Inductive Approach to Minimizing Algorithmically the Given Objective Function

Galina Digo, Natalya Digo

Institute for Automation and Control processes, Far Eastern Division of Russian Academy of Sciences, Radio Street 5, Vladivostok, 690041, Russia

bernatsk@iacp.dvo.ru , digo@iacp.dvo.ru

Abstract. *The multidimensional global optimization problems particularity consists in its unsolvable at general case and multiextremal objective function. The situation when objective function is given algorithmically over n-dimensional parallelepiped, it satisfies the Lipschitz condition with unknown Lipschitz constant and its values reception (calculation in some point of feasible domain of optimizing parameters) requires of considerable calculating resources is considered. A priori information about Lipschitzian of objective function allows using algorithms variant from exhaustive search, in particular, bisection method, based on non-uniform covering of feasible domain and used the supposition about objective function minimum estimate existence in the feasible domain. In addition unknown constant Lipschitz estimation problem arises. The cases using global estimation of constant Lipschitz evaluating over all feasible domain and local estimations of constant Lipschitz calculating over all separated subdomains of feasible domain are considered. On basis of inductive approach possibility of alternate transition from local information to global information (and on the contrary) at adaptive estimation of local Lipschitz constants over different subsets of current partition (dividing) of search domain is suggested.*

Keywords

Multidimensional global optimization, algorithmically the given objective function, bisection method, non-uniform covering of feasible domain, inductive approach, adaptive estimation of global and local Lipschitz constants.

1 Introduction

Search problems of global extremum arise in areas of scientific and techniques such as designing, control, modeling of real processes or appearances, data analysis. The multidimensional global optimization problems particularity consists in its unsolvable at general case and multiextremal objective function. Difficulty of their numerical solving is connected with high dimension and a priori information deficiency about the objective function. Such problems solving algorithms represent iterative process generating sequence of points by prescribed rules set and including account conclusion criterion [1]. Global extremum is looked for of all found local solutions but search only of local solution part is probable with proof that remaining local solutions have not effect on ultimate result. So the all these methods are reductive to objective function value estimation over the some subdomain of the feasible domain, and they differ by these points selection the way. Failing universal algorithms of global optimization for concrete problem solution method selection the basic support happens in accordance with characteristics of the objective function and the feasible domain.

Extremum search of giving algorithmically the objective function is successfully realized such methods as the random search or scanning method (exhaustion search) [2-3]. Random search algorithms have simple structure, are easily realized and have low-sensitivity over optimization domain dimension increase. In scanning method the feasible domain of search is partitioned on elementary cells, in each cell the point for calculating of objective function value is selected by specific algorithm. Extremum is found among the objective function values at these selected points. Dense position of points always guarantees the finding global extremum, but grid point's number is exponentially increased together with dimension growth.

A priori information about Lipschitzian of objective function allows using algorithms variant from exhaustive search, in particular, bisection method. It is based on non-uniform covering of feasible domain and uses supposition about objective function minimum estimation existence in feasible domain. In addition, unknown constant Lipschitz estimation problem arises. It is difficult complex global optimization problem, and approaches to it's depend both having a priori information about objective function and its calculation algorithm complexity. Algorithmic representation of objective function does not allow applying interval analysis techniques for such functions.

Current level of development of computer aids and information processing technologies allows solving multidimensional problems of optimization at multiprocessing computers by way of computation process parallelization.

The situation when objective function is algorithmically given over *n*-dimensional parallelepiped is considered. In the search t domain it satisfies the Lipschitz condition with unknown Lipschitz constant and its values reception (calculation in some point of feasible domain of optimizing parameters) requires of considerable calculating resources.

2 Theoretical Part

Let $\varphi(\mathbf{x})$ be an algorithmically given multidimensional multiextremal function over *n*-dimensional parallelepiped $P \subset R^n$. In the search domain this function satisfies a Lipschitz condition with unknown Lipschitz constant *L*. Minimization problem

$$\varphi_* = \min_{\mathbf{x} \in P} \varphi(\mathbf{x}), \quad P = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{a} \le \mathbf{x} \le \mathbf{b}\}, \tag{1}$$

does not have analitical solution. Here and below vector ineqaulity $\mathbf{x} \le \mathbf{z}$, where $\mathbf{x}, \mathbf{z} \in \mathbf{R}^n$, means that $x^i \le z^i, 1 \le i \le n$. Below under algorithmically given function it is understood function for which exists its values evaluation algorithm for any feasible value of the argument.

We assume that there exists objective function minimum estimation in search domain. Denote by $\mathbf{X}_*$ the solutions set of problem (1) and introduce for consideration the set of its ε-optimal (approached) solutions

$$X_*^{\varepsilon} = \{\mathbf{x} \in P : f(\mathbf{x}) \le f_* + \varepsilon\}. \tag{2}$$

It is obvious that $\mathbf{X}_* \subset \mathbf{X}_*^{\varepsilon} \subset P$. Then the definition with the given accuracy ε of global minimum value of the function $\varphi(\mathbf{x})$ implies finding at least one point $\mathbf{x}_r \in \mathbf{X}_*^{\varepsilon}$ in which this approximate value is reached [4].

Global minimum function given by (1) is found by bisection method [5]. As the additional information only its values are used, but their evaluation in points of feasible domain requires considerable computational resources. Such methods foresee the search space partition in fewer subdomains, the evaluation for these subdomains of lower bound of objective function taking into account the Lipschitz constant value or its estimation. Lipschitz constant estimation search itself is difficult complex global optimization problem [6].

Approaches to its solving must depend both having a priori information about objective function and its calculation algorithm complexity. The inductive approach [7] obeys by such requirements. It gives possibility of alternate transition from local information to global information (and on the contrary) and allows to estimate adaptively the global Lipschitz constant *L* valid over the whole feasible domain or the local Lipschitz constants L_i valid over different subsets (*n*-dimensional parallelepipeds) $P_i \subseteq P$ of current partition (dividing) of search domain [6, 8-9]. Such balancing between global and local information under uncertainty provides the search algorithm work acceleration.

Thus, for successful application of bisection method with respect to algorithmically given objective functions it is necessary to solve the following significant problems:

• to select Lipschitz constant estimation method which is necessary for reception lower bounds of objective function at each several subdomains of feasible domain;
• to accelerate process of required calculations.

Depending on existing a priori information about objective function and function values calculation algorithm complexity the cases using global estimation K of constant Lipschitz L evaluating over all feasible domain and local estimations K_i of constant Lipschitz L_i calculating over all separated subdomains of feasible domain [8] are considered.

Evidently, if $K < L$, then sought-for extremum value of function can turn out uncertain, and application of upper raw count can reduce to algorithm convergence speeding down. Proposed in [10] and generalized for multidimensional case in [11] an adaptive estimation of global Lipschitz constant is based on application of informational statistical algorithm of global optimization. It allows find the global estimation K by formula

$$K = K(k) = (r + C/k)\max(\lambda(k), \xi), \quad (3)$$

where $r > 1$, $C > 0$, $\xi > 0$ - small constant, $\lambda(k)$ - the current estimate of Lipschitz constant at an iteration $k \geq 1$ of an algorithm [11].

An initial value $\lambda(1)$ is found by formula

$$\lambda(1) = \left|\varphi(\mathbf{a}) - \varphi(\mathbf{b})\right| / \left\|\mathbf{a} - \mathbf{b}\right\|,$$

and the current value $\lambda(k)$ after performance the iteration $k \geq 1$ of the algorithm is adjustable by formula

$$\lambda(k+1) = \max\{\lambda(k), \max_{\mathbf{x},\mathbf{y} \in P_i, \mathbf{x} \neq \mathbf{y}} (\left|\varphi(\mathbf{x}) - \varphi(\mathbf{y})\right| / \left\|\mathbf{x} - \mathbf{y}\right\|)\}. \quad (4)$$

The parameters r, C, ξ from (3) are chosen (defined) taking into account having a priori information about properties of the objective function. They control estimate K, influence on the algorithm convergence speed. By decreasing r and C the search speed is accelerated but herewith method convergence probability to a point which is not the point of global minimum function $\varphi(\mathbf{x})$ increases.

Found in (3) the value $K = K(k+1)$ is used at calculation of lower bounds $g(P_i)$ for objective function $\varphi(\mathbf{x})$ on set of multidimensional subdomains $\{P_i\} \subseteq P$:

$$g(P_i) = \varphi(\mathbf{c}_i) - (L/2)\max_{1 \leq j \leq n}\left|b_i^j - a_i^j\right|,$$

where constant L is substituted by its estimation K.

Received at (3) the Lipschitz constant estimations for solution of problem (1)-(2) can turn out to be strongly overpriced and can be not reached over the some parallelepipeds $P_i \subset P$. Analogous situation is same possible in a priori known exact value L for the initial parallelepiped $P \subset R^n$. Then, even in case of known exact value of the Lipschitz constant, it can be needed go over to its estimate.

The passage to local estimations of Lipschitz constants allows more nearly estimating the behavior pattern of function $\varphi(\mathbf{x})$ over each several parallelepiped $P_i \subset P$ and allows escaping its strong overstating.

For local estimation calculation of unknown Lipschitz constant L two algorithms are selected [6]. At each from multidimensional parallelepipeds $P_i \subset P$ the estimate K_i of the local constant L_i is realized by one thing of the formula

$$K_i = K_i(k) = (r + C/k)\max\{\lambda_i(k), \gamma_i, \xi\}, \quad (5)$$

$$K_i = K_i(k) = \max\{(r + C/k)(\nu\mu + (1-\nu)\lambda_i(k)^2/\mu), \xi\}, \quad (6)$$

where $r > 1$, $C > 0$, $\xi > 0$ are analogous of the same name parameters used into (3). They exert significant influence on an algorithm convergence rate, $0 \le \nu \le 1$, and $\lambda_i(k)$ is the Lipschitz constant current estimate at kth iteration, $k \ge 1$, of the algorithm

$$\lambda_i(k) = |\varphi(\mathbf{a}_i) - \varphi(\mathbf{b}_i)| / \|\mathbf{a}_i - \mathbf{b}_i\|, \quad (7)$$

$$\gamma_i = \|\mathbf{a}_i - \mathbf{b}_i\| / \max_{1 \le j \le m} \|\mathbf{a}_j - \mathbf{b}_j\|, \quad (8)$$

$$\mu = \max_{1 \le i \le m} (|\varphi(\mathbf{a}_i) - \varphi(\mathbf{b}_i)| / \|\mathbf{a}_i - \mathbf{b}_i\|). \quad (9)$$

Unlike the algorithm represented by formula (5), at the expression (6) it is combined global and local information received earlier. The parameter ν provides balance between the global information (the μ parameter) and the local information ($\lambda_i^2(k)/\mu$ value). At both expressions local information about the objective function is used not only around local minimum but also over the whole search domain.

Series of step on the bisection method application for the problem solution (1) in the case adaptive estimation of global and local Lipschitz constants is particularized in [12]. Herewith for calculation process acceleration the encoded densely information about edge bisection numbers of the initial parallelepiped *P*, characteristic list of subdomains P_i in ascending order of values lower bounds $g(P_i)$ of the objective function, the procedure of local minimization for (1) [5], multisequencing technology on evidence, but at estimation of local Lipschitz constants and by stream are used.

3 Conclusion

During the search of global minimum of the given algorithmically objective function by covering methods, and in particularly by bisection method, the finding of Lipschitz constant estimation is one of the main questions.

In practical view point it is of interest adaptive estimation along the strike of algorithm working both global Lipschitz constant *L* over the whole feasible domain and local constants L_i over the separated subdomains obtaining at the current steps of partition. The alternate transition from local information to global information (and on the contrary) accelerates a search process of global minimum.

The solving delivered problem on basis of inductive approach allows not only to use having a priori and current information about objective function, but soundly to select the parameters included in formula for Lipschitz constants estimation taking into account dimension and particularity objective function.

4 Acknowledgements

The work is realized under financial support of 06-I-П15-054 grant of Far Eastern Division of Russian Academy of Sciences (program No15 Energetic, Mechanical Angering, Mechanics and Control Processes Division of Russian Academy of Sciences).

References

[1] Orlyanskaya I.V. Current Approaches to the Construction of Global Optimization Methods Investigated in Russia, 5, 2097-2108, 2002. http://zhurnal.ape.relarn.ru/articles/2002/189.pdf. (in Russian)

[2] Batischev D.I. Searching method of optimal designing, Moscow: Sovetskoe radio. 1975. (in Russian)

[3] Zhigljavsky A.A., Zilinskas A.G. Methods for finding of the global extremum, Moscow: Nauka.1991. (in Russian)

[4] Evtushenko Yu. Numerical methods for finding global extreme (case of a non-uniform mesh). Computational Mathematics and Mathematical Physics, - 1971. Vol. 11, N 6, pp. 38-54. (in Russian)

[5] Evtushenko Yu. and Ratkin V. (), Bisection method for global optimization, Izvestija Akademii Nauk AN USSR, Tehnicheskaya Kibernetika, - 1987. No 1, pp. 119-128 (in Russian)

[6] A. Molinaro, C. Pizzuti, Ya.D. Sergeyev. Acceleration tools for diagonal information global optimization algorithms // Computational Optimization and Applications. – 2001. Vol. 18, no 1, pp. 5-26.

[7] Ivakhnenko A. G., Madala H.R. Inductive learning algorithms for complex systems modeling.-London, Tokyo: CRC Press, 1994.

[8] Yaroslav D. Sergeyev and Dmitri E. Kvasov. Global Search Based on Efficient Diagonal Partitions and a Set of Lipschitz Constants // SIAM: Journal of Optimization. – 2006. Vol. 16, no 3, pp. 910-937.

[9] D. R. Jones, C. D. Perttunen, and B. E. Stuckman. Lipschitzian optimization without the Lipschitz constant // Journal of Optimization Theory and Applications. – 1993. – Vol. 79, no 1, pp. 157-181.

[10] Strongin R.G. Numerical Methods on Multiextremal Problems, Moscow: Nauka.1978. (in Russian)

[11] Kvasov D.E. and Sergeyev Ya.D. Multidimensional global optimization algorithm based on adaptive diagonal curves. Computational Mathematics and Mathematical Physics, - 2003. Vol. 43, no 1, pp. 42–59. (in Russian)

[12] Digo G.B., Digo N.B. Algorithmically given function global extremum search by the bisection methods and in the case of a non-uniform mesh: efficiency analysis // In: Proceeding of the VII International Conference “System Identification and Control Problems” SICPRO’08. Moscow, January 28-31, 2008. V.A. Trapeznikov Institute of Control Sciences/ Moscow: V.A. Trapeznikov Institute of Control Sciences, 2008. pp 512-525.