NONSMOOTH ANALYSIS AND OPTIMIZATION

Huang Liren

Thesis

Submitted to the Faculty of the Graduate School of

The Chinese University of Hong Kong

(Division of Mathematics)

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

May 1993

Acknowledgement

I would like to express my deepest gratitude to my supervisors Dr. K. F. Ng and Dr. W. L. Chan for their kind supervision and discussions during the course of my Ph. D. research work at the Chinese University of Hong Kong.

Contents

Abstract

A representation result of Cominetti and Correa's generalized second-order directional derivative is given in chapter 2 and then applied to obtain a Taylor theorem type result. A conjecture made by Cominetti and Correa concerning functions of the form

$$\max_{1 \le i \le n} g_i(x)$$

is proved under a strengthened assumption, but not true otherwise.

In chapter 3, we generalize and sharpen R. W. Chaney's results on unconstrained and constraïned second-order necessary and sufficient optimality conditions for general Lipschitz functions without the semismooth assumption.

Introduction

Since the pioneering works of F. Clarke and B. N. Pshenichnyi [4, 6][1] generalized directional derivatives have been studied, and successfully applied in various fields especially in optimization and control theory. The study of generalized *second-order* directional derivatives with its applications in the optimization theory is more recent.

In chapter 2, we consider a generalized second-order directional derivative defined by Cominetti and Correa [5]:

$$f^{\infty}(x;u,v) := \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{1}{ts}\left\{f(y+tu+sv) - f(y+tu) - f(y+sv) + f(y)\right\},$$

where f is a continuous function of a locally convex space X into $\mathbb{R}$ and $x, u, v \in X$. In Proposition 2.2.4 we represent $f^{\infty}(x;u,v)$ in the form of upper limit of the rates of changes of the lower or upper Dini directional derivatives:

$$f^{\infty}(x;u,v) = \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{D_{+}f(y+tv;u) - D_{+}f(x;u)}{t}$$

$$= \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{D^{+}f(y+tv;u) - D^{+}f(x;u)}{t}.$$

This result enables us to establish a generalized second-order Taylor expansions (Theorem 2.4.2 and 2.4.3) for nonsmooth functions. These extend the corresponding results of Cominetti and Correa who assume the function f is C^1. In section 6 of this chapter, we apply our results to a large class of functions (for example, convex and concave

1 References for this Introduction are listed in page 5. Separated references for each subsequent chapter of this thesis are listed at the end of that chapter.

functions) which are not covered by [5, Prop. 4.1]. Some applications to optimization theory are presented in section 7. In [5], a conjecture was made about the possible validity of

$$h^{\infty}(x;u,v) = \max_{1 \leq i \leq n} D^2 g_i(x;u,v),$$

where each g_i is C^2 and $D^2 g_i$ denotes the second-order directional derivative. In section 3 of this chapter, we provide an example [Example 2.3.3] which shows that the conjecture is incorrect. An affirmative answer is given in Corollary 2.3.5 and Corollary 2.3.7 under some strengthened conditions.

In chapter 3, we consider the necessary and sufficient conditions for optimal solutions in the following nonsmooth optimization problems $(\mathbb{P}_1)$ or $(\mathbb{P}_2)$:

(i) Problem $(\mathbb{P}_1)$ minimize $f(x)$

(ii) Problem $(\mathbb{P}_2)$ minimize $f(x)$

subject to $g_i(x) \leq 0$ for $i = 1, 2, \cdots, m$;

$g_i(x) = 0$ for $i = m+1, \cdots, m+p$,

where $x \in X$ and X is a normed space and f, g_i, $i = 1, 2, \cdots, m+p$, are locally Lipschitz functions of an open subset W of X into $\mathbb{R}$. In recent years, many of results on the above problems have been presented. A common feature of these results is the use of various kinds of conditions such as semismoothness, regularity or convexity to replace the assumption of differentiability. In order to weaken the above additional conditions, we make use Chaney's generalized lower and upper second-order directional derivatives introduced in [1], [2] (for definition see Definition 3.1.3). Note that if $\bar{x}$ is a local minimum point of f then the lower Dini directional derivative $D_+ f(\bar{x};u) \geq 0$ at each direction u. If $D_+f(\bar{x};u) = 0$ for some u, then we show (Theorem 3.2.2) that 0 belongs to Chaney's subdifferential $\partial_u f(\bar{x})$ and so $f''_-(\bar{x};0,u)$ is meaningful and $f''_-(\bar{x};0,u) \geq 0$. This result was proved by Chaney in [1] in the special case when $X = \mathbb{R}^n$ under an additional semismooth assumption of f. Now let T be the set of all vectors $w = (w_0, w_1, \cdots, w_{m+p}) \in \mathbb{R}^{m+p}$ with $\sum_{i=0}^{m+p} (w_i)^2 = 1$ and $w_i \geq 0$ for $i = 0, 1, \cdots, m$. Let L(x,w) denote the Lagrangean function on W × T defined by

$$L(x,w) := w_0 f(x) + \sum_{i=1}^{m+p} w_i g_i(x)$$

and define the function

$$G(x,f) := \max \{L(x,w) - w_0 f(\bar{x}) : w \in T\}$$

as in [1] and [2]. Then, for $\bar{x}$ to be a local minimum point of f for problem $(\mathbb{P}_2)$, the following conditions are shown to be necessary (Theorem 3.3.6):

$$\text{(i) } G''_-(\bar{x},0,u) \geq 0, \text{ and (ii) } L''_+(\bar{x},\bar{w},0,u) \geq 0$$

for some Lagrange multiplier $\bar{w}$ in $M_u(\bar{x})$ whenever $D_+G(\cdot,f)(\bar{x};u) = 0$. Further, for $X = \mathbb{R}^n$, the complementary results on sufficient optimality conditions for unconstrained/constrained problems are obtained (Theorem 3.2.8, Theorem 3.4.2 and Theorem 3.4.8) and thereby not only the related results in [2] are generalized (to not necessarily semismooth functions) but also the conclusions are considerably sharpened.

References

[1] R. W. Chaney, Second-order necessary conditions in constrained semismooth optimization, SIAM J. Control and Optim. Vol. 25, No. 4(1987), 1072-1081.

[2] R. W. Chaney, Second-order sufficient conditions in nonsmooth optimization, Math. Oper. Res. Vol. 13. No. 4 (1988), 660-673.

[3] R. W. Chaney, Second-order necessary conditions in semismooth optimization, Math. Programming 40(1988), 95-109.

[4] F. H. Clarke, Optimization and Nonsmooth Analysis, John Wiley, New York, 1983.

[5] R. Cominetti and R. Correa, A generalized second-order derivative in nonsmooth optimization, SIAM J. Control Optim. 28(4) (1990), 789-809.

[6] B. N. Pshenichnyi, Necessary Conditions for an Extremum, Marcel Dekker, New York, 1983.

[7] W. L. Chan, L. R. Huang and K. F. Ng, On generalized second-order derivatives and Taylor expansions in nonsmooth optimization, to appear in SIAM J. Control and Optim.

[8] L. R. Huang and K. F. Ng, Second-order necessary and sufficient conditions in nonsmooth optimization, submitted to Math. Programming.

Chapter 1. Some elementary results in nonsmooth analysis and optimization

In this chapter, we summarize some preliminary background material necessary for the later chapters. In section 1, we list some computation rules of "lim sup" and "lim inf" in the extended real field which will be made use in computing the generalized directional derivatives. In section 2, we review an elementary result with respect to the directional derivative of the so-called sup-type function

$$h(x):=\max\{f(x,t);\ t\in T\}.$$

In section 3, some elementary results in nonsmooth analysis and optimization theory are recalled for easy reference.

Throughout this chapter, X will be denoted as a locally convex space except when it is mentioned specially. $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty,+\infty\}$ will be the extended real field with the usual operations, order and topology familiar in convex analysis.

1. Some properties for "lim sup" and "lim inf"

Let W be a subset of a locally convex space X with closure denoted by $\overline{W}$. For $x \in \overline{W}$, let $\mathcal{U}_x$ be the neighbourhood system at x restricting to W. The upper limit and lower limit of f at x [5] are defined by

$$\limsup_{y\to x} f(x):=\inf_{U\in\mathcal{U}_x}\ \sup_{y\in U} f(y)$$

and

$$\liminf_{y\to x} f(x) := \sup_{U\in\mathcal{U}_x} \inf_{y\in U} f(y)$$

respectively. Clearly these definitions are unchange if $\mathcal{U}_x$ is replaced by the subfamily consisting of the sets of the form $W \cap V$, where V runs over a neighbourhood base at x. For a net $\{x_t;\ t \in \Gamma\}$ in W convergening to x, where Γ is a direct set, we can also form the upper and lower limits of $f(x_t)$ [5]. They are denoted by

$$\limsup_t f(x_t) := \inf_{T\in\Gamma} \sup_{t\geq T} f(x_t) \text{ and } \liminf_t f(x_t) := \sup_{T\in\Gamma} \inf_{t\geq T} f(x_t).$$

From the definitions, we obtain immediately the following

Lemma 1.1.1. [5] For any $x \in \overline{W}$,

(1) for any net $y_t \in W$ with $y_t \longrightarrow x$, one has

$$\limsup_t f(y_t) \leq \limsup_{y\to x} f(y) \text{ and } \liminf_t f(y_t) \geq \liminf_{y\to x} f(y).$$

(2) there exist nets y_t, $z_s \in W$ with $y_t \longrightarrow x$ and $z_s \longrightarrow x$ such that

$$\lim_t f(y_t) = \limsup_{y\to x} f(y) \text{ and } \lim_s f(z_s) = \liminf_{y\to x} f(y);$$

Lemma 1.1.2. [6, pp. 37] Suppose that f_1 and f_2 are functions of W into $\mathbb{R}$ and $x \in \overline{W}$.

(1) If $\limsup_{y\to x} f_1(y)$ and $\sup_{y\to x} f_2(y)$ are not simultaneously infinity, opposite in sign, then one has

$$\limsup_{y\to x} (f_1 + f_2)(y) \leq \limsup_{y\to x} f_1(y) + \limsup_{y\to x} f_2(y);$$

(2) If $\limsup_{y\to x} f_1(y)$ and $\limsup_{y\to x} f_2(y)$ are not simultaneously infinity, same in sign, then one has

$$\limsup_{y\to x} (f_1 - f_2)(y) \geq \limsup_{y\to x} f_1(y) - \limsup_{y\to x} f_2(y);$$

(3) If $\liminf_{y\to x} f_1(y)$ and $\liminf_{y\to x} f_2(y)$ are not simultaneously

infinity, opposite in sign, then one has

$$\liminf_{y\to x}(f_1+f_2)(y) \geq \liminf_{y\to x} f_1(y) + \liminf_{y\to x} f_2(y);$$

(4) If $\liminf_{y\to x} f_1(y)$ and $\liminf_{y\to x} f_2(y)$ are not simultaneously infinity, same in sign, then one has

$$\liminf_{y\to x}(f_1-f_2)(y) \leq \liminf_{y\to x} f_1(y) - \liminf_{y\to x} f_2(y);$$

(5) The following inequalities hold

$$\liminf_{y\to x}(f_1+f_2)(y) \leq \liminf_{y\to x} f_1(y) + \limsup_{y\to x} f_2(y)$$
$$\leq \limsup_{y\to x}(f_1+f_2)(y)$$

provided the right hand side of the first inequality is not of the form $\infty - \infty$.

2. The directional derivative of the sup-type function

Let T be a compact metric space, $\{f(\cdot,t);\ t \in T\}$ a collection of continuous functions of X into $\mathbb{R}$ and let h be defined by

$$h(x) := \max\{f(x,t);\ t \in T\},\quad (x \in X).$$

Let $x \in X$. By a slight abuse of notations we write $\mathbf{f}(x)$ for the indexed set $(f(\cdot,t))_{t\in T}$ of values of the function $t \mapsto f(x,t)$, and let $I(\mathbf{f}(x))$ denote the set of all t for which $h(x) = f(x,t)$.

Recall that the one side-directional derivative of a function $g: X \longrightarrow \mathbb{R}$ at x in the direction $v \in X$ is defined by

$$g'(x;v) := \lim_{s\downarrow 0} \frac{1}{s}\{g(x+sv) - g(x)\}.$$

Lemma 1.2.1. Suppose that $f(\cdot,\cdot)$ is continuous on $X \times T$. Then h is continuous on X.

Proof: Let $x \in X$. By the definition of h and the compactness of T, for any $y \in X$, there exists t_y such that $h(y) = f(y,t_y)$ and hence that

(1.2.1) $$h(y) - h(x) \le f(y,t_y) - f(x,t_y).$$

Also similarly, take $\lambda \in I(f(x))$, i.e. $h(x) = f(x,\lambda)$; hence

(1.2.2) $$h(y) - h(x) \ge f(y,\lambda) - f(x,\lambda).$$

Now we let $y \longrightarrow x$. By the compactness of T and considering a subnet if necessary, we can assume that $t_y \longrightarrow \tau$. Thus, by the continuity of $f(\cdot,\cdot)$ on $X \times T$, (1.2.1) and (1.2.2) we obtain

$$\limsup_{y} \{h(y) - h(x)\} \le \lim_{y\to x} \{f(y,t_y) - f(x,t_y)\}$$
$$= f(x,\tau) - f(x,\tau) = 0$$

and

$$\liminf_{y} \{h(y) - h(x)\} \ge \lim_{y\to x} \{f(y,\lambda) - f(x,\lambda)\}$$
$$= f(x,\lambda) - f(x,\lambda) = 0.$$

Combining the above two inequalities we see that h is continuous at x.

Lemma 1.2.2. Suppose that $f(\cdot,\cdot)$ is continuous on $X \times T$, $y \longrightarrow x$ in X and $t_y \in I(f(y))$. Let τ be a cluster point of t_y. Then $\tau \in I(f(x))$.

Proof: Take a subnet y_ν of y with $t_{y_\nu} \longrightarrow \tau$. Note that $t_{y_\nu} \in I(f(y_\nu))$ and that $y_\nu \longrightarrow x$ since $y \longrightarrow x$. Since h is continuous on X by Lemma 1.2.1 and $f(\cdot,\cdot)$ is on $X \times T$, we have

$$h(x) = \lim_{\nu} h(y_\nu) = \lim_{\nu} f(y_\nu,t_{y_\nu}) = f(x,\tau)$$

and so $\tau \in I(f(x))$. □

Proposition 1.2.3. Suppose that $f(\cdot,\cdot)$ is continuous on $X \times T$, $x \in X$ and $f(\cdot,t)$ is directionally differentiable at x for each $t \in T$.

Then for any $v \in X$, one has

$$h'(x;v) = \max_{t\in I(f(x))} f'(x,t)v$$

if (1) T is a finite set or (2) $f'(\cdot,\cdot)$ is continuous on $X \times T$.

Further, if X is the finite dimensional space $\mathbb{R}^n$ and (2) holds, then h is locally Lipschitz on $\mathbb{R}^n$.

Proof: Note that $I(f(x))$ is compact subset of T (and finite if T is finite). Thus, in either (1) or (2), there exists $\tau \in I(f(x))$ at which $f'(x,\cdot)v$ attains its maximum on $I(f(x))$:

$$f'(x,\tau)v = \max_{t\in I(f(x))} f'(x,t)v.$$

Then for any $\lambda > 0$, one has

$$f(x,\tau) = h(x) \text{ and } h(x+\lambda v) \geq f(x+\lambda v,\tau)$$

and so

$$\frac{1}{\lambda}\{h(x+\lambda v) - h(x)\}$$
$$\geq \frac{1}{\lambda}\{f(x+\lambda v,\tau) - f(x,\tau)\};$$

consequently

(1.2.3)
$$\liminf_{\lambda\downarrow 0} \frac{1}{\lambda}\{h(x+\lambda v) - h(x)\}$$
$$\geq \liminf_{\lambda\downarrow 0} \frac{1}{\lambda}\{f(x+\lambda v,\tau) - f(x,\tau)\}$$
$$= f'(x,\tau)v = \max_{t\in I(f(x))} f'(x,t)v,$$

where the first equality holds since $f(\cdot,\tau)$ is directionally differentiable at x by assumption. On the other hand, for any $t_\lambda \in I(f(x+\lambda v))$, one has

(1.2.4)
$$\frac{1}{\lambda}\{h(x+\lambda v) - h(x)\}$$
$$= \frac{1}{\lambda}\{f(x+\lambda v,t_\lambda) - f(x,t_\lambda)\} + \frac{1}{\lambda}\{f(x,t_\lambda) - h(x)\}$$
$$\leq \frac{1}{\lambda}\{f(x+\lambda v,t_\lambda) - f(x,t_\lambda)\}.$$

By considering subnets if necessary, we can assume that

(1.2.5)
$$t_\lambda \longrightarrow t_0$$

(hence $t_0 \in I(f(x))$ by Lemma 1.2.2) and

(1.2.6)
$$\limsup_{\lambda \downarrow 0} \frac{1}{\lambda}\{h(x+\lambda v) - h(x)\}$$
$$\leq \limsup_{\lambda \downarrow 0} \frac{1}{\lambda}\{f(x+\lambda v, t_\lambda) - f(x, t_\lambda)\}$$
$$= \lim_{\lambda \downarrow 0} \frac{1}{\lambda}\{f(x+\lambda v, t_\lambda) - f(x, t_\lambda)\}.$$

Now we suppose that (1) holds, that is, T is a finite set. In this case, (1.2.5) ensures that there exist infinitely many of $t_\lambda = t_0$. Thus, (1.2.6) can be rewritten by

$$\limsup_{\lambda \downarrow 0} \frac{1}{\lambda}\{h(x+\lambda v) - h(x)\}$$
$$= \lim_{\lambda \downarrow 0} \frac{1}{\lambda}\{f(x+\lambda v, t_\lambda) - f(x, t_\lambda)\}$$
$$= \lim_{\lambda \downarrow 0} \frac{1}{\lambda}\{f(x+\lambda v, t_0) - f(x, t_0)\}$$
$$= f'(x, t_0)v \leq \max_{t \in I(f(x))} f'(x,t)v$$

where the last equality follows from the directional differentiability of $f(\cdot, t)$ on X for any $t \in T$ and the last inequality holds since $t_0 \in I(f(x))$. Together with (1.2.3) we conclude that

$$h'(x;v) = \lim_{\lambda \downarrow 0} \frac{1}{\lambda}\{h(x+\lambda v) - h(x)\} = \max_{t \in I(f(x))} f'(x,t)v.$$

Next we consider the case (2). By the Mean-Value Theorem (1.2.4) can be rewritten by

$$\frac{1}{\lambda}\{h(x+\lambda v) - h(x)\} \leq f'(x+\delta_\lambda v, t_\lambda)v, \text{ where } \delta_\lambda \in (0,\lambda).$$

By considering a subnet if necessary, we can assume that $t_\lambda \longrightarrow t_0$ and so $t_0 \in I(f(x))$ by part (2) of Lemma 1.2.2 since $t_\lambda \in I(f(x+\lambda v))$. Thus, it follows from the continuity of $f'(\cdot,\cdot)v$ on $X \times T$ that

$$\limsup_{\lambda \downarrow 0} \frac{1}{\lambda}\{h(x+\lambda v) - h(x)\}$$
$$\leq \limsup_{\lambda \downarrow 0} f'(x+\delta_\lambda v, t_\lambda)v$$
$$= f'(x, t_0)v \leq \max_{t \in I(f(x))} f'(x,t)v.$$

Therefore, together with (1.2.3) we also obtain

$$h'(x;v) = \lim_{\lambda\downarrow 0} \frac{1}{\lambda}\{h(x+\lambda v) - h(x)\} = \max_{t\in I(f(x))} f'(x,t)v.$$

For the last assumption of the Proposition 1.2.3, suppose on the contrary that there exist $x_0 \in \mathbb{R}^n$ and sequences y_k, $z_k \in \mathbb{R}^n$ with $\|y_k-x_0\|$, $\|z_k-x_0\| \longrightarrow 0$ such that

$$|h(y_k) - h(z_k)| > K\|y_k-z_k\|,$$

for each k. In general, we can further assume that $h(y_k) \geq h(z_k)$. Let

$$t_k = \|y_k-z_k\|,\ u_k = \frac{y_k-z_k}{\|y_k-z_k\|} \text{ and } h(y_k) = f(y_k,\lambda_k).$$

By considering a subsequences if necessary, we can assume that $u_k \longrightarrow u$ and $\lambda_k \longrightarrow \lambda$. In terms of these, one has

$$K < \frac{h(y_k) - h(z_k)}{\|y_k-z_k\|} \leq \frac{f(z_k+t_k u_k,\lambda_k) - f(z_k,\lambda_k)}{t_k}.$$

and so it follows from the continuity of $f'(\cdot,\cdot)$ and the Mean-Value Theorem that

$$+\infty \leq \lim_{n\to\infty} \frac{f(z_n+t_n u_n,\lambda_n) - f(z_n,\lambda_n)}{t_n} = f'(x_0,\lambda)u.$$

This is a contradiction. □

3. Some results in nonsmooth analysis and optimization

In this section, we give some well-known results in nonsmooth analysis and optimization for convenient background. The details of proofs are referred to, e.g. [1], [2].

Let $f\colon X \longrightarrow \mathbb{R}$ be a function. Recall that the Clarke's directional derivative at x in the direction u is defined by

$$f^0(x;u) := \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{1}{t}\{f(y+tu) - f(y)\} \quad (\leq +\infty)$$

and that the Clarke's subdifferential of f at x is the subset of the topological dual X^* of X defined by

$$\partial f(x) := \{x^* \in X^*;\ \langle x^*, v\rangle \leq f^0(x;v) \text{ for all } v \in X\}.$$

It is clear that $\partial f(x)$ is a w^*-closed convex subset of X^*. Recall that a function f of X into $\mathbb{R}$ is said to be locally Lipschitz at x if there exist a neighbourhood U of x and a continuous seminorm p on X such that

$$|f(y) - f(z)| \leq p(y-z)$$

for all y, z in U. Further, if it is locally Lipschitz at each point of X, then we call it a locally Lipschitz function on X.

The following result was given in [2, pp 25-27] for the case when X is a normed space.

Theorem 1.3.1. [2, pp. 25-27] Suppose that f is a real-valued function of X. Then the following hold:

(1) For any $\alpha \geq 0$ and u, $v \in X$, one has

(i) $f^0(x;\alpha u) = \alpha f^0(x;u)$;

(ii) if $f^0(x;u)$ and $f^0(x;v)$ are not simultaneously infinity, opposite in sign, then

$$f^0(x;u+v) \leq f^0(x;u) + f^0(x;v).$$

(2) $f^0(\cdot,u)$ is upper semicontinuous for any $u \in X$. Furthermore, if f is locally Lipschitz on X, then $f^0(x;\cdot)$ is a locally Lipschitz function for any x and $f^0(\cdot,\cdot)$ is upper semicontinuous.

(3) If X is a normed space and f is a locally Lipschitz function on X, then the multi-function $x \mapsto \partial f(x)$ locally takes values in a w^*-compact convex set, that is, for any $x \in X$, there exist a neighbourhood U of x and a w^*-compact convex subset K of X^* such that

$\partial f(z) \subseteq K$ for all $z \in U$.

Proof: (1) (i) follows from the definition of f^0 and (ii) follows from applying (1) of Lemma 1.1.2 to the following equality:

$$\frac{1}{t}\{f(y+t(u+v)) - f(y)\} = \frac{1}{t}\{f(y+tu+tv) - f(y+tu)\} + \frac{1}{t}\{f(y+tu) - f(y)\}$$

(We note that, $y \longrightarrow x$ if and only if $y+tu \longrightarrow x$ because $t \downarrow 0$).

(2) By definition of f^0, for any $\varepsilon > 0$ there exist an open neighbourhood U of x and $\delta > 0$ such that

$$\frac{1}{t}\{f(y+tu) - f(y)\} < \varepsilon + f^0(x;u) \ (\le +\infty)$$

for all $y \in U$ and $0 < t < \delta$. Since U is an open neighbourhood of each z in U, this implies that

$$f^0(z;u) = \lim \sup_{\substack{y \to z \\ t \downarrow 0}} \frac{1}{t}\{f(y+tu) - f(y)\} \le f^0(x;u) + \varepsilon$$

and so $\lim \sup_{z \to x} f^0(z;u) \le f^0(x;u)$ since ε is arbitrary. This shows the first result. For the second, since f is locally Lipschitz on X, $f^0(x;u)$ is finite for any x and u. By (ii) of this theorem we have

(1.1.1) $$f^0(x;u) - f^0(x;v) \le f^0(x;u-v) = \lim \sup_{\substack{y \to x \\ t \downarrow 0}} \frac{1}{t}\{f(y+t(u-v)) - f(y)\}.$$

For any $x \in X$, applying the Lipschtz condition of f, we can find an open neighbourhood U of x and a continuous seminorm p such that

$$|f(z) - f(w)| \le p(z-w)$$

for all z, $w \in U$. Then, for all u, v in X and $y \in U$ consider $t > 0$ small enough such that $y+t(u-v) \in U$. Then

$$|f(y+t(u-v)) - f(y)| \le tp(u-v).$$

It follows from (1.1.1) (applied to z in place of x) that

$$f^0(z;u) - f^0(z;v) \le p(u-v)$$

valid for all z in U. This proves the second result, and further implies that

$$\limsup_{\substack{z\to x\\ u\to v}} f^0(z;u) \le \limsup_{z\to x} f^0(z;v) + \limsup_{u\to v} p(u-v) \le f^0(x;v) + 0,$$

showing the last assertion of this part.

(3) See [2, proposition 2.1.2]. □

Theorem 1.3.2. [2,Proposition 2.3.3] Suppose that f_1, f_2 are the locally Lipschitz functions of a normed space X into ℝ. Let $f = f_1 + f_2$. Then for any $x \in X$, one has

$$\partial f(x) \subseteq \partial f_1(x) + \partial f_2(x). \quad \square$$

Ekeland's variational principle. [3 or 2, Theorem 7.5.1] Suppose that (X,ρ) is a complete metric space and f is a lower semicontinuous function which is bounded below. For $x \in X$ and $\varepsilon > 0$, if one has

$$f(x) \le \inf\{f(y);\ y \in X\} + \varepsilon,$$

then for any $\lambda > 0$, there exists $x_0 \in X$ such that

(i) $\rho(x_0,x) \le \lambda$;

(ii) $f(x_0) \le f(x)$;

and

(iii) $f(x_0) \le f(y) + \frac{\varepsilon}{\lambda}\rho(y,x_0)$. □

In the following, we give some results in nonsmooth optimization. Theorem 1.3.3 is given in [2, Proposition 2.3.2] in the case when X is a normed space.

Theorem 1.3.3. [2, Proposition 2.3.2] Suppose that f is a function of X into $\mathbb{R}$ and x_0 is a local minimum (resp local maximum) for f, that is, there exists a neighbourhood U of x_0 such that

$$f(x_0) \leq f(x) \quad (\text{resp } f(x) \leq f(x_0))$$

for all $x \in U$. Then

$$f^0(x_0;u) \geq 0$$

for all $u \in X$, equivalently, $0 \in \partial f(x_0)$ by definition of ∂f.

Proof: First we consider the case when x_0 is a local minimum. Then it follows from the definitions that one has

$$0 \leq \limsup_{t \downarrow 0} \frac{1}{t}\{f(x_0+tu) - f(x_0)\} \leq f^0(x_0;u)$$

for any $u \in X$. Secondly, we consider the case of that x_0 is a local maximum. For any $u \in X$, $t > 0$ and $x_t = x_0-tu$, one has $x_t \longrightarrow x_0$ as $t \downarrow 0$; hence

$$f^0(x_0;u) \geq \limsup_{t \downarrow 0} \frac{1}{t}\{f(x_t+tu) - f(x_t)\}$$
$$= \limsup_{t \downarrow 0} \frac{1}{t}\{f(x_0) - f(x_t)\} \geq 0. \qquad \square$$

Recall that the lower Dini directional derivative $D_-f(x;v)$ of $f: X \longrightarrow \mathbb{R}$ at x in the direction v is defined by

$$D_-f(x;v) := \liminf_{t \downarrow 0} \frac{1}{t}\{f(x+tu) - f(x)\}.$$

Ioffe's Proposition. [4, Proposition 1] Let $x_0 \in \mathbb{R}^n$ and f be a locally Lipschitz function of $\mathbb{R}^n$ into $\mathbb{R}$. If $D_-f(x_0;v) \geq 0$ for all v in $\mathbb{R}^n$, then for any $\varepsilon > 0$, there exists $\delta > 0$ such that

(1.3.1) $$f(x_0) < f(x) + \varepsilon\|x-x_0\|$$

for all $0 < \|x-x_0\| \leq \delta$.

Proof: If the desired conclusion is false, then there exist $\varepsilon > 0$ and a sequence $z_k \in \mathbb{R}^n$ with $x_0 \neq z_k$ such that $z_k \longrightarrow x_0$ and

$$f(x_0) \geq f(z_k) + \varepsilon\|z_k - x_0\|.$$

Let $t_k = \|z_k-x_0\|$ and $u_k = \frac{z_k-x_0}{t_k}$. Then by considering a subsequence if necessary, we can assume that u_k converges to some unit vector u. Thus, one has

$$D_-f(x_0;u) \leq \liminf_{k\to\infty} \frac{1}{t_k}\{f(x_0+t_ku) - f(x_0)\}$$
$$\leq \liminf_{k\to\infty} \frac{1}{t_k}\{f(z_k) - f(x_0)\} + \limsup_{k\to\infty} \frac{1}{t_k}\{f(x_0+t_ku) - f(z_k)\} \leq -\varepsilon,$$

where the second inequality follows from (5) of Lemma 1.1.2 and the third inequality follows from (1.3.1) and the Lipschitz condition of f (notice that $f(x_0+t_ku) - f(z_k) \leq Lt_k\|u-u_k\|$ for some Lipschitz constant L). This contradicts the assumption of the proposition. □

We end this section by Clarke's nonsmooth Lagrange multiplier rule. Let X be a normed space and f a locally Lipschitz function of X into $\mathbb{R}$. Further, we let g be a locally Lipschitz function from X into $\mathbb{R}^n$ and C a closed subset of X.

Now we consider the minimization problem with constraint:

(P) minimizing $\{f(x);\ x \in Q\}$,

where $Q: = \{x \in C;\ g(x) \leq 0\}$. Define the distance function $d_C: X \longrightarrow \mathbb{R}$ by

$$d_C(x): = \inf\{\|x-y\|;\ \forall y \in C\}.$$

In terms of this, we have the following Lagrange multiplier rule:

Theorem 1.3.4. [2, Theorem 6.1.1] Suppose that x_0 is a local minimum for the problem (P), that is, there exists a neighbourhood U of x_0 such that for any $x \in U \cap Q$, one has $f(x_0) \leq f(x)$. Then there exists a multiplier $(\lambda, \gamma) \in \mathbb{R} \times \mathbb{R}^n$ with $\lambda, \gamma_i \geq 0$, $1 \leq i \leq n$ and $\lambda + \sum_{i=1}^{n} \gamma_i = 1$ such that

$$\gamma g(x_0) = 0 \text{ and } 0 \leq L^0(x_0;v)$$

for any $v \in X$, where L is the Lagrangian function defined by

$$L(x) := \lambda f(x) + \gamma g(x) + \alpha d_C(x)$$

and α is constant strictly larger than a Lipschitzian constant for both f and g on a neighbourhood of x_0. □

References

[1] J. P. Aubin and I. Ekeland, Applied Nonlinear Analysis, John Wiley & Sons, 1984.

[2] F. H. Clarke, Optimization and Nonsmooth Analysis, John Wiley & Sons, 1983.

[3] I. Ekeland, Nonconvex minimization problems, Bulletin of the American Math. Society, Vol.1, No.3, (1979), 443-474.

[4] A. D. Ioffe, Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps, Nonlinear Analysis, Theory, Methods & Applications, 8(5) (1984), 517-593.

[5] G. Kothe, Toplogical Vector Spaces I, Springer-Verlag, 1969.

[6] H. L. Royden, Real Analysis, Collier-Macmillan, 1972.

[7] H. H. Schaefer, Topological Vector Spaces, Springer-Verlag, GTM, 1970.

Chapter 2. On generalized second-order derivatives and Taylor expansions in nonsmooth optimization

1. Introduction

In this chapter we study a generalized second-order directional derivative $f^{\infty}(x;u,v)$ recently introduced by Cominetti and Correa [2]. Based on their work, in proposition 2.2.4 we represent $f^{\infty}(x;u,v)$ in the form of the upper limit of the rates of changes of the first order Dini directional derivatives. This representation enables us to establish a second-order Taylor expansions (Theorems 2.4.2 and 2.4.3) for nonsmooth functions. These extend the corresponding results of Cominetti and Correa who assumed the C^1-condition. In §6 we apply our results to a large class of functions (e.g. convex and concave functions) which are not covered by [2. Prop. 4.1]. Applications to optimization theory are presented in § 7.

In [2], a conjecture was made about the possible validity of $h^{\infty}(x;u,v) = \max_{1\leq i\leq n} D^2 g_i(x;u,v)$, where each g_i is C^2 and $D^2 g_i$ denotes the second-order directional derivative. Example 2.3.3 shows that the conjecture is incorrect and an affirmative answer is given in Corollary 2.3.5 and Corollary 2.3.7 under strengthened but similar conditions.

2. Dini-directional derivatives, Clarke's directional derivatives and generalized second-order directional derivatives.

Let X be a locally convex space and $f: X \longrightarrow \mathbb{R}$ a function. We consider the extended real field $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ with the usual operations, order and topology familiar in convex analysis. Denote the upper and lower Dini-directional derivatives by

$$D^+f(x;v) := \limsup_{t\downarrow 0} \frac{1}{t} (f(x+tv) - f(x)),$$

$$D_+f(x;v) := \liminf_{t\downarrow 0} \frac{1}{t} (f(x+tv) - f(x)),$$

and the upper and lower Clarke's directional derivatives at x along the direction $v \in X$ by

$$f^0(x;v) := \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{1}{t}(f(y+tv) - f(y))$$

and

$$f_0(x;v) := \liminf_{\substack{y\to x\\ t\downarrow 0}} \frac{1}{t}(f(y+tv) - f(y)).$$

If $X = \mathbb{R}$ and $v = 1$, we shall write $D^+f(x)$ for $D^+f(x;v)$ and similarly for $D_+f(x), f^0(x)$. We shall often make use of the elementary computation rules (see, Proposition 1.1.2) for lim sup and lim inf without further comments, e.g., if $f = f_1 - f_2$ then

$$D_+f(x) \le D_+f_1(x) - D_+f_2(x),$$

provided that the two terms on the right are finite. Also $D_+f(x) \le D^+f_1(x) - D^+f_2(x)$ with similar provisions.

Furthermore, as in [2], [4] and [5] we define the upper and lower generalized second-order directional derivatives at x in the direction $(u,v) \in X \times X$ by

$$f^\infty(x;u,v) := \limsup_{\substack{y\to x\\ t,s\downarrow 0}} \frac{1}{st}\Big\{f(y+tu+sv) - f(y+tu) - f(y+sv) + f(y)\Big\}$$

and

$$f_\infty(x;u,v) := \liminf_{\substack{y\to x\\ t,s\downarrow 0}} \frac{1}{st}\Big\{f(y+tu+sv) - f(y+tu) - f(y+sv) + f(y)\Big\}.$$

If f is a C^2-function, then applying the Mean Value Theorem to the function $F(t) := f(y+sv+tu) - f(x+tu)$, we verify that f^∞ and

f_∞ are just the second-order directional derivatives of f (or see Prop. 2.2.4 later). For the sake of convenience, we list some of their properties in the following Proposition (their proof is given in Lemma A1 of the Appendix). For further properties of f^∞ we refer to [2], [4].

Proposition 2.2.1 [2]. Let $f\colon X \longrightarrow \mathbb{R}$ and $x \in X$. Then:

(i) The map $(u,v) \longmapsto f^\infty(x;u,v)$ is symmetric, and sublinear on each variable separately.

(ii) The map $y \mapsto f^\infty(y;u,v)$ is upper semi-continuous at x for every $(u,v) \in X \times X$.

(iii) $f^\infty(x;u,-v) = f^\infty(x;-u,v) = (-f)^\infty(x;u,v) = -f_\infty(x;u,v)$.

Before studying the relationship among these directional derivatives, we give a few lemmas which will often be used in the sequel. Lemma 2.2.2 has appeared in [2, Lemmas 1.4, 1.5] and its proof is given in Lemma A2 of the Appendix.

Lemma 2.2.2. Let $f\colon X \longrightarrow \mathbb{R}$ be a continuous function, $x,\ v \in X$, and $t > 0$. Then there exists $\alpha \in (0,t)$ such that

$$\frac{f(x+tv) - f(x)}{t} \le D_+f(x+\alpha v;v).$$

Consequently,

$$\limsup_{y\to x} D_+f(y;v) = \limsup_{y\to x} D^+f(y;v) = \limsup_{y\to x} f^0(y;v) = f^0(x;v).$$

Remark: Let $f = -g$. Then we have

$$\frac{g(x+tv) - g(x)}{t} \ge D^+g(x+\alpha v;v),$$

and the corresponding results for $f_0(x;v)$ follows.

From Lemma 2.2.2 we have the following

Lemma 2.2.3. Suppose that $f: X \longrightarrow \mathbb{R}$ is continuous and x, u, $v \in X$. Then for any $t_0 > 0$, $t \in (0,t_0)$ and $s \in \mathbb{R}$ there exists $\alpha \in (0,t)$ such that

(2.2.1) $$\frac{1}{t}[f(x+sv+tu) - f(x+sv) - f(x+tu) + f(x)]$$

$$\leq D^+f(x+\alpha u+sv;u) - D^+f(x+\alpha u;u)$$

and

(2.2.2) $$\frac{1}{t}[f(x+sv+tu) - f(x+sv) - f(x+tu) + f(x)]$$

$$\leq D_+f(x+\alpha u+sv;u) - D_+f(x+\alpha u;u)$$

if $D^+f(\cdot;u)$ and $D_+f(\cdot;u)$ are finite on the segments $(x,x+t_0u)$ and $(x+sv,x+sv+t_0u)$.

Remark: If we let $f = -g$, then we have

(2.2.1)' $$\frac{1}{t}[g(x+sv+tu) - g(x+sv) - g(x+tu) + g(x)]$$

$$\geq D_+g(x+\alpha u+sv;u) - D_+g(x+\alpha u;u)$$

and

(2.2.2)' $$\frac{1}{t}[g(x+sv+tu) - g(x+sv) - g(x+tu) + g(x)]$$

$$\geq D^+g(x+\alpha u+sv;u) - D^+g(x+\alpha u;u)$$

Proof of Lemma 2.2.3: Let us fix an arbitrary $s \in \mathbb{R}$ and denote the left number of (2.2.1) by

$$\frac{\Phi(t) - \Phi(0)}{t}$$

where $\Phi(t) := f(x+sv+tu) - f(x+tu) = \Phi_1(t) - \Phi_2(t)$ with the obvious meaning of Φ_1, Φ_2. If $t_0 > 0$ and $t \in (0,t_0)$ then, by Lemma 2.2.2, there exists $\alpha \in (0,t)$ such that

$$\frac{\Phi(t) - \Phi(0)}{t} \leq D_+\Phi(\alpha),$$

where $D_+\Phi(\alpha)$ denotes $D_+\Phi(\alpha;1)$ for short. By assumption $D^+\Phi_1(\alpha)$,

$D^+\Phi_2(\alpha)$ are finite and it follows that

$$\frac{\Phi(t) - \Phi(0)}{t} \le D_+\Phi(\alpha) \le D^+\Phi_1(\alpha) - D^+\Phi_2(\alpha)$$
$$= D^+f(x+\alpha u+sv;u) - D^+f(x+\alpha u;u).$$

This proves (2.2.1), and similarly one can prove (2.2.2) because

$$D_+\Phi(\alpha) \le D_+\Phi_1(\alpha) - D_+\Phi_2(\alpha)$$

as the two terms on the right are finite. □

Recall that f is regular at x [1] if the one sided directional derivative

$$f'(x;v) = \lim_{t\downarrow 0} \frac{1}{t}(f(x+tv) - f(x)),$$

exists and $f'(x;v) = f^0(x;v)$ for all v.

Proposition 2.2.4. Let $f\colon X \longrightarrow \mathbb{R}$ be a continuous function. Let x, u, v $\in$ X and suppose that $f^0(\cdot;u)$, $D^+f(\cdot;u)$ and $D_+f(\cdot;u)$ are finite near x. Then one has

(2.2.3) $(f^0(\cdot;u))^0(x;v)$

$$\le f^\infty(x;u,v) = (D^+f(\cdot;u))^0(x;v) = (D_+f(\cdot;u))^0(x;v),$$

that is,

(2.2.4) $\displaystyle \limsup_{\substack{y\to x\\ s\downarrow 0}} \frac{1}{s}(f^0(y+sv;u) - f^0(y;u))$

$$\le f^\infty(x;u,v) = \limsup_{\substack{y\to x\\ s\downarrow 0}} \frac{1}{s}(D_+f(y+sv;u) - D_+f(y;u))$$
$$= \limsup_{\substack{y\to x\\ s\downarrow 0}} \frac{1}{s}(D^+f(y+sv;u) - D^+f(y;u)).$$

Dually one also has

(2.2.5) $(f_0(\cdot;u))_0(x;v)$

$$\ge f_\infty(x;u,v) = (D_+f(\cdot;u))_0(x;v) = (D^+f(\cdot;u))_0(x;v)$$

if $f_0(\cdot;u)$, $D_+f(\cdot;u)$ and $D^+f(\cdot;u)$ are finite near x.

Furthermore, if f is regular near x, then the inequality in

(2.2.3) becomes an equality.

Proof: We need only to prove (2.2.3) as (2.2.5) will then follow by considering $-f = g$ (the assertion for the regular case is evident from (2.2.4) because then $D_+f(y+sv;u) = f^o(y+sv;u)$ for all y near x and small v). By Lemma 2.2.2 we have

$$\limsup_{z\to y} D^+f(z + sv;u) = f^o(y + sv;u).$$

Thus, since $f^o(\cdot;u)$ is finite near x, it follows from the subadditivity of lim sup that

$$f^o(y + sv;u) - f^o(y;u) \le \limsup_{z\to y} (D^+f(z + sv;u) - D^+f(z;u)).$$

This implies that

(2.2.6) $\quad (f^o(\cdot;u))^o(x,v)$

$$= \limsup_{\substack{y\to x\\ s\downarrow 0}}\frac{1}{s}(f^o(y + sv;u) - f^o(y;u))$$

$$\le \limsup_{\substack{y\to x\\ s\downarrow 0}} \limsup_{z\to y}\frac{1}{s}(D^+f(z + sv;u) - D^+f(z;u))$$

$$\le \limsup_{\substack{y\to x\\ s\downarrow 0}}\frac{1}{s}(D^+f(y + sv;u) - D^+f(y;u))$$

$$= (D^+f(\cdot;u))^o(x;v),$$

showing the inequality in (2.2.3).

On the other hand, since $D^+f(\cdot;u)$ and $D_+f(\cdot;u)$ are finite near x, one has, by the subadditivity of lim sup,

$$D^+f(y+sv;u) - D^+f(y;u)$$

$$\le \limsup_{t\downarrow 0} \frac{1}{t}[f(y+sv+tu) - f(y+sv) - f(y+tu) + f(y)]$$

and also

$$D_+f(y+sv;u) - D_+f(y;u)$$

$$\le \limsup_{t\downarrow 0} \frac{1}{t}[f(y+sv+tu) - f(y+sv) - f(y+tu) + f(y)].$$

These imply that

(2.2.7) $\quad (D^+f(\cdot;u))^o(x;v)$

$$= \limsup_{\substack{y\to x\\ s\downarrow 0}} \frac{1}{s}(D^+f(y+sv;u) - D^+f(y;u))$$

$$\le \limsup_{\substack{y\to x\\ s,t\downarrow 0}} \frac{1}{st}[f(y+sv+tu) - f(y+sv) - f(y+tv) + f(y)]$$

$$= f^{\infty}(x;u,v)$$

and, similarly,

(2.2.8) $(D_+f(\cdot;u))^0(x;v)$

$$= \limsup_{\substack{y\to x\\ s\downarrow 0}} \frac{1}{s}(D_+f(y+sv;u) - D_+f(y;u))$$

$$\le \limsup_{\substack{y\to x\\ t,s\downarrow 0}} \frac{1}{ts}[f(y+sv+tu) - f(y+sv) - f(y+tu) + f(y)]$$

$$= f^{\infty}(x;u,v).$$

By definition and (2.2.2) of Lemma 2.2.3,

(2.2.9) $f^{\infty}(x;u,v)$

$$= \limsup_{\substack{y\to x\\ t,s\downarrow 0}} \frac{1}{ts}[f(y+sv+tu) - f(y+sv) - f(y+tu) + f(y)]$$

$$\le \limsup_{\substack{y\to x\\ t,s\downarrow 0}} \frac{1}{s}(D_+f(y+\alpha u+sv;u) - D_+f(y+\alpha u;u)] \quad \alpha \in (0,t)$$

$$= \limsup_{\substack{y\to x\\ s\downarrow 0}} \frac{1}{s}(D_+f(y + sv;u) - D_+f(y;u))$$

$$= (D_+f(\cdot;u))^0(x;v),$$

where we have written α for $\alpha = \alpha(y,s,u,v)$ for the sake of simplicity in notations. Similarly,

(2.2.10) $$f^{\infty}(x;u,v) \le (D^+f(\cdot;u))^0(x;v).$$

Together with (2.2.6), (2.2.7), (2.2.8) and (2.2.9), we have (2.2.3). □

Remark: There are examples of Lipschitz functions on an interval, say [0,b], which fail to be right-differentiable at infinitely many points near 0. Thus the representation given in the preceding proposition is valid, but cannot be expressed in the form of (3) in [2, Proposition 1.3]. For example, write

$$(0,1/2\pi] = \bigcup_{k=1}^{\infty}[x_{k+1},x_k],\ x_k = 1/2k\pi.$$

Define $f(0) = 0$, $f(x_k) = 0$ and

$$f(x) = (x-x_{k+1})(x_k-x)\sin (x-x_{k+1})^{-1}$$

if $x \in (x_{k+1}, x_k)$. Then $f'_+(x)$ does not exist at each x_k.

In view of proposition 2.2.4 we introduce the following generalized second directional derivative in line of Clarke's derivatives as an alternative to $f^{\infty}(x;u,v)$.

Definition 2.2.5 Let $f\colon X \longrightarrow \mathbb{R}$, x, $v \in X$, and suppose that $f^0(x;v)$ and $f_0(\cdot;v)$ are finite near x. Then the upper and lower generalized second directional derivatives are defined respectively by

$$f^{00}(x;u,v) := \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{1}{t}(f^0(y+tu;v) - f^0(y;v))$$

and

$$f_{00}(x;u,v) := \liminf_{\substack{y\to x\\ t\downarrow 0}} \frac{1}{t}(f_0(y+tu;v) - f_0(y;v)).$$

It is easy to see that the function $u \longmapsto f^{00}(x;u,v)$ is sublinear and the function $x \longmapsto f^{00}(x;u,v)$ is upper semi-continuous. Furthermore, if f is continuous and $f^0(x;v)$, $f_0(x;v)$ are finite near x, then from the above proposition we have

$$f_{\infty}(x;u,v) \le f_{00}(x;u,v) \le f^{00}(x;u,v) \le f^{\infty}(x;u,v);$$

and $f^{00}(x;u,v) = f^{\infty}(x;u,v)$ if f is regular near x.

In § 7 we shall give applications of f^{∞} and f^{00} in the second-order necessary optimality condition for constrained problem.

3. On Cominetti and Correa's conjecture.

In this section we study second-order directional derivative of the function h of the form

$$h(x) = \max\{g_1(x),\ g_2(x),\ \cdots,\ g_n(x)\} \qquad (x \in X)$$

where each g_i is a real-valued function on X. Note that $h = f \circ g$ if one writes $g = (g_1, g_2, \cdots, g_n)$ and defines

$$f(a) = \max_{i \in I} \{a_i\}, \text{ for any } a = (a_1, \cdots, a_n) \in \mathbb{R}^n$$

where $I: = \{1, 2, \cdots, n\}$. Let $I(a)$ denote the subset of I consisting of all i for which $f(a) = a_i$.

For x, u, v $\in$ X we shall write $H(x;u,v) \underset{=}{\propto} 0$ to denote the following condition:

$$(g_i'(x;u) - g_j'(x;u))(g_i'(x;v) - g_j'(x;v)) \le 0$$

for all i, j $\in I(g(x))$, and $H(x;u,v) \propto 0$ to denote the condition that the strict inequality holds for all distinct i, j $\in I(g(x))$.

Suppose each g_i is a C^2-function with the usual second-order directional derivative at x with respect to the directions u, v denoted by $D^2g_i(x;u,v)$. Cominetti and Correa conjectured in [2] that if $\{g_i'(x);\ i \in I(g(x))\}$ is affinely independent and if $H(x;u,v) \underset{=}{\propto} 0$, then the following formula holds

$$h^{\infty}(x;u,v) = \max_{i \in I(g(x))} D^2g_i(x;u,v). \tag{2.3.1}$$

This is incorrect as shown by Example 2.3.3 below, but true if the condition is strengthened to $H(x;u,v) \propto 0$ (Corollary 2.3.5).

In the following we first consider a property related to the set I(a).

Lemma 2.3.1. Suppose that X is a locally convex space and g is an arbitrary continuous function of X into $\mathbb{R}^n$ denoted by $g = (g_1, \cdots, g_n)$. Then for the above f and for any $x \in X$, there exists a neighbourhood W of x such that

$$I(g(y)) \subseteq I(g(x))$$

for all $y \in W$.

Proof: We fix $i \in I(g(x))$. Then, for each $j \in I \setminus I(g(x))$,

$$g_j(\cdot) < g_i(\cdot)$$

at x and hence on a neighbourhood W_j of x. Do this for each such j and let W denote the intersection of W_j's. Then W has the required property: if $y \in W$ and $j \notin I(g(x))$, then $g_j(y) < g_i(y)$ showing that $j \notin I(g(y))$. □

Lemma 2.3.2. Suppose that $g'(\cdot) = (g'_1(\cdot), \cdots, g'_n(\cdot))$ is continuous near x and $\{g'_i(x);\ i \in I(g(x))\}$ are affinely independent. Then either there exists a neighbourhood W of x such that the following condition $H(y;u,v) \underline{\propto} 0$ holds for each $y \in W$:

$$(g'_i(y;v) - g'_j(y;v))(g'_i(y;u) - g'_j(y;u)) \leq 0, \ \forall i,\ j \in I(g(y))$$

or

$$h^{\infty}(x;u,v) = +\infty.$$

Proof. By the continuity of g' at x it is easy to show that there exists a neighbourhood W_1 of x such that

$$\{g'_i(y);\ i \in I(g(x))\}$$

are affinely independent for all $y \in W_1$. Now if for each neighbourhood U of x, there exists $y \in U \cap W_1$ so that the condition $H(y;u,v) \underline{\propto} 0$ is not satisfied, then by [2, Proposition 3.9] or Lemma A4. $h^{\infty}(y;u,v) = +\infty$. Hence by the upper semicontinuity of $h^{\infty}(\cdot;u,v)$,

$$h^{\infty}(x; u, v) = +\infty. \qquad \square$$

Example 2.3.3. Let $g = (g_1, g_2, g_3)$ with the C^2-functions $g_1(x,y,z) = \xi(x) + x + y$, $g_2(x,y,z) = x + 2y$ and $g_3(x,y,z) = 4(x-y) + z$ for all $x, y, z \in \mathbb{R}$, where

$$\xi(x) = \begin{cases} x^5 \sin \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Let $\bar{x} = (0,0,0)$. Since $g_1, g_2, g_3 = 0$ at $\bar{x}$, $I(g(\bar{x})) = \{1,2,3\}$. Further, $g_1'(\bar{x}) = (1,1,0)$, $g_2'(\bar{x}) = (1,2,0)$ and $g_3'(\bar{x}) = (4,-4,1)$. Thus, $g_1'(\bar{x})$, $g_2'(\bar{x})$ and $g_3'(\bar{x})$ are linearly independent. Let

$$u = (1,0,0) \text{ and } v = (1,1,0).$$

Then

$$g_1'(\bar{x};u) - g_2'(\bar{x};u) = 1 - 1 = 0 \text{ and } g_1'(\bar{x};v) - g_2'(\bar{x};v) = 2 - 3 < 0.$$

Similarly we can verify, for all other pairs of distinct i, j, that $g_i'(\bar{x};u) - g_j'(\bar{x};u)$ and $g_i'(\bar{x};v) - g_j'(\bar{x};v)$ are of opposite signs (or zero); that is, the Cominetti and Correa's condition $H(\bar{x};u,v) \underline{\propto} 0$ is satisfied. But, in constrast to their conjecture, (2.3.1) does not hold. In fact, we will prove that

$$\text{(2.3.2)} \qquad h^{\infty}(\bar{x}; u, v) = +\infty.$$

Let $P_n := (x_n, y_n, 0)$ with

$$x_n = -\frac{1}{2n\pi + \frac{\pi}{2}} \quad \text{and} \quad y_n = \frac{1}{(2n\pi + \frac{\pi}{2})^5} = -x_n^5.$$

Then,

$$g_1(P_n) = -x_n^5 + x_n + y_n = g_2(P_n).$$

This implies that $I(g(P_n)) = \{1,2\}$ because $x_n \leq 2\, y_n$. Further,

$$g_1'(P_n) = (1 - 5(2n\pi + \tfrac{\pi}{2})^{-4}, 1, 0)$$

and so

$$g_1'(P_n;u) - g_2'(P_n;u) = -5(2n\pi + \tfrac{\pi}{2})^{-4} < 0$$

and

$$g_1'(P_n;v) - g_2'(P_n;v) = 2 - 5(2n\pi+\frac{\pi}{2})^{-4} - 3 < 0.$$

Thus the condition $H(P_n;u,v) \underset{-}{\propto} 0$ does not hold for all n. Since $P_n \longrightarrow \bar{x}$, it follows from Lemma 2.3.2 that (2.3.2) must hold.

The above example actually shows that for $n \geq 2$ (if $n = 2$, we ignore g_3), the condition $H(x;u,v) \underset{-}{\propto} 0$ is not sufficient for

$$h^{\infty}(x;u,v) = \max_{i\in I(g(X))} D^2 g_i(x;u,v).$$

We shall show however that the strengthened condition $H(x;u,v) \propto 0$ will be sufficient. Before our proof we recall an elementary fact (see, Proposition 1.2.3) that if each g_i is directionally differentiable at x, then one has

$$h'(x;u) = \max_{i\in I(g(x))} g_i'(x;u) \tag{2.3.3}$$

for all $u \in X$.

Proposition 2.3.4. Suppose that each g_i is a C^1-function (that is, continuous Gâteaux differentiable function) at x. If for all $i, j \in I(g(x))$, $i \neq j$, one has

$$[g_i'(x;u) - g_j'(x;u)][g_i'(x;v) - g_j'(x;v)] < 0, \tag{2.3.4}$$

then

$$h^{\infty}(x;u,v) \leq \max_{i\in I(g(X))} g_i^{\infty}(x;u,v).$$

Proof. By Proposition 2.2.4, we take a net $(z_\upsilon,\lambda_\upsilon)_\upsilon \in X \times \mathbb{R}_+$ written for short (z,λ) with $z \longrightarrow x$ and $\lambda \downarrow 0$ such that

$$h^{\infty}(x;u,v) = \lim_{\substack{z\to x\\ \lambda\downarrow 0}} \frac{1}{\lambda}(h'(z+\lambda u;v) - h'(z;v))$$

$$= \lim_{\substack{z\to x\\ \lambda\downarrow 0}} \frac{1}{\lambda}\left(\max_{i\in I(g(z+\lambda u))} g_i'(z+\lambda u;v) - \max_{i\in I(g(z))} g_i'(z;v)\right).$$

In view of lemma 2.3.1, we can assume that

$$I(g(z)),\ I(g(z+\lambda u)) \subseteq I(g(x)). \tag{2.3.5}$$

Since I is a finite set and considering a subnet if necessary we can assume without loss of generality that

$$h'(z;v) = \max_{i\in I(g(z))} g_i'(z;v) = g_1'(z;v) \text{ say,}$$

and

$$h'(z+\lambda u;v) = \max_{i\in I(g(z+\lambda u))} g_i'(z+\lambda u;v) = g_{i_0}'(z+\lambda u;v)$$

for some $i_0 \in I(g(z+\lambda u))$ and for all (z,λ).

We claim that there exists a subnet (z_s,λ_s) of (z,λ) such that

$$g_{i_0}'(z_s;v) - g_1'(z_s;v) \le 0.$$

In this case we will then obtain

$$\begin{aligned} h^{\infty}(x;u,v) &= \lim_s \frac{1}{\lambda_s}[g_{i_0}'(z_s+\lambda u;v) - g_{i_0}'(z_s;v) \\ &\qquad + g_{i_0}'(z_s;v) - g_1'(z_s;v)] \\ &\le \limsup_s \frac{1}{\lambda_s}[g_{i_0}'(z_s+\lambda u;v) - g_{i_0}'(z_s;v))] \\ &\le g_{i_0}^{\infty}(x;u,v). \end{aligned}$$

By (2.3.5), $i_0 \in I(g(x))$ and so we are done.

If our claim is false then, by considering a subnet if necessary, we assume that for all (z,λ)

(2.3.6) $$g_{i_0}'(z;v) - g_1'(z;v) > 0$$

where

(2.3.7) $$i_0 \in I(g(z+\lambda u)) \text{ and } 1 \in I(g(z))$$

for all (z,λ). It follows that $g_{i_0}'(x;v) - g_1'(x;v) \ge 0$ and from (2.3.4) that the strict inequality must hold and

(2.3.8) $$g_{i_0}'(x;u) - g_1'(x;u) < 0.$$

Since g_i is C^1, the formula (2.3.8) can be rewritten as

$$\lim_{\substack{y\to x\\ t\downarrow 0}} \frac{1}{t}[g_{i_0}(y+tu) - g_{i_0}(y) - g_1(y+tu) + g_1(y)] < 0$$

and so we can choose a neighbourhood W of x and $\delta > 0$ such that

(2.3.9) $\quad (g_{i_0}(y+tu) - g_{i_0}(y) - g_1(y+tu) + g_1(y)) < 0$

for all $y \in W$ and $0 < t < \delta$. Without loss of generality, we can assume that

$$(z,\lambda) \in W \times (0,\delta).$$

From (2.3.6) and the choice of 1, we see that $i_0 \notin I(g(z))$ and so $g_{i_0}(z) < g_1(z)$ for all (z,λ). Thus, together with (2.3.9) we conclude that

$$g_{i_0}(z+\lambda u) - g_1(z+\lambda u)$$

$$< g_{i_0}(z+\lambda u) - g_{i_0}(z) - g_1(z+\lambda u) + g_1(z) < 0.$$

But this is imposible since

$$i_0 \in I(g(z+\lambda u)). \quad \square$$

Corollary 2.3.5. Suppose that each g_i is a C^2-function at x, $1 \le i \le n$, and the derivatives $\{g_i'(x);\ i \in I(g(x))\}$ are affinely independent. If for all i, $j \in I(g(x))$, $i \neq j$, one has

$$(g_i'(x;u) - g_j'(x;u))(g_i'(x;v) - g_j'(x;v)) < 0,$$

then

$$h^{\infty}(x;u,v) = \max_{i\in I(g(X))} D^2 g_i(x;u,v).$$

Proof: Since each g_i is a C^2-function,

$$g_i^{\infty}(x;u,v) = D^2 g_i(x;u,v).$$

It follows from Proposition 2.3.4 that

$$h^{\infty}(x;u,v) \le \max_{i\in I(g(X))} D^2 g_i(x;u,v).$$

But the assumption on affinely independence ensures

$$h^{\infty}(x;u,v) \ge \max_{i\in I(g(X))} D^2 g_i(x;u,v)$$

by [2, Prop. 3.7 and 3.8] or Lemma A3. $\square$

For normed spaces, we have another sufficient condition result for the similar representation of h^{∞}:

Proposition 2.3.6. Let X be a normed space, u, x ∈ X and g_i, $1 \le i \le n$, be C^2-functions at x. Suppose that W is a neighbourhood of x such that for all y ∈ W and i, j ∈ I(g(y)), one has $g_i'(y;u) = g_j'(y;u)$. Then

$$h^{\infty}(x;u,u) = \max_{i\in I(g(x))} D^2 g_i(x;u,u).$$

Proof: Note first that since g_i, $1 \le i \le n$, are C^2-functions, $g_i'(\cdot;u)$ are continuous on some neighbourhood $W_1 \subseteq W$ of x [1, p.32, cor.]. Consequently by Lemma 2.3.1 and (2.3.3) $h'(\cdot;u)$ is also continuous. Next we show that

$$D^+(h'(\cdot;u))(z;u) = \max_{i\in I(g(z))} D^2 g_i(z;u,u) \tag{2.3.10}$$

for any $z \in W_1$. To do this, we choose a subnet $t_\upsilon > 0$ written for short t such that

$$D^+(h'(\cdot;u))(z;u) = \lim_{t\downarrow 0} \frac{1}{t}\{h'(z+tu;u) - h'(z;u)\}$$
$$= \lim_{t\downarrow 0} \frac{1}{t}\{\max_{i\in I(g(z+tu))} g_i'(z+tu;u) - \max_{i\in I(g(z))} g_i'(z;u)\}.$$

By Lemma 2.3.1 we can assume that $I(g(z+tu)) \subseteq I(g(z))$ and $z+tu \in W_1$. Since I is a finite set and considering a subnet if necessary we can assume without loss of generality that there exists $i_z \in I(g(z+tu)) \subseteq I(g(z))$ such that

$$\max_{i\in I(g(z+tu))} g_i'(z+tu;u) = g_{i_z}'(z+tu;u)$$

for all t. By assumption, $g_i'(z;u) = g_{i_z}'(z;u)$ so

$$\max_{i\in I(g(z))} g_i'(z;u) = g_{i_z}'(z;u))$$

it follows that

$$D^+(h'(\cdot;u))(z;u) = \lim_{t\downarrow 0} \frac{1}{t}\{g_{i_z}'(z+tu;u) - g_{i_z}'(z;u)\}$$

$$= D^2 g_{i_z}(z;u,u).$$

Now if (2.3.10) is not true there must exists $i \in I(g(z))$ such that

$$D^2 g_i(z;u,u) > D^2 g_{i_z}(z;u,u),$$

since $g_i'(y;u) = g_{i_z}'(y;u)$ by assumption, it follows that

$$g_i'(z+\tau u;u) - g_{i_z}'(z+\tau u;u) > 0$$

for all small enough $\tau > 0$. Now we choose a small enough t from our net $\{t_\upsilon\}$ and recall that $i_z \in I(g(z+tu))$. But

$$\int_0^t (g_i - g_{i_z})'(z+\tau u;u)d\tau = g_i(z+tu) - g_{i_z}(z+tu) > 0,$$

contrading the given assumption. Thus (2.3.10) is proved. On the other hand, we have by Proposition 2.2.4 that

$$h^{\infty}(x;u,u) = (h'(\cdot;u))^0(x;u) = \lim\sup_{z\to x} D^+(h'(\cdot;u))(z;u)$$

$$= \lim\sup_{z\to x} \max_{i\in I(g(z))} D^2 g_i(z;u,u) = D^2 g_{i_0}(x;u,u)$$

for some $i_0 \in I(g(x))$, where the last equality is valid because of Lemma 2.3.1 and the fact that I is a finite set. Since

$$h^{\infty}(x;u,u) \geq D^+(h'(\cdot;u))(x;u)$$

and, by (2.3.10)

$$D^+(h'(\cdot;u))(x;u) = \max_{i\in I(g(x))} D^2 g_i(x;u,u),$$

it follows that

$$h^{\infty}(x;u,u) = \max_{i\in I(g(x))} D^2 g_i(x;u,u). \qquad \square$$

Corollary 2.3.7. Let X be a normed space, u, $x \in X$ and g_i, $1 \leq i \leq n$, be C^2-functions at x. Suppose further that

$$\{g_i'(x): i \in I(g(x))\}$$

are affinelly independent. Then

$$h^{\infty}(x;u,u) = \max_{i\in I(g(x))} D^2 g_i(x;u,u)$$

if and only if there exists a neighbourhood W of x such that

$$g_i'(y;u) = g_j'(y;u)$$

for all $y \in W$ and $i, j \in I(g(y))$.

Proof: The sufficency follows from Proposition 2.3.6. Conversely if $h^{\infty}(x;u,u) = \max_{i\in I(g(x))} D^2 g_i(x;u,u)$, then $h^{\infty}(x;u,u)$ is finite and hence, by Lemma 2.3.2, there exists a neighbourhood W of x such that the condition $H(y;u,u) \underline{\propto} 0$ holds for each $y \in W$. This implies immediately that

$$g_i'(y;u) - g_j'(y;u) = 0$$

for all $y \in W$ and $i, j \in I(g(y))$. □

4. Generalized Second-order Taylor expansion.

Suppose that X and f are as in § 1, we define the generalized Hessian [2] of f at x by

$$\partial^2 f(x)(u) := \{x^* \in X^*;\ \langle x^*, v\rangle \le f^{\infty}(x;u,v) \text{ for all } v \in X\},$$

where the symbol X^* denotes the dual space of X. It is easy to see that $\partial^2 f(x)(u)$ is a closed convex subset of X^* with respect to the w^*-topology. If f is twice C-differentiable at x [2], that is, $f^{\infty}(x;\cdot,v)$ (or equivalently $f^{\infty}(x;v,\cdot)$) is lower semi-continuous for each $v \in X$, then one has

(2.4.1) $$f^{\infty}(x;u,v) = \sup\langle \partial^2 f(x)(u), v\rangle.$$

Now for a, u, v and $x \in X$, we consider the following new kinds of first and second-order directional derivatives respectively defined by

$$f_a^0(x;u) := \limsup_{\substack{\lambda\to 0\\ s\downarrow 0}} \frac{1}{s}\left\{f(x+\lambda a+su) - f(x+\lambda a)\right\},$$

$$f^{0}_{+a}(x;u) := \limsup_{s,\lambda\downarrow 0} \frac{1}{s}\Big\{f(x+\lambda a+su) - f(x+\lambda a)\Big\},$$

$$f^{0}_{-a}(x;u) := \limsup_{\substack{\lambda\uparrow 0\\ s\downarrow 0\\ \lambda+s\le 0}} \frac{1}{s}\Big\{f(x+\lambda a+su) - f(x+\lambda a)\Big\},$$

$$f_{0,a}(x;u) := -(-f)^{0}_{a}(x;u), \quad f_{0,+a}(x;u) := -(-f)^{0}_{+a}(x;u)$$

and

$$f^{\infty}_{a}(x;u,v) := \limsup_{\substack{\lambda\to 0\\ s,t\downarrow 0}} \frac{1}{st}\Big\{f(x+\lambda a+tu+sv) - f(x+\lambda a+tu) -f(x+\lambda a+sv) + f(x+\lambda a)\Big\},$$

$$f_{\infty,a}(x;u,v) := \liminf_{\substack{\lambda\to 0\\ s,t\downarrow 0}} \frac{1}{st}\Big\{f(x+\lambda a+tu+sv) - f(x+\lambda a+tu) -f(x+\lambda a+sv) + f(x+\lambda a)\Big\},$$

(f^{0}_{a} and f^{∞}_{a} are different from f^{0} and f^{∞} as here we only consider

$$x + \lambda a \longrightarrow x$$

along the direction a).

In terms of f^{0}_{a}, f^{0}_{+a}, and f^{0}_{-a}, we have the following

Lemma 2.4.1. Suppose that $f: X \longrightarrow \mathbb{R}$ is a continuous function. Then one has

$$f^{0}_{+a}(x;a) = \limsup_{\lambda\downarrow 0} D^{+}f(x+\lambda a;a),$$

$$f^{0}_{-a}(x;a) = \limsup_{\lambda\uparrow 0} D^{+}f(x+\lambda a;a),$$

and

$$f^{0}_{a}(x;a) = \limsup_{\lambda\to 0} D^{+}f(x+\lambda a;a)$$

Proof: Let x, a $\in$ X, and $\lambda < 0$, $s > 0$. By Lemma 2.2.2 there exists $\alpha \in (0,s)$ such that

$$\frac{1}{s}\Big\{f(x+\lambda a+sa) - f(x+\lambda a)\Big\} \le D^{+}f(x+\lambda a+\alpha a;a).$$

Hence, by the definition of f^{0}_{-a} we have

$$f^{0}_{-a}(x;a) \le \limsup_{\lambda\uparrow 0} D^{+}f(x+\lambda a;a).$$

But

$$\limsup_{\lambda\uparrow 0} D^+f(x+\lambda a;a) = \limsup_{\lambda\uparrow 0} \limsup_{s\downarrow 0} \frac{1}{s}\left\{f(x+\lambda a+sa) - f(x+\lambda a)\right\}$$

$$\le \limsup_{\substack{\lambda\uparrow 0\\ s\downarrow 0\\ \lambda+s\le 0}} \frac{1}{s}\left\{f(x+\lambda a+sa) - f(x+\lambda a)\right\} = f^0_{-a}(x;a).$$

So we have $f^0_{-a}(x;a) = \limsup_{\lambda\uparrow 0} D^+f(x+\lambda a;a)$. Similarly, we have $f^0_{+a}(x;a) = \limsup_{\lambda\downarrow 0} D^+f(x+\lambda a;a)$. Thus, one has

$$f^0_a(x;a) = \limsup_{\lambda\to 0} D^+f(x+\lambda a;a). \quad \square$$

The following Theorem provides an answer to the question of Cominetti and Correa [2] about Taylor's expansion.

Theorem 2.4.2. Let $f\colon [x,y] \longrightarrow \mathbb{R}$ be a continuous function on a line segment in a locally convex space X. Suppose that $D^+f(\cdot;y-x)$ is finite, upper semi-continuous on (x,y) and $f^0_{+(y-x)}(x;y-x)$, $f^0_{-(y-x)}(y;y-x)$ are finite. Then there exists $t_0 \in (0, 1)$ such that

$$\text{(2.4.2)} \quad \tfrac{1}{2}f^{\infty}_{y-x}(x+t_0(y-x);y-x,y-x) \ge f(y) - f(x) - f^0_{+(y-x)}(x;y-x)$$
$$\ge \tfrac{1}{2}f_{\infty,y-x}(x+t_0(y-x);y-x,y-x).$$

Hence, we also have

$$\text{(2.4.3)} \quad \tfrac{1}{2}f^{\infty}_{y-x}(x+t_0(y-x);y-x,y-x) + f^0_{y-x}(x;y-x) \ge f(y) - f(x)$$
$$\ge f_{0,y-x}(x;y-x) + \tfrac{1}{2}f_{\infty,y-x}(x+t_0(y-x);y-x,y-x).$$

The following theorem is a corollary of Theorem 2.4.2.

Theorem 2.4.3. Suppose that the assumptions in Theorem 2.4.2 hold. If in addition f is defined on X and is twice C-differentiable at each point of (x,y), then there exists $t_0 \in (0,1)$ such that

$$(2.4.4)\quad f(y) - f(x) - f^0_{+(y-x)}(x;y-x) \in \overline{\tfrac{1}{2}\langle\partial^2 f(x+t_0(y-x))(y-x),y-x\rangle};$$

and also

$$(2.4.5)\quad f(y) - f(x) \in \overline{\langle\partial f(x),y-x\rangle + \tfrac{1}{2}\langle\partial^2 f(x+t_0(y-x))(y-x),y-x\rangle}$$

if $\partial f(x)$ is nonempty and $f^0(x;y-x) = \sup_{x^* \in \partial f(x)} \langle x^*,y-x\rangle$, where $\partial f(x)$ denotes the Clarke's subdifferential and the "bar" denotes the closure of the set. The bar is superfluous if f is $C^{1,1}$ [2] on (x,y).

Indeed, granting Theorem 2.4.2, we have

$$\tfrac{1}{2}f^{\infty}(x+t_0(y-x);y-x,y-x) + f^0(x;y-x)$$
$$\geq f(y) - f(x) \geq f_0(x;y-x) + \tfrac{1}{2}f_{\infty}(x+t_0(y-x);y-x,y-x)$$

by (2.4.3). Thus, by (2.4.1) and our assumptions for any $\varepsilon > 0$ there exist $x_1^* \in \partial f(x)$ and $x_2^* \in \partial^2 f(x+t_0(y-x))(y-x)$ such that

$$f(y) - f(x) \leq \langle x_1^*+\tfrac{1}{2}x_2^*,y-x\rangle + \varepsilon.$$

Similarly, since $f_{\infty}(x+t_0(y-x);y-x,y-x) = -f^{\infty}(x+t_0(y-x);y-x,x-y)$ and $f_0(x;y-x) = -f^0(x;x-y)$, there exist

$$z_1^* \in \partial f(x) \text{ and } z_2^* \in \partial^2 f(x+t_0(y-x))(y-x)$$

such that

$$f(x) - f(y) \leq \langle z_1^*+\tfrac{1}{2}z_2^*,x-y\rangle + \varepsilon.$$

Hence we can choose $\lambda \in (0,1)$ such that

$$f(y) - f(x) = \langle(\lambda z_1^*+(1-\lambda)x_1^*)+\tfrac{1}{2}(\lambda z_2^*+(1-\lambda)x_2^*),y-x\rangle + (1-\lambda)\varepsilon - \lambda\varepsilon.$$

Since $\partial f(x)$ and $\partial^2 f(x+t_0(y-x))(y-x)$ are convex,

$$\lambda z_1^*+(1-\lambda)x_1^* \in \partial f(x) \text{ and } \lambda z_2^*+(1-\lambda)x_2^* \in \partial^2 f(x+t_0(y-x))(y-x).$$

We then have

$$f(y) - f(x) \in \overline{\langle\partial f(x),y-x\rangle + \tfrac{1}{2}\langle\partial^2 f(x+t_0(y-x))(y-x),y-x\rangle}$$

as required to show for (2.4.5). Similarly one can prove (2.4.4). Thus it remains only to prove Theorem 2.4.2.

5. Detailed Proof of Theorem 2.4.2.

This section is entirely devoted to the proof of theorem 2.4.2. Let γ: $[0,1] \longrightarrow \mathbb{R}$ be defined by $\gamma(t) = f(x+t(y-x))$. Then it is easy to see that

$$D^{+}f(x+t(y-x);y-x) = D^{+}\gamma(t;1),$$

$$f^{o}_{-(y-x)}(y;y-x) = \gamma^{o}_{-}(1;1) \ (= \gamma^{o}_{-1}(1;1)),$$

$$f^{o}_{+(y-x)}(x;y-x) = \gamma^{o}_{+}(0;1) \ (= \gamma^{o}_{+1}(0;1)),$$

$$f^{o}_{y-x}(x;y-x) = \gamma^{o}(0;1), \ f_{o,y-x}(x;y-x) = \gamma_{o}(0;1)$$

and

$$f^{\infty}_{y-x}(x+t_{o}(y-x);y-x,y-x) = \gamma^{\infty}(t_{o};1,1),$$

$$f_{\infty,y-x}(x+t_{o}(y-x);y-x,y-x) = \gamma_{\infty}(t_{o};1,1).$$

Then Theorem 2.4.2 can be rewritten as

Theorem 2.5.1. Let γ: $[0,1] \longrightarrow \mathbb{R}$ be a continuous function. Suppose that $D^{+}\gamma(\cdot;1)$ is finite, upper semi-continuous on $(0,1)$ and $\gamma^{o}_{-}(1;1)$, $\gamma^{o}_{+}(0;1)$ are finite. Then there exists $t_{o} \in (0,1)$ such that

$$(2.5.1) \quad \tfrac{1}{2}\gamma^{\infty}(t_{o};1,1) \ge \gamma(1) - \gamma(0) - \gamma^{o}_{+}(0;1) \ge \tfrac{1}{2}\gamma_{\infty}(t_{o};1,1)$$

and so

$$(2.5.2) \quad \tfrac{1}{2}\gamma^{\infty}(t_{o};1,1) + \gamma^{o}(0;1) \ge \gamma(1) - \gamma(0) \ge \gamma_{o}(0;1) + \tfrac{1}{2}\gamma_{\infty}(t_{o};1,1).$$

To show Theorem 2.5.1, we need the following

Lemma 2.5.2. Suppose that the function h: $[0,1] \longrightarrow \mathbb{R}$ is upper semi-continuous on $(0,1)$ with $h(0) = h(1)$ and

$$\limsup_{t\downarrow 0} h(t) = h(0), \ \limsup_{t\uparrow 1} h(t) = h(1).$$

Then one of the following properties holds:

(i) h attains a local maximum at some $t_0 \in (0,1)$;

(ii) there exists $t_0 \in (0,1)$ such that h is decreasing on $[0,t_0)$ and increasing on $(t_0,1]$.

Proof: By assumptions, h is upper semi-continuous on [0,1]. Suppose (i) does not hold. Then h is neither decreasing on [0,1) nor increasing on (0,1], for otherwise the assumptions of the lemma would imply that h is a constant function. Thus there exist $t_1, t_2 \in [0,1)$ with $t_1 < t_2$ and $h(t_1) < h(t_2)$. We then claim that h is increasing on $(t_2,1]$. In fact if there exist $t_3, t_4 \in (t_2,1]$ with $t_3 < t_4$ such that $h(t_3) > h(t_4)$ the upper semi-continuity of h on $[t_1,t_4]$ will imply that (i) holds at some interior point of $[t_1,t_4]$.

Let t_0 denote the greatest lower bound of the non-empty set $T: = \left\{t \in (0,1);\ h \text{ is increasing on } (t,1]\right\}$. Then $t_0 \le t_2 < 1$ and also $t_0 \ne 0$ because h is not increasing on (0,1]. Note further that h is decreasing on $[0,t_0)$ for otherwise one can show as above that there exist $\bar{t}_1, \bar{t}_2 \in [0,t_0)$ with $\bar{t}_1 < \bar{t}_2$, $h(\bar{t}_1) < h(\bar{t}_2)$ and hence that h is increasing on $(\bar{t}_2,1]$, contradicting the definition of t_0. It is now clear that t_0 has the properties required in (ii). □

Now we prove Theorem 2.5.1. Define the function $h: [0,1] \longrightarrow \mathbb{R}$ by

$$h(t) := \gamma(t) - \gamma(1) + (1-t)\xi(t) + (1-t)^2[\gamma(1) - \gamma(0) - \gamma_+^0(0;1)],$$

where

$$\xi(t) := \begin{cases} \gamma_+^0(0;1) & t = 0 \\ D^+\gamma(t;1) & 0 < t < 1. \\ \gamma_-^0(1;1) & t = 1 \end{cases}$$

Then by the finitness assumption of $\gamma_-^0(1;1)$ and $\gamma_+^0(0;1)$ it follows from Lemma 2.4.1 that $h(0) = h(1) = 0$,

$$\limsup_{t\downarrow 0} h(t) = \limsup_{t\downarrow 0} D^+\gamma(t;1) - \gamma_+^0(0;1) = 0 = h(0)$$

and

$$\limsup_{t\uparrow 1} h(t) = 0 = h(1).$$

Further h is upper semi-continuous on (0,1) since $D^+\gamma(\cdot;1)$ is assumed upper semi-continuous on (0,1). Thus, Lemma 2.5.2 is applicable to h and so there exists $t_0 \in (0,1)$ such that either (i) h attains a local maximum at t_0 or (ii) h is decreasing on $[0,t_0)$ and increasing on $(t_0,1]$.

(I) Suppose (i) holds. Then we have

(a) $0 \geq D^+h(t_0;1)$

and

(b) $0 \leq h^0(t_0;1)$

by Theorem 1.3.3. Note that, by subadditivity,

$$h^0(t_0;1) \leq \gamma^0(t_0;1) - D^+\gamma(t_0;1) + (1-t_0)(D^+\gamma(\cdot;1))^0(t_0;1)$$
$$- 2(1-t_0)[\gamma(1) - \gamma(0) - \gamma_+^0(0;1)],$$

where the first two terms can be cancelled out because

$$\limsup_{t\to t_0} D^+\gamma(t;1) = \gamma^0(t_0;1)$$

by Lemma 2.2.2 and $\limsup_{t\to t_0} D^+\gamma(t;1) \leq D^+\gamma(t_0;1)$ by the upper semi-continuity assumption of $D^+\gamma(\cdot;1)$. Hence (b) and (2.2.3) of Proposition 2.2.4 imply that

(2.5.3) $\gamma(1) - \gamma(0) - \gamma_+^0(0;1) \leq \frac{1}{2}(D^+\gamma(\cdot;1))^0(t_0;1) = \frac{1}{2}\gamma^\infty(t_0;1,1).$

This verifies one inequality required in (2.5.1). The other inequality in (2.5.1) follows similarly from (a) because, by (2.2.5) of proposition 2.2.4, one has

$$D_+(D^+\gamma(\cdot;1))(t_0;1) \geq \gamma_\infty(t_0;1,1)$$

and, by elementary computation rules for D^+ and D_+, that

(2.5.4) $D^+h(t_0;1) \geq D^+\gamma(t_0;1) - D^+\gamma(t_0;1) + (1-t_0)D_+(D^+\gamma(\cdot;1))(t_0;1)$
$$- 2(1-t_0)[\gamma(1) - \gamma(0) - \gamma_+^0(0;1)].$$

(II) We next consider the case when (ii) holds: h is decreasing on $[0,t_0)$ and increasing on $(t_0,1]$. Take a sequence $t_n \uparrow t_0$ and note that

$(\bar{a})$ $0 \geq D^+\gamma(t_n;1)$

for each n and

$(\bar{b})$ $0 \leq h^0(t_0;1)$.

As done above $(\bar{b})$ ensures that (2.5.3) holds while $(\bar{a})$ implies that

$$\gamma(1) - \gamma(0) - \gamma_+^0(0;1) \geq \frac{1}{2}D_+(D^+\gamma(\cdot;1))(t_n;1) \geq \frac{1}{2}\gamma_\infty(t_n;1,1)$$

because (2.5.4) holds with t_0 replaced by t_n. Since $\gamma_\infty(\cdot;1,1)$ is lower semi-continuous (Proposition 2.2.1) we have the other inequality required in (2.5.1) in addition to (2.5.3). □

6. Corollaries of Theorem 2.4.2 and Theorem 2.4.3.

Corollary 2.6.1. [2, Prop. 4.1.] Suppose that $f: X \longrightarrow \mathbb{R}$ is continuously Gâteaux differentiable and twice C-differentiable on a segment $[x,y] \subseteq X$. Then there exists $t_0 \in (0,1)$ such that

$$f(y) - f(x) - f'(x;y-x) \in \frac{1}{2}\overline{\langle\partial^2 f(x+t_0(y-x))(y-x),y-x\rangle}.$$

If f is $C^{1,1}$ on $[x,y]$, then the closure can be ignored.

Proof: Since f is continuously Gâteaux differentiable at each point of $[x,y]$, it satisfies the assumptions in Theorem 2.4.2. Now apply Theorem 2.4.3. □

Corollary 2.6.2. Suppose that $f\colon X \longrightarrow \mathbb{R}$ is continuous at each point of a segment $[x,y]$. Then f satisfies (2.4.2) in each of the following cases:

(i) $D^+f(\cdot;y-x)$, $f^{\infty}_{y-x}(\cdot;y-x,y-x)$ and $f_{\infty,y-x}(\cdot;y-x,y-x)$ are finite on (x,y) and $f^0_{+(y-x)}(x;y-x)$, $f^0_{-(y-x)}(y;y-x)$ are finite;

(ii) $D^+f(z;y-x) = f^0_{y-x}(z;y-x)$ at each point of (x,y) and $f^0_{+(y-x)}(x;y-x)$, $f^0_{-(y-x)}(x;y-x)$ are finite;

(iii) f is regular in the Clarke's sense at each point of (x,y) and $f^0_{+(y-x)}(x;y-x)$, $f^0_{-(y-x)}(y;y-x)$ are finite;

Proof: Suppose that (i) is true. Let $\gamma(t) := f(x+t(y-x))$, $t \in (0,1)$. Clearly, it suffices to show that $D^+\gamma(\cdot;1)$ is upper semi-continuous on $(0,1)$. Now $D^+\gamma(t;1)$ and $\gamma^{\infty}(t;1,1)$ are finite for any $t \in (0,1)$. Take a finite number $K > \gamma^{\infty}(t;1,1)$. Then, by (2.2.4) of Proposition 2.2.4, there exists $\delta > 0$ such that

$$K > \frac{1}{\lambda}\left\{D^+\gamma(t'+\lambda;1) - D^+\gamma(t';1)\right\}$$

whenever $|t'-t| < \delta$ and $0 < \lambda < \delta$. Passing to the limits as $\lambda \downarrow 0$ and $t' \longrightarrow t$ it follows that

$$0 \ge \limsup_{\substack{t'\to t\\ \lambda\downarrow 0}} \left\{D^+\gamma(t'+\lambda;1) - D^+\gamma(t';1)\right\}$$

and so

$$0 \ge \limsup_{\lambda\downarrow 0} D^+\gamma(t+\lambda;1) - D^+\gamma(t;1).$$

Thus,

$$\text{(2.6.1)} \quad D^+\gamma(t;1) \ge \limsup_{\lambda\downarrow 0} D^+\gamma(t+\lambda;1) = \limsup_{t'\downarrow t} D^+\gamma(t';1).$$

Similarly, since $\gamma_{\infty}(t;1,1)$ is finite, one can apply (2.2.5) of Proposition 2.2.4 to show that

$$0 \le \liminf_{\substack{t'\to t\\ \lambda\downarrow 0}} \left\{D^+\gamma(t'+\lambda;1) - D^+\gamma(t';1)\right\}.$$

Letting $\tau' = t' + \lambda$ we then obtain

$$0 \ge \limsup_{\substack{\tau'\to t\\ \lambda\downarrow 0}} \left\{D^+\gamma(\tau'-\lambda;1) - D^+\gamma(\tau';1)\right\}$$

$$\ge \limsup_{\lambda\downarrow 0} D^+\gamma(t-\lambda;1) - D^+\gamma(t;1)$$

and so

$$\text{(2.6.2)}\quad D^+\gamma(t;1) \ge \limsup_{\lambda\downarrow 0} D^+\gamma(t-\lambda;1) = \limsup_{t'\uparrow t} D^+\gamma(t';1).$$

Together with (2.6.1) we have

$$D^+\gamma(t;1) \ge \limsup_{t'\to t} D^+\gamma(t';1),$$

showing that $D^+\gamma(\cdot;1)$ is upper semi-continuous on (0,1).

In the case (ii) $D^+f(\cdot;y-x)$ is upper semi-continuous on (x,y) since $f^0_{y-x}(z;y-x)$ is clearly so. Consequently Theorem 2.4.2 is applicable.

For the case (iii), let $z = x + t(y-x)$, $t \in (0,1)$. Then, by the regularity of f, Lemmas 2.2.2 and 2.4.1 one has

$$f'(z;y-x) = f^0(z;y-x) = \limsup_{z'\to z} f'(z';y-x)$$

$$\ge \limsup_{t'\to t} f'(x+t'(y-x);y-x) = f^0_{y-x}(z;y-x)$$

showing that

$$f'(z;y-x) = f^0_{y-x}(z;y-x)$$

for any $z \in (x,y)$. Thus, the result holds from the case (ii). □

Corollary 2.6.3. Let $-f\colon [x,y] \longrightarrow \mathbb{R}$ satisfy the assumptions in Theorem 2.4.2. Then there exists $t_0 \in (0,1)$ such that

$$\text{(2.6.1)}\quad \frac{1}{2}f^{\infty}_{y-x}(x+t_0(y-x);y-x,y-x) \ge f(y) - f(x) - f_{0,+(y-x)}(x;y-x)$$

$$\ge \frac{1}{2}f_{\infty,y-x}(x+t_0(y-x);y-x,y-x)$$

and so (2.4.3) holds, where $f_{0,+(y-x)}(x;y-x) = -(-f)^0_{+(y-x)}(x;y-x)$.

Proof: By Theorem 2.4.2, we have

$$\frac{1}{2}(-f)^{\infty}_{y-x}(x+t_0(y-x);y-x,y-x) \geq f(x) - f(y) - (-f)^{0}_{+(y-x)}(x;y-x)$$
$$\geq \frac{1}{2}(-f)_{\infty,y-x}(x+t_0(y-x);y-x,y-x)$$

and so, by elementary results similar to (iii) of proposition 2.2.1,

$$\frac{1}{2}f^{\infty}_{y-x}(x+t_0(y-x);y-x,y-x) \geq f(y) - f(x) - f_{0,+(y-x)}(x;y-x)$$
$$\geq \frac{1}{2}f_{\infty,y-x}(x+t_0(y-x);y-x,y-x).$$

This implies immediately that

$$\frac{1}{2}f^{\infty}_{y-x}(x+t_0(y-x);y-x,y-x) + f^{0}_{y-x}(x;y-x) \geq f(y) - f(x)$$
$$\geq f_{0,y-x}(x;y-x) + \frac{1}{2}f_{\infty,y-x}(x+t_0(y-x);y-x,y-x). \quad \square$$

Remark 1: By [1, prop. 2.3.6], it follows from part (ii) of Corollary 2.6.2 and Corollary 2.6.3 that a convex function satisfies (2.4.2) and a concave function satisfies (2.6.1) respectively, and both satisfy (2.4.3).

Remark 2: In each of the cases (i) — (iii), it is well-known that f can fail to have Gâteaux derivative at some points so [2. Prop. 4.1] is not applicable.

7. Some applications in optimization.

Definition 2.7.1. Let $f: X \longrightarrow \mathbb{R}$ and $x \in X$. $\partial^2 f(x)$ will be said to be positively definite [2] if $f_{\infty}(x;u,u) > 0$ for every $u \in X$, $u \neq 0$. Furthermore, a function $f: X \longrightarrow \mathbb{R}$ is called twice uniformly locally Lipschitzian at x [2] if there exist neighbourhoods X_0 of x and U of zero such that $f^{\infty}(X_0;U,U)$ is bounded in $\mathbb{R}$. This condition implies in particular that f is twice C-differentiable at each point x_0 in X_0 because then, for each $u \in U$, the sublinear map $v \longmapsto f^{\infty}(x_0;u,v)$ is bounded on U and hence continuous on X.

Proposition 2.7.2. Let $x \in X = \mathbb{R}^n$, $f\colon X \longrightarrow \mathbb{R}$ be locally Lipschitz near x and twice uniformly locally Lipschitzian at x. If $f^0_{+u}(x;u) \geq 0$ for all $u \in X$, then a sufficient condition for x to be a strict local minimum point of f is that $\partial^2 f(x)$ is positively definite.

Proof: By assumption, take a constant $M > 1$ and neighbourhoods X_0 of x and U of zero such that

(2.7.1) $$|f^{\infty}(X_0;U,U)| < M$$

and that f on X_0 is Lipschitz. Let $B\colon = \{u \in X;\ \|u\| = 1\}$ and $u \in B$. By the strict positivity of $f_{\infty}(x;u,u)$ and the lower semi-continuity of $f_{\infty}(\cdot;u,u)$ one has a convex neighbourhood W(u) of x contained in X_0 and $1 > \delta(u) > 0$ such that

$$f_{\infty}(y;u,u) > \delta(u)$$

for all $y \in W(u)$. Let $1 > \lambda > 0$ with $\lambda u \in U$ and $U(u) = \frac{\lambda\delta(u)}{8M}U$. For any $v \in u + U(u)$, $y \in W(u)$, it follows from proposition 2.1.1 that

$$\begin{aligned} f_{\infty}(y;v,v) &= f_{\infty}(y;u+(v-u),u+(v-u)) \\ &\geq f_{\infty}(y;u,u) + f_{\infty}(y;v-u,v-u) + 2f_{\infty}(y;u,v-u) \\ &> \delta(u) - \frac{\lambda^2\delta(u)^2}{8^2M^2}\cdot M - \frac{\delta(u)}{4} \geq \frac{\delta(u)}{2} > 0. \end{aligned}$$

Since $X = \mathbb{R}^n$, by the compactness of B we can choose m neighbourhoods $u_1 + U(u_1), \cdots, u_m + U(u_m)$ whose union covers B. Let

$$W = \bigcap_{i=1}^{m} W(u_i) \text{ and } \delta = \min_{1\leq i\leq m}\{\delta(u_i)\}$$

Then for any $v \in B$, $y \in W$,

$$f_\infty(y;v,v) > \delta/2;$$

consequently $f_\infty(y;v,v) > 0$ for all $v \in X$, and $y \in W$. In view of the assumption (2.7.1), it follows from part (i) of Coorollary 2.6.2, that for any $y \in W$, $y \neq x$, there exists $t_0 \in (0,1)$ such that

$$f(y) - f(x) \geq f^0_{+(y-x)}(x;y-x) + \frac{1}{2}f_\infty(y+t_0(x-y);y-x,y-x) > 0$$

because $f^0_{+(y-x)}(x;y-x) \geq 0$ and $y+t_0(x-y) \in W$. Therefore x is a strictly local minimal point. □

Remark: The preceding proposition can be deduced from [2, Prop. 5.2] because the twice uniformly locally Lipschitzian of f implies $f \in C^{1,1}$ [18]. We are indebted to the referee for the reference [18].

Let $f\colon X \longrightarrow \mathbb{R}$ and $g\colon X \longrightarrow \mathbb{R}^n$ be locally Lipschitz functions, C be a closed subset of X.

Now we consider the minimization problem with constaint:

(ℙ) $\min\{f(x);\ x \in Q\}$,

where $Q\colon = \{x \in C \text{ and } g(x) \leq 0\}$. If x_0 is a solution of problem (ℙ), then by Theorem 1.3.4 there exists a multiplier

$$(\lambda,\gamma) \in \mathbb{R}^1 \times \mathbb{R}^n$$

with λ, $\gamma_i \geq 0$, $1 \leq i \leq n$ and $\lambda + \sum_{i=1}^{n}\gamma_i = 1$ such that

(2.7.2) $\gamma g(x_0) = 0 \text{ and } 0 \leq L^0(x_0;u)$

for any $u \in X$, where

$$L(x)\colon = \lambda f(x) + \gamma g(x) + \alpha d_Q(x)$$

and α is a Lipschitzian constant for both f and g on a neighbourhood of x_0.

Propposition 2.7.3. Suppose that x_0 is a solution of the problem ($\mathbb{P}$). Let A: = $\{v;\ \gamma g(v) \geq 0\}$ with the contingent cone $T_A(x_0)$ [3]. Then

(i) $L^{\infty}(x_0;u,u) \geq 0$, for any u in $T_A(x_0)$ with $D_+L(x_0;u) = 0$;

(ii) $L^{oo}(x_0;u,u) \geq 0$, for any u in $T_A(x_0)$ with $L^o(x_0;u) = 0$.

Proof: (i) Since $f(x_0) \leq f(x)$ for any $x \in Q$, by [1, Prop. 2.4.3] $f + \alpha d_Q$ attains a local minimum at x_0. Let $u \in T_A(x_0)$ with $D_+L(x_0;u) = 0$ and take sequences $u_i \to u$ and $t_i \downarrow 0$ with $x_0 + t_iu_i \in A$. Therefore, one has

$$\begin{aligned}
&L(x_0+t_iu_i) - L(x_0)\\
&= \lambda f(x_0+t_iu_i) + \gamma g(x_0+t_iu_i) + \alpha d_Q(x_0+t_iu_i) - \lambda f(x_0)\\
&= \lambda\left\{f(x_0+t_iu_i) + \alpha d_Q(x_0+t_iu_i) - f(x_0)\right\} + \gamma g(x_0+t_iu_i)\\
&\qquad + (1-\lambda)\alpha d_Q(x_0+t_iu_i) \geq 0.
\end{aligned}$$

By Lemma 2.2.2 there exists $\tau_i \in (0,t_i)$ such that

$$D_+L(x_0+\tau_iu;u) \geq \frac{1}{t_i}(L(x_0+t_iu) - L(x_0)) \geq 0$$

Therefore $\limsup_{\tau\downarrow 0} \frac{1}{\tau}D_+L(x_0+\tau u;u) \geq 0$. Since $D_+L(x_0;u) = 0$ it follows from Proposition 2.2.4 that

$$\begin{aligned}
L^{\infty}(x_0;u,u) &= \limsup_{\substack{y\to x_0\\ t\downarrow 0}} \frac{1}{t}(D_+L(y+tu;u) - D_+L(y;u))\\
&\geq \limsup_{t\downarrow 0} \frac{1}{t}(D_+L(x_0+tu;u) - D_+L(x_0;u))\\
&= \limsup_{t\downarrow 0} \frac{1}{t}D_+L(x_0+tu;u) \geq 0.
\end{aligned}$$

(ii) By definition 2.2.5 and similar proof of part (i), one has

$$\begin{aligned}
L^{oo}(x_0;u,u) &= \limsup_{\substack{y\to x_0\\ t\downarrow 0}} \frac{1}{t}(L^o(y+tu;u) - L^o(y;u))\\
&\geq \limsup_{t\downarrow 0} \frac{1}{t}(L^o(x_0+tu;u) - L^o(x_0;u))
\end{aligned}$$

$$\geq \limsup_{t \downarrow 0} \frac{1}{t} D_+ L(x_o + tu; u) \geq 0. \qquad \square$$

References

[1] F. H. Clarke, Optimization and Nonsmooth Analysis, John Wiley, New York, 1983.

[2] R. Cominetti and R. Correa, A generalized second-order derivative in nonsmooth optimization, SIAM J. Control Optim., Vol.28, No. 4(1990), 789-809.

[3] J-P. Aubin and I. Ekeland, Applied Nonlinear Analysis, John Wiley, New York, 1984.

[4] J-B. Hiriart-Urruty, S-J. Strodiot and V. Hien Hguyen, Generalized Hessian matrix and second-order optimality conditions for problems with $c^{1,1}$ data, Appl. Math. Optim, 11(1984), 43-56.

[5] A. Auslender and R. Cominetti, A comparative study of multifunction differentiability with applications in mathematical programming, Math. Oper. Res. 16 (1991), 240-258.

[6] A. Ben-Tal and J. Zowe, A unified theory of first and second order conditions for extremum problems in topological vector spaces, Math. Programming Stud., 19 (1982), 39-76.

[7] R. W. Chaney, Second-order sufficiency conditions for nondifferentiable programmming problems, SIAM J. Control Opt. 20 (1982), 20-33.

[8] R. W. Chaney, Second-order necessary conditions in constrained semismooth optimization, SIAM J. Control Opt, 25 (1987), 1072-1081.

[9] A. D. Ioffe, Approximate subdifferentials and applications. I: The finite dimensional theory, Trans. Amer. Math. Soc. 281

(1984), 389-416.

[10] A. D. Ioffe, Calculus of Dini subdifferentials of functions and contient coderivatives of set-valued maps, Nonlinear Anal. Theor. Meth. Appl. 8 (1984), 517-539.

[11] M. Pappalardo, Tangent cones and derivatives, J. Optim. Th. Appl. 70 (1991), 97-107.

[12] B. N. Pshenichnyi, Necessary Conditions for an Extremum, Marcel Dekker, New York, 1983.

[13] R. T. Rockafellar, Directionally Lipschitzian functions and subdifferential Calculus, Proc. London Math. Soc. (3) 39 (1979), 331-355.

[14] R. T. Rockafellar, Generalized directional derivatives and subgradients of nonconvex functions, Canad. J. Math. Soc. 37 (1980), 257-280.

[15] R. T. Rockafellar, First- and second-order epi-differentiability in nonlinear programming, Trans. Amer. Math. Soc. 307 (1988), 75-108.

[16] R. T. Rockafellar, Generalized second derivatives of convex functions and saddle functions, Trans. Amer. Math. Soc. 322 (1990), 51-77.

[17] A. J. White, Real Analysis: an Introduction, Addison-Wesley, 1968.

[18] R. Cominetti, Equivalence between the classes of $C^{1,1}$ and twice locally Lipschitzian functions, in Thesis, Université Blaise Pascal (1989).

Chapter 3. Second-order necessary and sufficient conditions in nonsmooth optimization

1. Introduction.

In this Chapter, we consider a locally Lipschitz real-valued function f on a normed space X. It is well-known that if $\bar{x}$ is local minimum point of f then the lower Dini-directional derivative $D_+f(\bar{x};u)$ must be nonnegative at each direction u. If $D_+f(\bar{x};u) = 0$ for some u, we show (Theorem 3.2.2) that 0 must be in the Chaney's subdifferential $\partial_u f(\bar{x})$ of f at $\bar{x}$ in the direction u and the second order directional derivative $f''_-(\bar{x},0,u) \geq 0$ (see below for definitions). This result was proved by Chaney in [5] in the special case when $X = \mathbb{R}^n$ under an additional semismooth assumption of f. Likewise, for $\bar{x}$ to be a local minimum point of f with inequality and equality constraints the following conditions are shown to be necessary (in Theorem 3.3.6) : (i) $G''_-(\bar{x},0,u) \geq 0$, and (ii) $L''_+(\bar{x},\bar{w},0,u) \geq 0$ for some Lagrange multiplier $\bar{w}$ in $M_u(\bar{x})$ whenever $D_+G(\cdot,f)(\bar{x};u) = 0$. Further, for $X = \mathbb{R}^n$, the complementary results on sufficient optimality conditions for unconstrained/constrained problems are obtained and thereby the related results in [6] are not only generalized (to not necessarily semismooth functions) but also the conclusions are considerably sharpened. In constrast to Chaney's approach, our arguments rely heavily on Ekeland's variational principle [1] and a result of Ioffe's Proposition [10] (as well as its generalization presented in lemma 3.4.1).

We turn now to some definitions and notations which are mostly taken from Chaney's papers [5, 6]. Let X be a normed space, W an open subset of X and f a locally Lipschitz function of X into $\mathbb{R}$. Denote the unit ball of X by B_1, the open and closed balls centred at x with radius δ by

$$B(x,\delta): = \{y;\ \|y - x\| < \delta\} \text{ and } B[x,\delta]: = \{y;\ \|y - x\| \le \delta\}$$

respectively. Recall that the lower and upper Dini-directional derivatives at $x \in X$ in the direction $u \in X$ are defined by

$$D_+f(x;u) = \liminf_{\substack{u' \to u \\ t \downarrow 0}} \frac{1}{t}\{f(x+tu') - f(x)\}$$

and

$$D^+f(x;u) = \limsup_{\substack{u' \to u \\ t \downarrow 0}} \frac{1}{t}\{f(x+tu') - f(x)\}$$

respectively. In the case of f being locally Lipschitz function, we have

$$D_+f(x;u) = \liminf_{t \downarrow 0} \frac{1}{t}\{f(x+tu) - f(x)\}$$

and

$$D^+f(x;u) = \limsup_{t \downarrow 0} \frac{1}{t}\{f(x+tu) - f(x)\}.$$

Definition 3.1.1. Let u be a nonzero vector in X. Suppose that the sequence $\{x_k\}$ in X converges to x. We say that x_k converges to x in the direction u, denoted by $x_k \longrightarrow_u x$, if the sequence $\frac{x_k - x}{\|x_k - x\|}$ converges to $u/\|u\|$.

Definition 3.1.2. Let u be a nonzero vector in X. As in [5,6] define the subset $\partial_u f(x)$ of the dual space X^* of X by

$$\partial_u f(x) := \{x^*: \text{there exist sequences } x_k \text{ and } x_k^* \in \partial f(x_k) \text{ such that } x_k \rightarrow_u x \text{ and } x_k^* \longrightarrow x^* \text{ in norm respectively}\},$$

where $\partial f(y)$ denotes the Clarke's subdifferential of f at y. Thus, $\partial_u f(x)$ is a subset of $\partial f(x)$; and $\partial_u f(x)$ is nonempty if $X = \mathbb{R}^n$. We emphasize that though the convergence of $x_k^* \in \partial f(x_k)$ are usually considered in the w^*-topology of X^* in similar situations, but here, it is considered in norm topology of X^*.

Definition 3.1.3. Let u be a nonzero vector in X. Suppose that $x^* \in \partial_u f(x)$. Then $f''_-(x,x^*,u)$ is defined to be the infimum of all numbers

$$\liminf \frac{1}{t_k^2}\{f(x_k) - f(x) - x^*(x_k-x)\},$$

taken over all triples of sequences x_k, x_k^*, and t_k for which

(a) $t_k > 0$ for each k and x_k converges to x,

(b) t_k converges to 0 and $(x_k-x)/t_k$ converges to u,

(c) x_k^* converges to x^* with x_k^* in $\partial f(x_k)$ for each k.

Similarly, we define $f''_+(x, x^*, u)$ to be the supremum of all numbers

$$\limsup \frac{1}{t_k^2}\{f(x_k) - f(x) - x^*(x_k-x)\},$$

taken over all triples of sequences x_k, x_k^*, and t_k for which (a), (b), and (c) above all hold.

Remark: By (b), we see that

$$(x_k-x)/\|x_k-x\| = [(x_k-x)/t_k]\cdot[t_k/\|x_k-x\|] \longrightarrow u/\|u\|,$$

that is, x_k converges to x in the direction u. Thus,

$$\liminf \frac{1}{t_k^2}\{f(x_k) - f(x) - x^*(x_k-x)\}$$

$$= \liminf \frac{1}{\|x_k-x\|^2}\{f(x_k) - f(x) - x^*(x_k-x)\}\cdot\frac{\|x_k-x\|^2}{t_k^2}$$

$$= \|u\|^2 \liminf \frac{1}{\|x_k-x\|^2}\{f(x_k) - f(x) - x^*(x_k-x)\}.$$

Hence, $f''_-(x,x^*,u)$ equals to the infimum of all numbers

$$\|u\|^2 \liminf \{f(x_k) - f(x) - x^*(x_k-x)\}/\|x_k-x\|^2,$$

taken over the set of all sequences x_k such that both

(a') x_k converges to x in the direction u

and

(b') there exists a sequence $x_k^* \in \partial f(x_k)$ converging to x^*.

2. Second-order necessary and sufficient conditions without constraint.

In this section we consider the problem of minimizing $f(x)$, over all x in W.

Lemma 3.2.1. Suppose that $f(x) \geq f(\bar{x})$ for all $x \in B[\bar{x},\delta]$. Let $0 \neq u \in X$, $t > 0$, $\alpha > 1$ and $0 < \varepsilon < (\alpha\|u\|)^2$ such that

(3.2.1) $$f(\bar{x}+tu) - f(\bar{x}) \leq t\varepsilon \text{ and } t(\|u\| + \varepsilon^{1/2}) < \delta.$$

Then there exist $z \neq \bar{x}$ in X and $z^* \in \partial f(z)$ such that

(i) $\|z-\bar{x}-tu\| \leq t\varepsilon^{1/2}\alpha^{-1}$ $(< t\varepsilon^{1/2})$,

(ii) $f(z) \leq f(\bar{x}+tu)$

and

(iii) $\|z^*\| \leq \alpha\varepsilon^{1/2}$.

Proof. Let B denote the closed ball with center $\bar{x}+tu$ and radius $t\varepsilon^{1/2}$. Then B is contained in $B(\bar{x},\delta)$ by the second inequality in (3.2.1) and hence

$$f(\bar{x}+tu) \le \inf_{x\in B} f(x) + t\varepsilon$$

by other assumptions of theorem. It follows from the Ekeland's variational principle [1] with $\lambda = t\varepsilon^{1/2}\alpha^{-1}$ that there exists $z \in B$ satisfying (i), (ii) and

(iv) $f(z) \le f(y) + (\alpha\varepsilon^{1/2})\|z-y\|$ for all $y \in B$.

By (i), $z \ne \bar{x}$ since $\varepsilon < (\alpha\|u\|)^2$, and z is in the interior of B since $\alpha > 1$. By basic calculus for subdifferentials (see, Theorem 1.3.2 and 1.3.3) it then follows from (iv) that $0 \in \partial f(z) + \alpha\varepsilon^{1/2}B_1^*$ where B_1^* denotes the unit ball in X^*. Thus (iii) holds for some $z^* \in \partial f(z)$.

□

The following theorem 3.2.2 provides first and second-order necessary conditions in nonsmooth optimization without constraint.

Theorem 3.2.2. Suppose that $\bar{x}$ is a local minimum point for f and $u \in X$ with norm 1 such that $D_+f(\bar{x};u) = 0$. Then $0 \in \partial_u f(\bar{x})$, and $f''_-(\bar{x},0,u) \ge 0$.

Proof. Let $\alpha = 2$ and $\varepsilon \in (0,1)$. Since $D_+f(\bar{x};u) < \varepsilon$, there exists an arbitrarily small $t > 0$ satisfying (3.2.1) with $\delta > 0$ being the same as in Lemma 3.2.1. Thus there exist z, z^* satisfying the properties stated in Lemma 3.2.1. Take a sequence $\varepsilon_k \downarrow 0$. We apply the above to $t_k \downarrow 0$ (for $t_k = t$) to obtain z_k, z_k^* satisfying the properties stated in Lemma 3.2.1 (for z and z^*). In particular (i) reads as

$$\|z_k-\bar{x}-t_k u\| \le t_k \varepsilon_k^{1/2}.$$

Dividing both side by t_k it follows that

$$[z_k - \bar{x}]/t_k \longrightarrow u,$$

showing that $z_k \longrightarrow \bar{x}$ in the direction u. Since $z_k^* \in \partial f(z_k)$ and $\|z_k^*\| \le 2\varepsilon_k^{1/2}$ by (iii), it follows that $0 \in \partial_u f(\bar{x})$. Further, $f''_-(\bar{x},0,u)$ is then defined by definition 3.1.3 and in fact it is non-negative since $\bar{x}$ is a local minimum point. □

In order to compare Chaney's theorem [7, theorem 1] with our result, we let $X = \mathbb{R}^n$ and define the sets in $\mathbb{R}^n$ by

$$D^*(x,f) := \{u \in \mathbb{R}^n;\ \exists\delta(u) > 0 \text{ such that } v\cdot u \le 0 \text{ for all } \|w - u\| \le \delta(u) \text{ and } v \in \partial_w f(x)\}$$

and

$$D^\#(x,f) = \{u \in \mathbb{R}^n;\ \langle v,u\rangle \le 0 \text{ for all } v \in \partial_u f(x)\}.$$

Thus, one has $D^*(x,f) \subseteq D^\#(x,f)$.

Lemma 3.2.3. (i) For any $x, u \in \mathbb{R}^n$, there exist w^+ and w_+ in $\partial_u f(x)$ such that

$$\langle w^+,u\rangle = D^+f(x;u) \text{ and } \langle w_+,u\rangle = D_+f(x;u).$$

(ii) $D^*(x,f) \subseteq D^\#(x,f) \subseteq \{u \in \mathbb{R}^n;\ D_+f(x;u) \le 0\}$.

(iii) If $\bar{x}$ is a local minimum point for $f(x)$, over all x in $\mathbb{R}^n$, then

$$D^*(\bar{x},f) \subseteq D^\#(\bar{x},f) \subseteq \{u \in \mathbb{R}^n;\ D_+f(\bar{x};u) = 0\}.$$

Proof. By Lebourg's mean valued theorem [9, theorem 2.3.7] for any $x, u \in \mathbb{R}^n$ and $t > 0$, there exist some $a_t \in (0, t)$ and some $w_t \in \partial f(x+a_t u)$ such that

$$\frac{1}{t}\{f(x+tu) - f(x)\} = \langle w_t, u\rangle.$$

Hence, one has

$$D^+f(x;u) = \limsup_{t\downarrow 0} \frac{1}{t}\{f(x+tu) - f(x)\} = \limsup_{t\downarrow 0} \langle w_t, u\rangle.$$

Since the multifunction $x \longrightarrow \partial f(x)$ is closed and locally takes values in a compact set by (3) of Theorem 1.3.1, we can choose a sequence $t_n \downarrow 0$ such that $\lim w_{t_n} = w^+ \in \partial f(x)$ and

$$D^+f(x;u) = \lim_{n\to\infty} \langle w_{t_n}, u\rangle = \langle w^+, u\rangle.$$

Since $x + a_{t_n} u$ converges to x in the direction u, one has $w^+ \in \partial_u f(x)$.

Similarly, we can show the corresponding result for $D_+f(x;u)$. Thus, we have shown (i). (ii) follows immediately from (i), and (iii) from (ii) as $D_+f(\bar{x};v) \geq 0$ for all v in $\mathbb{R}^n$ if $\bar{x}$ is a local minimum point of f. □

Therefore, from above Lemma 3.2.3 and Theorem 3.2.2, we arrive at the following result which was proved by Chaney [7, theorem 1] under additional assumption that f is semismooth [11].

Corollary 3.2.4. Suppose that $\bar{x}$ is an unconstrained local minimizer for $f(x)$, over x in $W \subseteq \mathbb{R}^n$. If u belongs to $D^*(\bar{x},f)$, then 0 belongs to $\partial_u f(\bar{x})$ and $f''_-(\bar{x},0,u) \geq 0$.

Since the local minimality assumptions of f at $\bar{x}$ clearly implies $D_+f(\bar{x};v) \geq 0$ for all v, Theorem 3.2.6 below strengthens the first conclusion of Theorem 3.2.2 for the special case when $X = \mathbb{R}^n$. To prepare the proof we need a technical result which follows from Lemma 3.2.1 immediately.

Lemma 3.2.5. Let X be a normed space. Suppose that $\bar{x}$ is a local minimum point for f and $u \in X$ with norm 1 such that $D_+f(\bar{x};u) < \varepsilon$, where $0 \leq \varepsilon$. Let $\alpha^2 > 1, \varepsilon$. Then there exist arbitrarily small $t > 0$, $z \in X \setminus \{\bar{x}\}$ and $z^* \in \partial f(z)$ such that

(i) $\|z-\bar{x}-tu\| \leq t\varepsilon^{1/2}/\alpha$,

(ii) $f(z) \leq f(\bar{x}+tu)$

and

(iii) $\|z^*\| \leq \alpha\varepsilon^{1/2}$.

Specializing in the case when $X = \mathbb{R}^n$, we have:

Theorem 3.2.6. Suppose that $f: \mathbb{R}^n \longrightarrow \mathbb{R}$ is a locally Lipschitz function and $D_+f(\bar{x};v) \geq 0$ for all $v \in \mathbb{R}^n$. For u in $\mathbb{R}^n$ with norm 1, if $D_+f(\bar{x};u) = 0$, then $0 \in \partial_u f(\bar{x})$.

Proof. Let $\alpha > 1$. Since $D_+f(\bar{x};v) \geq 0$ for all $v \in \mathbb{R}^n$, one has by Ioffe's proposition that for any $\varepsilon_k \downarrow 0$ with $0 < \varepsilon_k < \alpha^2/2$, there exists $\delta_k \downarrow 0$ such that $F_{\varepsilon_k}(\bar{x}) \leq F_{\varepsilon_k}(x)$ for all $\|x-\bar{x}\| \leq \delta_k$, where $F_{\varepsilon_k}(x) := f(x) + \varepsilon_k\|x-\bar{x}\|$. Thus, if u is a unit vector with $D_+f(\bar{x};u) = 0$ then

$$D_+F_{\varepsilon_k}(\bar{x};u) = \liminf_{t\downarrow 0} \frac{1}{t}\{F_{\varepsilon_k}(\bar{x}+tu) - F_{\varepsilon_k}(\bar{x})\}$$
$$= \liminf_{t\downarrow 0} \frac{1}{t}\{f(\bar{x}+tu) - f(\bar{x})\} + \varepsilon_k t\} = \varepsilon_k < 2\varepsilon_k$$

for all k. By Lemma 3.2.5 one can find $t_k \downarrow 0$, $z_k \in X \setminus \{\bar{x}\}$, $z_k^* \in \partial F_{\varepsilon_k}(z_k)$ satisfying the properties stated in Lemma 3.2.5 (for function F_{ε_k} instead of f and $2\varepsilon_k$ instead of ε). In particular (i) reads

$$\|z_k-\bar{x}-t_k u\| \le t_k(2\varepsilon_k)^{1/2}/\alpha < t_k(2\varepsilon_k)^{1/2}.$$

Dividing by t_k it follows that

$$[z_k - \bar{x}]/t_k \longrightarrow u$$

showing that $z_k \longrightarrow \bar{x}$ in the direction u. Since

$$z_k^* \in \partial F_{\varepsilon_k}(z_k) \subseteq \partial f(z_k) + \varepsilon_k B_1^* \text{ and } \|z_k^*\| \le \alpha(2\varepsilon_k)^{1/2}$$

by (iii), it follows that $0 \in \partial_u f(\bar{x})$. □

The converse of Theorem 3.2.6 is false as the following example shows.

Example 3.2.7. Let ξ: $\mathbb{R} \longrightarrow \mathbb{R}$ be defined by

$$\xi(x) = \begin{cases} x^2 \sin x^{-1} & x \ne 0 \\ 0 & x = 0 \end{cases}$$

and f: $\mathbb{R} \longrightarrow \mathbb{R}$ by $f(x) = \log(1+|x|+|\xi(x)|)$.

Let $x_k = 1/2k\pi$, k = ±1, ±2, $\cdots$. Since the derivative of f at

$$x \notin \{0,\ 1/k\pi : k = \pm 1, \pm 2, \cdots\}$$

is

$$f'(x) = \frac{1}{1+|x|+|\xi(x)|}\{\mathrm{sgn}(x) + 2x|\sin x^{-1}| - \mathrm{sgn}(\sin x^{-1})\cdot\cos x^{-1}\}.$$

One has by [8, Lemma 1.5] that,

$$\begin{aligned} f^0(x_k;1) &= \limsup_{x\to x_k} f'(x;1) \\ &= \left(1+\frac{1}{2k\pi}\right)^{-1}\left[1 - \liminf_{x\to x_k} \mathrm{sgn}(\sin x^{-1})\right] = 2\cdot\left(1+\frac{1}{2k\pi}\right)^{-1}; \end{aligned}$$

$$f^0(x_k;-1) = \limsup_{x\to x_k} f'(x;-1)$$

$$= -(1+\frac{1}{2k\pi})^{-1}[1 - \limsup_{x\to x_k} \mathrm{sgn}(\sin x^{-1})] = 0,$$

for all $k = 1, 2, \cdots$. Hence, the subdifferential of f at x_k is

$$\partial f(x_k) = [0,2(1+1/2k\pi)^{-1}],\ k = 1, 2, \cdots.$$

Similarly, we have

$$\partial f(x_k) = [-2(1+1/2k\pi)^{-1},0],\ k = -1, -2, \cdots.$$

This implies that

$$0 \in \partial_1 f(0) \text{ and } 0 \in \partial_{-1} f(0).$$

But

$$D_+f(0;1) = \liminf_{t\downarrow 0} \frac{1}{t}\log(1+t+t^2|\sin t^{-1}|) = \lim_{t\downarrow 0} \frac{1}{t}\log(1+t) = 1$$

and

$$D_+f(0;-1) = \liminf_{t\downarrow 0} \frac{1}{t}\log(1+t+t^2|\sin (-t^{-1})|) = \lim_{t\downarrow 0} \frac{1}{t}\log(1+t) = 1.$$

Theorem 3.2.8. (Second-order sufficient conditions without constraint) Suppose that $f: \mathbb{R}^n \longrightarrow \mathbb{R}$ is a locally Lipschitz function. Suppose that $D_+f(\bar{x};u) \geq 0$ for all unit vectors u in $\mathbb{R}^n$. If $f''_-(\bar{x},0,u) > 0$ for all unit vectors u in $\mathbb{R}^n$ for which $D_+f(\bar{x};u) = 0$, then there exists $\delta > 0$ such that $f(x) > f(\bar{x})$ for all $\|x-\bar{x}\| \leq \delta$.

Proof. Suppose that the desired conclusion is false. Then, by continuity, one has for each $\gamma > 0$ that f attains its minimum on $B[\bar{x},\gamma]$ at some point $y_\gamma \in B[\bar{x},\gamma] \setminus \{\bar{x}\}$. We have two cases to consider.

(I) Suppose that there exists a sequence $\gamma_k \downarrow 0$ such that each y_{γ_k} is in the open ball $B(\bar{x},\gamma_k)$. Then $0 \in \partial f(y_{\gamma_k})$ by Theorem 1.3.3 and the definition of ∂f. By passing to a subsequence, we can assume that

y_{γ_k} converges to $\bar{x}$ in a direction u. Hence, $0 \in \partial_u f(\bar{x})$. By assumption,

$$0 \le D_+f(\bar{x};u) \le \liminf_{k\to\infty} [f(y_{\gamma_k}) - f(\bar{x})]/\|y_{\gamma_k} - \bar{x}\| \le 0$$

since $f(y_{\gamma_k}) \le f(\bar{x})$ for all k. Thus $D_+f(\bar{x};u) = 0$ and so $f''_-(\bar{x},0,u) > 0$ by hypothesis. But by definition of f''_-,

$$f''_-(\bar{x},0,u) \le \liminf_{k\to\infty} [f(y_{\gamma_k}) - f(\bar{x})]/\|y_{\gamma_k} - \bar{x}\|^2 \le 0.$$

Therefore this case cannot happen.

(II) Suppose that there exists $\gamma' > 0$ with

(3.2.1) $$\|y_\gamma - \bar{x}\| = \gamma$$

for all $0 < \gamma \le \gamma'$. Since $D_+f(\bar{x};u) \ge 0$ for all unit vectors in $\mathbb{R}^n$ and by Ioffe's proposition, for any $\varepsilon > 0$, there exists $\gamma'' > 0$ such that

(3.2.2) $$f(\bar{x}) \le f(x) + \varepsilon\|x - \bar{x}\|$$

for all $\|x-\bar{x}\| \le \gamma''$. We let

$$\gamma_0 = \min\{\gamma', \gamma''\} \text{ and } \gamma_k = \gamma_0/2^k,\ k = 1, 2, \cdots.$$

By (3.2.1), (3.2.2) and the definition of y_γ, one has

(1) $\|y_{\gamma_{k+1}} - \bar{x}\| = \gamma_{k+1} = \gamma_k/2$ for $k = 0, 1, \cdots$.

(2) $f(y_{\gamma_{k+1}}) \le f(\bar{x}) \le f(x) + \varepsilon\gamma_k$ for all $x \in B[\bar{x},\gamma_k]$ and $k = 0, 1, \cdots$.

It follows from the Ekeland's variational principle (with $\lambda = \gamma_{k+2}$) that there exists $x_{k+1} \in B[\bar{x},\gamma_k]$ such that

(i) $\|x_{k+1} - y_{\gamma_{k+1}}\| \le \gamma_{k+2}$ $(= \gamma_k/4)$;

(ii) $f(x_{k+1}) \le f(y_{\gamma_{k+1}})$;

(iii) $f(x_{k+1}) \le f(x) + 4\varepsilon\|x - x_{k+1}\|$

for all $x \in B[\bar{x},\gamma_k]$. It follows from (2) and (ii) that

(3.2.3) $$f(x_{k+1}) \le f(\bar{x})$$

and from (1) and (i) that $x_{k+1} \ne \bar{x}$ and $x_{k+1} \in B(\bar{x},\gamma_k)$. Together with (iii) we have $0 \in \partial f(x_{k+1}) + 4\varepsilon B_1^*$ by Theorem 1.3.3 and 1.3.2 and so we obtain

(3.2.4) $$y_k^* \in \partial f(x_{k+1}) \text{ with } \|y_k^*\| \le 4\varepsilon.$$

Thus we have constructed for any $\varepsilon > 0$, sequences x_k and $y_k^* \in \partial f(x_k)$ such that $\bar{x} \ne x_k \to \bar{x}$, $f(x_k) \le f(\bar{x})$ and $\|y_k^*\| \le 4\varepsilon$.

Now we let $\varepsilon_n \downarrow 0$. We can inductively choose a sequence $x(n)$ convergent to $\bar{x}$ in some direction u and a sequence $y^*(n) \in \partial f(x_n)$ such that $f(x(n)) \le f(\bar{x})$ and $y^*(n)$ converges to 0 (and so $0 \in \partial_u f(\bar{x})$).

Note that for this u, one has

$$D_+f(\bar{x};u) \le \liminf_{n\to\infty} \frac{1}{\|x(n) - \bar{x}\|}\{f(x(n)) - f(\bar{x})\} \le 0$$

showing that $D_+f(\bar{x};u) = 0$ by assumption. By Theorem 3.2.6, $0 \in \partial_u f(\bar{x})$ and $f''_-(\bar{x},0,u)$ is meaningfully defined and

$$f''_-(\bar{x},0,u) \le \liminf_{n\to\infty}[f(x(n)) - f(\bar{x})]/\|x(n)-\bar{x}\|^2 \le 0$$

since $f(x_n) \le f(\bar{x})$ for all n. This contradicts an given assumption of the theorem and so we complete the proof. □

Observe that if $v\cdot u \ge 0$ for all unit vector u in $\mathbb{R}^n$ and all v in $\partial_u f(\bar{x})$, then $D_+f(\bar{x};u) \ge 0$ for all unit vectors in $\mathbb{R}^n$. This is because, for any unit vector u in $\mathbb{R}^n$ and by lemma 3.2.3 there exists $w_+ \in \partial_u f(\bar{x})$ such that $D_+f(\bar{x};u) = \langle w_+,u\rangle$. Theorem 1 of Chaney in [6] is a weak form of the following result where be assumed the following stronger condition in place of (ii):

(ii)* $f''_-(\bar{x},0,u) > 0$ for all unit vectors u in $\mathbb{R}^n$ for which

$0 \in \partial_u f(\bar{x})$.

Corollary 3.2.9. Let $\bar{x} \in \mathbb{R}^n$, and suppose that

(i) $v \cdot u \geq 0$ for all unit vectors u in $\mathbb{R}^n$ and v in $\partial_u f(\bar{x})$.

(ii) $f''_-(\bar{x},o,u) > 0$ for all unit vectors u in $\mathbb{R}^n$ for which

$$D_+ f(\bar{x};u) = 0.$$

Then there exists $\delta > 0$ such that $f(x) > f(\bar{x})$ for all $0 < \|x-\bar{x}\| \leq \delta$.

We end this section with an example of non-semismooth function which in particular shows the situation that our Theorem 3.2.2 and Theorem 3.2.8 can be applied but not [5, Theorem 1] and [6, Theorem 1].

Example 3.2.10. Let $f(x) := g(x) + x^2$, where

$$g(x) = \begin{cases} x^2 |\sin \frac{1}{x}| & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

Then f, g are Locally Lipschitz functions that are not semismooth at 0 [11]. Note that $f'(x) = 2x|\sin \frac{1}{x}| - \operatorname{sgn}(\sin \frac{1}{x}) \cdot \cos \frac{1}{x} + 2x$ for all $x \neq 0$. Let $x_k = -\frac{1}{2k\pi+\pi+(1/k)}$, $k = 1, 2, \cdots$. Then x_k converges to 0 in the direction -1 and $f'(x_k)$ ($\in \partial f(x_k)$) $\longrightarrow 1 \in \partial f(0)$. So $1 \in \partial_{-1} f(0)$. Hence f does not satisfies the conditions of [5, Theorem 1] and [6, Theorem 1]. Since $|g(x)| \leq x^2$, $f(x) \geq 0$ and so $x = 0$ is a minimum point of f. Note that

$$(3.2.4) \qquad D_+ f(0;\pm 1) = \liminf_{t \downarrow 0} \frac{1}{t}\{t^2 |\sin \frac{1}{\pm t}| + t^2\} = 0.$$

Thus one may apply either Theorem 3.2.2 or Theorem 3.2.6 to conclude that $0 \in \partial_{\pm 1} f(0)$. Furthermore, that $x = 0$ is a local minimum point of f can also be seen from Theorem 3.2.8.

3. Second-order necessary conditions with constraints

Let f, g_1, $\cdots$, g_m, $\cdots$, g_{m+p} be real-valued locally Lipschitz functions on an open set W in a normed space X. We consider the following optimization problem $\mathbb{P}(X)$:

$$\text{minimize } f(x)$$

$$\text{subject to } g_i(x) \le 0 \text{ for } i = 1, 2, \cdots, m;$$

$$g_i(x) = 0 \text{ for } i = m+1, \cdots, m+p.$$

In particular, if $X = \mathbb{R}^n$, the n-dimensional real Euclidean space, then above optimization problem is denoted by $\mathbb{P}(\mathbb{R}^n)$. We shall prove necessary conditions theorems for $\mathbb{P}(X)$ and sufficient conditions theorems for problem $\mathbb{P}(\mathbb{R}^n)$, which extend the theorems of Chaney [5, Theorem 2] and [6, Theorem 4, 5] to the case without his assumption of semismoothness. Throughout S denotes the set of all points in W which are feasible for problem $\mathbb{P}(X)$. For $x \in S$, $K_S(x)$ denotes the contingent cone [9] of S at x. I(x) denotes the set of all "active indices" i with $1 \le i \le m$ such that $g_i(x) = 0$. NI(x) denotes the set of all $1 \le i \le m$ such that $g_i(x) < 0$. For convenience, we agree to adopt the same definitions for I(x), NI(x) even if x is not feasible. It is well-known and can easily be verified that each u in $K_S(\bar{x})$ satisfies the "tangential constraints":

$$\begin{cases} D_+g_i(\bar{x};u) \le 0, \text{ for all } i \in I(\bar{x}) \\ D_+g_k(\bar{x};u) \le 0 \le D^+g_k(\bar{x};u) \text{ for all } k = m+1, \cdots, m+p \end{cases}$$

Let T be the set of all vectors $w = (w_0, w_1, \cdots, w_{m+p})$ in

$\mathbb{R}^{1+m+p}$ such that $\sum_{i=0}^{m+p}(w_i)^2 = 1$ and $w_i \geq 0$ for $i = 0, 1, \cdots, m$. Let $L(x,w)$ denote the Lagrangean function on $W \times T$ defined by

$$L(x,w) := w_0 f(x) + \sum_{i=1}^{m+p} w_i g_i(x).$$

We fix $\bar{x} \in S$ and, as in [5, 6], define the function $G_{\bar{x}}$ from W to $\mathbb{R}$ by

$$G_{\bar{x}}(x,f) := \max\{L(x,w) - w_0 f(\bar{x}) : w \in T\}$$

That is,

$$(3.3.1) \qquad G_{\bar{x}}(x,f) = \max\{\sum_{i=0}^{m+p} w_i g_i(x) : w \in T\},$$

where $g_0(x) := f(x) - f(\bar{x})$. Let

$$T_1(\bar{x}) := \{w \in T : w_i = 0, \forall i \in NI(\bar{x})\}$$
$$= \{w \in T : w_i g_i(\bar{x}) = 0, \forall i = 1, 2, \cdots, m+p\}.$$

Thus,

$$G_{\bar{x}}(x,f) = \max\{\sum_{i=0}^{m+p} w_i g_i(x) : w \in T_1(\bar{x})\}.$$

If $\bar{x}$ is clear from the text, we shall simply write G for $G_{\bar{x}}$, and T_1 for $T_1(\bar{x})$. Let

$$M(x) := \{w \in T : 0 \in \partial L(\cdot,w)(x) \text{ and } w_i g_i(x) = 0 \text{ for } i = 1, \cdots, m\}$$

and for any $u \in X$, let

$$M_u(x) := \{w \in M(x) : 0 \in \partial_u L(\cdot,w)(x)\}$$

and

$$M_-(x) := \{w \in T : 0 \leq D_+ L(\cdot,w)(x;v) \text{ for all } v \in X$$
$$\text{and } w_i g_i(x) = 0 \text{ for } i = 1, \cdots, m\}.$$

Note: If the equality constraints are not appear (i.e., $p = 0$) and $g_i(\bar{x}) < 0$ for all $i = 1, \cdots, m$, then the problem $\mathbb{P}(X)$ becomes an unconstrained problem already studied in section 2. We henceforth assume that there is j with $1 \le j \le m + p$ such that $g_j(\bar{x}) = 0$. With this provision, the maximum in (3.3.1) can only possibly be attained at those w in $T_1(\bar{x})$.

Lemma 3.3.1. Let $g_j(\bar{x}) = 0$ for some j with $1 \le j \le m + p$. Then, for all x near $\bar{x}$, one has

$$\begin{aligned} G(x,f) &= \max \{w_0[f(x) - f(\bar{x})] + \sum_{i=1}^{m+p} w_i g_i(x) : w \in T\} \\ &= \max \{w_0[f(x) - f(\bar{x})] + \sum_{i=1}^{m+p} w_i g_i(x) : w \in T_1(\bar{x})\} \\ &= \max \{ \sum_{i=0}^{m+p} w_i g_i(x) : w \in T_1(\bar{x})\} \end{aligned}$$

where $g_0(x) := f(x) - f(\bar{x})$. Moreover, for x near $\bar{x}$, if $w \in T$ is such that

$$G(x,f) = w_0[f(x) - f(\bar{x})] + \sum_{i=1}^{m+p} w_i g_i(x), \tag{3.3.1$'$}$$

then $w \in T_1(\bar{x})$.

Proof. Clearly we need only prove the last assertion. For each $i \in NI(\bar{x})$ we have $g_i(\bar{x}) < g_j(\bar{x})$ and so $g_i(\cdot) < g_j(\cdot)$ on some ball $B(\bar{x},\delta_i)$ with $\delta_i > 0$. Take $\delta = \min \{\delta_i : i \in NI(\bar{x})\}$. Let $x \in B(\bar{x},\delta)$ and $w \in T$ satisfy (3.3.1)$'$. Then, since

$$g_i(x) < g_j(x) \text{ and } 0 \le w_i,$$

it follows from the maximality of G(x,f) in its definition that w_i must be zero. This argument is valid for all i in $NI(\bar{x})$ so $w \in T_1(\bar{x})$. □

For the following lemma we take constant $\gamma > 0$ with the following property:

$$\sum_{i=0}^{m+p} |w_i - s_i| \le \gamma\Big(\sum_{i=0}^{m+p} |w_i - s_i|^2\Big)^{1/2} \ (= \gamma\|w-s\|)$$

for any w, $s \in \mathbb{R}^{1+m+p}$.

Lemma 3.3.2. Suppose that M is a Lipschitzian constant of f, g_i, $i = 1, \cdots, m+p$, on a neighbourhood $B(\bar{x},\delta)$ for some $\delta > 0$. Then for any w, $s \in T$, $t > 0$, $\|v\| = 1$ with y, $y+tv \in B(\bar{x},\delta)$, we have

(i) $L(y+tv,w) - L(y,w) - [L(y+tv,s) - L(y,s)] \le \gamma Mt\|w-s\|$;

(ii) $L^0(\cdot,w)(x;v) - L^0(\cdot,s)(x;v) \le \gamma M\|w-s\|$.

Proof. By definitions one has

$$L(y+tv,w) - L(y+tv,s) = (w_0 - s_0)f(y+tv) + \sum_{i=1}^{m+p} (w_i - s_i)g_i(y+tv)$$

and

$$L(y,w) - L(y,s) = (w_0 - s_0)f(y) + \sum_{i=1}^{m+p} (w_i - s_i)g_i(y).$$

Hence the left menber of (i) is

$$(w_0 - s_0)[f(y+tv) - f(y)] + \sum_{i=1}^{m+p} (w_i - s_i)[g_i(y+tv) - g_i(y)$$

$$\le \sum_{i=0}^{m+p} |w_i - s_i| Mt \le \gamma Mt\Big(\sum_{i=0}^{m+p} |w_i - s_i|^2\Big)^{1/2} = \gamma Mt\|w-s\|;$$

proving (i). (ii) follows immediately from (i) by taking upper limits in

$$\frac{1}{t}\{L(y+tv,w) - L(y,w)\} \le \frac{1}{t}\{L(y+tv,s) - L(y,s)\} + \gamma M\|w-s\|. \qquad \square$$

Our next lemma deals with a relationship between G and L with regard to their subdifferentials.

Lemma 3.3.3. Let $g_j(\bar{x}) = 0$ for some j with $1 \le j \le m+p$. Then, for some $\delta > 0$ and all $x \in B(\bar{x},\delta)$, there exists $w \in T_1(\bar{x})$ such that

(i) $G(x,f) = L(x,w) - w_0 f(\bar{x})$.

(ii) $G^0(\cdot,f)(x;v) \le L^0(\cdot,w)(x;v)$ for all $v \in X$.

(iii) $\partial G(\cdot,f)(x) \subseteq \partial L(\cdot,w)(x)$.

Remark. By Lemma 3.3.1, there exists $\delta > 0$ such that for each $x \in B(\bar{x},\delta)$ there exists $w \in T_1(\bar{x})$ with property (i). The point is that we want this w concurently satisfies (ii) and (iii).

Proof of Lemma 3.3.3. Let $\delta > 0$ be such that each $x \in B(\bar{x},\delta)$ satisfies (i) with some $w \in T_1(\bar{x})$. Thus, for $x \in B(\bar{x},\delta)$ y near x and $t > 0$ near zero such that $y + tv$, $y \in B(\bar{x},\delta)$, there exists $w = w(y,t) \in T_1(\bar{x})$ such that

$$G(y+tv,f) = L(y+tv,w) - w_0 f(\bar{x}). \tag{3.3.2}$$

Since, by definition

$$G(y,f) \ge L(y,w) - w_0 f(\bar{x}),$$

it follows that

$$\frac{1}{t}\{G(y+tv,f) - G(y,f)\} \le \frac{1}{t}\{L(y+tv,w) - L(y,w)\}$$

and consequently

$$G^0(\cdot,f)(x;v) \le \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{1}{t}\{L(y+tv,w(y,t)) - L(y,w(y,t))\}$$

$$= \lim_{n\to\infty} \frac{1}{t_n}\{L(y_n+t_n v,w^n) - L(y_n,w^n)\}, \tag{3.3.3}$$

where $\{y_n\}$, $\{t_n\}$ are appropriate sequences satisfying the last equality with $y_n \longrightarrow x$ and $t_n \downarrow 0$ (and w^n stands for $w(y_n,t_n)$). By compactness of $T_1(\bar{x})$, we assume without of generality that w^n converges to some $\bar{w} \in$

$T_1(\bar{x})$. By (3.3.3) and Lemma 3.3.2 we then have

$$G^0(\cdot,f)(x;v) \le \limsup_{n\to\infty} \frac{1}{t_n}\{[L(y_n+t_nv,\bar{w}) - L(y_n,\bar{w})] + \gamma Mt_n\|w^n-\bar{w}\|\}$$
$$\le L^0(\cdot,\bar{w})(x;v).$$

Thus (ii) holds if w is replaced by $\bar{w}$. With the same replacement (i) and (iii) also hold. Indeed, (iii) follows from (ii) by definitions. For (i), we note from the definition of w^n and (3.3.2) that

$$L(y_n+t_nv,w^n) - w_0^nf(\bar{x}) = G(y_n+t_nv,f) \ge L(y_n+t_nv,u) - u_0f(\bar{x})$$

for all $u = (u_i) \in T$. By continuities, it follows that

$$L(x,\bar{w}) - w_0f(\bar{x}) \ge L(x,u) - u_0f(\bar{x})$$

showing that $G(x,f) = L(x,\bar{w}) - \bar{w}_0f(\bar{x})$. □

Finally, for the sake of easy reference, we list below a few preparative results from Chaney [5, Lemma 1] and for the sake of self-contain, we give their proofs in Lemma A5 of Appendix.

Lemma 3.3.4. [5, lemma 1]. (i) $\bar{x}$ is a strictly local solution to problem $\mathbb{P}(X)$ if and only if we have $G(x,f) > 0 = G(\bar{x},f)$ for all x near $\bar{x}$ with $x \ne \bar{x}$.

(ii) If $\bar{x}$ is a local solution to problem $\mathbb{P}(X)$, then $G(x,f) \ge 0 = G(\bar{x},f)$ for all x near $\bar{x}$.

We are now ready to present the following technical result which will play important role for both necessary and sufficient conditions theorems.

Theorem 3.3.5. Let u be a unit vector in X and suppose that $g_j(\bar{x}) = 0$ for some j.

(i) If $0 \in \partial_u G(\cdot,f)(\bar{x})$ then $M_u(\bar{x})$ is nonempty.

In fact, if $\{x_k\}$ is a sequence convergent to $\bar{x}$ in the direction of u and $\{x_k^*\}$ is a sequence converget to zero and if $x_k^* \in \partial G(\cdot,f)(x_k)$ for each k then there exists sequence $\{w^k\}$ in $T_1(\bar{x})$ with a cluster point $\bar{w}$, and there exists a sequence $\{y_k^*\}$ such that

(a) $G(x_k,f) = w_0^k(f(x_k) - f(\bar{x})) + \sum_{i=1}^{m+p} w_i^k[g_i(x_k)]$;

(b) $y_k^* \in \partial L(\cdot,\bar{w})(x_k)$ for each k;

(c) $\|y_k^*\| \longrightarrow 0$ as $k \longrightarrow \infty$.

Consequently $0 \in \partial_u L(\cdot,\bar{w})(\bar{x})$ and $\bar{w} \in M_u(\bar{x})$.

(ii) If $X = \mathbb{R}^n$ and if u is a unit vector such that $D_+G(\cdot,f)(\bar{x};u) = 0$ then $M_u(\bar{x}) \supseteq M_-(\bar{x})$.

Proof (i). By Lemma 3.3.3, for all large enough k, there exists $w^k \in T_1(\bar{x})$ such that

$$(3.3.4) \qquad \begin{cases} G(x_k,f) = L(x_k,w^k) - w_0^k f(\bar{x}) \\ G^0(\cdot,f)(x_k,v) \le L^0(\cdot,w_k)(x_k;v) \text{ for all } v \in X. \\ \partial G(\cdot,f)(x_k) \subseteq \partial L(\cdot,w^k)(x_k) \end{cases}$$

Thus, by the first equality, (a) holds. By compactness of $T_1(\bar{x})$, we can assume that $\{w^k\}$ converges to some $\bar{w} \in T_1(\bar{x})$. By Lemma 3.3.2 (and the positive homogenuity), we have, for each $v \in X$, that

$$(3.3.5) \qquad L^0(\cdot,w^k)(x_k;v) \le L^0(\cdot,\bar{w})(x_k;v) + \gamma M\|w^k-\bar{w}\|\|v\|.$$

Then

$$(3.3.6) \qquad \partial L(\cdot,w^k) \subseteq \partial L(\cdot,\bar{w})(x_k) + \gamma M\|w^k-\bar{w}\|B_1^*.$$

Indeed, if x^* is a member of the set on the left then we have from

(3.3.5) that

$$x^*(v) - L^0(\cdot,\overline{w})(x_k;v) \le \gamma M\|w^k-\overline{w}\|\,\|v\|$$

for all $v \in X$. By the Separation Theorem, there exists $y^* \in \gamma M\|w^k-\overline{w}\|B_1^*$ such that $x^*(v) - y^*(v) \le L^0(\cdot,\overline{w})(x_k;v)$, for all $v \in X$, i.e. $x^* - y^* \in \partial L(\cdot,\overline{w})(x_k)$, proving (3.3.6).

By definiion of x_k^* it follows from (3.3.4) and (3.3.6) that there exists $y_k^* \in \partial L(\cdot,\overline{w})(x_k)$ such that $x_k^* \in y_k^* + \gamma M\|w^k-\overline{w}\|B_1^*$. Passing to the limits as $k \longrightarrow \infty$ we see that $\|y_k^*\| \longrightarrow 0$.

(ii) Let $w \in M_-(\overline{x})$: $w_i g_i(\overline{x}) = 0$ for all $i = 1, \cdots, m + p$ and $D_+L(\cdot,w)(\overline{x};v) \ge 0$ for all $v \in X = \mathbb{R}^n$. Then $L(\overline{x},w) = w_0 f(\overline{x})$ and, by Lemma 3.3.1, $G(\overline{x},f) = 0$. By definition of G, we also have, for each $t > 0$, that

$$G(\overline{x}+tu,f) \ge L(\overline{x}+tu,w) - w_0 f(\overline{x})$$

i.e.

$$G(\overline{x}+tu,f) - G(\overline{x},f) \ge L(\overline{x}+tu,w) - L(\overline{x},w).$$

Dividing by t and taking lower limits, it follows that

$$D_+G(\cdot,f)(\overline{x};u) \ge D_+L(\cdot,w)(\overline{x};u).$$

Since the left member is assumed zero, the right member must also be zero as $D_+L(\cdot,w)(\overline{x};v) \ge 0$ for all v by assumption. Now, by Theorem 3.2.6 (applicable for $X = \mathbb{R}^n$), $0 \in \partial_u L(\cdot,w)(\overline{x})$. Hence $w \in M_u(\overline{x})$. □

Theorem 3.3.6. (Second-order necessary condition with constraint) Suppose that $\overline{x}$ is a local minimum solution of problem $\mathbb{P}(X)$ and that u is a unit vector with $D_+G(\cdot,f)(\overline{x};u) = 0$ (a fortiori, if $0 \in \partial_u G(\cdot,f)(\overline{x})$). Then $G''_-(\overline{x},0,u) \ge 0$ and there exists a Lagrange multiplier $\overline{w}$ in $M_u(\overline{x})$ such that $L''_+(\overline{x},\overline{w},0,u) \geq 0$.

Remark 1. In constrast to theorem 3.2.2, the conclusion in theorem 3.3.5 cannot be strengthened to $L''_-(\overline{x},\overline{w},0,u) \geq 0$ even if $X = \mathbb{R}^n$, and f and all g_i are semismooth (see [5]).

Remark 2. If the equality constraints are not appear (i.e. p = 0) and $g_i(\overline{x}) < 0$ for all $i = 1, 2, \cdots, m$, then all x near to $\overline{x}$ are feasible for problem ($\mathbb{P}$) and hence $f(x) \geq f(\overline{x})$ by assumption. Therefore, similar arguments given for Lemma 3.3.1 show that $G(x,f) = f(x) - f(\overline{x})$ and $L(x,w) = f(x)$ for $w = (1, 0, \cdots, 0)$ the only element in $M(\overline{x})$. Thus, in this case theorem 3.3.6 follows from theorem 3.2.2.

Proof of Theorem 3.3.6. In view of the preceding Remark 2, we can assume that there is j with $g_j(\overline{x}) = 0$. By Lemma 3.3.4 (ii), $\overline{x}$ is a local minimum point (with value 0) for $G(\cdot,f)$ as it is so for problem ($\mathbb{P}$). By assumption of $D_+G(\cdot,f)(\overline{x};u) = 0$ it follows from Theorem 3.2.2 that

$$0 \in \partial_u G(\cdot,f)(\overline{x}), \text{ and } G''_-(\overline{x},0,u) \geq 0.$$

Take a sequence $\{x_k\}$ convergent to $\overline{x}$ in the direction u, and take $x^*_k \in \partial G(\cdot,f)(x_k)$ for each k such that $\|x^*_k\| \longrightarrow 0$. In view of Theorem 3.3.5, we can further assume that there are sequence $\{w^k\}$ in $T_1(\overline{x})$ convergent to $\overline{w}$ and a sequence $\{y^*_k\}$ such that (a), (b), (c) of Theorem 3.3.5 (i) hold. In particular, by (b) and (c), $0 \in \partial_u L(\cdot,\overline{w})(\overline{x})$ and consequently $\overline{w} \in M_u(\overline{x})$ as $\overline{w} \in T_1(\overline{x})$. It remains to show that $L''_+(\overline{x},\overline{w},0,u) \geq 0$. In view of the definition of L''_+ it suffices to show that there exists a subsequence x_{k_n} of x_k such that $L(x_{k_n},\overline{w}) - L(\overline{x},\overline{w}) \geq 0$ for all n. To prove this we suppose on the contrary that, for all large k, one has $L(x_k,\overline{w}) - L(\overline{x},\overline{w}) < 0$, that is,

$$\overline{w}_0[f(x_k) - f(\overline{x})] + \sum_{i=1}^{m+p} \overline{w}_i g_i(x_k) < 0$$

because $L(\overline{x},\overline{w}) = \overline{w}_0 f(\overline{x})$ as $\overline{w} \in T_1(\overline{x})$. Writing g_0 for the function $f(\cdot) - f(\overline{x})$, we see that there exist some index, say I $(0 \le I \le m + p)$, such that

(3.3.7) $$\overline{w}_I[g_I(x_k)] < 0$$

for infinitely many k. By considering a subsequence if necessary, we can assume that this is so for all k. Since $w^k \longrightarrow \overline{w}$, we can further assume that

(3.3.8) $$w_I^k[g_I(x_k)] < 0$$

for all k.

We next claim that there exists some index J with $J \ne I$ such that $g_J(x_k) \ge 0$ for infinitely many k. In fact, if not then there exists K such that

(3.3.9) $$g_i(x_k) < 0, \text{ for all } i \in \{0, 1, 2, \cdots, m + p\} \setminus \{I\}$$

and all $k \ge K$. But, as $w \in T_1(\overline{x})$ and

(3.3.9)′ $$0 \le G(x_k,f) = \sum_{i=0}^{m+p} w_i^k g_i(x_k)$$

(by a remark at the beginning of our proof and property (a) of Theorem 3.3.5 (i)), it follows that $p \ne 0$ and so

(3.3.10) $$\begin{cases} w_I^k = 0 \text{ for } i = 0, 1, \cdots, m \text{ with } i \ne I \\ w_I^k \le 0 \text{ for } i = m+1, \cdots, m+p \text{ with } i \ne I \end{cases}$$

by virtue of the maximality of $G(x_k,f)$ in its definition. If $p = 1$ and $I = m+1$ that (3.3.8), (3.3.9), (3.3.10) are clearly not consistent and therefore we can suppose that $p > 1$ and some $i \ne I$ with $m+1 \le i \le m+p$.

We define a new multiplier $w = (w_0, \cdots, w_{w+p}) \in T$ by

$$w_I = 0,\ w_i = -[(w_I^k)^2 + (w_i^k)^2]^{1/2}$$

and all other coodinations coincide with that of w^k. Then it follows from (3.3.8) and (3.3.9) that

$$w_i[g_i(x_k)] > w_i^k[g_i(x_k)] + w_I^k[g_I(x_k)]$$

and so

$$\sum_{i=0}^{m+p} w_i g_i(x_k) > \sum_{i=0}^{m+p} w_i^k g_i(x_k) = G(x_k, f)$$

contradicting the maximality of $G(x_k,f)$ in its definition. Therefore our claim must stand and we then take some x_k satisfying (3.3.8) and $J \neq I$ such that $g_J(x_k) \geq 0$. One defines a new multiplier $w \in T$ with $w_I = 0$, and

$$w_J = [(w_I^k)^2 + (w_J^k)^2]^{1/2}$$

and a contradiction is obtained same as above. □

If we consider the problem $\mathbb{P}$ in n-dimensional real Euclidean $X = \mathbb{R}^n$, that is, the functions f, g, i = 1, 2, $\cdots$, m + p are defined on $X = \mathbb{R}^n$, then we have the following strengtheneed result:

Theorem 3.3.7. (Second-order necessary condition with constraints in $\mathbb{R}^n$) Suppose that $\bar{x}$ is a local minimum solution of problem $\mathbb{P}(\mathbb{R}^n)$ and that u is a unit vector with $D_+G(\cdot,f)(\bar{x};u) = 0$. Then $G_-''(\bar{x},0,u) \geq 0$, $M_u(\bar{x}) \supseteq M_-(\bar{x})$ and there exists a Lagrange multiplier $\bar{w}$ in $M_u(\bar{x})$ such that $L_+''(\bar{x},\bar{w},0,u) \geq 0$.

Proof. Again we can assume that $g_j(\bar{x}) = 0$ for some j. Now the theorem follows from Theorem 3.3.6 together with Theorem 3.5 (ii). □

In view of Lemma 2.2.3, the preceding theorem 3.3.6 certainly implies the following

Theorem 3.3.8. Suppose that $\overline{x}$ provides a local solution to problem $\mathbb{P}(\mathbb{R}^n)$ and let $u \in D^*(\overline{x},G)$. Then there exists a Lagrange multiplier $\overline{w}$ in $M(\overline{x})$ such that (i) $0 \in \partial_u L(\overline{x},\overline{w})$ and (ii) $L_+''(\overline{x},\overline{w},0,u) \geq 0$.

Remark. This theorem was proved by Chaney [5, Theorem 2] under additional assumption that f is semismooth.

4. Sufficient Conditions Theorem with Constraints.

We begin with the following result which was proved by Ioffe's Proposition in the special case of unconstrained problems.

Lemma 3.4.1. Suppose that f is a locally Lipschitz function of $\mathbb{R}^n$ into $\mathbb{R}$. Suppose that $D_+f(\overline{x};u) \geq 0$ for any $u \in K_s(\overline{x})$. Then $\overline{x}$ provides a strictly local minimum for the following problems: for any $\varepsilon > 0$,

$$\mathbb{P}(\varepsilon) \quad \begin{cases} \text{minimize} & F_\varepsilon(x) := f(x) + \varepsilon\|x-\overline{x}\| \\ \text{subject to} & g_i(x) \leq 0 \quad \text{for } i = 1, 2, \cdots, m, \\ & g_i(x) = 0 \quad \text{for } i = m+1, \cdots, m+p. \end{cases}$$

Proof. Suppose that the desired conclusion is false. Then there exists a feasible sequence z_k convergent to $\overline{x}$ (say in direction u without loss of generality) such that for all k, one has

$$f(z_k) - f(x) \le -\varepsilon\|z_k-\bar{x}\|.$$

Clearly $u \in K_S(\bar{x})$. Then with $t_k = \|z_k-\bar{x}\|$, one has

$$D_+f(\bar{x};u) \le \liminf_{k\to\infty} \frac{1}{t_k}\{f(z_k) - f(\bar{x})\} \le -\varepsilon,$$

contradicting a given assumption. □

We next prepare a technical result for the proof of our main results on second-order sufficient conditions theorems with constraints.

Lemma 3.4.2. Let $\bar{x} \in S \subseteq \mathbb{R}^n$ such that $D_+f(\bar{x};v) \ge 0$ for all $v \in K_S(\bar{x})$. Then

(i) $D_+G(\cdot,f)(\bar{x};v) \ge 0$ for all $v \in \mathbb{R}^n$;

(ii) For each unit vector u with $D_+G(\cdot,f)(\bar{x};u) = 0$, one has that $0 \in \partial_u G(\cdot,f)(\bar{x})$, $M_u(\bar{x}) \supseteq M_-(\bar{x})$ and $M_u(\bar{x})$ is nonempty.

Remark. If there are no equality constraints (i.e.,p = 0) and $g_i(\bar{x}) < 0$ for all $i = 1, \cdots, m$, then $K_S(\bar{x}) = \mathbb{R}^n$. It then follows from assumption that $D_+f(\bar{x};v) \ge 0$ for all $v \in \mathbb{R}^n$, and hence that

$$f(\cdot) - f(\bar{x}) > g_i(\cdot)$$

for all i and x near $\bar{x}$. Consequently, by definition of G,

$$G(x,f) = f(x) - f(\bar{x}).$$

Thus (i) holds by assumption. Note also that $T_1(\bar{x})$ is the singleton consisting of $\bar{w} = (1, 0, \cdots, 0)$, $L(x,\bar{w}) = f(x)$ and therefore $M_-(\bar{x}) = \{\bar{w}\}$. By (i) and Theorem 3.2.6, the hypothesis

$$D_+G(\cdot,f)(\bar{x};u) = 0$$

implies that $0 \in \partial_u G(\cdot,f)(\bar{x}) = \partial_u L(\cdot,\bar{w})(\bar{x}) = \partial_u f(\bar{x})$ and hence that $M_u(\bar{x}) \supseteq M_-(\bar{x})$ (so equal to the singleton $\{\bar{w}\}$).

Proof of Lemma 3.4.2. In view of the preceding remark, we can suppose that $g_j(\bar{x}) = 0$ for some j with $1 \le j \le m + p$. Let $\varepsilon > 0$. By Lemmas 3.4.1, 3.3.4 (i), $\bar{x}$ provides a strict local minimum for $G(\cdot,F_\varepsilon)$ without constraints so $D_+G(\cdot,F_\varepsilon)(\bar{x};v) \ge 0$ for all v. By definitions, it is easy to verify that $G(\bar{x},f) = 0 = G(\bar{x},F_\varepsilon)$,

$$G(\bar{x}+tv,F_\varepsilon) \le G(\bar{x}+tv,f) + \varepsilon t\|v\|$$

and

$$D_+G(\cdot,F_\varepsilon)(\bar{x};v) \le D_+G(\cdot,f)(\bar{x};v) + \varepsilon\|v\|$$

for all $t > 0$ and $v \in \mathbb{R}^n$. Therefore, it follows that

$$0 \le D_+G(\cdot,f)(\bar{x};v) + \varepsilon\|v\|$$

and (i) is proved as ε is arbitrary.

(ii) follows from (i), Theorem 3.2.6 and parts (i), (ii) of Theorem 3.3.5. □

Our proof of theorem 3.4.4 later will be based on the following generalization of theorem 3.2.8.

Theorem 3.4.3 (Second-order sufficient conditions with constraints I). Let $\bar{x} \in S \subseteq \mathbb{R}^n$ such that $D_+f(\bar{x};u) \ge 0$ for all $u \in K_S(\bar{x})$. Suppose that for every unit vector u with $D_+G(\cdot,f)(\bar{x};u) = 0$, we have $G''_-(\bar{x},0,u) > 0$. Then there exists $\delta > 0$ such that $f(x) > f(\bar{x})$ for all $x \in B[\bar{x},\delta] \cap S$ and $x \ne \bar{x}$.

Proof. By assumptions, Lemma 3.4.2 (i) and Theorem 3.2.8, $\bar{x}$

provides a strictly local minimum for $G(\cdot,f)$ and hence for the problem $\mathbb{P}(\mathbb{R}^n)$ by part (i) of Lemma 3.3.4. □

We now state the main result of this section.

Theorem 3.4.4 (Second-order sufficient conditions with constraints IIa). Let $\bar{x} \in S \subseteq \mathbb{R}^n$ such that $D_+f(\bar{x};v) \geq 0$ for all $v \in K_S(\bar{x})$. Suppose that for every unit vector u with properties $D_+G(\cdot,f)(\bar{x};u) = 0$, $M_u(\bar{x})$ nonempty and $M_u(\bar{x}) \supseteq M_-(\bar{x})$, we have either (1) $G''_-(\bar{x},0,u) > 0$ or (2) $L''_-(\bar{x},w,0,u) > 0$ for all w in $M_u(\bar{x})$. Then there exists $\delta > 0$ such that $f(x) > f(\bar{x})$ for all $x \in B[\bar{x},\delta] \cap S$ and $x \neq \bar{x}$.

By part (ii) of Lemma 3.4.2 the preceding theorem is equivalent to the following formally weaker one:

Theorem 3.4.5 (Second-order sufficient conditions with constraints IIb). Let $\bar{x} \in S \subseteq \mathbb{R}^n$ and suppose that $D_+f(\bar{x};v) \geq 0$ for all $v \in K_S(\bar{x})$. Suppose also that for every unit vector u with

$$D_+G(\cdot,f)(\bar{x};u) = 0,$$

we have either (1) $G''_-(\bar{x},0,u) > 0$ or (2) $L''_-(\bar{x},w,0,u) > 0$ for all w in $M_u(\bar{x})$. Then there exists $\delta > 0$ such that $f(x) > f(\bar{x})$ for all $x \in B[\bar{x},\delta] \cap S$ and $x \neq \bar{x}$.

Proof of Theorem 3.4.4. In view of Lemma 3.4.2, it suffices to prove theorem 3.4.5 which in turn follows immediately from Theorem 3.4.3 and the following Lemma (the inclusion of part (i) is to facilitate the comparison of our Theorem 3.4.4 with a result of Chaney).

Suppose that u is a unit vector for which $M_u(\bar{x})$ is nonempty. Following [6], the modified lower second-order directional derivative $G''_{-m}(\bar{x},0,u)$ of G at $\bar{x}$, 0 in the direction u is defined to be the infimum of all numbers

$$\liminf\ [G(x_k,f) - G(\bar{x},f)]/\|x_k-\bar{x}\|^2,$$

taken over the set of all sequences $\{x_k\}$ such that both

(a') $\{x_k\}$ converges to x in the direction u and

(b') There exist w in $M_u(\bar{x})$ and a sequence x_k^* converging to 0 such that $x_k^* \in \partial L(x_k,w)$ for all k.

Lemma 3.4.6. Let u be a unit vevtor, and $0 \in \partial_u G(\cdot,f)(\bar{x})$ (e.g., by Lemma 3.4.2, this condition will be satisfied if $D_+G(\cdot,f)(\bar{x};u) = 0$ and $D_+f(\bar{x};v) \geq 0$ for all $v \in K_S(\bar{x})$). Then

(i) $G''_-(\bar{x},0,u) \geq G''_{-m}(\bar{x},0,u)$

and

(ii) $G''_-(\bar{x},0,u) \geq L''_-(\bar{x},\bar{w},0,u)$

for some $\bar{w} \in M_u(\bar{x})$.

Proof. We suppose that $G''_-(\bar{x},0,u)$ is finite and let $\lambda > G''_-(\bar{x},0,u)$. By the Remark following definition 3.1.3, there exist a sequence $\{x_k\}$ convergent to $\bar{x}$ in the direction u, and a sequence $\{x_k^*\}$ convergent to zero with each $x_k^* \in \partial G(\cdot,f)(x_k)$, such that

(3.4.1) $$\lambda > \liminf_{k\to\infty} \{[G(x_k,f) - G(\bar{x},f)]/\|x_k-\bar{x}\|^2\}.$$

By Theorem 3.3.5 (i) there exist $\overline{w} \in T_1(\overline{x})$ with $0 \in \partial_u L(\overline{x},\overline{w})$ (i.e. $\overline{w} \in M_u(\overline{x})$) and a sequence $\{y_k^*\}$ convergent to zero such that each $y_k^* \in \partial L(\cdot,\overline{w})(x_k)$. Then, by definition of G''_{-m}, right member of (3.4.1) is greater than or equal to $G''_{-m}(\overline{x},0,u)$. Therefore (i) follows. Moreover, as in the proof for Theorem 3.3.5 (ii), one has

$$G(x_k,f) - G(\overline{x},f) \geq L(x_k,\overline{w}) - L(\overline{x},\overline{w}).$$

Dividing by $\|x_k-\overline{x}\|^2$ and taking lower limits, we see that the right member of (3.4.1) is greater then or equal to $L''_-(\overline{x},\overline{w},0,u)$, and so (ii) follows. □

To compare the Theorem 3.4.4 with theorem 5 of [6], we define $D(\overline{x})$ in $\mathbb{R}^n$ by

$$D(\overline{x}) := \{u \in \mathbb{R}^n : D_+f(\overline{x};u) \leq 0,\ D_+g_i(\overline{x};u) \leq 0,\ \forall i \in I(\overline{x}),$$
$$\text{and } D_+g_i(\overline{x};u) \leq 0 \leq D^+g_i(\overline{x};u) \text{ for } i = m+1, \cdots, m+p.\}.$$

If for each $i = m+1, \cdots, m+p$, the function g_i has the directional derivative at $\overline{x}$, that is the limit

$$g'_i(\overline{x};u) := \lim_{t \downarrow 0} \frac{1}{t}\{g_i(\overline{x}+tu) - g_i(\overline{x})\}$$

exists for all $u \in \mathbb{R}^n$, then the set $D(\overline{x})$ becomes

$$\{u \in \mathbb{R}^n : D_+f(\overline{x};u) \leq 0,\ D_+g_i(\overline{x};u) \leq 0,\ \forall i \in I(\overline{x}),$$
$$g'_i(\overline{x};u) = 0 \text{ for } i = m+1, \cdots, m+p.\},$$

the one considered by Chaney [6].

Lemma 3.4.7. If $u \in \mathbb{R}^n$ with $D_+G(\cdot,g_0)(\overline{x};u) = 0$, then $u \in D(\overline{x})$.

Proof. As before we can suppose that $g_j(\overline{x}) = 0$ for some j with $1 \leq j \leq m+p$. Recalling from Lemma 3.3.1 that

$$G(x,f) = \max \{ \sum_{i=0}^{m+p} w_i g_i(x) : w \in T,\ w_i g_i(\bar{x}) = 0,\ \forall i = 1, 2, \cdots, m+p\}$$

for all x near $\bar{x}$. Considering all w of the form $(w_0, w_1, \cdots, w_{m+p})$ whose coordinates all zero except at a coordinate i and at this i,

$$w_i = \begin{cases} 1 & \text{if } i = 0 \text{ or } i \in I(\bar{x}); \\ \pm 1 & \text{if } m + 1 \le i \le m + p, \end{cases}$$

it follows that $G(x,f) \ge f(x) - f(\bar{x}),\ g_i(x),\ \pm g_k(x)$ for all $i \in I(\bar{x})$ and $k = m+1, \cdots, m+p$. Since $G(\bar{x},f) = 0 = g_i(\bar{x}) = g_k(\bar{x})$, it follows from the definitions that

$$D_+G(\cdot,f)(\bar{x};u) \ge D_+f(x;u),\ D_+g_i(\bar{x};u),\ D_+(\pm g_k)(\bar{x};u).$$

Since $D_+(-g_k)(\bar{x};u) = -D^+g_k(\bar{x};u)$, the required result follows. □

In view of this lemma and since each u in $K_S(\bar{x})$ satisfies the tangential constraints as noted at the beginning of section 3, Theorem 3.4.4 clearly strengthens Chaney's theorem 5 in [6] even under his additional semismoothness assumption as he replaces (1) of Theorem 3.4.4 by a stronger and somewhat unnatural condition:

(1^*) $G''_-(\bar{x},0,u) > 0$.

Likewise, his Theorem 4 can also be strengthened by the following

Theorem 3.4.8 (Second-order sufficient conditions with constraints III). Let f, g_i, $i = 1, 2, \cdots, m+p$, be locally Lipschitz functions at $\bar{x} \in S$ and $M_-(\bar{x})$ nonempty. If for any $u \in D(\bar{x}) \cap K_S(\bar{x})$ with $M_u(\bar{x}) \supseteq M_-(\bar{x})$ we have $L''_-(\bar{x},w,0,u) > 0$ for some $w \in M_-(\bar{x})$, then there exists $\delta > 0$ such that $f(x) > f(\bar{x})$ for every $x \in B(\bar{x},\delta) \cap S$ with $x \ne \bar{x}$.

Proof. Suppose that the desired conclusion is false. Then there exists a feasible sequence $\{x_k\}$ convergent to $\bar{x}$ in the direction u such that $f(x_k) \le f(\bar{x})$. Then $u \in D(\bar{x}) \cap K_S(\bar{x})$. Now we claim $M_-(\bar{x}) \subseteq M_u(\bar{x})$. Let $w \in M_-(\bar{x})$. Then $D_+L(\cdot,w)(\bar{x},v) \ge 0$ for all $v \in \mathbb{R}^n$. But

$$D_+L(\cdot,w)(\bar{x},u) \le \liminf_{k\to\infty} \frac{1}{\|x_k-\bar{x}\|}\{L(x_k,w) - L(\bar{x},w)]\} \le 0$$

since $L(x_k,w) \le w_0f(x_k) \le w_0f(\bar{x}) = L(\bar{x},w)$. Thus, $D_+L(\cdot,w)(\bar{x};u) = 0$. By Theorem 3.2.6 we have $0 \in \partial_uL(\cdot,w)(\bar{x})$ and so $w \in M_u(\bar{x})$. This implies that $M_-(\bar{x}) \subseteq M_u(\bar{x})$. Then by assumption there exists $\bar{w} \in M_-(\bar{x})$ such that $L_-''(\bar{x},\bar{w},0,u) > 0$. On the other hand, for any $w \in M_-(\bar{x})$, since $D_+L(\cdot,w)(\bar{x};v) \ge 0$ for all $v \in \mathbb{R}^n$, it follows from Ioffe's Proposition that for any $\varepsilon_i \downarrow 0$, there exists $\delta_i \downarrow 0$ such that

$$F_{\varepsilon_i}(x) := L(x,w) + \varepsilon_i\|x-\bar{x}\|$$

attains a minimum $\bar{x}$ on $B[\bar{x},\delta_i]$. Now for each i, we can find a x_{k_i} such that $x_{k_i} \in B[\bar{x},\frac{1}{4}\delta_i]$ and

$$L(x_{k_i},w) \le L(\bar{x},w) \le L(x,w) + \varepsilon_i\|x-\bar{x}\|$$

for all $x \in B[\bar{x},\delta_i]$. Hence

$$L(x_{k_i},w) \le L(x,w) + 2\varepsilon_i\|x_{k_i}-\bar{x}\|$$

for all $x \in B[\bar{x},2\|x_{k_i}-\bar{x}\|]$. By Ekeland variation principle with

$$\lambda = \varepsilon_i^{1/2}\|x_{k_i}-\bar{x}\|/2,$$

we have $z_{k_i} \in B[\bar{x},2\|x_{k_i}-\bar{x}\|]$ such that

(i) $\|z_{k_i}-x_{k_i}\| \le \varepsilon_i^{1/2}\|x_{k_i}-\bar{x}\|/2$;

(ii) $L(z_{k_i},w) \le L(x_{k_i},w)$;

and

(iii) $L(z_{k_i},w) \le L(x,w) + 4\varepsilon_i^{1/2}\|x-z_{k_i}\|$ for all $x \in B[\bar{x},3\|x_{k_i}-\bar{x}\|]$.

From (i) we we have $\bar{x} \ne z_k$ and $z_{k_i} \in B(\bar{x},3\|x_{k_i}-\bar{x}\|)$. Thus, from (iii) w obtain $0 \in \partial L(\cdot,w)(z_{k_i}) + 4\varepsilon_i^{1/2}B_i^*$. Thus, there exists

$$z_{k_i}^* \in \partial L(\cdot w)(z_{k_i}) \text{ with } \|z_{k_i}^*\| \le 4\varepsilon_i^{1/2}$$

and so $z_{k_i}^* \to 0 \in \partial L(\cdot,w)(\bar{x})$. Note that by (i)

$$(z_{k_i}-x_{k_i})/\|x_{k_i}-\bar{x}\| \to 0.$$

Then

$$(z_{k_i} - \bar{x})/\|z_{k_i}-\bar{x}\| = \frac{[(z_{k_i}-x_{k_i})/\|x_{k_i}-\bar{x}\|] + [(x_{k_i}-\bar{x})/\|x_{k_i}-\bar{x}\|]}{\|[(z_{k_i}-x_{k_i})/\|x_{k_i}-\bar{x}\|] + [(x_{k_i}-\bar{x})/\|x_{k_i}-\bar{x}\|]\|} \to u.$$

Hence it follows from the definition of L_-'', (ii) and $L(x_{k_i},w) \le L(\bar{x},w)$ that

$$L_-''(\bar{x},w,0,u) \le \liminf_{i\to\infty} \frac{1}{\|z_{k_i}-\bar{x}\|^2}\{L(z_{k_i},w) - L(\bar{x},w)\} \le 0.$$

Since w is arbitrary in $M\_(\bar{x})$, this is a contradiction. □

The following result was proved by Chaney [6, Theorem 4] under additional assumption that f, g_i, i = 1, 2, ⋯, m+p, are semismooth at $\bar{x}$.

Corollary 3.4.9. Let f, g_i, i = 1, 2, ⋯, m+p, be locally Lipschitz functions at $\bar{x} \in S$ and $M(\bar{x})$ nonempty. Suppose that the function $L(\cdot,w)$ is regular at $\bar{x}$ for every w in $M(\bar{x})$. If for any $u \in D(\bar{x}) \cap K_S(\bar{x})$ with $M_u(\bar{x}) = M(\bar{x})$ we have $L_-''(\bar{x},w,0,u) > 0$ for some $w \in M(\bar{x})$, then there exists $\delta > 0$ such that $f(x) > f(\bar{x})$ for every

$x \in B(\bar{x},\delta) \cap S$ with $x \neq \bar{x}$.

Proof. In the case of regularity, we have $M(\bar{x}) = M_{-}(\bar{x})$. Then,

$$M_{-}(\bar{x}) \geq M_{u}(\bar{x})$$

for any $u \in \mathbb{R}^n$. Hence, if $u \in D(\bar{x}) \cap K_S(\bar{x})$ with $M_u(\bar{x}) \geq M_{-}(\bar{x})$, then

$$M_u(\bar{x}) = M_{-}(\bar{x}) \ (= M(\bar{x}))$$

and so by the assumption we have $D_{+}L(\bar{x},w,0,u) > 0$ for some $w \in M_{-}(\bar{x})$ $(= M(\bar{x}))$. By Theorem 3.4.8 we complete the proof. □

References

[1] J-P Aubin and I. Ekeland, Applied Nonlinear Analysis, John Wiley & Sons, 1984.

[2] A. Ben-Tal, Second-order and related extremality conditions in nonlinear programming, J. Optim. Theory Appl, 31(1980), 143-165.

[3] A. Ben-Tal and Zowe, A unified theory of first and second order conditions for extremum problem in topological vector spaces, Math. Programming Stud. 19(1982), 39-76.

[4] A. Ben-Tal and Zowe, Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems. Math. Programming 24 (1982), 70-91.

[5] R. W. Chaney, Second-order necessary conditions in constrained semismooth optimization, SIAM J. Control and Optim. Vol. 25, No. 4(1987), 1072-1081.

[6] R. W. Chaney, Second-order sufficient conditions in nonsmooth optimization, Math. Oper. Res. Vol. 13. No. 4(1988), 660-673.

[7] R. W. Chaney, Second-order necessary conditions in semismooth optimization, Math. Programming 40(1988), 95-109.

[8] R. Cominetti and R. Correa, A generalized second-order derivative in nonsmooth optimization, SIAM J. Control and Optim. Vol. 28, No. 4(1990), 789-809.

[9] F. H. Clarke, Optimization and Nonsmooth Analysis. Wiley-Interscience, New York, 1983.

[10] A. D. Ioffe, Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps, Nonlinear Analysis,

Theory, Methods & Applications, Vol. 8. No. 5(1984), 517-539.

[11] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, SIAM J. Control and Optim. Vol. 15, No. 6 (1977), 959-972.

[12] R. T. Rockafellar, Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives, Math. Oper. Res. vol.14. No. 3(1989), 462-484.

[13] M. Studniarski, Second-order necessary conditions for optimality in nonsmooth nonlinear programming, J. Math. Anal. Appl., Vol. 154, No. 2(1991), 303-317.

Appendix

For convenience of reference, we review here some known results with proofs that have been used by us in the preceding chapters.

Let X be a locally convex space and $f^{\infty}(x;u,v)$ be as in chapter 2.

Lemma A1. [2]. Let $f: X \longrightarrow \mathbb{R}$ and $x \in X$. Then:

(i) The map $(u,v) \longmapsto f^{\infty}(x;u,v)$ is symmetric, and sublinear on each variable separately.

(ii) The map $y \longmapsto f^{\infty}(y;u,v)$ is upper semi-continuous at x for every $(u,v) \in X \times X$.

(iii) $f^{\infty}(x;u,-v) = f^{\infty}(x;-u,v) = (-f)^{\infty}(x;u,v) = -f_{\infty}(x;u,v)$.

Proof: By definition

$$f^{\infty}(x;u,v) := \limsup_{\substack{y\to x \\ t,s\downarrow 0}} \square_f(y,t,s,u,v),$$

where

$$\square_f(y,t,s,u,v) = \frac{1}{st}\Big\{f(y+tu+sv) - f(y+tu) - f(y+sv) + f(y)\Big\}.$$

Then for any $\alpha > 0$ and $z, w \in X$,

$$(4.1) \qquad \square_f(y,t,s,u,\alpha v) = \alpha\square_f(y,t,\alpha s,u,v)$$

$$= \alpha\square_f(y,t,\lambda,u,v), \qquad (\lambda = \alpha s)$$

$$(4.2) \qquad \square_f(y,t,s,u,z+w)$$

$$= \frac{1}{st}\Big\{f(y+tu+s(z+w)) - f(y+tu) - f(y+s(z+w)) + f(y)\Big\}$$

$$= \frac{1}{st}\Big\{f((y+sz)+tu+sw)-f((y+sz)+tu)-f((y+sz)+sw)+f(y+sz)\Big\}$$

$$+ \frac{1}{st}\Big\{f(y+tu+sz) - f(y+tu) - f(y+sz) + f(y)\Big\}$$

$$= \square_f(y+sw,t,s,u,z) + \square_f(y,t,s,u,w)$$

and

(4.3) $$\square_f(y,t,s,u,-v)$$

$$= \frac{1}{st}\left\{f(y+tu-sv) - f(y+tu) - f(y-sv) + f(y)\right\}$$

$$= \frac{1}{st}\left\{f(\xi) - f(\xi+sv) - f(\xi-tu) + f(\xi-tu+sv)\right\}$$

$$= \square_f(\xi,t,s,-u,v),$$

where $\xi = y+tu-sv$. Similarly, let $\zeta = y-sv$; then one has

(4.4) $$\square_f(y,t,s,u,-v)$$

$$= \frac{1}{st}\left\{f(\zeta+tu) - f(\zeta+tu+sv) - f(\zeta) + f(\zeta+sv)\right\}$$

$$= \square_{(-f)}(\zeta,t,s,u,v).$$

(i) The symmetry of $f^{\infty}(x;\cdot,\cdot)$ follows from the definition and the sublinearity of $f^{\infty}(x;u,\cdot)$ follows from taking upper limits on (4.1) and (4.2).

(iii) The last equality follows easily from definitions. The other equalities follow from (4.3) and (4.4).

(ii) The proof is similar to the proof of (2) of Theorem 1.3.1. □

Lemma A2. Let $f\colon X \longrightarrow \mathbb{R}$ be a continuous function, x, $v \in X$, and $t > 0$. Then there exists $\alpha \in (0,t)$ such that

$$\frac{f(x+tv) - f(x)}{t} \le D_+f(x+\alpha v;v).$$

Consequently,

$$\limsup_{y\to x} D_+f(y;v) = \limsup_{y\to x} D^+f(y;v) = \limsup_{y\to x} f^0(y;v) = f^0(x;v).$$

Remark: If let $f = -g$, then we have

$$\frac{g(x+tv) - g(x)}{t} \ge D^+g(x+\alpha v;v),$$

and so the corresponding results for $f_0(x;v)$.

Proof: By introducing

$$h(s) = f(x+sv) - \frac{s}{t}[f(x+tv) - f(x)],$$

we have

$$h(\alpha+\lambda) - h(\alpha) = f(x+\alpha v+\lambda v) - f(x+\alpha v) - \frac{\lambda}{t}[f(x+tv) - f(x)]$$

and so

$$\frac{f(x+tv) - f(x)}{t} \le D_+f(x+\alpha v;v)$$

is equivalent to

$$D_+h(\alpha;1) = \lim\inf_{\lambda\downarrow 0} \frac{1}{\lambda}\{h(\alpha+\lambda) - h(\alpha)\} \ge 0.$$

Since $h(0) = h(t)$, there exists $t_0 \in (0,t]$ being a global maximum point of h. Let $t_1 \in (0,t_0)$. Then we either have $D_+h(t_1;1) \ge 0$, in which case we are done by taking $\alpha = t_1$, or $D_+h(t_1;1) < 0$ in which case h attains a local minimum at some $\alpha \in (t_1,t_0)$ and clearly $D_+h(\alpha;1) \ge 0$. The proof of the first conclusion is completed.

Now for any $y \in X$ and $t > 0$, we apply our first assertion to obtain $\alpha \in (0,t)$ such that

$$\frac{f(y+tv) - f(y)}{t} \le D_+f(y+\alpha v;v) \le D^+f(y+\alpha v;v).$$

Taking upper limits in the above inequalities, we conclude that

$$f^0(x;v) \le \lim\sup D_+f(y;v) \le \lim\sup_{y\to x} D^+f(y;v).$$

On the other hand, the opposite inequalities must also hold because by $D_+f(y;v) \le D^+f(y;v) \le f^0(y;v)$ and the upper semicontinuity of $f^\infty(\cdot,v)$ (see Theorem 1.3.1. (2)), we have

$$\lim\sup_{y\to x} D_+f(y;v) \le \lim\sup_{y\to x} D^+f(y;v)$$
$$\le \lim\sup_{y\to x} f^0(y;v) \le f^0(x;v). \qquad \square$$

Let g, h and I(g(x)) be as in chapter 2 (pp. 27-28).

Lemma A3. [2, Prop. 3.7 and 3.8] Suppose that each g_i is a C^2-function at x, $1 \le i \le n$, and the derivatives

$$\{g_i'(x);\ i \in I(g(x))\}$$

are affinely independent. Then for any $u, v \in X$,

$$h^{\infty}(x;u,v) \ge \max_{i \in I(g(x))} D^2 g(x;u,v).$$

Proof: Let $j \in I(g(x))$. Since $\{g_i'(x);\ i \in I(g(x))\}$ are affinely independent, the following system

$$\{g_j'(x) - g_k'(x);\ k \in I(g(x)),\ k \ne j\}$$

is linearly independent. Then there exists $w \in X$ such that

$$g_j'(x;w) - g_k'(x;w) = 1$$

for all $k \in I(g(x))$ and $k \ne j$. Consequently, there exists $M > 0$ such that

$$g_j(x+tw) > g_i(x+tw)$$

for all $i \ne j$, $1 \le i \le n$ and all $t \le M$ since $j \in I(g(x))$. By continuity, it follows that the equality

$$g_j(\cdot) > g_k(\cdot) \qquad (k = 1, 2, \cdots, n;\ k \ne j)$$

on an open neighbourhood U_t of $x + tw$. Consequently, by definition of h, one has

$$g_j = h \text{ on } U_t$$

and so $D^2 g_j(x+tw;u,v) = h^{\infty}(x;u,v)$. By passing to the limits as $t \downarrow 0$ it follows from Proposition 2.2.4 that

$$D^2 g_j(x;u,v) \le h^{\infty}(x;u,v). \qquad \square$$

Lemma A4. [2. Prop. 3.9] Let $u, v \in X$. If each g_i is a C^2-function at x, $1 \le i \le n$, $\{g_k'(x);\ k \in I(g(x))\}$ are affinely independent and there are $i, j \in I(g(x))$ such that

$$[g_i'(x;v) - g_j'(x;v)][g_i'(x;u) - g_j'(x;u)] > 0,$$

then $h^{\infty}(x;u,v) = +\infty$.

Proof: Since $h^{\infty}(x;u,v) = h^{\infty}(x;-u,-v)$ by (iii) of Lemma A1, we may assume (by changing u to -u and v to -v if necessary) that

$$g_i'(x;v) > g_j'(x;v) \quad \text{and} \quad g_i'(x;u) > g_j'(x;u).$$

Since $\{g_k'(x),\ k \in I(g(x))\}$ are affinely independent,

$$\{g_k'(x) - g_j'(x) \colon k \neq j \text{ and } k \in I(g(x))\}$$

are linearly independent. Thus, there exists $w \in X$ such that

$$(4.5) \qquad g_i'(x;w) - g_j'(x;w) = 0$$

and

$$(4.6) \qquad g_k'(x;w) - g_j'(x;w) = -1,\ k \neq i,\ j \text{ and } k \in I(g(x)).$$

It follows from a standard implicit function theorem [3, Theorem 5.2] that we may find a path $x(t) \in X$ and $\delta > 0$ such that $x(0) = x$, $x'(0) = w$ and

$$g_l(x(t)) = g_l(x) + tg_l'(x;w),\ l \in I(g(x))$$

for all $|t| < \delta$. Thus, by (4.5) and (4.6) we obtain

$$(4.7) \qquad g_i(x(t)) = g_j(x(t)) > g_k(x(t))$$

for every $k \in I(g(x))$, $k \neq i$, j and $|t| \le \delta$. By taking smaller δ if necessary we can further assume the above inequality in (4.7) holds even if $k \notin I(g(x))$ (because g_j, g_k are continuous and $g_i(x(t)) > g_k(x(t))$ at $t = 0$). Consequently,

$$I(g(x(t))) = \{i,\ j\}.$$

Since $g_i'(x;u) > g_j'(x;u)$, by continuity of $g_i'(\cdot;u)$ and $g_j'(\cdot;u)$ the same holds with x replaced by x(t) provided $t > 0$ is small. Then it follows from Lemma 2.3.1 and the definition of

$$(g_i'(x(t)) - g_j'(x(t)))u$$

we may find $\varepsilon_t > 0$ such that

$$I(g(x(t)+\tfrac{s}{2}u)) = \{i\} \quad \text{and} \quad I(g(x(t)-\tfrac{s}{2}u)) = \{j\}$$

for all $s \in (0,\varepsilon_t)$ and t near zero.

Take $s_t \in (0,\varepsilon_t)$ with $s_t \downarrow 0$ and set

$$y(t) = x(t) - \frac{s_t}{2} u.$$

Then

$$\lim_{t\downarrow 0} h'(y(t)+s_t u;v) = \lim_{t\downarrow 0} g_i'(y(t)+s_t u;v) = g_i'(x;v),$$

$$\lim_{t\downarrow 0} h'(y(t);v) = \lim_{t\downarrow 0} g_j'(y(t);v) = g_j'(x;v),$$

and since $g_i'(x;v) > g_j'(x;v)$ we conclude from Proposition 2.2.4 that

$$h^{\infty}(x;u,v) \ge \limsup_{t\downarrow 0} \frac{h'(y(t)+s_t u;v) - h'(y(t);v)}{s_t} = +\infty,$$

because the numerator converges to the finite limit

$$g_i'(x;v) - g_j'(x;v) > 0. \qquad \square$$

Let the problem $\mathbb{P}(X)$ and S, T be defined as in chapter 3 (p.64) and let $\bar{x} \in S$. Also define the function $G(x,f)$ as in pages 64-65.

Lemma A5. [1, Lemma 1] (i) $\bar{x}$ is a strictly local solution to problem $\mathbb{P}(X)$ if and only if we have $G(x,f) > 0 = G(\bar{x},f)$ for all x near $\bar{x}$ with $x \neq \bar{x}$.

(ii) If $\bar{x}$ is a local solution to problem $\mathbb{P}(X)$, then $G(x,f) \ge 0 = G(\bar{x},f)$ for all x near $\bar{x}$.

Proof: (ii) Suppose that $\bar{x}$ is a local solution to the problem $\mathbb{P}(X)$. Then there exists a neighbourhood U of $\bar{x}$ such that

(a) for each feasible x in U, $f(x) \ge f(\bar{x})$;

(b) for each unfeasible point x in U, there exists i such that $g_i(x) > 0$ if $1 \le i \le m$ or $g_i(x) \neq 0$ if $m+1 \le i \le m+p$.

For the case (a), we pick $w = (1, 0, \cdots, 0)$ and note that

$$G(x,f) \geq L(x,w) - f(\bar{x}) \geq f(x) - f(\bar{x}) \geq 0$$

by definitions. For the case (b), we choose $w \in T$ with

$$w_j = 0 \text{ for all } j \neq i \text{ and } w_i = \begin{cases} 1 & \text{if } 1 \leq i \leq m \\ \operatorname{sgn}(g_i(x)) & \text{if } m+1 \leq i \leq m+p \end{cases}.$$

Then we also have

$$L(x,w) - w_0 f(\bar{x}) = w_i g_i(x) > 0$$

so that $G(x,f) > 0$. That $G(\bar{x},f) = 0$ follows from the definition of G and noting that $\bar{x}$ is a feasible point (so that $\sum_{i=1}^{m+p} w_i g_i(\bar{x}) \leq 0$ for all $w \in T$).

(i) The "only if" part follows similarly as (ii). Conversely, suppose there exists a neighbourhood U of $\bar{x}$ such that $G(x,f) > 0 = G(\bar{x},f)$ for all $x \in U \setminus \{\bar{x}\}$. Then for any feasible point x in $U \setminus \{\bar{x}\}$ and $w \in T$, one has

$$\sum_{i=1}^{m+p} w_i g_i(x) \leq 0$$

so that

$$0 < G(x,f) = \max\{w_0(f(x) - f(\bar{x})) + \sum_{i=1}^{m+p} w_i g_i(x);\ w \in T\}$$
$$\leq \max\{w_0(f(x) - f(\bar{x}));\ w \in T\}.$$

This implies that $f(x) > f(\bar{x})$ for all feasible points x in $U \setminus \{\bar{x}\}$.

□

References

[1] R. W. Chaney, Second-order necessary conditions in constrained semismooth optimization, SIAM J. Control Opt, 25 (1987), 1072-1081.

[2] R. Cominetti and R. Correa, A generalized second-order derivative in nonsmooth optimization, SIAM J. Control Optim., Vol.28, No. 4(1990), 789-809.

[3] M. R. Hestenes, Optimization Theory, John Wiley & Sons, 1975.