Journal of Convex Analysis
Volume 3 (1996), No.1, 55–61

A Note on the Characterization of the Global Maxima of a (Tangentially) Convex Function Over a Convex Set

J.-B. Hiriart-Urruty
Université Paul Sabatier 118, route de Narbonne,
31062 Toulouse cedex, France.
e-mail: jbhu@cict.fr

Yuri S. Ledyaev
Steklov Mathematics Institute, Vavilov St. 42,
Moscow 117966, Russia.
e-mail: ledyaev@de.mian.su

Received May 31, 1994
Revised manuscript received February 1, 1995

Dedicated to R. T. Rockafellar on his 60th Birthday

We consider the problem of characterizing points $\bar{x}$ in a convex set C which globally maximize an objective function f over C. When f is convex, we show how the first order necessary condition extended to those $x \in C$ at the same level as $\bar{x}$ is necessary and sufficient for $\bar{x}$ being a global maximum of f over C.This improves a recent result by A. Strekalovski who was the first to propose such type of global optimality condition, but whose characterization required to scan all the points (even out of C) at the same level as $\bar{x}$. Next we extend the obtained characterization to the case where the objective function is just tangentially convex and an appropriate qualification condition holds.

1. Characterizing the global maxima of a convex function over a convex set

Let X be a (real) Banach space, X^* its topological dual space ; when $s \in X^*$ and $x \in X$, we use the notation $\langle s, x \rangle$ for $s(x)$. Let $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a *lower semicontinuous convex* function, and let C be a nonempty *closed convex* set contained in the interior of the domain of f. We are interested in characterizing those $\bar{x}$ in C *maximizing globally* f on C, i.e. satisfying :

$$f(x) \leq f(\bar{x}) \, for \; all \, x \in C. \tag{1}$$

The convexity of C is not so important here since we know that the maximum of a convex function (resp. of a continuous convex function) over a set S is unaltered if we substitute the convex hull of S (resp. the closed convex hull of S) for S.

If $\bar{x} \in C$ is a maximum (even local) of f on C and if f is Gâteaux- differentiable at $\bar{x}$, then $Df(\bar{x})$ (the Gâteaux derivative of f at $\bar{x}$) necessarily lies in the normal cone to C at $\bar{x}$:

ISSN 0944-6532 / \$ 2.50

$$Df(\bar{x}) \in N(C, \bar{x}). \tag{2}$$

This necessary condition for maximality holds true whenever f is differentiable at $\bar{x}$, be it convex or not. When, like in our case, f is convex, we have

$$\partial f(\bar{x}) \subset N(C, \bar{x}) \tag{3}$$

at any maximum (even local) $\bar{x}$ of f on C (here, $\partial f(\bar{x})$ denotes as usual the subdifferential of f at $\bar{x}$). Already at this stage the convexity of f has played its role : a natural extension of (2) would be $\partial f(\bar{x}) \cap \ N(C, \bar{x}) \neq \emptyset$, while we get in (3) much more : the whole subdifferential of f at $\bar{x}$ is contained in $N(C, \bar{x})$.

Condition (3) is not sufficient - far from it- for $\bar{x}$ being a maximum (even local) of f on C. To characterize a global maximum of f on C, we need "to globalize" (3) somehow or other. The first approach in that direction is treated in [2] where the global maxima of f on C are *characterized* with the help of "ε - enlargements" of the subdifferential of f and of the normal cone to C. Indeed, denoting by $\partial_\varepsilon f(\bar{x})$ the ε - subdifferential of f at $\bar{x}$ and by $N_\varepsilon(C, \bar{x})$ the set of ε - normal directions to C at $\bar{x}$, we have : $\bar{x} \in C$ is a global maximum of f on C if and only if :

$$\partial_\varepsilon f(\bar{x}) \subset N_\varepsilon(C, \bar{x}) \quad for \ \ all \ \varepsilon > 0. \tag{4}$$

Since $\partial f(\bar{x}) = \bigcap_{\varepsilon > 0} \partial_\varepsilon f(\bar{x})$ and $N(C, \bar{x}) = \bigcap_{\varepsilon > 0} N_\varepsilon(C, \bar{x})$, the condition (3) is precisely the limiting case $\varepsilon \downarrow 0$ in (4). In (4) the parameter $\varepsilon > 0$ helped to globalize the (purely local) condition (3).

It turns out there is another way of globalizing the local condition (3) : by imposing the same condition as (3) to all the points x in the level-surface of f at level $f(\bar{x})$. This approach is due to A. Strekalovski ([3], [4]) and his characterization can be stated as follows. Consider $\bar{x} \in C$ such that

$$\begin{gathered} -\infty \leq \inf_X f < f(\bar{x}); \text{ then } \bar{x} \text{ is a global maximum of } f \text{ on } C \text{ if only if} \\ \partial f(x) \subset N(C, x) \text{ for all } x \text{ satisfying } f(x) = f(\bar{x}). \end{gathered} \tag{5}$$

The trouble in that formulation is that one needs to consider, a priori, normal cones to C at points x which do not necessarily lie in C ($d \in N(C, x)$ still means in that case : $\langle d, c - x \rangle \leq 0$ for all $c \in C$), which is unnatural. The improvement of Strekalovski's result we propose below consists in showing that a global maximum $\bar{x}$ of f on C is still characterized by (5) restricted to those x *in* C at the same level as $\bar{x}$.

Theorem 1.1. *Consider a point $\bar{x} \in C$ such that*

$$-\infty \leq \inf_C f < f(\bar{x}).$$

Then $\bar{x}$ is a global maximum of f on C if and only if

$$\partial f(x) \subset N(C, x) \ for \ all \ x \ in \ C \ satisfying \ f(x) = f(\bar{x}). \tag{6}$$

Proof. [*Condition* (6) *is necessary*]. Let $s \in \partial f(x)$ we have :

$$\langle s, x^{'} - x\rangle \leq f(x^{'}) - f(x) \text{ for all } x^{'} \in X.$$

If $x \in C$ and $f(x) = f(\bar{x})$, we necessarily have
$\langle s, x^{'} - x\rangle \leq f(x^{'}) - f(\bar{x}) \leq 0$ for all $x \in C$, that is : $s \in N(C, \bar{x})$.

[*Condition* (6) *is sufficient*]. We will use here an equivalent form of condition (6) :
for all x in C satisfying $f(x) = f(\bar{x})$, we have $\sup_{c \in C} f'(x, c - x) \leq 0$, (*)

i.e., there is no direction d tangent to C at x for which $f'(x, d) > 0$ (here, $f'(x, .)$ denotes the directional derivative of f at x ; the assumptions made ensure that $f'(x, .)$ is a continuous function).

Suppose the contrary holds : there is some $\tilde{x} \in C$ for which $f(\tilde{x}) > f(\bar{x})$. We aim at contradicting (*).

Due to the assumption $\inf_C f < f(\bar{x})$, there exists $x_* \in C$ such that
$f(x_*) < f(\bar{x})$. Connect x_* and $\tilde{x}$ by a line-segment. Since f restricted to this line remains convex (continuous) and $f(x_*) < f(\bar{x}) < f(\tilde{x})$, there is a unique point x_0 on the interval $(x_*, \tilde{x})$ at which $f(x_0) = f(\bar{x})$. Moreover, the restriction of f to the interval $(x_*, \tilde{x})$ obviously increases with a positive right-derivative, in particular $f'(x_0, \tilde{x} - x_0 > 0$. We therefore have found x_0 in C such that $f(x_0) = f(\bar{x})$ and $\sup_{c \in C} f'(x_0, c - x_0) > 0$, which contradicts (*). □

Strekalovski's condition (5) has been substantially improved in (6) by only considering the points at the same level $f(\bar{x})$ which lie in C. There is however a slight difference in the assumptions : Strekalovski's condition, applied for $\bar{x}$, requires there exists some $x_o \in X$ such that $f(x_o) < f(\bar{x})$, while our condition requires there exist $c_o \in C$ such that $f(c_o) < f(\bar{x})$. Both requirements are very mild and easily understandable. They cannot be removed : if, for example, $\bar{x}$ is the unique minimum on X of a differentiable f and if $\bar{x}$ lies in the interior of C, condition (5) (or (6)) is obviously satisfied. The next simple example shows the difference between the two assumptions.

Example 1.2. Let $f : (\xi_1, \xi_2) \in R^2 \mapsto f(\xi_1, \xi_2) = \xi_2^2 - \xi_1 + 1$ and let $C := \{0\} \times [-1, +1]$. Consider $\bar{x} = (0, 0)$; we have : $f(\bar{x}) = \inf_C f > \inf_{R^2} f$.

It is easy to verify that our condition (6) is satisfied, but Theorem 1.2 cannot be invoked; indeed $\bar{x}$ is not a global maximum of f on C. As for condition (5), it is not satisfied : at $x = (1, 1)$ which is at the same level $f(\bar{x})$ (but not in C), the gradient of f at x (actually the unique element of $\partial f(x)$) does not belong to $N(C, x)$.

2. Extending the characterization of global maxima to tangentially convex objective functions

In this section we extend Theorem 1.2 to (global) maximization problems with possibly nonconvex objective functions. More precisely we remove the convexity condition on f and replace it by the following one :

f is *locally Lipschitz* and *strictly tangentially convex* (or regular in Clarke's terminology [1, p.39]) on the interior of its domain, i.e. : for all x in the interior of the domain of f,
(i) The usual one-sided directional derivative $d \mapsto f'(x,d)$ exists ;
(ii) $f'(x,.)$ is the support function of Clarke's generalized gradient $\partial f(x)$ of f at x.
As simple examples show it, one cannot get away from the convexity (and continuity) of the "tangent approximations" $d \mapsto f'(x,d)$ of f at x, at least if one wishes to have the *same* condition be necessary *and* sufficient for global optimality.

We assume moreover X is a (real) *Hilbert space* ; we denote by $\langle .,. \rangle$ the inner product in X and by $||.||$ the associated norm.

Theorem 2.1. *Assume the following qualification condition holds at $\bar{x} \in C$:* $(QC)_{\bar{x}}$
For all $x \in C$ such that $f(x) = f(\bar{x})$, it there is $c_x \in C$ such that

$$f'(x, c_x - x) < 0.$$

Then $\bar{x}$ is a global maximum of f on C if and only if

$$\partial f(x) \subset N(C,x) for\ all\ x\ in\ C\ satisfying\ f(x) = f(\bar{x}). \tag{6}'$$

The difference between (6) and $(6)'$ is that, now, $\partial f(x)$ stands for the generalized gradient of f at x. The statement of Theorem 2.1 is similar to that of Theorem 1.1 but the lack of convexity of f compels us to use different techniques in the proof. Before going further, some comments on $(QC)_{\bar{x}}$ are in order :

- What does $(QC)_{\bar{x}}$ mean ? It means that from any point x in C at the same level as $\bar{x}$, there is a direction of descent (for f) directed towards a point of C. It is clear that $(QC)_{\bar{x}}$ cannot be completely removed : suppose, for example, f is strictly differentiable at a point $\bar{x}$ lying in the interior of C, there is no other point in C at the same level as $\bar{x}$, and $Df(\bar{x}) = 0$; then $(6)'$ is satisfied but $\bar{x}$ need not necessarily be a maximum of f over C.
- $(QC)_{\bar{x}}$ is truly a generalization of the qualification condition encountered in section 1 for convex f. Indeed, if f is convex and if there is some $c_o \in C$ such that $f(c_o) < f(\bar{x})$, then
$$f'(x, c_o - x) \leq f(c_o) - f(x) = f(c_o) - f(\bar{x})$$
for all x in C at the same level as $\bar{x}$.
- Suppose f is continuously differentiable on the interior of its domain. Then $(QC)_{\bar{x}}$ and $(6)'$ take respectively the following form :
$$\inf_{c \in C} \langle Df(x), c \rangle \ < \ \langle Df(x), x \rangle$$
and
$$\langle Df(x), x \rangle = \max_{c \in C} \langle Df(x), c \rangle$$
(for all $x \in C$ at the same level as $\bar{x}$). Hence, under $(QC)_{\bar{x}}$, for all x in C at the same level as a global maximum $\bar{x}$ of f on C, the width of C in the $Df(x)$ direction is non-null.

Proof of Theorem 2.1. [*Condition* (6)$'$ *is necessary*]. Let $x \in C$ satisfy $f(x) = f(\bar{x})$. Since x is also a global maximum of f on C,

$$f(x + t(c - x)) \leq f(x) \text{ for all } \ 0 < t \leq 1 \text{ and } c \in C.$$

Consequently

$$f'(x, c - x) \leq 0 \text{ for all } c \in C,$$

and because $f'(x, .)$ is the support function of $\partial f(x)$

$$\langle s, c - x \rangle \leq 0 \text{ for all } c \in C \text{ and } s \in \partial f(x)$$

i.e.,

$$\partial f(x) \subset N(C, x).$$

[*Condition* (6)$'$ *is sufficient*]. Suppose for contradiction there is some $\tilde{x} \in C$ such that $f(\tilde{x}) > f(\bar{x})$ and let $D := C \cap \{x \in X \mid f(x) \leq f(\bar{x})\}$. We consider the following minimization problem :

$$\text{Minimize } \frac{1}{2} \parallel x - \tilde{x} \parallel^2 \ \text{ subject to } x \in D.$$

In general this problem has no solution since D is only a closed subset of X. But, according to Ekeland's variational principle (see [1, p. 266] for example), for any $\varepsilon > 0$ there exists $\hat{x}_\varepsilon \in D$ such that the function

$$\varphi_\varepsilon : x \longmapsto \varphi_\varepsilon(x) := \frac{1}{2}||x - \tilde{x}||^2 + \varepsilon||x - \hat{x}_\varepsilon||$$

attains its minimum on D at $\hat{x}_\varepsilon$.

Claim 1 : $f(\hat{x}_\varepsilon) = f(\bar{x})$ for $\varepsilon > 0$ small enough.
If not there exists $(\varepsilon_k)_k \downarrow 0$ such that $\hat{x}_k := \hat{x}_{\varepsilon_k}$ satisfies

$$f(\hat{x}_k) < f(\bar{x}) \ for \ all \ k. \tag{7}$$

We set $x_{k,t} := \hat{x}_k + t(\tilde{x} - \hat{x}_k)$ for $0 < t \leq 1$. Since $x_{k,t} \in C$ (because both $\hat{x}_k$ and $\tilde{x}$ are in C and C is convex) and $f(x_{k,t}) \leq f(\bar{x})$ for t small enough (because of the continuity of f at $\hat{x}_k$ and (7)), we can say :

$$\exists \bar{t}_k > 0 \text{ such that } x_{k,t} \in D \ for \ all \ 0 < t \leq \bar{t}_k.$$

But $\hat{x}_k$ is, by definition, a minimum of the function $x \longmapsto \frac{1}{2}||x - \tilde{x}||^2 + \varepsilon||x - \hat{x}_k||$ on D ; thus :

$$\frac{1}{2}||x_{k,t} - \tilde{x}||^2 + \varepsilon_k||x_{k,t} - \hat{x}_k|| \geq \frac{1}{2}||\hat{x}_k - \tilde{x}||^2 \text{ for all } 0 < t \leq \bar{t}_k. \tag{8}$$

By plugging the expression $x_{k,t} = \hat{x}_k + t(\tilde{x} - \hat{x}_k)$ into (8), we derive :

$$\frac{1}{2}(1 - t)^2||\hat{x}_k - \tilde{x}||^2 + t\varepsilon_k||\tilde{x} - \hat{x}_k|| \geq \frac{1}{2}||\hat{x}_k - \tilde{x}||^2$$

or

$$\frac{1}{2}t^2||\hat{x}_k - \tilde{x}||^2 - t||\hat{x}_k - \tilde{x}||^2 + t\varepsilon_k||\tilde{x} - \hat{x}_k|| \geq 0 \text{ for all } 0 < t \leq \bar{t}_k.$$

Since $||\hat{x}_k - \tilde{x}|| > 0$ it follows that for $t > 0$ small enough

$$||\hat{x}_k - \tilde{x}|| \leq \varepsilon_k + \frac{t}{2}||\hat{x}_k - \tilde{x}||.$$

By letting $t \downarrow 0$ we obtain $||\hat{x}_k - \tilde{x}|| \leq \varepsilon_k$.
Now, since $\hat{x}_k$ satisfies (7) and $(\hat{x}_k)_k$ converges to $\tilde{x}$, the continuity of f implies that $f(\tilde{x}) \leq f(\bar{x})$, which contradicts the choice of $\tilde{x}$.
Thus we have proved that $f(\hat{x}_\varepsilon) = f(\bar{x})$ for $\varepsilon > 0$ small enough. We henceforth consider such $\varepsilon > 0$.

Claim 2 : $\partial f(\hat{x}_\varepsilon \subset N(C, \hat{x}_\varepsilon)$ leads to a contradiction.

Consider for that the function

$$g : x \longmapsto g(x) := max\{\varphi_\varepsilon(x) - \varphi_\varepsilon(\hat{x}_\varepsilon), f(x) - f(\bar{x})\}.$$

We have : $g(\hat{x}_\varepsilon) = 0$ and $g(x) \geq 0$ for all $x \in C$; we reformulate that by saying that $\hat{x}_\varepsilon$ solves the following minimization problem :

$$(\hat{P}) \text{ Minimize } g(x) \text{ subject to } x \in C.$$

Then a necessary condition is as follows : there exists $\lambda_0 \geq 0$ and $\lambda \geq 0$ adding up to 1 such that

$$\lambda_0\varphi_\varepsilon^{'}(\hat{x}_\varepsilon, c - \hat{x}_\varepsilon) + \lambda f'(\hat{x}_\varepsilon, c - \hat{x}_\varepsilon) \geq 0 \text{ for all } c \in C$$

We claim that $\lambda_0 > 0$ necessarily. If not we would have

$$f'(\hat{x}_\varepsilon, c - \hat{x}_\varepsilon) \geq 0 \text{ for all } c \in C,$$

which contradicts $(QC)_{\bar{x}}$.

Thus

$$\frac{\lambda}{\lambda_0}f'(\hat{x}_\varepsilon, c - \hat{x}_\varepsilon) \geq -\varphi_\varepsilon'(\hat{x}_\varepsilon, c - \hat{x}_\varepsilon) \text{ for all } c \in C \tag{9}$$

The calculation of $\varphi_\varepsilon'(x, .)$ at $x = \hat{x}_\varepsilon$ is easy :

$$\forall d \in H, \varphi_\varepsilon'(\hat{x}_\varepsilon, d) =< \hat{x}_\varepsilon - \tilde{x}, d > +\varepsilon||d||.$$

Now, by choosing $c = \tilde{x}$ in (9) we get at

$$\frac{\lambda}{\lambda_0}f'(\hat{x}_\varepsilon, \tilde{x} - \hat{x}_\varepsilon) \geq ||\tilde{x} - \hat{x}_\varepsilon||^2 - \varepsilon||\tilde{x} - \hat{x}_\varepsilon||. \tag{10}$$

But since $f'(\hat{x}_\varepsilon, c - \hat{x}_\varepsilon) \leq 0$ for all $c \in C$ (as said earlier this is an equivalent form of the inclusion $\partial f(\hat{x}_\varepsilon) \subset N(C, \hat{x}_\varepsilon)$), the inequality (10) implies that $||\hat{x}_\varepsilon - \tilde{x}|| \leq \varepsilon$. So again

$lim_{\varepsilon\downarrow 0}\hat{x}_{\varepsilon} = \tilde{x}$, whence $f(\tilde{x}) \leq f(\bar{x})$, which contradicts the inquality $f(\tilde{x}) > f(\bar{x})$ assumed from the beginning. □

Remark 2.2. The second part of the proof above shows that if f is just locally Lipschitz (not necessarily strictly tangentially convex) and if we assume $(QC)_{\bar{x}}$ (expressed with Clarke's generalized directional derivative of f instead of f), then $(6)'$ is a *sufficient* condition for $\bar{x}$ being a global maximum of f on C.

To end this note we mention a global optimality condition in nonconvex optimization of a similar philosophy to the ones presented here. Suppose X is a (real) Banach space, C a closed convex set of X, and $f : C \to \mathbb{R}$ is continuous (actually more general X and C could be considered, C just connected for example). Consider $\bar{x} \in C$ satisfying the following :

$$Every\ x \in C\ for\ which\ f(x) = f(\bar{x})\ is\ a\ local\ maximum\ of\ f\ on\ C. \tag{11}$$

Then $\bar{x}$ is a global maximum of f on C. Thus (11) is a necessary and sufficient condition for $\bar{x}$ being a global maximum. The proof is left as an exercise to the reader.

Acknowledgment. We would like to thank the referees for the comments and improvements they suggested. In particular, the present proof of Theorem 1.1, shorter than our earlier one, is due to one of them.

References

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

[2] J.-B. Hiriart-Urruty: From convex optimization to nonconvex optimization. Part I : Necessary and sufficient conditions for global optimality, in Nonsmooth Optimization and Related Topics, Plenum Press (1989) 219–239.

[3] A. Strekalovski: On questions of global extremum in nonconvex extremal problems, Izvestia Vyssh. Uchebnyh Zavedenii 8 (1990) 74–80 [in Russian].

[4] A. Strekalovski: On search of global maximum of convex functions on a constraint set, J. of Comp. Mathematics and Math. Physics 33 (1993) 349–363 [in Russian, to be translated into English].

[5] P.-J. Laurent: Approximation et Optimisation, Hermann, Paris 1972.

HIER :

Leere Seite
62