PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série, tome 76(90) (2004), 101–109

ON A CLASS OF NONCONVEX PROBLEMS WHERE ALL LOCAL MINIMA ARE GLOBAL

Leo Liberti

Communicated by Gradimir Milovanović

ABSTRACT. We characterize a class of optimization problems having convex objective function and nonconvex feasible region with the property that all local minima are global.

1. Introduction

The nonlinear optimization problem we address in this paper is of the form

$$\begin{cases} \min_{\underline{x}\in V} f(x) \\ \underline{g}(\underline{x}) = 0 \\ \underline{h}(\underline{x}) \leqslant 0 \end{cases}$$

where $\underline{x} \in \mathbb{R}^n$, $f : V \to \mathbb{R}$ with $V \subseteq \mathbb{R}^n$, and $\underline{g}, \underline{h}$ are systems of equations and inequalities, respectively. We are interested in finding the global minimum $\hat{\underline{x}}$.

Most Branch-and-Bound algorithms designed to solve this kind of problems need to compute the lower bound to the objective function value in each Branch-and-Bound region. The lower bound is calculated by solving a convex relaxation of the original problem, the reason for this being that in a convex problem all local minima are global ones; hence a local solver can be used to obtain a guaranteed lower bound to the objective function value of the original problem in the region of interest [**SP99, Lib04, Lib05**].

However, the notion of globality of all local minima applies to many problem instances [**Bon98**], apart from the class of convex problems. This suggests the use of nonconvex relaxations (having the same local-to-global minimality property) which may be much tighter than an ordinary convex relaxations, and thus speed up considerably the global optimization software acting on the nonlinear problem.

Here, we follow a topological approach. We prove that certain convex functions defined on nonconvex sets satisfying a special set of conditions have the desired property. There is some relation between this work and [**Rap91**].

2000 *Mathematics Subject Classification*: Primary 26B25; Secondary 32F17, 52A30.

2. Convexity over Path-connected Sets

Let $\mathbb{F}$ be any set and $\mathbb{G}$ be a totally ordered set. We refer to elements of $\mathbb{F}^n$ as vectors or points (and write them as $\underline{x}$) and to elements of $\mathbb{G}$ as scalars or numbers (and write them simply as x). If we refer to a function with a vector notation, as in $\underline{\gamma}$, it means that the function returns a vector of some kind. Conversely, referring to a function with a scalar notation, as in f, points out the fact that f returns a scalar.

Since in what follows both $\mathbb{F}$ and $\mathbb{G}$ are assumed to have the same topological properties as the real numbers, for simplicity we shall use $\mathbb{R}$ instead of $\mathbb{F}$ and $\mathbb{G}$.

A *path* between two points $\underline{x}, \underline{y}$ in $\mathbb{R}^n$ is a continuous function $\underline{\gamma} : [0, 1] \to \mathbb{R}^n$ such that $\underline{\gamma}(0) = \underline{x}$ and $\underline{\gamma}(1) = \underline{y}$. We shall sometimes refer to a path γ (without the vector notation), meaning the set

$$\gamma = \{\underline{u} \in V \mid \exists \alpha \in [0, 1](\underline{u} = \underline{\gamma}(\alpha))\}.$$

Basically, γ is the set of all points on the path. A subset $U \in \mathbb{R}^n$ is *path-connected* if for any two points in U there is a path between them which is also contained in U. A *subpath* δ of a path γ is a path-connected subset of γ. A *segment* between two points $\underline{x}, \underline{y}$ is a path σ such that, for all $\alpha \in [0, 1]$,

$$\underline{\sigma}(\alpha) = \alpha\underline{x} + (1 - \alpha)\underline{y}.$$

A segment is sometimes also called a straight path.

A set $U \subseteq \mathbb{R}^n$ is *convex* if for any two points $\underline{x}, \underline{y}$ in U the segment σ between them is a subset of U. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *convex* over a path γ connecting two points $\underline{x}$ and $\underline{y}$ if and only if

$$\forall \alpha \in [0, 1]\ (f(\underline{\gamma}(\alpha)) \leqslant \alpha f(\underline{x}) + (1 - \alpha)f(\underline{y})).$$

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *convex* over a subset $U \subseteq \mathbb{R}^n$ if and only if U is convex and for each segment $\sigma \subset U$ connecting two points $\underline{x}$ and $\underline{y}$ in U, f is convex on σ.

Definition 2.1. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *superconvex* over a subset $U \subseteq \mathbb{R}^n$ if and only if: (i) U is a path-connected set; (ii) for any segment $\sigma \subset U$, f is convex on σ; and (iii) for each pair of points $\underline{x}, \underline{y} \in U$ there is a path $\gamma \subset U$ between $\underline{x}$ and $\underline{y}$ such that f is convex over γ.

Convex functions have the useful property that all local minima are also global on the convex set they are defined over. We shall show in the next section that superconvex functions also share this property, with the notable difference that the only condition on U is that it is path-connected (and thus, possibly nonconvex).

Example 2.1. One example of a nonconvex optimization problem which is superconvex is the following (see Fig. 1):

$$\begin{cases} \min_{x,y} x^2 + y^2 \\ y \leqslant \sqrt{1 - x^2} \\ y \geqslant \sqrt{\frac{9}{4} - x^2} - \frac{11}{10} \\ -1 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1. \end{cases}$$

The feasible region U defined by the constraints is nonconvex, but it is easy to check that x^2+y^2 is convex on all segments in U and for each pair of points in U there is a path in U such that the function is convex over that path. As a consequence, all minima of x^2+y^2 over U are global. In particular, there is only one local minimum at $x=0, y=2/5$, which is also global. On the other hand, the convex envelope of U is

$$\begin{cases} \min_{x,y} x^2+y^2 \\ y \leqslant \sqrt{1-x^2} \\ -1 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1, \end{cases}$$

that has minimum at $x=0, y=0$. Thus, superconvexity is a useful concept in that in some instances it is a weaker condition than convexity.

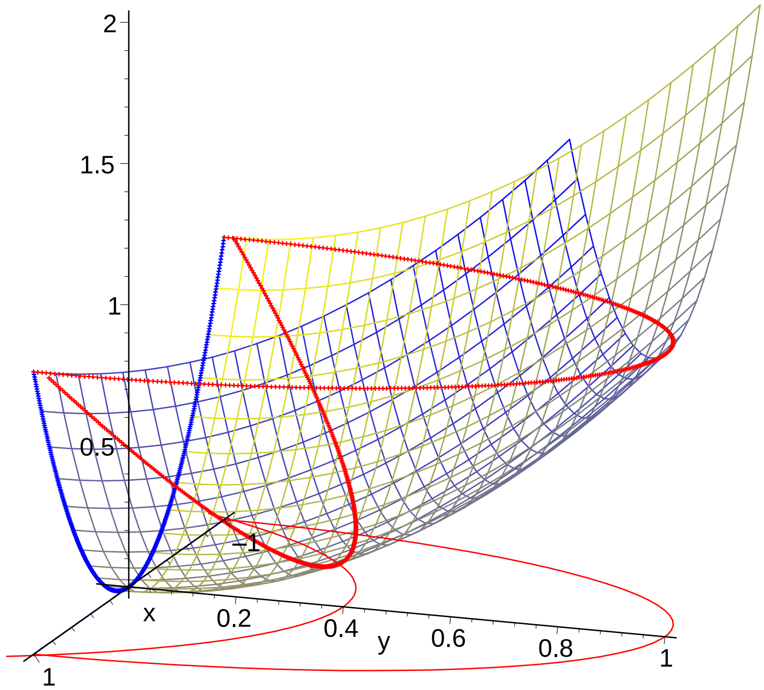

FIGURE 1. Example of a superconvex optimization problem which is not convex.

PROPOSITION 2.1. *Let $f:\mathbb{R}^n \to \mathbb{R}$ be superconvex over $U \subseteq \mathbb{R}^n$. For each convex subset $V \subseteq U$, f is convex over V.*

PROOF. This is trivial, as any segment in V is also a segment in U. □

PROPOSITION 2.2. *Let $f:\mathbb{R}^n \to \mathbb{R}$ a function which is non-constant over a subset $U \subseteq \mathbb{R}^n$. Then there exists at least one path $\gamma \subset U$ such that f is not convex over γ.*

PROOF. Since f is not constant, there are at least two points, $\underline{x}$ and $\underline{y}$ in U such that $f(\underline{x}) \neq f(\underline{y})$. Assume, without loss of generality, that $f(\underline{x}) < f(\underline{y})$. Let σ be the segment connecting $\underline{x}$ to $\underline{y}$, and τ be the inverse segment connecting $\underline{y}$ to $\underline{x}$. Consider the path

$$\underline{\gamma}(\alpha) = \begin{cases} \underline{\sigma}(2\alpha) & \alpha \in [0, 1/2) \\ \underline{\tau}(2\alpha - 1) & \alpha \in [1/2, 1] \end{cases}$$

Since for $\alpha = 1/2$ we have

$$f(\underline{\gamma}(1/2)) = f(\underline{y}) > f(\underline{x}) = f(\underline{\gamma}(0)) = f(\underline{\gamma}(1)),$$

f is clearly not convex over γ. □

Whilst the technique employed to prove the last proposition may seem to point out to a rather unlikely case instead of a general rule, it is true that for any "reasonable" function there are in fact many paths on which the function is not convex.

This may seem irrelevant since we put no condition on the function in the statement of the proposition above, nor on the set over which it is defined. The point we wish to stress is that even though a function may be convex on a set U, it is very likely that there will be a path connecting two points in U such that the curve is *not* convex on that path. The main implication of this point is the following:

> *It is not true that if f is convex over U then, for any path-connected subset V of U, f is superconvex on V.*

This, to practical purposes, is probably the weakest point of the whole concept of superconvexity, for its main purpose is to be a parallel notion to that of convexity, without requiring the underlying set to be convex. It now turns out that even though a function is convex over a set, we cannot be sure of its superconvexity over the subsets. It is therefore essential that we check that, for any given f, the path-connected subset $V \subseteq U$ we want f to be superconvex over satisfies the conditions imposed by the definition, i.e., that f is convex on any segment in V and that for any two points in V there is a path γ connecting them such that f is convex over γ.

Fig. 2 gives a hint as to which are the typical situations where this may happen: although the surface is convex on the (convex) set U, it is evidently non-convex over the path p.

3. The Notion of Minimum

First we need to recall some basic definitions from topology.

DEFINITION 3.1. Let U be a non-empty subset of $\mathbb{R}^n$.

(1) We say $\underline{x}$ is an *interior point* of U if and only if there is an open subset W of $\mathbb{R}^n$ such that $\underline{x} \in W$ and $W \subseteq U$. We call the set of all interior points of U the *interior* of U and we refer to it as $I(U)$.

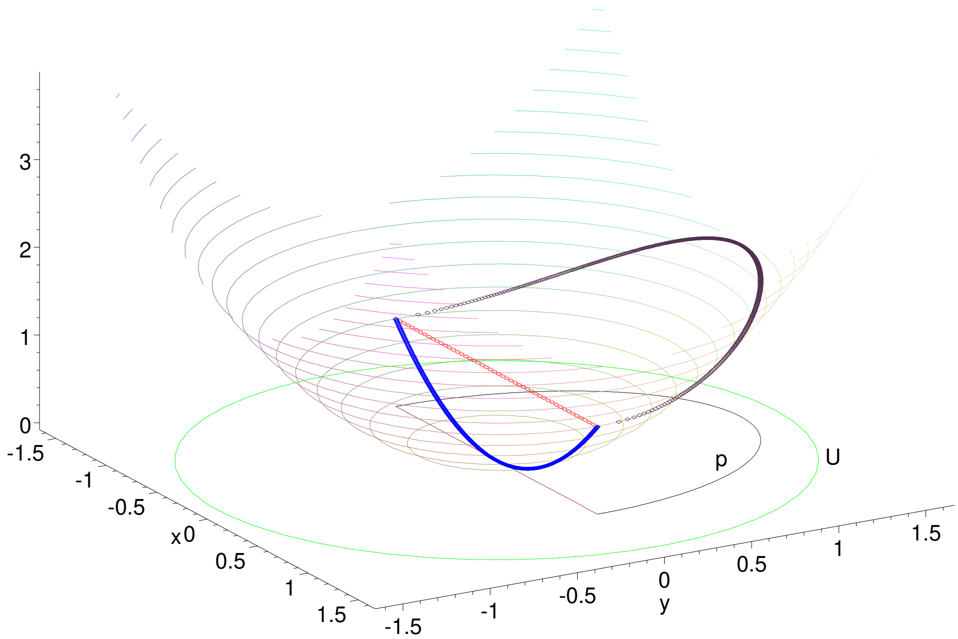

FIGURE 2. Example of a convex function which is not superconvex.

(2) We say $\underline{x}$ is a *border point* of U if and only if $\underline{x}$ is not an interior point and there is an open subset W of $\mathbb{R}^n$ such that $W \cap U \neq \emptyset$. We call the set of all border points of U the *border* of U and we refer to it as ∂U.
(3) We say $\underline{x}$ is a *limit point* of U if and only if for all open subsets $W \subseteq \mathbb{R}$ there is $\underline{y} \in W \cap U$ such that $\underline{y} \neq \underline{x}$.
(4) The *closure* $\overline{U}$ of U is the set $I(U) \cup \partial U$.

DEFINITION 3.2. A subset $U \subseteq \mathbb{R}^n$ is *bounded* if and only if, for any open set V such that $\underline{0} \in V$, there is a number $a > 0$ such that $U \subseteq \{a\underline{x} \mid \underline{x} \in V\}$.

Note that the above definition only makes sense when U is a subset of a topological linear space, i.e. a topological space which is also a vector space where vector addition and scalar multiplication are continuous under the topology. It turns out that $\mathbb{R}^n$ is a topological linear space.

LEMMA 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. For any bounded subset $U \subseteq \mathbb{R}^n$ we have $f(\overline{U}) = \overline{f(U)}$.*

PROOF. Let $t \in f(\overline{U})$ and claim $t \in \overline{f(U)}$. Since $t \in f(\overline{U})$, there is $\underline{x} \in \overline{U}$ such that $f(\underline{x}) = t$. In case $\underline{x} \in U$, then trivially $t \in f(U) \subseteq \overline{f(U)}$, so assume $\underline{x} \notin U$. Then $\underline{x}$ must be a limit point of U, hence there is a sequence $\langle \underline{x}_n \rangle \subseteq U$ which converges to $\underline{x}$. Since f is continuous the sequence $\langle f(\underline{x}_n) \rangle$ converges to $f(\underline{x}) = t$. Now, for all integers n we have $\underline{x}_n \in U$ and hence $f(\underline{x}_n) \in f(U)$. We conclude that t is a limit point for $f(U)$, and $t \in \overline{f(U)}$ as claimed.

Now let $t \in \overline{f(U)}$, and claim $t \in f(\overline{U})$. As before, in case $t \in f(U)$ the claim is trivially true, for there is $\underline{x} \in U$ such that $f(\underline{x}) = t$ and since $U \subseteq \overline{U}$, $\underline{x} \in \overline{U}$ whence the claim. So assume $t \notin f(U)$, so that t must be a limit point of $f(U)$. Hence there is a sequence $\langle t_n \rangle \subseteq f(U)$ which converges to t, and a sequence $\langle \underline{x}_n \rangle \subseteq U$ such that for all n, $f(\underline{x}_n) = t_n$. By the Bolzano-Weierstrass theorem, every bounded sequence has a convergent subsequence; and since U is bounded, the sequence $\langle \underline{x}_n \rangle$ is also bounded. So let $M \subseteq \mathbb{N}$ be an index subset such that the subsequence $\langle \underline{x}_m \mid m \in M \rangle$ converges to $\underline{x} \in \overline{U}$. But as $\langle t_n \rangle$ is a convergent sequence, all of its subsequences must converge to the same limit t. In particular its subsequence $\langle f(\underline{x}_m) \rangle$ converges to t, and by continuity of f we get $f(\underline{x}) = t \in f(\overline{U})$. □

One important issue when dealing with convexity is that of continuity. It is a well known fact that if a function is convex on a set, then it is also continuous on that set.

THEOREM 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex on $U \subseteq \mathbb{R}^n$. Then f is continuous on the interior $I(U)$.*

For the proof of this theorem we refer the reader to [**BS79**, Theorem 3.1.3, p. 82]. Superconvexity has this interesting property, as well.

LEMMA 3.2. *Let S be an open subset of $\mathbb{R}^n$. Then for all $\underline{x} \in S$ there is $\varepsilon_{\underline{x}} > 0$ such that*

$$S = \bigcup_{\underline{x} \in S} B(\underline{x}, \varepsilon_{\underline{x}})$$

where $B(\underline{x}, \varepsilon_{\underline{x}})$ is the open ball centred at $\underline{x}$ with radius $\varepsilon_{\underline{x}}$.

PROOF. Since S is open and $\mathbb{R}^n$ is a metric space with the usual metric, for all $\underline{x}$ there is $\varepsilon_{\underline{x}} > 0$ such that $B(\underline{x}, \varepsilon_{\underline{x}}) \subset S$. It is evident that

$$S \subseteq \bigcup_{\underline{x} \in S} B(\underline{x}, \varepsilon_{\underline{x}}).$$

Now pick $y \in \bigcup_{\underline{x} \in S} B(\underline{x}, \varepsilon_{\underline{x}})$. Then there must be $\underline{w} \in S$ such that $y \in B(\underline{w}, \varepsilon_{\underline{w}})$. But since $B(\underline{w}, \varepsilon_{\underline{w}}) \subset S$, also $y \in S$, as claimed. □

THEOREM 3.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be superconvex on $U \subseteq \mathbb{R}^n$. Then f is continuous on the interior of U.*

PROOF. Since $I(U)$ is open, by Lemma 3.2 it can be written as

$$I(U) = \bigcup_{\underline{x} \in I(U)} B(\underline{x}, \varepsilon_{\underline{x}}).$$

By Proposition 2.1, for all $\underline{x} \in I(U)$, f is convex on $B(\underline{x}, \varepsilon_{\underline{x}})$, hence by Theorem 3.1 f is continuous on $B(\underline{x}, \varepsilon_{\underline{x}})$, hence f is continuous at every $\underline{x} \in I(U)$, which means that f is continuous on $I(U)$. □

Note that in the last proof we have explicitly used the usual metric on $\mathbb{R}^n$, whereas we had previously employed techniques involving only the usual topology on $\mathbb{R}^n$. This means that the previous proof may not apply so generally as the most of the other proofs given throughout this paper.

We now turn our attention to the definition of a local minimum.

DEFINITION 3.3. Let T be a non-empty subset of $\mathbb{R}$. We say $m \in \mathbb{R}$ is a *lower bound* (respectively *upper bound*) for T if and only if for all $x \in T$ we have $x \geqslant m$ (resp. $x \leqslant m$). $m \in \mathbb{R}$ is a *strict lower bound* (respectively *strict upper bound*) for T if and only if for all $x \in T$ we have $x > m$ (resp. $x < m$).

Let T be any non-empty subset of $\mathbb{R}$ that has a lower bound. It is a well known theorem of basic analysis that T has a greatest lower bound [**Spi92**].

DEFINITION 3.4. Let T be a non-empty subset of $\mathbb{R}$. We say that $m \in \mathbb{R}$ is a *lower extremum* (resp. *upper extremum*) for T if and only if m is a lower bound (resp. upper bound) for T and $m \in T$. $m \in \mathbb{R}$ is *strict lower extremum* (resp. *strict upper extremum*) for T if and only if m is a strict lower bound (resp. strict upper bound) for T and $m \in T$.

DEFINITION 3.5. Let $f : \mathbb{R}^n \to \mathbb{R}$ and let $U \subseteq \mathbb{R}^n$. We say that f attains a *local minimum* (resp. *local maximum*) at $\underline{x}$ over U if and only if there is a non-empty open subset $V \subseteq U$ such that $\underline{x} \in \overline{V}$ and $f(\underline{x})$ is a lower extremum (resp. upper extremum) for $f(\overline{V})$. f attains a *strict local minimum* (resp. *strict local maximum*) at $\underline{x}$ if and only if there is a non-empty open subset $V \subseteq U$ such that $\underline{x} \in \overline{V}$ and $f(\underline{x})$ is a strict lower extremum (resp. strict upper extremum) for $f(\overline{V})$.

Note that in our definition the point $\underline{x}$ at which f attains the local maximum or the local minimum over U need not be in U: it suffices that it be in the closure $\overline{U}$.

THEOREM 3.3. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous and bounded below over a bounded subset $U \subseteq \mathbb{R}^n$. Then f attains a local minimum over U.*

PROOF. We know that the set $f(U)$ has greatest lower bound t and that $t \in \overline{f(U)}$. Since f is continuous, by Lemma 3.1 we gather $t \in f(\overline{U})$, so there must be $\underline{x} \in \overline{U}$ such that $f(\underline{x}) = t$. We claim that f attains a local minimum at $\underline{x}$ over U. Let $V = I(\overline{U})$, then $\underline{x} \in \overline{V}$ and $f(\underline{x})$ is a lower extremum for $f(\overline{V})$. $\square$

4. The Main Result

We will first present a classic result.

THEOREM 4.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex over U. Then f cannot attain more than one strict local minimum over U.*

Provided U is bounded, by Theorems 3.1, 3.3 and 4.1 a convex function f on U attains exactly one strict local minimum in U.

We now build up the proof to the result which is the aim of this paper: that if f is superconvex over U, f cannot attain more than one strict local minimum over U.

Lemma 4.1. *Let* $\underline{x} \neq \underline{y}$ *are two strict local minima for* $f : \mathbb{R}^n \to \mathbb{R}$ *over a path-connected set* $U \subseteq \mathbb{R}^n$. *Then for any path* γ *between* $\underline{x}$ *and* $\underline{y}$ *such that* $\gamma \subseteq U$ *there is a* $\underline{v} \in \gamma$ *such that*

$$f(\underline{v}) > \max\{f(\underline{x}), f(\underline{y})\}.$$

Proof. First we show that $\underline{x}$ and $\underline{y}$ are two strict local minima for f over γ. Since $\underline{x}$ is a strict local minimum for f over U, let V be an open set such that $\underline{x} \in \overline{V}$ and $f(\underline{x})$ is a strict lower extremum for $f(\overline{V})$. Let γ_1 be the subpath of γ given by $\overline{V} \cap \gamma$. Consider the interior of γ_1. It is a non-empty open subset of γ (it is open in the topology induced by $\mathbb{R}^n$ on γ), $\underline{x}$ is in its closure and $f(\underline{x})$, being a strict lower extremum for the whole of $f(\overline{V})$ is obviously a strict lower extremum for $f(\overline{\gamma_1})$. Hence $\underline{x}$ is a strict local minimum for f over γ, and the same goes for $\underline{y}$. Now let

$$\underline{w} = \begin{cases} \underline{x} & \text{if } f(\underline{x}) \geqslant f(\underline{y}) \\ \underline{y} & \text{if } f(\underline{x}) < f(\underline{y}). \end{cases}$$

Since $\underline{w}$ is a strict local minimum, there is a non-empty open subset $\gamma_1 \subseteq \gamma$ such that $\underline{w} \in \overline{\gamma_1}$ and $f(\underline{w})$ is a strict lower extremum of $f(\overline{\gamma_1})$. Since γ_1 is a path and one of its endpoints is $\underline{w}$, we can choose $\underline{v} = \underline{\gamma_1}(1/2)$ so that

$$f(\underline{v}) > f(\underline{w}) \geqslant \max\{\underline{\gamma}(0), \underline{\gamma}(1)\},$$

as claimed. □

Lemma 4.2. *For any* $a, b \in \mathbb{R}$, *for all* $\alpha \in [0, 1]$ *we have*

$$\alpha a + (1 - \alpha)b \leqslant \max\{a, b\}.$$

Proof. Suppose $\alpha \in (0, 1)$. If $a \geqslant b$ we have

$$\begin{aligned} a &\geqslant b \\ (1-\alpha)a &\geqslant (1-\alpha)b \\ a &\geqslant \alpha a + (1-\alpha)b \end{aligned}$$

as claimed. The case $b > a$ is similar. If $\alpha = 0$ we have $b \leqslant \max\{a, b\}$ and if $\alpha = 1$ we get $a \leqslant \max\{a, b\}$. Both cases are trivially true. □

Theorem 4.2. *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be superconvex over* $U \subseteq \mathbb{R}^n$. *If* $\underline{x}$ *is a strict local minimum for* f *over* U, *then it is the only one.*

Proof. Suppose, to get a contradiction, that there is $\underline{y} \neq \underline{x} \in U$ which is also a strict local minimum for f over U. Since f is superconvex, there is a path γ between $\underline{x}$ and $\underline{y}$ such that f is convex on γ. By applying Lemma 4.1 we find a $\underline{v} \in \gamma$ such that $f(\underline{v}) > \max\{f(\underline{x}), f(\underline{y})\}$, say $\underline{v} = \underline{\gamma}(\alpha_0)$. By definition of convexity on a path, we get that

$$\forall \alpha \in [0, 1]\ (f(\underline{\gamma}(\alpha)) \leqslant \alpha f(\underline{x}) + (1 - \alpha) f(\underline{y})).$$

In particular, for $\alpha = \alpha_0$, we get

$$f(\underline{v}) = f(\underline{\gamma}(\alpha_0)) \leqslant \alpha_0 f(\underline{x}) + (1 - \alpha_0) f(\underline{y}),$$

then we apply Lemma 4.2 to get

$$f(\underline{v}) \leqslant \max\{f(\underline{x}), f(\underline{y})\},$$

which contradicts Lemma 4.1. Hence the result. □

5. Conclusion

We have shown that every superconvex function attains one strict local minimum over its definition set, hence that minimum is also the global minimum. Superconvex functions may therefore be globally minimized by using a local optimization software rather than a global one. Since there are superconvex functions which are not convex, this fact can be used to calculate tighter guaranteed lower bounds to nonlinear problems for use with Branch-and-Bound type algorithms. The results obtained in this paper, however, are only theoretical in nature, and at this stage we are not able to provide an automatic procedure for constructing superconvex relaxations of arbitrary functions.

References

[Bon98] G. Bonanno. *On a class of functionals whose local minima are global*, J. Global Optimization 12 (1998), 101–104.

[BS79] M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley, Chichester, 1979.

[KF75] A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis*, Dover, New York, 1975.

[Lib04] L. Liberti, *Reformulation and Convex Relaxation Techniqies for Global Optimization*, Ph.D. thesis, Imperial College London, March 2004.

[Lib05] L. Liberti, *Writing global optimization software*, in: L. Liberti and N. Maculan (eds), *Global Optimization: from Theory to Implementation*, Kluwer, Dordrecht, to appear.

[Rap91] T. Rapcsak, *Geodesic convexity in optimization*, J. Optim. Theory Appl. 69 (1991), 169–183.

[SP99] E. M. B. Smith and C. C. Pantelidis, *A symbolic reformulation/spatial branch-and-bound algorithm for the global optimization of nonconvex MINLPs*, Computers and Chemical Engineering 23 (1999), 457–478.

[Spi92] M. Spivak, *Calculus*, 13th edition, Addison-Wesley, London, 1992.

DEI, Politecnico di Milano
P.zza L. Da Vinci 32, I-20133
Italia
liberti@elet.polimi.it

(Received 07 10 2003)