambda (j = 1, 2, \ldots, J)$, $\gamma_k^\lambda (k \in K(\bar{x}))$, $\eta_l^\lambda (l = 1, 2, \ldots, L)$, such that

$$\begin{aligned} 0 \in -\xi + \lambda B^* + \sum_{i \in I(\bar{x})} \alpha_i^\lambda \nabla g_i(\bar{x}) + \sum_{j=1}^{L} \beta_j^\lambda \nabla h_j(\bar{x}) \\ + \sum_{k \in K(\bar{x})} \gamma_k^\lambda \partial^\diamond \phi_k(\bar{x}) + \sum_{l=1}^{L} \eta_l^\lambda \partial^\diamond \psi_l(\bar{x}) + N(\bar{x}, \Omega). \end{aligned}$$

Since the (NNAMCQ) holds for problem (P'), $\{(\alpha^\lambda, \beta^\lambda, \gamma^\lambda, \eta^\lambda)\}$ must be bounded. Without loss of generality, we may assume that $\{(\alpha^\lambda, \beta^\lambda, \gamma^\lambda, \eta^\lambda)\}$ converges. The proof of the lemma is completed after taking limits as $\lambda \to 0$, by virtue of the weak* compactness of the Michel–Penot subdifferentials (see Proposition 2.3). ∎

THEOREM 4.10 (KKT condition under the metric regularity CQ when the objective function is Fréchet differentiable). *Let $\bar{x}$ be a local optimal solution of* (P).

Assume that f is Fréchet differentiable at $\bar{x}$, the constraint functions satisfy assumption (A) *, and there exists a vector that is hypertangent to Ω at $\bar{x}$. If C is metrically regular at $\bar{x}$, then there exist scalars $\alpha_i \geq 0 (i \in I(\bar{x})), \beta_j (j = 1, \ldots, J), \gamma_k \geq 0 (k \in K(\bar{x})), \eta_l (l = 1, 2, \ldots, L)$ such that*

$$0 \in \nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \alpha_i \nabla g_i(\bar{x}) + \sum_{j=1}^{L} \beta_j \nabla h_j(\bar{x})$$
$$+ \sum_{k \in K(\bar{x})} \gamma_k \partial^{\diamond} \phi_k(\bar{x}) + \sum_{l=1}^{L} \eta_l \partial^{\diamond} \psi_l(\bar{x}) + N(\bar{x}, \Omega).$$

Proof. Since f is Fréchet differentiable at $\bar{x}$, we have

$$\lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|} = 0.$$

Since $\bar{x}$ is a local solution to (P), one has

$$\begin{aligned} \limsup_{x \to \bar{x}, x \in C} \frac{-\langle \nabla f(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|} &\leq \limsup_{x \to \bar{x}, x \in C} \frac{f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|} \\ &\leq \lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|} \\ &= 0. \end{aligned}$$

That is, $-\nabla f(\bar{x}) \in N^F(\bar{x}, C)$. The proof of the theorem follows by applying Lemma 4.9, the expression of the Fréchet normal cone to the constraint region. ◻

5. Multiplier rules for the case $I = J = 0$. In this section we consider problem (P) in the case where all constraint functions are Lipschitz and the objective function f is Fréchet differentiable. Under this assumption, we derive multiplier rules without requiring the existence of a hypertangent to the abstract constraint set Ω. Note that in Asplund space, the results are sharper since the limiting subdifferentials and the limiting normal cones instead of the Clarke generalized gradients and the Clarke normal cones are used.

First we prove the following formula for the Fréchet normal cone to the feasible region with $I = J = 0$ and then we use the result to derive the multiplier rules.

LEMMA 5.1. *Let $\bar{x}$ be a feasible solution of* (P) *with $I = J = 0$. Assume that $\phi_k (k \in K(\bar{x})), \psi_l (l = 1, 2, \ldots, L)$ are Lipschitz near $\bar{x}$ and $\phi_k (k \notin K(\bar{x}))$ are continuous at $\bar{x}$. If C is metrically regular at $\bar{x}$, then*

$$N^F(\bar{x}, C) \subset \left\{ \sum_{k \in K(\bar{x})} \gamma_k \partial \phi_k(\bar{x}) + \sum_{l=1}^{L} \eta_l \partial \psi_k(\bar{x}) + N(\bar{x}, \Omega) : \gamma_k \geq 0, k \in K(\bar{x}) \right\},$$

where C denotes the feasible region of (P) *with $I = J = 0$.*

Moreover, if X is an Asplund space, then

$$N^F(\bar{x}, C) \subset \left\{ \sum_{k=1}^{K} \gamma_k \hat{\partial} \phi_k(\bar{x}) + \hat{\partial} \langle \eta, \psi \rangle(\bar{x}) + \hat{N}(\bar{x}, \Omega) : \gamma_k \geq 0, k \in K(\bar{x}) \right\}.$$

Proof. Let ξ be any element in $N_C^F(\bar{x})$. Then for any $\lambda_\nu \downarrow 0$, there exists $\delta > 0$ such that

$$\langle \xi, x' - \bar{x} \rangle \le \lambda_\nu \|x' - \bar{x}\| \qquad \forall x' \in C \cap (\bar{x} + \delta B).$$

That is, $\bar{x}$ is a local solution to the following problem:

$$\begin{array}{ll} \min & -\langle \xi, x' \rangle + \lambda_\nu \|x' - \bar{x}\| \\ \text{s.t.} & x' \in C. \end{array}$$

Since the objective function of the above problem is Lipschitz continuous, by virtue of Proposition 2.7, $\bar{x}$ is a local solution to the following problem:

$$\min \quad -\langle \xi, x' \rangle + \lambda_\nu \|x' - \bar{x}\| + L d_C(x'),$$

where $L \ge \|\xi\| + \lambda_\nu$ for all $\nu = 1, 2, \ldots$. By metrical regularity, $\bar{x}$ is a local solution to the following problem:

$$\begin{array}{ll} \min & -\langle \xi, x' \rangle + \lambda_\nu \|x' - \bar{x}\| + L\mu \left(\sqrt{K} \max\limits_{k \in K(\bar{x})} \{\phi_k(x'), 0\} + \|\psi(x')\| \right) \\ \text{s.t.} & x' \in \Omega. \end{array}$$

Or equivalently, $\bar{x}$ is a local solution to the following problem:

$$\min -\langle \xi, x' \rangle + \lambda_\nu \|x' - \bar{x}\| + M \left(\max_{k \in K(\bar{x})} \{\phi_k(x'), 0\} + \|\psi(x')\| \right) + \tilde{L} d_\Omega(x'),$$

with $M = L\mu\sqrt{K}$ and $\tilde{L}$ being the Lipschitz constant of the objective function of the previous optimization problem.

If X is an Asplund space, then by the sum rule for limiting subdifferentials (Proposition 2.5(ii)),

$$0 \in -\xi + \lambda_\nu B^* + M\hat{\partial}\varphi \circ (\phi, \psi)(\bar{x}) + \hat{N}(\bar{x}, \Omega),$$

where $\varphi(u, v) = \max_{k \in K(\bar{x})}\{u_k, 0\} + \|v\|$. By the chain rule,

$$\xi \in \lambda_\nu B^* + M \cup_{(\gamma,\eta) \in \hat{\partial}\varphi(\phi(\bar{x}), \psi(\bar{x}))} \hat{\partial}\langle (\gamma, \eta), (\phi, \psi) \rangle (\bar{x}) + \hat{N}(\bar{x}, \Omega).$$

That is, there exists $(\gamma_\nu, \eta_\nu) \in \hat{\partial}\varphi(\phi(\bar{x}), \psi(\bar{x}))$ such that

$$\xi \in \lambda_\nu B^* + M\hat{\partial}\langle (\gamma_\nu, \eta_\nu), (\phi, \psi) \rangle (\bar{x}) + \hat{N}(\bar{x}, \Omega).$$

Since φ is Lipschitz, by virtue of Proposition 2.3, (γ_ν, η_ν) is a bounded sequence in R^{K+L} and one can assume that $(\gamma_\nu, \eta_\nu) \to (\gamma, \eta)$ for some $(\gamma, \eta) \in \hat{\partial}\varphi(\phi(\bar{x}), \psi(\bar{x}))$. Hence,

$$\begin{aligned} \xi &\in \lambda_\nu B^* + M\hat{\partial}\langle (\gamma_\nu, \eta_\nu), (\phi, \psi) \rangle (\bar{x}) + \hat{N}(\bar{x}, \Omega) \\ &\subseteq \lambda_\nu B^* + M[\hat{\partial}\langle (\gamma, \eta), (\phi, \psi) \rangle (\bar{x}) + \hat{\partial}\langle (\gamma_\nu, \eta_\nu) - (\gamma, \eta), (\phi, \psi) \rangle (\bar{x})] + \hat{N}(\bar{x}, \Omega) \\ &\subseteq \lambda_\nu B^* + M\hat{\partial}\langle (\gamma, \eta), (\phi, \psi) \rangle (\bar{x}) + M\|(\gamma_\nu, \eta_\nu) - (\gamma, \eta)\| L_{(\phi,\psi)} B^* + \hat{N}(\bar{x}, \Omega). \end{aligned}$$

Taking limits as $\nu \to \infty$, by virtue of the weak* sequential closedness of limiting subdifferentials, one has

$$\xi \in M\hat{\partial}\langle (\gamma, \eta), (\phi, \psi) \rangle (\bar{x}) + \hat{N}(\bar{x}, \Omega)$$

for some

$$(\gamma,\eta) \in \hat{\partial}\varphi(\phi(\bar{x}),\psi(\bar{x})) = \left\{(\gamma,\eta) : \sum_{k\in K(\bar{x})} \gamma_k = 1, \gamma_k \geq 0 k \in K(\bar{x}), \eta \in B_L\right\}.$$

The case where X is a general Banach space can be proved similarly. □

THEOREM 5.2 (KKT condition when $I = I = 0$ under the metric regularity CQ). *Let $\bar{x}$ be a local optimal solution of* (P) *with $I = J = 0$. Assume that f is Fréchet differentiable at $\bar{x}$, $\phi_k(k \in K(\bar{x})), \psi_l(l = 1, 2, \ldots, L)$ are Lipschitz near $\bar{x}$ and $\phi_k(k \notin K(\bar{x}))$ are continuous at $\bar{x}$. If C is metrically regular at $\bar{x}$, then there exist $\gamma_k \geq 0(k \in K(\bar{x})), \eta_l(l = 1, 2, \ldots, L)$ such that*

$$0 \in \nabla f(\bar{x}) + \sum_{k\in K(\bar{x})} \gamma_k \partial\phi_k(\bar{x}) + \sum_{l=1}^{L} \eta_l \partial\psi_l(\bar{x}) + N(\bar{x},\Omega).$$

Moreover, if X is a Asplund space and C is metrically regular at $\bar{x}$, then there exist $\gamma_k \geq 0(k \in K(\bar{x})), \eta_l \in R(l = 1, 2, \ldots, L)$ such that

$$0 \in \nabla f(\bar{x}) + \sum_{k\in K(\bar{x})} \gamma_k \hat{\partial}\phi_k(\bar{x}) + \hat{\partial}\langle\eta,\psi\rangle(\bar{x}) + \hat{N}(\bar{x},\Omega).$$

Proof. Since f is Fréchet differentiable at $\bar{x}$, as in the proof of Theorem 4.10,

$$-\nabla f(\bar{x}) \in N^F(\bar{x}, C).$$

The proof of the theorem follows by applying Lemma 5.1, the expression of the Fréchet normal cone to the constraint region. □

Remark 2. Sufficient conditions for metrical regularity in the case $I = J = 0$ include the following:

(i) (see [10, Theorem 3].) The constraint region is defined by a system of linear equalities and inequalities, i.e.,

$$C := \{x \in X : \langle x_k^*, x\rangle = 0, k = 1, \ldots, K, \langle y_l^*, x\rangle \leq 0 l = 1, \ldots, L\}$$

for some $x_k^* \in X^*(k = 1, \ldots, K), y_l^* \in X^*(l = 1, \ldots, L)$.

(ii) In Banach space [4, Theorem 6.6.1], the (NNAMCQ) in the Clarke generalized gradient form is satisfied, i.e.,

$$0 \in \sum_{k\in K(\bar{x})} \gamma_k \partial\phi_k(\bar{x}) + \sum_{l=1}^{L} \eta_l \partial\psi_k(\bar{x}) + N(\bar{x},\Omega),$$
$$\gamma_k \geq 0 \ \forall k \in K(\bar{x})$$

implies that $\gamma_k = 0, k \in K(\bar{x}), \eta_l = 0, l = 1, 2, \ldots, L$. In Asplund space [18, Corollary 6.2], the (NNAMCQ) in the limiting subdifferential form is satisfied, i.e.,

$$0 \in \sum_{k\in K(\bar{x})} \gamma_k \hat{\partial}\phi_k(\bar{x}) + \hat{\partial}\langle\eta,\psi\rangle(\bar{x}) + \hat{N}(\bar{x},\Omega),$$
$$\gamma_k \geq 0 \ \forall k \in K(\bar{x})$$

implies that $\gamma_k = 0, k \in K(\bar{x}), \eta_l = 0, l = 1, 2, \ldots, L$.

THEOREM 5.3 (KKT condition when $I = J = 0$ under the calmness CQ). *Let $\bar{x}$ be a local optimal solution of* (P) *with $I = J = 0$. Assume that f is Fréchet differentiable at $\bar{x}$, $\phi_k(k \in K(\bar{x})), \psi_l(l = 1, 2, \ldots, L)$ are Lipschitz near $\bar{x}$ and $\phi_k(k \notin K(\bar{x}))$ are continuous at $\bar{x}$. If* (P) *is calm at $\bar{x}$, then the conclusions of Theorem* 5.2 *hold.*

Proof. By the definition of calmness, $(x, u) = (\bar{x}, 0)$ is a local solution to

$$\begin{aligned} \min \quad & f(x) + \mu(\|u\| + \|\psi(x)\|) \\ \text{s.t.} \quad & \phi(x) + u \le 0, \\ & x \in \Omega. \end{aligned}$$

Since $\phi(x) - \phi^+(x) \le 0$ and $\phi(\bar{x}) - \phi^+(\bar{x}) = 0$, $\bar{x}$ is also a local solution of the following problem:

$$\begin{aligned} \min \quad & f(x) + \mu\left(\sqrt{K}\max\{\phi_1(x), \ldots, \phi_K(x), 0\} + \|\psi(x)\|\right) \\ \text{s.t.} \quad & x \in \Omega. \end{aligned}$$

Case 1. X is a general Banach space. It is easy to see that $(x, r, s) = (\bar{x}, 0, 0)$ is a local solution to the following problem:

$$\begin{aligned} \min \quad & f(x) + \mu\left(\sqrt{K}r + \sum_{l=1}^{L} s_l\right) \\ \text{s.t.} \quad & r \ge \phi_k(x),\ k = 1, \ldots, K, \\ & r \ge 0, \\ & s_l \ge \psi_l(x), \quad l = 1, 2, \ldots, L, \\ & s_l \ge -\psi_l(x), \quad l = 1, 2, \ldots, L, \\ & x \in \Omega. \end{aligned}$$

It is straightforward to verify that the (NNAMCQ) for the above problem is satisfied and the KKT condition follows from Theorem 5.2 and in Remark 2(ii).

Case 2. X is an Asplund space. Equivalently, $\bar{x}$ is a local solution to the following problem:

$$\min \quad f(x) + \mu\left(\sqrt{K}\max_{k \in K(\bar{x})}\{\phi_k(x), 0\} + \|\psi(x)\|\right) + \delta_\Omega(x),$$

where $\delta_\Omega(x)$ is the indicator function of a set Ω defined by

$$\delta_\Omega(x) := \begin{cases} 0 & \text{if } x \in \Omega, \\ +\infty & \text{if } x \notin \Omega. \end{cases}$$

Since f is Fréchet differentiable and $G(x) := \mu(\sqrt{K}\max_{k \in K(\bar{x})}\{\phi_K(x), 0\} + \|\psi(x)\|)$ is Lipschitz near $\bar{x}$, one has

$$\begin{aligned} 0 &\in \nabla f(\bar{x}) + \partial^F(G + \delta_\Omega)(\bar{x}) \quad \text{(by Proposition 2.5(i))} \\ &\subseteq \nabla f(\bar{x}) + \hat{\partial}(G + \delta_\Omega)(\bar{x}) \\ &\subseteq \nabla f(\bar{x}) + \hat{\partial}G(\bar{x}) + \hat{N}(\bar{x}, \Omega) \quad \text{(Proposition 2.5(ii))}. \end{aligned}$$

The remaining proof follows by using the sum rules and the chain rules as in the proof of Lemma 5.1. ∎

Proof of Theorem 1.2. Suppose X is a Banach space. If the (NNAMCQ) in the Clarke generalized gradient form does not hold, then the Fritz John condition holds with $\lambda = 0$. Otherwise if the (NNAMCQ) in the Clarke generalized gradient form holds, then by Remark 2 and Theorem 5.2, the Fritz John condition holds with $\lambda = 1$.

Similarly suppose that X is an Asplund space. If the (NNAMCQ) in the limiting subdifferential form as in Remark 2 does not hold, then the required Fritz John condition holds with $\lambda = 0$. Otherwise if the (NNAMCQ) in the limiting subdifferential form holds, then by Remark 2 and Theorem 5.2, the required Fritz John condition holds with $\lambda = 1$.

Acknowledgments. The author would like to thank Jay Treiman for his suggestions on replacing the Clarke generalized gradient by the Michel–Penot subdifferential in Theorem 1.1 and those in section 4 in an earlier version.

REFERENCES

[1] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming Theory and Algorithms*, 2nd ed., John Wiley & Sons, New York, 1993.
[2] J.M. Borwein and Q. Zhu, *A survey of subdifferential calculus with applications*, Nonlinear Anal., 38 (1999), pp. 687–773.
[3] F.H. Clarke, *A new approach to Lagrange multipliers*, Math. Oper. Res., 1 (1976), pp. 165–174.
[4] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.
[5] F.H. Clarke, Yu. S. Ledyaev, R.J. Stern, and P.R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.
[6] F.H. Clarke, R. J. Stern, and P.R. Wolenski, *Subgradient criteria for monotonicity, the Lipschitz condition, and convexity*, Canad. J. Math., 45 (1993), pp. 1167–1183.
[7] S. Di, *Classical optimality conditions under weaker assumptions*, SIAM J. Optim., 6 (1996), pp. 178–197.
[8] L.A. Fernández, *On the limits of the Lagrange multiplier rule*, SIAM Rev., 39 (1997), pp. 292–297.
[9] H. Halkin, *Implicit functions and optimization problems without continuous differentiability of the data*, SIAM J. Control, 12 (1974), pp. 229–236.
[10] A.D. Ioffe, *Regular points of Lipschitz functions*, Trans. Amer. Math. Soc., 251 (1979), pp. 61–69.
[11] A.D. Ioffe, *A Lagrange multiplier rule with small convex-valued subdifferentials for nonsmooth problems of mathematical programming involving equality and nonfunctional constraints*, Math. Programming, 58 (1993), pp. 137–145.
[12] A. Jourani, *Constraint qualifications and Lagrange multipliers in nondifferentiable programming problems*, J. Optim. Theory Appl., 81 (1994), pp. 533–548.
[13] P.D. Loewen, *Optimal Control via Nonsmooth Analysis*, CRM Proc. Lecture Notes 2, AMS, Providence, RI, 1993.
[14] O.L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York, 1969; reprint, SIAM, Philadelphia, 1994.
[15] P. Michel and J.-P. Penot, *Calcul sous-différentiel pour des fonctions lipschitziennes et non lipschitziennes*, C.R. Acad. Sci. Paris Sér. I Math, 12 (1984), pp. 269–272.
[16] B.S. Mordukhovich, *On necessary conditions for an extremum in nonsmooth optimization*, Soviet Math. Dokl., 283 (1985), pp. 215–220.
[17] B.S. Mordukhovich and Y. Shao, *Nonsmooth sequential analysis in Asplund spaces*, Trans. Amer. Math. Soc., 348 (1996), pp. 1235–1280.
[18] B.S. Mordukhovich and Y. Shao, *Stability of set-valued mappings in infinite dimensions: Point criteria and applications*, SIAM J. Control Optim., 35 (1997), pp. 285–314.
[19] J.S. Treiman, *Shrinking generalized gradients*, Nonlinear Anal., 12 (1988), pp. 1429–1450.