The 7th Seminar on Functional Analysis and its Applications
1-2 March 2023
Imam Khomeini International University
Qazvin, Iran

GENERALIZED FJ AND KKT CONDITIONS IN NONSMOOTH NONCONVEX OPTIMIZATION

JAVAD KOUSHKI[1,*], MAJID SOLEIMANI-DAMANEH[2]

[1] *Department of Applied Mathematics, Faculty of Mathematics, K. N. Toosi University of Technology, Tehran, Iran*
j.koushki@kntu.ac.ir
[2] *School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran*
m.soleimani.d@ut.ac.ir

Abstract. In this talk, we investigate optimality conditions for nonsmooth nonconvex optimization problems by means of generalized Fritz John (FJ) and Karush-Kuhn-Tucker (KKT) conditions. We obtain alternative-type optimality conditions, which could be helpful in analyzing duality results and sketching numerical algorithms.

1. Introduction

FJ and KKT conditions play a central role in optimization (both theoretically and numerically). Many researchers have examined these conditions under different assumptions. Consider the following optimization problem with inequality constraints and a nonempty geometric constraint set $X \subseteq \mathbb{R}^n$:

$$\begin{array}{ll} \min & f(x) \\ s.t. & g_i(x) \leq 0,\ i = 1, 2, \cdots, m, \\ & x \in X, \end{array} \tag{1.1}$$

2020 *Mathematics Subject Classification. Primary 90C26; Secondary 90C30*
Key words and phrases. Nonsmooth optimization, Nonconvex optimization, FJ conditions, KKT conditions.
[*] Speaker.

in which $f, g_i : \mathbb{R}^n \longrightarrow \mathbb{R},\ i = 1, 2, \ldots, m$, are real-valued functions. We show the feasible solutions set of (1.1) by

$$F := \{x \in X : g_i(x) \leq 0,\ i = 1, 2, \cdots, m\},$$

and the set of indices of the active constraints at $\bar{x} \in F$ by

$$I(\bar{x}) := \{i \in \{1, 2, \cdots m\} : g_i(\bar{x}) = 0\}.$$

For a set $K \subseteq \mathbb{R}^n$, the nonnegative polar cone of K, the tangent cone of K at $y \in clK$, and the normal cone of K at $y \in clK$, denoted by K°, $T_K(y)$, and $N_K(y)$, respectively, and defined as

$$K^\circ := \{z \in \mathbb{R}^n : z^T y \geq 0,\ \forall y \in K\},$$

$$T_K(y) := \Big\{ d \in \mathbb{R}^n : \exists \Big(t_\nu > 0,\ \{y^\nu\} \subseteq K \Big)\ s.t.\ \ y^\nu \longrightarrow y,\ t_\nu(y^\nu - y) \rightarrow d \ \Big\},$$

$$N_K(y) = -[T_K(y)]^\circ.$$

If the functions appeared in problem (1.1) are differentiable, $\bar{x} \in F$ is said to be an FJ point of (1.1) if there are non-negative coefficients $\lambda_0, \lambda_i \geq 0,\ i \in I(\bar{x})$, not all zero, such that

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) \in [T_X(\bar{x})]^\circ. \tag{1.2}$$

If $\bar{x} \in intX$, then $T_X(\bar{x}) = \mathbb{R}^n$ and $[T_X(\bar{x})]^\circ = \{0_n\}$. In this case, the above-mentioned FJ condition is reduced to the well-known classic form. Also, if $\lambda_0 \neq 0$, then we reach the KKT condition.

In the following, we present Flores-Bazan and Mastroeni's definition [1] of the FJ and KKT points, which takes into account any arbitrary set $B \subseteq \mathbb{R}^n$ instead of the tangent cone.

Definition 1.1. [1] Let $B \subseteq \mathbb{R}^n$ be a given nonempty set. Assuming differentiability of f and g_i's, a vector $\bar{x} \in F$ is called a

(i) B-FJ point of (1.1) if there exist scalars $\lambda_0, \lambda_i \geq 0,\ i \in I(\bar{x})$, not all zero, satisfying

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

(ii) B-KKT point of (1.1) if there exist scalars $\lambda_i \geq 0,\ i \in I(\bar{x})$, satisfying

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

2. Alternative-type FJ and KKT optimality conditions

Let $Z \subseteq \mathbb{R}^n$. The interior, the closure, the relative interior, and the boundary of Z are denoted by $int\, Z$, $cl\, Z$, $ri\, Z$, and $bd\, Z$, respectively. The convex hull and the cone generated by Z are denoted by $conv\, Z$, and $cone\, Z$, respectively. Recall that $cone\, Z := \bigcup_{t \geq 0} tZ$.

Definition 2.1. [3] Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be locally Lipschitz at $\bar{x} \in \mathbb{R}^n$. The generalized directional derivative of f at $\bar{x} \in \mathbb{R}^n$ in direction $d \in \mathbb{R}^n$ is defined by

$$f^\circ(\bar{x}; d) := \limsup_{\substack{y \longrightarrow \bar{x} \\ t \downarrow 0}} \frac{f(y + td) - f(y)}{t}.$$

Moreover, the Clarke subdifferential (generalized gradient) of f at $\bar{x} \in \mathbb{R}^n$ is the set

$$\partial f(\bar{x}) := \{\xi \in \mathbb{R}^n : f^\circ(\bar{x}; d) \geq \xi^T d, \quad \forall d \in \mathbb{R}^n\}.$$

Theorem 2.2. [3, Theorem 5.1.6] *Suppose that $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is locally Lipschitz at $\bar{x} \in \mathbb{R}^n$ and attains its local minimum over the set $C \subseteq \mathbb{R}^n$ at $\bar{x}$. Then*

$$0 \in \partial f(\bar{x}) + N_C(\bar{x}).$$

Let $B \subseteq \mathbb{R}^n$ be a given nonempty set. Consider the following sublinear problem corresponding to Problem (1.1) and given B:

$$\begin{array}{lll} \mu := & \inf & f^\circ(\bar{x}; d) \\ & s.t. & d \in G_o(\bar{x}), \end{array} \tag{2.1}$$

where $G_o(\bar{x}) := \{d \in cl\, conv B : g_i^\circ(\bar{x}; d) < 0, \ \forall i \in I(\bar{x})\}$. Set $\mu := +\infty$ whenever $G_o(\bar{x}) = \emptyset$.

Definition 2.3. [2] Suppose that $f, g_i, \ i \in I(\bar{x})$, are locally Lipschitz at $\bar{x} \in F$. The vector $\bar{x}$ is called a

(i) FJ point of (1.1) if there exist $\lambda_0, \lambda_i \geq 0, \ i \in I(\bar{x})$, not all zero, $\bar{\xi} \in \partial f(\bar{x})$, and $\bar{\zeta}_i \in \partial g_i(\bar{x}), \ i \in I(\bar{x})$, such that

$$\lambda_0 \bar{\xi} + \sum_{i \in I(\bar{x})} \lambda_i \bar{\zeta}_i \in B^\circ. \tag{2.2}$$

(ii) KKT point of (1.1) if there exist $\bar{\xi} \in \partial f(\bar{x}), \ \bar{\zeta}_i \in \partial g_i(\bar{x}), \ \lambda_i \geq 0$; $i \in I(\bar{x})$, such that

$$\bar{\xi} + \sum_{i \in I(\bar{x})} \lambda_i \bar{\zeta}_i \in B^\circ. \tag{2.3}$$

The next results have been reported in our recent work, [2]. Based on the following Theorem, we can derive an alternative-type FJ optimality condition.

Theorem 2.4. *Suppose that $\bar{x} \in X$ and $f, g_i, \ i \in I(\bar{x})$, are locally Lipschitz at $\bar{x}$. Then one and only one of the following two statements is true.*

(i) *There exists $d \in cl\, conv B$ such that*

$$\begin{array}{l} f^\circ(\bar{x}; d) < 0, \\ g_i^\circ(\bar{x}; d) < 0, \quad i \in I(\bar{x}). \end{array} \tag{2.4}$$

(ii) *$\bar{x}$ is a FJ point of (1.1).*

Furthermore, if B is a cone, then

$$(i) \Longleftrightarrow \mu = -\infty.$$

In consequence of Theorem 2.4, Corollary 2.5 provides an FJ necessary optimality condition.

Corollary 2.5. *Assume that $\bar{x} \in F$ and f, g_i, $t = 1, 2, \cdots, m$, are locally Lipschitz at $\bar{x}$. Furthermore, suppose that $cl\, conv B \subseteq T_X(\bar{x})$. Then the FJ condition (2.2) is satisfied if x is a local optimal solution to (1.1).*

Corollary 2.6 provides a KKT necessary optimality conditions under some assumptions. The first assumption is related to the set B, and second one is corresponding to the feasibility of the sublinear Problem (2.1).

Corollary 2.6. *Let $\bar{x} \in F$ be given. Assume that f, g_i, $t = 1, 2, \cdots, m$, are locally Lipschitz at $\bar{x}$. Furthermore, assume that $cl\, conv B \subseteq T_X(\bar{x})$. If $\bar{x}$ is a local optimal solution to (1.1), then under either (a) or (b) $\bar{x}$ is a KKT point of (1.1).*

(a) $conv\{\xi_i : \xi_i \in \partial g_i(\bar{x}),\ i \in I(\bar{x})\} \cap B^\circ = \emptyset$;
(b) *There exists some $d \in cl\, conv B$ such that $g_i^\circ(\bar{x}; d) < 0$, for any $i \in I(\bar{x})$.*

Theorem 2.7 provides a necessary and sufficient condition equivalent to KKT conditions. Given $\Omega \subseteq \mathbb{R}^n$, define

$$F(\Omega) := \left\{ \begin{pmatrix} f^\circ(\bar{x}; d) \\ g^\circ_{I(\bar{x})}(\bar{x}; d) \end{pmatrix} : d \in \Omega \right\}.$$

in which, $g^\circ_{I(\bar{x})}(\bar{x}; d)$ is a $|I(\bar{x})|$-vector whose components are $g_i^\circ(\bar{x}; d)$, $i \in I(\bar{x})$.

Theorem 2.7. *Let $\bar{x} \in F$ and $B \subseteq \mathbb{R}^n$ be a nonempty cone. Assume that f, g_i, $i = 1, 2, \cdots, m$, are locally Lipschitz at $\bar{x}$. Then $\bar{x}$ is a KKT point for (1.1) if and only if*

$$cl\Big[F(cl\, conv B) + \big(\mathbb{R}_+ \times \mathbb{R}_+^{I(\bar{x})}\big)\Big] \bigcap -\big(\mathbb{R}_{++} \times \{0\}\big) = \emptyset.$$

More results will be presented in the related talk.

References

1. Flores-Bazan, F., Mastroeni, G.: Characterizing FJ and KKT conditions in nonconvex mathematical programming with applications, SIAM J. Optim., **25**, 647-676 (2015)
2. Koushki, J., Soleimani-damaneh, M.: Characterization of generalized FJ and KKT conditions in nonsmooth nonconvex optimization, J Glob Optim **76**, 407-431 (2020)
3. Makela, M. M., Neittaanmaki, P.: Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control, World Scientific, Singapore (1992)