Journal of Global Optimization **14:** 79–101, 1999.

On Optimization Properties of Functions, with a Concave Minorant

OLEG KHAMISOV
Siberian Power Energy Institute, Lermontov Street 130, Irkutsk, 664033, Russia

(Received 14 August 1995; Accepted 31 July 1998)

Abstract. We give a definition of the class of functions with a concave minorant and compare these functions with other classes of functions often used in global optimization, e.g. weakly convex functions, d.c. functions, Lipschitzian functions, continuous and lower semicontinuous functions. It is shown that the class of functions with a concave minorant is closed under operations mainly used in optimization and how a concave minorant can be constructed for a given function.

Key words: Concave minorant (c.m.), d.c. Functions, Global minimum, Lower semicontinuous (l.s.c) functions, Weakly convex functions

1. Introduction

At the present time the widest class of objective functions considered in global optimization consists of locally Lipschitzian functions. Many methods have been developed for solving Lipschitzian optimization problems (see, for example, Evtushenko 1972; Galperin 1988; Hansen and Jaumard 1995; Meewella and Mayne 1988; Nefedov 1987; Pijavskii 1967; Pintér 1996; Wood 1992, and many others; see also Horst and Tuy 1996 and references therein). However, the efficiency of these methods essentially depends on the estimation accuracy of the Lipschitz constant. Some methods for the estimation of Lipschitz constants are given in Strongin (1978), Sukharev (1989) and Pintér (1996). Nevertheless, the problem of estimating a Lipschitz constant quite exactly is rather difficult itself.

Another class is formed by functions, which can be represented as a difference of two convex functions. Such functions are called d.c. functions (‘d.c.’ stands for the abbreviation of ‘difference of two convex’). A representation of a d.c. function as a difference of two convex functions is usually called a d.c. decomposition. It is known that each d.c. function is a locally Lipschitzian function. The inverse statement is not always true (see Hirriart-Urruty 1985). Hence, the d.c. functions form a very important subclass of locally Lipschitzian functions. The special structure of d.c. functions allows one to construct more effective global minimization methods than those for Lipschitzian functions (Horst and Tuy 1996; Thach and Tuy 1992; Tuy 1987). The main difficulty here is how to construct an effective d.c. decomposition for a given d.c. function. A general theory of d.c. optimization including duality, minimization methods and different ways of d.c. decomposition is described in Tuy (1995). Special d.c. minimization methods together with effective d.c. decompositions and duality with zero gap for the so called low rank nonconvex functions are described in Konno, et al. (1997) and Thach and Konno (1993). If a given function $f(x)$ is a low rank nonconvex function, then this means that $f(x)$ is almost convex or convex in almost all variables. So, the underlying convexity is the most important property here. Similarly, one can consider high rank nonconvex functions or low rank convex functions, i.e. functions that are nonconvex in almost all variables. It seems that in this case the most important and useful property is concavity. For example, it is easier to find a global minimum of a concave function over a simplex under a reverse convex constraint than to find a global minimum of a convex function over the same feasible set.

In this paper we study functions that can be expressed as the maximum of a family of continuous concave functions:

$$f(x) = \max_{y \in R} \varphi(x, y), \tag{1.1}$$

where $R \subset E^n$ is a compact set and $\varphi(x, y)$ is a continuous function which is concave in x.

Similar approaches were developed earlier by different authors (see Baritompa 1994; Baritompa and Cutler 1994; Breiman and Cutler 1993). Dolecki (1978) tried to develop a duality theory based on representation (1.1) under some additional (and restrictive) assumptions.

Here we describe only practically implementable approaches. In Bulatov (1987), extensions of cutting plane methods from concave minimization to minimization of functions of type (1.1) were given. The most similar approach was described in Norkin (1992). In our paper we present some new results partly connected with the investigations of V. Norkin. The concept of a support function was introduced in Zabotin and Khabibulin (1975) in order to develop general optimality conditions for constrained extremal problems. A survey of approaches based on representation (1.1), where $\varphi(x, y)$ is not concave in x is given in Avriel et al. (1988).

The paper is organized in the following way. In Section 2 we give the definition of functions with a concave minorant and compare them with other classes of functions. Section 3 describes how a concave minorant can be constructed for a given function. Section 4 is devoted to minimization of functions with a concave minorant over a compact set. In Section 5 we show how a dual lower bound can be used for the problem described in Section 4. Section 6 contains several examples showing different properties of special concave minorants.

2. Definition and comparison with other classes of functions

Let a set $R \subset E^n$ and a real valued function $f(x),\quad f : R \to E^1$ be given.

DEFINITION 2.1. A function $f(x)$ is said to have a concave minorant on R if there exists a function $\varphi(x, y),\quad \varphi : E^n \times R \to E^1$, continuous in x for any fixed y, such that

1. $\varphi(x, y)$ is concave in x;
2. $f(x) \geqslant \varphi(x, y) \quad \forall (x, y) \in R \times R;$ (2.1)
3. $f(y) = \varphi(y, y) \quad \forall y \in R.$ (2.2)

The function $\varphi(x, y)$ is called a concave minorant of $f(x)$, constructed at the point $y \in R$. The set of all functions $f(x),\ f : R \to E^1$, which have a concave minorant on R is denoted by $CM(R)$ and each function $f \in CM(R)$ is called c.m. function on R.

Below we assume that $R \subset E^n$ is a compact set. The functional class $CM(R)$ is quite large, since it is not difficult to see that any Lipschitzian function $f(x)$ is also a c.m. function with

$$\varphi(x, y) = f(y) - L\|x - y\|, \tag{2.3}$$

where L is the Lipschitz constant.

Let $f \in CM(R)$ with the concave minorant $\varphi(x, y)$. Then obviously

$$f(x) = \max_{y \in R} \varphi(x, y). \tag{2.4}$$

Since $\varphi(x, y)$ is continuous in x, then due to representation (2.4), $f(x)$ must be a lower semicontinuous (l.s.c.) function as an upper envelope of a family of continuous functions. Hence, we can formulate the following Proposition.

PROPOSITION 2.1. *Every function $f \in CM(R)$ is l.s.c. on R.*

Thus, a c.m. function is a l.s.c. function which can be considered as a pointwise maximum of a family of continuous concave functions. Then the question arises: what is the difference between a c.m. function and a l.s.c. function? The following example shows, that not every l.s.c. function is a c.m. function.

EXAMPLE 2.1. Define on $R = [0, 2] \subset E^1$ the univariate function $f(x)$

$$f(x) = \begin{cases} 1 - \sqrt{1 - x^2}, & 0 \leqslant x \leqslant 1 \\ 2 - \sqrt{1 - (x - 2)^2}, & 1 < x \leqslant 2. \end{cases}$$

It is not difficult to see, that $f(x)$ is l.s.c. on R, but it is impossible to construct a concave minorant of $f(x)$ at the point $y = 1$. Hence, there are some ‘points of dissimilarity’ between c.m. and l.s.c. functions.

DEFINITION 2.2. Let $R \subset E^n$ be a compact set and $f(x)$ a l.s.c. function on R. The point $y \in R$ is called a c.m. point of $f(x)$ if $f(x)$ has a concave minorant at y.

Let us introduce the following set:

$$\text{dom}\,(f, R) = \{x \in R : f(x) < +\infty\}$$

and let us denote by $cl D$ the closure of a set $D \subset E^n$.

THEOREM 2.1. *The set of all c.m. points of a l.s.c. function $f(x)$ on R is dense in $dom(f, R)$.*

Proof. Recall that a set A is dense in a set B if $cl A \supset B$. Denote by M the set of all c.m. points of f(x) on R. Let $y \in dom(f, R)$. To prove that $y \in cl M$ we use the Ekeland Theorem (Ekeland and Temam 1976) which states the following. Let for some point $z \in R$ and $\varepsilon > 0$ the inequality

$$f(z) \leqslant \inf\{f(x) : x \in R\} + \varepsilon \tag{2.5}$$

be fullfiled, then for any $\lambda > 0$ there exists a point $v \in R$, such that

$$f(v) \leqslant f(z), \tag{2.6}$$

$$\|z - v\| \leqslant \lambda \tag{2.7}$$

and for all $x \in R$ we have

$$f(x) + \left(\frac{\varepsilon}{\lambda}\right)\|x - v\| \geqslant f(v).$$

Since $y \in \text{dom}(f, R)$ then inequality (2.5) holds for $z = y$ and $\varepsilon = f(y) - \inf\{f(x) : x \in R\}$. Due to the Ekeland Theorem for any $\lambda_k = 1/k$ there exists the point $v^k \in R$, such that

$$\|y - v^k\| \leqslant \lambda_k, \tag{2.8}$$

$$f(x) + \left(\frac{\varepsilon}{\lambda_k}\right)\|x - v^k\| \geqslant f(v^k), \quad \forall x \in R. \tag{2.9}$$

It follows from (2.9) that at each point v^k the function $f(x)$ has a concave minorant

$$\varphi(x, v^k) = f(v^k) - \left(\frac{\varepsilon}{\lambda_k}\right)\|x - v^k\|, k = 1, 2, \ldots$$

Hence, v^k is a c.m. point of $f(x)$ on R. By virtue of (2.8) point y is the limit point of the sequence $\{v^k\}$ of c.m. points, so $y \in cl M$. □

Thus, an arbitrary l.s.c. function on R has a concave minorant almost everywhere and if $f(x)$ does not have a concave minorant at some point $z \in R$, then due to (2.6), (2.7) there exists a c.m. point $v \in R$, 'not far from z', with a possibly less functional value. From (2.8) and (2.9) we can draw the following conclusion.

PROPOSITION 2.2. *Every local minimum of a l.s.c. function over the compact set R is a c.m. point.*

Now we can say that there is a very close relationship between c.m. functions and l.s.c. functions. For a c.m. function we have the max representation (2.4). How strongly can it be destroyed when we move from a c.m. function to a l.s.c. one? The following result is in fact the Korovkin theorem proved in Kutateladze and Rubinov (1976).

THEOREM 2.2. *A finite function $f : R \to E^1$, is l.s.c. on a compact set R if and only if it can be represented as the upper envelope (pointwise supremum) of a nonempty family of continuous concave functions $\varphi(x, y),\ y \in Y$:*

$$f(x) = \sup_{y \in Y} \varphi(x, y). \tag{2.10}$$

□

From the aforesaid, we can conclude that the difference between a c.m. function and a l.s.c. function is exactly the difference between the operations max and sup over a family of continuous concave functions.

Now we can also see an analogy between l.s.c. convex functions and l.s.c. functions. Both are obtained as pointwise supremums of some auxiliary families of continuous functions : for l.s.c. convex functions it is the family of affine functions and for l.s.c. (nonconvex) functions it is the family of concave functions.

Let us now see how we can move (more or less gradually) from the convex l.s.c. functions to the l.s.c. functions.

Our first step leads to the so-called ρ-convex functions (Vial 1983).

DEFINITION 2.3. Let $f : D \to E^1$ be a real valued function on a convex subset D of E^n. f is said to be ρ-convex if there exists some real number ρ, such that $\forall x_1, x_2 \in D, \forall\lambda \in [0, 1]$

$$f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda) f(x_2) - \rho\lambda(1-\lambda)\|x_1 - x_2\|^2.$$

if $\rho > 0$, the function is said to be strongly convex;
if $\rho = 0$, the function is convex;
if $\rho < 0$, the function is said to be weakly convex.

We are interested in the third (nonconvex, $\rho < 0$) case, i.e. in weakly convex functions. The following two results are important for our consideration (for the proofs, see Vial 1983).

PROPOSITION 2.3. *A function $f : E^n \to E^1$ is ρ-convex if and only if for any $x \in E^n$ there is some $\xi \in E^n$ such that*

$$f(x) \geqslant f(y) + \xi^T(x - y) + \rho\|x - y\|^2, \quad \forall y \in E^n.$$

Hence, a weakly convex function $f(x)$ is a c.m. function with the minorant

$$\varphi(x, y) = f(y) + \xi^T(x - y) + \rho\|x - y\|^2, \rho < 0. \tag{2.11}$$

Therefore, a weakly convex function is an upper envelope of the family of concave functions and each concave auxiliary function is of the type (2.11).
From the following statement we can see also how much convexity was ‘lost’ in the movement from convex functions to weakly convex functions.

PROPOSITION 2.4. *A function $f : E^n \to E^1$ is ρ-convex if and only if there exists a convex function $h(x)$, $h : E^n \to E^1$ such that*

$$f(x) = h(x) + \rho\|x\|^2. \tag{2.12}$$

We still have the convexity property, but now in the two opposite directions, since for weakly convex functions the constant ρ is negative. In other words we have ‘some amount’ of convexity ‘for’ and the ‘other amount’ of convexity ‘against’ and this ‘negative amount’ of convexity is of a rather particular form: $\rho\|x\|^2,\ \rho < 0$.

Let us consider now the following generalization of the weakly convex functions : in decomposition (2.12) we substitute the term $\rho\|x\|^2$ by an arbitrary concave function. In this way we arrive at the very important concept of d.c. functions.

DEFINITION 2.4. Let $D \subset E^n$ be a convex set. A function $f : D \to E^1$ is called d.c. on D if there are two convex functions $p : D \to E^1$, $q : D \to E^1$ such that

$$f(x) = p(x) - q(x),\ \forall x \in D. \tag{2.13}$$

A function that is d.c. on E^n will be called a d.c. function.

The main properties of d.c. functions were studied in Hirriart-Urruty (1985), Horst and Tuy (1996) and Tuy (1995). Let $f(x)$ be a d.c. function and the decomposition (2.13) be known. From the convexity of $p(x)$ it follows that the function

$$\varphi(x, y) = p(y) + \xi^T(x - y) - q(x), \quad \xi \in \partial p(y), \tag{2.14}$$

where $\partial p(y)$ denotes the subdifferential of p at the point y, is a concave minorant of $f(x)$ at y. Therefore, every d.c. function is a c.m. function.

Let us formulate analogue of Theorem 2.2 for d.c. functions. Consider the following generalization of function (2.14)

$$\psi(x, y) = c(y)^T(x - y) + r(y) - q(x), \tag{2.15}$$

where $c(y) \in E^n, r(y) \in E^1,\ y \in Y$, Y is some nonempty set and $q(x)$ is a continuous convex function. Note, that in the definition of $\psi(x, y)$ we use arbitrary functions $r(y)$ and $c(y)$. Taking the supremum of $\psi(x, y)$ over $y \in Y$, we obtain

$$\begin{aligned}\sup_{y\in Y}\psi(x, y) &= \sup_{y\in Y}\{c(y)^T(x - y) + r(y) - q(x)\}\\ &= \sup_{y\in Y}\{c(y)^T(x - y) + r(y)\} - q(x)\\ &= v(x) - q(x),\end{aligned} \tag{2.16}$$

where

$$v(x) = \sup_{y \in Y} \{c(y)^T (x - y) + r(y)\}$$

is a convex function as the pointwise supremum of a family of affine functions. Combining (2.14), (2.15) and (2.16), we have the following:

PROPOSITION 2.5. *A function $f : E^n \to E^1$ is a d.c. function if and only if*

$$f(x) = \sup_{y \in Y} \psi(x, y),$$

where $\psi(x, y)$ is a function of type (2.15).

The next example shows that there is a gap between d.c. and c.m. functions.

EXAMPLE 2.2. Consider the following univariate function

$$f(x) = \begin{cases} 1, & x < 0 \\ 0, & x = 0 \\ 2, & x > 0 \end{cases}$$

It is easy to see that the function

$$\varphi(x, y) = \begin{cases} \min\{1, x/y\}, & y < 0 \\ 0, & y = 0 \\ \min\{2, 2x/y\}, & y > 0 \end{cases}$$

is a concave minorant for $f(x)$. Hence, $f(x)$ is a c.m. function, but not a d.c. function, since every d.c. function is locally Lipschitzian (Hirriart-Urruty 1985) and, therefore, continuous.

Denote by $\mathrm{Lip}(R, L)$ the set of all Lipschitzian functions defined on a compact set R with the Lipschitz constant L. We have already seen that each $f \in \mathrm{Lip}(R, L)$ is also a c.m. function. The inverse statement is not true (see Example 2.2). On the other hand, since every l.s.c. function is 'almost' a c.m. function, some connection between locally Lipschitzian functions and c.m. functions can be described by the following proposition.

PROPOSITION 2.6. *Let $f : R \to E^1$ be a l.s.c. function on R, $f(x) > -\infty, \forall x \in R$, R be a compact set. Then there exists a sequence of functions $f_k : R \to E^1$ such that $\forall x \in R$*

$$f_1(x) \leqslant f_2(x) \leqslant \dots ,$$

$$f(x) = \lim_{k \to \infty} f_k(x)$$

and $f_k \in Lip(R)$, where $Lip(R) = \cup_{L>0} Lip(R, L)$.

The proof of this proposition can be found in Natanson (1977, Theorem 10, pp. 467–468).

Note that continuous functions as a subclass of l.s.c. function can be represented as a pointwise supremum of a family concave functions $\varphi(x, y),\ \ y \in Y,\ Y \neq \emptyset$. The theorem below (Norkin 1992) gives conditions under which the pointwise supremum of a family of concave functions is a continuous function.

THEOREM 2.3. *Let function* $f : E^n \to E^1$ *be representable as*

$$f(x) = \sup_{y \in Y} \varphi(x, y),$$

where $\varphi(x, y),\ y \in Y, Y \neq \emptyset$ *is a family of equicontinuous concave functions. Then* $f(x)$ *is continuous.* □

Let us make some concluding remarks.

We started from a l.s.c. convex function, which is the upper envelope of a family of affine functions. Then we saw that a weakly convex function is the upper envelope of a family of concave continuous functions of the type (2.11). Substituting the term $\rho\|x - y\|^2$ in (2.11) by an arbitrary concave function we obtained concept of d.c. functions. Only the linear part of the concave minorant (2.14) depends on y. If the concave part also depends on y, then the upper envelope of the family of concave equicontinuous functions is continuous and the upper envelope of a family of arbitary concave continuous functions is lower semicontinuous.

In other words, if we may say that a convex l.s.c. function has a ‘positive amount of convexity’. Then the weakly convex function still has a ‘positive amount of convexity’, but it has some ‘negative amount of convexity’ (or some ‘amount of concavity’), although of a rather particular form. For the d.c. function the ‘negative amount of convexity’ is allowed to be of a rather general form. We say that for the d.c. function ‘positive’ and ‘negative amount of convexity’ have equal rights of ‘existence’. For the continuous and l.s.c. functions we have information only about the ‘negative amount of convexity’.

The gap between Lipschitzian and l.s.c. functions is (more or less) characterized by Proposition 2.6.

In each case the connection with the concept of a c.m. function was shown. From the practical point of view, c.m. functions are very general, but they can be used in the situations when no additional information exept the c.m. property is available.

3. Main properties of c.m. functions

Let a family of c.m. functions be given. We will say that an operation (e.g. supremum, composition, etc.) preserves the c.m. property of the family of c.m. functions if the result of the operation is still a c.m. function. The purpose of this Section is to describe operations, which are often used in mathematical programming and preserve the c.m. property.

The next proposition follows directly from the Definition 2.1. We use the term nonnegative combination to denote a linear combination with nonnegative coefficients.

PROPOSITION 3.1. *Let c.m. functions* $f(x)$, $f_i(x)$, $i = 1, \dots, m$ *be given. Then the following statements are true.*
(i) Any nonnegative combination of the functions $f_i(x)$ *is a c.m. function;*
(ii) $\max_{1 \leqslant i \leqslant m} f_i(x)$ *and* $\min_{1 \leqslant i \leqslant m} f_i(x)$ *are c.m. functions;*
(iii) $f^+(x) = \max\{0, f(x)\}$, $f^-(x) = \min\{0, f(x)\}$ *are c.m. functions.*

Let us introduce the following:

DEFINITION 3.1. A function $f : E^n \to E^1$ is called monotoneously nondecreasing if for any $x, y \in E^n$ such that $x_i \geqslant y_i$, $i = 1, \dots, n$, f satisfies $f(x) \geqslant f(y)$.

If $g_1(x), \dots, g_m(x)$ are concave functions on E^n and $h : E^m \to E^1$ is a monotoneously nondecreasing concave function, then the composite function $f(x) = h(g_1(x), \dots, g_m(x))$ is a concave function (Bazaraa and Shetty 1979). From this property we easily derive

PROPOSITION 3.2. *Let* $g_i : E^n \to E^1$, $i = 1, \dots, m$ *and* $h : E^m \to E^1$ *be c.m. functions. If for any fixed* $y \in E^m$ *there exists a monotoneously nondecreasing concave minorant* $\varphi_h(x, y)$ *of the function* h, *then* $f(x) = h(g(x))$ *is a c.m. function.*

Proof. Let $\varphi_i(x, y)$ be concave minorants of the functions $g_i(x)$, i.e. $\forall (x, y) \in E^n \times E^n$

$$g_i(x) \geqslant \varphi_i(x, y), \quad i = 1, \dots, m, \tag{3.1}$$

$$g_i(y) = \varphi(y, y), \quad i = 1, \dots, m \tag{3.2}$$

and let $\varphi_h(u, v)$, $(u, v) \in E^m \times E^m$ be in u a monotoneously nondecreasing concave minorant of $h(u)$. Denote by $g(x)$ the vector function $g(x) = (g_1(x), \dots, g_m(x))$ and by $\varphi(x, y) = (\varphi_1(x, y), \dots, \varphi_m(x, y)$. Then $\forall (x, y)$

$$f(x) = h(g(x)) \geqslant \varphi_h(g(x), g(y)) \tag{3.3}$$

From the monotonicity of φ_h and (3.1) we have

$$\varphi_h(g(x), g(y)) \geqslant \varphi_h(\varphi(x, y), g(y)) = \varphi_f(x, y), \tag{3.4}$$

where $\varphi_f(x, y)$ is a concave function continuous in x. Similarly,

$$f(y) = h(g(y)) = \varphi_h(g(y), g(y)) = \varphi_h(\varphi(y, y), g(y)) = \varphi_f(y, y). \tag{3.5}$$

Hence, $\varphi_f(x, y)$ is a concave minorant of $f(x)$ and, therefore, $f(x)$ is a c.m. function. □

COROLLARY 3.1. *Let* $f_i : E^n \to E^1,\ i = 1, \dots, k$ *be c.m. functions. Then the function*

$$f(x) = \sum_{i=1}^{k} [\max\{0, f_i(x)\}]^q, \quad q \geqslant 1$$

is a c.m. function.

COROLLARY 3.2. *Let* $g_0(t)$ *be a positive univariate c.m. function. Then the functions* $\ln(g_0(t)),\ \exp(g_0(t)),\ \sqrt{(g_0(t))}$ *are c.m. functions.*

Propositions 3.1 and 3.2 describe quite general properties of c.m. functions. If we have some (nonconvex) function how can we recognize whether this function is a c.m. function and, moreover, if yes, then how can we construct its concave minorant? We conclude this Section with a description of a rather wide subclass of the class of c.m. functions and give rules for the construction of a concave minorant for a function from this subclass.

DEFINITION 3.2. If a function $f : R \to E^1,\ R \subset E^n$ satisfies

$$f \in CM(R), \qquad -f \in CM(R), \tag{3.6}$$

then f is called a c.m. symmetric function on R.

The set of all c.m. symmetric functions on R is denoted by $CMS(R)$. If $f \in CMS(E^n)$, then f is called a c.m. symmetric function. It follows from (3.6) that for every $f \in CMS(R)$ there exist a function $\varphi^-(x, y), \varphi^- : E^n \times R \to E^1$ continuous and concave in x and a function $\varphi^+(x, y), \varphi^+ : E^n \times R \to E^1$ continuous and convex in x, such that

$$\varphi^-(x, y) \leqslant f(x) \leqslant \varphi^+(x, y), \quad \forall x \in R,\ y \in R, \tag{3.7}$$

$$\varphi^-(y, y) = f(y) = \varphi^+(y, y), \quad \forall y \in R. \tag{3.8}$$

Function $\varphi^+(x, y)$ is called a convex majorant and $\varphi^-(x, y)$ is called a concave minorant of the function $f(x)$. Because of Proposition 2.1 a c.m. symmetrical function is continuous.

PROPOSITION 3.3. *Let* $f \in CMS(R)$ *and* $f_i \in CMS(R)$, $i = 1, \dots, m$, $m > 1$, *and* R *be a compact set. Then*

(i) $\sum_{i=1}^{m} \lambda_i f_i \in CMS(R), \lambda_i \in E^1$;
(ii) $f^2 \in CMS(R)$;
(iii) $f_1 \cdot f_2 \in CMS(R)$;
(vi) if $f(x) > 0, \forall x \in R$, *then* $1/f(x) \in CMS(R)$;

Proof. Denote by $\varphi^+(x,y)$, $\varphi^-(x,y)$ a convex majorant and a concave minorant of $f(x)$ and by $\varphi_i^+, \varphi_i^-(x,y)$ a convex majorant and a concave minorant of the functions $f_i(x)$, respectively.

(i) It is not difficult to see that the function

$$\sum_{i:\lambda_i>0} \lambda_i \varphi_i^-(x,y) + \sum_{i:\lambda_i<0} \lambda_i \varphi_i^+(x,y)$$

is a concave minorant and

$$\sum_{i:\lambda_i>0} \lambda_i \varphi_i^+(x,y) + \sum_{i:\lambda_i<0} \lambda_i \varphi_i^-(x,y)$$

is a convex majorant of $\sum_{i=1}^m \lambda_i f_i(x)$.

(ii) Since

$$[f(x) - f(y)]^2 \geqslant 0, \quad \forall x, y \in R,$$

we have

$$f^2(x) \geqslant 2f(x)f(y) - f^2(y).$$

Hence, a concave minorant of $f^2(x)$ can be determined in the following way:

$$\psi^-(x,y) = \begin{cases} 2f(y)\varphi^-(x,y) - f^2(y), & f(y) > 0 \\ 2f(y)\varphi^+(x,y) - f^2(y), & f(y) < 0 \end{cases}.$$

Since $f(x)$ is continuous, there exist numbers a and b such that

$$a < f(x) < b, \quad \forall x \in R. \tag{3.9}$$

From the convexity of function $g(t) = t^2$ we have for any fixed $\hat{t} \in [a,b]$

$$g(t) \leqslant \max\{l_1(t), l_2(t)\}, \quad \forall t\ [a,b] \tag{3.10}$$

where

$$l_1(t) = \hat{t}^2 + \frac{a^2 - \hat{t}^2}{a - \hat{t}}(t - \hat{t}), \tag{3.11}$$

$$l_2(t) = \hat{t}^2 + \frac{b^2 - \hat{t}^2}{b - \hat{t}}(t - \hat{t}). \tag{3.12}$$

Let $\hat{t} = f(y)$ for some fixed $y \in R$. Then due to (3.10)

$$f^2(x) = g(f(x)) \leqslant \max\{l_1(f(x)), l_2(f(x))\}. \tag{3.13}$$

Denote by $K_1 = a^2 + f^2(y)$, by $K_2 = b^2 + f^2(y)$ and consider functions

$$p_1(x, y) = \begin{cases} f^2(y) + K_1(\varphi^+(x, y) - f(y)), & K_1 > 0 \\ f^2(y) + K_1(\varphi^-(x, y) - f(y)), & K_1 \leqslant 0 \end{cases}, \tag{3.14}$$

$$p_2(x, y) = \begin{cases} f^2(y) + K_2(\varphi^+(x, y) - f(y)), & K_2 > 0 \\ f^2(y) + K_2(\varphi^-(x, y) - f(y)), & K_2 \leqslant 0 \end{cases}. \tag{3.15}$$

It is obvious, that functions p_i, $i = 1, 2$ are convex in x for fixed y. Now, from (3.11)–(3.15) we obtain that the function

$$\psi^+(x, y) = \max\{p_1(x, y), p_2(x, y)\}$$

is a convex majorant of $f^2(x)$.

(iii) Since

$$f_1(x) \cdot f_2(x) = \frac{[f_1(x) + f_2(x)]^2}{4} - \frac{[f_1(x) - f_2(x)]^2}{4},$$

this statement follows from the proved statements (i) and (ii).

(vi) Let $f(x) > 0 \forall x \in R$. Due to the convexity of function $g(t) = 1/t$ when $t > 0$, we have

$$g(t) \geqslant \frac{1}{\hat{t}} - \frac{1}{\hat{t}^2}(t - \hat{t}), \hat{t} > 0. \tag{3.16}$$

Substituting t by $f(x)$ and $\hat{t}$ by $f(y)$ we obtain

$$\frac{1}{f(x)} \geqslant \frac{1}{f(y)} - \frac{1}{f^2(y)}(f(x) - f(y)).$$

Hence, a concave minorant of $1/f(x)$ can be defined as

$$\tau^-(x, y) = \frac{1}{f(y)} - \frac{1}{f^2(y)}(\varphi^+(x, y) - f(y)).$$

The convex majorant $\tau^+(x, y)$ can be derived in a similar way. □

Note, that the proof of the theorem can be practically used to construct a concave minorant for almost every c.m. symmetrical function given the minorants of the underlying functions.

4. An approach to the global minimization of c.m. function over a compact set

We introduce here the notion of a c.m. programming problem that is based on the definition of a c.m. (c.m. symmetrical) function. Many other authors earlier considered approaches based on the representation of the objective function (as well as of the constraints) as an upper envelope of some set of functions. In this paper we consider only continuous concave auxiliary functions. A very similar approach was described in Norkin 1992.

Consider the following mathematical programming problem

$$\min f(x), \tag{4.1}$$

$$g_i(x) \leqslant 0, \quad i = 1, \dots, m, \tag{4.2}$$

$$h_i(x) = 0, i = 1, \dots, l, \tag{4.3}$$

$$x \in R, \tag{4.4}$$

where $R \subset E^n$ is a compact set, $f, g_i \in CM(R),\ i = 1, \dots, m,\ h_i \in CMS(R),\ i = 1, \dots, l$.

Assume that the feasible set in (4.1)–(4.4) is nonempty. Since every c.m. function on R is l.s.c. on R and a c.m. symmetrical function on R is continuous, the feasible domain in problem (4.1)–(4.4) is compact and due to the Weierstrass Theorem problem (4.1)–(4.4) has a finite solution. We call this problem a c.m. programming problem. Obviously, a c.m. programming problem can be highly multiextremal.

First we consider a more simple case when $m = 0,\ l = 0$ and R is convex. Let $x^1, x^2, \dots, x^k$ be some points in R. Then, it follows from (2.2) that

$$f(x) \geqslant \max_{1 \leqslant j \leqslant k} \varphi(x, x^j) = f_k(x), \quad \forall x \in R \tag{4.5}$$

where $\varphi(x, y)$ is a concave minorant of $f(x)$. We call the problem

$$\min\{f_k(x) : x \in R\} \tag{4.6}$$

an approximating c.m. problem since due to (4.5)

$$f^* = \min_{x \in R} f(x) \geqslant \min_{x \in R} f_k(x) = f_k^*. \tag{4.7}$$

Problem (4.6) is again a c.m. programming problem and, therefore, multiextremal, but here we have the advantage in the special form of the objective function. More exactly, it was shown in Tuy (1987) that $f_k(x)$ is a d.c. function since

$$f_k(x) = f_k^+(x) - f_k^-(x),$$

where

$$f_k^+(x) = -\min_{1 \leqslant i \leqslant k} \left\{ \sum_{j=1, j \neq i}^{k} \varphi(x, x^j) \right\},$$

$$f_k^-(x) = -\sum_{j=1}^{k} \varphi(x, x^j)$$

and both $f_k^+(x)$ and $f_k^-(x)$ are convex. Introducing now an additional variable x_{n+1} we can reduce the approximating c.m. problem to the following one:

$$\min\{x_{n+1} - f_k^-(x)\}, \tag{4.8}$$

$$f_k^+(x) \leqslant x_{n+1}, \tag{4.9}$$

$$x \in R. \tag{4.10}$$

The feasible domain in problem (4.8)–(4.10) is convex and the objective function is concave. Hence, (4.8)–(4.10) is a concave programming problem. Thus, the problem of global minimization of a c.m. function over a convex set can be approximated by a sequence of concave programming problems. This fact seems to be the natural generalization of the approximation of a convex programming problem by a sequence of linear programs. Therefore, we can say that in general almost each mathematical programming problem can be approximated by a sequence of auxiliary problems, which are not more complicated than the concave programming problem.

A general c.m. programming problem can be transformed to the problem of the minimization of a c.m. function over a compact set by using, for example, the Huard's method of centers.

5. A dual lower bounding for general c.m. programming problems

In this section we assume that the set R in (4.4) is a polytope with known vertices, $R = co\{z^1, \dots, z^N\}, z^i \in E^n, i = 1, \dots, N, N \geqslant n + 1$. The system of constraints (4.2)–(4.3) in problem (4.1)–(4.4) can be rewriten as one constraint

$$g(x) \leqslant 0, \tag{5.1}$$

with

$$g(x) = \max\{g_1(x), \dots, g_m(x), \mid h_1(x) \mid, \dots, \mid h_l(x) \mid\}.$$

Then, problem (4.1)–(4.4) is equivalent to

$$\min f(x), \tag{5.2}$$

$$g(x) \leqslant 0, \tag{5.3}$$

$$x \in R \tag{5.4}$$

and $f, g \in CM(R)$. Denote by $\varphi(x, y)$ a concave minorant of the function $f(x)$ and by $\psi(x, y)$ a concave minorant of the function $g(x)$.

Let $x^0 \in R$ be an arbitrary point. The optimal value φ_0^* of the problem

$$\min \varphi(x, x^0), \tag{5.5}$$

$$g(x) \leqslant 0, \tag{5.6}$$

$$x \in R \tag{5.7}$$

gives a lower bound for the optimal value f^* of problem (5.2)–(5.4), $\varphi_0^* \leqslant f^*$. Moreover, if x^1 is an optimal solution of (5.5)–(5.7), $\varphi(x^1, x^0) = \varphi_0^*$, then $g(x^1) \leqslant 0$. In other words x^1 is feasible, which is not less important than optimality, since the constraint function $g(x)$ can be highly nonconvex. From the feasibility of x^1 we also get the upper bound: $f^* \leqslant f(x^1) = f^1$. Consider now the problem

$$\min \varphi(x, x^1), \tag{5.8}$$

$$g(x) \leqslant 0, \tag{5.9}$$

$$f(x) \leqslant f^1 - \varepsilon, \tag{5.10}$$

$$x \in R \tag{5.11}$$

where ε is a given accuracy. Let x^2 be an optimal solution of (5.8)–(5.11). Then $f(x^2) \leqslant f(x^1) - \varepsilon$, so the current soluion is improved. If problem (5.8)–(5.11) is infeasible then x^1 is an ε-optimal solution.

Problems (5.5)–(5.7) and (5.8)–(5.11) form the main parts of the following procedure.

Step 0. Take a point $x^0 \in R$ and solve problem (5.5)–(5.7). If this problem is infeasible then the initial problem (5.2)–(5.4) is also infeasible. Otherwise we obtain an optimal solution x^1. Set $k = 1$, $f^1 = f(x^1)$.

Step k (k $\geqslant 1$). Solve the problem

$$\min \varphi(x, x^k), \tag{5.12}$$

$$g(x) \leqslant 0, \tag{5.13}$$

$$f(x) \leqslant f^k - \varepsilon, \tag{5.14}$$

$$x \in R. \tag{5.15}$$

If this problem is infeasible then stop: x^k is an ε-optimal solution. Otherwise, denote by x^{k+1} the optimal solution of (5.12)–(5.15). Set $k = k + 1$ and repeat the iteration.

Since the feasible set in problem (5.2)–(5.4) is compact and $f(x)$ is a c.m. (hence l.s.c.) function, the minimal value (if it exists) is finite.Therefore, the above described procedure is also finite, since at every iteration the record f^k is improved at least by the accuracy ε.

Note, that both problems (5.5)–(5.7) and (5.12)–(5.15) are special problems of the following type

$$\min p(x), \tag{5.16}$$

$$q(x) \leqslant 0, \tag{5.17}$$

$$x \in R = co\{z^1, \ldots, z^N\}, \quad N \geqslant n+1, \tag{5.18}$$

where $p(x)$ is a concave and $q(x)$ is a c.m. functions (for problem (5.12)–(5.15) we can take $q(x) = \max\{g(x), f(x) - f^k + \varepsilon\}$). In the above described procedure problem (5.16)–(5.18) plays a very essential role. If we can solve this problem, we have a finite minimization procedure for (5.2)–(5.4). Below, we describe a quite simple approach, which in general gives a lower bound for the minimal value of (5.16)–(5.18) and can be used as a lower bounding procedure for a branch and bound minimization method for (5.16)–(5.18).

Assume that $q(x)$ is concave. Due to the concavity of $p(x)$ at least one of the optimal solutions of (5.16)–(5.18) is reached at a vertex of the feasible set

$$G = \{x \in R : q(x) \leqslant 0\}. \tag{5.19}$$

Due to the concavity of $q(x)$ the vertex set $V(G)$ can be determined by an outer approximation procedure (see Horst and Tuy 1996), but in this case instead of a cutting plane we have the cutting surface $q(x) = 0$. Since the vertices $V(G)$ that belong to the cutting surface (i.e. new vertices) are obtained as intersections of edges of R with surface $q(x) = 0$ (see Horst et al. 1995) one can use, for example, the Thieu–Tam–Ban procedure or other procedures (see again Horst et al. 1995) to determine $V(G)$.

Unfortunately, this approach may be not valid for problems with more than one concave constraint. For instance, consider the problem

$$\min p(x), \tag{5.20}$$

$$q_i(x) \leqslant 0, \quad i = 1, \ldots, k, \quad k > 1 \tag{5.21}$$

$$x \in R = \text{co}\{z^1, \ldots, z^N\}, \quad N \geqslant n+1, \tag{5.22}$$

where $q_i(x)$ are concave functions. If we use one of the outer approximation methods mentioned above, then after k steps we obtain the set

$$S = \text{co}\{z^{k,1}, \ldots, z^{k,N_k}\}$$

where $z^{k,i}, i = 1, \ldots, N_k$ are final set of vertices and S may not coincide with the set set $K = \text{co}\{x \in R : q_i(x) \leqslant 0, i = 1, \ldots, k\}$. In general $S \supset K$. Hence, by the outer approximation approach one can get only a lower bound $\underline{p} = \min_{1\leqslant j\leqslant N_k} p(z^{k,j})$ for the optimal value of problem (5.20)–(5.22).

However, due to the concavity of all functions $p(x), q_i(x), i = 1, \ldots, k$ one can use the dual bound θ^* to improve $\underline{p}$ (this idea was also suggested earlier in Norkin 1992),

$$\theta^* = \sup_{\mu\geqslant 0} \theta(\mu),$$

$$\theta(\mu) = \min_{x\in R} L(x, \mu),$$

$$L(x, \mu) = p(x) + \sum_{i=1}^{k} \mu_i q_i(x).$$

Since $L(x, \mu)$ is concave in x and $R = \text{co}\{z^1, \ldots, z^N\}$, then

$$\theta(\mu) = \min_{x\in R} \left\{ p(x) + \sum_{i=1}^{k} \mu_i q_i(x) \right\} = \min_{1\leqslant j\leqslant N} \left\{ p(z^j) + \sum_{i=1}^{k} \mu_i q_i(z^j) \right\}.$$

Hence, the dual problem

$$\theta^* = \sup_{\mu\geqslant 0} \min_{1\leqslant j\leqslant N} \left\{ p(z^j) + \sum_{i=1}^{k} \mu_i q_i(z^j) \right\}. \tag{5.23}$$

is equivalent to the linear programming problem

$$\sup \mu_0,$$

$$p(z^j) + \sum_{i=1}^{k} \mu_i q_i(z^j) \geqslant \mu_0, \quad j = 1, \ldots, N,$$

$$\mu_i \geqslant 0, \quad i = 1, \ldots, k.$$

Now, let us come back to problem (5.16)–(5.18). Since $q(x)$ is a c.m. function, problem (5.16)–(5.18) can be seen as a problem with an infinite number of concave constraints in contrast to problem (5.20)–(5.22). In this case one can approximate problem (5.16)–(5.18) by problem (5.20)–(5.22) for some $q_i(x)$ and k and obtain a lower bound through (5.23). If the obtained lower bound is not satisfactory (in some sense) then add a new constraint $q_{k+1}(x) \leqslant 0$ to problem (5.20)–(5.22), set

$k = k + 1$ and solve again (5.23). These calculations can be repeated untill a stopping criterion is fulfilled.

Here, we suggest a generalization of the outer approximation approach with dual lower bounds (5.23). Let $\eta(x, y)$ be a concave minorant of $q(x)$. Then the suggested outer approximation procedure is as follows:

Step 0. Take $x^0 \in R$, define the initial vertex set $V^0 = \{z^1, \dots, z^N\}$. Determine

$$\overline{z}^0 = \underset{1 \leqslant j \leqslant N}{\operatorname{argmin}}\, p(z^j),$$

$$\theta_0^* = p(\overline{z}^0).$$

If $q(\overline{z}^0) \leqslant 0$ then stop: $\overline{z}^0$ is the optimal solution.

Step k ($k \geqslant 1$). Determine

$$\theta_k^* = \sup_{\mu \geqslant 0} \min_{z^j \in V^{k-1}} \left\{ p(z^j) + \sum_{i=1}^{k-1} \mu_i \eta(z^j, \overline{z}^i) \right\}, \tag{5.24}$$

and $\mu^k : \theta(\mu^k) = \theta_k^*$.

Determine

$$\overline{z}^k = \underset{z^j \in V^{k-1}}{\operatorname{argmin}} \{ p(z^j) + \sum_{i=1}^{k-1} \mu_i^k \eta(z^j, \overline{z}^i) \}.$$

If $q(\overline{z}^k) \leqslant 0$ then stop: $\overline{z}^k$ is the optimal solution. Otherwise, determine the new vertex set V^k by enumerating the vertices of the set S^k

$$S^k = R^{k-1} \bigcap \{x : \eta(x, \overline{z}^k) \leqslant 0\},$$

where $R^{k-1} = co\{V^{k-1}\}$. Set $k = k + 1$ and repeat the iteration.

If

$$\theta_{k+1}^* - \theta_k^* \geqslant \delta, \tag{5.25}$$

then we have a good improvement of θ_k^*. Therefore, this procedure can be used for the bounding operation in a branch and bound method: as soon as (5.25) is violated we obtain a lower bound θ_{k+1}^* for the optimal value of (5.16)–(5.18).

The advantage of the described approach consists in the following. We can estimate the minimal value of problem (5.2)–(5.4) from below even if no feasible point is known.

6. On the use of properties of $\varphi^+(x, y)$ and $\varphi^-(x, y)$

In this section we consider several examples with different assumptions on $f(x)$ which imply some usefull properties of the corresponding functions $\varphi^+(x, y)$ and $\varphi^-(x, y)$.

Assume that $f(x)$ is a continuously differentiable function, $f \in C^1$, which has a convex majorant function $\varphi^+(x, y)$ and a concave minorant function $\varphi^-(x, y)$:

$$\varphi^-(x, y) \leqslant f(x) \leqslant \varphi^+(x, y), \quad \forall x \in R, y \in R, \tag{6.1}$$

$$\varphi^-(y, y) = f(y) = \varphi^+(y, y), \quad \forall y \in R. \tag{6.2}$$

We start with function $\varphi^-(x, y)$.

PROPOSITION 6.1. *Let $f(x)$ be a differentiable c.m. function defened on a compact set R, $int(R) \neq \emptyset$.Then for every critical point $y \in int(R)$ of function $f(x)$ we have*

$$\max_{x \in R} \varphi^-(x, y) = \varphi^-(y, y) = f(y).$$

Proof. From (6.1) it follows that the function

$$F_y(x) = f(x) - \varphi^-(x, y) \tag{6.3}$$

achieves its global minimum at the support point y. Assume that $y \in \text{int}(R)$. Since $f(x)$ is differentiable and $\varphi^-(x, y)$ is concave in x, the necessary optimality condition has the form (see, for example, Mine and Fukushima 1981)

$$\nabla f(y) \in \partial\varphi^-(y, y), \tag{6.4}$$

where $\partial\varphi^-(y, y)$ denotes the superdifferential of $\varphi^-(x, y)$ at the point y. If y is a critical point of $f(x)$ then $\nabla f(y) = 0$. Hence, from (6.4) we have

$$0 \in \partial\varphi^-(y, y) \tag{6.5}$$

The latter inclusion means that $\varphi^-(x, y)$ attains its maximum at the point y. Therefore, the proposition is proved. □

From (6.2) it follows that

$$f(y) = \varphi^-(y, y) \leqslant \max_{x \in R} \varphi^-(x, y) = g^-(y). \tag{6.6}$$

Therefore,

$$f(y) \leqslant g^-(y), \quad \forall y \in R, \tag{6.7}$$

$$f(y) = g^-(y), \quad y \in \text{int}(R), \qquad \nabla f(y) = 0. \tag{6.8}$$

Consider now the problem

$$\min f(x), \tag{6.9}$$

$$x \in R. \tag{6.10}$$

Assume, that there exists at least one global minimum x^* of problem (6.9)–(6.10) such that $x^* \in \text{int}(R)$. Determine $g^-(y)$ by (6.6). Then from (6.8) it follows that problem (6.9)–(6.10) is equivalent to the following one:

$$\min g^-(y), \tag{6.11}$$

$$y \in R. \tag{6.12}$$

This substitution makes sense if problem (6.11)–(6.12) is easier than (6.9)–(6.10).

EXAMPLE 6.1. Assume $f(x) = \| x \|^2 - h(x)$, $h(x)$ is a convex finite function and it is known that $x^* \in \text{int}(R)$ for the corresponding problem (6.9)–(6.10). Then

$$\varphi^-(x, y) = \| y \|^2 + 2y^T(x - y) - h(x) = 2y^T x - \| y \|^2 - h(x)$$

is a concave minorant of $f(x)$ and

$$\begin{aligned} g^-(y) &= \max_{x \in R} \varphi^-(x, y) = - \| y \|^2 + \max_{x \in R}\{2y^T x - h(x)\} \\ &= - \| y \|^2 + q(y), \\ q(y) &= \max_{x \in R}\{2y^T x - h(x)\}. \end{aligned} \tag{6.13}$$

Since $q(y)$ is a convex function, $g^-(y)$ is again a d.c. function. However, in $f(x)$ we have the convex quadratic part and an arbitrary concave, whereas in $g^-(y)$ these parts (in some sense) are interchanged. Since the concave part in $g^-(y)$ is now quadratic, one can use this advantage to design an efficient minimization procedure for the corresponding problem (6.11)–(6.12). The price for a such an interchange is the convex auxiliary problem in (6.13).

Let us now show how the convex majorant function $\varphi^+(x, y)$ can also be used. Below we do not assume that the global minimum in problem (6.9)–(6.10) belongs to $\text{int}(R)$.

From (6.2) we again have

$$f(y) = \varphi^+(y, y) \geqslant \min_{x \in R} \varphi^+(x, y) = g^+(y). \tag{6.14}$$

If x^* is a global minimum of $f(x)$ over R, then

$$f^* = f(x^*) \leqslant \varphi^+(x, x^*), \quad \forall x \in R.$$

Hence,

$$f(x^*) \leqslant \min_{x\in R} \varphi^+(x, x^*) = g^+(x^*). \tag{6.15}$$

Therefore (6.14) and (6.15) imply that $f(x^*) = g^+(x^*)$ and problem (6.9)–(6.10) is equivalent to

$$\min g^+(y), \tag{6.16}$$

$$y \in R. \tag{6.17}$$

EXAMPLE 6.2. Let $f(x) = h(x) - \sum_{i=1}^{p} \lambda_i x_i^2, \lambda_i > 0, i = 1, \dots, p, p \leqslant n, h(x)$ is a convex finite function. Then

$$\varphi^+(x, y) = h(x) - 2\sum_{i=1}^{p} \lambda_i x_i y_i + \sum_{i=1}^{p} \lambda_i y_i^2$$

and

$$g^+(y) = \min_{x\in R} \varphi^+(x, y) = \sum_{i=1}^{p} \lambda_i y_i^2 + \min_{x\in R} \left\{ h(x) - 2\sum_{i=1}^{p} \lambda_i x_i y_i \right\}.$$

Hence, if $p < n$ then we obtain a reduction in dimension in the corresponding problem (6.16)–(6.17) in a rather simple way.

EXAMPLE 6.3. Let now $f(x) = \sum_{i=1}^{p} \lambda_i x_i^2 - h(x), \lambda_i > 0, i = 1, \dots, p, p \leqslant n, h(x)$ is a convex finite function. Then

$$\begin{aligned}
\varphi^-(x, y) &= -\sum_{i=1}^{p} \lambda_i y_i^2 + 2\sum_{i=1}^{p} \lambda_i x_i y_i - h(x) \\
&= -\sum_{i=1}^{p} \lambda_i y_i^2 + 2\sum_{i=1}^{p} \lambda_i x_i y_i - h(x) + \sum_{i=1}^{p} \lambda_i x_i^2 - \sum_{i=1}^{p} \lambda_i x_i^2 \\
&= \sum_{i=1}^{p} \lambda_i x_i^2 - h(x) + \sum_{i=1}^{p} \lambda_i (x_i - y_i)^2 \\
&= f(x) - \sum_{i=1}^{p} \lambda_i (x_i - y_i)^2.
\end{aligned}$$

Therefore,

$$f(x) - \varphi^-(x, y) = \sum_{i=1}^{p} \lambda_i (x_i - y_i)^2. \tag{6.18}$$

Determine

$$\overline{x} = argmin_{x \in R} \varphi^{-}(x, y). \tag{6.19}$$

Then

$$f^{*} - \varphi^{-}(\overline{x}, y) \leqslant f(\overline{x}) - \varphi^{-}(\overline{x}, y) = \sum_{i=1}^{p} \lambda_i (\overline{x}_i - y_i)^2. \tag{6.20}$$

Assume now that p is small, $p \ll n$. The error (6.20) of the auxiliary problem (6.19) depends only on the convex part. Therefore, it is reasonable to minimize the error by, for example, bisection with respect to the convex variables only. This is unusual, but it seems unavoidable for creating an efficient algorithm, especially when problem (6.19) is easy to solve (for example, if R is a simplex).

References

Avriel, M., Diewert, W., Schaible, S. and Zang, I. (1988), *Generalized Concavity*. Plenum Press, New York and London.

Baritompa, W. (1994), Accelerations for a variety of global optimization methods, *Journal of Global Optimization* 4: 37–45.

Baritompa, W. and Cutler, A. (1994), Accelerations for global optimization optimization covering methods using second derivatives, *Journal of Global Optimization* 4: 329–341.

Bazaraa, M. and Shetty, C. (1979), *Nonlinear Programming Theory and Algorithms*. John Wiley and Sons, New York.

Breiman, L. and Cutler, A. (1993), A deterministic algorithm for global optimization, *Mathematical Programming* 58: 179–199.

Bromberg, M. and Chang, T.-S. (1992), One dimensional global optimization using linear lower bounds, in C.A. Floudas and P.M. Pardalos (eds.), *Recent Advances in Global Optimization*, pp. 200–220. Kluwer Academic Publishers, Dordrecht/Boston/London.

Bulatov, V. (1987), Methods for solving multiextremal problems (global search), in B.A. Beljtiukov and V.P. Bulatov (eds.), *Methods of Numerical Analysis and Optimization*, pp. 133–187. Nauka, Novosibirsk.

Dolecki, S. and Kurcyusz, S. (1978), On Φ convexity in extremal problems, *SIAM J. Control and Optimization* 16(2): 277–300.

Ekeland, I. and Temam, R. (1976), *Convex Analysis and Variational Problems*. Elsevier, New York.

Evtushenko, Yu. (1971), Numerical methods for finding global extrema, *USSR Comput. Math. Math. Phys.* 11(6): 38–54.

Galperin, E. (1988), Precision, complexity and computational schemes of the cubic algorithm, *J.O.T.A.* 57(2): 223–238.

Hansen, P. and Jaumard, B. (1995), Lipschitz optimization, in R. Horst and P. Pardalso (eds), *Handbook of Global Optimization*, pp. 407–493. Kluwer Academic Publishers, Dordrecht/Boston/London.

Hirriart-Urruty, J.-B. (1985), Generalized differentiability, duality and optimization for problems dealing with differences of convex functions, *Lecture Notes in Economics and Mathematical Systems* 256: 37–70.

Horst, R. and Tuy, H. (1996), *Global Optimization (Deterministic Approaches)*, 3rd, revised and enlarged edition. Springer Verlag, Berlin.

Horst, R., Pardalos, P. and Thoai, N.V. (1995), *Introduction to Global Optimization*. Kluwer Academic Publishers, Dordrecht/Boston/London.

Konno, H., Thach, P.T. and Tuy, H. (1997), *Optimization on Low Rank Nonconvex Structures*. Kluwer Academic Publishers, Dordrecht/Boston/London.

Kutateladze, S. and Rubinov, A. (1976) *Minkowski Duality and its Applications*. Nauka, Novosibirsk (in Russian).

Meewella, C. and Mayne, D. (1988), An algorithm for global optimization of Lipschitz continuous functions, *J.O.T.A.* 57(2): 307–322.

Mine, H. and Fukushima, M. (1981), A minimization method for the sum of a convex and a continuously differentiable function, *J.O.T.A.* 33(1): 9–23.

Natanson, I. (1977) *Theorie der funktionen einer reelen Veränderlichen*. Verlag Harri Deutsch, Zurich.

Nefedov, V. (1987), The search for a global maximum of a function of several variables in a set specified by a constraints of the inequality type, *USSR Comput. Math. Math. Phys.* 27(1): 23–32.

Norkin, V. (1992), Pijavskii's method for solving the general global optimization problem, *USSR Comput. Math. Math. Phys.* 32(7): 873–886.

Papadimitriou, C. and Steiglitz, K. (1982), *Combinatorial Optimization. Algorithms and Complexity*. Prentice-Hall, Englewood Cliffs, N.J.

Pijavskii, S. (1967), An algorithm for searching for the global minimum of a function, *Theory of the Optimal Solutions*, Vol. 2, pp. 13–24. Institute of Cybernetics (in Russian).

Pintér, J.D. (1995), *Global Optimization in Action*. Kluwer Academic Publishers, Dordrecht/Boston/London.

Strongin, R. (1978), *Numerical Methods in Multiextremal Problems*. Nauka, Moscow (in Russian).

Sukharev, A. (1989), *Minimax Algorithms in the Problems of Numerical Analysis*. Nauka, Moscow (in Russian).

Thach, P. and Tuy, H. (1992), Relief indicator method – a new approach in global optimization, in V.P. Bulatov (ed.), *Optimization, Models, Methods, Solutions*, pp. 221–235. Nauka, Novosibirsk, (in Russian).

Thach, P. and Konno, H. (1993), A generalized Dantzig–Wolfe decomposition principle for a class of nonconvex programming problems, *Mathematical Programming, Series B* 62: 239–260.

Tuy, H. (1987), Convex programs with an additional reverse convex constraint, *J.O.T.A.* 52(3): 463–486.

Tuy, H. (1995), D.C. optimization: theory methods and algorithms, in R. Horst and P. Pardalos (eds), *Handbook of Global Optimization*, pp. 149–216. Kluwer Academic Publishers, Dordrecht/Boston/London.

Vial, J.-P. (1983), Strong and weak convexity of sets and functions, *Mathematics of Operations Research* 8(2): 231–259.

Wood, G. (1992), The bisection method in higher dimensions, *Mathematical Programming* 55: 319–337.

Zabotin, Ya. and Khabibulin, R. (1975), Nonlinear support functionals and optimization problems, *Investigations in Applied Mathematics*, Vol. 3, pp. 60–67. University of Kazan (in Russian).