Homogeneous algorithms of global optimization

S. M. Elsakov*, V. I. Shiryaev†

*South Ural State University, esergeym@mail.ru
†South Ural State University, vis@prima.susu.ac.ru

Algorithms solving a problem (1) are considered

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

where $f(x): R^n \to R$ - objective function, $X \subset R^n$ - feasible set. It is supposed that algorithms of solving the problem (1) can be presented in the form of the following principal algorithm:

Algorithm 1.

Step 1. To choose the points $x_1, ..., x_M$, to calculate values of objective function $f(x_1), ..., f(x_M)$. To put $k = M$.

Step 2. To calculate

$$x_{k+1} = \operatorname*{argabsmax}_{x \in X} P\left(x, \{x_i, f(x_i)\}_{i=1}^{k}\right), \tag{2}$$

to calculate $f(x_{k+1})$.

Step 3. If some condition of stop φ is satisfied, then stop, otherwise put $k = k + 1$ and go to Step 2.

Definition 1 *The algorithm is homogeneous if for objective functions $f(x)$ and $f(x)+c$ sequences $\{x_k\}$ constructed by Algorithm 1 will coincide.*

Henceforth only homogeneous algorithms will be considered. It is supposed that the information about the points in which objective function is calculated is represented in the form of two functions $m_k(x)$ and $s_k(x)$ which satisfy the following requirements: A1. $m_k(x_i) = f(x_i),\ i = 1, k$ A2. $s_k(x_i) = 0,\ i = 1, k$ A3. $s_k(x) > 0,\ x \neq x_i,\ i = 1, k$ A4. $m_k(x)$ and $s_k(x)$ satisfy Lipchitz condition with constants L_m and L_s accordingly. It allows to present (2) in the form of

$$x_{k+1} = \operatorname*{argabsmax}_{x \in X} P\left(s_k(x), m_k(x)\right). \tag{3}$$

Functions $m_k(x)$ and $s_k(x)$ map multidimensional feasible set X on a plane. We will consider that the image of set X on this plane is some set $SM = \left\{(s,m) \in R^2 \mid s = s_k(x), m = m_k(x),\ x \in X\right\}$.

Theorem 1 *If the function $P(s,\ m)$ is doubly continuously differentiable, for any point $(s*, m*) \in R^2$, there exists the set SM, such that $(s*, m*) = \operatorname*{arg\,absmax}_{(s,m)\in SM} P(s, m)$ and for functions $m_k(x)$ and $s_k(x)$ the following conditions are satisfied: A5. $m(x, x_1,\ \dots x_k, f(x_1) + c, \dots f(x_k) + c) = m(x, x_1,\ \dots x_k, f(x_1), \dots f(x_k)) + c$; A6. $s(x, x_1,\ \dots x_k, f(x_1) + c, \dots f(x_k) + c) = s(x, x_1,\ \dots x_k, f(x_1), \dots f(x_k))$, then for any homogeneous algorithm criterion $P(s_k(x),\ m_k(x))$ can be presented in a form $P(s_k(x),\ m_k(x)) = C \cdot m_k(x) + p(s_k(x))$, where $C = const$ and $p(\cdot)$ - doubly continuously differentiable function.*

Theorem 1 allows formulating the condition of convergence for homogeneous algorithms. Convergence here is understood as coincidence of a set of limiting points of sequence $\{x_k\}$ and set of global minima of objective function.

Theorem 2 *In order that the set of limiting points generated by homogeneous algorithm with criterion $P(s_k(x), m_k(x))$ coincides with the set of global minima of Lipchitz function $f(x)$ with Lipchitz constant L on compact X it is sufficient that function $p(\cdot)$ is Lipchitz and $-\max_{x \in X}(P(s_k(x), m_k(x))) \leq \min_{x \in X} \max_{i=\overline{1,k}} (f(x_i) - K\|x - x_i\|)$, where $K = const$ and $K > L$.*

Theorem 3 *If for function $s_k(x)$ the condition*

$$s_k(x) \geq \min_{i=\overline{1,k}} \|x - x_i\|, \tag{4}$$

is satisfied, then the set of limiting points generated by algorithm with criterion $P(s_k(x),\ m_k(x)) = 2Ks_k(x) - m_k(x)$, where $K > L + L_m$, will coincide with the set of global minima of Lipchitz function $f(x)$ with Lipchitz constant L.

Theorem 4 *If the condition of kind $\min_{i=\overline{1,k}} \|x_i - x_{k+1}\| < \varepsilon$ is chosen as the criterion of a stop, then the homogeneous algorithm will provide the solution of the problem (1) with accuracy by value of objective function not worse than $L\frac{L_p+L_m}{K-L}\varepsilon$.*

These theorems allow to reduce the construction of algorithm of global optimization to a problem of choosing functions $m_k(x)$ and $s_k(x)$ satisfying the conditions A1)-A6), (4). Thus it is required to consider the necessity of solving the problem (2) on each iteration of algorithm. Formally, the problem (2) is also a problem of global optimization, however, it represents a simpler problem of global optimization rather than (1) because objective function is calculated much faster. Besides algorithms of its solving can use the additional information about the structure of $P(s_k(x), m_k(x))$, for example, all points of local minima can be easily found, the information on $P(s_k(x), m_k(x))$, obtained on the previous steps can be effectively used, the function can be constructed as differentiable, etc.

The analysis of existing algorithms of global optimization shows that the most effective algorithms (by quantity of calculations of objective function) demand the use of functions $m_k(x)$ and $s_k(x)$, for which solving of the problem (2) takes a lot of time. Thus, for fast calculated objective functions it is appropriate to choose algorithms which demand a greater number of calculations of objective function in order to minimize optimization time [2, 5, 6, 9], but having little time to solve the problem (2), and for problems with objective functions which time of calculation is very essential, it is necessary to choose algorithms with a minimum quantity of references to objective function [1, 3, 4, 7, 8].

As an example it is offered to consider functions $m_k(x) = \sum\limits_{i=1}^{k} c_i \|xr_i\|^3 + \sum\limits_{i=1}^{d} b_i x_i + b_0$ and $s_k(x) = \min\limits_{i=1,k} \|xx_i\|$, and as a method of solving the problem (2) to choose the method offered in [5]. For testing test functions of the kind:

$$f(x) = -\left\{ \begin{array}{c} \left(\sum\limits_{i=1}^{7} \sum\limits_{j=1}^{7} A_{ij} a_{ij}(x) + B_{ij} b_{ij}(x) \right)^2 + \\ \left(\sum\limits_{i=1}^{7} \sum\limits_{j=1}^{7} C_{ij} a_{ij}(x) + D_{ij} b_{ij}(x) \right)^2 \end{array} \right\}^{0.5},$$

were used where $a_{ij}(x) = \sin(i\pi x_1)\sin(j\pi x_2)$, $b_{ij}(x) = \cos(i\pi x_1)\cos(j\pi x_2)$, A_{ij}, B_{ij}, C_{ij}, D_{ij} - are uniformly distributed variables on $[-1; 1]$. The comparison was made on the sample of a hundred functions. The minimum coefficient was being selected wherein the algorithm in all cases found a minimum. Then the average quantity of the iterations spent on finding a global minimum with the obtained coefficient was calculated. Accuracy set in $\varepsilon = 0.01$. To solve the problem (2) the algorithm on the basis of adaptive-diagonal curves [5] with constant coefficient K=500 was used. On average, to solve the problem the algorithm needed 154 references to objective function. Application has demanded on the average 418 calculations of objective function under the same conditions of algorithm on the basis of adaptive diagonal curves. The suggested algorithm demands in 2.7 times less references to objective function when solving test problems.

References

[1] Crino, S., Brown, D.E. Global Optimization With Multivariate Adaptive Regression Splines// Systems, Man, and Cybernetics, Part B, IEEE Transactions on , vol.37, no.2, 2007, pp.333-340.

[2] Evtushenko Yu. G. , Malkova V. U. , Stanevichyus A. A. Parallelization of the Global Extremum Searching Process// Automation and Remote Control, Vol. 68, No 5, 2007, pp. 787-798.

[3] Gutmann, H. A Radial Basis Function Method for Global Optimization // Journal of Global Optimization 19, 3, 2001, pp. 201-227.

[4] Jones D. R. A Taxonomy of Global Optimization Methods Based on Response Surfaces // Journal of Global Optimization, Vol. 21, No. 4., 2001, pp. 345-383.

[5] Kvasov D.E., Sergeyev Ya.D. Multidimensional global optimization algorithm based on adaptive diagonal curves// Comput. Maths. Math. Phys., 43(1), 2003, pp. 40-56.

[6] Pijavski S.A. An algorithm for finding the absolute extremum of a function// USSR Comput. Math. and Math. Phys., 2, 1972, pp. 57-67.

[7] Regis R. G., Shoemaker C. A. A Stochastic Radial Basis Function Method for the Global Optimization of Expensive Functions// INFORMS JOURNAL ON COMPUTING, Vol. 19, No. 4, 2007, pp. 497-509.

[8] Shubert B. A sequential method seeking the global maximum of a function// SIAM J. Numer. Anal., 9, 1972, 379-388.

[9] Strongin R.G., Sergeyev Ya.D. Global Optimization with Non-Convex Constraints. Sequential and Parallel Algorithms. Kluwer Academic Publishers. Dordrecht. The Netherlands, 2000, 728 pp.