Sociedad de Estadística e Investigación Operativa
Top (2005) Vol. 13, No. 1, pp. 127–143

Necessary Optimality Conditions and Saddle Points for Approximate Optimization in Banach Spaces

J. Dutta
Department of Mathematics, Indian Institute of Technology, Kanpur
Kanpur-208016, India
E-mail: jdutta@iitk.ac.in

Abstract

In this article we study approximate optimality in the setting of a Banach space. We study various solution concepts existing in the literature and develop very general necessary optimality conditions in terms of limiting subdifferentials. We also study saddle point conditions and relate them to various solution concepts.

Key Words: Approximate optimization, limiting subdifferential, locally Lipschitz function, approximate saddle points.

AMS subject classification: 90C29, 49J52.

1 Approximate Optimization : Basic Motivation

To begin the study of any mathematical subject it is important to have some motivation for such a study. The problem to determine a point in a given set which minimizes a given numerical function over that set remains an interesting as well as challenging area of study. Consider the problem

$$\min f(x), \quad x \in X.$$

Most of the practical algorithms will try to find a point $\bar{x} \in X$ such that $\nabla f(\bar{x}) = 0$. In most cases this may not be easy and one may have to remain content with finding an x^* such that

$$\|\nabla f(x^*)\| < \varepsilon.$$

Part of this research was carried out when the author was a post-doctoral fellow at UAB, Barcelona by the Grant No. SB99-B0771103B of the Spanish Ministry of Education and Culture. The hospitality and the facilities provided at CODE, UAB is gratefully acknowledged.

The author is very thankful to the anonymous referees whose suggestions have enhanced the presentation of the paper.

Manuscript received: July 2003. Final version accepted: January 2005.

where $\varepsilon > 0$ is a preset value which is usually very small. Thus the point x^* may not be exactly a critical or an optimal point but may still serve our purpose from the practical point of view. Consider another example

$$\min f(x) = \frac{1}{x} \quad \text{subject to} \quad x \in [1, +\infty).$$

In this case we are seeking to find a point in the convex set $[1, +\infty)$ which minimizes the convex function $f(x) = \frac{1}{x}$ over it. Though the function has an infimum over this set which is zero but there is no $x \in [1, +\infty)$ such that $f(x) = 0$. But for a given $\varepsilon > 0$ in this case one will find $x^* \in [1, +\infty)$ such that $f(x^*) < \varepsilon$. This simple idea leads to the notion of an approximate minimum. In most decision making processes involving optimization the decision maker usually is satisfied with an approximate minima with a sufficient degree of accuracy. In this article we are concerned with the following problem (MP)

$$\min f(x), \quad \text{subject to} \quad x \in S, \tag{MP}$$

with $f : X \to R$ and $S \subseteq X$, where X is a Banach space. We shall denote by X^* the dual space of X and by $\langle ., . \rangle$ the canonical pairing between X^* and X. The norm of X will be denoted by $\|.\|$ and the dual space X^* will be equipped with the weak-star topology. As usual B and B^* stands for the unit balls in X and X^*, respectively. We shall denote the interior, closure and the convex hull of a set S by int$S$, cl$S$ and co$S$ respectively. We shall denote by cl^* the weak-star closure in X^*. We will also denote by $\overline{R}$ the extended real line $R \cup \{+\infty, -\infty\}$ wherever required. As usual we shall denote by $dom f$ the domain of an extended real-valued function $f : X \to \overline{R}$.

In this article we aim to develop a very general and sharp necessary optimality condition for an approximate minimum of a locally Lipschitz program. The necessary conditions will be in terms of the limiting subdifferential of Mordukhovich (1985). From the calculus rules of such subdifferentials it will be clear that the subclasses of Banach spaces called Asplund Space (see for example Phelps (1993)), are the most natural settings for studying such sharp optimality conditions. The paper has been organized as follows. In Section 2 we introduce the relevant solution concepts and various definitions from the non-smooth analysis of Mordukhovich. In Section 3 we present various optimality conditions. We begin by presenting some results for the general program (MP) and then consider the case where the set S is described by equality and inequality constraints. In Section 4 we describe the notion of an approximate saddle point and its relation to certain approximate solution concepts.

There has been a considerable interest among researchers in the theoretical study of the notion of an approximate minimum. One of the earliest analysis was carried out by Loridan (1982) and Loridan and Morgan (1983) where various solution concepts have been introduced and necessary optimality conditions derived using the Clarke subdifferential. There has been some further works due to Bustos (1989), Bustos (1994), and Liu (1991). For example in Bustos (1994) new notions of ε-subgradients were defined for locally Lipschitz functions and optimality conditions were derived in terms of these ε-subgradients. Very recently Mordukhovich and Wang (2002) have introduced a new variational principle using the tools of nonconvex variational analysis and have used it to study optimality conditions in approximate optimization. The above list is by no means exhaustive and some relevant references will also be given in connection with various results in this article.

2 Basic Definitions and Results

We will begin with the definition of some relevant solution concepts.

Definition 2.1. Consider the problem (MP) given as

$$\min f(x) \quad \text{subject to} \quad x \in S,$$

where $f : X \to R$ and $S \subseteq X$. Let $\varepsilon \geq 0$ be given then a point $x^* \in S$ is said to be an ε-minimum (MP) if

$$f(x^*) \leq f(x) + \varepsilon, \quad \forall x \in S. \tag{2.1}$$

If $\varepsilon = 0$ this reduces to the usual notion of a minimum. Thus in the study of approximate optimality the case $\varepsilon > 0$ is of interest. Loridan(1982) introduced another notion of an approximate solution called ε-quasiminimum. We provide the definition below:

Definition 2.2. Let $\varepsilon \geq 0$ be given. Then $x^* \in S$ is said to be an ε-quasiminimum of (MP) if

$$f(x^*) \leq f(x) + \sqrt{\varepsilon}\|x - x^*\|, \quad \forall x \in S. \tag{2.2}$$

A point $x^* \in S$ is called a regular ε-minimum of (MP) if x^* is an ε-minimum and an ε-quasiminimum. Thus an ε-quasiminimum in effect is an exact minimum of a slighlty perturbed objective function over the same constraint set. Thus x^* is a solution of the problem

$$\min_{x \in S}(f(x) + \sqrt{\varepsilon}\|x - x^*\|).$$

It is interesting to note that if a function $f : X \to R$ has an ε-quasiminimum at x^* then the function is *calm from below* at x^* (see Rockafellar and Wets (1998) for the definition of calmness). Conversely if a function is calm from below at x^* with modulus $\alpha > 0$ then x^* is a α-quasiminimum of f in a local sense.

Very recently Huang and Yang (2001) also studied the above two notions of approximate minimum. In the definition of an ε-quasimininimum Huang and Yang (2001) have replaced $\sqrt{\varepsilon}$ with ε. But we will consider the definition as given in Loridan (1982). The above two definitions are in fact motivated by the Ekeland Variational principle which lies at the heart of the study of approximate optimization. We state below the Ekeland Variational principle (see for example Phelps (1993)).

Theorem 2.1 (Ekeland Variational Principle). *Let $f : X \to R \cup \{+\infty\}$ be a lower-semicontinuous function bounded below, where X is a Banach space. Let $\varepsilon \geq 0$ and $x^* \in X$ be such that*

$$f(x^*) \leq \inf_{x \in X} f(x) + \varepsilon.$$

For any $\lambda > 0$ there exists $x_\lambda \in X$ such that

i) $f(x_\lambda) \leq f(x^*)$.

ii) $\|x_\lambda - x^*\| \leq \lambda$.

iii) $f(x_\lambda) \leq f(x) + \dfrac{\varepsilon}{\lambda}\|x - x_\lambda\|$.

Let us now consider the program (MP) where the set S is now described through some inequality and equality constraints. Thus S can be represented as

$$S = \{x \in X : g_i(x) \leq 0, i = 1, \ldots, m, \quad h_j(x) = 0, j = 1, \ldots, r\}.$$

where $g_i : X \to R$ and $h_j : X \to R$. We can further consider the following set. Given $\varepsilon \geq 0$ the ε-feasible set for (MP) is denoted by S_ε is given by

$$S_\varepsilon = \{x \in X : g_i(x) \leq \varepsilon, i = 1, \ldots, m, \quad -\varepsilon \leq h_j(x) \leq \varepsilon, j = 1, \ldots r\}.$$

This leads to the following notion of an almost ε-minimum of (MP) due to Loridan (1982).

Definition 2.3. Let us consider the problem (MP). A point $x_\varepsilon \in S_\varepsilon$ is said to be an *almost-ε-minimum* of (MP) if

$$f(x_\varepsilon) \leq f(x) + \varepsilon, \quad \forall x \in S. \tag{2.3}$$

An important notion related to a convex function is that of the ε-subdifferential. Consider a proper convex function $f : X \to \overline{R}$ and let $x_0 \in dom f$. Then given $\varepsilon \geq 0$ the ε-subdifferential of f at x_0 is given as

$$\partial_\varepsilon f(x_0) = \{\xi \in X^* : f(x) - f(x_0) \geq \langle \xi, x - x_0 \rangle - \varepsilon, \forall x \in X\}.$$

If f is a proper and lower-semicontinuous convex function then it is well known that $\partial_\varepsilon f(x_0) \neq \emptyset$ for all $x \in dom f$ (see for example Phelps (1993)). Moreover for any lower-semicontinuous convex function $f : X \to R$, a point $x_0 \in dom f$ is an ε-minimum of f if and only if $0 \in \partial_\varepsilon f(x_0)$.

The *upper Dini directional derivative* of $f : X \to R$ at x and in the direction v is given as

$$f_d^+(x, v) = \limsup_{\lambda \downarrow 0} \frac{f(x + \lambda v) - f(x)}{\lambda}.$$

If f is locally Lipschitz then $v \mapsto f_d^+(x, v)$ is also locally Lipschitz but need not be convex.

We now need to introduce the notion of the contingent cone or the Bouligand Tangent which is an important tool in expressing optimality conditions. Let $S \subseteq X$ and $x_0 \in S$. An element $v \in X$ is said to be a tangent to the set S at x_0 if there exists a sequence $v_n \to v$ and $t_n \downarrow 0$ such that $x_0 + t_n v_n \in S$. Equivalently we can say that v is a tangent to S at x_0 if there exists a sequence $x_n \to x_0$ with $x_n \in S$ for each n and $\lambda_n > 0$ such that $\lambda_n(x_n - x_0) \to v$. The set of all such v forms a closed cone denoted by $T(S, x_0)$ and is termed as the *contingent cone* or the *Bouligand tangent cone.* If S is convex then $T(S, x_0)$ is convex but in general it need not be convex. Again the convexity of the set S guarantees that $T(S,x_0) = \mathrm{clcone}(S-x_0)$ (for example see Jahn (1996)). For a convex set S the following set

$$N(S,x_0) = \{\xi \in X^* : \langle \xi, x-x_0\rangle \leq 0, \quad \forall x \in S\}$$

is said to be the normal cone to the convex set S at the point x_0.

We will now introduce the various definitions and results from the non-smooth analysis of Mordukhovich.

Definition 2.4. For a set valued map $F : X \to X^*$ we denote by

$$\limsup_{x\to x_0} F(x)$$

the sequential Kuratowski-Painleve upper limit with respect to the norm topology on X and weak-star toplology on X^*, which is given as

$$\limsup_{x\to x_0} F(x) = \{x^* \in X^* : \exists \quad \text{sequences} \quad x_k \to x_0, \quad \text{and} \quad x_k^* \xrightarrow{w^*} x^*, \quad \text{with} \quad x_k^* \in F(x_k), \quad \forall k = 1,2,\ldots\},$$

where $\xrightarrow{w^*}$ denotes convergence in the weak-star topology of X^*.

Definition 2.5 (Mordukhovich and Shao (1996)). Let S be a non-empty subset of X and let $\varepsilon \geq 0$. Given $x \in \mathrm{cl}S$ the non-empty set

$$N_\varepsilon^F(S,x) = \left\{ x^* \in X^* : \limsup_{y\to x, y\in S} \frac{\langle x^*, y-x\rangle}{\|y-x\|} \leq \varepsilon \right\}$$

is called the set of Frechet ε-normals to S at x. When $\varepsilon = 0$ then the above set is a cone is called the set of Frechet normals and is denoted by $N^F(S,x)$.

Let $x_0 \in \mathrm{cl}S$. The non-empty cone

$$N_L(S,x_0) = \limsup_{x\to x_0, \varepsilon\to 0} N_\varepsilon^F(S,x)$$

is called the limiting normal cone or the Mordukhovich normal cone to S at x_0.

It is important to note that the set of Frechet ε-normals is a convex set for every $\varepsilon \geq 0$ but the limiting normal cone is in general non-convex. For more details on limiting normals see for example Mordukhovich (1985) and Mordukhovich and Shao (1996). For a treatment of limiting normals in finite dimension see for example Mordukhovich (1994). When S is a convex set then the limiting normal cone reduces to the standard normal cone of convex analysis which has been defined earlier. Moreover it is important to note that if X is an Asplund space then we have

$$N_L(S, x_0) = \limsup_{x \to x_0} N^F(S, x).$$

Definition 2.6. Let $f : X \to \overline{R}$ be a given function and $x_0 \in \text{dom} f$. The set

$$\partial_L f(x_0) = \{x^* \in X^* : (x^*, -1) \in N_L(\text{epi} f, (x_0, f(x_0))\}$$

is called the limiting subdifferential or the Mordukhovich subdifferential of f at x_0. If $x_0 \notin \text{dom} f$ then we set $\partial_L f(x_0) = \emptyset$.

Moreover if X is an Asplund space and $f : X \to R$ is locally Lipschitz around x_0 then we have

$$\partial^\circ f(x_0) = cl^* \text{co} \partial_L f(x_0),$$

where $\partial^\circ f(x_0)$ denotes the Clarke subdifferential of f at x_0. For more details on the Clarke subdifferential see Clarke (1983).

In Mordukhovich and Shao (1996) there is an interesting construction of an ε-subdifferential for a lower-semicontinuous function which we provide below.

Definition 2.7. Let $f : X \to \overline{R}$ be a lower-semicontinuous function. Then for $\varepsilon \geq 0$ the ε-subdifferential (or the limiting ε-subdifferential) is given as

$$\partial^L_\varepsilon f(x_0) = \limsup_{x \xrightarrow{f} x_0} \partial^F_\varepsilon f(x),$$

where $\partial^F_\varepsilon f(x)$ denotes the Frechet ε-subdifferential of f at $x \in dom f$ is given as follows

$$\partial^F_\varepsilon f(x) = \left\{ x^* \in X^* : \liminf_{u \to x} \frac{f(u) - f(x) - \langle x^*, u - x \rangle}{\|u - x\|} \geq -\varepsilon \right\}.$$

Further $\xrightarrow{f}$ denotes f dependent convergence i.e. $x \to x_0$ and $f(x) \to f(x_0)$.

In Mordukhovich and Shao (1996) it has been mentioned that the above construction is originally due to Jofre et al. (1996) where they studied the ε-convexity of extended real-valued functions in terms of the ε-monotonicity of their subdifferentials. The notion of ε-convexity will be defined in Section 3 of this paper and its use in the study of approximate optimization will be explored. The following result due to Mordukhovich and Shao (1996) makes the ε-subdifferential an important tool to represent optimality conditions for approximate minima in both constrained and unconstrained case.

Lemma 2.1. *Let X be an Asplund space and let $f : X \to \overline{R}$ be lower-semicontinuous around $x_0 \in domf$. Then*

$$\partial_\varepsilon^L f(x_0) = \partial_L f(x_0) + \varepsilon B^*$$

Using Proposition 2.5 and Lemma 6.3 in Mordukhovich and Shao (1996) we have the following lemmas.

Lemma 2.2. *Let X be an Asplund space and let $x_0 \in X$. Let $f_i : X \to \overline{R}$, $i = 1, 2$ be lower-semicontinuous and one of these is Lipschitz near x_0 (i.e. locally Lipschitz at x_0) then one has*

$$\partial_L(f_1 + f_2)(x_0) \subseteq \partial_L f_1(x_0) + \partial_L f_2(x_0).$$

Lemma 2.3. *Let X be an Asplund space and $f : X \to R$ be locally Lipschitz near x_0 and let $g : R \to R$ be locally Lipschitz near $y_0 = f(x_0)$. Then*

$$\partial_L(g \circ f)(x_0) \subseteq \bigcup_{y^* \in \partial_L g(y_0)} \partial_L(y^* f)(x_0)$$

3 Optimality Conditions

In this section we shall discuss various necessary and sufficient conditions for the existence of an approximate minimum of the problem (MP). One of the first detailed study of optimality conditions for approximate optimization was done by Loridan (1982) where he developed necessary conditions for problems with objective functions which are directionally differentiable. He had also studied the Lagrange multiplier rule for almost-ε-minimum for locally Lipschitz functions. Hiriart-Urruty (1982) had developed a Lagrangian multiplier rule for ε-minimization of a convex programming problem in terms of the ε-subdifferentials of the related functions. Strodiot et al. (1983) had also studied Lagrange multiplier rules for the ε-minimization of a convex programming problems in terms of the ε-subdifferential of the functions involved. Very recently Hamel (2001) also developed a Karush-Kuhn-Tucker type condition for a locally Lipschitz program in terms of the Clarke subdifferential of the fuctions involved.

We first present the following simple result without proof.

Proposition 3.1. *Consider the problem* (MP) *where $f : X \to R$ is a locally Lipschitz function. If $x_0 \in S$ is an ε-quasiminimum of f over S then following results hold,*

$$f_d^+(x_0, v) + \sqrt{\varepsilon}\|v\| \geq 0, \quad \forall v \in T(S, x_0)$$

and

$$f^\circ(x_0, v) + \sqrt{\varepsilon}\|v\| \geq 0, \quad \forall v \in T(S, x_0),$$

where $f^\circ(x_0, v)$ denote the Clarke directional derivative of f at x_0 in the direction f (see Clarke (1983)).

We will now introduce the notion of ε-convexity due to Jofre et al. (1996).

Definition 3.1. Given $\varepsilon \geq 0$, the function $f : X \to R$ is called ε-convex if for $x, y \in X$ and $\lambda \in (0, 1)$ we have

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) + \varepsilon\lambda(1 - \lambda)\|x - y\|.$$

Proposition 3.2. *Let us consider the problem* (MP) *where f is locally Lipschitz and $\sqrt{\varepsilon}$-convex over X and S is a convex set. Assume that there exists $x_0 \in S$ such that*

$$f_d^+(x_0, v) + \sqrt{\varepsilon}\|v\| \geq 0, \quad \forall v \in T(S, x_0).$$

Then x_0 is a 4ε-quasiminimum of (MP).

Proof. Since f is $\sqrt{\varepsilon}$-convex we have for any $x \in S$ and $\lambda \in (0, 1)$

$$f(x_0 + \lambda(x - x_0)) \leq \lambda f(x) + (1 - \lambda)f(x_0) + \sqrt{\varepsilon}\lambda(1 - \lambda)\|x - x_0\|.$$

This shows that

$$f_d^+(x_0, x - x_0) \leq f(x) - f(x_0) + \sqrt{\varepsilon}\|x - x_0\|$$

Since S is convex we have $S - x_0 \subseteq T(S, x_0)$. This shows that

$$-\sqrt{\varepsilon}\|x - x_0\| \leq f(x) - f(x_0) + \sqrt{\varepsilon}\|x - x_0\|.$$

This shows that for all $x \in S$ one has

$$f(x_0) \leq f(x) + \sqrt{4\varepsilon}\|x - x_0\|.$$

This proves the result. □

Theorem 3.1. *Let X be an Asplund Space. Consider the problem* (MP) *where $f : X \to R$ is a locally Lipschitz function and S is a closed and proper subset of X. Let $x_0 \in S$ be an ε-minimum for* (MP)*. Then there exists $y_0 \in S$ such that y_0 is also an ε-minimum for* (MP) *such that $\|x_0 - y_0\| \leq 1$ and*

$$0 \in \partial_\varepsilon^L f(y_0) + N_L(S, y_0).$$

If $S = X$ then f can be considered to be an extended-valued proper lower-semicontinuous function bounded below and one has $0 \in \partial_\varepsilon^L f(y_0)$.

Proof. Since $x_0 \in S$ is an ε-minimum of (MP) it is clear that x_0 is an ε-minimum of the unconstrained problem

$$\min_{x \in X}(f(x) + \delta_S(x)),$$

where δ_S is the indicator function of S. Now by considering $\lambda = 1$ in the Ekeland Variational principle one can assert the existence of $y_0 \in S$ such that $\|y_0 - x_0\| \leq 1$ and

$$f(y_0) \leq f(x) + \delta_S(x) + \varepsilon\|x - y_0\|$$

Thus by using Proposition 7.8 in Mordukhovich and Shao (1996) we have

$$0 \in \partial_L(f + \delta_S + \|\varepsilon - y_0\|)(y_0).$$

Now by using Lemma 2.2 and using the fact that $\partial_L \delta_S(y_0) = N_L(S, y_0)$ from Mordukhovich and Shao (1998) we have that

$$0 \in \partial_L f(y_0) + N_L(S, x_0) + \varepsilon B^*.$$

Now by using Lemma 2.1 we have the result.

When $S = X$ and f is proper lower-semicontinuous and bounded below then one can immediately apply the Ekeland's variational principle with $\lambda = 1$ and then use Lemma 2.1 to come to the conclusion. □

Remark 3.1. It is important to observe that the whole emphasis on the study of ε-minimum is to get a better idea about the true minimum of the problem. Let $f : R^n \to R$ be a convex function which is bounded below and let us consider a sequence $\varepsilon_k \to 0$. Let for each k, x_k denote the ε_k-minimum of f. Now by setting $\lambda_k = \sqrt{\varepsilon_k}$ in the Ekeland Variational Principle, we can assert the existence of y_k such that $\|x_k - y_k\| \leq \sqrt{\varepsilon_k}$ and

$$0 \in \partial f(y_k) + \sqrt{\varepsilon_k} B^*$$

Let $y_k \to y_0$, then as $\varepsilon_k \to 0$ by using the fact that ∂f is locally bounded set-valued map with a closed graph, eventually we have

$$0 \in \partial f(y_0)$$

Thus y_0 is a minimum for f. Thus at least in the convex case the true minimum can sometimes be thought of as a limit of a sequence of approximate minimizers.

We will now present the last result of this section where we will consider S to be described by equality and inequality constraints.

Theorem 3.2. *Let X be an Asplund Space. Consider the problem* (MP) *where $f : X \to R$ is locally Lipschitz which is bounded below on X. The set S is given as*

$$S = \{x \in X : g_i(x) \leq 0, i = 1, \ldots, m \quad h_j(x) = 0, j = 1, \ldots, m\},$$

where $g_i : X \to R$ and $h_j : X \to R$ are locally Lipschitz functions. Then given any $\varepsilon > 0$ and x_0 an ε-minimum of (MP) *there exists x^* which is an almost-ε-minimum for* (MP) *with $\|x_0 - x^*\| \leq \sqrt{\varepsilon}$ and there exists $\lambda_i(\varepsilon) \geq 0$, $i = 1, \ldots, m$ and $\mu_j(\varepsilon) \in R$, $j = 1, \ldots, r$ such that*

$$0 \in \partial_L f(x^*) + \sum_{i=1}^{m} \partial_L(\lambda_i(\varepsilon) g_i)(x^*) + \sum_{j=1}^{r} \partial_L(\mu_j(\varepsilon) h_j)(x^*) + \sqrt{\varepsilon} B^*.$$

Proof. The proof follows along the lines of Theorem 5.1 of Loridan (1982). In this direction we shall introduce the penalty function for the problem (MP) and this is given as

$$f_r(x) = f(x) + \sum_{i=1}^{m} \frac{1}{r_i} [\max\{0, g_i(x)\}]^2 + \sum_{j=1}^{r} \frac{1}{k_j} [h_j(x)]^2,$$

where $r_i > 0$ for all i and $k_j > 0$ for all j. Since f is bounded below it is clear the f_r is bounded below and locally Lipschitz on X. Then from the Ekeland Variational Principle we have that there exists x^* such that

$$f_r(x^*) \leq f_r(x) + \varepsilon, \quad \forall x \in X \tag{3.1}$$

and

$$f_r(x^*) \leq f_r(x) + \sqrt{\varepsilon}\|x - x^*\|, \quad \forall x \in X. \tag{3.2}$$

It is clear from (3.1)that

$$f(x^*) \leq f_r(x^*) \leq f_r(x) + \varepsilon, \quad \forall x \in X.$$

Thus for all $x \in S$ we have

$$f(x^*) \leq f(x) + \varepsilon. \tag{3.3}$$

Further observe that $\inf_{x \in X} f(x) \leq f(x^*)$. Now $f_r(x^*) \leq \inf_{x \in S} f(x) + \varepsilon$. This shows us that

$$\sum_{i=1}^{m} \frac{1}{r_i}[\max\{0, g_i(x^*)\}]^2 + \sum_{j=1}^{r} \frac{1}{k_j}[h_j(x^*)]^2 \leq \alpha + \varepsilon, \tag{3.4}$$

where $0 \leq \alpha = \inf_{x \in S} f(x) - \inf_{x \in X} f(x)$. Let us consider $r_0(\varepsilon) = k_0(\varepsilon) = \dfrac{\varepsilon^2}{(\alpha + \varepsilon)}$. Hence for any real positive number $r_p \leq r_0(\varepsilon)$ and $k_p \leq k_0(\varepsilon)$, it is clear from (3.4) that $g_i(x^*) \leq \varepsilon$ for all $i = 1, \ldots, m$ and $-\varepsilon \leq h_j(x^*) \leq \varepsilon$ for all $j = 1, \ldots, r$. This shows that $x^* \in S_\varepsilon$. This along with (3.3) shows that x^* is an almost-ε-minimum of (MP). Now using (3.2), Proposition 7.8 in Mordukhovich and Shao (1996), Lemma 2.2 and Lemma 2.3 we have

$$0 \in \partial_L f(x^*) + \sum_{i=1}^{m} \partial_L \left(\frac{2g_i^+(x^*)}{r_i} g_i \right)(x^*) + \sum_{j=1}^{r} \partial_L \left(\frac{2h_j(x^*)}{k_j} h_j \right)(x^*) + \sqrt{\varepsilon} B^*,$$

where $g_i^+(x^*) = \max\{0, g_i(x^*)\}$, for all i. Hence the result holds with $\lambda_i(\varepsilon) = \dfrac{2g_i^+(x^*)}{r_i} \geq 0$ and $\mu_j(\varepsilon) = \dfrac{2h_j(x^*)}{k_j}$. □

Remark 3.2. It is interesting to note that Mordukhovich and Wang (2002) have characterized an Asplund space in terms of a subdifferential variational principle and used it to develop necessary conditions for the existence of

ε-minimum of a locally Lipschitz program in an Asplund Space. However in this article we use the Ekeland Variational Principle and characterize an almost ε-minimum for a locally Lipschitz program in an Asplund space as is evident from the above theorem. It is further interesting to note that the above theorem provides a much sharper necessary optimality condition in the setting of an Asplund space than that given in Loridan (1982) (Theorem 5.1) in terms of the Clarke subdifferential.

4 Approximate Saddle Points

In this section we shall study the approximate optimality conditions in terms of the well known Lagrangian function. We shall present some saddle point results related to ε-minimization of a convex program.

Let us consider the program (MP) where $f : X \to R$ and the set S is described by inequality constraints and is given as $S = \{x \in X : g_i(x) \leq 0, \quad i = 1, \ldots, m\}$. The Lagrangian function associated with such a program (MP) is defined as a function $L : X \times R^m_+ \to R$ given as

$$L(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle,$$

where $g(x) = (g_1(x), \ldots, g_m(x))$. A point $(x_0, \lambda_0) \in X \times R^m_+$ is said to be an *ε-saddle point* if

$$L(x_0, \lambda) - \varepsilon \leq L(x_0, \lambda_0) \leq L(x, \lambda_0) + \varepsilon,$$

for any $x \in X$ and $\lambda \in R^m_+$.

Loridan (1982) also introduced the notion of an *ε-quasisaddle point* as follows. A point $(x_0, \lambda_0) \in X \times R^m_+$ is said to be an ε-quasisaddle point if

$$L(x_0, \lambda) - \sqrt{\varepsilon}\|\lambda - \lambda_0\| \leq L(x_0, \lambda_0) \leq L(x, \lambda_0) + \sqrt{\varepsilon}\|x - x_0\|,$$

We now present the following result.

Theorem 4.1. *Consider the program* (MP) *given as*

$$\min f(x) \quad \textit{subject to} \quad g_i(x) \leq 0, i = 1, \ldots, m.$$

Let f and each g_i , $i = 1, 2, \ldots, m$ be convex functions and let $\varepsilon \geq 0$ be given. Assume that the Slater Constraint Qualification holds, i.e. there exist $x^* \in X$ such that $g_i(x^*) < 0$ for all $i = 1, \ldots, m$. Then we have the following, *i) Let* $x_0 \in S$ *be an* ε*-minimum of* (MP)*. Then there exists* $\lambda_0 \in R^m_+$ *such that* (x_0, λ_0) *is an* ε*-saddle point of* $L(x, \lambda)$ *and* $\varepsilon + \langle \lambda_0, g(x_0) \rangle \geq 0$.

ii) Let $x_0 \in S$ *be an* ε*-quasiminimum of* (MP)*. Then there exists* $\lambda_0 \in R^m_+$ *such that* (x_0, λ_0) *is an* ε*-quasisaddle point of* $L(x, \lambda)$.

Proof. Since x_0 is an ε-minimum of (MP) the following system

$$(f(x) - f(x_0) + \varepsilon, g(x)) \in -\mathrm{int}\,(R_+ \times R^m_+), \quad x \in X.$$

has no solution. Thus by standard separation argument we have that there exists $(0, 0) \neq (\tau_0, \lambda_0) \in (R_+ \times R^m_+)$ such that

$$\tau_0(f(x) - f(x_0)) + \tau_0\varepsilon + \langle \lambda_0, g(x) \rangle \geq 0, \tag{4.1}$$

for all $x \in X$. Assume now that $\tau_0 = 0$. Since there exists $x^* \in S$ such that $g_i(x^*) < 0$. This shows that we have $\langle \lambda_0, g(x^*) \rangle < 0$. This contradicts (4.1). Thus $\tau_0 \neq 0$ and we may without loss of generality consider that $\tau_0 = 1$. Thus we have

$$f(x) - f(x_0) + \varepsilon + \langle \lambda_0, g(x) \rangle \geq 0. \tag{4.2}$$

If we now put $x = x_0$ then we have from (4.2), $\varepsilon + \langle \lambda_0, g(x_0) \rangle \geq 0$.
From (4.2) we can see that

$$f(x) + \langle \lambda_0, g(x) \rangle + \varepsilon \geq f(x_0). \tag{4.3}$$

Since $\langle \lambda_0, g(x_0) \rangle \leq 0$, (4.3) reduces to

$$f(x) + \langle \lambda_0, g(x) \rangle + \varepsilon \geq f(x_0) + \langle \lambda_0, g(x_0) \rangle.$$

Thus we have

$$L(x_0, \lambda_0) \leq L(x, \lambda_0) + \varepsilon, \quad \forall x \in X.$$

For any $\lambda \in R^m_+$ we clearly have that $\langle \lambda, g(x_0) \rangle \leq 0$. Thus we have

$$f(x_0) + \langle \lambda, g(x_0) \rangle \leq f(x_0).$$

Since $\varepsilon \geq 0$ we have that

$$f(x_0) + \langle \lambda, g(x_0) \rangle - \varepsilon \leq f(x_0) - \varepsilon.$$

Since we have $\varepsilon + \langle \lambda_0, g(x_0) \rangle \geq 0$ we conclude that

$$L(x_0, \lambda) - \varepsilon \leq L(x_0, \lambda_0), \quad \forall \lambda \in R^m_+.$$

The part (*ii*) can also be proved in a similar fashion and hence we omit it. □

Proposition 4.1. *Let us consider the convex program* (MP) *with S as described in the previous theorem. Let $\varepsilon \geq 0$ be given and $x_0 \in S$ be an ε-minimum of* (MP). *Then there exists scalars $\varepsilon_0 \geq 0$ and $\varepsilon_i \geq 0$, $i = 1, \dots, m$ such that $\varepsilon_0 + \sum_{i=1}^{m} \varepsilon_i = \varepsilon$ and scalars $\lambda_i \geq 0$, $i = 1, \dots, m$ such that*

i) $0 \in \partial_{\varepsilon_0} f(x_0) + \sum_{i=1}^{m} \partial_{\varepsilon_i}(\lambda_i g_i)(x_0)$

ii) $\varepsilon + \sum_{i=1}^{m} \lambda_i g_i(x_0) \geq 0$.

Proof. It follows from Theorem 3.1 and the sum rule for ε-subdifferentials (see for example Hiriart-Urruty (1982)). □

Remark 4.1. Let us observe that the necessary conditions stated above need not be sufficient. This fact will be made clear by the next theorem. However if $\sum_{i=1}^{m} \lambda_i g_i(x_0) = 0$ and $x_0 \in S$, then the above condition is also sufficient.

The converse of Theorem 4.1 may not always hold but we have the following for part *i)* of Theorem 4.1 and for a somewhat converse of part *ii)* of Theorem 4.1 see Theorem 6.2 in Loridan (1982).

Theorem 4.2. *Let us consider $(x_0, \lambda_0) \in X \times R_+^m$ such that x_0 is an ε_1-minimum of $L(., \lambda_0)$, $\lambda_0 \in R_+^m$ and λ_0 is an ε_2-maximum of $L(x_0, .)$, $x \in X$. Then x_0 is an almost-$(\varepsilon_1 + \varepsilon_2)$-minimum of* (MP).

Proof. Since $\lambda_0 \in R_+^m$ is an ε_2-maximum of $L(x_0, \lambda)$ we have that for any $\lambda \in R_+^m$,

$$L(x_0, \lambda) - \varepsilon_2 \leq L(x_0, \lambda_0).$$

This reduces to

$$\langle (\lambda - \lambda_0), g(x_0) \rangle \leq \varepsilon_2, \quad \forall \lambda \in R_+^m. \tag{4.4}$$

We claim that $x_0 \in S_{\varepsilon_2}$, where $S_{\varepsilon_2} = \{x \in X : g_i(x) \leq \varepsilon_2, \quad \forall i = 1, \dots, m\}$. On the contrary let us assume that $x_0 \notin S_{\varepsilon_2}$. Thus we have $\varepsilon_2 e - g(x_0) \notin R_+^m$, where $e = (1, \dots, 1) \in R_+^m$. Thus by standard separation argument we can deduce that there exists $0 \neq p \in R_+^m$ such that

$$\langle p, \varepsilon_2 e \rangle - \langle p, g(x_0) \rangle < 0.$$

Since $p \neq 0$ we can consider without loss of generality that $\langle p, e \rangle = 1$. Thus we have that

$$\varepsilon_2 - \langle p, g(x_0) \rangle < 0.$$

Hence we have that

$$\langle p, g(x_0) \rangle > \varepsilon_2 \tag{4.5}$$

Now as $\lambda_0 \in R^m_+$ and $p \in R^m_+$ it is clear that $\lambda_0 + p \in R^m_+$. Thus from we (4.4) we have that

$$\langle p, g(x_0) \rangle \leq \varepsilon_2$$

This is clearly a contradiction to (4.5). Hence $x_0 \in S_{\varepsilon_2}$ and thus $x_0 \in S_{\varepsilon_1 + \varepsilon_2}$, where

$$S_{\varepsilon_1+\varepsilon_2} = \{x \in X : g_i(x) \leq \varepsilon_1 + \varepsilon_2, \quad \forall i = 1, \ldots, m\}.$$

Now as x_0 is an ε_1-minimum of $L(x, \lambda_0)$ we have for all $x \in X$,

$$f(x) - f(x_0) + \langle \lambda_0, g(x) - g(x_0) \rangle + \varepsilon_1 \leq 0.$$

For any $x \in S$, where S is the feasible set of (MP), we have $\langle \lambda_0, g(x) \rangle \leq 0$ and thus the above expression reduces to

$$f(x) - f(x_0) - \langle \lambda_0, g(x_0) \rangle + \varepsilon_1 \geq 0.$$

Again from (4.4), setting $\lambda = 0$ we have that $\langle \lambda_0, g(x_0) \rangle \geq -\varepsilon_2$. This clearly shows that

$$f(x) - f(x_0) \geq -(\varepsilon_1 + \varepsilon_2), \quad \forall x \in S.$$

This proves the result. □

References

Bustos M. (1989). Solution Approchées de Problèmes de Calcul des Variations. *Revista de Matemáticas Aplicadas* 10, 95-113.

Bustos M. (1994). ε-Gradients pour les Fonctions Localements Lipschitziennes et Applications. *Numerical Functional Analysis and Applications* 15, 435-453.

Clarke F.H. (1983). *Optimization and Nonsmooth Analysis.* Wiley-Interscience.

Hamel A. (2001). An ϵ-Lagrange Multiplier Rule for a Mathematical Programming Problem on Banach Spaces. *Optimization* 49, 137-149.

Hiriart-Urruty J.-B. (1982). ε-Subdifferential Calculus. In: Aubin J.-P and Vinter R.B. (eds.), *Convex Analysis and Optimization*, Research Notes in Mathematics Series 57. Pitman, 43-92.

Huang X.X. and Yang X.Q. (2001). Approximate Optimal Solutions and Nonlinear Lagrange Functions. *Journal of Global Optimization* 21, 51-65.

Jahn J. (1996). *Introduction to the Theory of Nonlinear Optimization.* Springer Verlag.

Jofre A., Luc D. T. and Thera M. (1996). ε-Subdifferential Calculus for Nonconvex Functions and ε-Monotonicity. *Comptes Rendus de l'Académie des Sciences* 323, 735-740.

Loridan P. (1982). Necessary Conditions for ε-Optimality. *Mathematical Programming Study* 19, 140-152.

Loridan P. and Morgan, J. (1983). Penalty Functions in ε-Programming and ε-Minimax Problems. *Mathematical Programming* 26, 213-231.

Liu J. C. (1991). ε-Duality Theorems for Non-Differentiable, Non-Convex Multiobjective Programming. *Journal of Optimization Theory and Applications* 69, 153-167.

Mordukhovich B. (1985). On Necessary Conditions for an Extremum in Nonsmooth Optimization. *Soviet Math Doklady* 32, 215-220.

Mordukhovich B. (1994). Generalized Differential Calculus for Nonsmooth and Set-Valued Mappings. *Journal of Mathematical Analysis and Applications* 183, 250-288.

Mordukhovich B. and Shao Y. (1996). Nonsmooth sequential Analysis in Asplund Spaces. *Transactions of American Mathematical Society* 348, 1235-1280.

Mordukhovich B. and Wang B.(2002). Necessary Suboptimality and Optimality Conditions via Variational Principles. *SIAM Journal of Control and Optimization* 41, 623-640.

Phelps R.R. (1993). *Convex Functions, Monotone Operators and Differentiability. Lecture Notes in Mathematics*, 1364. Springer Verlag.

Rockafellar R.T. and Wets R.J.B. (1998). *Variational Analysis.* Springer Verlag.

Strodiot J.-J., Nguyen V.H. and Heukemes N. (1983). ε-Optimal Solutions in Nondifferentiable Convex Programming and Some Related Questions. *Mathematical Programming* 25, 307-328.