Journal of Global Optimization **12:** 415–434, 1998.

Global Optimality Conditions for Nonconvex Optimization

ALEXANDER S. STREKALOVSKY
Mathematical Faculty, Irkutsk State University, Irkutsk, Russia
e-mail:strekal@math.isu.runnet.ru

(Received 15 September 1994 accepted 2 September 1997)

Abstract. In this paper we give an analytical equivalent for the inclusion of a set to the Lebesque set of a convex function. Using this results, we obtain global optimality conditions (GOC) related to classical optimization theory for convex maximization and reverse-convex optimization. Several examples illustrate the effectiveness of these optimality conditions allowing to escape from stationary points and local extremums.

Key words: Lebesque set, Normal cone, Extreme points, Subdifferential.

1. Introduction

Nowadays there is a renewal of interest and a revival of activity in global optimization (GO) [2–16] due to two kinds of incentives: more precise demands from the world of applications, and more pertinent theoretical contributions from mathematicians. However, in contrast to the fact, that at present there is a huge number of papers devoted to GO (see, for example, the references in [8]), there does not exist a generally accepted GO conditions theory allowing to construct GO algorithms capable of solving large dimensional GO problems, although a hope to do it is glimmering yet. Taking into account the situation in global optimization we decided to inform the Western 'GO society' on the approach proposed in [11–16], since the papers [12–14] (in Russian) are not fully available to Western mathematicians. Surprisingly enough the simple notion of convexity plays a crucial role in the development of the theory and the practice of optimization during the second part of the 20th century. As it is well-known the Fermat and Lagrange principles remain basic for solution methods in mathematical programming, mainly for convex problems, where every local solution turns out to be global one. Nevertheless, this classical apparatus turns out to be inoperative for nonconvex problems where there may be 'many' local extremums which differ from global ones as it is in 'reverse convex' optimization, where convexity is present, but in a reverse sense (concave minimization or convex maximization, reverse convex problems and d.c. programming problems). Regardless of the importance of classical theory, the development of 'reverse convex optimization' took the way of branch and bounds, bisection's and cut's ideas etc., which stand further from the Classical Principles of Extremum

Theory. However, there may exist a way for GO keeping on the classical road. In order to analyze the situation it would be pertinant to recall a few basic features of Classical Optimality Conditions (GOC).

To begin with, consider the characterization of a global solution $z \in D$ to a convex problem

$$f(x) \downarrow \text{min}, \quad x \in D, \tag{1}$$

($f(\cdot)$ and D are convex) given by the condition

$$< \nabla f(z), x - z > \geqslant 0 \quad \forall x \in D. \tag{2}$$

This means that in order to verify whether a point z is a global solution to (1), we have to solve the linearized problem

$$< \nabla f(z), x > \downarrow \text{min}, \quad x \in D; \tag{3}$$

and after this, we have to verify the inequality

$$< \nabla f(z), x(z) - z > \geqslant 0, \tag{4}$$

where $x(z)$ is a solution of (3).

Hence, the meaning of optimality condition (OC) (2) in particular consists in reducing the original problem (1) to a simplier one (3) (with a linear objective function).

Secondly, it is well-known that if Inequality (4) is violated, that is,

$$< \nabla f(z), x(z) - z > < 0,$$

then one can form a feasible point $x(\alpha) = \alpha x(z) + (1 - \alpha)z, \alpha \in]0, 1[$, which is better than $z : f(x(\alpha)) < f(z)$. In other words, OC (2) allows us to decide whether a feasible point z is a global solution to (1) or not, and if not, (2) enables us to construct a better feasible point. In the sequel, this property of the OC will be called the Algorithmic Property.

It can be readily seen that the classical local OC:

$$< \nabla f(z), x - z > \leqslant 0 \quad \forall x \in D, \tag{5}$$

for a convex maximization problem

$$f(x) \uparrow \text{max}, \quad x \in D, \tag{6}$$

(where f and D are convex) concerves the Algorithmic Property of OC (2). The same can be said about Rockafeller's OC (1970 [2])

$$\partial f(z) \subset N(z/D), \tag{7}$$

since (7) can be expressed as

$$\forall z^* \in \partial f(z) :< z^*, x - z > \leqslant 0 \quad \forall x \in D. \tag{7'}$$

As far as we know, (7) was the first OC which distinguishes itself from classical OC that have the character of nonempty intersections (or of existence)

$$\partial f(z) \cap N(z/D) \neq \emptyset. \tag{8}$$

The latter condition is generated by the geometry of intersections

$$D_0 \cap D = \emptyset, \tag{9}$$

where $D_0 = \{x \in R^n / f(x) > f(z)\}$ for (6).

It is clear that if z is a global solution of (6), then (9) holds. But for arbitrary sets D_0 and D we are not able to give any analytical equivalent to (9) which can lead us to OC with the Algorithmic Property. By replacing D_0 and D by corresponding convex conic approximations K_0 and K_1 [2–6] we derive from (9)

$$K_0 \cap K_1 = \emptyset. \tag{10}$$

When we apply an appropriated separation theorem at this point we get (8). Hence, one can say that the classical OC theory is based on Separation Theorems [2–7]. However, note that Rockafeller's OC (7) cannot be obtained in this way and we must be very careful to interpret (7) as a classical OC. On the other hand, as it was noted long ago (see the references in [5–11]) for the reverse convex problems (as in (6)) the geometry of intersections ((9) and (10)) can not aid in an analytical characterization of a global solution. Thus, it would be resonable to abandon the geometry of intersections and to pass to more suitable in the case geometry of inclusions. For instance, for (6) we have

$$D \subset C, \tag{11}$$

where C is the complement of D_0:

$$C = \{x \in R^n / f(x) \leqslant f(z)\}.$$

However, when changing the geometry, we need another analytical apparatus, different from the separation idea and appropriated to the inclusion geometry. Recall several results from Convex Analysis that give various analytical criterions of inclusion to a closed convex set.

Let Ω^0 be the polar to a set $\Omega \subset R^n$, $0 \in \Omega$

$$\Omega^0 = \{x^* \in R^n :< x^*, x >\leqslant 1 \quad \forall x \in \Omega\},$$

and $\sigma(x^*/S)$ be the support function to $S \subset R^n$:

$$\sigma(x^*/S) = \sup_x \{< x^*, x > /x \in S\}.$$

It is well-known [2, 5] that if C is a closed convex set and D is an arbitrary set in R^n, $z \in C \cap D$, then the inclusion $D \subset C$ is equivalent to each of the dual conditions

$$(C - z)^0 \subset (D - z)^0, \tag{12}$$

$$\sigma(x^*/D) \leqslant \sigma(X^*/C) \quad \forall x^* \in R^n : ||x^*|| = 1. \tag{13}$$

Further, using the useful apparatus of supporting hyperplanes one obtains that any closed convex set C in R^n is the intersection of all closed halfspaces generated by the supporting hyperplanes to the set C and containing this set. As a consequence, an arbitrary set $D \subset R^n$ is contained in a closed convex set C iff D belongs to every supporting halfspace to the set C. We prefer to express this under the form [11–16] more appropriated to the theory of OC. Given two nonempty sets D and C in R^n, one of which (assume C) is closed and convex. Then the inclusion $D \subset C$ holds iff the dual inclusion

$$N(y/C) \subset N(y/D) \quad \forall y \in bdC \tag{14}$$

takes place. Here bdC is the boundary of C.

It can be readily seen that the dual inclusions (12)–(14) are related to each other and that they are almost obviously equivalent. Nevertheless, (14) works only with boundary points of C ((14) stresses this fact) and not with the global information about C and D given by (12) and (13). On the other hand, we would not consider (12)–(14) and, moreover, the inclusion $D \subset C$ as some OC, because these results exist independently of any extremum problem and have their own value in Convex Analysis [1–10]. Besides, the dual inclusions (12)–(14) play the same role for non-convex problems as do Separation Theorems and the Minkowski-Farkash Theorem in the Classical Extremum Theory. That is, they provide the analytical apparatus characterizing the geometrical fact (10). However, nobody estimates Separation Theorems as OC. Then the following question is natural. What can be called an OC for a problem of interest? To our opinion, it should be an analytical expression satisfying the following requirements:

a) This must be an analytical condition involving only the data of the problem under investigation.
b) It must have a relation with Classical Extremum Theory.
c) This condition must reduce the original problem to a problem or a family of problems, which are, in a sense, simpler to solve than the original problem.
d) This condition must possess the Algorithmic Property, that is, in the case when the condition is violated, there is a rule allowing to construct a feasible point which is better than the point where the condition does not hold.

For instance, it is easy to see that OC (2) satisfies all four properties (a)-(d) for Problem (1). Similarly, the KKT-OC

$$\left.\begin{aligned} &\nabla f(z) + \textstyle\sum_1^m \lambda_i \nabla g_i(z) = 0, \\ &\lambda_i g_i(z) = 0, \quad \lambda_i \geqslant 0, \end{aligned}\right\} \tag{15}$$

for the mathematical programming problem

$$f(x) \downarrow \min, \quad g_i(x) \leqslant 0, i = 1, \ldots, m, \tag{16}$$

satisfies the requirements (a)-(d).

Actually, one may say that Problem (16) is reduced (by means of (15)) to solving the linearized problem

$$\left.\begin{aligned} &< \nabla f(z), x \downarrow \text{min}, \\ &< \nabla g_i(x), x > \leqslant 0, i \in I; \end{aligned}\right\}$$

$$I = \{i \in \{1, \ldots, m\} / g_i(z) = 0\}$$

which is obviously simpler than the original Problem (16), and so on.

Similarly, the GOC of Hiriart-Urruty for Problem (6) using the ϵ-subdifferential and the ϵ-normal cone [5–7]

$$\partial_\epsilon f(z) \subset N_\epsilon(z|D), \quad \forall_\epsilon \geqslant 0, \tag{17}$$

possesses all the features (a)-(d). Clearly, when $\epsilon = 0$, we readily obtain (7).

In [11–16] another approach was proposed for constructing GOC. For instance, if z is a global solution to Problem (6), then

$$\left.\begin{aligned} &\forall y : f(y) = f(z), && \forall y^* \in \partial f(y) \\ &\langle y^*, x - y\rangle \leqslant 0, && \forall x \in D \end{aligned}\right\} \tag{18}$$

or (which is equivalent)

$$\partial f(y) \in N(y|D), \quad \forall y : f(y) = f(z). \tag{18$'$}$$

By setting $y = z$ we again obtain Rockafellar's Condition (7) which characterizes only local maxima in (6). However, under the assumption

$$\exists v \in R^n : -\infty < f(v) < f(v) < +\infty$$

OC (18) turns out to be sufficient for z to be global maximizer to (6). It can be readily seen that a difficulty of using Condition (18) consists in the choice of a point y or several points y^i, $i = 1 \ldots r$, on the level surface $U = \{y \in R^n | f(y) = f(z)\}$. By analyzing GOC (18) it is easy to note that in order to escape from a stationary point z, we have to solve the linearized problems (in the smooth case)

$$\langle f'(y^i), x\rangle \to \text{max},\ x \in D;$$

and subsequently we have to verify the inequalities

$$\langle f'(y^i), \bar{x}^i - y^i\rangle \leqslant 0,$$

where $\bar{x}^i \in D$ is a solution to corresponding linearized problem.

In order to facilitate the choice of these points y^i for the case of the functions $f(\cdot)$ with the compact Lebesque set $S(f, z)$ we here propose to use only extreme points of $S(f, z)$. Hence, this paper aims at several objectives. The first is to give an introduction to the results from [11–16] and to future papers that we intend to publish. On the other hand, we display a 'geometrical' proof of the GO theory constructed in [11–16]. This proof is completely different from those presented in [15] and [16] and supplies convex analysis tools concerning the inclusions into a convex set.

In addition, Proposition 1 (see below) describes the elements from the polar $(S-z)^0$ of the Lebesque set $S = S(f, z)$ of a convex function $f : X \to R \cup \{+\infty\}$. Using this result, Proposition 2 gives the analytical characterization of the geometrical fact of inclusion of a set into the Lebesque set $S(f, z)$. This makes it possible to characterize a global solution for convex maximization problems (Section 2) and the reverse-convex optimization (Section 3) which is shown by the examples (Sections 2 and 3).

2. Polars and sets inclusions

Let X and X^* be a pair of dual locally convex linear topological spaces (see [1]). Define the polars A° and B° of sets $A \subset X$ and $B \subset X^*$ respectively, as

$$A^\circ = \{x^* \in X^* | \langle x^*, x\rangle \leqslant 1 \quad \forall x \in A\},$$
$$B^\circ = \{x \in X | \langle x^*, x\rangle \leqslant 1 \qquad \forall x^* \in B\},$$

and the bipolars by the equalities $A^{\circ\circ} = (A^\circ)^\circ, B^{\circ\circ} = (B^\circ)^\circ$.

It is well-known [2] that the polar is always the convex and closed set containing 0. Therefore, if $0 \in A$ and A is convex and closed then $A^{\circ\circ} = A$. In the general case (see [2–5]) one can write $A^\circ = D^\circ$, where

$$D = \text{cl}[\text{co}(A \cup \{0\})],$$

and since $D^{\circ\circ} = D$, we have

$$A^{\circ\circ} = \text{cl}[\text{co}(A \cup \{0\})].$$

Further, it is clear that if $A, C \subset X, C \neq \emptyset$ and $A \subset C$ then $C^\circ \subset A^\circ$. The inverse is not trivial.

LEMMA 1 [2]. *Suppose sets A, $C \subset X$, $A \neq \emptyset$, C is convex and closed and $z \in C$. Then the following two inclusions are equivalent:*

$$a) \quad A \subset C;$$

$$b) \quad (C - z)^\circ \subset (A - z)^\circ.$$

□

Denote by $N(y|A)$ the set

$$\{x^* \in X^* | \langle x^*, x - y\rangle \leqslant 0 \quad \forall x \in A\}$$

(even if y does not belong to A) and

$$M(z|A) = \{x^* \in (A-z)^\circ| \quad \exists u \in A : \langle x^*, u-z\rangle > 0\}.$$

Then, one has

$$(A-z)^\circ = N(z|A) \cup M(z|A) \tag{19}$$

and the assertion of below turns out to be almost obvious.

LEMMA 2. *Let the conditions of Lemma 1 hold. Then $A \subset C$ iff*

$$\left.\begin{array}{ll} i) & N(z|C) \subset N(z|A), \\ ii) & M(z|C) \subset (A-z)^\circ. \end{array}\right\} \tag{20}$$

□

Further, consider a proper convex function $f : X \to R \cup \{+\infty\}$ (see [2–5]) s.t.

$$-\infty \leqslant \inf(f, X) < f(z) < +\infty. \tag{21}$$

Assume, that its Lebesque set

$$S(f,z) \triangleq \{x \in X | f(x) \leqslant f(z)\} \subset \text{ intdom} f \tag{22}$$

is compact. Note that closed, compact, etc., are understood in the weak topology $\sigma(X, X^*)$ [1].

PROPOSITION 1. *Let $y^* \in X^*$ and suppose there exists $u \in S \triangleq S(f,z)$ s.t.*

$$\langle y^*, u-z\rangle > 0.$$

Then the inclusion $y^* \in (S-z)^\circ$ holds iff there exists an extreme point $y \neq z$ of S and a number $\lambda > 0$ s.t.

$$\left.\begin{array}{l} f(y) = f(z), \quad \lambda y^* \in \partial f(y), \\ 0 < \langle y^*, y-z\rangle = \max_x\{\langle y^*, x-z\rangle | x \in S\} \leqslant 1. \end{array}\right\} \tag{23}$$

Proof. a) If y^* satisfies (23), then

$$0 \geqslant f(x) - f(y) \geqslant \langle \lambda y^*, x-y\rangle \quad \forall x \in S.$$

Since $\lambda > 0$, we have

$$\langle y^*, y\rangle \geqslant \langle y^*, x\rangle \quad \forall x \in S.$$

Therefore, due to the inequalities in (23) we obtain

$$1 \geqslant \langle y^*, y-z\rangle \geqslant \langle y^*, x-z\rangle \quad \forall x \in S.$$

Thus, $y^* \in (S-z)^\circ$.

b) Necessity. 1) Let $y^* \in (S-z)^\circ$ and $\exists u \in S : \langle y^*, u-z\rangle > 0$. Since S is compact, there exists $y \in S$:

$$\langle y^*, y\rangle = \max_x\{\langle y^*, x\rangle | x \in S\}. \tag{24}$$

Hence,

$$0 < \langle y^*, u-z\rangle \leqslant \max_x\{\langle y^*, x-z\rangle | x \in S\} = \langle y^*, y-z\rangle \leqslant 1.$$

2) It is well-known (see [8]) that the maximum of a convex function over convex sets is attained at an extreme point. Therefore, y is an extreme point of S and $f(y) = f(z)$.

3) Introduce two sets

$$A = \{(x,\eta) \in X \times R | f(x) - f(z) \leqslant \eta\},$$
$$B = \{(x,0) \in X \times R | \langle y^*, x-y\rangle \geqslant 0\}.$$

With the help of the Conditions (21) and (22) it is easy to see that A and B are convex and nonempty and besides

$$\text{int}A \triangleq \{(x,\eta) | x \in \text{ intdom} f, f(x) < f(z) + \eta\} \neq \emptyset. \tag{25}$$

Let us show that $B \cap \text{int}A = \emptyset$. Actually, if there exists $(w, 0) \in B \cap \text{int}A$, then

$$\langle y^*, w-y\rangle > 0 > f(w) - f(z).$$

However, due to (24) the latter is impossible because $w \in S$ and $\langle y^*, w\rangle > \langle y^*, y\rangle$.

4) Hence, the sets A and B are separable (see [1–5]), i.e., $\exists(z^*, \beta) \neq 0 \in X^* \times R$, $\exists \gamma \in R$, such that,

$$\left.\begin{array}{lll} \text{i)} & \langle z^*, x\rangle + \beta\eta \leqslant \gamma & \forall (x,\eta) \in A, \\ \text{ii)} & \langle z^*, x\rangle + \beta\eta < \gamma & \forall (x,\eta) \in \text{int}A, \\ \text{iii)} & \langle z^*, x\rangle \geqslant \gamma & \forall x : \langle y^*, x-y\rangle \geqslant 0. \end{array}\right\} \tag{26}$$

If $\beta > 0$ we derive from (26, i) with $x = y$, that

$$\langle z, y\rangle + \beta\eta \leqslant \gamma, \forall \eta : \eta \geqslant f(y) - f(z) = 0,$$

which is impossible.

If $\beta = 0$ it results from (26, ii and iii) that

$$\langle z^*, x\rangle < \gamma \quad \forall x \in \text{ intdom} f,$$
$$\langle z^*, x\rangle \geqslant \gamma \quad \forall x : \langle y^*, x-y\rangle \geqslant 0.$$

Since $\{x | f(x) \leqslant f(z)\} \subset \text{ intdom} f,$ then $\exists \lambda > 0 : v \triangleq y + \lambda(u-z) \in \text{ intdom} f$ and therefore,

$$\langle y^*, v-y\rangle = \lambda\langle y^*, u-v\rangle > 0.$$

Hence, $\langle z^*, v\rangle \geqslant \gamma$. On the other hand, $\langle z^*, v\rangle < \gamma$, since $v \in$ intdom f. Thus, the case when $\beta = 0$ is also impossible and therefore, $\beta < 0$.

5) By dividing (26) by $|\beta|$ and setting $x^* = |\beta|^{-1} z^*, \alpha = \gamma|\beta|^{-1}$, we have

$$\left.\begin{array}{lll} \text{i)} & \langle x^*, x\rangle - \eta \leqslant \alpha & \forall (x,\eta) \in A, \\ \text{ii)} & \langle x^*, x\rangle - \eta < \alpha & \forall (x,\eta) \in \text{int}A, \\ \text{iii)} & \langle x^*, x\rangle \geqslant \alpha & \forall x : \langle y^*, x - y\rangle \geqslant 0. \end{array}\right\} \tag{27}$$

If $x^* = 0$ it follows from (27, ii and iii) that

$$-\eta < \alpha \leqslant 0 \quad \forall \eta : \exists x \in \text{ intdom} f, \quad f(x) - f(z) < \eta.$$

Due to (21) there exists $x_1 \in D : f(x_1) < f(z)$. Therefore, $\exists \eta : 0 > \eta > f(x_1) - f(z)$ and then $0 < -\eta < \alpha \leqslant 0$ which is impossible. Hence, $x^* \neq 0$.

6) By setting $x = y$, $\eta = f(y) - f(z) = 0$ in (27, i and ii) we obtain $\alpha \leqslant \langle x^*, y\rangle \leqslant \alpha$, that is, $\alpha = \langle x^*, y\rangle$. Furthermore, since $x^* \neq 0 \neq y^*$ and by virtue of (27, iii) one has

$$\langle x^*, x\rangle \geqslant \langle x^*, y\rangle \quad \forall x : \langle y^*, x\rangle \geqslant \langle y^*, y\rangle.$$

Hence, there exists $\lambda > 0 : y^* = \lambda x^*$. On the other hand, it follows from (27, i) that

$$\langle x^*, x - y\rangle \leqslant \eta \quad \forall (x,\eta) : f(x) - f(z) \leqslant \eta.$$

By setting $\eta = f(x) - f(z) = f(x) - f(y)$ we have

$$\langle x^*, x - y\rangle \leqslant f(x) - f(y) \quad \forall x \in \text{ intdom } f.$$

Hence, $x^* \in \partial f(y)$ and the proof is completed. □

Now we are in the position to obtain the basic result for the sequel of the paper. Let Ext$C$ be the set of extreme points of C.

PROPOSITION 2. *Let $f : X \to R \cup \{+\infty\}$ be a proper convex function whose Lebesque set $S(f,z) = \{x \in X | f(x) \leqslant f(z)\}$ is compact.*

If a nonempty set $A \subset X$ belongs to $S(f,z)$ i.e.,

$$A \subset \{x \in X | f(x) \leqslant f(z)\}, \tag{28}$$

then

$$\partial f(y) \subset N(y|A) \quad \forall y \in \text{Ext } S(f,z). \tag{29}$$

If, in addition, Assumption (21) holds, then Condition (29) implies Inclusion (28).

Proof. 1) Since $\forall y \in$ Ext $S(f,z)$, equality $f(y) = f(z)$ holds and from (28) we derive

$$\forall y : f(y) = f(z), \quad \forall x^* \in \partial f(y), \quad \forall x \in A,$$
$$0 \geqslant f(x) - f(z) = f(x) - f(y) \geqslant \langle x^*, x - y\rangle.$$

This means that $x^* \in N(y|A)$.

2) If (21) holds then according to Lemma 2 Inclusion (28) is equivalent to (20) where $C = S(f,z) \triangleq S$. Besides, in virtue of the known (see [3]) equality $N(z|S) =$ cone $\partial f(z)$ which holds due to (21), from (20, i) we derive $\partial f(z) \subset N(z|A)$. Now consider the second Inclusion (20, ii), $M(z|S) \subset (A-z)^\circ$. According to Proposition 1 the latter inclusion can be expressed as follows. For all $y \in$ Ext S and for all $x^* \in \partial f(y)$ and some $\alpha > 0$, such that,

$$\langle \alpha x^*, y - a\rangle = 1, \tag{30}$$

the inequality

$$\langle \alpha x^*, x - z\rangle \leqslant 1 \quad \forall x \in A \tag{31}$$

takes place.

Subtracting (30) from (31) and dividing by $\alpha > 0$ we have $\langle x^*, x - y\rangle \leqslant 0 \quad \forall x \in A$, i.e. $x^* \in N(y|A)$. The proof is completed. □

3. Convex maximization over a feasible set

This section is devoted to the investigation of the convex maximization problem

$$f(x) \to \max, \quad x \in D, \tag{P}$$

where $f : X \to R \cup \{+\infty\}$ is a convex closed function and D is a set from X, co$D \neq X$. Let us denote $D_o = \{x \in X | f(z) < f(x)\}$ and assume that

$$C \triangleq X \backslash D_o = \{x \in X | f(x) \leqslant f(z)\} \subset \text{intdom}\, f, \tag{32}$$
$$C \text{ is compact.} \tag{33}$$

THEOREM 1. *If* $z \in D$ *is a global maximizer of Problem (P)* ($z \in$ Argmax$(f,D) =$ Argmax(P)) *then*

$$\left.\begin{array}{l} \forall y : f(y) = f(z), y \in \text{Ext}C, \quad \forall y^* \in \partial f(y), \\ \langle y^*, x - y\rangle \leqslant 0 \quad \forall x \in D. \end{array}\right\} \tag{34}$$

If in addition,

$$\exists v \in X : f(v) < f(z) < +\infty, \tag{35}$$

the Condition (34) becomes sufficient for z *to be a global solution to (P).*

Proof. It is easy to see that the inclusion $z \in$ Argmax(P) is equivalent to $D_o \cap D = \emptyset$, or $D \subset C$. By using the convexity and the compactness of C and with the help of Proposition 2 we obtain the assertion of the theorem. □

Remarks. 1) One can see that Assumption (35) is pertinent for the sufficiency. Actually, for example, if $X = R^n$ and f is differentiable then all global minimizers of f over the whole space R^n satisfy the condition

$$f'(z) = 0,$$

and, as a consequence, Condition (34) trivially holds.

2) Obviously, Assumption (35) is equivalent to

$$\partial f(y) \cap \{0\} = \emptyset \quad \forall y : f(y) = f(z) \tag{36}$$

which in the differentiable case gives

$$f'(y) = 0 \quad \forall y : f(y) = f(z). \tag{36'}$$

3) From (34) with $y = x$ and for differentiable f we obtain the well-known local OC

$$\langle f'(z), x - z \rangle \leqslant 0 \quad \forall x \in D, \tag{37}$$

usually proved for convex D. Here we do not need the convexity of D. Hence, GOC (34) is connected with the Classical Extremum Theory.

4) In the non-smooth case it follows from (34) with $y = z$ that

$$\partial f(z) \subset N(z|D). \tag{38}$$

This necessary local OC was proved in [2] for a convex D. The convexity of D is not obligatory here. In addition, it is clear that (38) is a particular case of the more general and more informative Optimality Condition (34) in which the inclusion $\partial f(y) \subset N(y|D)$ holds $\forall y \in$ Ext $S(f, z)$ (consequently, $f(y) = f(z)$).

EXAMPLE 1. Let, in $(P), X = R, f(x) = (x^2 - 2)$ and $D = [-2, -1]$. It is easy to see that the 'classical' OC (36) holds at two points $z_1 = 1, z_2 = -2$.

a) Obviously, there only exists one point $y_1 = -1, y_1 \neq z_1$ and $y_1 \in$ Ext $S(f, z_1)$. But for $u = -1\frac{1}{2}$ Condition (34) is violated:

$$\langle f'(y_1), u_1 - y_1 \rangle = -2.(-\tfrac{1}{2}) > 0.$$

Consequently, z_1 is not a global solution.

b) $z_2 = -2$. As above, there exists the unique point $y_2 = 2 \neq z_2 : y_2 \in$ Ext $S(f, z_2)$. However, the OC holds.

$$\langle f'(y_2), x - y_2 \rangle \leqslant 0 \quad \forall x \in \text{co}D.$$

It means that z_2 is the global solution.

Note, that the verification of Optimality Condition (34) is rather difficult. In order to make (34) more manageable it is possible to transform it into another form.

THEOREM 2. *Let the assumptions of Theorem 1 hold. In order for a point z to be a global solution to Problem (P), it is necessary, and with the Assumption (35) sufficient, that*

$$\forall y \in \text{Ext}\, S(f,z), \forall y^* \in \partial f(y),$$

and for every maximizing sequence $\{x^k\}$ of the linearized problem

$$\langle y^*, x\rangle \to \max,\ x \in D, \tag{39}$$

the following condition holds:

$$\lim_{k\to\infty} \langle y^*, x^k - y\rangle \leqslant 0. \tag{40}$$

□

It can be readily seen that Problem (39) is simpler than the original problem (P), since the objective function of (39) is linear. For instance, in the case of convex D, Problem (39) turns out to be convex. Hence, in this case, Problem (39) turns out to be solvable with the aid of standard optimization methods [5]. On the other hand, the investigation of Problem (39)

$$\forall y \in\ \text{Ext}\, S(f,z), \quad \forall y^* \in \partial f(y),$$

is a hard task. However, in order to convince ourselves that z is not a global maximizer to (P), it suffices to find a single triplet (y, y^*, u)

$$y \in\ \text{Ext}\, S(f,z), \quad y^* \in \partial f(y), \quad u \in D,$$

such that,

$$\langle y^*, u - y\rangle > 0.$$

This enables us, as the examples below show, to simplify the investigation of a local solution considerably.

EXAMPLE 2. Consider the problem (P), where $x \in R^2$,

$$f(x) = \max\{|x_1|, |x_2|\} - 1, \quad D = \bigcap_{i=1}^{4} D_i,$$

$$\begin{aligned}
D_1 &= \{x = (x_1,x_2) | 0 \leqslant x_i \leqslant 1, \quad i = 1,2\},\\
D_2 &= \{x = (x_1,x_2) | -\tfrac{1}{2} \leqslant x_1 \leqslant 0, \quad 0 \leqslant x_2 \leqslant \tfrac{1}{2}\},\\
D_3 &= \{x = (x_1,x_2) | -1 \leqslant x_i \leqslant 0, \quad i = 1,2\},\\
D_4 &= \{x = (x_1,x_2) | 0 \leqslant x_1 \leqslant 1\tfrac{1}{2}, \quad -1\tfrac{1}{2} \leqslant x_2 \leqslant 0\}.
\end{aligned}$$

It is not difficult to see that the Lebesque set

$$S(f,z) = \{x \in R^2 | f(x) \leqslant f(z)\}$$

contains the simple set of extreme points

$$\text{Ext}\, S(f,z) = \{y = (y_1, y_2) : |y_1| = |y_2|\},$$

that consists of four points, the vertices of the rectangle. Further, numbering these vertices correspondingly to the orthant numbers, we obtain the equalities

$$\begin{aligned} \partial f(y^1) &= \text{co}\{(1,0);(0,1)\}; \\ \partial f(y^2) &= \text{co}\{(-1,0);(0,1)\}; \\ \partial f(y^3) &= \text{co}\{(-1,0);(0,-1)\}; \\ \partial f(y^4) &= \text{co}\{(1,0);(0,-1)\}. \end{aligned}$$

At all other points of the level surface

$$U(z) = \{y \in R^2 | f(y) = f(z)\}, z \neq 0,$$

the function $f(\cdot)$ is differentiable, and its gradient is equal to one of the four vectors: (1,0), (0.1), (-1,0), (0,-1). Let us show, how, with the help of Theorem 3, it is possible to reach the global maximum f over D beginning at and of arbitrary feasible point. Let z_0 be $(-\frac{1}{2}, \frac{1}{2})$. It follows from the above that $y_0^* = (-1,0) \in \partial f(z^0)$. Then, it can be readily seen that the point $z_1 = (-1,0) \in D_3$ is the solution of the problem

$$\langle y_0^*, x\rangle \to \max, \quad x \in D.$$

Besides, the point z_1 turns out to be stationary, since $f'(z_1) = (1,0)$ and

$$\langle f'(z_1), x - z_1\rangle \leqslant 0 \quad \forall x \in D.$$

Nevertheless, if we consider only one $y \in S(f, z_1)$, namely $y = (-1,-1)$ with $y_1^* = (0,-1) \in \partial f(y)$, one obtains $z_2 = (0, 1-\frac{1}{2})$ as the solution of the linearized problem

$$\langle y_1^*, x\rangle \to \max, \quad x \in D.$$

Now let us prove that $z_2 \in \text{Argmax}(f, D)$. Actually,

$$\text{Ext}\, S(f,z_2) = \left\{\left(1\tfrac{1}{2}, 1\tfrac{1}{2}\right); \left(-1\tfrac{1}{2}, 1\tfrac{1}{2}\right); \left(-1\tfrac{1}{2}, -1\tfrac{1}{2}\right); \left(1\tfrac{1}{2}, -1\tfrac{1}{2}\right)\right\}.$$

In addition, it can be readily seen that $\forall y \in \text{Ext}\, S(f, z_2)$, $\forall y^* \in \partial f(y)$ the solution $u(y, y^*)$ of the problem

$$\langle y^*, x\rangle \to \max, \quad x \in D,$$

exists and verifies the condition

$$\langle y^*, u(y,y^*) - y\rangle \leqslant 0.$$

Hence, z_2 is the global solution.

4. Minimization over the supplement of a convex set

Consider the problem

$$f(x) \to \min,\ x \in D, \tag{PR}$$

where D has the non-empty compact convex supplement C, such that,

$$C \triangleq X \backslash D, \quad z \in (D \cap \mathrm{cl}C) \subset \mathrm{bd}D,$$

cl$A$, and bd$A$ being the closure and the boundary of a set A from a Banach space, respectively. In addition, assume that

$$\left.\begin{array}{l} -\infty < f(z) = f^{**}(z) < +\infty, \\ D_o = \{x \in X | f(x) < f(z)\} \subset \mathrm{intdom} f^{**}, \end{array}\right\} \tag{41}$$

where

$$f^{**}(x) = \sup_{x* \in X*} \{< x^*, x > -f^*(x^*)\}, \quad f^*(x^*) = \sup_{x \in X} \{< x^*, x > -f(x)\}$$

are the bipolar and the polar functions of $f(\cdot)$ (see [2–5]). In particular, f may be convex and finite. Note, that Problem (PR) differs from Problem (P) because the feasible set of (PR) is the supplement of a convex set, and therefore, it might happen that co$D = X$. In addition, the objective function is not convex or concave, but has its proper bipolar f^{**}. Recall the following fact, mentioned in Introduction.

PROPOSITION 3 ([12],[14]). Let C be a closed convex set and let D be an arbitrary set from X, s.t. $C \cap D \neq \emptyset$. Then the inclusion $D \subset C$ holds if and only if

$$N(y|C) \subset N(y|D) \quad \forall y \in bdC. \tag{42}$$

THEOREM 3. i) *If* $z \in \mathrm{bd}D$ *is a global minimizer in* (PR) *(*$z \in$ Argmin(PR)*), then*

$$\left.\begin{array}{ll} \forall y \in \mathrm{Ext}\, C, & \forall y^* \in N(y|C), \\ \langle y^*, x - y \rangle \leqslant 0 & \forall x : f(x) \leqslant f(z). \end{array}\right\} \tag{43}$$

ii) *If, in addition,*

$$\exists v \in X : f(v) < f(z), \tag{44}$$

then (43) becomes sufficient for z *to be a global solution of* (PR).

iii) *In particular, from (43) with* $y = z$ *it follows*

$$N(z|C) \subset \mathrm{cone}\, \partial f^{**}(z). \tag{45}$$

Proof. i) If $z \in$ Argmin(PR), then $D \cap D_o \neq \emptyset$, which is equivalent to

$$D_o \subset C. \tag{46}$$

Since C is closed and convex one has

$$\text{clco}D_o \subset C, \tag{47}$$

where $\text{clco}D_o = \{x \in X | f^{**}(x) \leqslant f(z)\}$. Then due to Proposition 3 we obtain

$$N(y|C) \subset N(y|\text{clco}D_o) \quad \forall y \in \text{bd}C. \tag{48}$$

Condition (48) can obviously be expressed as

$$\left.\begin{array}{ll} \forall y \in \text{bd}C & \forall y^* \in N(y|C), \\ \langle y^*, x - y\rangle \leqslant 0 & \forall x : f^{**}(x) \leqslant f(z), \end{array}\right\} \tag{49}$$

which is equivalent to (43) in virtue of (41).

ii) Now let (43) and (44) hold. As above, (43) implies Condition (49) which is equivalent to (48) and (47). Let us prove that (47) implies (46). Actually, due to (44) the set

$$\text{int}(\text{clco}D_o) = \{x \in X | f^{**}(x) < f(z)\}$$

is nonempty and open. Furthermore, since Inclusion (47) holds we have

$$\text{int}(\text{clco}D_o) \subset \text{ int}C \subset C.$$

Besides, Assumption (41) allows us to conclude that $D_o \subset \text{int}(\text{clco}D_o)$ and consequently (46) also takes place. The latter means that $z \in$ Argmin(PR).

iii) Since (see [2–5]) $N(z|\text{ clco}D^o) = \text{cone}\,\partial f^{**}(z)$ Inclusion (48) with $y = z$ implies (45). □

Let us make several remarks concerning the obtained condition.

1) It is well-known (see [2–5]) that from the classical OC for Problem (PR) it follows that

$$\partial f(z) \cap N(z|K) \neq \emptyset, \tag{50}$$

where K is a local conical approximation of D. Condition (43) is completely different because it works with the supplement C of the feasible set D. This can easily be seen when one sets (43) as

$$N(y|C) \subset N(y|S(f,z)) \quad \forall y \in \text{bd}C, \tag{51}$$

where $S(f,z) = \{x \in X | f(x) \leqslant f(z)\}$. For example, for differentiable f, (45) implies

$$\langle f'(z), x - z\rangle \leqslant 0 \quad \forall x \in C. \tag{52}$$

It seems, that this condition is known and even classical. However, taking into account that (52) does not hold for the elements of the feasible set D but only for the points of its supplement $C = X\backslash D$, it becomes clear that (43) is not really ordinary because it works with unfeasible elements. In addition, Condition (45) (hence (52)) is the particular case ($y = z$) of the more general and the more informative ($\forall y \in \text{bd}C$) OC (43).

2) On the other hand, (43) conserves the Algorithmic Property of the classical conditions in the following sense. In order to verify (52) we have to solve the linearized problem

$$\langle f'(z), x\rangle \to \max, \quad x \in C.$$

If now, $\forall y \in \text{ bd}C$ and $\forall y^* \in N(y|C)$, we consider the problem

$$\langle y^*, x\rangle \to \max, \;\; f(x) \leqslant f(z), \tag{53}$$

then it follows from (43) that whenever $\{x^k\}$ is a maximizing sequence of Problem (53) we have

$$\lim_{k\to+\infty}\langle y^*, x^k - y\rangle \leqslant 0.$$

3) Besides if there exists a single triplet (y, y^*, u)

$$y \in \text{ bd}C, \quad y^* \in N(y|C), \quad f(u) \leqslant f(z),$$

for which we have $\langle y^*, u - y\rangle > 0$ then in virtue of (43) z is not a global solution of Problem (PR).

EXAMPLE 3. Define $f(\cdot)$ and D in (PR) as follows

$$f(x) = \max\{e^x, 1 - x, -2x\}, \quad D \triangleq R\backslash] - 1, 1[.$$

Obviously, $C =] - 1, 1[$ and $\text{bd}C = \text{Ext } C = \{-1; 1\}$. Denote $z_1 = -1, z_2 = 1$. Then

$$\partial f(z_1) = [-2, -1], \partial f(z_2) = \{f'(z_2)\} = e.$$

It can easily be seen that z_1 and z_2 verify the classical condition (50) so that they are stationary in the classical sense (where $K(z_1) =] - \infty, -1[, K(z_2) =]1, +\infty[$). Let us apply Condition (43).

1) For $z_2 = 1$ we have $f(z_2) = e$ and one can see that $u = -\frac{e}{2}$ is the solution of the problem $(y^* = -1/2 \in N(y|C))$

$$\langle y^*, x\rangle \to \max, \quad f(x) \leqslant f(z_2),$$

In addition, we have

$$\langle y^*, u - y\rangle = \left(-\tfrac{1}{2}\right)\left(-\tfrac{e}{2} + 1\right) > 0,$$

and hence, z_2 is not a global solution.

2) For z_1 consider the two problems

$$\langle y_1^*, x\rangle \to \max, \quad f(x) \leqslant f(z_1) = 2, \tag{54}$$

$$\langle y_2^*, x\rangle \to \max, \quad f(x) \leqslant 2. \tag{55}$$

$$y_i^* \in N(y^i|C), i = 1,2; y^1 = -1, y^2 = 1, \quad \{y^1, y^2\} = \text{bd}C.$$

Since $N(y^1|C) =]-\infty, 0], N(y^2|C) =]0, +\infty]$, it is possible to take $y_i^* = -1, y_2^* = 1$. Then we see that $u_1 = -1$ and $u_2 = \ln 2$ are the solutions of Problems (54) and (55), respectively. In addition,

$$\langle y_1^*, u_1 - y_1\rangle = 0,$$
$$\langle y_2^*, u_2 - y_2\rangle < 0.$$

Thus, Condition (43) holds $\forall y \in \text{bd}C$ and therefore, z_2 is the global solution. □

Now, consider the problem with reverse convex constraint

$$f(x) \to \min, \quad g(x) \geqslant 0, \tag{56}$$

where f satisfies (41) and $g : X \to R \cup \{+\infty\}$ is a convex closed function, such that,

$$C \triangleq \{x \in X | g(x) \leqslant 0\} \subset \text{intdom } g, \tag{57}$$

$$-\infty \leqslant \inf(g, X) < g(z) = 0. \tag{58}$$

THEOREM 4. i) *If z is a global solution of Problem (56), $g(z) = 0$, then*

$$\left.\begin{array}{l} \forall y : g(y) = 0, \quad \forall y^* \in \partial f(y), \\ \langle y^*, x - y\rangle \leqslant 0 \quad \forall x : f(x) \leqslant f(z). \end{array}\right\} \tag{59}$$

Under Assumption (41) Condition (59) becomes sufficient for z to be a global minimizer to (56).

ii) *If set C defined by (57) is compact then in (59) it is sufficient to use only the points $y \in$* Ext C.

iii) *In particular from (59) with $y = z$ it follows that*

$$\partial g(z) \subset \text{cone } \partial f^{**}(z). \tag{60}$$

Proof. The assertion i) follows from Theorem 3 and the equalities (see [3]):

$$N(y|C) = \text{cone } \partial g(y) \quad \forall y : g(y) = 0, \tag{61}$$

that are true because of (58).

ii) It suffices to recall Proposition 2.

iii) It results from (61) and (45) of Theorem 3. □

Remarks. 1) It can be readily seen when $X = R^n$ and f and g are differentiable that it follows from (60)

$$g'(z) = \lambda f'(z), \quad \lambda > 0. \tag{62}$$

The latter is the well-known local optimality condition. Hence, there is a certain relation between (59) and the classical OC for Problem (56).

2) On the other hand, it is not difficult to see that Condition (60) is not standard because $\forall y^* \in \partial g(z)$ it states the existence of a representation $y^* = \lambda x^*, \lambda > 0, x^* \in \partial f^{**}(z)$, while the classical OC assert the existence of only one of such representations.

3) Besides, (60) is only a particular case $(y = z)$ of the more deep and more informative OC (59) $(\forall y \in \text{ Ext } C)$.

4) Clearly, for verifying Condition (59),

$$\forall y : g(y) = 0, \quad \forall y^* \in \partial g(y),$$

one has to solve the linearized problem

$$\langle y^*, x\rangle \to \max, f(x) \leqslant f(x), \tag{63}$$

(which is simpler than the original problem (56)) and after that one has to verify the inequality

$$\lim_{k\to\infty} \langle y^*, x^k - y\rangle \leqslant 0,$$

for a maximizing (for (63)) sequence $\{x^k\}$.

5) Finally, in order to show that z is not a global solution it suffices to find a single triplet $(y, y^*, u) : g(y) = 0, y^* \in \partial g(y), f(u) \leqslant f(z)$, s.t.

$$\langle y^*, u - y\rangle > 0.$$

EXAMPLE 4. Consider the problem

$$f(x) \triangleq \left(x_1 - \tfrac{1}{2}\right)^2 + \left(x^2 + \tfrac{1}{2}\right)^2 \to \min, \quad g(x) \triangleq |x| - \sqrt{2} \geqslant 0.$$

Since $g(\cdot)$ is differentiable : $g'(x) = x/|x|$ for $x \neq 0$, it results from (62) with $z = (\zeta_1, \zeta_2)$

$$\frac{(\zeta_1, \zeta_2)}{|z|} = 2\lambda \left(\zeta_1 - \tfrac{1}{2}; \zeta_2 + \tfrac{1}{2}\right). \tag{64}$$

Taking into account, that $|z| = \sqrt{2}$, we have

$$\alpha\zeta_1 = \zeta_1 - 1, \quad \alpha\zeta_{21} = \zeta_2 + 1, \quad \alpha > 0.$$

Consequently, $\zeta_1 = -\zeta_2 = (2-\alpha)^{-1}$. Further, since $\zeta_1^2 + \zeta_2^2 = 2$, we obtain a quadratic equation for α. Solving this equation we have $\alpha_1 = 3, \alpha_2 = 1$. Hence, there are two critical points

$$z_1 = (1, -1), \quad z_2 = (1, -1).$$

Now, apply GOC (59). In order to do this consider the points $y = (0, -\sqrt{2})$, $g(y) = 0$, $u = (0, \frac{5}{2})$ and $f(u) = \frac{17}{4} < \frac{9}{2} = f(z_1)$. Then, we have

$$\left\langle \frac{g'(y)}{|y|}, u - y \right\rangle = \tfrac{5}{2} - \sqrt{2} > 0.$$

The latter means that the point z_1 is not a global solution. Hence, the point z_2 is a global minimizer because Equation (64) has only two solutions.

EXAMPLE 5. Let $x \in R^2$ and consider the problem

$$f(x) = (x_1 - 1)^2 + x_2 \to \text{min}, \quad g(x) = x_2^2 - x_1 - 2 \geqslant 0.$$

It can be readily seen that the point $z = (-2, 0)$ verifies the classical condition (62). However, for (y, u)

$$y = \tfrac{1}{2}\left(1, \sqrt{10}\right), g(y) = 0 = g(z), u = \tfrac{1}{2}\left(1, \tfrac{3}{2}\sqrt{10}\right), f(u) \leqslant f(z),$$

we obtain

$$\langle g'(y), u - y \rangle = \frac{10}{4} > 0.$$

Hence, according to Theorem 4, z cannot be a global solution.

5. Conclusion

In this paper, using the subdifferential $\partial f(y)$, at an extreme point y of the compact

$$S = S(f, z) = \{x \in X / f(x) \leqslant f(z)\}$$

of a convex function

$$f : X \to R \cup \{+\infty\}$$

over a Banach space X, the following results have been obtained:

- the characterization of an element $y^* \in X^*$, which belongs to the polar $(S-z)^o$ but is not contained in the normal cone $N(z, S)$;
- the characterization of the inclusion of a set $A \subset X$ into $S(f, z)$;
- global optimality conditions for convex maximization, which generalized the classical optimality conditions for the problem of interest;
- global optimality conditions for reverse convex problems, that are connected with Classical Extremum Theory;

– all theoretical results were illustrated by numerical examples.

References

1. Rudin, W. (1973), *Functional Analysis*, McGraw-Hill.
2. Rockafellar, R.T. (1970), *Convex Analysis*, Princeton University Press.
3. Ioffe, A.D. and Tikhomirov, V.M. (1979), *Theory of Extremal Problems*. North-Holland.
4. Alexeev, V., Tikhomirov, V. and Fomine, S. (1982), *Optimal Control*, Mir, Moscow.
5. Hiriart-Urruty, J.-B. and Lemaréchal, C. (1993), *Convex Analysis and Minimization Algorithms*, Volumes 1 and 2, Springer Verlag.
6. Hiriart-Urruty, J.B. and Lemaréchal, C. (1990), *Testing Necessary and Sufficient Conditions for Global Optimality in the Problem of Maximizing a Convex Quadratic Function over a Convex Polyhedron*, Seminar of Numerical Analysis, University Paul Sabatier, Toulouse, pp. 1–34.
7. Hiriart-Urruty, J.B., (1989), From Convex Optimization to Nonconvex Optimization, Part 1: Necessary and Sufficient Conditions for Optimality, in *Nonsmooth Optimization and Related Topics*, Plenum Press, pp. 219–239.
8. Horst, R. and Tuy, H. (1990), *Global Optimization, a Deterministic Approach*, Springer Verlag.
9. Horst, R., Pardalos, P.M. and Thoai, N.V. (1995), *Introduction to Global Optimization*, Kluwer Academic Publishers, Dordrecht/Boston/London.
10. Horst, R. and Pardalos, P.M. (eds.) (1995) *Handbook of Global Optimization*, Kluwer Academic Publishers, Dordrecht/Boston/London.
11. Strekalovsky, A. (1987), *On the Problem of Global Extremum*. Soviet Math. Doclady of Science Academy, Vol. 35, pp. 194–198.
12. Strekalovsky, A. (1990), *On the Problem of Global Extremum in Non-convex Extremal Problems*. Izvestia VUZov, seria Mathematics, No. 8, pp. 74–80.
13. Strekalovsky, A. (1992), *On the Global Extremum Condition in a Nonconvex Minimization Problem*. Izvestia VUZov, seria Mathematics, No.2, pp. 94–96.
14. Strekalovsky, A. (1992), *On the Problem of Global Extremum Conditions in Nonconvex Optimization Problems*. Voprosy Kibernetiki, Analisis of large systems, Russian Academy of Sciences, Scientific Moscow, pp. 178–197.
15. Strekalovsky, A. (1993), On Optimization Problems over Supplements of Convex Sets. *Cybernetics and System Analysis* 1, 113–126.
16. Strekalovsky, A. (1993), On Global Maximum Search of Convex Function over a Feasible set. *Journal of Numerical Mathematics and Mathematical Physics* 3, 349–363.