

## ON NONSMOOTH OPTIMALITY THEOREMS FOR ROBUST OPTIMIZATION PROBLEMS

GUE MYUNG LEE AND PHẠM TIẾN SƠN

ABSTRACT. In this paper, we prove a necessary optimality theorem for a nonsmooth optimization problem in the face of data uncertainty, which is called a robust optimization problem. Recently, the robust optimization problems have been intensively studied by many authors. Moreover, we give examples showing that the convexity of the uncertain sets and the concavity of the constraint functions are essential in the optimality theorem. We present an example illustrating that our main assumptions in the optimality theorem can be weakened.

### 1. Introduction

Let  $X$  be a Banach space, and let functions  $f, g_i: X \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$  be given. Consider the following optimization problem with inequality constraints:

$$(P) \quad \inf\{f(x) \mid g_i(x) \leq 0, i = 1, \dots, m\}.$$

This problem in the face of data uncertainty in the constraints can be written by the following optimization problem:

$$(UP) \quad \inf\{f(x) \mid g_i(x, v_i) \leq 0, i = 1, \dots, m\},$$

where  $v_i$  is an uncertain parameter and  $v_i \in \mathcal{V}_i$  for some sequentially compact topological space  $\mathcal{V}_i$  and  $g_i: X \times \mathcal{V}_i \rightarrow \mathbb{R}$  is a function. Robust optimization, which has emerged as a powerful deterministic approach for studying optimization problems under uncertainty [3, 5, 6, 7], associates with the uncertain program (UP) its robust counterpart [2, 4, 16],

$$(RP) \quad \inf\{f(x) \mid g_i(x, v_i) \leq 0, \forall v_i \in \mathcal{V}_i, i = 1, \dots, m\},$$

---

Received April 1, 2013; Revised July 5, 2013.

2010 *Mathematics Subject Classification.* Primary 90C22, 90C25, 90C46.

*Key words and phrases.* robust optimization problem, Lagrange multipliers, locally Lipschitz functions, generalized gradients, necessary optimality conditions.

The first author was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. 2011-0018619).

The second author was supported by the National Foundation for Science and Technology Development (NAFOSTED), Vietnam, Grant 101.04-2013.07.

where the uncertain constraints are enforced for every possible value of the parameters within their prescribed uncertainty sets  $\mathcal{V}_i$ ,  $i = 1, \dots, m$ . The problem (RP) can be understood as the robust case (the worst case) of (UP). So optimizing (UP) with (RP) can be regarded as the robust approach (worst approach) for (UP). For the importance of practical view, many authors have intensively studied robust optimization [2-7, 14-17, 19]. Very recently, Jeyakumar, Li and Lee [17] proved a differentiable Karush-Kuhn-Tucker (KKT for short) optimality theorem for (RP) when involved functions in (RP) are continuously differentiable. The aim of this paper is to extend the KKT optimality theorem in [17] to a nonsmooth and robust optimization problem and to give examples to show that some assumptions in our KKT optimality theorem are essential for getting an uncertain vector  $v_i$  of the uncertain set of (RP) in the theorem.

In this paper, we give a nonsmooth KKT optimality theorem for (RP) when involved functions in (RP) are locally Lipschitz and to present examples showing that the convexity assumption of the uncertain sets and the concavity assumption of the constraint functions are essential for our KKT optimality theorem for (RP). Moreover, we give an example illustrating that the assumptions in the nonsmooth KKT optimality theorem can be weakened.

The paper is organized as follows. In Section 2, we recall some basic properties of the generalized gradient for the locally Lipschitz function and prove useful results on generalized gradients for constraint functions (see Theorem 2.2). In Section 3, we state and prove optimality theorems in Karush-Kuhn-Tucker formulations for (RP). Some two counter-examples for our robust KKT necessary optimality theorem and an illustrating example for an extension of the theorem are given in Section 4.

## 2. Preliminaries

### 2.1. Generalized gradients

Let  $U$  be an open subset of a Banach space  $X$ , and let a function  $f: U \rightarrow \mathbb{R}$  be given. We shall suppose that  $f$  is Lipschitz on  $U$ ; i.e., that for some constant  $L > 0$ , for all  $x$  and  $y$  in  $U$ , we have

$$|f(x) - f(y)| \leq L\|x - y\|.$$

**Definition 2.1.** For each  $d \in X$ , the *generalized directional derivative* of  $f$  at  $x$  in the direction  $d$ , denoted  $f^0(x; d)$ , is given by

$$f^0(x; d) = \limsup_{h \rightarrow 0, t \rightarrow 0+} \frac{f(x + h + td) - f(x + h)}{t}$$

(Here of course  $h$  belongs to  $X$  and  $t$  to  $(0, \infty)$ ).

We also shall denote the usual *one-sided directional derivative* of  $f$  at  $x$  by  $f'(x; d)$ . Thus

$$f'(x; d) = \lim_{t \rightarrow 0+} \frac{f(x + td) - f(x)}{t},$$

whenever this limit exists.

In the sequel,  $X^*$  denotes the (continuous) dual space of  $X$  and  $\langle \cdot, \cdot \rangle$  is the duality pairing between  $X$  and  $X^*$ . The norm of an element  $\xi$  of  $X^*$ , denoted  $\|\xi\|^*$ , is given by

$$\|\xi\|^* := \sup\{\langle \xi, d \rangle \mid d \in X, \|d\| \leq 1\}.$$

However, all statements involving a topology on  $X^*$  are with respect to the weak\* topology, unless otherwise stated.

**Definition 2.2.** The *generalized gradient* of  $f$  at  $x$ , denoted by  $\partial f(x)$ , is the (nonempty) set of all  $\xi$  in  $X^*$  satisfying the following condition

$$f^0(x; d) \geq \langle \xi, d \rangle \quad \text{for all } d \in X.$$

The following theorem summarizes some fundamental results in the calculus of generalized gradients (for more details, see [8, 9, 10, 11, 18]).

**Theorem 2.1.** *The following statements hold:*

- (i)  $\partial f(x)$  is a nonempty, convex, weak\* compact subset of  $X^*$  and  $\|\xi\|^* \leq L$  for every  $\xi$  in  $\partial f(x)$ .
- (ii) The function  $d \mapsto f^0(x; d)$  is convex and satisfies

$$|f^0(x; d)| \leq L\|d\|.$$

- (iii) For every  $d$  in  $X$ , one has

$$f^0(x; d) = \max\{\langle \xi, d \rangle \mid \xi \in \partial f(x)\}.$$

- (iv) Let  $\psi: U \rightarrow \mathbb{R}$  be another locally Lipschitz function. Then

$$\partial(f + \psi) \subset \partial f + \partial \psi.$$

## 2.2. Max functions

In this subsection, we study the generalized gradient of a type of function which is important in optimization.

Let  $\mathcal{V}$  be a sequentially compact topological space, and let  $U$  be an open subset of a Banach space  $X$ . Suppose that we are given a function  $g: U \times \mathcal{V} \rightarrow \mathbb{R}$ ,  $(x, v) \mapsto g(x, v)$ , satisfying the following conditions:

### Assumptions

- (A1)  $g(x, v)$  is upper semicontinuous in  $(x, v)$ .
- (A2)  $g$  is Lipschitz in  $x$ , uniformly for  $v$  in  $\mathcal{V}$ ; i.e., for some  $L > 0$ , for all  $x, y \in U$  and  $v \in \mathcal{V}$ , one has

$$|g(x, v) - g(y, v)| \leq L\|x - y\|.$$

- (A3)  $g_x^0(x, v; \cdot) = g_x'(x, v; \cdot)$ , the derivatives being with respect to  $x$ .
- (A4) the generalized gradient  $\partial_x g(x, v)$  with respect to  $x$  is weak\* upper semicontinuous in  $(x, v)$ .

*Remark 2.2.* In a suitable setting, assumptions (A2), (A3), and (A4) will follow if the function  $g$  is convex in  $x$  and continuous in  $v$ . These conditions on the function  $g$  also hold when the derivative  $\nabla_x g(x, v)$  with respect to  $x$  exists and is continuous in  $(x, v)$ .

We define a function  $\psi: U \rightarrow \mathbb{R}$  via

$$\psi(x) := \max\{g(x, v) \mid v \in \mathcal{V}\},$$

and we observe that our assumptions (A1)-(A4) imply that  $\psi$  is defined and finite (with the maximum defining  $\psi$  attained) on  $U$ . It also follows readily that  $\psi$  is Lipschitz on  $U$  (of rank  $L$ ), since each  $g(\cdot, v)$  is.

Let

$$\mathcal{V}(x) := \{v \in \mathcal{V} \mid g(x, v) = \psi(x)\}.$$

It is easy to see that  $\mathcal{V}(x)$  is nonempty and closed for each  $x$  in  $U$ .

The following lemma, which is a nonsmooth version of Danskin's theorem for max-functions, makes connection between the functions  $\psi'(x; d)$  and  $g_x^0(x, v; d)$ .

**Lemma 2.3.** *Under the assumptions (A1)-(A4), the usual one-sided directional derivative  $\psi'(x; d)$  exists, and satisfies*

$$\begin{aligned} \psi'(x; d) = \psi^0(x; d) &= \max\{g_x^0(x, v; d) \mid v \in \mathcal{V}(x)\} \\ &= \max\{\langle \xi, d \rangle \mid \xi \in \partial_x g(x, v), v \in \mathcal{V}(x)\}. \end{aligned}$$

*Proof.* See [10, Theorem 2] (see also [8, Theorem 2.1], [12]). □

The following result will be useful in our later analysis.

**Theorem 2.4.** *In addition to the basic assumptions (A1)-(A4), suppose that  $\mathcal{V}$  is convex, and that  $g(x, \cdot)$  is concave on  $\mathcal{V}$ , for each  $x \in U$ . Then the following statements hold*

- (i) *The set  $\mathcal{V}(x)$  is convex and sequentially compact.*
- (ii) *The set*

$$\partial_x g(x, \mathcal{V}(x)) := \{\xi \mid \exists v \in \mathcal{V}(x) \text{ such that } \xi \in \partial_x g(x, v)\}$$

*is convex and weak\* compact.*

- (iii)  *$\partial\psi(x) = \{\xi \mid \exists v \in \mathcal{V}(x) \text{ such that } \xi \in \partial_x g(x, v)\}$ .*

*Proof.* (i) **The set  $\mathcal{V}(x)$  is convex.** In fact, let  $v^1, v^2 \in \mathcal{V}(x) \subset \mathcal{V}$  and  $\alpha \in (0, 1)$ . Then  $\alpha v^1 + (1 - \alpha)v^2 \in \mathcal{V}$  and  $g(x, v^1) = g(x, v^2) = \psi(x)$ . As the function  $g(x, \cdot)$  is concave, we see that

$$\psi(x) \geq g(x, \alpha v^1 + (1 - \alpha)v^2) \geq \alpha g(x, v^1) + (1 - \alpha)g(x, v^2) = \psi(x).$$

Therefore  $g(x, \alpha v^1 + (1 - \alpha)v^2) = \psi(x)$ , and so  $\alpha v^1 + (1 - \alpha)v^2 \in \mathcal{V}(x)$ .

**The set  $\mathcal{V}(x)$  is sequentially compact.** Suppose that we have a sequence of points  $v^k \in \mathcal{V}(x) \subset \mathcal{V}, k \geq 1$ . We may assume that  $v^k$  converges to some

$v \in \mathcal{V}$ , since  $\mathcal{V}$  is sequentially compact. Then, by definition,  $g(x, v^k) = \psi(x)$  for all  $k \geq 1$ . As  $g(x, \cdot)$  is upper semicontinuous, we see that

$$\psi(x) = \lim_{k \rightarrow \infty} g(x, v^k) \leq g(x, v) \leq \psi(x).$$

Consequently,  $g(x, v) = \psi(x)$ , and so  $v \in \mathcal{V}(x)$ .

(ii) For short, let us denote

$$\Gamma := \{\xi \mid \exists v \in \mathcal{V}(x) \text{ such that } \xi \in \partial_x g(x, v)\}.$$

**The set  $\Gamma$  is convex.** In fact, let  $\xi^1, \xi^2 \in \Gamma$  and  $\alpha \in (0, 1)$ . Then, by the definition, there exist  $v^1, v^2 \in \mathcal{V}(x)$  such that  $\xi^1 \in \partial_x g(x, v^1)$  and  $\xi^2 \in \partial_x g(x, v^2)$ . By Item (i), the set  $\mathcal{V}(x)$  is convex. Hence  $\alpha v^1 + (1 - \alpha)v^2 \in \mathcal{V}(x)$ . Note that

$$\begin{aligned} g(x, \alpha v^1 + (1 - \alpha)v^2) &= \psi(x) = g(x, v^1) = g(x, v^2) \\ &= \alpha g(x, v^1) + (1 - \alpha)g(x, v^2). \end{aligned}$$

Therefore, we have for any  $d \in X$ ,

$$\begin{aligned} &g_x^0(x, \alpha v^1 + (1 - \alpha)v^2; d) \\ &= g'_x(x, \alpha v^1 + (1 - \alpha)v^2; d) \\ &= \lim_{t \rightarrow 0^+} \frac{g(x + td, \alpha v^1 + (1 - \alpha)v^2) - g(x, \alpha v^1 + (1 - \alpha)v^2)}{t} \\ &\geq \lim_{t \rightarrow 0^+} \frac{\alpha g(x + td, v^1) + (1 - \alpha)g(x + td, v^2) - \alpha g(x, v^1) - (1 - \alpha)g(x, v^2)}{t} \\ &= \alpha \lim_{t \rightarrow 0^+} \frac{g(x + td, v^1) - g(x, v^1)}{t} + (1 - \alpha) \lim_{t \rightarrow 0^+} \frac{g(x + td, v^2) - g(x, v^2)}{t} \\ &= \alpha g'_x(x, v^1; d) + (1 - \alpha)g'_x(x, v^2; d) \\ &= \alpha g_x^0(x, v^1; d) + (1 - \alpha)g_x^0(x, v^2; d) \\ &\geq \alpha \langle \xi^1, d \rangle + (1 - \alpha) \langle \xi^2, d \rangle \\ &= \langle \alpha \xi^1 + (1 - \alpha)\xi^2, d \rangle. \end{aligned}$$

By the definition of the generalized gradient, then

$$\alpha \xi^1 + (1 - \alpha)\xi^2 \in \partial_x g(x, \alpha v^1 + (1 - \alpha)v^2).$$

Thus  $\alpha \xi^1 + (1 - \alpha)\xi^2 \in \Gamma$ , and so the set  $\Gamma$  is convex.

**The set  $\Gamma$  is weak\* compact.** We first prove that  $\Gamma$  is closed. In fact, let  $\xi^k \in \Gamma, k \geq 1$ , be a sequence weak\* converging to  $\xi$  in  $X^*$ . Then, by the definition, there exists  $v^k \in \mathcal{V}(x)$  such that  $\xi^k \in \partial_x g(x, v^k)$ . Since the set  $\mathcal{V}(x)$  is sequentially compact, without loss of generality, we may assume that  $v^k$  converges to  $v \in \mathcal{V}(x)$ . It follows from Assumption (A4) that  $\xi \in \partial_x g(x, v)$  and so  $\xi \in \Gamma$ . Hence  $\Gamma$  is weak\* closed.

Now it follows from Assumption (A2) that

$$\begin{aligned} |g_x^0(x, v; d)| &\leq \limsup_{h \rightarrow 0, t \rightarrow 0^+} \left| \frac{g(x+h+td, v) - g(x+h, v)}{t} \right| \\ &\leq L\|d\|, \quad \forall v \in \mathcal{V}, \quad \forall d \in X. \end{aligned}$$

By the definition of the generalized gradient, we get

$$\|\xi\| \leq L, \quad \forall \xi \in \partial_x g(x, v) \text{ and } \forall v \in \mathcal{V}.$$

This inequality, together with Alaoglu's theorem, implies that  $\Gamma$  is weak\* compact.

(iii) We first prove the inclusion

$$\Gamma \subset \partial\psi(x).$$

In fact, we have for any  $v \in \mathcal{V}(x)$  and  $d \in X$ ,

$$\begin{aligned} g_x^0(x, v; d) = g'_x(x, v; d) &= \lim_{t \rightarrow 0^+} \frac{g(x+td, v) - g(x, v)}{t} \\ &\leq \limsup_{t \rightarrow 0^+} \frac{\psi(x+td) - \psi(x)}{t} \leq \psi^0(x; d). \end{aligned}$$

By Theorem 2.1(iii), therefore

$$\langle \xi, d \rangle \leq \psi^0(x; d) \quad \text{for all } \xi \in \partial_x g(x, v).$$

Consequently, we derive

$$\Gamma \subset \partial\psi(x).$$

Now we prove the inverse inclusion. Let  $\xi$  belong to  $\partial\psi(x)$ ; that is

$$\psi^0(x; d) \geq \langle \xi, d \rangle \quad \text{for all } d \in X.$$

We deduce from Lemma 2.3 that

$$\max\{\langle \eta, d \rangle \mid \eta \in \partial_x g(x, v), v \in \mathcal{V}(x)\} = \psi^0(x; d) \geq \langle \xi, d \rangle \quad \text{for all } d \in X.$$

Consequently,

$$\max\{\langle \eta - \xi, d \rangle \mid \eta \in \partial_x g(x, v), v \in \mathcal{V}(x)\} \geq 0 \quad \text{for all } d \in X.$$

In particular, we get

$$\inf_{\|d\| \leq 1} \max_{\eta \in \Gamma} \langle \eta - \xi, d \rangle \geq 0.$$

By Item (ii), the set  $\Gamma$  is convex and weak\* compact. Next we apply the ‘‘lop-sided’’ minimax theorem (see, for example, [1]) to deduce the existence of an element  $\bar{\eta}$  of  $\Gamma$  such that

$$\inf_{\|d\| \leq 1} \langle \bar{\eta} - \xi, d \rangle = \inf_{\|d\| \leq 1} \max_{\eta \in \Gamma} \langle \eta - \xi, d \rangle \geq 0.$$

Note that

$$\inf_{\|d\| \leq 1} \langle \bar{\eta} - \xi, d \rangle = -\|\bar{\eta} - \xi\|^*.$$

Therefore,

$$-\|\bar{\eta} - \xi\|^* = \inf_{\|d\| \leq 1} \max_{\eta \in \Gamma} \langle \eta - \xi, d \rangle \geq 0,$$

which implies that  $\bar{\eta} = \xi$ , and we obtain the inclusion  $\partial\psi(x) \subset \Gamma$ . □

### 3. Necessary conditions for optimality in Lagrange multiplier form

#### 3.1. Necessary optimality theorems

Let  $X$  be a Banach space, and let  $f, \psi_i: X \rightarrow \mathbb{R}, i = 1, \dots, m$ , be locally Lipschitz functions. We consider the following nonsmooth optimization problem:

$$\begin{aligned} \text{(P)} \quad & \text{Minimize} && f(x) \\ & \text{subject to} && \psi_i(x) \leq 0, \quad i = 1, \dots, m. \end{aligned}$$

Let

$$\mathcal{C} := \{x \in X \mid \psi_i(x) \leq 0, i = 1, \dots, m\}$$

be the constrain set of the problem (P).

**Definition 3.1.** We say that  $x^*$  is a *local minimizer* of (P) if  $x^* \in \mathcal{C}$  and there exists a neighbourhood  $U$  of  $x^*$  such that for all  $x \in \mathcal{C} \cap U$ , one has  $f(x) \geq f(x^*)$ .

Let us decompose  $I := \{1, \dots, m\}$  into two index sets  $I = I_1(x) \cup I_2(x)$ , where  $I_1(x) = \{i \in I \mid \psi_i(x) = 0\}$  and  $I_2(x) = I \setminus I_1(x)$ .

The first optimality relation in general form is the following. This result is in the spirit of those of J. B. Hiriart-Urruty [13]:

**Theorem 3.1.** *If  $x^* \in \mathcal{C}$  is a local minimizer of (P), then the following system*

$$\begin{cases} f^0(x^*; d) < 0, \\ \psi_i^0(x^*; d) < 0, \quad \forall i \in I_1(x^*), \end{cases}$$

has no solution  $d \in X$ .

*Proof.* Suppose not, then there exists  $d \in X$  such that

$$f^0(x^*; d) < 0 \quad \text{and} \quad \psi_i^0(x^*; d) < 0, \quad \forall i \in I_1(x^*).$$

We have

$$\begin{aligned} \limsup_{t \rightarrow 0+} \frac{f(x^* + td) - f(x^*)}{t} &\leq \limsup_{h \rightarrow 0, t \rightarrow 0+} \frac{f(x^* + h + td) - f(x^* + h)}{t} \\ &= f^0(x^*; d) < 0 \end{aligned}$$

(Note that the function  $f$  is locally Lipschitz). By definition of upper limit, there exists a sequence  $t_k > 0$  with  $\lim_{k \rightarrow \infty} t_k = 0$  such that

$$\lim_{k \rightarrow \infty} \frac{f(x^* + t_k d) - f(x^*)}{t_k} = \limsup_{t \rightarrow 0+} \frac{f(x^* + td) - f(x^*)}{t} < 0.$$

Consequently, there exists an integer number  $K_1 > 0$  such that

$$f(x^* + t_k d) < f(x^*), \quad \forall k \geq K_1.$$

On the other hand, we have for all  $i \in I_1(x^*)$ ,

$$\begin{aligned} \limsup_{k \rightarrow \infty} \frac{\psi_i(x^* + t_k d) - \psi_i(x^*)}{t_k} &\leq \limsup_{h \rightarrow 0, t \rightarrow 0^+} \frac{\psi_i(x^* + h + td) - \psi_i(x^* + h)}{t} \\ &= \psi_i^0(x^*; d) < 0. \end{aligned}$$

Then there exists an integer number  $K_2 > 0$  such that

$$\psi_i(x^* + t_k d) < \psi_i(x^*) = 0, \quad \forall i \in I_1(x^*) \quad \text{and} \quad \forall k \geq K_2.$$

Moreover, as  $\psi_i(x^*) < 0$  for all  $i \in I_2(x^*)$ , we see that there exists an integer number  $K_3 > 0$  such that

$$\psi_i(x^* + t_k d) < 0, \quad \forall i \in I_2(x^*) \quad \text{and} \quad \forall k \geq K_3.$$

Therefore,  $x^* + t_k d \in \mathcal{C}$  for all  $k \geq \max\{K_1, K_2, K_3\}$ , which contradicts the optimality of  $x^*$ .  $\square$

**Definition 3.2.** We define a *Nonsmooth Mangasarian-Fromovitz constraint qualification* (NMFCQ) at  $x \in \mathcal{C}$  as follows:

$$\exists d \in X \text{ such that } \psi_i^0(x; d) < 0, \quad \forall i \in I_1(x).$$

The following result is a necessary condition for optimality in Lagrange multiplier form originally due to Clarke [9]. But by using Theorem 3.1, we give self-contained proof for the result.

**Theorem 3.2.** *Let  $x^* \in \mathcal{C}$  be a local minimizer of (P). Then, there exist  $\lambda_i \geq 0, i = 0, \dots, m$ , with  $\sum_{i=0}^m \lambda_i = 1$ , such that*

$$\begin{aligned} 0 &\in \lambda_0 \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial \psi_i(x^*), \\ 0 &= \lambda_i \psi_i(x^*), \quad i = 1, \dots, m. \end{aligned}$$

Moreover, if we further assume that the *Nonsmooth Mangasarian-Fromovitz constraint qualification* (NMFCQ) at  $x^*$  holds, then

$$\begin{aligned} 0 &\in \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial \psi_i(x^*), \\ 0 &= \lambda_i \psi_i(x^*), \quad i = 1, \dots, m. \end{aligned}$$

*Proof.* As  $x^* \in \mathcal{C}$  is a local minimizer of (P), there exists a neighbourhood  $U$  of  $x^*$  such that

$$\psi_i(x) \leq 0, \quad i = 1, \dots, m, \quad x \in U, \quad \implies \quad f(x) \geq f(x^*).$$

Note that if  $I_1(x^*) = \emptyset$ , then  $\psi_i(x^*) < 0$  for all  $i = 1, \dots, m$ , and so the theorem holds with  $\lambda_0 = 1$  and  $\lambda_i = 0, i = 1, \dots, m$ .

Without loss of generality, we may assume that  $I_1(x^*) \neq \emptyset$ .



In view of Theorem 2.1(ii), the functions  $d \mapsto f^0(x^*, d)$  and  $d \mapsto \psi_i^0(x^*, d)$  are convex. By Theorem 3.1, it follows from Gordan's alternative theorem in [20] that there exist  $\lambda_0 \geq 0$  and  $\lambda_i \geq 0, i \in I_1(x^*)$ , not all zero, such that

$$\lambda_0 f^0(x^*; d) + \sum_{i \in I_1(x^*)} \lambda_i \psi_i^0(x^*; d) \geq 0, \quad \text{for all } d \in X.$$

It is immediate from Theorems 2.1(iii) that

$$\begin{aligned} & \lambda_0 \max\{\langle \xi_0, d \rangle \mid \xi_0 \in \partial f(x^*)\} \\ & + \sum_{i \in I_1(x^*)} \lambda_i \max\{\langle \xi_i, d \rangle \mid \xi_i \in \partial \psi_i(x^*)\} \geq 0 \quad \text{for all } d \in X. \end{aligned}$$

Hence

$$\max_{\substack{\xi_0 \in \partial f(x^*) \\ \xi_i \in \partial \psi_i(x^*)}} \langle \lambda_0 \xi_0 + \sum_{i \in I_1(x^*)} \lambda_i \xi_i, d \rangle \geq 0, \quad \text{for all } d \in \mathbb{B},$$

where  $\mathbb{B} := \{d \in X \mid \|d\| \leq 1\}$  is the unit ball of  $X$ . This is equivalent to

$$\inf_{d \in \mathbb{B}} \max_{\substack{\xi_0 \in \partial f(x^*) \\ \xi_i \in \partial \psi_i(x^*)}} \langle \lambda_0 \xi_0 + \sum_{i \in I_1(x^*)} \lambda_i \xi_i, d \rangle \geq 0.$$

On the other hand, by Theorem 2.1(i), the sets  $\partial f(x^*)$  and  $\partial \psi_i(x^*)$  are convex and weak\* compact. Hence, by the ‘‘lop-sided’’ minimax theorem (see, for example, [1]), there exist elements  $\bar{\xi}_0 \in \partial f(x^*)$  and  $\bar{\xi}_i \in \partial \psi_i(x^*), i \in I_1(x^*)$ , such that

$$\inf_{d \in \mathbb{B}} \langle \lambda_0 \bar{\xi}_0 + \sum_{i \in I_1(x^*)} \lambda_i \bar{\xi}_i, d \rangle = \max_{\substack{\xi_0 \in \partial f(x^*) \\ \xi_i \in \partial \psi_i(x^*)}} \inf_{d \in \mathbb{B}} \langle \lambda_0 \xi_0 + \sum_{i \in I_1(x^*)} \lambda_i \xi_i, d \rangle \geq 0.$$

Note that

$$\inf_{d \in \mathbb{B}} \langle \lambda_0 \bar{\xi}_0 + \sum_{i \in I_1(x^*)} \lambda_i \bar{\xi}_i, d \rangle = -\|\lambda_0 \bar{\xi}_0 + \sum_{i \in I_1(x^*)} \lambda_i \bar{\xi}_i\|^*$$

Therefore,

$$-\|\lambda_0 \bar{\xi}_0 + \sum_{i \in I_1(x^*)} \lambda_i \bar{\xi}_i\|^* = \max_{\substack{\xi_0 \in \partial f(x^*) \\ \xi_i \in \partial \psi_i(x^*)}} \inf_{d \in \mathbb{B}} \langle \lambda_0 \xi_0 + \sum_{i \in I_1(x^*)} \lambda_i \xi_i, d \rangle \geq 0.$$

Consequently,  $-\|\lambda_0 \bar{\xi}_0 + \sum_{i \in I_1(x^*)} \lambda_i \bar{\xi}_i\|^* = 0$ , and we get

$$0 \in \lambda_0 \partial f(x^*) + \sum_{i \in I_1(x^*)} \lambda_i \partial \psi_i(x^*).$$

Thus, by letting  $\lambda_i := 0$  for all  $i \in I_2(x^*)$ , we see that

$$\begin{aligned} 0 & \in \lambda_0 \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial \psi_i(x^*), \\ 0 & = \lambda_i \psi_i(x^*), \quad i = 1, \dots, m. \end{aligned}$$

We now assume that the Nonsmooth Mangasarian-Fromovitz constraint qualification (NMFCQ) at  $x^*$  holds. Then  $\lambda_0 > 0$ . In fact, if it is not true then  $\lambda_i \geq 0, i \in I_1(x^*)$ , not all zero, and

$$0 \in \sum_{i \in I_1(x^*)} \lambda_i \partial \psi_i(x^*) \subset \partial \left( \sum_{i \in I_1(x^*)} \lambda_i \psi_i \right) (x^*)$$

(The second relation follows from Theorem 2.1(iv)). Consequently,

$$\sum_{i \in I_1(x^*)} \lambda_i \psi_i^0(x^*; d) \geq 0, \quad \text{for all } d \in X,$$

which contradicts (NMFCQ). Hence we may assume that  $\lambda_0 = 1$ , and so

$$\begin{aligned} 0 &\in \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial \psi_i(x^*), \\ 0 &= \lambda_i \psi_i(x^*), \quad i = 1, \dots, m. \end{aligned}$$

□

### 3.2. Robust necessary optimality theorems

Consider the following nonsmooth robust optimization problem:

$$\begin{aligned} \text{(RP)} \quad &\text{Minimize} && f(x) \\ &\text{subject to} && g_i(x, v_i) \leq 0, \quad \forall v_i \in \mathcal{V}_i, \quad i = 1, \dots, m, \end{aligned}$$

where  $f: X \rightarrow \mathbb{R}$  is a locally Lipschitz function, and  $g_i: X \times \mathcal{V}_i \rightarrow \mathbb{R}$  is a function satisfying Assumptions (A1)-(A4) in Subsection 2.2 with respect to a sequentially compact topological space  $\mathcal{V}_i$ . Let

$$\mathcal{C} := \{x \in X \mid g_i(x, v_i) \leq 0, \forall v_i \in \mathcal{V}_i, i = 1, \dots, m\}$$

be the constrained set of the problem (RP).

**Definition 3.3.** We say that  $x^*$  is a *robust local minimizer* of (RP) if  $x^* \in \mathcal{C}$  and there exists a neighbourhood  $U$  of  $x^*$  such that for all  $x \in \mathcal{C} \cap U$ , one has  $f(x) \geq f(x^*)$ .

For each  $i = 1, \dots, m$ , let  $\psi_i: X \rightarrow \mathbb{R}$  be the function given by

$$\psi_i(x) := \max\{g_i(x, v_i) \mid v_i \in \mathcal{V}_i\}.$$

Then it is clear that

$$\mathcal{C} = \{x \in X \mid \psi_i(x) \leq 0, i = 1, \dots, m\}.$$

Let us decompose  $I := \{1, \dots, m\}$  into two index sets  $I = I_1(x) \cup I_2(x)$ , where  $I_1(x) = \{i \in I \mid \psi_i(x) = 0\}$  and  $I_2(x) = I \setminus I_1(x)$ . For each  $i \in I_1(x)$  we put

$$\mathcal{V}_i(x) := \{v_i \in \mathcal{V}_i \mid g_i(x, v_i) = \psi_i(x)\}.$$

Now we define an *Extended Nonsmooth Mangasarian-Fromovitz constraint qualification* (ENMFCQ) at  $x \in \mathcal{C}$  as follows:

$$\exists d \in X \text{ such that } g_{ix}^0(x, v_i; d) < 0, \forall v_i \in \mathcal{V}_i(x), \forall i \in I_1(x),$$

where  $g_{ix}^0(x, v_i; d)$  denotes the generalized directional derivative of  $g_i$  with respect to  $x$ .

The following result is a nonsmooth generalization of robust KKT necessary optimality theorem originally due to Jeyakumar, Li and Lee [17].

**Theorem 3.3.** *In addition to the basic assumptions (A1)-(A4), suppose that  $\mathcal{V}_i$  is convex, and that the function  $g_i(x, \cdot)$  is concave on  $\mathcal{V}_i$ , for each  $x \in X$  and for each  $i = 1, \dots, m$ . If  $x^* \in \mathcal{C}$  be a robust local minimizer of (RP), then there exist  $v_i \in \mathcal{V}_i(x^*)$  and  $\lambda_i \geq 0, i = 1, \dots, m$ , with  $\sum_{i=0}^m \lambda_i = 1$ , such that*

$$\begin{aligned} 0 &\in \lambda_0 \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial_x g_i(x^*, v_i), \\ 0 &= \lambda_i g_i(x^*, v_i), \quad i = 1, \dots, m. \end{aligned}$$

Moreover, if we further assume that the *Extended Nonsmooth Mangasarian-Fromovitz constraint qualification* (ENMFCQ) at  $x^*$  holds, then

$$\begin{aligned} 0 &\in \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial_x g_i(x^*, v_i), \\ 0 &= \lambda_i g_i(x^*, v_i), \quad i = 1, \dots, m. \end{aligned}$$

*Proof.* By Theorem 3.2, there exist  $\lambda_i \geq 0, i = 0, \dots, m$ , with  $\sum_{i=0}^m \lambda_i = 1$ , such that

$$\begin{aligned} 0 &\in \lambda_0 \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial \psi_i(x^*), \\ 0 &= \lambda_i \psi_i(x^*), \quad i = 1, \dots, m. \end{aligned}$$

On the other hand, it follows from Theorem 2.4(iii) that

$$\partial \psi_i(x^*) = \{\xi_i \mid \exists v_i \in \mathcal{V}_i(x^*) \text{ such that } \xi_i \in \partial_x g_i(x^*, v_i)\}, \quad i = 1, \dots, m.$$

Therefore there are  $v_i \in \mathcal{V}_i(x^*)$  satisfying the following conditions

$$\begin{aligned} 0 &\in \lambda_0 \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial_x g_i(x^*, v_i), \\ 0 &= \lambda_i g_i(x^*, v_i), \quad i = 1, \dots, m. \end{aligned}$$

We now assume that the *Extended Nonsmooth Mangasarian-Fromovitz constraint qualification* (ENMFCQ) at  $x^*$  holds. Then  $\lambda_0 > 0$ . In fact, if it is not true then  $\lambda_i \geq 0, i \in I_1(x^*)$ , not all zero, and

$$0 \in \sum_{i \in I_1(x^*)} \lambda_i \partial_x g_i(x^*, v_i) \subset \partial_x \left( \sum_{i \in I_1(x^*)} \lambda_i g_i \right) (x^*, v_i)$$

(The second relation follows from Theorem 2.1(iv)). Consequently,

$$\sum_{i \in I_1(x^*)} \lambda_i g_{ix}^0(x^*, v_i; d) \geq 0, \quad \text{for all } d \in X,$$

which contradicts (ENMFCQ). Hence we may assume that  $\lambda_0 = 1$ , and so

$$\begin{aligned} 0 &\in \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial_x g_i(x^*, v_i), \\ 0 &= \lambda_i g_i(x^*, v_i), \quad i = 1, \dots, m. \end{aligned} \quad \square$$

*Remark 3.4.* (i) Let us consider the case  $f(x) := \max_{v_0 \in \mathcal{V}_0} g_0(x, v_0)$ , where  $\mathcal{V}_0$  is a sequentially compact topological space and  $g_0: X \times \mathcal{V}_0 \rightarrow \mathbb{R}$  is a function. Then, by the same method in the proof of Theorem 3.3, we get the following: *In addition to the basic assumptions (A1)-(A4), suppose that  $\mathcal{V}_i$  is convex, and that  $g_i(x, \cdot)$  is concave on  $\mathcal{V}_i$ , for each  $x \in U$ , and for each  $i = 0, 1, \dots, m$ . Let  $x^* \in \mathcal{C}$  be a robust local minimizer of (RP). Then there exist  $v_i \in \mathcal{V}_i(x^*)$  and  $\lambda_i \geq 0$ ,  $\sum_{i=0}^m \lambda_i = 1$ , such that*

$$\begin{aligned} 0 &\in \lambda_0 \partial_x g_0(x^*, v_0) + \sum_{i=1}^m \lambda_i \partial_x g_i(x^*, v_i), \\ 0 &= \lambda_i g_i(x^*, v_i), \quad i = 1, \dots, m. \end{aligned}$$

(ii) Theorem 3.3 can be weakened as follows: *In addition to the basic assumptions (A1)-(A4), suppose that for each  $i \in I_1(x^*)$ , the set  $\partial_x g_i(x^*, \mathcal{V}_i(x^*))$  is convex. If  $x^* \in \mathcal{C}$  is a robust local minimizer of (RP), then there exist  $v_i \in \mathcal{V}_i(x^*)$  and  $\lambda_i \geq 0$ ,  $\sum_{i=1}^m \lambda_i = 1$ , such that*

$$\begin{aligned} 0 &\in \lambda_0 \partial f(x^*) + \sum_{i=1}^m \lambda_i \partial_x g_i(x^*, v_i), \\ 0 &= \lambda_i g_i(x^*, v_i), \quad i = 1, \dots, m. \end{aligned}$$

#### 4. Examples

Now we give an example showing that the assumption that  $\mathcal{V}_i$  is convex is essential for Theorem 3.3.

**Example 4.1.** Let  $x := (x_1, x_2) \in \mathbb{R}^2$  and  $v := (v_1, v_2) \in \mathcal{V} := \{v \in \mathbb{R}^2 \mid v_1^2 + v_2^2 = 1, v_1 \leq 0 \text{ or } v_2 \leq 0\}$ . It is clear that the set  $\mathcal{V}$  is non-convex. Consider the functions

$$\begin{aligned} f(x) &:= -x_1 - x_2, \\ g(x, v) &:= x_1 v_1 + x_2 v_2 - 1. \end{aligned}$$

Let

$$\mathcal{C} := \{x \in \mathbb{R}^2 \mid \psi(x) \leq 0\},$$

where

$$\psi(x) := \max_{v \in \mathcal{V}} g(x, v) = \begin{cases} \|x\| - 1 & \text{if } x_1 \leq 0 \text{ or } x_2 \leq 0, \\ x_2 - 1 & \text{if } 0 \leq x_1 \leq x_2, \\ x_1 - 1 & \text{if } 0 \leq x_2 \leq x_1. \end{cases}$$

Then  $x^* := (1, 1) \in \mathcal{C}$  is an optimal solution of the problem:

$$\begin{aligned} \min & f(x) \\ \text{subject to} & g(x, v) \leq 0, \quad \forall v \in \mathcal{V}. \end{aligned}$$

We have  $\mathcal{V}(x^*) = \{(1, 0), (0, 1)\}$ . Moreover, it is easy to check that (ENMFCQ) holds at  $x^*$  and there no exist  $v \in \mathcal{V}(x^*)$  and  $\lambda_0, \lambda_1 \geq 0$ , with  $\lambda_0 + \lambda_1 = 1$ , such that

$$\begin{aligned} \lambda_0 \nabla f(x^*) + \lambda_1 \nabla_x g(x^*, v) &= 0, \\ \lambda_1 g(x^*, v) &= 0. \end{aligned}$$

where  $\nabla_x g(x, v)$  denotes the gradient of  $g$  with respect to  $x$ .

Also, we present an example illustrating that the assumption that  $g_i(x, \cdot)$  is concave is necessary for Theorem 3.3.

**Example 4.2.** Let  $x := (x_1, x_2) \in \mathbb{R}^2$  and  $v \in \mathcal{V} := [\frac{\pi}{2}, 2\pi] \subset \mathbb{R}$ . Consider the functions

$$\begin{aligned} f(x) &:= -x_1 - x_2 \\ g(x, v) &:= x_1 \cos v + x_2 \sin v - 1. \end{aligned}$$

It is clear that  $g(x, \cdot)$  is not concave, if  $(x_1, x_2) = (1, 0)$ .

Let

$$\mathcal{C} := \{x \in \mathbb{R}^2 \mid \psi(x) \leq 0\},$$

where

$$\psi(x) := \max_{v \in \mathcal{V}} g(x, v) = \begin{cases} \|x\| - 1 & \text{if } x_1 \leq 0 \text{ or } x_2 \leq 0, \\ x_2 - 1 & \text{if } 0 \leq x_1 \leq x_2, \\ x_1 - 1 & \text{if } 0 \leq x_2 \leq x_1. \end{cases}$$

Then  $x^* := (1, 1) \in \mathcal{C}$  is an optimal solution of the problem:

$$\begin{aligned} \min & f(x) \\ \text{subject to} & g(x, v) \leq 0, \quad \forall v \in \mathcal{V}. \end{aligned}$$

We have  $\mathcal{V}(x^*) = \{\frac{\pi}{2}, 2\pi\}$ . Moreover, it is easy to check that (ENMFCQ) holds at  $x^*$  and there no exist  $v \in \mathcal{V}(x^*)$  and  $\lambda_0, \lambda_1 \geq 0$ , with  $\lambda_0 + \lambda_1 = 1$ , such that

$$\begin{aligned} \lambda_0 \nabla f(x^*) + \lambda_1 \nabla_x g(x^*, v) &= 0, \\ \lambda_1 g(x^*, v) &= 0. \end{aligned}$$

Now we give an example illustrating Remark 3.4(ii). Here it is worth while noticing that the function  $g(x, \cdot)$  in the example is not concave.

**Example 4.3.** Let  $x := (x_1, x_2) \in \mathbb{R}^2$  and  $v \in \mathcal{V} := [0, 2\pi] \subset \mathbb{R}$ . Consider the functions

$$\begin{aligned} f(x) &:= -x_1 \\ g(x, v) &:= x_1 \cos v + x_2 \sin v - 1. \end{aligned}$$

It is clear that  $g(x, \cdot)$  is not concave, if  $(x_1, x_2) = (1, 0)$ .

Let

$$\mathcal{C} := \{x \in \mathbb{R}^2 \mid \psi(x) \leq 0\},$$

where

$$\psi(x) := \max_{v \in \mathcal{V}} g(x, v) = \|x\| - 1.$$

Then  $x^* := (1, 0) \in \mathcal{C}$  is an optimal solution of the problem:

$$\begin{aligned} \min & f(x) \\ \text{subject to} & g(x, v) \leq 0, \quad \forall v \in \mathcal{V}. \end{aligned}$$

We have  $\mathcal{V}(x^*) = \{0, 2\pi\}$  and  $\nabla_x g(x^*, \mathcal{V}(x^*)) = \{(1, 0)\}$ . Moreover, it is easy to check that (ENMFCQ) holds at  $x^*$  and if we let  $\lambda_0 = \lambda_1 := \frac{1}{2}$  then

$$\begin{aligned} \lambda_0 \nabla f(x^*) + \lambda_1 \nabla_x g(x^*, v) &= 0, \\ \lambda_1 g(x^*, v) &= 0. \end{aligned}$$

for all  $v \in \mathcal{V}(x^*)$ .

**Acknowledgments.** This research was performed while the second author had been staying at Department of Applied Mathematics, Pukyong National University, Busan, Korea. He wish to thank the first author and other professors in the department for hospitality and support.

## References

- [1] J. P. Aubin and I. V. Ekeland, *Applied Nonlinear Analysis*, John Wiley & Sons, Inc. 1984.
- [2] A. Beck and A. Ben-Tal, *Duality in robust optimization: primal worst equals dual best*, Oper. Res. Lett. **37** (2009), no. 1, 1–6.
- [3] A. Ben-Tal and A. Nemirovski, *Robust optimization methodology and applications*, Math. Program. **92** (2002), no. 3, Ser B, 453–480.
- [4] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*, Princeton Series in Applied Mathematics, 2009.
- [5] A. Ben-Tal and A. Nemirovski, *Selected topics in robust convex optimization*, Math. Program. **112** (2008), no. 1, Ser. B, 125–158.
- [6] D. Bertsimas and D. Brown, *Constructing uncertainty sets for robust linear optimization*, Oper. Res. **57** (2009), no. 6, 1483–1495.
- [7] D. Bertsimas, D. Pachamanova, and M. Sim, *Robust linear optimization under general norms*, Oper. Res. Lett. **32** (2004), no. 6, 510–516.
- [8] F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.
- [9] ———, *A new approach to Lagrange multipliers*, Math. Oper. Res. **1** (1976), no. 2, 165–174.
- [10] ———, *Generalized gradients of Lipschitz functionals*, Adv. in Math. **40** (1981), no. 1, 52–67.

- [11] ———, *Optimization and Nonsmooth Analysis*, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.
- [12] J. M. Danskin, *The Theory of Max-min and Its Application to Weapons Allocation Problems*, Springer-Verlag, New York, 1967.
- [13] J. B. Hiriart-Urruty, *Tangent cones, generalized gradients and mathematical programming in Banach spaces*, Math. Oper. Res. **4** (1979), no. 1, 79–97.
- [14] V. Jeyakumar and G. Y. Li, *Characterizing robust set containments and solutions of uncertain linear programs without qualifications*, Oper. Res. Lett. **38** (2010), no. 3, 188–194.
- [15] V. Jeyakumar and G. Y. Li, *Robust Farkas' lemma for uncertain linear systems with applications*, Positivity **15** (2011), no. 2, 331–342.
- [16] ———, *Strong duality in robust convex programming: complete characterizations*, SIAM J. Optim. **20** (2010), no. 6, 3384–3407.
- [17] V. Jeyakumar, G. Y. Li, and G. M. Lee, *Robust duality for generalized convex programming problems under data uncertainty*, Nonlinear Anal. **75** (2012), no. 3, 1362–1373.
- [18] G. Lebourg, *Generic differentiability of Lipschitzian functions*, Trans. Amer. Math. Soc. **256** (1979), 125–144.
- [19] G. Y. Li, V. Jeyakumar, and G. M. Lee, *Robust conjugate duality for convex optimization under uncertainty with application to data classification*, Nonlinear Anal. **74** (2012), no. 6, 2327–2341.
- [20] Olvi L. Mangasarian, *Nonlinear Programming*, Classics in Applied Mathematics, 10. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

GUE MYUNG LEE  
DEPARTMENT OF APPLIED MATHEMATICS  
PUKYONG NATIONAL UNIVERSITY  
BUSAN 608-737, KOREA  
*E-mail address:* gmlee@pknu.ac.kr

PHẠM TIẾN SƠN  
CENTER OF RESEARCH AND DEVELOPMENT  
DUY TAN UNIVERSITY  
K7/25, QUANG TRUNG, DANANG, VIETNAM  
AND  
DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF DALAT  
1, PHU DONG THIEN VUONG, DALAT, VIETNAM  
*E-mail address:* sonpt@dlu.edu.vn