A Trajectory Algorithm Based on the Gradient Method I. The Search on the Quasioptimal Trajectories

E. G. STURUA and S. K. ZAVRIEV
Moscow State University, Faculty of Computational, Mathematics & Cybernetics, SU-119899, U.S.S.R.

(Received: 22 April 1991; accepted: 16 May 1991)

Abstract. The global optimization problem is considered under the assumption that the objective function is convex with respect to some variables. A finite subgradient algorithm for the search of an ε-optimal solution is proposed. Results of numerical experiments are presented.

Key words. Projected subgradient method, global optimization, trajectory algorithms.

1. Introduction

Consider the global optimization problem

$$F(x) \to \min_{x \in X}, \tag{1}$$

where the objective function $F(\cdot)$ is Lipschitz continuous with the constant $L > 0$ on X, $X \subset E_k$ is convex and compact, E_k is a k-dimensional Euclidean space.

Denote

$$X_0^\varepsilon = \{x \in X \mid F(x) \leq \min_{x' \in X} F(x') + \varepsilon\}, \ \varepsilon \geq 0,$$

$$X_0^0 = X_0,$$

X_0^ε is the set of ε-optimal solutions of the problem (1),

X_0 is the optimal set.

To construct a trajectory algorithm (see [1]) for the global ε-optimization we need

(i) a class $\mathfrak{C}$ of curves C, $C \subset X$, satisfying

$$C \cap X_0^\varepsilon \neq \emptyset \quad \forall C \in \mathfrak{C}$$

where curve $C \subset E_k$ may be considered as some trajectory $C = \{c(t), t \in T\}$;

(ii) the very algorithm searching a global extremum on a curve (trajectory) $C \in \mathfrak{C}$.

Usually the class $\mathfrak{C}$ of trajectories is defined by ordinary differential equations

$$\begin{gathered} \dot{x} = G(x), \quad x(0) = x_0, \\ \mathfrak{C} = \{x(t), \quad t \in [0, \infty] \mid \dot{x} = G(x), \quad x(0) \in X^1\}, \end{gathered} \tag{2}$$

and as searching algorithms finite-difference analogues of (2) are considered (see [1–6]). In the present paper we consider a different approach to the definition of $\mathfrak{C}$. Thus, we have to construct special algorithms searching the global optimum on trajectories $C \in \mathfrak{C}$.

ASSUMPTION 1.1. *Let in problem* (1)

$$X = Y \times Z , \tag{3}$$

where $Y \subset E_{k_1}$, $Z \subset E_{k_2}$ *are compact,* $k_1 + k_2 = k$, Y *is convex,* Z *is simply connected.*

Set

$$x = (y, z) \in Y \times Z = X ,$$

$$F(x) = F(y, z) ,$$

$$Y_0^\beta(z) = \{y \in Y \mid F(y, z) \leq \min_{y' \in Y} F(y', z) + \beta\}, \ \beta \geq 0 ,$$

$$Y_0^\beta(z) = Y_0(z), \ z \in Z .$$

DEFINITION 1.2. The trajectory $\{y(z), z \in Z\}$ is said to be a quasioptimal trajectory of problem (1) (in respect to the decomposition (3)) if

$$y(z) \in Y_0(z) \quad \forall z \in Z . \qquad \square$$

DEFINITION 1.3. The finite set $Z^* = \{z^1, \ldots, z^N\}$ is called a γ-chain on Z, if Z^* has the following properties:

$$z^i \in Z , \quad i = 1, \ldots, N ;$$

$$\min_{1 \leq i \leq N} \|z - z^i\| < \gamma \quad \forall z \in Z ;$$

$$\| z^i - z^{i+1} \| < 2\gamma , \quad i = 1, \ldots, N - 1 . \qquad \square$$

DEFINITION 1.4. The finite trajectory $\{y(z), z \in Z^*\}$, $Z^* = \{z^1, \ldots, z^N\} \supset Z$, is said to be (α, β, γ)-quasioptimal trajectory of problem (1), $\alpha, \beta \geq 0$, $\gamma > 0$, if

$$Z^* \text{ is a } \gamma\text{-chain on } Z$$

and

$$\rho(y(z^i), Y_0^\beta(z^i))^2 \leq \alpha , \quad i = 1, \ldots, N ;$$

$$\rho(x, A) = \inf \|x - y\|, x \in E, A \subset E . \qquad \square$$

LEMMA 1.5. (1) *Every quasioptimal trajectory* $C = \{y(z), z \in Z^*\}$ *of problem* (1) *has the property that*

$$C \cap X_0 \neq \emptyset .$$

(2) *Every* (α, β, γ)*-quasioptimal trajectory* $C = \{y(z), z \in Z^*\}$ *of problem* (1) *has the property that*

$$C \cap X_0^\varepsilon \neq \emptyset ,$$

where $\varepsilon = L\alpha^{1/2} + \beta + L\gamma$.

Proof. Let us fix arbitrary $\varepsilon > 0$ and $\alpha, \beta \geq 0, \gamma > 0$, satisfying the estimation

$$L\alpha^{1/2} + \beta + L\gamma \leq \varepsilon .$$

Let $C = (y(z), z \in Z^*)$ be an arbitrary (α, β, γ)-quasioptimal trajectory of (1) and $x_0 = (y_0, z_0)$ be an optimal solution of (1),

$$x_0 = (y_0, z_0) \in X_0 .$$

By Definition 1.4 the finite set $Z^* \subset Z$ is a γ-chain on Z, thus there exists $z_0^* \in Z^*$ such that

$$\| z_0^* - z_0 \| < \gamma .$$

Set

$$\varphi(z) = \min_{y \in Y} F(x, y), \ z \in Z .$$

It is easily seen that $\varphi(\cdot)$ is Lipschitz continuous with the constant $L > 0$ on Z and

$$\varphi(z_0) = \min_{x \in X} F(x) .$$

Hence, we have

$$\varphi(z_0^*) \leq \varphi(z_0) + L \| z_0^* - z_0 \| < \min_{x \in x} F(x) + L\gamma .$$

By Definition 1.4 there exists $y_0(z_0^*) \in Y_0^\beta(z_0^*)$ such that

$$\| y_0(z_0^*) - y(z_0^*) \| \leq \alpha^{1/2} .$$

Applying the definition of $Y_0^\beta(z_0^*)$, we obtain

$$\begin{aligned} F(y(z_0^*), z_0^*) &\leq F(y_0(z_0^*), z_0^*) + L \| y(z_0^*) - y_0(z_0^*) \| \\ &\leq \varphi(z_0^*) + \beta + L\alpha^{1/2} \leq \min_{x \in X} F(x) + L\gamma + \beta + L\alpha^{1/2} \\ &\leq \min_{x \in X} (x) + \varepsilon . \end{aligned}$$

Therefore, we get

$$(y(z_0^*), x_0^*) \in X_0^\varepsilon$$

and

$$C \cap X_0^\varepsilon \neq \emptyset . \qquad \square$$

In Section 2 the finite subgradient algorithm for the convex parametric programming problem is considered. The infinite algorithms of this type were studied in [7, 8]. In Section 3 the subgradient parametric programming algorithm is applied to search an ε-optimal solution of the problem (1) on an (α, β, γ)-quasioptimal trajectory. Some results of numerical experiments are presented in Sections 4, 5.

2. The Finite Subgradient Algorithm for the Parametric Optimization

Consider the parametric programming problem: to search $x(t) \in X_0(t)$,

$$X_0(t) = \{x \in X \mid \Phi(x, t) = \min_{x' \in X} \Phi(x', t)\}, \ t \in T, \tag{4}$$

where the objective function $\Phi(\cdot, t)$ is continuous on $X \times T$, $X \subset E_k$, $T \subset E_S$ are compact, X is convex, T is simply connected, diam $X = \sup_{x', x'' \in X} \| x' - x'' \| = D$.

Set

$$X_0^\beta(t) = \{x \in X \mid \Phi(x, t) \leq \min_{x' \in X} \Phi(x', t) + \beta\}, \ t \in T, \ \beta \geq 0 .$$

DEFINITION 2.2. The discrete trajectory $\{x(t), t \in T^*\}$, $T^* = \{t^1, \ldots, t^N\}$, is called an (α, β, γ)-optimal solution of (4) if

$$T^* \text{ is a } \gamma\text{-chain on } T$$

and

$$\rho(x(t^i), X_0^\beta(t^i))^2 \leq \alpha, \quad i = 1, \ldots, N .$$

ASSUMPTION 2.3. *Let in* (4) *the objective function* $\Phi(\cdot, t)$ *be convex on* X *for every* $t \in T$, $L > 0$ *and*

$$|\Phi(x, t') - \Phi(x, t'')| \leq L \| t' - t'' \| \quad \forall t', t'' \in T$$

Let

$$\partial_x^{(\delta)} \Phi(x, t)$$

denote the set of all δ*-subgradients of the convex function* $\Phi(\cdot, t)$ *in the point* x (*see* [9]), $\delta \geq 0$.

Consider the following algorithm:

$$x^{n+1} = \mathrm{pr}_X(x^n - a\xi^n), \quad n = 1, \ldots, N-1, \tag{5}$$

where $\mathrm{pr}_X(\cdot)$ is the projector on X,

$$\xi^n \in \partial_x^{(\delta)} \Phi(x^n, t^n), \quad n = 1, \ldots, N-1,$$

$T^* = \{t^1, \ldots, t^N\}$ is a γ-chain on T,

a, δ, γ – are parameters, $a, \gamma > 0$, $\delta \geq 0$;

$x^1 \in X$ is an initial point.

Let us fix arbitrary α_0, β_0, $\gamma_0 > 0$. Our aim is to obtain some $(\alpha_0, \beta_0, \gamma_0)$-optimal solution of (4).

THEOREM 2.4. *Let Assumption* 2.3 *be fulfilled and in* (5)

$$\|\xi^n\| \leq K, \quad n = 1, \ldots, N-1.$$

Suppose that the parameters a, δ, γ *of algorithm* (5) *satisfy*

$$\begin{cases} \theta_0 + 2a\delta + a^2K^2 + \dfrac{8LD^2}{\beta_0}\gamma + 16\dfrac{L^2D^2}{\beta_0^2}\gamma^2 \leq \alpha_0, \\ 2a\delta + a^2K^2 + \dfrac{8LD^2}{\beta_0}\gamma + 16\dfrac{L^2D^2}{\beta_0^2}\gamma^2 \leq \dfrac{2a\beta_0}{D}\theta_0^{1/2}, \\ 0 < \gamma \leq \gamma_0,\ \delta \geq 0,\ a > 0, \end{cases} \tag{6}$$

for some $\theta_0 > 0$.

Suppose that the initial point $x^1 \in X$ *satisfies*

$$\rho(x^1, X_0^{\beta_0}(t^1))^2 \leq \alpha_0.$$

Then every trajectory $\{x(t^n) = x^n, n = 1, \ldots, N\}$ *of the algorithm* (5) *is an* $(\alpha^0, \beta_0, \gamma_0)$*-optimal solution of* (4), *i.e.*,

$$\rho(x^n, X_0^{\beta_0}(t^n))^2 \leq \alpha_0, \quad n = 1, \ldots, N. \qquad \square$$

REMARK 2.5. It is easily seen that for any α_0, β_0, $\gamma_0 > 0$ there exist a, δ, γ satisfying (6).

The proof is in the Appendix.

3. The Global Optimization Algorithm

Consider the global optimization problem (1).

ASSUMPTION 3.1. *Let the objective function* $F(\cdot, z)$ *be convex on* Y *for every* $z \in Z$.

The algorithm searching on an $(\alpha_0, \beta_0, \gamma_0)$-quasioptimal trajectory is as follows:

$$\begin{aligned} & y^{n+1} = \mathrm{pr}_Y(y^n - a\xi^n), \\ & x^{n+1} = (y^{n+1}, z^{n+1}), \\ & \bar{x}^{n+1} = \begin{cases} \bar{x}^n, & \text{if } F(\bar{x}^n) \leq F(x^{n+1}); \\ x^{n+1}, & \text{in opposite case}; \end{cases} \\ & n = 1, \ldots, N-1, \end{aligned} \tag{7}$$

where $\{z^1, \ldots, z^N\}$ is a γ-chain on Z;

$\xi^n \in \partial F_y^{(\delta)}(y^n, z^n)$, $n = 1, \ldots, N-1$,

parameters a, $\gamma > 0$, $\delta \geq 0$;

$y^1 \in Y$ is an initial point.

Let us fix an arbitrary $\varepsilon > 0$. Applying Theorem 2.4 and Lemma 1.5, we obtain

THEOREM 3.2. *Let Assumptions* 1.1, 3.1 *be fulfilled and in* (7)

$$\| \xi^n \| \leq K, \quad n = 1, \ldots, N-1.$$

Suppose that parameters a, δ, γ, *of the algorithm* (7) *satisfy*

$$\theta_0 + 2a\delta + a^2K^2 + \frac{8LD^2}{\beta_0}\gamma + 16\frac{L^2D^2}{\beta_0^2}\gamma^2 \leq \alpha_0,$$

$$2a\delta + a^2K^2 + \frac{8LD^2}{\beta_0}\gamma + 16\frac{L^2D^2}{\beta_0^2}\gamma^2 \leq \frac{2a\beta_0}{D}\theta_0^{1/2},$$

$$L\alpha_0^{1/2} + \beta_0 + L\gamma_0 \leq \varepsilon,$$

$$0 < \gamma \leq \gamma_0, \; a > 0, \; \delta \geq 0, \; \theta_0 > 0,$$

$$D = \text{diam } Y,$$

for some α_0, β_0, γ_0, $\theta_0 > 0$, *and the initial point* y^1 *satisfies*

$$\rho(y^1, Y_0^{\beta_0}(z^1))^2 < \alpha_0$$

(*it is obvious that* $y^1 \in Y_0^{\varepsilon}(z^1)$).

Then every trajectory of the algorithm (7) *has the following properties*:

(i) $\{y^n(z^n) = y^n, n = 1, \ldots, N\}$ *is an* $(\alpha_0, \beta_0, \gamma_0)$-*quasioptimal trajectory of* (1);

(ii) $\bar{x}^N \in X_0^{\varepsilon}$.

4. Numerical Examples

Conditions on parameters of algorithm (7) are rather restricting. If we set, for example, $D = 10$, $L = 100$ and $\varepsilon = 0.05$, then to satisfy the conditions we have to take very small values of a, δ and γ (for instance, $\gamma < 10^{-6}$). But analysing the proof of Theorem 2.4 it is easy to see that conditions (6) may be weakened and essentially greater values of parameters a, δ, γ may be used in algorithm (7). Numerical experiments show such abilities.

EXAMPLE 4.1. Let $F(\cdot)$ in (1) be defined by

$F(x_1, x_2) = x_1^4 + 4x_1^3 + 4x_1^2 + (x_2 - 2\sin(\pi/2x_1))^2$ (modified Treccani function [11]),

$$X = \{x = (x_1, x_2) \mid -3 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\},$$

$$X_0 = \{(-2, 0), (0, 0)\},$$

$$F^* = \min_{x \in X} F(x) = 0.$$

Set

$$y = x_2, \; z = x_1,$$

$$Y = [-1, 1], \; Z = [-3, 1]$$

(it is easily seen that diam $Y = D = 2$, $L \simeq 100$).

Set

$$z^1 = -3, \; z^{n+1} = z^n + \gamma, \; n = 1, \ldots, N,$$

$$x^1 = (-3.0, -0.95),$$

$$F(x^1) = F(y^1, z^1) \leq \min_{y \in Y} F(y, z^1) + 0.05,$$

$$\delta = 0.0.$$

Applying the algorithm (7), we get the results presented in Table I (see Figure 1).

EXAMPLE 4.2. Let $F(\cdot)$ in (1) be defined by

$$F(x_1, x_2) = x_2^2 - x_1 x_2 + 2x_1^2 - 1.05x_1^4 + \frac{1}{6}x_1^6$$ ("three-hump camel-back function" [2]),

$$X = \{x = (x_1, x_2) \mid -2.5 \leq x_1 \leq 2.5, \; -1 \leq x_2 \leq 1\},$$

$$X_0 = \{(0, 0)\}, \; F^* = 0.$$

Set

$$y = x_2, \; z = x_1,$$

$$Y = [-1, 1], \; Z = [-2.5, 2.5],$$

it is easily seen that diam $Y = D = 2$, $L \simeq 10$.

Table I

γ	a	N	$\bar{x}^N$	$F(\bar{x}^N)$
$8 \cdot 10^{-2}$	0.5	26	$(-0.039999, -0.125580)$	-0.006146
$4 \cdot 10^{-2}$	0.5	51	$(-4.768371 \cdot 10^{-7}, -1.490116 \cdot 10^{-6})$	$9.095568 \cdot 10^{-13}$
10^{-2}	0.5	201	$(-2.000000, 3.173947 \cdot 10^{-6})$	$3.637991 \cdot 10^{-12}$
10^{-3}	0.5	2001	$(2.799136 \cdot 10^{-5}, 8.793734 \cdot 10^{-5})$	$3.134153 \cdot 10^{-9}$
10^{-4}	0.1	20001	$(2.000034, 0.002620)$	$6.323266 \cdot 10^{-6}$
10^{-6}	0.05	2000001	$(-6.165428 \cdot 10^{-6}, -0.000584)$	$3.199272 \cdot 10^{-7}$

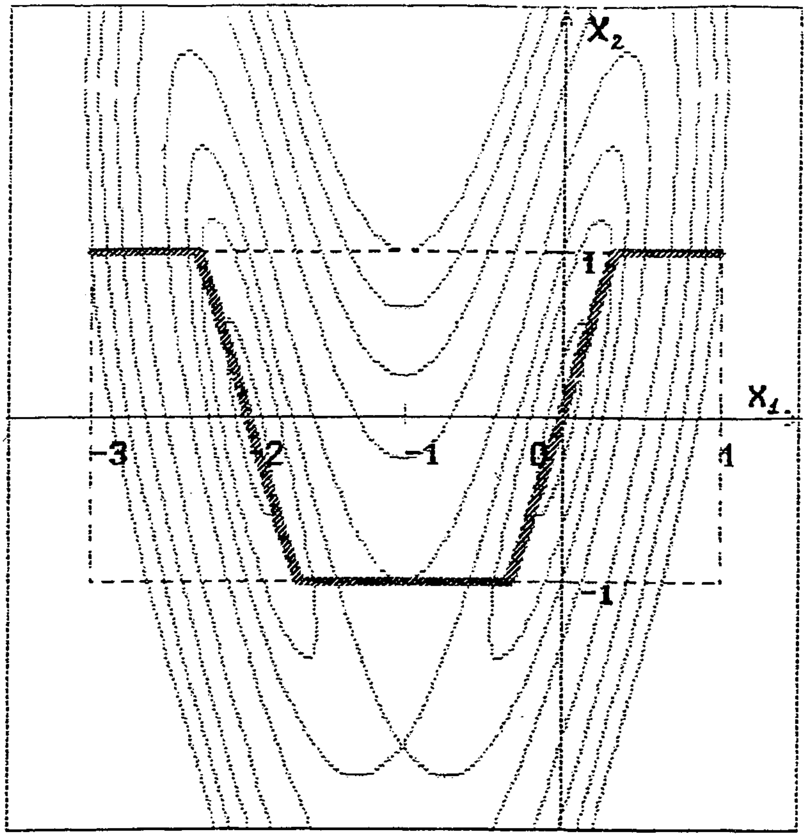

Fig. 1.

Set

$$z^1 = -3,\ z^{n+1} = z^n + \gamma\ , \quad n = 1, \dots, N\ ,$$

$$x^1 = (-2.5,\ -0.95)\ ,$$

$$F(x^1) = F(y^1, z^1) \leq \min_{y \in Y} F(y, z^1) + 0.05\ ,$$

$$\delta = 0.0\ .$$

Applying the algorithm (7), we get the results presented in Table II (see Figure 2).

Table II

γ	a	N	$\bar{x}^N$	$F(\bar{x}^N)$
$8 \cdot 10^{-2}$	0.5	32	$(-1.817941 \cdot 10^{-6}, -9.089708 \cdot 10^{-7})$	$5.783595 \cdot 10^{-12}$
$4 \cdot 10^{-2}$	0.5	63	$(2.682209 \cdot 10^{-7}, 1.341104 \cdot 10^{-7})$	$1.258992 \cdot 10^{-13}$
10^{-2}	0.5	251	$(-1.817941 \cdot 10^{-6}, -9.089708 \cdot 10^{-7})$	$5.783595 \cdot 10^{-12}$
10^{-3}	0.5	2501	$(4.626344 \cdot 10^{-6}, 2.313172 \cdot 10^{-6})$	$3.745536 \cdot 10^{-11}$
10^{-4}	0.1	25001	$(-9.912793 \cdot 10^{-6}, -0.000404)$	$1.601719 \cdot 10^{-7}$
10^{-6}	0.05	2500001	$(-6.165428 \cdot 10^{-6}, -9.308272 \cdot 10^{-5})$	$8.166523 \cdot 10^{-9}$

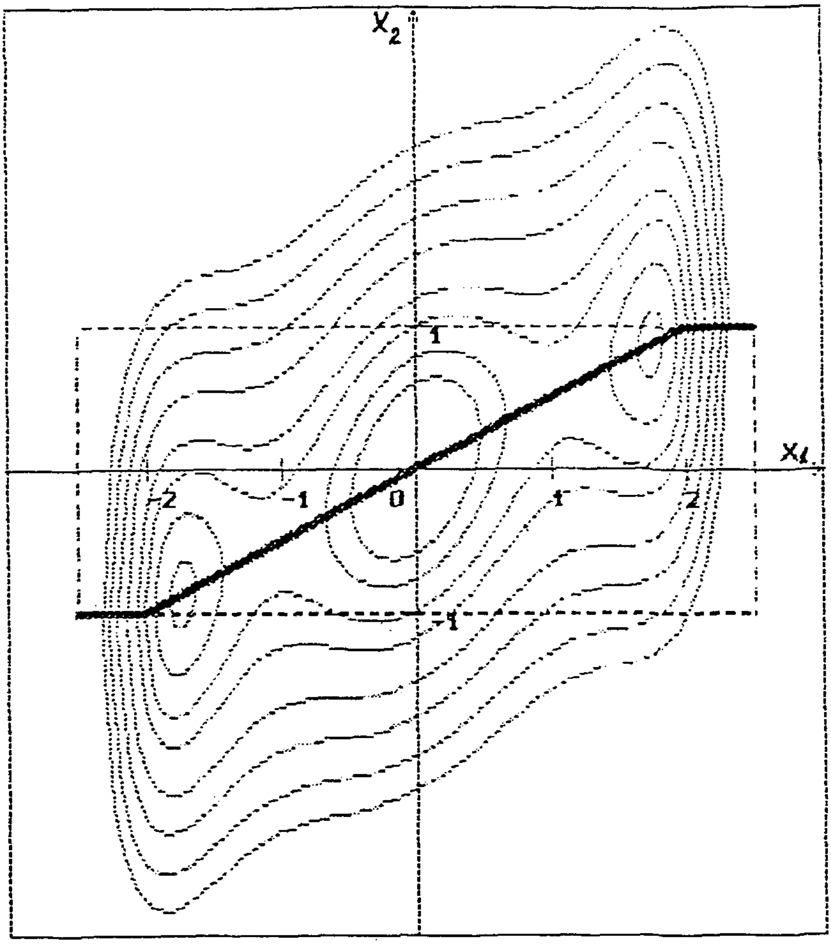

Fig. 2.

5. Conclusion

The algorithm (7) may be applied not only in problems (1) with the special "partly convex" structure. It is possible to use this algorithm, starting with all local optima $\{y_1^1, \ldots, y_S^1\}$ of the problem

$$F(y, z^1) \to \min_{y \in Y},$$

when Assumption 3.1 is not fulfilled. In this situation we obtain S trajectories

$$\{x_j^n = (y_j^n, z^n), \bar{x}_j^n, n = 1, \ldots, N\}, \quad j = 1, \ldots, S,$$

and set

$$\bar{x}^N = \underset{1 \le j \le S}{\operatorname{argmin}} F(\bar{x}_j^N).$$

EXAMPLE 5.1. Let $F(\cdot)$ in (1) be defined by

$$F(x_1, x_2) = 4x_2^2 - 2.1x_2^4 + \frac{1}{3} x_2^6 - x_1x_2 - 4x_1^2 + 4x_1^4 \quad \text{("six-hump camel-back function" [2])},$$

$$X = \{x = (x_1, x_2) \mid -1.5 \leq x_1 \leq 1.5,\ -2 \leq x_2 \leq 2\},$$

$$X_0 = \{x_0\},\ x_0 \simeq (-0.7,\ 0.1),\ F^* \simeq -1.031628.$$

Set

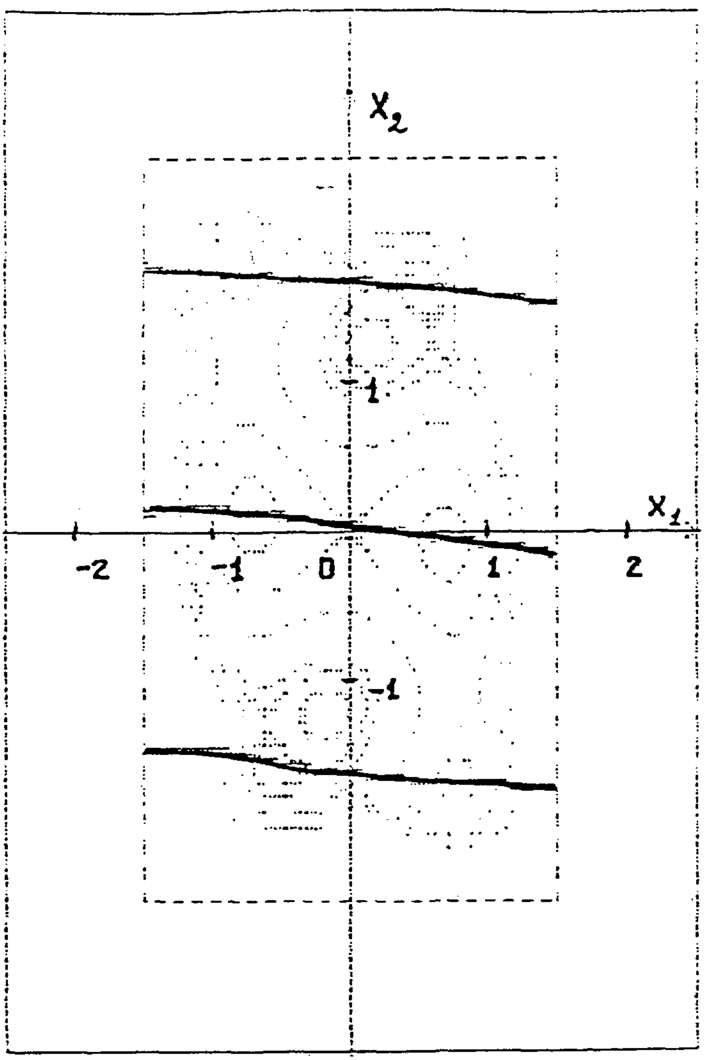

Fig. 3.

$$y = x_2, \; z = x_1,$$
$$Y = [-2, 2], \; Z = [-1.5, 1.5].$$
$$z^1 = -1.5,$$
$$y_1^1 = 0.2, \; y_2^1 = 1.8, \; y_3^1 = -1.5, \; S = 3$$

and

$$x^1 = (-1.5, 0.2), \; \delta = 0.0.$$

Applying the algorithm (7) with the starting points y_1^1, y_2^1 and y_3^1, we get the results, presented in Table III (see Figure 3).

Table III

γ	a	N	$\bar{x}^N$	$F(\bar{x}^N)$
10^{-1}	0.2	31	(−0.699999, 0.081739)	−1.030185
$5 \cdot 10^{-2}$	0.2	61	(−0.699999, 0.083760)	−1.030272
10^{-2}	0.2	301	(−0.710000, 0.090421)	−1.031568

6. Appendix

In this section we supply a proof to Theorem 2.4 (see [10]).

Set

$$\varphi(t) = \min_{x \in X} \Phi(x, t), \; t \in T.$$

LEMMA 6.1. *The function $\varphi(\cdot)$ is Lipschitz continuous with constant $L > 0$ on T.* □

LEMMA 6.2. *Let $f(\cdot)$ be convex on $X \subset E_k$, X is convex and compact.*

Then for every $\beta \geq 0$ and $x \in X$

$$\min(0, \min_{x' \in X} f(x') - f(x) + \beta) \leq -\frac{\beta}{D} \rho(x, X_\beta) = -\frac{\beta}{D} \| x - \mathrm{pr}_{X_\beta}(x) \|,$$

where $X_\beta = \{x \in X \mid f(x) \leq \min_{x' \in X} f(x') + \beta\}$. □

Set

$$h(A, B) = \max(\max_{x \in A} \rho(x, B), \max_{x \in B} \rho(x, A)),$$

where $A, B \subset X$.

COROLLARY 6.3. *For every $\beta > 0$, $t', t'' \in T$*

$$\rho(X_0^\beta(t'), X_o^\beta(t'')) \leq \frac{2LD}{\beta} \| t' - t'' \|.$$

Proof. Fix arbitrary $t', t'' \in T$, $\beta > 0$ and

$$x \in X_0^\beta(t') . \tag{8}$$

If $x \notin X_0^\beta(t'')$, then, applying Lemma 6.2, we get

$$-\Phi(\dot{x}, t'') - L \| t' - t'' \| + \varphi(t'') + \beta$$
$$\leqslant -\Phi(x, t'') + \varphi(t'') + \beta \leqslant -\frac{\beta}{D} \rho(x, X_0^\beta(t'')) .$$

By (8)

$$\Phi(x, t') \leqslant \varphi(t') + \beta ,$$

and, hence

$$-\varphi(t') - \beta - L \| t' - t'' \| + \varphi(t') + \beta \leqslant -\frac{\beta}{D} \rho(x, X_0^\beta(t'')) ,$$

$$\varphi(t'') - \varphi(t') - L \| t' - t'' \| \leqslant -\frac{\beta}{D} \rho(x, X_0^\beta(t'')) .$$

Now, we have by Lemma 6.1

$$-2L \| t' - t'' \| \leqslant -\frac{\beta}{D} \rho(x, X_0^\beta(t'')) ,$$

$$\rho(x, X_0^\beta(t'')) \leqslant \frac{2LD}{\beta} \| t' - t'' \| . \tag{9}$$

It is obvious, that if $x \in X_0^\beta(t'')$ then (9) is valid as well.
Thus,

$$\max_{x \in X_0^\beta(t')} \rho(x, X_0^\beta(t'')) \leqslant \frac{2LD}{\beta} \| t' - t'' \| ,$$

Similarly

$$\max_{x \in X_0^\beta(t'')} \rho(x, X_0^\beta(t')) \leqslant \frac{2LD}{\beta} \| t' - t'' \| ,$$

This completes the proof.

THEOREM 2.4. [10] *Let Assumption* 2.3 *be fulfilled and in* (5)

$$\| \xi^n \| \leqslant K , \quad n = 1, \ldots, N-1 . \tag{10}$$

Suppose that a, δ, γ *satisfy* (6) *for some* $\theta_0 > 0$ *and the initial point* $x^1 \in X$ *satisfies*

$$\rho(x^1, X_0^{\beta_0}(t^1))^2 \leqslant \alpha_0 . \tag{11}$$

Then every trajectory $\{x(t^n) = x^n, n = 1, \ldots, N\}$ *of* (5) *satisfies*

$$\rho(x^n, X_0^{\beta_0}(t^n))^2 \leqslant \alpha_0 , \quad n = 1, \ldots, N .$$

Proof. Fix an arbitrary trajectory $\{x^n\}$ of (5) and set

$$\theta_n = \rho(x^n, X_0^\beta(t^n))^2 , \quad n = 1, \ldots, N .$$

By (11) we have

$$\theta_1 \leqslant \alpha_0 . \tag{12}$$

Set

$$x^i(t^j) = \mathrm{pr}_{X_0^{\beta_0}(t^j)}(x^i) ,$$

$$\bar{x}^i(t^j) = \mathrm{pr}_{X_0^{\beta_0}(t^i)}(x^i(t^j)) , \quad i, j = 1, \ldots, N .$$

For any $n = 1, \ldots, N-1$ we have

$$\begin{aligned} \theta_{n+1} &= \| x^{n+1} - x^{n+1}(t^{n+1}) \|^2 \leqslant \| x^{n+1} - \bar{x}^{n+1}(t^n) \|^2 \\ &\leqslant (\| x^{n+1} - x^{n+1}(t^n) \| + \| x^{n+1}(t^n) - \bar{x}^{n+1}(t^n) \|)^2 \\ &\leqslant (\| x^{n+1} - x^n(t^n) \| + h(X_0^{\beta_0}(t^n), X_0^{\beta_0}(t^{n+1})))^2 . \end{aligned} \tag{13}$$

By assumption 2.3 and (10) we get

$$\begin{aligned} \| x^{n+1} - x^n(t^n) \|^2 &\leqslant \| \mathrm{pr}_X(x^n - a\xi^n) - x^n(t^n) \|^2 \leqslant \| x^n - a\xi^n - x^n(t^n) \|^2 \\ &\leqslant \| x^n - x^n(t^n) \|^2 - 2a\langle \xi^n, x^n - x^n(t^n) \rangle + a^2 \| \xi^n \|^2 \\ &\leqslant \theta_n + 2a\langle \xi^n, x^n(t^n) - x^n \rangle + a^2K^2 \\ &\leqslant \theta_n + 2a(\Phi(x^n(t^n), t^n) - \Phi(x^n, t^n) + \delta) + a^2K^2 . \end{aligned}$$

Then, it follows from Lemma 6.2 that

$$\Phi(x^n(t^n), t^n) - \Phi(x^n, t^n) \leqslant -\frac{\beta_0}{D} \| x^n(t^n) - x^n \| = -\frac{\beta_0}{D} \theta_n^{1/2} .$$

Thus,

$$\| x^{n+1} - x^n(t^n) \|^2 \leqslant \theta_n - \frac{2a\beta_0}{D} \theta_n^{1/2} + 2a\delta + a^2K^2 . \tag{14}$$

By Corollary 6.3, (14) and (13), we have

$$\theta_{n+1} \leqslant \theta_n - \frac{2a\beta_0}{D} \theta_n^{1/2} + 2a\delta + a^2K^2 + 2D \frac{2DL}{\beta_0} 2\gamma + \frac{16D^2L^2\gamma^2}{\beta_0^2}$$

$$\leqslant \begin{cases} \theta_0 + 2a\delta + a^2K^2 + \dfrac{8D^2L}{\beta_0} \gamma + \dfrac{16D^2L^2\gamma^2}{\beta_0^2} , & \text{if } \theta_n \leqslant \theta_0 ; \\ \theta_n - \dfrac{2a\beta_0}{D} \theta_0^{1/2} + 2a\delta + a^2K^2 + \dfrac{8LD^2}{\beta_0} \gamma + \dfrac{16D^2L^2\gamma^2}{\beta_0^2} , & \text{if } \theta_n > \theta_0 . \end{cases}$$

Applying (6), yields

$$\theta_{n+1} \leqslant \max(\alpha_0, \theta_n) , \quad n = 1, \ldots, N-1 .$$

Now it follows from (12) that

$$\theta_n \leqslant \alpha_0 , \quad n = 1, \ldots , N - 1 .$$

The theorem is established. □

References

1. Branin, F. and Hoo, S. (1972), A Method for Finding Multiple Extrema of a Function of *n* Variables, in F. A. Lootsma (ed.), *Numerical Methods to Nonlinear Optimization*, London, New-York: Academic Press, pp. 231–237.
2. Branin, F. H. (1972), Widely Convergent Methods for Finding Multiple Solutions of Simultaneous Nonlinear Equations, *IBM J. Res. Dev.* **16** (5), 504–522.
3. Gomulka, J. (1975), Remarks on Branin's Method for Solving Nonlinear Equations in L. C. W. Dixon and G. P. Szegö (eds.), *Towards Global Optimization*, Amsterdam, Oxford: North-Holland, New York: Amer. Elsevier, pp. 29–54.
4. Diener, I. (1986), Trajectory Nets Connecting All Critical Points of a Smooth Function, *Math. Progr.* **36** (3), 340–352.
5. Yamashita, H. (1979), A Continuous Path Method of Optimization and Its Applications to Global Optimization, *Survey of Mathematical Programming* (Proc. 9th International Math. Prog. Symp. Budapest, 1976), Amsterdam, North-Holland, V. 1., pp. 539–546.
6. Treccani, G. (1975), On the Convergence of Branin's Method: A Counter Example, in L. C. W. Dixon and G. P. Szegö (eds.), *Towards Global Optimization*, Amsterdam, Oxford: North-Holland, New York: Amer., Elsevier, pp. 107–116.
7. Eremin, I. I. and Mazurov, V. D. (1979), *Nonstationary Mathematical Programming Processes*, Nauka, Moscow (in Russian).
8. Gaivoronski, A. A. (1978), On Nonstationary Stochastic Programming Problems, *Kibernetika* **4**, 89–92.
9. Dem'janov, V. F. and Vasil'jev, K. S. (1985), *Nondifferentiable Optimization*, Nauka (in Russian).
10. Zavriev, S. K. (1989), The Finite ε-Subgradient Algorithm for the Approximate Solving of Parametrical Programming Problems, in P. S. Krasnoschokov and L. N. Korolev (eds.), *Computing Complexes and the Modelling of Complicated Systems*, MGU, Moscow, pp. 111–117 (in Russian).
11. Treccani, G., Trabattoni, L. and Szegö, G. P. (1972), A Numerical Method for the Isolation of Minima, in G. P. Szegö (ed.), *Minimization Algorithms Mathematic Theories and Computer Results*, London, New-York: Academic Press, pp. 239–289.