

ON CONVERGENCE OF A GLOBAL SEARCH STRATEGY FOR REVERSE CONVEX PROBLEMS

ALEXANDER STREKALOVSKY

Received 19 January 2003 and in revised form 10 February 2004

On the base of global optimality conditions for RCP we develop the global search strategy (GSS) and focus on the global convergence of GSS giving a variant of the proof.

1. Introduction

Nowadays specialists on optimization observe the persistent demands from the world of applications to create an effective apparatus for finding just a global solution to nonconvex problems in which there may exist local solutions located very far from a global one even up to the values of goal function.

As well-known, the conspicuous limitation of convex optimization methods applied to nonconvex problems is their ability of being trapped at a local extremum or even a critical point depending on a starting point. In other words, the classical apparatus shows itself inoperative for new problems arising from practice.

That is why, the development of nonconvex optimization took the way of tools initially generated in discrete optimization (DO) and for discrete optimization. So, DO gave some apparatus for continuous optimization (CO). Gradually, Branches & Bounds, cuts, and so forth ideas became very popular in nonconvex area of CO, although in some cases its turned out to be too much sensitive, for instance, with respect to the changing the size of a problem.

In [8, 9, 10, 11, 12] it was proposed other approach to d.c. programming problems based on global optimality conditions (GOC) that has proved its effectiveness for numerous continuous (even dynamical) and discrete optimization problems [11, 13, 14, 15].

Nevertheless, the theoretical substantiation of the approach stays in some cases uncompleted. This concerns, in particular, convergence proofs for global search strategies based on GOC.

This paper aims especially to fill the lacuna for RCP. On the other hand, we demonstrate below the importance of new notion in optimization—the notion of resolving set, or a “good” approximation of the level surface of a convex function $g(\cdot)$.

The crucial influence of the approximation on results of global search was transparently shown in [11, 12, 14, 15] by computational experiments. Here we would like

underline the value of new notion or the theoretical side of investigations and its resolving impact on the convergence proof of GSS.

The paper is organized as follows. First, we remember the GOC for the simplest case of reverse convex problems ($x \in R^n$):

$$f(x) \downarrow \min, \quad x \in S, g(x) \geq 0, \quad (P)$$

and give the necessary and sufficient conditions for a sequence $\{z^k\}$ of feasible points to be minimizing to (P). After this, we discuss some features of global search strategy (GSS) for (P), and finally we study the convergence of GSS independent of starting point choice.

2. Global optimality conditions and minimizing sequences

In the sequel we assume that the goal function of problem (P) is bounded from below on the feasible set, that is,

$$f_* := \inf_x \{f(x) \mid x \in S, g(x) \geq 0\} > -\infty. \quad (2.1)$$

In addition, suppose, the function $g(\cdot)$ is differentiable over some convex open set Ω containing S .

Besides, the following assumption seems to be natural for reverse convex problem (P) [1, 5, 6, 7, 16, 17]:

(G) there does not exist any global minimum point x_* , such that $g(x_*) > 0$.

In other words,

$$\text{Sol}(P) \cap \{x \mid g(x) > 0\} = \emptyset. \quad (2.2)$$

where $\text{Sol}(P) := \text{Argmin}(P)$.

Actually, if condition (G) is broken down, problem (P) ceases to be reverse convex and might be solved by a suitable convex optimization method, say, in the case of convexity of $f(\cdot)$ and S .

THEOREM 2.1 [9, 10]. *Let assumption (G) be fulfilled and a point z be a solution to (P) ($z \in \text{Sol}(P)$). Then the following condition holds*

(\mathcal{E}_0)

$$\forall y : g(y) = 0, \forall x \in S : f(x) \leq f(z), \quad \langle \nabla g(y), x - y \rangle \leq 0. \quad (2.3)$$

At first sight, condition (\mathcal{E}_0) has no relation to the classical extremum theory. But if one sets $y = z, S = R^n$, from (\mathcal{E}_0) it follows that z is a solution to the problem:

$$\langle \nabla g(z), x \rangle \uparrow \max, \quad f(x) \leq f(z), \quad (2.4)$$

and that is why the Lagrange rule takes place:

$$\lambda \nabla g(z) + \lambda_0 \nabla f(z) = 0, \quad (2.5)$$

with $\lambda_0 \geq 0$, $\lambda \leq 0$, $\lambda_0 + |\lambda| > 0$. Clearly, (2.5) is the well-known result from nonlinear programming [2, 3, 4, 18] for problem (P) with $S = R^n$. Hence (2.3) is connected with optimization Theory.

On the other hand, optimality condition (\mathcal{E}_0) possesses so-called algorithmic property (AP) as all classical optimality conditions. It means if optimality condition is broken down at a point z under study then there exists a procedure allowing to construct a feasible point which is better than the point z of interest. Indeed, if there is a pair (v, u) , such that $g(v) = 0$, $u \in S$, $f(u) \leq f(z)$ and $\langle \nabla g(v), u - v \rangle > 0$, then due to convexity of $g(\cdot)$ it results $g(u) > 0 = g(v)$. Thus, one has the feasible point $u \in S$, $g(u) > 0$, $f(u) \leq f(z)$. Therefore, in virtue of assumption (G) there is a possibility to decrease the value of objective function $f(u)$ by descending on the constraint $g(x) = 0$, that is, to construct a better feasible point $\hat{z} \in S$, such that $g(\hat{z}) = 0$, $f(\hat{z}) < f(u) \leq f(z)$. It can be carried out by one of classical optimization method, since the constraint $g \geq 0$ is not active at the point $u \in S$, $f(u) \leq f(z)$.

THEOREM 2.2 [9, 10]. *Suppose, in problem (P) the set S is convex and the following assumptions take place:*

(\mathfrak{K})

$$-\infty \leq \inf(g, R^n) < g(z) = 0; \quad (2.6)$$

(H)

$$\forall y \in S, g(y) = 0, \exists h = h(y) \in S, \quad \langle \nabla g(y), h - y \rangle > 0. \quad (2.7)$$

Then condition (\mathcal{E}_0) where $y \in S$ becomes sufficient for z being a global solution to (P).

Remark 2.3. Assumption (2.6) seems to be natural since it means, that there exist points in R^n , which are inadmissible with respect to the constraint $g(x) \geq 0$. Otherwise the constraint would be senseless.

Remark 2.4. Assumption (2.7) means, that if we throw away the constraint $f(x) \leq f(z)$, it becomes immediately possible to violate the basic inequality $\langle \nabla g(y), x - y \rangle \leq 0$ into condition (\mathcal{E}_0). So, assumption (H) turns out to be similar to the regularity (or normality) conditions in Lagrange multipliers rule, when, say, Slater conditions guarantee nontriviality ($\lambda_0 \neq 0$) of the multiplier corresponding to the objective function.

If $\lambda_0 = 0$, the multiplier rule becomes inconsistent, senseless, because it expresses only some property of constraints, for instance, the linear dependence of constraints gradients. But the goal function is not involved into optimality condition. It is clear, it should be there in, that is, the objective function must be in any optimality conditions in some influential form. In other words, it would be senseless, if optimality condition did not take into account the goal function. Fortunately, in our case it is not so.

Example 2.5. Consider the problem ($x \in R^2$):

$$\begin{aligned} f(x) &:= x_1^2 + (x_2 - 1)^2 \downarrow \min, & h_1(x) &:= x_1^2 + 4x_2^2 - 4 \leq 0, \\ h_2(x) &:= 2x_2 - x_1 \leq 0, & g(x) &:= 2x_1^2 - x_2 - 2 \geq 0. \end{aligned} \quad (2.8)$$

It can be shown, that the point $z = -\alpha(1, 1/2)$, $\alpha = 1/8(1 + \sqrt{65})$ verifies the KKT-theorem:

$$\begin{aligned} \nabla f(z) + \sum_{i=1}^2 \lambda_i \nabla h_i(z) - \mu \nabla g(z) &= 0, \\ \mu g(z) &= 0, \quad \lambda_i h_i(z) = 0, \quad i = 1, 2. \end{aligned} \tag{2.9}$$

However, for $y = (1, 0)$, $g(y) = 0 = g(z)$ and $u = (9/8, 0)$, $h_i(u) < 0$, $f(u) < f(z)$ we get $\langle \nabla g(y), u - y \rangle = 1/2 > 0$.

So, z is not global minimum.

Definition 2.6. (i) A sequence $\{z^k\} \subset R^n$ is called minimizing for problem (P), if three following conditions are fulfilled (see (2.1))

$$\lim_{k \rightarrow \infty} f(z^k) = f_* := \inf_x \{f(x) \mid x \in S, g(x) \geq 0\}; \tag{2.10}$$

$$z^k \in S, \quad k = 0, 1, 2, \dots; \tag{2.11}$$

$$\liminf_{k \rightarrow \infty} g(z^k) \geq 0. \tag{2.12}$$

(ii) The set of all minimizing sequences for (P) will be denoted by $\mathcal{M} = \mathcal{M}(P)$. Further assume, that

(G1) there no exists $\{z^k\} \in \mathcal{M}$, such that $\liminf_{k \rightarrow \infty} g(z^k) > 0$.

LEMMA 2.7. *Assumption (G1) is equivalent to the following condition:*

(G2)

$$\lim_{k \rightarrow \infty} g(z^k) = 0 \quad \forall \{z^k\} \in \mathcal{M}. \tag{2.13}$$

Proof. Clearly, (G2) implies (G1). Now, if (G1) holds, then

$$\liminf_{k \rightarrow \infty} g(z^k) \leq 0 \quad \forall \{z^k\} \in \mathcal{M}. \tag{2.14}$$

Thus, due to (2.12) one gets

$$\liminf_{k \rightarrow \infty} g(z^k) = 0 \quad \forall \{z^k\} \in \mathcal{M}. \tag{2.15}$$

Suppose $\limsup_{k \rightarrow \infty} g(z^k) > 0$. It means, there exists a subsequence $\{z^m\} \in \{z^k\}$, such that $\lim_{m \rightarrow \infty} g(z^m) > 0$. On the other hand, it can be readily seen, that $\{z^m\}$ verifies (2.10) and (2.11), as well as $\{z^k\}$. Therefore, $\{z^m\} \in \mathcal{M}$ what contradicts to (G1). \square

The following result is related to the duality theorem of Tuy [5, 16, 17] to which we will return later (see Section 4).

LEMMA 2.8. *Let assumption (G1) takes place. Then for every $\{z^k\} \in \mathcal{M}$ there exists a numerical sequence $\{\zeta_k\} \subset R$, such that $\zeta_k > 0$, $k = 1, 2, \dots$, $\zeta_k \downarrow 0(k \rightarrow \infty)$, and*

$$g(x) \leq g(z^k) + \zeta_k, \quad \forall x \in S: f(x) \leq f(z^k). \tag{2.16}$$

Proof. Suppose, for some sequence $\{z^k\} \in \mathcal{M}$ there no exists any sequence $\{\zeta_k\}$ verifying the conditions above. It means, that for every $k = 1, 2, \dots$ there exists $x^k \in R^n$, such that

$$x^k \in S, \quad f(x^k) \leq f(z^k), \quad (2.17)$$

$$g(x^k) \geq g(z^k) + \gamma_k, \quad (2.18)$$

where $\gamma_k > 0, k = 1, 2, \dots$

$$\lim \gamma_k = \gamma > 0. \quad (2.19)$$

Then, due to Lemma 2.7 it follows from (2.17) and (2.18)

$$\liminf_{k \rightarrow \infty} g(x^k) \geq \lim_{k \rightarrow \infty} g(z^k) + \lim \gamma_k = \gamma > 0. \quad (2.20)$$

It results from (2.17) that $\{x^k\}$ is also a minimizing sequence, since $\{z^k\} \in \mathcal{M}$. On the other hand, the latter inequalities contradict condition (G1). \square

Further, consider the function

$$\varphi(z) = \sup_{(x,y)} \{ \langle \nabla g(y), x - y \rangle \mid x \in S, f(x) \leq f(z), g(y) = g(z) \}. \quad (2.21)$$

Since $0 = \langle \nabla g(z), z - z \rangle \leq \varphi(z)$, one has

$$\varphi(z) \geq 0 \quad \forall z \in S. \quad (2.22)$$

THEOREM 2.9. *Let assumption (G1) takes place and $\{z^k\} \in \mathcal{M}$. Then the following condition holds*

(\mathcal{E})

$$\lim_{k \rightarrow \infty} \varphi(z^k) = 0. \quad (2.23)$$

Proof. According to Lemma 2.8 there exists a numerical sequence $\{\zeta_k\}, \zeta_k > 0, k = 0, 1, 2, \dots, \zeta_k \downarrow 0$, verifying (2.16). Therefore, for all $k = 0, 1, 2, \dots$, for all $y : g(y) = g(z^k)$, for all $x \in S, f(x) \leq f(z^k)$, in virtue of the convexity of $g(\cdot)$ one gets

$$\zeta_k \geq g(x) - g(z^k) = g(x) - g(y) \geq \langle \nabla g(y), x - y \rangle. \quad (2.24)$$

Whence, due to (2.21) and (2.22) it follows

$$0 \leq \varphi(z^k) \leq \zeta_k, \quad (2.25)$$

what proves (\mathcal{E}). \square

Now consider the following condition:

($\mathfrak{R}1$)

$$\|\nabla g(y)\| \geq \chi > 0 \quad \forall y : g(y) = g(z^k); \quad (2.26)$$

THEOREM 2.10. *Let the goal function $f(\cdot)$ be upper semicontinuous over a convex set S and assumption (H) of Theorem 2.2 takes place.*

Let in addition $(\mathfrak{R}1)$ and the condition (\mathfrak{E}) be fulfilled together with the condition:

$$\lim_{k \rightarrow \infty} g(z^k) \leq 0. \quad (2.27)$$

Then equality (2.10) holds. If (2.27) is verified as equality, then the sequence $\{z^k\}$ turns out to be minimizing to problem (P) : $\{z^k\} \in \mathcal{M}$.

Proof. (1) Suppose, (2.10) is broken down. Then for some subsequence $\{z^m\}$ of the sequence $\{z^k\}$ one has

$$\lim_{m \rightarrow \infty} f(z^m) > f_* + \gamma + \varepsilon, \quad (2.28)$$

where $\varepsilon, \gamma > 0$ and ε is rather small.

There then exists $w \in S, g(w) \geq 0$, such that

$$\lim_{m \rightarrow \infty} f(z^m) > f(w) + \gamma. \quad (2.29)$$

(2) Let show, that in this case one can find u , such that

$$u \in S, \quad g(u) > 0, \quad f(u) < \lim_{m \rightarrow \infty} f(z^m) - \gamma. \quad (2.30)$$

If $g(w) > 0$, it is done. When $g(w) = 0$, due to upper semicontinuity of $f(\cdot)$ there exists a neighborhood W of the point w , such that

$$f(x) < \lim_{m \rightarrow \infty} f(z^m) - \gamma \quad \forall x \in W. \quad (2.31)$$

On the other hand, in virtue of (H) one can find $h \in S$, such that $\langle \nabla g(w), h - w \rangle > 0$. Then for $x(\alpha) = \alpha h + (1 - \alpha)w \in S, \alpha \in]0, 1[$, because of the convexity of $g(\cdot)$ one gets

$$g(x(\alpha)) - g(w) \geq \langle \nabla g(w), x(\alpha) - w \rangle = \alpha \langle \nabla g(w), h - w \rangle > 0. \quad (2.32)$$

So, $g(x(\alpha)) > 0, x(\alpha) \in S, \alpha \in]0, 1[$.

Besides, for a rather small α we have $x(\alpha) \in W$, and due to (2.31)

$$f(x(\alpha)) < \lim_{m \rightarrow \infty} f(z^m) - \gamma. \quad (2.33)$$

Thus, (2.30) is proven with $u = x(\alpha)$ and α rather small.

(3) Now, due to (2.27) for some number \mathfrak{K} one has

$$g(z^m) < g(u) \quad \forall m \geq \mathfrak{K}. \quad (2.34)$$

Hence, there exists $y^m \in R^n$, such that

$$g(y^m) = g(z^m) =: g_m, \quad 1/2 \|y^m - u\|^2 = \inf_x \{1/2 \|x - u\|^2 : g(x) \leq g_m\}. \quad (2.35)$$

Besides,

$$\|y^m - u\| > 0, \quad (2.36)$$

since u do not belong to $\{x \mid g(x) \leq g_m\}$.

So, the KKT-theorem takes place, that is,

$$\exists \lambda_0 \geq 0, \exists \lambda \geq 0 : \lambda_0 + \lambda > 0, \quad \lambda_0(y^m - u) + \lambda \nabla g(y^m) = 0, \quad \lambda[g(y^m) - g_m] = 0. \quad (2.37)$$

If $\lambda_0 = 0$, then $\lambda > 0$ and $\nabla g(y^m) = 0$, what contradicts (R1). If $\lambda = 0$, then $\lambda_0 > 0$ and $y^m = u$, what is impossible due to (2.36).

So, $\lambda > 0$ and $g(y^m) = g_m := g(z^m)$. In this case it follows from (2.37)

$$\lambda_m \nabla g(y^m) = u - y^m, \quad \lambda_m = \lambda/\lambda_0 > 0. \quad (2.38)$$

Taking into account (R1) we get

$$\lambda_m^{-1} \|u - y^m\| = \|\nabla g(y^m)\| \geq \chi > 0, \quad (2.39)$$

whence it immediately follows

$$\langle \nabla g(y^m), u - y^m \rangle = \lambda_m^{-1} \|u - y^m\|^2 \geq \chi \|u - y^m\|. \quad (2.40)$$

On the other hand,

$$\langle \nabla g(y^m), u - y^m \rangle \leq \sup_x \{ \langle \nabla g(y^m), x - y^m \rangle \mid x \in S, f(x) \leq f(z^m) \} \leq \varphi(z^m). \quad (2.41)$$

Unifying the latter inequalities with (2.40) one finally has

$$0 < \chi \|y^m - u\| \leq \varphi(z^m). \quad (2.42)$$

Whence, due to (E) it follows $\lim y^m = u$. Then because of continuity of $g(\cdot)$ and (2.27), we obtain

$$0 \geq \lim g(z^m) = \lim g(y^m) = g(u) > 0, \quad (2.43)$$

what is impossible. \square

3. Global search strategy

In this section, we briefly repeat the basic positions of conceptual global search algorithm advanced in [15]. Theorems 2.1–2.10 suggest to consider the following problem:

$$\Psi(x, y) \triangleq \langle \nabla g(y), x - y \rangle \uparrow \max_{(x, y)}, \quad x \in S, g(x) \geq 0, g(y) = 0, \quad (3.1)$$

which is rather tight. That is why we decompose it into two consecutive problems:

$$\langle \nabla g(y), x \rangle \uparrow \max_x, \quad x \in S, \quad f(x) \leq f(z), \quad (3.2)$$

$$h_u(v) := \langle \nabla g(v), u - v \rangle \uparrow \max_v, \quad g(v) = 0, \quad (3.3)$$

where u is an approximate solution to (3.2). Supposing these problems solvable, in [15] we advanced the following.

Global search strategy. Let given sequences $\{\tau_k\}$, $\{\delta_k\}$, such that $\tau_k \downarrow 0$, $\delta_k \downarrow 0$, $\tau_k > 0$, $\delta_k > 0$, $k = 0, 1, 2, \dots$ and an initial point $x_0 \in S$, $g(x_0) \geq 0$.

Step 0. Set $k := 0$, $x^k := x_0$.

Step 1. Starting from the point $x^k \in S$, $g(x^k) \geq 0$, by means of one of local search method (LSM) obtain a τ_k -stationary point

$$z^k \in S, \quad g(z^k) = 0, \quad \zeta_k := f(z^k). \quad (3.4)$$

Step 2. Construct an approximation

$$\mathcal{A}_k := \{v^1, v^2, \dots, v^{N_k} \mid g(v^i) = 0, i = 1, \dots, N_k\} \quad (3.5)$$

of the level surface $g(y) = 0$.

Step 3. For all $i = 1, \dots, N_k$ find a δ_k -solution $u^i \in S$, $f(u^i) \leq \zeta_k$, of linearized problem (3.2):

$$\langle \nabla g(v^i), u^i \rangle + \delta_k \geq \sup_x \{ \langle \nabla g(v^i), x \rangle \mid x \in S, f(x) \leq \zeta_k \}. \quad (3.6)$$

Step 4. For all $i = 1, \dots, N_k$ find δ_k -solution w^i , $g(w^i) = 0$, of the Level problem (3.3) corresponding to u^i :

$$\langle \nabla g(w^i), u^i - w^i \rangle + \delta_k \geq \sup_v \{ \langle \nabla g(v), u^i - v \rangle \mid g(v) = 0 \}. \quad (3.7)$$

Step 5. Set

$$\eta_k := \langle \nabla g(w^j), u^j - w^j \rangle = \max_{1 \leq i \leq N_k} \langle \nabla g(w_i), u^i - w^i \rangle. \quad (3.8)$$

Step 6. If $\eta_k > 0$, then set $x^{k+1} := w^j$, $k := k + 1$ and loop to Step 1.

Step 7. If $\eta_k \leq 0$, then set $x^{k+1} := z^k$, $k := k + 1$ and loop to Step 1.

Remark 3.1. In order the description of GSS becomes more substantiated we assume, that

(HL) for all $\delta > 0$, for all $z \in S$, $g(z) \geq 0$ for all $v : g(v) = 0$, one can find a point $u \in S$, $f(u) \leq f(z)$, such that

$$\langle \nabla g(v), u \rangle + \delta \geq \sup_x \{ \langle \nabla g(v), x \rangle \mid x \in S, f(x) \leq f(z) \}; \quad (3.9)$$

(HU) for all $\delta > 0$, for all $u \in S$ one can find a point $w : g(w) = 0$, such that $\langle \nabla g(w), u - w \rangle + \delta \geq \sup_v \{ \langle \nabla g(v), u - v \rangle \mid g(v) = 0 \}$.

Remark 3.2. It can be easily seen that the sequence $\{z_k\}$ generated by GS strategy is a sequence of τ_k -stationary (critical) points due to describing of Step 1.

Remark 3.3. When $\eta_k > 0$ (Step 6), due to convexity of $g(\cdot)$, one has

$$0 < \langle \nabla g(w^j), u^j - w^j \rangle \leq g(u^j) - g(w^j) = g(u^j), \quad (3.10)$$

whence $g(x^{k+1}) \geq \eta_k > 0 = g(z^k) = g(w^i)$. On the other hand, from the description of Step 3 it follows

$$f(x^{k+1}) \leq \zeta_k := f(z^k), \quad x^{k+1} \in S, \quad g(x^{k+1}) > 0. \quad (3.11)$$

Thus, one constructed a feasible point $x^{k+1} \in S$, that is not worse than z^k , (since $f(x^{k+1}) \leq \zeta_k := f(z^k)$) and at which the constraint $g \geq 0$ is not active, $g(x^{k+1}) > 0$. In this case, starting new local search at x^{k+1} under the assumption of type (G), (G1), or, when constraint $g \geq 0$ is essential [1, 7, 16, 17], we will get z^{k+1} , $g(z^{k+1}) = 0$, with the property $f(z^{k+1}) < f(x^{k+1})$ so that

$$f(z^{k+1}) < f(x^{k+1}) \leq f(z^k). \quad (3.12)$$

This observation will have an important impact on the convergence of GS Strategy.

Thus, the strategy above becomes relaxing, that is, decreasing the value of $f(\cdot)$ at every iteration, when $\eta_k > 0$. Note, that we suppose to use some minimization method (say, Newton's method or an interior point method and so on) in order "to descent" on the surface $g = 0$ with the obligatory strict improving of the goal function and without taking into account the constraint $g \geq 0$ (so, a free (from the constraint $g = 0$) descent). Therefore, we call this LSM the free descent procedure. More strictly it can be reformulated as follows:

(FD) there exist $\mu > 0$, such that for all $x \in S$, $g(x) > 0$, for all $\tau > 0$ one can find a τ -critical point $\hat{z} = \hat{z}(x)$, such that

$$\hat{z} \in S, \quad g(\hat{z}) = 0, \quad f(\hat{z}) \leq f(x) - \mu g(x) + \tau. \quad (3.13)$$

It is easy to see, that (3.13) is equivalent to

$$g(x) \leq 1/\mu [f(x) - f(\hat{z})] + \tau/\mu. \quad (3.14)$$

Remark 3.4. Returning to the describing of the global search strategy, note that it can not be viewed as an algorithm, because on Step 1, it is not precised what kind of LS algorithm you have to use, as well as, on Step 3 for solving the linearized problem you are free to choose any suitable method, which must be, however very fast, since one has to repeat it several times at every iteration.

Nevertheless, GS strategy allows us not to lose from the view all basic moments of global search, while if we begin to precise some point, for instance, the linearized problem solving, we are risking to be lost in some particularities.

4. Convergence of global search strategy

As it was pointed out above, the verification of the fact, whether a feasible point z is a global solution or not, can be reduced to the solving problem (3.1), which in turn can be partially performed by global search strategy.

Clearly, the choosing methods for solving problems (3.2) and (3.3), as well as the local search method on Step 1 of GSS, must be skilled, but, nevertheless, is standard, in a sense, and "already seen" [2, 3, 4, 5, 6, 7, 18].

At the same time, the constructing of a “good” approximation \mathcal{A}_k of the level surface $g(x) = 0$ on Step 2 turns out to be of paramount importance from the view point of real global search, as the numerous computational experiments show [8, 11, 13, 14, 15].

If you are able to construct a “good” approximation on Step 2 of GSS, you can escape from any stationary point and finally to say that an obtained point z is an approximate global solution [11, 13, 14, 15].

Let us look at the situation from the theoretical point of view and show the resolving impact of a “good” approximation on the convergence of GS strategy.

In order to do it, set $\zeta := f(z)$ and consider an approximation

$$\mathfrak{R}(\zeta) := \mathcal{A} = \{v^1, \dots, v^N \mid g(v^i) = 0, i = 1, \dots, N = N(\zeta)\}. \quad (4.1)$$

Suppose, some points $u^i \in S$, $f(u^i) \leq \zeta$, $w^i \in R^n$, $g(w^i) = 0$ verify the inequalities (3.6) and (3.7) according to assumption (HL) and (HU), respectively, with $\delta_k := \delta$.

Besides, set, as above

$$\eta(\zeta) := \langle \nabla g(w^j), u^j - w^j \rangle = \max_{1 \leq i \leq N} \langle \nabla g(w_i), u^i - w^i \rangle. \quad (4.2)$$

Definition 4.1. Let given constants $\Delta, \delta, \varepsilon > 0$, $0 < \Theta < 1$.

(i) The approximation $\mathfrak{R}(\zeta)$ is called strictly $(\Delta, \delta, \varepsilon, \Theta)$ -resolving set, if from the fact that z is not a Δ -solution to problem (P), that is,

$$f(z) - \Delta > f_* \triangleq \inf_x \{f(x) \mid x \in S, g(x) \geq 0\}, \quad (4.3)$$

it follows two inequalities:

$$\eta(\zeta) > 0, \quad (4.4)$$

$$\eta(\zeta) > \Theta \varphi(z) - \varepsilon. \quad (4.5)$$

(ii) If inequalities (4.4) and (4.5) are fulfilled as nonstrict, then the approximation $\mathfrak{R}(\zeta)$ is called simply (or merely) resolving.

Let us consider together with problem (P) the dual (according Tuy [5, 7, 16, 17]) problem of convex maximization:

$$g(x) \uparrow \max, \quad x \in S, f(x) \leq \beta. \quad (Q_\beta)$$

Recall that due to Tuy in [5, Proposition 10, page 166] the point $z \in S$ is a solution to (P) if and only if

$$V(f(z)) := \sup_x \{g(x) \mid x \in S, f(x) \leq f(z) = f_*\} = 0, \quad (4.6)$$

where $V(\beta)$ is the optimal value function of problem (Q_β)

$$V(\beta) := \sup_x \{g(x) \mid x \in S, f(x) \leq \beta\}. \quad (4.7)$$

Assume, that the optimal value function $V(\beta)$ is upper Lipschitzian at f_* , that is, there exist two constants $\bar{\beta}$ and $M > 0$, such that for every $\beta: f_* \leq \beta < \bar{\beta}$ the following estimation holds

$$V(\beta) - V(f_*) \leq M(\beta - f_*). \quad (4.8)$$

Whence, in virtue of (4.6) it follows

$$\sup_x \{g(x) \mid x \in S, f(x) \leq f(z)\} \leq M(f(z) - f_*). \quad (4.9)$$

Note, that estimation (4.8), as well known [2, 3, 4], is rather natural for problem (Q_β) , since for a Lipschitzian function (in (Q_β) it is convex, hence, Lipschitzian), a bounded set S and a continuous function $f(\cdot)$ the stability property in a neighborhood of zero takes place for the optimal value function $V(\cdot)$. See, for instance, [2], Corollaries 1 and 2 from Theorem 6.3.2, as well as Sections 6.4 and 6.5.

Remember also, that any normality condition for problem (Q_β) , for instance, Slater or Mangasarian-Fromovitz regularity conditions, guarantees Lipschitzian property for the optimal value function, or the stability of the problem (Q_β) with respect to the simplest perturbation of inequality constraint [2].

Therefore, one can say, that assumption (4.8) by no means bounds the generality of consideration of problems (P) and (Q_β) .

In the sequel the following result will be rather useful.

LEMMA 4.2. *Let an approximation $\mathfrak{R}(\zeta)$ be strictly (merely) $(\Delta, \delta, \varepsilon, \Theta)$ -resolving. In addition, estimation (4.9) takes place with a constant $M > 0$ and $\varepsilon \geq M\Theta\Delta$. Then, from the (nonstrict) inequality (4.4) it follows the (nonstrict) inequality (4.5).*

Proof. First, from the definition of the function $\varphi(\cdot)$ and due to convexity of $g(\cdot)$ it follows

$$\begin{aligned} \varphi(z) &:= \sup_{x,y} \{ \langle \nabla g(y), x - y \rangle \mid x \in S, f(x) \leq f(z), g(y) = 0 \} \\ &\leq \sup_x \{g(x) \mid x \in S, f(x) \leq f(z)\}. \end{aligned} \quad (4.10)$$

Whence with the help of (4.9) one has

$$\varphi(z) \leq M(f(z) - f_*). \quad (4.11)$$

If now the inequality (4.5) is broken down, then

$$\eta(\zeta) + \varepsilon \leq \Theta M(f(z) - f_*). \quad (4.12)$$

On the other hand, in this case according to definition of strictly resolving set we have

$$\zeta := f(z) \leq f_* + \Delta. \quad (4.13)$$

Then, from (4.12) it follows

$$\eta(\zeta) \leq M\Theta\Delta - \varepsilon \leq 0, \quad (4.14)$$

what contradicts (4.4).

Clearly, in the case of merely resolving set and the nonstrict inequalities the proof is similar. \square

Notes. First, Lemma 4.2 gives the concordance condition for parameters of computation Δ , ε and Θ : $\varepsilon \geq M\Theta\Delta$, where Δ is the solving accuracy of problem (P), ε stands for exactness of the basic inequality (4.5), Θ is a share of solving the auxiliary problem (3.1).

Second, according to this lemma, when applying global search strategy one can look only for the number $\eta_k = \eta(\zeta_k)$, $\zeta_k := f(z^k)$, neglecting (4.5). This completely corresponds to the describing of GSS, in which one does not use the inequality (4.5). The latter will be without fail satisfied, if $\eta_k > 0$ ($\eta_k \geq 0$), when using strictly (merely) resolving set $\mathfrak{R}(\zeta_k)$ at every iteration of GSS on Step 2.

Therefore, for applying the GS strategy (or conceptual algorithm) for solving problem (P) the following assumption seems to be natural.

(HR) For all $(\Delta, \tau, \delta, \varepsilon, \Theta), \Delta, \tau, \delta, \varepsilon > 0$, $0 < \Theta < 1$, $\varepsilon \geq M\Theta\Delta$, for every τ -stationary point $z \in S$, $g(z) = 0$, $\zeta := f(z)$, which is not a Δ -solution to problem (P), one can construct strictly (merely) $(\Delta, \delta, \varepsilon, \Theta)$ -resolving approximation $\mathfrak{R}(\zeta)$.

Further, assume, that at every iteration of GSS on Step 2 one constructs strictly (merely) $(\Delta_k, \delta_k, \varepsilon_k, \Theta_k)$ -resolving set $\mathfrak{R}_k = \mathfrak{R}(\zeta_k)$.

Note, when using merely resolving sets we have to change in the describing GSS the strict inequality $\eta_k > 0$ (on Step 6) for nonstrict, and the inequality $\eta_k \leq 0$ (on Step 7) for the strict one correspondingly.

Since the notion of resolving approximation plays the crucial role in the proof of convergence of GS strategy, at each iteration of which on Step 2 one constructs a resolving set, we will call such conceptual algorithm shorter \mathfrak{R} -strategy.

Further, let us consider the following assumptions:

($\mathfrak{R}3$)

$$\|\nabla g(y)\| \geq \rho > 0 \quad \forall y : g(y) = 0; \quad (4.15)$$

(H1)

$$\forall y \in S, g(y) = 0, \quad \exists h = h(y) \in S, \quad \langle \nabla g(y), h - y \rangle > 0. \quad (4.16)$$

Besides, let us suppose that numerical sequences $\{\Delta_k\}$, $\{\tau_k\}$, $\{\delta_k\}$, $\{\varepsilon_k\}$, and $\{\Theta_k\}$ verify the following conditions:

$$\begin{aligned} \Delta_k, \tau_k, \delta_k, \varepsilon_k > 0, \quad 0 < \Theta_k < 1, \quad \varepsilon_k \geq \Theta_k M \Delta_k, \quad k = 0, 1, 2, \dots; \\ \tau_k \downarrow 0, \quad \delta_k \downarrow 0, \quad \varepsilon_k \downarrow 0 \quad (k \rightarrow \infty). \end{aligned} \quad (4.17)$$

THEOREM 4.3. *Let in problem (P) the goal function be continuous over an open domain Ω , containing a convex set S and $\{x \mid g(x) = 0\}$. Besides*

$$f_* := \inf_x \{f(x) \mid x \in S, g(x) \geq 0\} > -\infty, \quad (4.18)$$

where a convex function $g(\cdot)$ is differentiable over Ω .

In addition, suppose assumptions (HL), (HU), (HR), (FD) as well as the conditions (4.9), (4.15)–(4.17) take place.

Then the sequence $\{z^k\}$ generated by \mathfrak{R} -strategy turns out to be minimizing for problem (P).

Moreover every cluster point z of the sequence $\{z^k\}$ yields the infimum of $f(\cdot)$ over the feasible set of problem (P).

In the case of closed set S this cluster point turns out to be a global solution to (P).

Proof (for strictly resolving sets). For the case of merely resolving sets the proof is similar with the corresponding change of the sign of inequalities from strict for nonstrict and vice versa.

(a) For the case $\eta_k = \eta(\zeta_k) \leq 0$ for all $k \geq k_0 \geq 0$ from definition of resolving set it follows

$$f(z^k) \leq f_* + \Delta_k. \quad (4.19)$$

Since $\Delta_k \downarrow 0$ ($k \rightarrow \infty$) due to (4.17), the sequence $\{z^k\}$ turns out to be minimizing (see, definition of minimizing sequence, (2.10)–(2.12)):

$$\lim_{k \rightarrow \infty} f(z^k) = f_*, \quad g(z^k) = 0, \quad z^k \in S, \quad k = 0, 1, 2, \dots \quad (4.20)$$

(b) Consider now the case, when

$$\eta_k > 0, \quad k = 0, 1, 2, \dots \quad (4.21)$$

By constructing, z^{k+1} is a τ_{k+1} -critical point, $g(z^{k+1}) = 0$, obtained by some method of local search, starting from $x^{k+1} := u^{jk} \in S$, $f(x^{k+1}) \leq f(z^k)$. Besides,

$$\eta_k = \langle \nabla g(w^{jk}), x^{k+1} - w^{jk} \rangle = \max_{1 \leq i \leq N_k} \langle \nabla g(w^i), u^i - w^i \rangle > 0, \quad (4.22)$$

where $g(w^i) = g(w^{jk}) = 0$, $i = 1, \dots, N_k$, $k = 0, 1, 2, \dots$

Therefore, due to convexity of $g(\cdot)$ one has

$$0 < \eta_k \leq g(u^{jk}) - g(w^{jk}) = g(x^{k+1}). \quad (4.23)$$

In addition, in virtue of condition (FD)

$$f(z^{k+1}) \leq f(x^{k+1}) - \mu g(x^{k+1}) + \tau_{k+1}, \quad (4.24)$$

where $\mu > 0$, or, what is the same,

$$g(x^{k+1}) \leq \frac{1}{\mu} [f(x^{k+1}) - f(z^{k+1})] + \tau_{k+1}/\mu. \quad (4.25)$$

Thus, the numerical sequence $\{f(z^k)\}$ turns out to be monotonously decreasing, since

$$f(z^{k+1}) < f(x^{k+1}) \leq \eta_k := f(z^k). \quad (4.26)$$

Because $f(\cdot)$ is bounded from below over the feasible set of problem (P), the sequence $\{f(z^k)\}$ converges. Due to (4.26) it implies that $\{f(x^k)\}$ also converges to the same limit. Therefore, it follows from (4.25) that $g(x^k) \downarrow 0$. In turn, in virtue of (4.23) it implies

$$\eta_k \downarrow 0, \quad (k \rightarrow \infty). \quad (4.27)$$

On the other hand, since the approximation $\mathfrak{R}_k = \mathfrak{R}(\zeta_k)$ is strictly $(\Delta_k, \delta_k, \varepsilon_k, \theta_k)$ -resolving, according to Lemma 4.2 from the inequality $\eta_k > 0$ it follows

$$\eta_k > \theta_k \varphi(z^k) - \varepsilon_k \geq \theta \varphi(z^k) - \varepsilon_k. \quad (4.28)$$

From the latter inequality due to (4.17) and (4.27) one obtains

(E)

$$\lim_{k \rightarrow \infty} \varphi(z^k) = 0. \quad (4.29)$$

Therefore, according to Theorem 2.10 the sequence $\{z^k\}$ turns out to be minimizing, since $\lim_{k \rightarrow \infty} g(z^k) = 0$ by constructing, and from the regularity condition (R3) it follows the condition (R1) of Theorem 2.10.

(c) In general case the sequence $\{\eta_k\}$ is divided into two subsequences $\{\eta_{k_s}\}$ and $\{\eta_{k_t}\}$ by means of conditions $\eta_{k_s} \leq 0, \eta_{k_t} > 0$, so that $\{k_s\} \cup \{k_t\} = \{0, 1, 2, \dots\}$. Correspondently, the sequence $\{z^k\}$ is also decomposed into two subsequences $\{z^{k_s}\}$ and $\{z^{k_t}\}$. Moreover, the both subsequences are minimizing according, respectively, to parts (a) and (b) of the proof. Therefore, the whole sequence $\{z^k\}$ shows itself minimizing to problem (P). \square

To conclude, one can say a few words about the existence of the resolving set, the crucial role of which we observed during the proving of the convergence of GSS.

For some simple problems as presented for instance in [8, 12], we are able to construct the weakly resolving set, that is, when from inequality (4.3) it follows only (4.4) (without (4.5)).

Simultaneously, the existence of a weakly resolving set is shown in the sufficiency proof of Theorem 2.2 [8, 9, 10, 12]. Therein, supposing that there exists a feasible element $u \in S$, $g(u) \geq 0$, which is better than the point under study, $f(u) < f(z)$, one find a point y , $g(y) = 0$, such that $\langle \nabla g(y), u - y \rangle > 0$. (In the same manner Theorem 2.10 was proven.) It means, that the set $W = \{y\}$ of only one point, shows itself as the weakly resolving approximation for z .

Moreover, we are able to prove that if z is not a critical point in the sense, say, of the condition (see (2.5))

$$\langle \nabla f(z) - \lambda \nabla g(z), x - z \rangle \geq 0, \quad \forall x \in S, \quad (4.30)$$

where $g(z) = 0$, then the set $W = \{z\}$ consisting of only one point, turns out to be weakly resolving. But this is the topic for following papers (see [12]).

On the other hand, we can prove that if there exists a procedure for constructing a weakly (Δ, δ) -resolving set, then one can construct the strictly $(\Delta, \delta, \varepsilon, \theta)$ -resolving approximation for some θ , $0 < \theta < 1$, and $\varepsilon \geq \theta M \Delta > 0$.

As a consequence, we can replace assumption (HR) by weaker one concerning the construction of only a weak resolving approximation. How one can do it for convex maximization problems, reverse convex and d.c. minimization problems one can see in [8, 10, 11, 12, 13, 14, 15].

5. Conclusion

In this paper for a minimisation problem with one reverse convex constraint

- (i) after remember optimality conditions for a global solution;
- (ii) the necessary and sufficient conditions for a sequence of feasible points to be minimising are given;
- (iii) the global search strategy for finding a global solution was proposed;
- (iv) the new notion of resolving set has been advanced and some features of it were studied;
- (v) the convergence of proposed global search strategy was proved.

Acknowledgments

The author is really grateful to the referee for pertinent remarks and suggestions allowing to improve the presentation of the paper. The investigation was carried out under financial support of RFBR (Grant number 01-01-00048).

References

- [1] P. P. Bansal and S. E. Jacobsen, *An algorithm for optimizing network flow capacity under economies of scale*, J. Optimization Theory Appl. **15** (1975), no. 5, 565–586.
- [2] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons, New York, 1983.
- [3] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms. I. Fundamentals*, Grundlehren der Mathematischen Wissenschaften, vol. 305, Springer, Berlin, 1993.
- [4] ———, *Convex Analysis and Minimization Algorithms. II. Advanced Theory and Bundle Methods*, Grundlehren der Mathematischen Wissenschaften, vol. 306, Springer, Berlin, 1993.
- [5] R. Horst and P. M. Pardalos (eds.), *Handbook of Global Optimization*, Nonconvex Optimization and Its Applications, vol. 2, Kluwer Academic, Dordrecht, 1995.
- [6] R. Horst, P. M. Pardalos, and N. V. Thoai, *Introduction to Global Optimization*, Nonconvex Optimization and Its Applications, vol. 3, Kluwer Academic, Dordrecht, 1995.
- [7] R. Horst and H. Tuy, *Global Optimization. Deterministic Approaches*, Springer, Berlin, 1993.
- [8] A. S. Strekalovsky, *The Search for a global maximum of a convex functional on an admissible set*, Comput. Math. Math. Phys. **33** (1993), no. 3, 315–328.
- [9] ———, *Extremal problems on complements of convex sets*, Cybernet. Systems Anal. **1** (1994), 88–100.
- [10] ———, *Global optimality conditions for nonconvex optimization*, J. Global Optim. **12** (1998), no. 4, 415–434.

- [11] ———, *One way to construct a global search algorithm for d.c. minimization problems*, Nonlinear Optimization and Related Topics (Erice, 1998) (G. Di Pillo and F. Giannessi, eds.), Appl. Optim., vol. 36, Kluwer Academic, Dordrecht, 2000, pp. 429–443.
- [12] ———, *Elements of Nonconvex Optimization*, Nauka, Novosibirsk, 2003.
- [13] A. S. Strekalovsky and A. A. Kuznetsova, *The convergence of global search algorithm in the problem of convex maximization on admissible set*, Russian Math. (Iz. VUZ) **43** (1999), no. 12, 70–77 (Russian).
- [14] A. S. Strekalovsky, A. A. Kuznetsova, and T. V. Yakovleva, *On numerical solving of nonconvex optimization problems*, Sibirckii Jurnal Vichislitel'noi Matematiki **4** (2001), 185–199 (Russian).
- [15] A. S. Strekalovsky and I. Tsevendorj, *Testing the \mathfrak{R} -strategy for a reverse convex problem*, J. Global Optim. **13** (1998), no. 1, 61–74.
- [16] H. Tuy, *Convex programs with an additional reverse convex constraint*, J. Optim. Theory Appl. **52** (1987), no. 3, 463–486.
- [17] ———, *The complementary convex structure in global optimization*, J. Global Optim. **2** (1992), no. 1, 21–40.
- [18] F. P. Vasilév, *Numerical Methods for Solving of Extremal Problems*, Nauka, Moscow, 1988.

Alexander Strekalovsky: Institute of System Dynamics & Control Theory SB of RAS, 134 Lermontov Street, Irkutsk-33, 664033, Russia
E-mail address: strekal@icc.ru