Minimizing and Stationary Sequences

by

WONG OI PING

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

Master of Philosophy

Hong Kong

July, 1999

ACKNOWLEDGMENTS

I would like to take this opportunity to express my deepest gratitude to my supervisor, Prof. Kung-fu Ng, for his continual guidance and valuable suggestions. His encouragement and supervision has helped me a lot throughout the period of my postgraduate studies and in the preparation of this thesis.

Abstract

Let f be a continuous differentiable function on a closed subset of $\mathbb{R}^n$. Following the works on [9], we first study the relation between LP-minimizing (Levitin-Polyak) and stationary sequences of the problem in minimizing f on X. Conditions for equivalence on these sequential properties are obtained. Moreover, we investigate [14] the relation between minimizing and stationary sequence for the problem on a Banach space, where the extended real-valued function f is lower-semicontinuous and bounded below. Finally, through discussing some functions G and L for constrained problem, sufficient and necessary conditions for nonsmooth optimization is introduced from [15].

簡介

設 f 是定義在 R^n 的一個閉子集中連續可微函數。循著文章 [9] 中的結果，我們首先研究求 f 的極小值問題上 LP 極小化序列及穩定序列。並且得到這兩個性質的等價條件。此外，我們更在 [14] 文章中考察兩種序列在 Banach 空間，延實價下半連續和有下界函數的關係。最後，通過討論在帶約束性問題上的兩個函數 G 及 L，由文章 [15] 引入非光滑最優化的充份及必要條件。

Introduction

This thesis surveys some recent theoretical results on optimization theory. Our study is divided into two parts. The first one emphasizes on the relation between minimizing sequences and stationary sequences. The second part is to study the necessary and sufficient conditions on nonsmooth optimization.

The study of minimizing sequences and stationary sequences of a convex programs was first investigated by Auslender, Crouzeix, Angleraud and their colleagues ([1] [4] [5] [18]). Initially, they wanted to determine the conditions for ensuring a stationary sequence of a given function are minimizing for an algorithm. But their work was mainly on convex functions. After that, Chou, Huang, Ng and Pang ([9] [14] [15] [16] [17]) extended the result to a lower semicontinuous function on a Banach space. Also, they dealt with the equivalence on minimizing and stationary sequences for constrained and unconstrained convex program.

Many results on necessary and sufficient optimality conditions for nonsmooth optimization problem had been presented under various kinds of conditions, for example, regularity, convexity and semismoothness ([6] [7] [8]). Huang and Ng then extended the result to a Banach space for a locally Lipschitzian function. They had investigated necessary and sufficient second order constrained and unconstrained programs.

Based on [9], the first chapter is to study the relation between Levitin-Polyak minimizing (LP-minimizing) and stationary sequences in residual function approach. As a consequence of Theorem (1.2.1), whether an LP-minimizing sequence $\{x_k\}$ is stationary depends on the behaviour of the function f and the gradient ∇f near $\{x_k\}$. For example, if f and ∇f are uniformly continuous near an LP-minimizing sequence $\{x_k\}$, then $\{x_k\}$ is N-stationary. Conversely, a stationary sequence becomes an LP-minimizing sequence if the problem is H-metrically regular, which is given in Theorem (1.3.1). Moreover, we investigate the equivalence relation for a special type problem CQQSP given in Theorem (1.4.1). In the last section of this chapter, we discuss a descent algorithm of a convex optimization program. Its convergence depends on the error bound condition in Theorem (1.7.1).

We extend a function f to be lower semicontinuous extended real-valued on a Banach space X. Based on [14], relation between minimizing and stationary sequences by subdifferential approach is introduced in chapter 2. First of all, we define a subdifferential as a subset of a dual space X^* in Definition (2.1.1) and a minimizing sequence is in terms of subdifferential. A minimizing sequence $\{x_k\}$ is stationary if the subdifferential ∂f is uniformly upper semicontinuous near $\{x_k\}$, which is illustrated in Theorem (2.2.2). On the other hand, if the function is C-convex and the Banach space is reliable, then a stationary sequence is minimizing for the level sets which can be characterised by the error bound condition. The result is given in Theorem (2.4.1). Finally, we deal with a problem that whether a function is critical. If the function f is C-convex, has at least one critical point and all its stationary seqeuences are bounded, then it is critical as in Proposition (2.5.2). In addition, for a reliable space X and a critical function f, minimizing is a consequence of a stationary sequence. This property for a critical function is illustrated in Theorem (2.5.1).

In the last chapter, based on [15] we turn to our attention to study necessary and sufficient conditions in an optimization problem, where the objective function is locally Lipschitz real-valued on a Banach space X. We divided the problem into two cases, unconstrained and constrained optimization. For the unconstrained problem, the discussion is mainly on the Dini-directional derivatives. In the case of constrained one, some functions G and L are introduced. If the lower Dini-directional derivatives of G in a direction u equals to zero, then the necessary condition is given in Theorem (3.4.1) that G''_- is nonnegative in such direction. On the other hand, $\bar{x}$ is a local minimizer if G''_- is strictly greater than zero at $\bar{x}$ in some direction v for which the lower Dini-directional derivative of G at $\bar{x}$ in the direction v is zero. This is the sufficient condition in Theorem (3.5.1) for a constrained problem.

Contents

Chapter 1

LP-minimizing and Stationary Sequences

For constrained finite-dimensional problems, we discuss [9] the relation between an LP-minimizing (Levitin-Polyak minimizing) sequence and stationary sequence. Analysis is mainly through residual function approach and the theory of error bounds. The advantage is that we can handle infeasible sequence and constraints explicitly in practical problems. Moreover, specializations of the results to convex quadratically constrained convex spline (differential convex piecewise quadratic function) minimization problems is introduced.

1.1 Residual function

We consider the following constrained finite-dimensional differential problem:

$$\begin{aligned} &\text{Minimize} \quad f(x) \\ &\text{subject to} \quad x \in X \end{aligned} \tag{1.1}$$

where $f:\mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function and X is a nonempty closed subset of $\mathbb{R}^n$.

It is well known that for each $x \in \mathbb{R}^n$ there exists $z \in X$ such that $\|z - x\| = dist(x, X)$, where $dist(y, X) = \inf_{v \in X} \|v - y\|$. If X is assumed to be convex, then there is exactly one vector z in X satisfying the above property and will henceforth be denoted by $\Pi_X(x)$. The following result is well-known; for completeness, we include a proof here.

Proposition 1.1.1 *Let X be a nonempty closed convex set in $\mathbb{R}^n \backslash X$. The following statements are equivalent:*

(i) $z = \Pi_X(x)$;

(ii) $z \in X$ *and* $\langle z - x, y - z \rangle \geq 0$, $\forall y \in X$.

Proof Let $z, y \in X$. Define $\varphi : [0, 1] \to \mathbb{R}$ by

$$\varphi(t) = \frac{1}{2}\|[z + t(y - z)] - x\|^2, \quad \forall t \in [0, 1].$$

Clearly, φ is convex and

$$\varphi'(t) = \langle z - x + t(y - z), y - z \rangle \quad \forall t \in [0, 1]$$

where $\varphi'(0) = \langle z - x, y - z \rangle$ to be interpreted as right-hand side derivative. Now, if (i) holds, then $\varphi(t) \geq \varphi(0)$ for each t and so $\varphi'(0) \geq 0$, that is (ii) holds. Conversely, suppose (ii) is true. Then, with any $y \in X$ and φ be defined as before, it follows from the convexity of φ that

$$\varphi(1) - \varphi(0) \geq \varphi'(0) = \langle z - x, y - z \rangle \geq 0,$$

so $\frac{1}{2}\|y - x\|^2 \geq \frac{1}{2}\|z - x\|^2$, and (i) is seen to hold. □

The inequality appeared in (ii) of the proposition is sometimes referred in the literature as the variational inequality. One of its application is:

Proposition 1.1.2 *Let X be a nonempty closed convex set in $\mathbb{R}^n$. Let $x \in \mathbb{R}^n \backslash X$. Then*

$$\|\Pi_X x_1 - \Pi_X x_2\| \leq \|x_1 - x_2\| \quad \forall x_1, x_2 \in \mathbb{R}^n.$$

Proof Write $\bar{x}_1, \bar{x}_2$ respectively for $\Pi_X x_1$ and $\Pi_X x_2$. Applying (ii) of the above Proposition to $x = x_1$ and $y = \bar{x}_2$, one has

$$\langle \bar{x}_1 - x_1, \bar{x}_2 - \bar{x}_1 \rangle \geq 0.$$

Similarly, we also have

$$\langle x_2 - \bar{x}_2, \bar{x}_2 - \bar{x}_1 \rangle \geq 0.$$

Summing up, we have

$$\langle x_2 - x_1, \bar{x}_2 - \bar{x}_1 \rangle \geq \langle \bar{x}_2 - \bar{x}_1, \bar{x}_2 - \bar{x}_1 \rangle.$$

By the Schwarz inequality, it follows that

$$\|x_2 - x_1\| \geq \|\bar{x}_2 - \bar{x}_1\|.$$

□

As we do not assume that f attained minimum (unless specified), let

$$f_{inf} = \inf_{x \in X} f(x) \geq -\infty.$$

Definition 1.1.1 *A sequence $\{x_k\} \subseteq \mathbb{R}^n$ is said to be a Levitin-Polyak minimizing (LP minimizing) sequence for (1.1) if the following conditions are satisfied:*

(i) Asymptotically Optimal

$$\lim_{k \to \infty} f(x_k) = f_{inf}$$

and

(ii) Asymptotically Feasible

$$\lim_{k\to\infty} dist(x_k, X) = 0$$

where $dist(x, X)$ is the distance function from x to the set X measured in the Euclidean norm.

After defining a minimizing sequence, we introduce a stationary sequence in residual function approach instead of subdifferential approach. The advantage is that it is convenient in computing stopping rules in practical problems and allows the treatment of infeasible sequences. For a unconstrained problem that $X = \mathbb{R}^n$, as f is differentiable on $\mathbb{R}^n$, we simply define a sequence $\{x_k\}$ to be stationary if $\{\nabla f(x_k)\}$ converges to zero. In the case that X is a proper closed convex subset of $\mathbb{R}^n$, we have two kinds of stationary sequences(N-stationary and $\mathcal{N}$-stationary).

Definition 1.1.2 *The residual function $R_N : \mathbb{R}^n \to \mathbb{R}^n$ at $x \in \mathbb{R}^n$ is*

$$R_N(x) = x - \Pi_X(x - \nabla f(x))$$

where Π_X is the projection onto the set X.

It is easily seen that $X = \mathbb{R}^n$, $R_N(x) = \nabla f(x)$. Note also that if x is a local minimum of (1.1) and X is convex, then

$$\langle \nabla f(x), z - x\rangle = \lim_{t\downarrow 0} \frac{f(x + t(z - x)) - f(x)}{t} \geq 0 \quad \forall z \in X \tag{1.2}$$

and it follows from the well-known variational inequality that

$$x \in \Pi_X(x - \nabla f(x)) \tag{1.3}$$

implying that $R_N(x) = 0$. Conversely, if f is a convex function and $x \in \mathbb{R}^n$ with $R_N(x) = \nabla f(x) = 0$, then (1.3) holds. Consequently x is a global minimum of (1.1) because

$$f(z) \geq \langle \nabla f(x), z - x\rangle + f(x) \geq f(x) \quad \forall z \in X$$

by the virtue of the convexity of f.

Definition 1.1.3 *A sequence $\{x_k\} \subseteq \mathbb{R}^n$ is a naturally stationary (N-stationary) sequence if* $\lim_{k\to\infty} R_N(x_k) = 0$.

A N-stationary sequence $\{x_k\}$ is asymptotically feasible:

$$\lim_{k\to\infty} R_N(x_k) = 0 \implies x_k - \Pi_X(x_k - \nabla f(x_k)) \to 0 \quad \text{as}\, k \to \infty$$
$$\implies \lim_{k\to\infty} dist(x_k, X) = 0.$$

Also note that if X is convex, then by the variational principle of the Euclidean projector, for all $y \in X$ and $x \in \mathbb{R}^n$, one has

$$\begin{aligned} &(y - x + R_N(x))^T(\nabla f(x) - R_N(x)) \\ = &-(y - \Pi_X(x - \nabla f(x)))^T(x - \nabla f(x) - \Pi_X(x - \nabla f(x))) \\ \geq &\ 0. \end{aligned} \tag{1.4}$$

Definition 1.1.4 *A normal function $R_{\mathcal{N}} : \mathbb{R}^n \to \mathbb{R}^n$ is defined as*

$$R_{\mathcal{N}}(z) = \nabla f \circ \Pi_X(z) + z - \Pi_X(z) \quad for\, z \in \mathbb{R}^n.$$

Definition 1.1.5 *A sequence $\{x_k\} \subseteq \mathbb{R}^n$ is said to be normally stationary ($\mathcal{N}$-stationary) if*

(i) $\{x_k\}$ is asymptotically feasible such that $\lim_{k\to\infty} dist(x_k, X) = 0$

and

(ii) there exists a sequence $\{z_k\} \subseteq \mathbb{R}^n$ such that $\Pi_X(z_k) = \Pi_X(x_k)$ for each k and $\lim_{k\to\infty} R_{\mathcal{N}}(z_k) = 0$.

Consider a special case when $\{x_k\}$ is a constant sequence $\{x\}$: $x_k = x$ for all $k \in \mathbb{N}$. Then (i) simply says that $x \in X$. Moreover, assuming (i), (ii) is equivalent to the condition that $\{x_k\} = \{x\}$ is N-stationary as the following result shows

Proposition 1.1.3 *Let $x \in X$. Then the following statements are equivalent:*

(i) $\langle \nabla f(x), y - x\rangle \geq 0 \quad \forall y \in X$;

(ii) $x = \Pi_X(x - \nabla f(x))$;

(iii) $R_N(x) = 0$;

(iv) There exists $z \in \mathbb{R}^n$ *with* $\Pi_X z = x$ *such that* $R_{\mathcal{N}}(z) = 0$.

Proof The equivalence of (i) and (ii) is due to the variational inequality. That (ii) is equivalent to (iii) by definition of R_N.

(ii) $\Leftrightarrow$ (iv) Let $z = x - \nabla f(x)$. Then $\Pi_X z = x$ by (ii). By the definition of $R_{\mathcal{N}}$ it is east to check that $R_{\mathcal{N}}(z) = 0$. Conversely, suppose that (iv) holds for some $z \in \mathbb{R}^n$. Then, since $R_{\mathcal{N}}(z) = 0$, $z = x - \nabla f(x)$. Also, since $\Pi_X z = x$,

$$\langle x - z, y - x\rangle \geq 0 \quad \forall y \in X$$

by virtue of the variational inequality. Therefore, (i) is seen to be hold. □

For $z \in \mathbb{R}^n$, let $\bar{x} = \Pi_X(z)$, we have the following relation between R_N and $R_{\mathcal{N}}$,

$$\begin{aligned} R_N(\bar{x}) &= \bar{x} - \Pi_X(\bar{x} - \nabla f(\bar{x})) \\ &= \Pi_X(z) - \Pi_X(\bar{x} - \nabla f(\bar{x})) \\ &= \Pi_X(\bar{x} - \nabla f(\bar{x}) + \nabla f \circ \Pi_X(z) + z - \Pi_X(z)) - \Pi_X(\bar{x} - \nabla f(\bar{x})) \\ &= \Pi_X(\bar{x} - \nabla f(\bar{x}) + R_{\mathcal{N}}(z)) - \Pi_X(\bar{x} - \nabla f(\bar{x})). \end{aligned}$$

By the nonexpansiveness of projection, we then have

$$\|R_N(\bar{x})\| \leq \|R_{\mathcal{N}}(z)\|.$$

Therefore, for any sequence $\{z_k\} \subseteq \mathbb{R}^n$ and $\{x_k\} \subseteq X$ with $\bar{x}_k = \Pi_X(z_k)$ for each k, one has

$$\lim_{k\to\infty} R_N(\bar{x}_k) = 0 \qquad \text{if} \qquad \lim_{k\to\infty} R_{\mathcal{N}}(z_k) = 0. \tag{1.5}$$

Definition 1.1.6 *A function $h : \mathbb{R}^n \to \mathbb{R}^m$ is said to be uniformly continuous near a sequence $\{x_k\}$ if for all $\epsilon > 0$, there exists a $\delta > 0$ and for all k and y*

$$\|h(y) - h(x_k)\| \leq \epsilon \quad \textit{whenever} \quad \|y - x_k\| \leq \delta.$$

1.2 Minimizing sequences

In this section, we assumed that f_{inf} for (1.1) is finite: $f_{inf} > -\infty$. We want to investigate a question: if $\{x_k\}$ is an LP-minimizing sequence, is it necessarily an N-stationary sequence? The answer of this question is negative in general. The following proposition shows however that there exists a sequence near to $\{x_k\}$ which is LP-minimizing as well as N-stationary under a suitably uniform continuity of f.

Proposition 1.2.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function and X be an nonempty closed convex subset of $\mathbb{R}^n$. If f_{inf} is finite, $\{x_k\} \subseteq \mathbb{R}^n$ is an LP-minimizing sequence and f is uniformly continuous near $\{x_k\}$, this condition can be dropped if $\{x_k\}$ is assumed to be in X, then there exists a nearby feasible sequence $\{y_k\} \subseteq X$ such that*

(i) $\lim_{k\to\infty} \|x_k - y_k\| = 0$

(ii) $\lim_{k\to\infty} f(y_k) = f_{inf}$ *and*

(iii) $\lim_{k\to\infty} R_N(y_k) = 0$.

Proof For each k, let $\bar{x}_k = \Pi_X(x_k)$. As $\{x_k\}$ is asymptotically feasible, $\lim_{k\to\infty} \|x_k - \bar{x}_k\| = 0$. It then follows by the uniformly continuity of f near $\{x_k\}$ that $\lim_{k\to\infty} f(\bar{x}_k) = \lim_{k\to\infty} f(x_k) = f_{inf}$. Thus an LP-minimizing sequence since $\{\bar{x}_k\}$ is certainly feasible. Take an arbitrary sequence (ϵ_k) with $\epsilon_k > 0$ for each k such that

$$\lim_{k\to\infty} \epsilon_k = 0 \quad \text{and} \quad f(\bar{x}_k) < f_{inf} + \epsilon_k.$$

Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a function that $\psi(x) = f(x) + I_X(x)$ where

$$I_X(x) = \begin{cases} 0, & \text{if } x \in X \\ \infty & \text{otherwise.} \end{cases}$$

Since $\psi_{inf} = f_{inf}$ and $\{\bar{x}_k\}$ is feasible,

$$\psi(\bar{x}_k) = f(\bar{x}_k) < f_{inf} + \epsilon_k = \psi_{inf} + \epsilon_k \quad \text{for each } k.$$

By the Ekeland's variational principle in Lemma (3.2.1), there exists a sequence $\{y_k\} \subseteq \mathbb{R}^n$ with

$$g_k(y) = \sqrt{\epsilon_k}\|y - y_k\| \quad \text{for } y \in \mathbb{R}^n,$$

such that

(a) $\|\bar{x}_k - y_k\| \leq \sqrt{\epsilon_k}$

(b) $\psi(y_k) \leq \psi(\bar{x}_k)$

(c) y_k is a global minimizer of $\psi + g_k$.

By (a), $\lim_{k\to\infty} \|x_k - y_k\| = 0$ for ϵ_k converges to zero and $\lim_{k\to\infty} \|x_k - \bar{x}_k\| = 0$. $y_k \in X$ as $\psi = \infty$ outside X. By the definition of ψ and (c), y_k is a global minimizer of

$$\begin{aligned} &\text{minimize} \quad f(x) + g_k(x) \\ &\text{subject to} \quad x \in X. \end{aligned}$$

Since $\lim_{k\to\infty} f(\bar{x}_k) = f_{inf}$, it follows from (b) that $\lim_{k\to\infty} f(y_k) = f_{inf}$.
For $y \in X$, the directional derivative of $f + g_k$ at y_k along $y - y_k$ must be nonnegative. So there exists $w_k \in \partial(f + g_k)(y_k)$ (Clarke's subdifferential) such that $(w_k)^T(y - y_k) \geq 0$. By Proposition (1.1.1), $y_k = \Pi_X(y_k - w_k)$. By Lemma

(3.2.3), w_k can be written as $\partial f(y_k) + z_k$ with $\|z_k\| \le \sqrt{\epsilon_k}$. Since Π_X is non-expansive by Proposition (1.1.2), one has

$$\begin{aligned}\|y_k - \Pi_X(y_k - \partial f(y_k)\| &= \|\Pi_X(y_k - w_k) - \Pi_X(y_k - \partial f(y_k)\| \\ &\le \|w_k - \partial f(y_k)\| \\ &= \|z_k\| \\ &\le \sqrt{\epsilon_k} \to 0\end{aligned}$$

that is

$$\lim_{k\to\infty} R_N(y_k) = 0.$$

□

Noticed that $f_{inf} > -\infty$ is essential for applying Ekeland's variational principle in Proposition (1.2.1). There is a counterexample to illustrate it: $X = \mathbb{R}$, we want to minimize $f(x) = x, \forall x \in \mathbb{R}$. Clearly, any sequence converging to $-\infty$ is minimizing, but $f^{'}(x) = 1 \neq 0$ for all x.

With the above proposition, we immediate have the following main result in this section.

Theorem 1.2.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable and X be an nonempty closed convex subset of $\mathbb{R}^n$. Assumed $f_{inf} > -\infty$. Suppose a sequence $\{x_k\} \subseteq \mathbb{R}^n$ is an LP-minimizing sequence, f and ∇f are uniformly continuous near $\{x_k\}$. Then $\{x_k\}$ is N-stationary.*

Proof Follow the proof in Proposition (1.2.1), we have a nearby N-stationary sequence $\{y_k\}$. Then by the uniformly continuity property of f, ∇f near $\{x_k\}$

and $\lim_{k\to\infty} \|x_k - y_k\| = 0$,

$$\begin{aligned}
\|R_N(x_k) - R_N(y_k)\| &= \|[x_k - \Pi_X(x_k - \nabla f(x_k))] - [y_k - \Pi_X(y_k - \nabla f(y_k))]\| \\
&= \|x_k - y_k - [\Pi_X(x_k - \nabla f(x_k)) - \Pi_X(y_k - \nabla f(y_k))]\| \\
&\leq \|x_k - y_k\| + \|x_k - y_k - \nabla f(x_k) + \nabla f(y_k)\| \\
&\leq 2\|x_k - y_k\| + \|\nabla f(x_k) - \nabla f(y_k)\| \\
&\to 0 \quad \text{as } k \to \infty.
\end{aligned}$$

Hence, $\lim_{k\to\infty} R_N(x_k) = 0$. □

1.3 Stationary sequences

In this section, we deal with a problem when does a stationary sequence becomes an LP-minimizing sequence. We used error bounds for the level sets of (1.1). For the analysis purpose, an assumption that f is convex for (1.1) is needed.

Definition 1.3.1 *The minimizing problem (1.1) is said to have H-metrically regular level sets, or in short is H-metrically regular if for each scalar $\lambda > f_{inf}$, there exists $c > 0$ and $0 < \gamma \leq 1$ such that*

$$dist(x, L(\lambda)) \leq cr_\gamma(x), \quad \forall x \in X \tag{1.6}$$

where $L(\lambda)$ is the λ-level set of (1.1),

$$L(\lambda) = \{x \in X \mid f(x) \leq \lambda\}$$

and $r_\gamma(x)$ is the residual for $L(\lambda)$,

$$r_\gamma(x) = \max([(f(x) - \lambda)_+]^\gamma, (f(x) - \lambda)_+).$$

Theorem 1.3.1 *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differential convex function and X is a nonempty closed convex subset of $\mathbb{R}^n$. If a sequence $\{x_k\} \subseteq \mathbb{R}^n$ satisfies at least one of the conditions:*

(i) $\{x_k\}$ *is a* $\mathcal{N}$*-stationary sequence and* $\{x_k\}$ *is feasible to (1.1);*

(ii) $\{x_k\}$ *is a* $\mathcal{N}$*-stationary sequence and* f *is uniformly continuous near* $\{x_k\}$*;*

(iii) $\{x_k\}$ *is a N-stationary sequence,* ∇f *is uniformly continuous near* $\{x_k\}$ *and* $\{\nabla f(x_k)\}$ *is bounded,*

then $\{x_k\}$ *is an LP-minimizing sequence whenever (1.1) is H-metrically regular.*

Proof Suppose (i) holds, we need to show $\lim_{k\to\infty} f(x_k) = f_{inf}$. Assume on the contrary that there exists $\lambda \in \mathbb{R}$ such that $\liminf_{k\to\infty} f(x_k) > \lambda > f_{inf}$. Since the level set $L(\lambda)$ is nonempty closed convex, for each k there exists $y_k \in L(\lambda)$ such that $dist(x_k, L(\lambda)) = \|x_k - y_k\|$. Also, $f(y_k) = \lambda$. Otherwise for X is convex, there exists $w_k \in L(\lambda)$ for each k where $w_k \neq x_k, y_k$ which lies on the line segment joining x_k and y_k in X and $\|w_k - x_k\| < \|x_k - y_k\|$. Noted that by the gradient inequality for convex function f , we have

$$\lambda = f(y_k) \geq f(x_k) + \nabla f(x_k)^T(y_k - x_k). \tag{1.7}$$

For $\{x_k\}$ is $\mathcal{N}$-stationary, there exists a sequence $\{z_k\} \subseteq \mathbb{R}^n$ such that $x_k = \Pi_X(z_k)$ for all k with

$$\lim_{k\to\infty} R_{\mathcal{N}}(z_k) = 0.$$

Then it follows by the definition of $R_{\mathcal{N}}(z_k)$,

$$\begin{aligned}
&(y_k - x_k)^T(\nabla f(x_k) - R_{\mathcal{N}}(z_k)) \\
&\qquad = (y_k - x_k)^T(\nabla f(x_k) - (\nabla f \circ \Pi_X(z_k) + z_k - \Pi_X(z_k))) \\
&\qquad = (y_k - x_k)^T(\nabla f(x_k) - z_k + \Pi_X(z_k) - \nabla f \circ \Pi_X(z_k)) \\
&\qquad = (y_k - x_k)^T(x_k - z_k) \\
&\qquad \geq 0.
\end{aligned} \tag{1.8}$$

It becomes

$$\nabla f(x_k)^T(y_k - x_k) \geq R_{\mathcal{N}}(z_k)^T(y_k - x_k).$$

By (1.7),

$$f(x_k) - \lambda \leq -R_{\mathcal{N}}(z_k)^T(y_k - x_k);$$

Since $dist(x_k, L(\lambda)) = \|y_k - x_k\|$ and $f(x_k) > \lambda$, by the Cauchy-schwartz inequality and (1.6), the inequality becomes

$$\begin{aligned} f(x_k - \lambda &\leq c\,\|R_{\mathcal{N}}(z_k)\|\, dist(x_k, L(\lambda)) \\ &\leq c\,\|R_{\mathcal{N}}(z_k)\|\, \max(\,f(x_k) - \lambda,\, (f(x_k) - \lambda)^{\gamma}\,). \end{aligned}$$

Dividing both sides by $f(x_k) - \lambda$, one has

$$1 \leq c\|R_{\mathcal{N}}(z_k)\| \max(\,1, (f(x_k) - \lambda)^{\gamma-1}\,).$$

As $\liminf_{k\to\infty} f(x_k) > \lambda$ and $\{R_{\mathcal{N}}(z_k)\}$ converges to zero, a contradiction to the above inequality is obtained by letting $k \to \infty$.

Assumed (ii) that $\{x_k\}$ is $\mathcal{N}$-stationary. There exists a sequence $\{z_k\} \subseteq \mathbb{R}^n$ such that $\Pi_X(z_k) = \Pi_X(x_k)$ with $\lim_{k\to\infty} R_{\mathcal{N}}(z_k) = 0$. Let $\bar{x}_k = \Pi_X(x_k)$ for each k, then $\Pi_X(z_k) = \bar{x}_k$ for each k and $\lim_{k\to\infty} R_{\mathcal{N}}(z_k) = 0$. That means $\{\bar{x}_k\}$ is also a $\mathcal{N}$-stationary sequence. It follows that $\{\bar{x}_k\}$ satisfies (i) and thus

$$\lim_{k\to\infty} f(\bar{x}_k) = f_{inf}.$$

Moreover, f is uniformly continuous near $\{x_k\}$ and the asymptotically feasibility of $\{x_k\}$, $\lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} f(\bar{x}_k) = f_{inf}$.

Supposed (iii) that $\{x_k\}$ is a N-stationary sequence and so asymptotically feasible. Let $\bar{x}_k = \Pi_X(x_k)$ for each k. It follows that $\{\bar{x}_k\}$ is a N-stationary sequence as f is uniformly continuous near $\{x_k\}$.

Next we need to show that f is uniformly continuous near $\{x_k\}$ first. By the mean-value theorem, for all $y \in \mathbb{R}^n$, there exists $0 \leq \zeta \leq 1$ such that

$$\begin{aligned} f(y) - f(x_k) &= \nabla f(x_k + \zeta_k(y - x_k))^T(y - x_k) \\ &= [\nabla f(x_k + \zeta_k(y - x_k)) - \nabla f(x_k)]^T(y - x_k) + \nabla f(x_k)^T(y - x_k). \end{aligned}$$

One can easily see that for $\{\nabla f(x_k)\}$ is bounded and ∇f is uniformly continuous near $\{x_k\}$, f is thus uniformly continuous near $\{x_k\}$ from the above inequality. We then follow the above proof that suppose $\liminf_{k\to\infty} f(\bar{x}_k) > \bar{\lambda} > f_{inf}$ for some scalar $\bar{\lambda} \in \mathbb{R}^n$. Then there exists $\bar{y}_k \in L(\bar{\lambda})$ such that $dist(\bar{x}_k,\, L(\bar{\lambda})) = \|\bar{x}_k - \bar{y}_k\|$. By equation (1.4) for $\{x_k\}$ is N-stationary, we get the following inequality instead of (1.8)

$$(y_k \;-\; \bar{x}_k \,+\, R_N(\bar{x}_k))^T(\nabla f(\bar{x}_k) \;-\; R_N(\bar{x}_k)) \,\geq\, 0. \tag{1.9}$$

Also, by the gradient inequality for convex function f the inequality (1.7) becomes

$$\lambda \,=\, f(\bar{y}_k) \,\geq\, f(\bar{x}_k) \,+\, \nabla f(\bar{x})^T(\bar{y}_k \;-\; \bar{x}_k). \tag{1.10}$$

Then

$$\begin{aligned} f(\bar{x}_k) \;-\; \lambda &= f(\bar{x}_k) \;-\; f(\bar{y}_k) \\ &\leq -(\bar{y}_k \;-\; \bar{x}_k)^T(\nabla f(\bar{x}_k)) \\ &\leq -R_N(\bar{x}_k)^T(\bar{y}_k \;-\; \bar{x}_k) \,+\, R_N(\bar{x}_k)^T\nabla f(\bar{x}_k) \;-\; R_N(\bar{x}_k)^T R_N(\bar{x}_k) \\ &\leq \|R_N(\bar{x}_k)\|\, dist(\bar{x}_k, L(\lambda)) \,+\, R_N(\bar{x}_k)^T\nabla f(\bar{x}_k) \\ &= c\,\|R_N(\bar{x}_k)\|\, \max\{(f(\bar{x}_k) \;-\; \bar{\lambda}), (f(\bar{x}_k) \;-\; \bar{\lambda})^\gamma\} \,+\, R_N(\bar{x}_k)^T\nabla f(\bar{x}_k), \end{aligned}$$

where the first inequality follows by (1.10), the second inequality follows by (1.9) and the third one follows by H-metrically regularity of (1.1). Similarly, for $f(\bar{x}_k) > \bar{\lambda}$, we divide both sides by $f(\bar{x}_k) - \lambda$,

$$1 \,\leq\, c\,\|R_N(\bar{x}_k)\|\, \max\{1,\, (f(\bar{x}_k) \;-\; \bar{\lambda})^{\gamma-1}\} \,+\, R_N(\bar{x}_k)^T\nabla f(\bar{x}_k)/(f(\bar{x}_k) \;-\; \lambda)$$

As $\{x_k\}$ is N-stationary, $\{\nabla f(\bar{x}_k)\}$ is bounded and ∇f is uniformly continuous near $\{x_k\}$, we obtain a contradiction as before by letting $k \to \infty$. Therefore, $\lim_{k\to\infty} f(\bar{x}_k) \,=\, f_{inf}$. Consequently, $\lim_{k\to\infty} f(x_k) \,=\, f_{inf}$ which follows by the uniformly continuity of f near $\{x_k\}$. □

1.4 On the equivalence of minimizing and stationary sequence

A piecewise quadratic function $g : \mathbb{R}^n \to \mathbb{R}$ is a continuous function and there exists finitely many convex polyhedra P_i, $i = 1, \ldots, p$ for some positive integer p, such that $\bigcup_{i=1}^{p} P_i = \mathbb{R}^n$ and g is quadratic on each P_i. Similarly, a piecewise linear function $h : \mathbb{R}^n \to \mathbb{R}^m$ is defined as a vector function which is continuous and there exists finitely many convex polyhedra S_i, $i = 1, \ldots, s$ for some positive integer s, such that $\bigcup_{i=1}^{s} S_i = \mathbb{R}^n$ and g is affine on each S_i. From [20], we know that a piecewise linear function is a globally Lipschitz function. It can be easily seen that for a differentiable piecewise quadratic function f, it's gradient ∇f is piecewise linear and hence globally Lipschitz.

Lemma 1.4.1 *If $f : \mathbb{R}^n \to \mathbb{R}$ is a convex piecewise quadratic function and*

$$X = \{x \in \mathbb{R}^n \mid g_i(x) \le 0, \; i = 1, \ldots, m\}$$

where g_i is a convex quadratic function and m is a positive integer, then the optimization problem (1.1) has H-metrically regular level sets.

Proof Let $\lambda > f_{inf}$, $\{P_i \mid i = 1, \ldots, p\}$ be the finite convex polyhedra where $\bigcup_{i=1}^{p} P_i = \mathbb{R}^n$ and f is quadratic on each P_i. Suppose f equals to a quadratic function q_i on each P_i. We can see that

$$L(\lambda) = \bigcup_{i=1}^{p} L_i(\lambda)$$

where $L_i(\lambda)$ is the intersection of a level set on P_i and the feasible set X, that is

$$L_i(\lambda) = \{x \in X \cap P_i \mid f(x) \le \lambda\}.$$

By the error bound in [21], for each i where $L_i(\lambda) \ne \emptyset$, there exist positive constants c_i and γ_i with $0 < \gamma_i < 1$ such that

$$dist(x, L_i(\lambda)) \le c_i \max\left([q_i(x) - \lambda)_+]^{\gamma_i}, (q_i(x) - \lambda)_+ \right) \quad \forall x \in X \cap P_i.$$

As

$$dist(x, L(\lambda)) \leq dist(x, L_i(\lambda)) \quad \forall x \in X$$

and q_i coincides with f on P_i for each i, we have

$$dist(x, L(\lambda)) \leq c_i \max\left([f(x) - \lambda)_+]^{\gamma_i}, (f(x) - \lambda)_+ \right) \quad \forall x \in X \cap P_i.$$

By taking

$$c = \max\{c_i \mid i = 1, \ldots, p\} \quad \text{and} \quad \gamma = \min\{\gamma_i \mid i = 1, \ldots, p\},$$

we thus get

$$dist(x, L(\lambda)) \leq c \max\left([f(x) - \lambda)_+]^{\gamma}, (f(x) - \lambda)_+ \right) \quad \forall x \in X.$$

□

A convex quadratically constrained quadratic spline (convex differentiable piecewise quadratic function f) program: (CQQSP)

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad G_i(x) = \frac{1}{2}x^T C_i x + a_i^T x + b_i \leq 0 \quad for\, i = 1, \ldots, m \end{aligned} \qquad (1.11)$$

where C_i is a $n \times n$ symmetric positive semidefinite matrix, a_i is a n-vector and b_i is a scalar. Clearly, problem (1.11) is H-metrically regular by Lemma (1.4.1).

In the proof of Theorem (1.3.1), we have already prove that if ∇f is uniformly continuous near a sequence $\{x_k\}$ and $\{\nabla f(x_k)\}$ is bounded, then f is uniformly continuous near $\{x_k\}$. Consequently, one has the follow theorem:

Theorem 1.4.1 *Suppose f_{inf} is finite for the minimizing problem CQQSP (1.11) and a sequence $\{x_k\} \subseteq \mathbb{R}^n$ has a property that $\{\nabla f(x_k)\}$ is bounded. Then the sequence $\{x_k\}$ is N-stationary if and only if it is LP-minimizing.*

Proof ($\Rightarrow$) Let $\{x_k\}$ be a N-stationary sequence for problem (1.11). By the assumption that $\{\nabla f(x_k)\}$ is bounded and ∇f is uniformly continuous near $\{x_k\}$

which follows from the property that f is a differentiable piecewise quadratic function, the condition in Theorem (1.3.1) (iii) is satisfied. Also, the problem CQQSP (1.11) is H-metrically regular by Lemma (1.4.1) and the feasible set is nonempty closed convex. Hence, $\{x_k\}$ is an LP-minimizing sequence by Theorem (1.3.1).

($\Leftarrow$) Suppose $\{x_k\}$ is an LP-minimizing sequence. Firstly, ∇f is uniformly continuous near $\{x_k\}$. Together with $\{\nabla f(x_k)\}$ is bounded, f is uniformly continuous near $\{x_k\}$. Moreover, with the assumption that f_{inf} is finite and the feasible set is nonempty closed convex, applied Theorem (1.2.1), $\{x_k\}$ is N-stationary. $\square$

Lemma 1.4.2 *If $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and X is a closed convex subset of $\mathbb{R}^n$ and $c > 0$, the following properties (i) and (ii) are equivalent.*

(i) For each $x \in X$ with $f(x) > f_{inf}$, there exists $\bar{x} \in X$ with $f(\bar{x}) < f(x)$ and

$$\|\bar{x} - x\| \leq c(f(x) - f(\bar{x})).$$

(ii) For all $\lambda > f_{inf}$,

$$dist(x, L(\lambda)) \leq c\,(f(x) - \lambda)_+, \quad \forall x \in X.$$

Proof (i) $\Rightarrow$ (ii) Let $x \in X$ and $f(x) > \lambda > f_{inf}$. Consider the problem of projecting x onto $L(\lambda)$:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}(z - x)^T(z - x) \\ \text{subject to} \quad & z \in L(\lambda). \end{aligned}$$

Let $\hat{x} = \Pi_{L(\lambda)}(x)$. One has $f(\hat{x}) = \lambda$, otherwise there exist $y \in L(\lambda)$ such that $\|y - \hat{x}\| < \|\hat{x} - x\|$. Suppose $\bar{x} \in X$ such that $f(\bar{x}) < f(\hat{x})$ and

$$\|\bar{x} - \hat{x}\| \leq c\,(f(\hat{x}) - f(\bar{x})).$$

From the result of convex analysis [19], there exists a nonnegative scalar τ and a vector $u \in \partial f(\hat{x})$ such that for all $z \in X$,

$$(z - \hat{x})^T(\hat{x} - x + \tau u) \geq 0. \tag{1.12}$$

By putting $z = \bar{x}$, one has

$$\tau u^T(\bar{x} - \hat{x}) \geq -(\bar{x} - \hat{x})^T(\hat{x} - x).$$

From the definition of u, we have

$$f(\bar{x}) - f(\hat{x}) \geq u^T(\bar{x} - \hat{x}).$$

Then

$$\tau(f(\bar{x}) - f(\hat{x})) \geq -(\bar{x} - \hat{x})^T(\hat{x} - x).$$

Therefore,

$$\tau \leq c\,\|\hat{x} - x\|.$$

By putting $z = x$ in (1.12), we deduce

$$\|\hat{x} - x\|^2 \leq \tau u^T(x - \hat{x}) \leq \tau(f(x) - f(\hat{x})).$$

As $x \neq \hat{x}$ and $f(\hat{x}) = \lambda < f(x)$, we have

$$\|\hat{x} - x\| \leq c\,(f(x) - \lambda)_+.$$

Thus (ii) holds.

(ii) $\Rightarrow$ (i) Let $x \in X$ with $f(x) > \lambda > f_{inf}$ for some λ. Let $\bar{x}$ be the Euclidean projection of x onto the level set $L(\lambda)$. Then $\bar{x} \in X$ and $f(\bar{x}) = \lambda < f(x)$. By (ii), we have

$$\|\bar{x} - x\| = dist(x, L(\lambda)) \leq c\,(f(x) - \lambda)_+ = c\,f(x) - f(\bar{x}).$$

That is, (i) holds. $\square$

Theorem 1.4.2 *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is convex ,continuous differentiable function and X is a nonempty closed convex subset of $\mathbb{R}^n$ with the property that f_{inf} finite. Let $c > 0$ be a constant such that for each $x \in X$ satisfying $f(x) > f_{inf}$, there exists $\bar{x} \in X$ with $f(\bar{x}) < f(x)$ and*

$$\|\bar{x} - x\| \leq c(f(x) - f(\bar{x})).$$

For an arbitrary sequence $\{x_k\}$ for which ∇f is uniformly continuous near $\{x_k\}$ and the sequence $\{\nabla f(x_k)\}$ is bounded, then $\{x_k\}$ is N-stationary if and only if it is LP minimizing.

Proof ($\Rightarrow$) Assumed $\{x_k\}$ is a N-stationary sequence. The property in Lemma (1.4.2) (i) is satisfied, so the problem is H-metrically regular. Also, by the assumption that $\{\nabla f(x_k)\}$ is bounded and ∇f is uniformly continuous near $\{x_k\}$, it follows by the Theorem (1.3.1) that $\{x_k\}$ is an LP-minimizing sequence.

($\Leftarrow$) Suppose $\{x_k\}$ is a LP-minimizing sequence. Firstly, f is uniformly continuous near $\{x_k\}$ which follows by the boundedness of $\{\nabla f(x_k)\}$ and the uniformly continuity of ∇f near $\{x_k\}$. Also, f_{inf} is finite. With the assumption that ∇f is uniformly continuous near $\{x_k\}$, $\{x_k\}$ is N-stationary by Theorem (1.2.1). □

1.5 Complementarity conditions

Consider X is a closed convex cone in the (1.1). Let X^* be the dual cone of X such that

$$X^* \equiv \{u \in \mathbb{R}^n \mid u^T v \geq 0 \text{ for all } v \in X\}.$$

If x is a local minimum of the problem

$$\begin{aligned} &\text{Minimize} \quad f(x) \\ &\text{subject to} \quad x \in X \end{aligned} \tag{1.13}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and X is a nonempty closed convex cone of $\mathbb{R}^n$, then

$$\nabla f(x)^T(y - x) \geq 0 \quad \forall y \in X.$$

It can be easily seen that by the property of a cone X, the complementarity system holds:

$$x^T \nabla f(x) = 0 \tag{1.14}$$

for all $x \in X$ and $\nabla f(x) \in X^*$.
On the contrary, suppose f is a convex function and a vector $x \in \mathbb{R}^n$ satisfying (1.14), then

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

That is x is a global minimum of the problem (1.13).

We want to see that whether the above property (1.14) is satisfied by a sequence $\{x_k\}$ instead of a vector x.

Theorem 1.5.1 *Suppose that X is a nonempty closed convex cone in $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. Let $\{x_k\} \subseteq X$ be a feasible sequence,*

(i) If $f_{inf} > -\infty$, ∇f is uniformly continuous near $\{x_k\}$ and $\{x_k\}$ is LP minimizing for the problem (1.13), then there exists a sequence $\{w_k\} \subseteq X^*$ such that

$$\lim_{k \to \infty} (x_k)^T w_k = 0 \quad \textit{and} \quad \lim_{k \to \infty} (\nabla f(x_k) - w_k) = 0. \tag{1.15}$$

(ii) Conversely, suppose that f is a convex function and the problem (1.13) is H-metrically regular. If there exists a sequence $\{w_k\} \subseteq X^*$ having the properties (1.15), then $\{x_k\}$ is an LP-minimizing for the problem (1.13).

Proof (i) In the proof of Proposition (1.2.1), we deduced two sequences of vectors $\{y_k\} \subseteq X$ and $\{w_k\} \subseteq \mathbb{R}^n$ satisfying the following properties

(a) $y_k = \Pi_X(y_k - w_k)$ for each k;

(b) $\lim_{k\to\infty}(y_k - x_k) = 0$;

(c) $\lim_{k\to\infty} f(y_k) = f_{inf}$;

(d) $\lim_{k\to\infty}(w_k - \nabla f(y_k)) = 0$.

Noted that

$$\nabla f(x_k) - w_k = (\nabla f(x_k) - \nabla f(y_k)) + (\nabla f(y_k) - w_k).$$

Since ∇f is uniformly continuous near $\{x_k\}$, by (b) and (d), let k goes to infinity we have

$$\lim_{k\to\infty} (\nabla f(x_k) - w_k) = 0.$$

Next, we are going to prove $\lim_{k\to\infty}(x_k)^T w_k = 0$. Noticed that

$$f(y_k) - f(x_k) = (y_k - x_k)^T \nabla f(x_k) + o(\|x_k - y_k\|)$$

where $o(t)$ converges to zero if $t \downarrow 0$. Also, $\{y_k\}$ is LP-minimizing by (c), let k goes to infinity it then follows that

$$\lim_{k\to\infty} (x_k - y_k)^T \nabla f(x_k) = 0. \tag{1.16}$$

On the other hand, for X is a convex cone,

$$(y_k)^T w_k = 0$$

Thus,

$$\begin{aligned}
(x_k)^T(w_k) &= (x_k - y_k)^T w_k + (y_k)^T w_k \\
&= (x_k - y_k)^T w_k \\
&= (x_k - y_k)^T(w_k - \nabla f(y_k)) + (x_k - y_k)^T \nabla f(y_k) + (x_k - y_k)^T(\nabla f(y_k) - \nabla f(x_k)).
\end{aligned}$$

Then by the uniformly continuity of ∇f near $\{x_k\}$, (b), (d) and (1.16), we establishes

$$\lim_{k\to\infty} (x_k)^T w_k = 0.$$

(b) Suppose on the contrary that there exists a scalar α such that $\liminf_{k\to\infty} f(x_k) > \alpha > f_{inf}$. For the problem (1.13) is H-metrically regular, we can let

$$dist(x_k, L(\alpha)) \le cr_\gamma(x_k) \quad \text{for each}\, k$$

where $r_\gamma(x_k)$ is the residual for $L(\alpha)$,

$$r_\gamma(x_k) = \max([(f(x_k) - \alpha)_+]^\gamma, (f(x_k) - \alpha)_+).$$

Moreover, there exists $y_k \in X$ for each k such that

$$f(y_k) = \alpha \quad \text{and} \quad dist(x_k, L(\alpha)) = \|x_k - y_k\|. \tag{1.17}$$

Since f is convex,

$$\nabla f(x_k)^T (x_k - y_k) \ge f(x_k) - \alpha. \tag{1.18}$$

Also, with $(w_k)^T y_k = 0$,

$$\begin{aligned}\nabla f(x_k)^T (x_k - y_k) &= (w_k)^T (x_k - y_k) + (\nabla f(x_k) - w_k)^T (x_k - y_k) \\ &\le |(w_k)^T x_k| + \|\nabla f(x_k) - w_k\| \|y_k - x_k\|.\end{aligned}$$

Thus by (1.18) and (1.17),

$$f(x_k) - \alpha \le |(w_k)^T x_k| + \|\nabla f(x_k) - w_k\| dist(x_k, L(\alpha).$$

As $f(x_k) - \alpha \neq 0$ for all k, we then divide both sides by $f(x_k) - \alpha$, one has

$$1 \le c\|\nabla f(x_k) - w_k\| \max(1, (f(x_k) - \alpha)^{\gamma-1}) + \frac{|(w_k)^T x_k|}{f(x_k) - \alpha}.$$

By letting k goes to infinity, the right hand side of the inequality becomes 0 while the left hand equals to 1. □

1.6 Subdifferential-based stationary sequence

Let ∂ denote the Clark's subdifferential that

$$f^0(x;u) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{1}{t}\{f(y+tu) - f(y)\} \quad \text{for } x, u \in X$$

and

$$\partial f(x) = \{x^* \in X^* \mid x^*(\cdot) \leq f^0(x;\cdot) \quad \text{on } X\}.$$

Suppose that X is a closed convex subset in $\mathbb{R}^n$ and I_X is the indicator function of X, such that

$$I_X(x) = \begin{cases} 0, & \text{if } x \in X \\ \infty & \text{if } x \notin X. \end{cases}$$

Note that I_X is a convex function. By [10] ∂I_X is defined as

$$\partial I_X(x) = \{u \in \mathbb{R}^n \mid u^T(y-x) \leq I_X(y) - I_X(x) \quad \forall y \in X\}.$$

We use ∂_ϵ to denote the ϵ-subdifferential [12]. The subdifferential $\partial_\epsilon I_X(x)$ of I_X at $x \in X$ is

$$\partial_\epsilon I_X(x) = \{u \in \mathbb{R}^N \mid u^T(y-x) \leq \epsilon \quad \forall y \in X\}.$$

Consider the function f and X in (1.1) are both convex such that

$$\begin{aligned} &\text{Minimize} \quad f(x) \\ &\text{subject to} \quad x \in X \end{aligned} \tag{1.19}$$

where f:$\mathbb{R}^n \to \mathbb{R}$ is continuously differentiable convex function and X is an nonempty closed convex subset in $\mathbb{R}^n$. Let a function $\psi : \mathbb{R}^n \to \mathbb{R}$ be such that

$$\psi(x) \equiv f(x) + I_X(x) \quad \forall x \in \mathbb{R}^n.$$

Definition 1.6.1 *A sequence $\{x_k\} \subseteq \mathbb{R}^n$ is said to be AC-stationary (AC for Auslender and Crouzeix) if*

(i) $\{x_k\}$ *is asymtotically feasible;*

(ii) *for each* k *with* $\bar{x}_k = \Pi_X(x_k)$*, there exists* $\alpha_k \in \partial\psi(\bar{x}_k)$ *and the sequence* $\{\alpha_k\}$ *converges to zero.*

Similarly, AC_ϵ-stationary sequence is defined as follows:

Definition 1.6.2 *A sequence* $\{x_k\} \subseteq \mathbb{R}^n$ *is said to be* AC_ϵ*-stationary if*

(i) $\{x_k\}$ *is asymptotically feasible;*

(ii) *for each* k *with* $\bar{x}_k = \Pi_X(x_k)$*, there exists* $\alpha_k \in \partial_{\epsilon_k}\psi(\bar{x}_k)$ *for some sequence of nonnegative scalars* $\{\epsilon_k\}$ *which converges to zero. The sequence* $\{\alpha_k\}$ *converges to zero as well.*

For a sequence which is AC-stationary, it is $\mathcal{N}$-stationary. Such relation is illustrated by the following theorem. In addition, the relation between N-stationarity, AC-stationarity and AC_ϵ-stationarity is mentioned with some criteria.

Theorem 1.6.1 *Suppose that* X *is a nonempty closed convex subset of* $\mathbb{R}^n$ *and* $f : \mathbb{R}^n \to \mathbb{R}$ *is a convex, continuously differentiable function. Let* $\{x_k\}$ *be a sequence in* $\mathbb{R}^n$*, we have the following properties:*

(i) *The sequence* $\{x_k\}$ *is AC-stationary if and only if it is* $\mathcal{N}$*-stationary.*

(ii) *If* $\{x_k\}$ *is AC-stationary and each* x_k *is feasible, then* $\{x_k\}$ *is N-stationary.*

(iii) *If* $\{x_k\}$ *is AC-stationary and* ∇f *is uniformly continuous near* $\{x_k\}$*, then* $\{x_k\}$ *is N-stationary.*

(iv) *If* $\{x_k\}$ *is N-stationary,* $\{\nabla f(x_k)\}$ *is bounded and* ∇f *is uniformly continuous near* $\{x_k\}$*, then* $\{x_k\}$ *is* AC_ϵ*-stationary.*

Proof (i) Let $\{x_k\}$ be a AC-stationary sequence. Write $\bar{x}_k = \Pi_X(x_k)$. Let $\phi = f + I_X$. Since f is differentiable,

$$\partial\phi(x) = \nabla f(x) + \partial I_X(x) \quad \forall x \in \mathbb{R}^n.$$

Suppose that a sequence $\{a_k\}$ have a property that $a_k \in \partial\phi(\bar{x}_k)$ for each k and a_k converging to zero. Take $b_k \in \partial I_X(\bar{x}_k)$ such that $a_k = \nabla f(\bar{x}_k) + b_k$. Also, for each k define $z_k = \bar{x}_k + b_k$. Then by definition of subdifferential of I_X

$$(\bar{x}_k - z_k)^T(y - \bar{x}_k) \geq I_X(\bar{x}_k) - I_X(y) = 0 \quad \forall y \in X.$$

It follows from the variational inequality, $\Pi_X(z_k) = \bar{x}_k$ for all k. Moreover, it can be easily seen that $R_{\mathcal{N}}(z_k) = a_k$ and hence converges to zero. This establishes the "only if" part. Conversely, suppose $\{x_k\}$ is a $\mathcal{N}$-stationary sequence. Then there exists $\{z_k\}$ such that $\Pi_X(z_k) = \Pi_X(x_k)$ for all k with $\lim_{k\to\infty} R_{\mathcal{N}}(z_k) = 0$. Write $\Pi_X(x_k) = \bar{x}_k$ for each k. Take $a_k = \nabla f(\bar{x}_k) + (z_k - x_k)$. It should be note that $z_k - \bar{x}_k \in \partial I_X(\bar{x})$ as $\Pi_X(z_k) = \bar{x}_k$ and the variational inequality $(\bar{x}_k - z_k)^T(y - \bar{x}_k) \geq 0$ for all $y \in X$. Hence $a_k = R_{\mathcal{N}}(z_k)$ converges to zero.

(ii) Since $\{x_k\}$ is a AC-stationary sequence, it follows from (i) that $\{x_k\}$ is also $\mathcal{N}$-stationary. Then for each x_k is feasible, there exists a sequence $\{z_k\} \subseteq \mathbb{R}^n$ satisfying $\Pi_X(z_k) = x_k$ for each k and

$$\lim_{k\to\infty} R_{\mathcal{N}}(z_k) = 0.$$

Consequently, by (1.5), $\lim_{k\to\infty} R_N(x_k) = 0$.

(iii) Similar to the proof in (ii), $\{x_k\}$ is $\mathcal{N}$-stationary. It follows that there exists a sequence $\{z_k\} \subseteq \mathbb{R}^n$ such that $\Pi_X(z_k) = \Pi_X(x_k) = \bar{x}_k$ for each k and

$$\lim_{k\to\infty} R_{\mathcal{N}}(z_k) = 0.$$

Once again, by (1.5),

$$\lim_{k\to\infty} R_N(\bar{x}_k) = 0.$$

As ∇f is uniformly continuous near $\{x_k\}$, by the definition of the residual function,

$$\lim_{k\to\infty} R_N(x_k) = \lim_{k\to\infty} R_N(\bar{x}_k) = 0.$$

(iv) Suppose $\{x_k\}$ is a N-stationary sequence, it is asymptotically feasible. Let $\bar{x}_k = \Pi_X(x_k)$ for each k. Since $\|x_k - \bar{x}_k\|$ converges to zero, it follows from the uniformly continuity of ∇f near $\{x_k\}$ and the boundedness of $\{\nabla f(x_k)\}$, one has $\{\nabla f(\bar{x}_k)\}$ is bounded and

$$\lim_{k\to\infty} R_N(\bar{x}_k) = 0. \tag{1.20}$$

Let

$$\epsilon_k = |R_N(\bar{x}_k)^T \nabla f(\bar{x}_K)| \quad \text{for each } k$$

the sequence $\{\epsilon_k\}$ converges to zero. By (1.4),

$$(R_N(\bar{x}_k) - \nabla f(\bar{x}_K))^T (y - \bar{x}_k) \le R_N(\bar{x}_k)^T (\nabla f(\bar{x}_k) - R_N(\bar{x}_k)) \le \epsilon_k$$

for all $y \in X$. This means $R_N(\bar{x}_k) \in \partial_{\epsilon_k} \psi(\bar{x}_k)$ for all k. Together with (1.20), the result follows. □

1.7 Convergence of an Iterative Algorithm

In this section, we consider a decent iterative method to solve an constrained minimization problem. Our goal is to show that the H-metric regularity condition in (1.6) is played an important role in the study of LP-minimizing sequences which is essential for the convergency of these method.

Now we consider the following problem

$$\begin{aligned} &\text{Minimize} \quad f(x) \\ &\text{subject to} \quad x \in X \end{aligned} \tag{1.21}$$

where X is a closed nonempty convex subset in $\mathbb{R}^n$ and f is a convex, continuously differentiable function on $\mathbb{R}^n$ to $\mathbb{R}$.

We are going to introduce an iterative decent method for solving this problem. This method generate a sequence of feasible vectors $\{x_k\}$ in X with decreasing objective function value $\{f(x_k)\}$. At each iteration, an iterater x_{k+1} is deduced by solving a convex subprogram in the form of

$$\begin{cases} \text{minimize} & \nabla f(x_k)^T d + \frac{1}{2} d^T M d \\ \text{subject to} & x_k + d \in X \\ \text{and} & \|d\| \leq \delta. \end{cases} \tag{1.22}$$

where M is a positive semidefinite matrix and δ is a positive scalar. This yield a feasible decent direction d_k for (1.21) at x_k. It then follow by the Armijo line search on f at x_k along d_k to get x_{k+1}.

Lemma 1.7.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function and X be a closed convex subset of $\mathbb{R}$. Also, let M be a symmetric positive semidefinite matrix. If $\bar{d}$ is an optimal solution of (1.22), then $\nabla f(x)^T \bar{d} \leq 0$ for all $x \in X$. In addition, if x has a property that $\nabla f(x)^T \bar{d} = 0$, then x is an optimal solution of f on X.*

Proof Since $\bar{d}$ is a local minimum of (1.22), we have

$$\langle \nabla f(x), \bar{d} \rangle \leq \langle \nabla f(x), \bar{d} \rangle + \frac{1}{2} \langle \bar{d}, M\bar{d} \rangle \leq \langle \nabla f(x), 0 \rangle = 0.$$

In the case that $\langle \nabla f(x), \bar{d} \rangle = 0$, let ϕ be a convex function defined on the convex set $(X - x) \cap \delta\mathcal{B}$ where $\phi(d) = \langle \nabla f(x), d \rangle + \frac{1}{2} \langle d, Md \rangle$. Since $\langle \bar{d}, Md \rangle \geq 0$ for all

$d \in \mathbb{R}^n$, $\langle \bar{d}, Md \rangle = 0$ for all $d \in \mathbb{R}^n$. For $\bar{d}$ is a minimum point of ϕ, one has

$$\langle \nabla f(x), d \rangle = \langle \nabla f(x), d \rangle + \langle \bar{d}, Md \rangle = \phi^0(\bar{d}, d) \geq 0$$

for feasible d. Therefore, x is a minimum point of f on X. □

Lemma 1.7.2 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable convex function, X be a closed convex subset of $\mathbb{R}^n$ and $\sigma \in (0,1)$. If $\nabla f(x)^T d < 0$, then for each $\rho \in (0,1)$, there exists an positive integer m such that*

$$f(x + \rho^m d) - f(x) \leq \sigma \rho^m \nabla f(x)^T d.$$

Proof Suppose that for all $m \in \mathbb{N}$,

$$f(x + \rho^m d) - f(x) > \sigma \rho^m \langle \nabla f(x), d \rangle.$$

Then dividing both sides by ρ^m and passing to limit, we get $\langle \nabla f(x), d \rangle \geq \sigma \langle \nabla f(x), d \rangle$. This contradicts the assumption that $\langle \nabla f(x), d \rangle < 0$ and $\sigma \in (0,1)$. □

A Descent Algorithm.

Step 1. Let $\rho, \sigma \in (0,1)$ be given scalar and $\delta > 0$ be an arbitrary constant. Initially, let x_0 be a given vector in X and M_0 be a symmetric positive semidefinite matrix. Set $k = 0$.

Step 2. Solve the convex subprogram in the variable $d \in \mathbb{R}^n$:

$$\begin{cases} \text{minimize} & \nabla f(x_k)^T d + \frac{1}{2} d^T M d \\ \text{subject to} & x_k + d \in X \\ \text{and} & \|d\| \leq \delta. \end{cases} \tag{1.23}$$

Let d_k be an optimal solution of this subprogram, such minimizer must exist as the objective function is quadratic and the feasible set is nonempty convex. Then it follows that $\nabla f(x_k)^T d_k \leq 0$. In case that $\nabla f(x_k)^T d_k = 0$, then x_k is an optimal solution of (1.21) by Lemma (1.7.1) and the iteration is terminated.

Otherwise, continue the iteration for $\nabla f(x_k)^T d_k < 0$.

Step 3. Let m_k be the smallest nonnegative integer m such that

$$f(x_k + \rho^m d) - f(x_k) \leq \sigma\rho^m \langle \nabla f(x_k), d\rangle. \tag{1.24}$$

Such integer always exists by Lemma (1.7.2). Set $\tau_k = \rho^{m_k}$ and $x_{k+1} = x_k + \tau_k d_k$.

Step 4. Test x_{k+1} whether it satisfies the prescribed stopping rule. If so, stop the iterations and x_{k+1} is a desired solution of the problem (1.21). Otherwise, choose a symmetric positive semidefinite matrix M_{k+1} and turn to step 2 with k replaced by $k+1$.

Theorem 1.7.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable convex function and X be a closed convex subset of $\mathbb{R}^n$. Also, let $\{M_k\}$ be a sequence of symmetric positive semidefinite matrices. Suppose that $\{x_k\}$ is an infinite sequence of vectors generated by the iterative descent algorithm. Assume that*

(i) $\{M_k\}$ is bounded;

(ii) ∇f is uniformly continuous near $\{x_k\}$;

(iii) for each $\lambda > f_{inf}$, there exists scalars $c > 0$ and $\gamma \in (0,1)$ such that for all k,

$$dist(x_k, L(\lambda)) \leq c\max\{(f(x_k) - \lambda)_+, [(f(x_k) - \lambda)_+]^\gamma\}.$$

Then $\{x_k\}$ is a minimizing sequence such that $\lim_{k\to\infty} f(x_k) = f_{inf}$.

Proof By (1.24), we can see that $\{f(x_k)\}$ is a decreasing sequence. Then the limit $\lim_{k\to\infty} f(x_k)$ exists and not smaller that f_{inf}. Suppose on the contrary that

$$\lim_{k\to\infty} f(x_k) > \lambda > f_{inf}$$

for some scalar λ. As $L(\lambda)$ is a closed convex subset of $\mathbb{R}^n$, for each vector $x \notin L(\lambda)$ there exists a unique vector $y \in L(\lambda)$ such that $\|x - y\| = dist(x, L(\lambda))$ and $f(y) = \lambda$. Therefore, for each x_k there exists $y_k \in L(\lambda)$ so that

$$\|x_k - y_k\| = dist(x_k, y_k) \leq c \max\{f(x_k) - \lambda)_+, [(f(x_k) - \lambda)_+]^\gamma\}.$$

If follows that the sequence $\{y_k - x_k\}$ is bounded. Thus, we define the scalar

$$\delta_k \equiv \begin{cases} 1 & \text{if } \|y_k - x_k\| \leq 1 \\ \|x_k - y_k\|^{-1} & \text{otherwise.} \end{cases}$$

Then the vector $d = \delta_k(y_k - x_k)$ is feasible to the subprogram (1.23). The boundedness of $\{x_k - y_k\}$ also implies

$$0 < \inf_{k\to\infty} \delta_k \leq \sup_{k\to\infty} \delta_k \leq 1.$$

Hence, the sequence $\{\delta_k^{-1}\}$ is also bounded.

Noted that $\nabla f(x_k)$ is the unique vector in $\mathbb{R}^n$ satisfying the gradient inequality

$$\nabla f(x_k)^T y - x_k \leq f(y) - f(x_k) \quad \text{for all } y \in \mathbb{R}^n. \tag{1.25}$$

Also, d_k is a minimum of the optimization subprogram (1.23), one has

$$d - d_k^T \nabla f(x_k) + M_k d_k \geq 0. \tag{1.26}$$

Combining the inequalities (1.25) and (1.26), for each k,

$$\begin{aligned} \lambda - f(x_k) &= f(y_k) - f(x_k) \\ &\geq \langle \nabla f(x_k), y_k - x_k \rangle \\ &\geq \frac{1}{\delta_k} d_k^T \nabla f(x_k) + M_k d_k - y_k - x_k^T M_k d_k. \end{aligned} \tag{1.27}$$

As d_k is an optimal solution of the subprogram (1.23), for each $k \in \mathbb{N}$,

$$\nabla f(x_k)^T d_k + \frac{1}{2} d_k^T M_k d_k \leq 0. \tag{1.28}$$

Then,

$$f(x_{k+1}) - f(x_k) \leq \sigma\tau_k \nabla f(x_k)^T d_k < 0.$$

Since $\lim_{k\to\infty} f(x_k)$ exists and hence $\lim_{k\to\infty}[f(x_{k+1}) - f(x_k)] = 0$, for $\sigma \in (0,1)$

$$\lim_{k\to\infty} \tau_k\langle\nabla f(x_k), d_k\rangle = 0. \tag{1.29}$$

If $\inf_{k\to\infty} \tau_k > 0$, then $\lim_{k\to\infty} f(x_k)^T d_k = 0$. If $\inf_{k\to\infty} \tau_k = 0$ and suppose the sequence $\{\tau_k\}$ itself converges to zero. Then $\lim_{k\to\infty} m_k = \infty$. We claim that

$$\lim_{k\to\infty} \nabla f(x_k)^T d_k = 0. \tag{1.30}$$

Indeed, from the boundedness of $\{d_k\}$ and the uniformly continuity of ∇f near $\{x_k\}$, we have

$$\lim_{k\to\infty} \frac{f(x_k + \rho^{m_k-1}d_k) - f(x_k) - \rho^{m_k-1}\langle\nabla f(x_k), d_k\rangle}{\rho^{m_{k-1}}} = 0. \tag{1.31}$$

Note that from the definition of m_k and the gradient inequality we obtain

$$f(x_k + \rho^{m_k-1}d_k) - f(x_k) - \rho^{m_k-1}\nabla f(x_k)^T d_k > \rho^{m_k-1}(\sigma - 1)\nabla f(x_k)^T d_k > 0.$$

Dividing the above inequality by ρ^{m_k-1} we easily deduce (1.30).

Consequently, we shown that regardless of the infimum value of the sequence $\{\tau_k\}$, (1.30) holds. There is no loss of generality to assume that (1.29). By (1.28) and the positive semidefinite of each M_k, the sequence of $\{\langle d_k, M_k d_k\rangle\}$ converges to zero. Furthermore, since $\{M_k\}$ is bounded, which implies that the eigenvalues of M_k are bounded above, it can be easily shown that the sequence $\{M_k d_k\}$ also converges to the zero vector.

Since $\{x_k - y_k\}$ and $\{\delta_k^{-1}\}$ are bounded, passing to limit (1.27), it leads to a contradiction as the left hand converges to a negative limit whereas the right hand side converges to zero. □

Chapter 2

Minimizing And Stationary Sequences In Nonsmooth Optimization

Let f be a bounded below, lower semicontinuous function from a Banach space into $\mathbb{R}\cup\{+\infty\}$. In this chapter, we study [14] the relation between minimizing and stationary sequence for the problem of minimizing f. The stationary condition is in terms of subdifferential, the function f is not convex or smooth on a Banach space.

2.1 Subdifferential

We consider lower semicontinuous extended real-valued function f defined on a Banach Space X whose dual space is denoted as X^*. Let

$$dom f \equiv \{x \in X \mid -\infty < f(x) < +\infty\}.$$

We are to study the problem:

$$\begin{aligned} &\text{Minimize} \quad f(x) \\ &\text{subject to} \quad x \in X \end{aligned}$$

Noticed that f is nonsmooth, nonconvex and X is a general Banach space.

Definition 2.1.1 *Given a class of normed vector space $\mathcal{X}$ and a class of function $\mathcal{F}(X)$ from X to $\mathbb{R} \cup \{+\infty\}$, for any $X \in \mathcal{X}$, $f \in \mathcal{F}(X)$ and $x \in X$, the subdifferential of f at x $\partial f(x)$ is defined as a subset of X^* which satisfied the following properties:*

(i) If $f(x) = \infty$, then $\partial f(x)$ is an empty set;

(ii) if x is a local minimizer of f, then $0 \in \partial f(x)$;

(iii) if f is convex, then $\partial f(x) = \{x^* \mid f(\cdot) \geq x^*(\cdot) + f(x) - \langle x^*, x\rangle \text{ on } X\}$;

(iv) if a function $g \in \mathcal{F}(X)$ coincide f on some neighborhood of x, then $\partial f(x) = \partial g(x)$;

(v) if $f(\cdot) = cg(A(\cdot) + b)$ where A is a linear map from X to Y such that $A(X) = Y$, $c \in \mathbb{R}_+$ and $b \in Y$, then $\partial f(x) = c\partial g(A(x) + b) \circ A$;

(vi) if $X = Y \times Z$, for $y \in Y$ and $z \in Z$ with $f(y, z) = g(y) + h(z)$, then $\partial f(y, z) = \partial g(y) \times \partial h(z)$.

Definition 2.1.2 *Given a Banach space X, $\mathcal{F}(X)$ a class of function from X to $\mathbb{R} \cup \{+\infty\}$ and a subdifferential ∂ on X and $\mathcal{F}(X)$. X is said to be reliable space or the subdifferential ∂ is reliable on X if for any lower semicontinuous function $f \in \mathcal{F}(X)$, any convex Lipschitzian function g on X and $x \in domf$ is a local minimizer of $f + g$, then for each $\epsilon > 0$,*

$$0 \in \partial f(u) + \partial g(v) + \epsilon \mathcal{B}^*,$$

for some $u, v \in \mathcal{B}(x, \epsilon)$ such that $|f(u) - f(x)| < \epsilon$.

Definition 2.1.3 *Given a set X, a class of function $\mathcal{F}(X)$ from X to $\mathbb{R} \cup \{+\infty\}$ and a subdifferential on X and $\mathcal{F}(X)$. A sequence $\{x_k\} \subseteq X$ is defined to be a stationary sequence if there exists a sequence $\{x_k^*\} \subseteq X^*$ such that*

$$x_k^* \in \partial f(x_k) \text{ for each } k \quad \text{and} \quad x_k^* \to 0 \text{ as } k \to \infty.$$

Definition 2.1.4 *A sequence $\{x_k\}$ is said to be minimizing for a function f on X if*

$$\{f(x_k)\} \to f_{inf} \equiv \inf_{x \in X} f(x).$$

2.2 Stationary and minimizing sequences

The following theorem told us that for a minimizing sequence, there always exists a nearby sequence which is stationary as well as minimizing.

Theorem 2.2.1 *Suppose that X is a Banach space and $\mathcal{F}(X)$ is a class of function from X to $\mathbb{R} \cup \{+\infty\}$. Let f is bounded below, lower semicontinuous on X and the subdifferential ∂ is reliable on X. If $\{x_k\}$ is a minimizing sequence for f, then there exists sequences $\{y_k\}$ and $\{y_k^*\} \subseteq X^*$ satisfied $y_k^* \in \partial f(x_k)$ for each k with the following properties,*

$$(a) \quad \lim_{k \to \infty} \|x_k - y_k\| = 0; \tag{2.1}$$

$$(b) \quad \lim_{k \to \infty} f(y_k) = f_{inf};$$

$$(c) \quad \lim_{k \to \infty} y_k^* = 0. \tag{2.2}$$

Proof As $\{x_k\}$ is a minimizing sequence, there exists a sequence $\{\epsilon_k\} \downarrow 0$ satisfied

$$f(x_k) \leq f_{inf} + \epsilon_k \quad \forall k.$$

Then, we apply Ekeland's variational principle in Lemma (3.2.1) for $\lambda_k = \epsilon_k^{1/2}$, there exists $z_k \in X$ such that

(i) $\|x_k - z_k\| \leq \lambda_k$;

(ii) $f(z_k) \leq f(x_k)$;

(iii) $f(z_k) = \inf_{x \in X}\{f(x) + \lambda_k\|x - z_k\|\}$.

It follows from (iii), we have

$$0 \in \partial(f + \lambda_k\| \cdot - z_k\|)(z_k). \tag{2.3}$$

Since ∂ is reliable on X, for $\lambda_k\| \cdot - z_k\|$ is a convex Lipschitzian function on X, (2.3) becomes

$$0 \in \partial f(y_k) + \lambda_k \mathcal{B}^*$$

for some $y_k \in X$ with $\|y_k - z_k\| \leq \lambda_k$ and $|f(y_k) - f(z_k)| \leq \lambda_k$. We do this for each k, then one has a sequence $\{y_k\}$ with the properties:

(1) $\|y_k - z_k\| \leq \lambda_k$;

(2) $|f(y_k) - f(z_k)| \leq \lambda_k$;

(3) $y_k^* \in \partial f(y_k)$ with $\|y_k^*\| \leq \lambda_k$.

for each k. Noted that $\lim_{k\to\infty} \lambda_k = \lim_{k\to\infty} \epsilon_k^{1/2} = 0$. Then, $\|x_k - y_k\| \leq \|x_k - z_k\| + \|z_k - y_k\| \leq 2\lambda_k$ by (i) and (1) and so $\lim_{k\to\infty} \|x_k - y_k\| = 0$. Also, by (2) and (iii) $f(y_k)$ converges to f_{inf}. Moreover, $\lim_{k\to\infty} y_k^* = 0$ which follows by (3). □

Definition 2.2.1 *Suppose X, Y are metric spaces and F is a multifunction from X to Y. F is said to be uniformly upper semicontinuous near a sequence $\{x_k\}$ if for any $\epsilon > 0$, there exists a $\delta > 0$ and $N \in \mathbb{N}$ such that*

$$F(x) \subseteq F(x_n) + \mathcal{B}(0, \epsilon)$$

for all $n \geq N$ and $\|x - x_n\| < \delta$.

Noticed that a subdifferential ∂ on a set X in which introduced a class of extended real-valued function $\mathcal{F}(X)$ is a multifunction. With the assumption that a subdifferential is uniformly upper continuous near a minimizing sequence, the minimizing sequence is then stationary. The fact is illustrated by the following theorem.

Theorem 2.2.2 *Let X be a Banach space which is reliable for the subdifferential ∂ and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a bounded below lower semicontinuous function. Also, suppose that $\{x_k\}$ is a minimizing sequence for f and the subdifferential ∂f is uniformly upper semicontinuous near a sequence $\{x_k\}$. Then it is a stationary sequence such that there exists $\{x_k^*\} \subseteq X^*$ with $x_k^* \to 0$.*

Proof Let $\epsilon_k \downarrow 0$. By the uniformly upper semicontinuity of ∂f near $\{x_k\}$, there exist $\delta_k > 0$ and $N_k \in \mathbb{N}$ such that

$$\partial f(x) \subseteq \partial f(x_n) + \mathcal{B}(0, \epsilon_k)$$

for all $n > N_k$ and $\|x - x_n\| < \delta_k$. Then we arrange $\delta_k \downarrow 0$. By Theorem (2.2.1), there exist $\{y_k\}$ with $\lim_{k\to\infty} \|x_k - y_k\| = 0$. Then we can take $n_k \geq N_k$ such that

$$\|x_n - y_n\| < \delta_k \quad \forall n > n_k$$

thus,

$$y_n^* \in \partial f(y_n) \subseteq \partial f(x_n) + \mathcal{B}(0, \epsilon_k) \quad \forall n > n_k.$$

We can certainly arrange n_k to strictly increase as k increase. Hence, for $n_k < n \leq n_{k+1}$ we pick $x_n^* \in \partial f(x_n)$ such that

$$\|x_n^* - y_n^*\| \leq \epsilon_k.$$

Therefore, one has

$$\{\|x_n^* - y_n^*\|\} \to 0.$$

It follows from (2.2) that $y_k^* \to 0$ and so $\{x_n^*\} \to 0$. □

2.3 C-convex and BC-convex function

Definition 2.3.1 *A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is critically convex or in short C-convex if for any stationary sequences $\{x_k\}$ and $\{y_k\}$ with $x_k \neq y_k$ for all k, we have*

$$\lim_{k\to\infty} \frac{|f(x_k) - f(y_k)|}{\|x_k - y_k\|} = 0. \tag{2.4}$$

Lemma 2.3.1 *If f is a convex function, then it is C-convex.*

Proof Let $\{x_k\}$ and $\{y_k\}$ are stationary sequences for f. Then there exist sequences $\{x_k^*\} \subseteq X^*$ and $\{y_k^*\} \subseteq X^*$ such that $x_k^* \to 0$ and $y_k^* \to 0$ as $k \to \infty$. Since f is convex, it follows from the definition of subdifferential Definition (2.1.1) (iii) that

$$x_k^*(x_k - y_k) \geq f(x_k) - f(y_k) \geq y_k^*(x_k - y_k).$$

Hence, for $\{x_k^*\}$ and $\{y_k\}$ are bounded, one has

$$|f(x_k) - f(y_k)| \leq (\|x_k^*\| \vee \|y_k^*\|)\|x_k - y_k\|.$$

Thus, $\lim_{k\to\infty} |f(x_k) - f(y_k)| \, \|x_k - y_k\|^{-1} = 0.$ □

Lemma 2.3.2 *Let f be a differentiable function on X. Suppose that for any stationary sequences $\{x_k\}$ and $\{y_k\}$ of f, any sequence $\{z_k\}$ such that $z_k \in [x_k, y_k]$ for each k is stationary. Then f is C-convex.*

Proof Since f is differentiable, by mean value theorem there exists a sequence $\{z_k\}$ with $z_k \in [x_k, y_k]$ for each k such that

$$|f(x_k) - f(y_k)| = |f'(z_k)(x_k - y_k)| \leq \|f'(z_k)\| \, \|x_k - y_k\|$$

for each k. Since $\{f'(z_k)\} \to 0$ by the assumption,

$$\lim_{k\to\infty} \frac{|f(x_k) - f(y_k)|}{\|x_k - y_k\|} = 0.$$

□

Lemma 2.3.3 *Any quadratic function is a C-convex function.*

Proof For a quadratic function, it is differentiable and its derivative is an affine function. So the assumption in Lemma (2.3.2) holds and thus a quadratic function is C-convex. □

Definition 2.3.2 *A function f is said to be boundedly critically convex or in short BC-convex if for any pair of bounded stationary sequences $\{x_k\}$ and $\{y_k\}$ with $x_k \neq y_k$ for all k, the following equality holds:*

$$\lim_{k\to\infty} \frac{|f(x_k) - f(y_k)|}{\|x_k - y_k\|} = 0. \tag{2.5}$$

Noted that a class of C-convex function on X is a subset of a class of BC-convex function on X as if any pair of stationary sequences satisfied equation (2.4), those bounded pair do has such property.

2.4 Minimizing sequences in terms of sublevel sets

Let $\lambda \in \mathbb{R}$ and $L(\lambda)$ be the level set of function f for a set X, such that

$$L(\lambda) \equiv \{x \in X \mid f(x) \leq \lambda\}.$$

As mentioned before, we assume f is a lower semicontinuous function and thus the level set is closed for each $\lambda \in \mathbb{R}$.

Theorem 2.4.1 *Let X be a reliable Banach space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a C-convex and bounded below lower semicontinuous function. Also, we assumed that a sequence $\{x_k\}$ is stationary. Then $\{x_k\}$ is minimizing if and only if there exists a sequence $\{\lambda_k\} \subseteq \mathbb{R}$ in which $\lambda_k \downarrow f_{inf}$ satisfied the inequality*

$$\sup_{k\in\mathbb{N}} dist(x_k\,,\, L(\lambda_k)) < +\infty. \tag{2.6}$$

Proof ($\Rightarrow$) Let $\lambda_k = f(x_k)$ for each k. Then $\{\lambda_k\} \downarrow f_{inf}$ for $\{x_k\}$ is minimizing and the inequality (2.6) holds as supremum equals to zero.

($\Leftarrow$) Let $\{\lambda_k\} \downarrow f_{inf}$ and $\{x_k\}$ is a stationary sequence such that there exists a sequence $\{x_k^*\} \to 0$ and $x_k^* \in \partial f(x_k)$ for each k. Then for each k we take $y_k \in L(\lambda_k)$ satisfying

$$\|x_k - y_k\| < dist(x_k, L(\lambda_k)) + \frac{1}{k}. \tag{2.7}$$

Noted that $f(y_k) \leq \lambda_k$, $y_k \in X$ for each k and $\{\lambda_k\} \downarrow f_{inf}$, so $\{f(y_k)\} \to f_{inf}$ that is $\{y_k\}$ is minimizing. By Theorem (2.2.1), there exists a nearby sequence $\{z_k\} \subseteq X$ with the following properties:

(i) $\lim_{k\to\infty} \|y_k - z_k\| = 0$;

(ii) $\{z_k\}$ is stationary;

(iii) $\{z_k\}$ is minimizing.

Let $K \subseteq \mathbb{N}$ such that $x_k \neq z_k$ for all $k \in K$. As f is C-convex, there exists $\{\delta_k\}_{k\in K} \subseteq \mathbb{R}_+$ which is decreasing such that

$$|f(x_k) - f(z_k)| \leq \delta_k \|x_k - z_k\| \quad \text{for each}\, k \in K.$$

For each $k \in \mathbb{N}\backslash K$ such that $x_k = z_k$, one has $f(x_k) = f(z_k)$. Consequently, we can take $\{\delta_k\}_{k\in\mathbb{N}} \downarrow 0$ such that

$$|f(x_k) - f(z_k)| \leq \delta_k(\|x_k - y_k\| + \|y_k - z_k\|) \quad \forall k \in \mathbb{N}. \tag{2.8}$$

Let $k \to \infty$, if follows from (2.8), (i) and (2.7), $\{x_k\}$ is also minimizing as $\{z_k\}$ is. □

Definition 2.4.1 *The ϵ-subdifferential (or ϵ-approximate subdifferential) of a function $f : X \to \mathbb{R} \cup \{+\infty\}$ at x is defined as*

$$\partial_\epsilon f(x) \equiv \{x^* \in X^* \mid x^*(y - x) \leq f(y) - f(x) + \epsilon \quad \forall y \in X\}.$$

Definition 2.4.2 *A sequence $\{x_k\}$ is said to be a ϵ-stationary sequence (or ϵ-approximate stationary sequence) if there exists a sequence $\{\epsilon_k\}$ with $\epsilon_k > 0$ and $\{x_k^*\}$ with $x_k^* \in \partial_\epsilon f(x_k)$ such that $x_k^* \to 0$.*

Lemma 2.4.1 *Suppose X is a Banach space and $f : X \to \mathbb{R} \cup \{+\infty\}$. If there exists a sequences $\{\epsilon_k\}$ and $\{x_k\}$ is ϵ-stationary, then $\{x_k\}$ is minimizing if and only if for any $\lambda > f_{inf}$, we have*

$$\sup_{n \in \mathbb{N}} dist(x_k, L(\lambda)) < +\infty. \tag{2.9}$$

Proof ($\Rightarrow$) Let $\lambda > f_{inf}$. Since $\{x_k\}$ is minimizing, $dist(x_k, L(\lambda)) = 0$ for all sufficient large k. So (2.9) holds.

($\Leftarrow$) Suppose on the contrary that there exist scalars γ, μ such that $\gamma > \mu > f_{inf}$ and a subsequence $\{x_{n_k}\}$ that $f(x_{n_k}) > \gamma$ for each k. Then we take a sequence $\{y_k\} \subseteq L(\mu)$ such that for each k,

$$\|x_k - y_k\| < dist(x_{n_k}, L(\mu)) + \frac{1}{k}. \tag{2.10}$$

It follows from the definition of $\partial_\epsilon f$ that

$$\begin{aligned} \|x_{n_k}^*\| \, \|x_{n_k} - y_k\| &\geq x_{n_k}^*(x_{n_k} - y_k) \\ &\geq f(x_{n_k}^*) - f(y_k) - \epsilon_{n_k} \\ &> \gamma - \mu - \epsilon_{n_k}. \end{aligned}$$

for all k. However, contradiction on the last inequality occurs as $x_{n_k}^* \to 0$, $\epsilon_{n_k} \to 0$ and $\|x_{n_k} - y_k\|$ is bounded by (2.10). $\square$

Theorem 2.4.2 *Suppose X is a Banach space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function. Let $\{x_k\}$ be an ϵ-stationary such that there exist $\{x_k^*\}$ and $\{\epsilon_k\}$ satisfied $\epsilon_k \geq 0$ for each k, $\{\epsilon_k\} \to 0$, $x_k^* \in \partial f(x_k)$ for each k and $\{x_k^*\} \to 0$. One has the conclusion that*

(i) the sequence $\{x_k\}$ is minimizing

if and only if any one of the following condition is satisfied:

(ii) for any $\lambda > f_{inf}$*,*

$$\sup_{k\in\mathbb{N}} dist(x_k, L(\lambda)) < +\infty;$$

(iii) there exists a sequence $\{\lambda_k\} \downarrow f_{inf}$ *such that*

$$\sup_{k\in\mathbb{N}} dist(x_k, L(\lambda_k)) < +\infty;$$

(iv) for any $\lambda > f_{inf}$*, there exists a scalar* $c > 0$ *such that for all* $k \in \mathbb{N}$

$$dist(x_k, L(\lambda)) \leq c(f(x_k) - \lambda)_+; \tag{2.11}$$

(v) for any $\lambda > f_{inf}$*, there exists a function* $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ *with* $\limsup_{r\to\infty} r^{-1}\mu(r) < \infty$ *and for all* $k \in \mathbb{N}$*,*

$$dist(x_k, L(\lambda)) \leq \mu(f(x_k) - \lambda). \tag{2.12}$$

Proof (ii) $\Leftrightarrow$ (i) This follows from Lemma (2.4.1).
(i) $\Rightarrow$ (iii) Let $\{x_k\}$ is minimizing, take $\lambda_k = f(x_k)$ for each k and $dist(x_k, L(\lambda_k)) = 0$ for all $k \in \mathbb{N}$. Thus (iii) holds.
(iii) $\Rightarrow$ (ii) For any $\lambda > f_{inf}$, as $\{\lambda_k\} \downarrow f_{inf}$ there exists $N \in \mathbb{N}$ for all $n \geq N$ such that $\lambda > \lambda_n \geq f_{inf}$ and

$$dist(x_n, L(\lambda)) \leq dist(x_n, L(\lambda_n)) \quad \text{for each } n \geq N.$$

Thus,

$$\sup_{n\geq N} dist(x_n, L(\lambda)) \leq \sup_{n\geq N} dist(x_n, L(\lambda_n)) < +\infty.$$

Also, $\sup_{n\in\mathbb{N}, n\leq N} dist(x_k, L(\lambda)) < +\infty$. Thus, $\sup_{n\in\mathbb{N}} dist(x_k, L(\lambda)) < +\infty$.
(i) $\Rightarrow$ (iv) Let $\lambda > f_{inf}$. As $\{x_k\}$ is minimizing, $dist(x_k, L(\lambda)) = 0$ for sufficient large k. So we can take a scalar $c > 0$ satisfied (2.11).

(iv) $\Rightarrow$ (v) By letting $\mu(r) = c\max(r,0)$, then $\limsup_{r\to\infty} r^{-1}\mu(r) < \infty$. Thus, (2.12) holds.

(v) $\Rightarrow$ (ii) Suppose on the contrary that there exists $\lambda_0 > f_{inf}$ such that

$$\sup_{k\in\mathbb{N}} dist(x_k, L(\lambda_0)) = +\infty.$$

Then by (2.12), we have $\sup_{k\in\mathbb{N}} \mu(f(x_k) - \lambda_0) = +\infty$. Since the function satisfies $\limsup_{r\to\infty} r^{-1}\mu(r) < \infty$, one has

$$\lim_{k\to\infty} f(x_k) = +\infty.$$

Take a sequence $\{y_k\} \subseteq L(\lambda_0)$ such that for each k,

$$\|x_k - y_k\| \le dist(x_k, L(\lambda_0)) + \frac{1}{k}.$$

Hence, by the definition of ∂_{ϵ_k}, it follows that

$$\begin{aligned}
\|x_k^*\|\, \mu(f(x_k) - \lambda_0) &\ge \|x_k^*\|\,(\|x_k - y_k\| - \frac{1}{k}) \\
&\ge x_k^*(x_k - y_k) - \frac{1}{k}\|x_k^*\| \\
&\ge f(x_k) - f(y_k) - \epsilon_k - \frac{1}{k}\|x_k^*\| \\
&\ge f(x_k) - \lambda_0 - \epsilon_k - \frac{1}{k}\|x_k^*\|.
\end{aligned}$$

It then let k large and divides both sides by $f(x_k) - \lambda_0 \neq 0$, one has

$$\|x_k^*\| \limsup_{n\in\mathbb{N}} (f(x_k) - \lambda_0)^{-1}\mu(f(x_k) - \lambda_0) \ge 1 - \frac{k\epsilon_k + \|x_k^*\|}{k(f(x_k) - \lambda_0)}.$$

For $\{x_k^*\} \to 0$ and the property of μ, contradiction on the above inequality occurs as $k \to \infty$ because the left side converges to zero while the right hand side converges to 1. □

2.5 Critical function

Definition 2.5.1 *Given a Banach space X, a class of function $\mathcal{F}(X)$ from X to $\mathbb{R}\cup\{+\infty\}$ and the subdifferential ∂, a function $f : X \to \mathbb{R}\cup\{+\infty\}$ is said to be critical if for any stationary sequence $\{x_k\}$ of f, the sequence $\{f(x_k)\}$ converges.*

Let Z be a set containing all critical points for f, that is for all $z \in Z$, $\partial f(z) = 0$.

Proposition 2.5.1 *For a critical function f, f is constant on Z.*

Proof Let $z, y \in Z$. Then we take a sequence $\{x_k\}$ in such a way that $x_{2p} = z$ and $x_{2p-1} = y$ for all $p \in \mathbb{N}$. Thus, $\{x_k\}$ is a stationary sequence. It then follows from the definition of a critical function, we have $f(y)$ and $f(z)$ must coincide with the limit of $\{f(x_k)\}$. □

Proposition 2.5.2 *Suppose a function f is C-convex, has at least a critical point z and all its stationary sequences are bounded. Then f is critical.*

Proof Let $\{x_k\}_{k\in\mathbb{N}}$ be a stationary sequence of f. We want to show that for any subsequence $\{x_i\}_{i\in I}$ of $\{x_k\}_{k\in\mathbb{N}}$ contains a subsequence $\{x_{i_j}\}_{j\in J}$ where $J \subseteq I$ such that $\{f(x_{i_j})\}_{j\in J}$ converges to $f(z)$. This means $\{f(x_k)\}$ converges to $f(z)$ and thus f is critical. If $x_i = z$ for infinitely many $i \in I$, then the claim follows. Otherwise, let $M \subseteq I$ be such that for all $m \in M$, $x_m \neq z$. Take a stationary sequence $\{y_m\}$ for which $y_m = z$ for all $m \in M$. Since the subsequence $\{x_m\}$ is bounded, we have

$$\|x_m - z\| \leq r \quad \forall m \in M$$

for some scalar $r > 0$. Also, for f is C-convex, there exists a sequence $\{\epsilon_m\} \downarrow 0$ such that

$$|f(x_m) - f(y_m)| = |f(x_m) - f(z)| \leq \epsilon_m r$$

It follows that the subsequence $\{f(x_m)\}$ converges to $f(z)$. Hence, $\{f(x_k)\}$ converges to $f(z)$ too. □

Instead of existing a critical point, similar result holds if there exists a stationary sequence whose functional values are bounded.

Proposition 2.5.3 *Suppose that a function f is C-convex, there exists a stationary sequence $\{z_k\}$ such that $\{f(z_k)\}$ is bounded and all the stationary sequences for f are bounded. Then f is critical.*

Proof The proof follows similar to that of Proposition (2.5.2). Let $\{x_k\}_{k\in\mathbb{N}}$ be a stationary sequence for f. Since $\{f(z_k)\}$ is bounded, there exists a subsequence $\{z_{k_n}\}$ such that $\{f(z_{k_n})\}$ converges. We claim that for any subsequence $\{x_i\}_{i\in I}$ of $\{x_k\}$, one can find a subsequence $\{x_{i_m}\}_{m\in M}$ of $\{x_i\}$ such that $\{f(x_{i_m})\}$ converges. This means $\{f(x_k)\}$ converges. If there exists a subsequence $\{x_{i_m}\}$ of $\{x_i\}$ such that $x_{i_m} = z_{k_m}$ for infinitely many m, then $\{f(x_{i_m})\}$ converges and the claim follows. Otherwise, let $M \subseteq I$ such that for all $m \in M \subseteq I$, $x_{i_m} \neq z_{k_m}$. As $\{x_{i_m}\}$ and $\{z_{k_m}\}$ are bounded, we have

$$\|x_{i_m} - z_{k_m}\| \leq r \quad \forall m \in M$$

for some scalar $r > 0$. Also, f is C-convex, one has

$$|f(x_{i_m}) - f(z_{k_m})| \leq \epsilon_k r$$

for some sequence $\{\epsilon_k\} \downarrow 0$. It follows that $\{f(x_{i_m})\}$ converges in $\mathbb{R}$ as $\{f(z_{k_m})\}$ converges. We have shown that for any stationary sequence $\{x_k\}$ and its subsequence $\{x_i\}_{i\in I}$, there exists a subsequence $\{x_{i_m}\}$ of $\{x_i\}_{i\in I}$ such that $\{f(x_{i_m})\}$ converges. That is $\{f(x_k)\}$ converges. □

Theorem 2.5.1 *Suppose that f is a bounded below lower semicontinuous function from a Banach space X to $\mathbb{R}\cup\{+\infty\}$. Also, let ∂ be a reliable subdifferential. Then, any stationary sequence for f is minimizing if f is critical.*

Proof Let $\{x_k\}$ be a stationary sequence. Since the function f is bounded below, a minimizing sequence for f always exists. Then by Theorem (2.1.2), there exists a sequence $\{y_k\} \subseteq X$ which is minimizing as well as stationary. Let a sequence $\{z_k\}$ be $z_{2k} = x_k$ and $z_{2k-1} = y_k$, then $\{z_k\}$ is a stationary sequence. By the assumption that f is critical, $\lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} f(y_k) = f_{inf}$. That is the sequence $\{x_k\}$ is minimizing. □

Chapter 3

Optimization Conditions

In this chapter, we discuss [15] the second-order optimization conditions for locally Lipschitz real-valued function f on a Banach Space X. In general, f is not necessarily smooth. Firstly, we will deal with unconstrainted problems. Then for constrainted problems some functions G and L will be introduced and optimization conditions are expressed in terms of G and L.

3.1 Introduction

First of all, let f be a locally Lipschitzian real-valued function on a Banach space X. We use $f^0(x;u)$ to denote the Clarke's directional derivative at x in the direction u and $\partial f(x)$ the Clarke's subdifferential at x:

$$f^0(x;u) = \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{1}{t}\{f(y+tu) - f(y)\} \quad \text{for}\, x, u \in X$$

$$\partial f(x) = \{x^* \in X^* \mid x^*(\cdot) \leq f^0(x;\cdot) \quad \text{on}\, X\}.$$

Recalled that the upper and lower Dini-directional derivative at $x \in X$ in the direction u is

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

and

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

Definition 3.1.1 *For f is a locally Lipschitz function, the upper and lower Dini-directional derivatives at $x \in X$ in the direction $u \in X$ are defined by*

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

and

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

Definition 3.1.2 *Suppose that a sequence $\{x_k\}$ in X converges to x. We defined that x_k converges to x in the direction $u \in X$ with $\|u\| \neq 0$, denoted by $\{x_k\} \to_u x$, if $x_k \neq x$ for each k and the sequence $((x_k - x)/\|x_k - x\|)$ converges to $u/\|u\|$.*

Definition 3.1.3 *Let u be a nonzero vector in X. The subset $\partial_u f(x)$ in the dual space X^* of X is defined by*

$$\begin{aligned}\partial_u f(x) =& \{x^* \in X^* \mid \text{there exists a sequence}\,(x_k) \subseteq X \text{ and } x_k{}^* \in \partial f(x_k)\,\text{for each } k\\ &\quad \text{such that}\,(x_k) \to_u x \text{ and } x_k^* \to x^* \text{ in norm}\}\\ =& \{x^* \in X^* \mid \text{there exists a sequence}\,(x_k) \subseteq X \text{ and}\,(x_k^*) \subseteq X^* \text{ such that}\\ &\quad x_k \to_u x,\ x_k^*(\cdot) \le f^0(x;\cdot)\ \text{on}\,X \text{ with } x_k^* \to x^*\}.\end{aligned}$$

With the above definitions, we are going to define the second-order directional derivative $f''_{-}(x, x^*, u)$ at x and x^* in the direction u:

Definition 3.1.4 *Let u be a nonzero vector in X. Suppose that $x \in X$ and $x^* \in \partial_u f(x)$, $f''_{-}(x, x^*, u)$ is the infimum of all extended real numbers*

$$\liminf_{k\to\infty} \frac{1}{t_k{}^2}\{f(x_k) - f(x) - x^*(x_k - x)\},$$

taken over all triples of sequences $\{x_k\}$, $\{x_k{}^*\}$, and $\{t_k\}$ for which

1. $t_k > 0$ *for each* k *and* $\{t_k\}$ *converges to* 0,

2. $\{x_k\}$ *converges to* x *and* $\{(x_k - x)/t_k\}$ *converges to* u,

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

Similarly, $f''_+(x, x^*, u)$ is the supremum of all extended real 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 the above conditions $1, 2$ and 3 all hold.

Remark In 2, $(x_k - x)/t_k$ converging to u implies x_k converging to x in the direction u as

$$(x_k - x)/\|x_k - x\| = [(x_k - x)/t_k][t_k/\|x_k - x\|] \to \frac{u}{\|u\|}.$$

Since

$$\begin{aligned}
&\liminf_{k\to\infty} \frac{1}{t_k^2}\{\, f(x_k) - f(x) - x^*(x_k - x)\,\} \\
&= \liminf_{k\to\infty} \frac{1}{\|x_k - x\|^2}\{f(x_k) - f(x) - x^*(x_k - x)\}\frac{\|x_k - x\|^2}{t_k^2} \\
&= \|u\|^2 \liminf_{k\to\infty} \frac{1}{\|x_k - x\|^2}\{f(x_k) - f(x) - x^*(x_k - x)\},
\end{aligned}$$

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

$$\|u\|^2 \liminf_{k\to\infty} \{f(x_k) - f(x) - x^*(x_k - x)\}/\|x_k - x\|^2,$$

taken over the set of all sequences (x_k) and (x_k^*) with the properties

1. (x_k) converges to x in the direction u.

2. there exists a sequence $\{x_k^*\} \subseteq X^*$ such that $x_k^* \in \partial f(x_k)$ converging to x^*.

3.2 Second-order necessary and sufficient conditions without constraint

In this section, we let W be an open subset of X. We are to study the problem of minimizing f over W. Before stating the necessary conditions for unconstraint problems, we have the following Ekeland' variational principle.

Lemma 3.2.1 *[10, Proposition 7.5.1] If u is a vector in X satisfying*

$$f(u) \leq f_{inf} + \epsilon$$

for some $\epsilon > 0$, then for each $\lambda > 0$ there exists a vector v in X such that

(i) $f(v) \leq f(u)$;

(ii) $\|u - v\| \leq \lambda$;

(iii) For all $w \in X$ with $w \neq v$, one has $f(w) + (\epsilon/\lambda)\|w - v\| > f(v)$.

Lemma 3.2.2 *[10, Proposition 2.3.2] Suppose f is locally Lipschitz and attains a local minimum or maximum at x, then $0 \in \partial f(x)$.*

Lemma 3.2.3 *[10, Proposition 2.3.3] Suppose $\{f_i \mid i = 1, \cdots, k\}$ is a family of locally Lipschitz functions, then*

$$\partial(\sum_{i=1}^{k} f_i)(x) \subseteq \sum_{i=1}^{k} \partial f_i(x).$$

Lemma 3.2.4 *[10, Proposition 2.3.7] Let x and y be vectors in X, and suppose that f is Lipschitz on an open set containing the line segment $[x, y]$. Then there exists a vector in (x, y) such that*

$$f(y) - f(x) \in \langle \partial f(u), y - x \rangle.$$

Lemma 3.2.5 *[10, Proposition 2.3.1] Let f be a locally Lipschitz function. For any scalar s, one has*

$$\partial(sf)(x) = s\partial f(x).$$

Lemma 3.2.6 *Suppose that $f(x) \geq f(\bar{x})$ for all $x \in \mathcal{B}[\bar{x}, \delta]$. Let $0 \neq u \in X$, $t > 0$, $\alpha > 1$ and $0 < \epsilon < (\alpha\|u\|)^2$ such that*

$$f(\bar{x} + tu) - f(\bar{x}) \leq t\epsilon \quad \text{and} \quad t(\|u\| + \epsilon^{1/2}) < \delta. \tag{3.1}$$

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

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

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

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

Proof By the assumption of the lemma,

$$f(\bar{x} + tu) \leq f(\bar{x}) + t\epsilon \leq \inf_{x \in \mathcal{B}} f(x) + t\epsilon$$

where $\mathcal{B}$ is the closed ball with center $\bar{x} + tu$ and radius $t\epsilon^{1/2}$. Note that $\mathcal{B} \subseteq \mathcal{B}(x, \delta)$ thanks to the second inequality in (3.1). Applied the Ekeland variational principle in Lemma (3.2.1) with $\lambda = t\epsilon^{1/2}\alpha^{-1}$, there exists $z \in \mathcal{B}$ satisfies (i), (ii) and

$$(iv) \quad f(z) \leq f(y) + (\alpha\epsilon^{1/2})\|z - y\| \quad \text{for all}\, y \in \mathcal{B}.$$

Since (i) and $\epsilon < (\alpha\|u\|)^2$, $z \neq \bar{x}$ and z is an interior point of $\mathcal{B}$ for $\alpha > 1$. Also, it follows from (iv) and calculus for subdifferential Lemma (3.2.2) and Lemma (3.2.3) that $0 \in \partial f(z) + \alpha\epsilon^{1/2}\mathcal{B}_1^*$ where $\mathcal{B}_1^*$ denotes the unit ball in X^*. Thus $\|z^*\| \leq \alpha\epsilon^{1/2}$ for some $z^* \in \partial f(z)$. □

Lemma 3.2.7 *Let $\epsilon > 0$ and $\alpha > 1, \sqrt{\epsilon}$. Let $\bar{x}$ be a local minimum of f, and let $u \in X$ be of norm 1 such that $D_-f(\bar{x};u) < \epsilon$. Then for an arbitrarily small $t > 0$, there exists $z \in X \setminus \{\bar{x}\}$ and $z^* \in \partial f(z)$ such that*

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

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

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

Proof Since $\bar{x}$ is a local minimum of f, $f(x) \geq f(\bar{x})$ for all $x \in \mathcal{B}[\bar{x}, \delta]$ for some $\delta > 0$. With $u \in X$ and $\epsilon > 0$, there exists $t_0 > 0$ such that

$$t_0(\|u\| + \epsilon^{1/2}) < \delta.$$

Also, for this t_0 and $D_-f(\bar{x};u) = \liminf_{t\downarrow 0} \frac{1}{t}\{f(\bar{x} + tu) - f(\bar{x})\} < \epsilon$, there exists t with $0 < t < t_0$ such that

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

That is, (3.1) in Lemma (3.2.6) is satisfied. Together with $0 < \epsilon < (\alpha\|u\|)^2$, the result (i), (ii) and (iii) all hold by Lemma (3.2.6). $\square$

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

Theorem 3.2.1 *Suppose that $\bar{x}$ is a local minimum point for f. Let $u \in X$ be of norm 1 such that $D_-f(\bar{x};u) = 0$. Then $0 \in \partial_u f(\bar{x})$ and $f''_-(\bar{x}, 0, u) \geq 0$.*

Proof Let $\alpha = 2$ and $\epsilon \in (0,1)$. Since $D_-f(\bar{x};u) < \epsilon$, by Lemma (3.2.7) we can take a sequence $(\epsilon_k) \downarrow 0$ and then get a sequence $(t_k) \downarrow 0$ to obtain (z_k), (z_k^*) with $z_k^* \in \partial f(z_k)$ for each k satisfying (i), (ii) and (iii) in Lemma(3.2.6). In particular (i) becomes

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

Dividing by t_k, it follows that

$$([z_k - \bar{x}]/t_k) \to u$$

showing that $(z_k) \to \bar{x}$ in the direction u. Since $z_k^* \in \partial f(z_k)$ and $\|z_k^*\| \leq 2\epsilon_k^{1/2}$ by (iii), we have $0 \in \partial_u f(\bar{x})$. Also, $f''_-(\bar{x}, 0, u) \geq 0$ by Definition (3.1.4) because $\bar{x}$ is a local minimum point. □

Definition 3.2.1 *Let $X = \mathbb{R}^n$. We define the sets $D^*(x, f)$ and $D^\sharp(x, f)$ in $\mathbb{R}^n$ by*

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

and

$$D^\sharp(x,f) = \{u \in \mathbb{R}^n \mid v \cdot u \leq 0 \quad \text{for all } v \in \partial_u f(x)\}.$$

By putting $w = u$ above, it is easily seen that $D^*(x, f) \subseteq D^\sharp(x, f)$.

Lemma 3.2.8 *(i) For any $x, u \in \mathbb{R}^n$, there exist w^+ and w_+ in $\partial_u f(x)$ such that $\langle w^+, u\rangle = D_+ f(x; u)$ and $\langle w_+, u\rangle = D_- f(x; u)$.*

(ii) $D^*(x, f) \subseteq D^\sharp(x, f) \subseteq \{u \in \mathbb{R}^n \mid D_- f(x; u) \leq 0\}$.

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

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

Proof By Lebourg's mean value theorem and Lemma (3.2.4) for any $x, u \in \mathbb{R}^n$ and $t > 0$, we can find $a_t \in (0, t)$ and $w_t \in \partial f(x + a_t u)$ such that

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

Taking the upper limits on both sides, it follows form the definition of D_+ that

$$D_+ f(x; u) = \limsup_{t \downarrow 0} \langle w_t, u\rangle.$$

Pick a sequence $(t_n) \downarrow 0$ such that $D_+f(x;u) = \lim_{n\to\infty}\langle w_{t_n}, u\rangle$. As the multifunction $x \to \partial f(x)$ is closed and locally takes values in a compact set [10], we can find a subsequence $(t_{n_k}) \downarrow 0$ with $\lim w_{t_n} = w^+ \in \partial f(x)$ and

$$D_+f(x;u) \;=\; \lim_{k\to\infty}\langle w_{t_{n_k}}, u\rangle \;=\; \langle w^+, u\rangle.$$

Also, $w^+ \in \partial_u f(x)$ as $(x + a_{t_{n_k}})$ converges to x in the direction u.
Similarly, one can show that $D_-f(x;u) = \langle w_+, u\rangle$ for some $w_+ \in \partial_u f(x)$. Therefore, (i) holds.

Let $u \in D^*(x,f)$. By (i) there exists $w^+ \in \partial_u f(x)$ such that $\langle w^+, u\rangle = D_+f(x;u)$. From the definition of $D^*(x,f)$, $\langle w^+, u\rangle \le 0$. Thus, $D^+f(s,u) \le 0$. That is, (ii) is also true.

Since $\bar{x}$ is a local minimum of f, $D_-f(\bar{x},v) \ge 0$ for all $v \in \mathbb{R}^n$. From (ii), one has $D_-f(\bar{x};u) \le 0$ for all $u \in D^\sharp(\bar{x},f) \supseteq D^*(\bar{x},f)$. Therefore, $D_-f(\bar{x};u) = 0$ for all $u \in D^\sharp(\bar{x},f) \supseteq D^*(\bar{x},f)$. Thus, (iii) follows. □

Corollary 3.2.1 *Suppose $\bar{x}$ is a local minimum point of locally Lipschitz function f over $W \subseteq \mathbb{R}^n$. If $u \in D^*(\bar{x},f)$ with norm 1, then $0 \in \partial_u f(\bar{x})$ and $f''_-(\bar{x},0,u) \ge 0$.*

Proof If $\bar{x}$ is a local minimum, it follows Lemma (3.2.8) (iii) that $D_-f(\bar{x};u) = 0$. Again, by Theorem (3.2.6), $0 \in \partial_u f(\bar{x})$ and $f''_-(\bar{x},0,u) \ge 0$. □

Definition 3.2.2 *Let S be an subset of X and $x \in S$, the contingent cone of S at x is defined as*

$$K_S(x) \;=\; \{u \mid \exists (x_k) \subseteq S \text{ such that } x_k \ne x \text{ for each } k \text{ and } x_k \to_u x\}.$$

Lemma 3.2.9 *Let S be an subset of X and $\bar{x} \in S$. Suppose f is locally Lipschitz defined on S into $\mathbb{R}$ and $D_-f(\bar{x},u) \ge 0$ for all $u \in K_S(\bar{x})$. Then we* *have for any $\epsilon > 0$, there exists $\delta > 0$ such that $f(\bar{x}) < f(x) + \epsilon\|x - \bar{x}\|$ for any $x \neq \bar{x}$ and $x \in \mathcal{B}(\bar{x}, \delta) \cap S$.*

Proof Suppose on the contrary that there exists a sequence $(x_k) \subseteq S$ with $x_k \neq \bar{x}$ and $x_k \rightarrow_u \bar{x}$ for some $u \in K_S(\bar{x})$ such that

$$f(x_k) - f(\bar{x}) \leq -\epsilon\|x_k - \bar{x}\|$$

for some $\epsilon > 0$. In particular, let $t_k = x_k - \bar{x}$

$$D_-f(\bar{x}; u) \leq \lim_{k\to\infty}\inf \frac{1}{t_k}[f(x_k) - f(\bar{x})] \leq -\epsilon.$$

This contradicts the assumption $D_-f(\bar{x}, u) \geq 0$. □

Similar to Theorem (3.2.1), the result $0 \in \partial_u f(\bar{x})$ still holds when we restrict to the case $X = \mathbb{R}^n$ and relax the minimality condition of $\bar{x}$ to $D_-f(\bar{x}; v) \geq 0$ for all $v \in \mathbb{R}^n$:

Theorem 3.2.2 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz function. If $D_-f(\bar{x}; v) \geq 0$ for all $v \in \mathbb{R}^n$ and $D_-f(\bar{x}; u) = 0$ for $u \in \mathbb{R}^n$ with $\|u\| = 1$, then $0 \in \partial_u f(\bar{x})$.*

Proof Let $\alpha > 1$. By Lemma (3.2.9) with $D_-f(\bar{x}; u) \geq 0$ for all $u \in \mathbb{R}^n$, one has for any $(\epsilon_k) \downarrow 0$ with $0 < \epsilon_k < \frac{1}{2}\alpha^2$, there exists $(\delta_k) \downarrow 0$ such that $F_{\epsilon_k}(\bar{x}) \leq F_{\epsilon_k}(x)$ for all $\|x - \bar{x}\| \leq \delta_k$ where $F_{\epsilon_k}(x) = f(x) + \epsilon_k\|x - \bar{x}\|$. It can be easily seen that F_{ϵ_k} is a locally Lipschitz function. If $u \in \mathbb{R}^n$ with norm equal to 1 and $D_-f(\bar{x}; u) = 0$, then

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

for all k. Applied Lemma (3.2.6) for $\bar{x}$ a local minimum of F_{ϵ_k} and $2\epsilon_k$ instead of ϵ, one can find $(t_k) \downarrow 0$, $z_k \in X \setminus \{\bar{x}\}$ and $y_k^* \in \partial F_{\epsilon_k}(z_k)$ satisfying the properties

(i) $\|z_k - \bar{x} - t_k u\| \leq \frac{t_k (2\epsilon_k)^{1/2}}{\alpha}$

(ii) $\|y_k^*\| \leq \alpha(2\epsilon_k)^{1/2}$.

Particularly, (i) becomes

$$\|z_k - \bar{x} - t_k u\| \leq \frac{t_k (2\epsilon_k)^{1/2}}{\alpha} < t_k (2\epsilon_k)^{1/2}$$

Then let $k \to \infty$, $([z_k - \bar{x}]/t_k) \to u$ showing that $(z_k) \to \bar{x}$ in the direction u. As

$$y_k^* \in \partial F_{\epsilon_k}(z_k) \subseteq \partial f(z_k) + \epsilon_k B_1^*$$

for sufficient large k and (ii), there exists $z_k^* \in \partial f(z_k)$ and $w_k^* \in \epsilon_k \mathcal{B}_1^*$ for large k such that

$$y_k^* = z_k^* + w_k^* \quad \textit{and} \quad \|z_k^* + w_k^*\| \leq \alpha(2\epsilon_k)^{1/2}.$$

Let k goes to infinity, then ϵ_k tend to zero. Consequently, both y_k^* and w_k^* go to zero by (ii) and $w_k^* \in \epsilon_k \mathcal{B}_1^*$. Hence, z_k^* goes to zero too, that is $0 \in \partial_u f(\bar{x})$. □

Theorem 3.2.3 (Second-order sufficient conditions without constraint) *Let $f : \mathbb{R}^n \to \mathbb{R}$ is a locally Lipschitz function. Suppose that $D_- f(\bar{x}, v) \geq 0$ for all unit vector v in $\mathbb{R}^n$. If $f''_-(\bar{x}, 0, u) > 0$ for all unit vector 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 $x \in \mathcal{B}(\bar{x}, \delta)$.*

Proof Suppose on the contrary that for each $\gamma > 0$, there exists $y_\gamma \in \mathcal{B}[\bar{x}, \gamma] \backslash \{\bar{x}\}$ at which f attained minimum on closed set $\mathcal{B}[\bar{x}, \gamma] \backslash \{\bar{x}\}$ such that $f(y_\gamma) \leq f(\bar{x})$ by the continuity of f.

Case (i) Suppose that there exists an sequence $(\gamma_k) \downarrow 0$ and $\|y_{\gamma_k} - \bar{x}\| < \gamma_k$. Applied Lemma (3.2.2), we have $0 \in \partial f(y_{\gamma_k})$ for each k. One can assume that by taking subsequence if necessary $y_{\gamma_k} \to_u \bar{x}$ for some direction u. Hence, $0 \in \partial_u f(\bar{x})$

and $f''_{-}(\bar{x}, 0, u)$ is meaningfully defined. By the assumption that $D_{-}f(\bar{x}; v) \geq 0$ for all $v \in \mathbb{R}^n$ and $f(y_{\gamma_k}) \leq f(\bar{x})$,

$$0 \leq D_{-}f(\bar{x}; u) \leq \lim_{k\to\infty}\inf [f(y_{\gamma_k}) - f(\bar{x})]/\|y_{\gamma_k} - \bar{x}\| \leq 0.$$

That is $D_{-}f(\bar{x}; u) = 0$. Similarly,

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

This contradicts $f''_{-}(\bar{x}, 0, u) > 0$ for which $D_{-}f(\bar{x}; u) = 0$.

Case(ii) Suppose there exists $\hat{\gamma} > 0$ with $\|y_\gamma - \bar{x}\| = \gamma$ for all $0 < \gamma \leq \hat{\gamma}$.
Since $D_{-}f(\bar{x}; u) \geq 0$ for all unit vector u in $\mathbb{R}^n$, we can apply Lemma (3.2.9): let $\epsilon > 0$ there exists $\delta > 0$ such that $f(\bar{x}) \leq f(x) + \epsilon\|x - \bar{x}\|$ for all $x \in \mathcal{B}[\bar{x}, \delta]$. Taking γ_0 =min$(\delta, \hat{\gamma})$ and $\gamma_k = \gamma_0/2^k$, for $k = 1, 2, \dots$, we have (y_{γ_k}) with the following properties:

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

(2) $f(y_{\gamma_{k+1}}) \leq f(\bar{x}) \leq f(x) + \epsilon\gamma_k$ for all $x \in \mathcal{B}[\bar{x}, \gamma_k]$ and $k = 0, 1, \dots$.

By Ekeland's variational principle Lemma (3.2.1) (with $\lambda = \gamma_{k+2}$), there exists $x_{k+1} \in \mathcal{B}[\bar{x}, \gamma_k]$ such that

(i) $\|x_{k+1} - y_{\gamma_{k+1}}\| \leq \gamma_{k+2}\,(= \frac{1}{4}\gamma_k)$

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

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

for all $x \in \mathcal{B}[\bar{x}, \gamma_k]$.

Therefore, we have $f(x_{k+1}) \leq f(\bar{x})$ by (2) and (ii), $x_{k+1} \neq \bar{x}$ and $x_{k+1} \in \mathcal{B}(\bar{x}, \gamma_k)$ which follow from (1) and (i). By (iii), Lemma (3.2.2) and Lemma (3.2.3), one has $0 \in \partial f(x_{k+1}) + 4\epsilon\mathcal{B}_1^*$. This means there exists $z_{k+1} \in \partial f(x_{k+1})$ with $\|z_{k+1}\| \leq 4\epsilon$. Thus, we have constructed for any $\epsilon > 0$, there exists a sequence (x_k) and $z_k \in \partial f(x_k)$ such that $x_k \neq \bar{x}$, $x_k \to \bar{x}$, $f(x_k) \leq f(\bar{x})$ and $\|z_k\| \leq 4\epsilon$.

Inductively, let $\epsilon_n \downarrow 0$, there exist a sequence $x^n \to_u \bar{x}$ for some u with $f(x^n) \leq f(\bar{x})$, $z^n \in \partial f(x^n)$ such that $z^n \to 0$. This means $0 \in \partial_u f(\bar{x})$ and thus $f''_-(\bar{x}, 0, u)$ is meaningfully defined.

Since $f(x^n) \leq f(\bar{x})$ for all n, we have

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

This contradicts the assumption with $D_- f(\bar{x}; u) = 0$, we have $f''_-(\bar{x}, 0, u) > 0$ where $D_- f(\bar{x}; u) = 0$ follows from

$$0 \leq D_- f(\bar{x}; u) \leq \liminf_{n\to\infty} [f(x^n) - f(\bar{x})]/\|x^n - \bar{x}\| \leq 0.$$

□

Corollary 3.2.2 *Let $\bar{x} \in \mathbb{R}^n$. If*

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

(ii) $f''_-(\bar{x}, 0, u) > 0$ for all unit vector u with $D_- f(\bar{x}; u) = 0$,

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

Proof It follows from Lemma (3.2.8) that for each $u \in \mathbb{R}^n$, there exists $w_+ \in \partial_u f(\bar{x})$ such that $D_- f(\bar{x}; u) = \langle w_+, u \rangle \geq 0$. Then the result holds by Lemma (3.2.8). □

3.3 The Lagrange and G-functions in constrained problems

Let $f, g_1, \ldots, g_m, \ldots, g_{m+p}$ be real-valued locally Lipschitz functions on an open subset U in a Banach Space X. We consider the following optimization problem:

$$\begin{aligned} \mathcal{P}(X): \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0 \quad \text{for}\, i = 1, 2 \ldots, m; \\ & g_i(x) = 0 \quad \text{for}\, i = m+1, \ldots, m+p. \end{aligned}$$

Let S be the feasible set of $\mathcal{P}(X)$ and $K_S(x)$ the contingent cone of S at x. Also, $I(x)$ denoted the set of active indices i at x with $1 \leq i \leq m$ such that $g_i(x) = 0$. $NI(x)$ is the set of all indices i with $1 \leq i \leq m$ such that $g_i(x) < 0$. With the above notations, one can easily see that for $\bar{x} \in S$ and for each $u \in K_S(\bar{x})$, we have the tangential constraints:

(i) For all $i \in I(\bar{x})$,

$$\begin{aligned} D_{-}g_i(\bar{x}; u) &= \liminf_{t \downarrow 0} \frac{1}{t}\{g_i(x \bar{+} tu - g_i(\bar{x})\} \\ &= \liminf_{t \downarrow 0} \frac{1}{t}\{g_i(\bar{x} + tu)\} \\ &\leq 0 \end{aligned}$$

(ii) For all $k = m+1, \ldots, m+p$,

$$\begin{aligned} D_{-}g_k(\bar{x}; u) &= \liminf_{t \downarrow 0} \frac{1}{t}\{g_k(\bar{x} + tu) - g_k(\bar{x})\} \\ &= \liminf_{t \downarrow 0} \frac{1}{t}\{g_k(\bar{x} + ti)\} \\ &\leq 0 \end{aligned}$$

and

$$\begin{aligned} D_{+}g_k(\bar{x}; u) &= \limsup_{t \downarrow 0} \frac{1}{t}\{g_k(\bar{x} + tu) - g_k(\bar{x})\} \\ &= \limsup_{t \downarrow 0} \frac{1}{t}\{g_k(\bar{x} + ti)\} \\ &\geq 0 \quad \text{as}\, u \in K_S(\bar{x}). \end{aligned}$$

Let W be the set of all vectors $w = (w_0, w_1, \ldots, w_{m+p})$ in $\mathbb{R}^{1+m+p}$ such that $\sum_{i=0}^{m+p}(w_i)^2 = 1$ and $w_i \geq 0$ for $i = 0, 1, \ldots, m$. Particularly, $W_1(x)$ is defined as

$$\begin{aligned} W_1(x) &= \{w \in W \mid w_i = 0, \quad \forall i \in NI(x)\} \\ &= \{w \in W \mid w_i g_i(x) = 0, \quad \forall i = 1, 2, \ldots, m + p\}. \end{aligned}$$

In order to apply the results in the preceding section for unconstrained problems, we introduce the Lagrange function L and $G_{\bar{x}}$:

Definition 3.3.1 *The Lagrange function L for $x \in U$ and $w \in W$ is defined by*

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

and for $x \in U$ and $\bar{x} \in S$

$$G_{\bar{x}}(x, f) = \max_{w \in W} \{L(x, w) - w_0 f(\bar{x})\}.$$

If we let $g_0(x) = f(x) - f(\bar{x})$, we have

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

For simplicity, we simply write G instead of $G_{\bar{x}}$ and W_1 for $W_1(\bar{x})$ whenever $\bar{x}$ is specified.

Remark If the equality constraints do not appear, that is $p = 0$ and $g_i(\bar{x}) < 0$ for all $i = 1, \ldots, m$, then all x near $\bar{x}$ are feasible by the locally continuous property of all $g_i s$ and so the problem $\mathcal{P}(X)$ becomes an unconstrained problem which already studied in Section 2. Therefore, we can assume $g_j(\bar{x}) = 0$ for some j with $1 \leq j \leq m + p$.

Lemma 3.3.1 *If $g_j(\bar{x}) = 0$ for some j with $1 \leq j \leq m+p$, then for all x near $\bar{x}$, we have*

$$\begin{aligned} G(x,f) &= \max_{w \in W} \{w_0[f(x) - f(\bar{x})] + \sum_{i=1}^{m+p} w_i g_i(x)\} \\ &= \max_{w \in W_1(\bar{x})} \{w_0[f(x) - f(\bar{x})] + \sum_{i=1}^{m+p} w_i g_i(x)\} \\ &= \max_{w \in W_1(\bar{x})} \{ \sum_{i=0}^{m+p} w_i g_i(x) \} \end{aligned}$$

where $g_0(x) = f(x) - f(\bar{x})$. Also, if for all x near $\bar{x}$ and $w \in W$, one has

$$G(x,f) \; = \; w_0[f(x) - f(\bar{x})] + \sum_{i+1}^{m+p} w_i g_i(x), \qquad (3.2)$$

then $w \in W_1(\bar{x})$.

Proof The first part is obvious and we need to prove the last assertion only. Let $i \in NI(\bar{x})$, we have $0 = g_j(\bar{x}) > g_i(\bar{x})$. Since g_i and g_j are locally Lipschitz functions, there exists $\delta_i > 0$ such that $g_j(x) > g_i(x)$ for all $x \in \mathcal{B}(\bar{x}, \delta_i)$ and each $i \in NI(\bar{x})$. Then we take $\delta = \min_{k \in NI(\bar{x})}\{\delta_k\}$. Let $y \in \mathcal{B}(\bar{x}, \delta)$ and y satisfies (3.2). We have

$$g_i(y) < g_j(y) \qquad \text{and} \qquad w_i \geq 0$$

Therefore, $w_i = 0$ by the maximality condition of G. Since it holds for all $i \in NI(\bar{x})$, we have $w \in W_1(\bar{x})$. □

Lemma 3.3.2 *(i) If $\bar{x}$ is a local solution to the problem $\mathcal{P}(X)$, then $G(x,f) \geq 0 = G(\bar{x}, f)$ for all x near $\bar{x}$, that is $\bar{x}$ is a local minimizer of G.*

(ii) $\bar{x}$ is a strict local solution to problem $\mathcal{P}(X)$ if and only if we have $G(x,f) > 0 = G(\bar{x}, f)$ for all x near $\bar{x}$ with $x \neq \bar{x}$, that is $\bar{x}$ is a strict local minimizer of G.

Proof (i) Suppose $\bar{x}$ is a local solution of $\mathcal{P}(X)$. Clearly,

$$G(\bar{x}, f) = \max_{w \in W}\{w_0 f(\bar{x}) + \sum_{i=1}^{m+p} w_i g_i(\bar{x}) - w_0 f(\bar{x})\} = 0.$$

Also, there exists an neighborhood U of $\bar{x}$ such that $f(x) \geq f(\bar{x})$ for all feasible $x \in U$. Particularly, if we take $w = (1, 0, 0, \dots) \in \mathbb{R}^{m+p+1}$, then for each feasible $x \in U$,

$$G(x, f) \geq L(x, w) - f(\bar{x}) \geq f(x) - f(\bar{x}) \geq 0 = G(\bar{x}, f).$$

If $x \in U$ is infeasible, then either there exists i with $1 \leq i \leq m$ such that $g_i(x) > 0$ or $g_i(x) \neq 0$ for some $i > m$. In this case, we choose $w_j = 0$ for all $j \neq i$ and $w_i = sgn(g_i(x))$. It follows that

$$G(x, f) \geq L(x, w) - w_0 f(\bar{x}) = w_i g_i(x) > 0 = G(\bar{x}, f).$$

(ii) The 'only if' part follows exactly as (i), but $f(x) > f(\bar{x})$ for all feasible x near $\bar{x}$. We are going to prove the 'if' part:

Suppose that there exists an neighborhood V of $\bar{x}$ such that $G(x, f) > 0 = G(\bar{x}, f)$ for all $x \in V \backslash \{\bar{x}\}$. Let x is feasible, $x \in V \backslash \{\bar{x}\}$ and $w \in W$, we have $\sum_{i=1}^{m+p} w_i g_i(x) \leq 0$. Thus,

$$\begin{aligned} 0 < G(x, f) &= \max_{w \in W}\{\, w_0[f(x) - f(\bar{x})] + \sum_{i=1}^{m+p} w_i g_i(x)\,\} \\ &\leq \max_{w \in W}\{\, w_0[f(x) - f(\bar{x})]\,\}. \end{aligned}$$

This implies $f(x) > f(\bar{x})$ for all feasible $x \in V \backslash \{\bar{x}\}$ as $w_0 \geq 0$. □

Lemma 3.3.3 *Let $x \in U$ and $w \in W$. Suppose that $G(x, f) = \sum_{i=0}^{m+p} w_i g_i(x) \geq 0$. Then $w_i g_i(x) \geq 0$ for all i.*

Proof It is obviously true for $i > m$, otherwise we can change the sign of w_i to obtain the maximality of $G(x, f)$.

For the case $i \leq m$, suppose on the contrary that $w_k g_k(x) < 0$ for some $1 \leq k \leq m$. We claimed that $w_i g_i(x) = 0$ for all $i > m$. If not, there exists $i > m$ such that $w_i g_i(x) > 0$. Then by changing w_k to zero and w_i to $(w_k^2 + w_i^2)^{1/2}$, it contradicts the maximality of $G(x, f)$. Therefore, $w_i g_i(x) = 0$ for all $i > m$. It follows that $\sum_{i=0}^{m+p} w_i g_i(x) = \sum_{i=0}^{m} w_i g_i(x) \geq 0$. Since $w_k g_k(x) < 0$, we can then find some $1 \leq j \leq m$ with $g_j(x) > 0$ and $w_j > 0$. Again, if we replace w_k and w_j into zero and $(w_k^2 + w_j^2)^{1/2}$ respectively, contradiction on maximality of $G(x, f)$ is then obtained as before. □

For the following lemma, we assume the inequality holds for some $\gamma > 0$:

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

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

Lemma 3.3.4 *Let M be a Lipschitz constant of $f, g_1, \ldots, g_{m+p}$ on $\mathcal{B}(\bar{x}, \delta)$ for some $\delta > 0$. Then for any $w, u \in W$, $t > 0$, $\|v\| = 1$ with $y, y + tv \in \mathcal{B}(\bar{x}, \delta)$, one has*

(i) $L(y + tv, w) - L(y, w) - [L(y + tv, u) - L(y, u)] \leq \gamma M t \|w - u\|$

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

Proof By definitions, we have

$$L(y + tv, w) - L(y + tv, u) = (w_0 - u_0) f(y + tv) + \sum_{i=1}^{m+p} (w_i - u_i) g_i(y + tv)$$

and

$$L(y, w) - L(y, u) = (w_0 - u_0) f(y) + \sum_{i=1}^{m+p} (w_i - u_i) g_i(y).$$

Therefore, by subtraction

$$\begin{aligned} &L(y+tv,w) - L(y,w) - [L(y+tv,u) - L(y,u)] \\ &= (w_0-u_0)[f(y+tv)-f(y)] + \sum_{i=1}^{m+p}(w_i-u_i)[g_i(y+tv)-g_i(y)] \\ &\le \sum_{i=1}^{m+p}|w_i-u_i|Mt \le \gamma Mt(\sum_{i=1}^{m+p}|w_i-u_i|^2)^{1/2} = \gamma Mt\|w-u\|. \end{aligned}$$

This proves (i). (ii) follows from (i) by taking upper limits with $y \to x$ and $t \downarrow 0$ in

$$\frac{1}{t}\{L(y+tv,w) - L(y,w)\} \le \frac{1}{t}\{L(y+tv,u) - L(y,u)\} + \gamma Mt\|w-u\|.$$

□

Lemma 3.3.5 *Suppose that $g_j(\bar{x}) = 0$ for some j with $1 \le j \le m+p$. Then for some $\delta > 0$ and for all $x \in \mathcal{B}(\bar{x},\delta)$, there exists $w \in W_1(\bar{x})$ 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) $\partial G(\cdot,f)(x) \subseteq \partial L(\cdot,w)(x)$.

Proof By Lemma (3.3.1), we already have (i) that for some $\delta > 0$ and for each $x \in \mathcal{B}(\bar{x},\delta)$, $G(x,f) = L(x,w_0) - w_0 f(\bar{x})$ for some $w \in W_1(\bar{x})$. The key point is to find some element in $W_1(\bar{x})$ so that it satisfies (i),(ii) and (iii).

Fixed $x \in \mathcal{B}(\bar{x},\delta)$, let y near x and $t > 0$ near zero such that $y+tv, y \in \mathcal{B}(\bar{x},\delta)$ where $v \in X$. It follows that there exists $u(y,t) \in W_1(\bar{x})$ such that

$$G(y+tv,f) = L(y+tv,u) - u_0 f(\bar{x}) \tag{3.3}$$

and

$$G(y,f) \ge L(y,u) - u_0 f(\bar{x}).$$

Thus,

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

Let $y \to x$, $t \downarrow 0$ and take upper limit, we have

$$\begin{aligned} G^0(\cdot,f)(x;v) &\le \limsup_{\substack{y\to x\\ t\downarrow 0}} \{ \frac{1}{t}[L(y+tv,u(y,t)) - L(y,u(y,t))] \} \\ &= \lim_{n\to\infty} \{ \frac{1}{t_n}[L(y_n+t_nv,u^n) - L(y_n,u^n)] \} \end{aligned} \tag{3.4}$$

where $(y_n),(u^n)$ and (t_n) are the appropriate sequences chosen to be satisfied the last equality with $y_n \to x$, $t_n \downarrow 0$ and $u^n = u^n(y_n,t_n)$ converges to some $w \in W_1(\bar{x})$ which follows from the compactness of $W_1(\bar{x})$. By Lemma (3.3.4) (i) and (3.4), we then have

$$\begin{aligned} G^0(\cdot,f)(x;v) &\le \limsup_{n\to\infty} \frac{1}{t_n}\{ [L(y_n+t_nv,w) - L(y_n,w)] + \gamma M t_n \|u^n - w\| \} \\ &= \limsup_{n\to\infty} \frac{1}{t_n}\{L(y_n+t_nv,w) - L(y_n,w)\} \\ &\le L^0(\cdot,w)(x;v) \end{aligned}$$

Since this hold for all $v \in X$, (ii) holds.

By the definition of subdifferential, if $x^* \in \partial G(\cdot,f)(x)$, then

$$\begin{aligned} x^*(v) &\le G^0(\cdot,f)(x;v) \quad \text{for all}\, v \in X \\ &\le L^0(\cdot,w)(x;v) \quad \text{for all}\, v \in X \end{aligned}$$

which follows by (ii). Thus, $x^* \in \partial L(\cdot,w)(x)$.

The remaining is to show that w satisfies (i). We noted that from the definition of u^n

$$\begin{aligned} L(y_n+t_nv,u^n) - u_0^n f(\bar{x}) &= G(y_n+t_nv,f) \\ &\ge L(y_n+t_nv,\mu) - \mu_0 f(\bar{x}) \end{aligned}$$

for all $\mu \in W_1(\bar{x})$. By continuities, it becomes

$$L(x,w) - w_0 f(\bar{x}) \ge L(x,\mu) - \mu f(\bar{x}) \quad \text{for all } \mu \in W_1(\bar{x}).$$

That is $G(x, f) = L(x, w) - w_0 f(\bar{x})$ and w satisfies (i) finally. □

Let

$$M(x) = \{ w \in W \mid 0 \in \partial L(\cdot, w)(x) \quad \text{and} \quad w_i g_i(x) = 0 \quad \text{for } i = 1, \ldots, m \}$$

and for any $u \in X$,

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

Also, let

$$M_-(x) = \{ w \in W \mid 0 \leq D_- L(\cdot, w)(x; v) \quad \text{for all } v \in X \quad \text{and} \quad w_i g_i(x) = 0 \quad \text{for } i = 1, \ldots, m \}.$$

Proposition 3.3.1 *Suppose that $g_j(\bar{x}) = 0$ for some j and u is a unit vector in X.*

(i) If $0 \in \partial_u G(\cdot, f)(\bar{x})$, then $M_u(\bar{x}) \neq \emptyset$.
In fact, if $0 \in \partial_u G(\cdot, f)(\bar{x})$ such that there exists $(x_k) \subseteq X$, $x_k \to_u \bar{x}$ and $x_k^* \in \partial G(\cdot, f)(x_k)$ for each k with $x_k^* \to 0$, then there exists $(w^k) \subseteq W_1(\bar{x})$ and $w^k \to \bar{w}$ for some cluster point $\bar{w}$ in $W_1(\bar{x})$. Furthermore, there exists (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) $(\|y_k^*\|) \to 0$ *as* $k \to \infty$.
Consequently, $0 \in \partial_u L(\cdot, \bar{w})(\bar{x})$ *or* $\bar{w} \in M_u(\bar{x})$.

(ii) If $X = \mathbb{R}^n$ and u be a unit vector such that $D_- G(\cdot, f)(\bar{x}; u) = 0$, then $M_-(\bar{x}) \subseteq M_u(\bar{x})$.

Proof (i) By Lemma (3.3.5), for sufficient large k, there exists w^k in $W_1(\bar{x})$ such that

$$\begin{aligned} G(x_k, f) &= L(x_k, w^k) - w_0^k f(\bar{x}) \\ G^0(\cdot, f)(x_k; v) &\leq L^0(\cdot, w^k)(x_k; v) \quad \text{for all}\, v \in X \\ \partial G(\cdot, f)(x_k) &\subseteq \partial L(\cdot, w^k)(x_k) \end{aligned} \tag{3.5}$$

Thus, (a) holds by the first equality. We can assume that taking subsequence if necessary w^k converges to some $\bar{w}$ in $W_1(\bar{x})$ as $W_1(\bar{x})$ is compact. By Lemma (3.3.4) (ii), for $w^k, \bar{w} \in \mathbb{R}^{1+m+p}$ and each $v \in X$,

$$L^0(\cdot, w^k)(x_k; v) \leq L^0(\cdot, \bar{w})(x_k; v) + \gamma M\|w^k - \bar{w}\|\|v\|$$

Then

$$\partial L(\cdot, w^k)(x_k) \subseteq \partial L(\cdot, \bar{w})(x_k) + \gamma M\|w^k - \bar{w}\|\mathcal{B}_1^* \tag{3.6}$$

which follows from Lemma (3.2.3).

Together with (3.5), (3.6) and the definition of x_k^*, there exists $y_k^* \in \partial L(., \bar{w})(x_k)$ for each k such that $x_k^* \in y_k^* + \gamma M\|w^k - \bar{w}\|\mathcal{B}_1^*$. Let k tends to infinity, we have $x_k^* \to 0$ and $w^k \to \bar{w}$. Therefore, $(\|y_k^*\|) \to 0$ as $k \to \infty$.

We have shown that $x_k \to_u \bar{x}$, $y_k^* \in \partial L(\cdot, \bar{w})(x_k)$, $y_k^* \to 0$ and $\bar{w} \in W_1(\bar{x})$, that is $0 \in \partial_u L(\cdot, \bar{w})(\bar{x})$ or $\bar{w} \in M_u(\bar{x})$.

(ii) Let w in $M_{-}(\bar{x})$. Then, for $\bar{x}$ is feasible $w_i g_i(\bar{x}) = 0$ for all $i = 1, 2, \ldots, m+p$ and $D_{-}L(\cdot, w)(\bar{x}; v) \geq 0$ for all $v \in X = \mathbb{R}^n$. It follows that $L(\bar{x}, w) = w_0 f(\bar{x})$, that is $G(\bar{x}, f) = 0$. By the definition of G, for $t > 0$, we have

$$G(\bar{x} + tu, f) \geq L(\bar{x} + tu, w) - w_0 f(\bar{x})$$

that is

$$G(\bar{x} + tu, f) - G(\bar{x}, f) \geq L(\bar{x} + tu, w) - w_0 f(\bar{x}).$$

If we divide both side by t and then take lower limits,

$$D_{-}G(\cdot, f)(\bar{x}; u) \geq D_{-}L(\cdot, w)(\bar{x}; u).$$

With the assumption that $D_{-}G(\cdot, f)(\bar{x}; u) = 0$ and $D_{-}L(\cdot, w)(\bar{x}; v) \geq 0$ for all $v \in \mathbb{R}^n$ and the above inequality, we thus have $D_{-}L(\cdot, w)(\bar{x}; u) = 0$. By Theorem (3.2.2), one has $0 \in \partial_u L(\cdot, w)(\bar{x})$. Therefore, $w \in M_u(\bar{x})$. □

3.4 Second-order necessary conditions for constrained problems

Theorem 3.4.1 (Second-order necessary condition with constraints) *Suppose that $\bar{x}$ is a local minimum of $\mathcal{P}(X)$ and u a unit vector such that $D_{-}G(\cdot, f)(\bar{x}; u) = 0$(or more generally $0 \in \partial_u G(\cdot, f)(\bar{x})$). Then $G''_{-}(\bar{x}, 0, u) \geq 0$ 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 Since $\bar{x}$ is a local minimum of $\mathcal{P}(X)$, it is also a local minimum of $G(x, f)$ which follows from Lemma (3.3.2) (i). By assumption that $D_{-}G(x, f)(\bar{x}; u) = 0$, one can apply Theorem (3.2.1) for G to get

$$0 \in \partial_u G(\cdot, f)(\bar{x}) \quad \text{and} \quad G''_{-}(\bar{x}, 0, u) \geq 0.$$

Then, one can take a sequence (x_k) converges to $\bar{x}$ in the direction u, $x_k^* \in \partial G(\cdot, f)(x_k)$ for each k and x_k^* converges to 0. By Proposition (3.3.1), there exists a sequence (w^k) in $W_1(\bar{x})$ which converges to some element $\bar{w}$ in $W_1(\bar{x})$ and $\bar{w} \in M_u(\bar{x})$.

Now, we want to prove $L''_{-}(\bar{x}, \bar{w}, 0, u) \geq 0$.

Suppose on the contrary that for all sufficient large k

$$L(x_k, \bar{w}) - L(\bar{x}, \bar{w}) < 0,$$

that is

$$\bar{w_0}[f(x_k) - f(\bar{x})] + \sum_{i=1}^{m+p} \bar{w_i} g_i(x_k) < 0 \tag{3.7}$$

as $L(\bar{x}, \bar{w}) = \bar{w_0} f(\bar{x})$ for $\bar{w}$ in $W_1(\bar{x})$. Let $g_0(\cdot) = f(\cdot) - f(\bar{x})$, in order to have (3.7), there exist h with $0 \leq h \leq m+p$ such that

$$\bar{w_h} g_h(x_k) < 0$$

for all infinitely many k. By considering subsequence if necessary, we can assume this hold for all k. Since (w^k) converges to $\bar{w}$, we have

$$w_h^k g_h(x_k) < 0 \quad \text{for some k.}$$

This contradicts Lemma (3.3.3) that $w_h^k g_h(x_k) \geq 0$ for $G(x_k, f) \geq 0$. $G(x_k, f) \geq 0$ because $\bar{x}$ is the local minimum of $G(\cdot, f)$ and $G(\bar{x}, f) = 0$. Hence, the result follows. □

Theorem 3.4.2 (Second-order necessary condition with constraints in $\mathbb{R}^n$) *If $\bar{x}$ is a local minimum of $\mathcal{P}(\mathbb{R}^n)$ and u is a unit vector with $D_-G(\cdot, f)(\bar{x}; u) = 0$, then $G''_-(\bar{x}, 0, u) \geq 0$, $M_-(\bar{x}) \subseteq M_-(\bar{x})$ and there exists an Lagrange multiplier $\bar{w}$ in $M_u(\bar{x})$ such that $L''_+(\bar{x}, \bar{w}, 0, u) \geq 0$.*

Proof For $X = \mathbb{R}^n$, $M_-(\bar{x}) \subseteq M_u(\bar{x})$ by Proposition (3.3.1) (ii) and the result follows form Theorem (3.2.9). □

3.5 Sufficient conditions for constrained problems

Lemma 3.5.1 *Let $\bar{x} \in S \subseteq \mathbb{R}^n$ such that $D_-f(\bar{x}; v) \geq 0$ for all $v \in K_S(\bar{x})$. Then*

(i) $D_-G(\cdot, f)(\bar{x}; v) \geq 0$ for all $v \in \mathbb{R}^n$

(ii) If $D_-G(\cdot, f)(\bar{x}; u) = 0$ for a unit vector u, then $0 \in \partial_u G(\cdot, f)(\bar{x})$, $M_-(\bar{x}) \subseteq M_u(\bar{x})$ and $M_u(\bar{x})$ is nonempty.

Remark If no equality constraint for $\bar{x}$, that is $p = 0$, $g_i(\bar{x}) < 0$ for $i = 1, 2, \ldots, m$, then $K_S(\bar{x}) = \mathbb{R}^n$. So, the assumption becomes $D_-f(\bar{x}, v) \geq 0$ for all $v \in \mathbb{R}^n$, or $f(x) - f(\bar{x}) \geq g_i(x)$ for all x near $\bar{x}$ and all i. Thus (i) holds as $G(x, f) = f(x) - f(\bar{x})$. Noticed that $W_1(\bar{x})$ contains $\bar{w} = \{1, 0, \ldots, 0\}$ only, $L(x, \bar{w}) - L(\bar{x}, \bar{w}) = f(x) - f(\bar{x}) \geq 0$ for all x near $\bar{x}$ and hence $M_-(\bar{x}) = \{\bar{w}\}$. By $D_-G(\cdot, f)(\bar{x}; \mu) \geq 0$ for all $\mu \in \mathbb{R}^n$ and the assumption in (ii), applied Theorem (3.2.2), one has $0 \in \partial_u G(\cdot, f)(\bar{x}) = \partial_u L(\cdot, \bar{w})(\bar{x}) = \partial_u f(\bar{x})$. Hence, $\{\bar{w}\} \subseteq M_-(\bar{x}) \subseteq M_u(\bar{x})$.

Proof By the definition of G,

$$G(\cdot, F_\epsilon) = \max_{w \in W}\{ w_0[F_\epsilon(\cdot) - F_\epsilon(\bar{x})] + \sum_{i=1}^{m+p} w_i g_i(\cdot)\}$$

where $F_\epsilon(\cdot) = f(\cdot) + \epsilon\|\cdot - \bar{x}\|$. As F_ϵ is locally Lipschitz function and $D_-f(\bar{x}; v) \geq 0$ for all $v \in K_S(\bar{x})$, Lemma (3.2.9) told us $\bar{x}$ is a strict local minimizer of F_ϵ. It is also a strict local minimizer of $G(\cdot, F_\epsilon)$ without constraints which follows from Lemma (3.3.2). Therefore,

$$D_-G(\cdot, F_\epsilon)(\bar{x}; v) \geq 0 \quad \text{for all}\, v \in \mathbb{R}^n. \tag{3.8}$$

On the other hand, $G(\bar{x}, f) = \max_{w \in W}\{\sum_{i=1}^{m+p} w_i g_i(\bar{x})\} = G(\bar{x}, F_\epsilon)$ and $\bar{x}$ is feasible, thus $G(\bar{x}, f) = 0 = G(\bar{x}, F_\epsilon)$.

Since

$$\begin{aligned} G(\bar{x} + tv, F_\epsilon) &= \max_{w \in W}\{w_0[F_\epsilon(\bar{x} + tv) - F_\epsilon(\bar{x})] + \sum_{i=1}^{m+p} w_i g_i(\bar{x} + tv)\} \\ &\leq \max_{w \in W}\{w_0[f_\epsilon(\bar{x} + tv) - f_\epsilon(\bar{x})] + \sum_{i=1}^{m+p} w_i g_i(\bar{x} + tv)\} + \epsilon\|tv\| \\ &= G(\bar{x} + tv, f) + \epsilon\|tv\| \end{aligned}$$

for all t and $v \in \mathbb{R}^n$, $D_-G(\cdot, F_\epsilon)(\bar{x}; v) \leq D_-G(\cdot, f)(\bar{x}; v) + \epsilon\|v\|$. Hence, $0 \leq D_-G(\cdot, f)(\bar{x}; v) + \epsilon\|v\|$ by (3.8). (i) is then proved as ϵ is arbitrary.

(ii) $G(\cdot, f)$ is a locally Lipschitz function as g_i, f are for all i. Together with (i) that $D_-G(\cdot, f)(\bar{x}; v) \geq 0$ for all $v \in \mathbb{R}^n$ and the assumption in (ii), one has $0 \in \partial_u G(\cdot, f)(\bar{x})$ by Theorem (3.2.2). Also, $M_-(\bar{x}) \subseteq M_u(\bar{x})$ and $M_u(\bar{x}) \neq \emptyset$ which follows from Proposition (3.3.1). $\square$

The following theorem is the main result in this section for unconstrained problem:

Theorem 3.5.1 (Second-order sufficient condition) *Let $\bar{x} \in S \subseteq \mathbb{R}^n$. If $D_-f(\bar{x}; u) \geq 0$ for all $u \in K_S(\bar{x})$ and $G''_-(\bar{x}, 0, v) > 0$ for which $D_-G(\cdot, f)(\bar{x}; v) = 0$, then there exists $\delta > 0$ such that $f(\bar{x}) < f(x)$ for all $x \in \mathcal{B}(\bar{x}, \delta) \cap S$ with $x \neq \bar{x}$.*

Proof By Lemma (3.5.1) (i) with $D_-f(\bar{x}; u) \geq 0$ for all $u \in K_S(\bar{x})$, $D_-G(\cdot, f)(\bar{x}, u) \geq 0$ for all $u \in \mathbb{R}^n$. Also, the assumption that $G''_-(\bar{x}, 0, v) > 0$ for which $D_-G(\cdot, f)(\bar{x}; v) = 0$, Theorem (3.2.3) told us $\bar{x}$ is a strict local minimum of $G(\cdot, f)$. Hence, it is a strict local minimum of $\mathcal{P}(\mathbb{R}^n)$ by Lemma (3.3.2) (ii). $\square$

Bibliography

[1] A. Auslender and J. P. Crouzeix, *Well behaved asymptotical convex functions*, Analyse Lineare, (1989), 101-112.

[2] A. Auslender, *How to deal with the unbounded in optimization: Theory and algorithms*, Mathematical Programming, Vol. 79 (1997), 3-18.

[3] A. Auslender, R. Cominetti and J. P. Crouzeix, *Convex functions with unbounded level sets and applications to duality theory*, SIAM Journal on Optimization, 3 (1993), 669-687.

[4] A. Auslender, *Convergence of critical sequences for variational inequalities with maximal monotone operators*, Applied Mathematics and Optimization, 28(1993), 161-172.

[5] A. Auslender, R. Cominetti and J.P. Crouzeix, *Convex functions with unbounded level sets and applications to duality theory*, SIAM J. Optimization, 3(1993), 669-687.

[6] A. Ben-Tal, *Second-order and related extremality conditions in nonlinear programming*, Journal of Optimization Theory and Applications 31(1980) 143-165.

[7] A. Ben-Tal and Zowe, *A unified theory of first and second order conditions for extremum problem in topological vector space*, Mathematical Programming Study 19 (1982) 39-76.

[8] A. Ben-Tal and Zowe, *Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems* Mathematical Programming 24 (1982) 70-91.

[9] Chin-Cheng Chou, Kung-Fu Ng and Jong-Shi Pang, *Minimizing and stationary sequences of Constrained Optimization Problem*, The Chinese Universtiy of Hong Kong, 1997.

[10] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.

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

[12] J.P. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms*, I & II, Springer-Verlag Berlin, 1993.

[13] J.S. Pang, S.P. Han, and N. Rangaraj, *Minimization of locally Lipschitzian functions*, SIAM Journal on Optimization 1(1991), pp.57-82.

[14] L.R. Huang, K.F. Ng and J.P. Penot, *On minimizing and critical sequences in nonsmooth optimization*, The Chinese Universtiy of Hong Kong.

[15] L.R. Huang and K.F. Ng, *Second-order necessary and sufficient conditions in nonsmooth optimization*, Math. Prog., 66(1994), 379-402.

[16] L.R. Huang and K.F. Ng., *On second order directional derivatives in nonsmooth optimization*, Recent Advances in Nonsmooth Optimization, D.Z. Du, L. Qi and R.S. Womersley (eds), world Scientific, Singapore (1995), 159-171.

[17] L.R. Huang and X.B. Li, *Minimizing sequences in nonsmooth optimization*, Preprint.

[18] P. Angleraud, *Caracterisation duale dubon comportement de fonctions convexes*, C.R.Acad. Sci. serie I, 314(1992), 583-586.

[19] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[20] T. Fujisawa and E.S. Kuh, *Piecewise linear theory of nonlinear networks*, SIAM Journal on Applied Mathematics, 22(1972), pp.307-328.

[21] T. Wang and J.S. Pang, *Global error bounds for convex quadratic inequality systems*, Optimization, 21 (1994), pp.1-12.