



DETERMINISTIC CHAOS PRODUCED BY LOCAL OPTIMISATION ALGORITHMS

Vitaly E. Podobedov

United Card Service, Smolnaya, 22, 125445 Moscow, Russia

E-mail: vetix@mail.ru

Received 12 October 2006 ; accepted 15 February 2007

Abstract. Deterministic chaos situations produced by local optimisation algorithms are considered. A family of test multi-extremal functions is proposed that allows getting the collections of functions with an arbitrary pre-defined number of the local optima. Special software is developed for generating test functions from the family, performing multiple local optimisations of different initial points, visualising the chaotic attraction regions. Properties of the regions of attraction are presented. Regularity of the attraction regions is required to be a special characteristic of the local optimisation algorithms.

Keywords: global and local optimisation, region of attraction, deterministic chaos.

1. Introduction

Many global optimisation algorithms combine global exploration of the objective function domain with a number of local optimisation searches: in this way, comprehensive global coverage is complemented by local algorithms efficiency.

Multistart algorithms are the widespread examples of such an approach. They use global covering points as the starting ones for the local optimisation algorithms, so a great number of local searches are performed.

Multiple local searches, however, often duplicate each other, starting from different points but converging to the same local optimum. These duplications can be considerable, so a number of the local searches must be optimised.

Special clustering techniques are often used for this purpose; they a priori create the clusters of potential starting points for local searches. Process of forming the clusters can be based, for example, on the distances between the points: the points inside a cluster are “close” to each other, and the clusters are “far” from each other [1].

Actually, the clusters are used here for optimisation of some real objects, namely the regions of attraction of local optima.

Definition 1. The region of attraction (RA) of a local optimum x^* is a set of the points such that the local search, starting from any of these points, converges to x^* .

Other different names used for RA are area, domain, or basin of attraction. Being the RA models, the clusters are

used instead of them in global optimisation algorithms. So, as only one local search must be ideally performed for each RA [2], then it is desirable to start only one local search from each cluster. Generally, if a new point is assigned to a cluster with a local search already performed from any of its points, then there is no need in starting a local search from this point.

However, the notions of RA and cluster are not sufficiently adequate. In particular, two typical problems are connected with clustering methods [1]: underclustering (a cluster may contain several Ras, so some local optima can be missed) and overclustering (RA may be divided into several clusters, so the same optimum can be located several times).

In addition, the “attraction region” term is sometimes applied not for an actual RA. For example, in [3, 4] it means, for simplicity, only a special object used in a test function construction, not any real RA.

Generally, Ras can have a very complex form, especially in the deterministic chaos situations. It is hardly or even impossible to produce any close approximation to Ras by means of clustering approaches. Thus, there cannot be a great sense in clustering in such cases.

The problem of describing the Ras for local optimisation algorithms is considered in the article. Simple Nelder-Mead and approximate gradient methods are applied many times to test functions from a family of functions with a fixed number of local optima. The phenomenon of deterministic chaos for corresponding Ras is demonstrated.

2. Test functions family

Let F_0 be a family of functions defined on $K = [0,1]^n$ [5]

$$F_0 = \left\{ f : f(\mathbf{x}) = \min_{i=1,\dots,m} (\rho(\mathbf{x}, \mathbf{p}_i) + c_i), \right. \\ \left. \mathbf{x}, \mathbf{p}_i \in K, c_i \geq 0, i = 1, \dots, m \right\}. \quad (1)$$

Here $\rho(\mathbf{x}, \mathbf{p}_i)$ is a distance between vectors \mathbf{x} and \mathbf{p}_i , eg

$$\rho = \rho_1(\mathbf{x}, \mathbf{p}_i) = \sum_{j=1}^n |q_{ij}(x_j - p_{ij})|^\beta, \text{ or} \quad (2.1)$$

$$\rho = \rho_2(\mathbf{x}, \mathbf{p}_i) = \max_{j=1,\dots,n} |q_{ij}(x_j - p_{ij})|^\beta, \quad (2.2)$$

$$\beta > 0, \mathbf{p}_i = (p_{i1}, \dots, p_{in}), q_{ij} > 0, i = 1, \dots, m, j = 1, \dots, n.$$

Note 1. An arbitrary function f from F_0 consists of m components f_i , $f_i(\mathbf{x}) = \rho(\mathbf{x}, \mathbf{p}_i) + c_i$, $i = 1, \dots, m$. So

$$f(\mathbf{x}) = \min_{i=1,\dots,m} f_i(\mathbf{x}) \text{ and } f_i(\mathbf{p}_i) = c_i.$$

Let S be the optimisation operator:

$$S(f) = \min_{\mathbf{x} \in K} f(\mathbf{x}).$$

Let $A = A(x_0, f)$, $x_0 \in K$, $f \in F_0$, be a local minimisation algorithm used to find $S(f)$, starting from an initial point x_0 . Then A will usually converge to one of the local minima of a function f .

Let us require that all points $\mathbf{p}_1, \dots, \mathbf{p}_m$ must be the local minima points, so next constraints must be satisfied:

$$f_i(\mathbf{p}_i) < f_j(\mathbf{p}_i), j \neq i, \text{ ie}$$

$$c_i < \min_{j \neq i} (\rho(\mathbf{p}_i, \mathbf{p}_j) + c_j), i = 1, \dots, m. \quad (3)$$

Now, let us define the test functions family F as one consisting of the functions satisfied by (1)–(3).

An arbitrary point \mathbf{x} from K belongs to one of the RAs, corresponding to the local minima $(\mathbf{p}_1, \dots, \mathbf{p}_m)$, or to none of them. In the former case, we can postulate that a non-empty RA exists for each \mathbf{p}_i , $i = 1, \dots, m$, due to (3). In the latter case, the local search A converges to any other point, ie on the boundary of K . In view of the fact that there are relatively few such points in most cases, we shall treat all such points as belonging to one complementary RA.

Definition 2. The complementary RA is a set of all the points such that the local search, starting from them, does not converge to any of the local minima $\mathbf{p}_1, \dots, \mathbf{p}_m$.

Thus we can split K into $m + 1$ RAs: m for the pre-defined local minima and 1 for the complementary RA.

Thus, we construct the family, each function of which has the known characteristics:

- Number of local minima: m ;
- Points of all local minima and their values: $(\mathbf{p}_1, \dots, \mathbf{p}_m), (c_1, \dots, c_m)$;

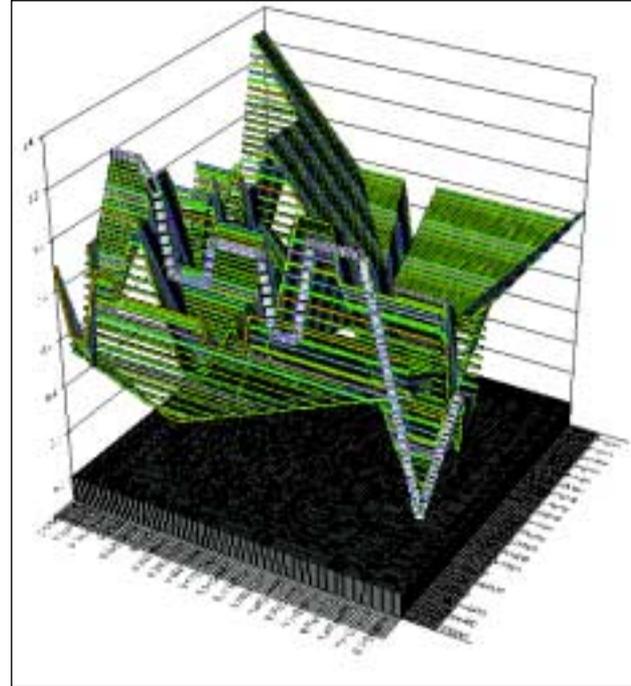


Fig 1. Test function from the family F for $\rho = \rho_2$

- Points of global minima \mathbf{p}^* and its value c^* :

$$(c^* = c_k = \min_{i=1,\dots,m} c_i), \quad (4)$$

$$\mathbf{p}^* = \mathbf{p}_k, \text{ for all } k, \text{ satisfied (4);}$$

- Number of regions of attractions: $m + 1$.

Also, the family F has other useful characteristics:

- Actual error of an optimisation algorithm can be determined for each function from F ,
- Collections of functions with the same number of local minima can be chosen.

Further in the article, the values $n = 2$, $m = 15$ are considered. Such small values are chosen in order to have good visualisation pictures. Fig 1 presents a function graph for a case when a distance ρ_2 is chosen, ie

$$f(x) = \min_{i=1,\dots,m} \max_{j=1,\dots,n} |q_{ij}(x_j - p_{ij})|^\beta + c_i.$$

2.1. Setting the parameters

We can previously set next parameters of the test functions family F :

- Dimension: n ;
- Number of local minima: m ;
- Metric: ρ_1 or ρ_2 ;
- α_j : maximal possible slope for j -th dimension, ie $q_{ij} \in [0, \alpha_j]$, $i = 1, \dots, m, j = 1, \dots, n$; (5)
- γ : maximal possible value for the local minima, ie $c_i \in [0, \gamma]$, $i = 1, \dots, m$; (6)
- β : a power in a distance function (2.1) or (2.2).

Also, we can set the coefficients q_{ij} the same for all dimensions in every component, ie $q_{ij} \equiv q_i$, $j = 1, \dots, n$, $i = 1, \dots, m$.

Next parameters are chosen randomly, by means of a random number generator:

- Local minimums ($\mathbf{p}_1, \dots, \mathbf{p}_m$), such that $\mathbf{p}_i \in K, i = 1, \dots, m$;
- Coefficients ($q_{11}, \dots, q_{mn}, c_1, \dots, c_m$), taking into account (5), (6).

3. Deterministic chaos

The phenomenon of deterministic chaos is well-known. It appears for deterministic systems or algorithms when a negligible change of their initial conditions can result in a considerable change of the results.

Deterministic chaos is demonstrated by:

- *Many dynamic systems.* Compound pendula, oscillating chemical reactions, predator-prey ecologies are the examples of chaotic systems. Having a fully deterministic behaviour and starting from very close initial conditions such systems quickly move to different states.
- *For numerical analysis algorithms.* Deterministic Newton’s method for solving the simple equations produces a chaotic picture for the regions of attraction of their roots (as shown in Fig 2).

In optimisation, however, deterministic chaos situations are practically unknown and uninvestigated [7].

We show here that a deterministic chaos can be observed for the local optimisation algorithms. We take perfectly deterministic local optimisation algorithm, without any random components. Then we see that very small changes of the algorithm’s starting point result in large changes of the solution, and RAs of the local optima are chaotic. Multi-extremum functions from the family F described by (1)–(3) are good testbeds for producing deterministic chaos and chaotic RAs.

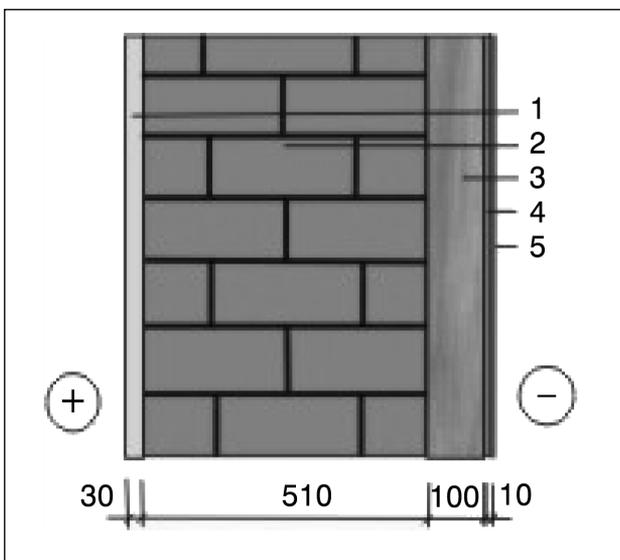


Fig 2. Regions of attraction for Newton’s algorithm of solving the system of two equations:
 $x - \sin(x) \cosh(y) = 0, y - \cos(x) \sinh(y) = 0$ [6]

3.1. Nelder-Mead method

Let us consider the Nelder-Mead method [8] designed for the minimisation of a function of n variables. It performs the comparison of function values at the $n + 1$ vertices of a general simplex, followed by the replacement of the vertex with the highest value by another point. The simplex adapts itself to the local landscape, and contracts to the final local minimum.

The Nelder-Mead method is the deterministic iterative algorithm, not depending on a random choice of any parameters. Does this algorithm converge to the local minimum closest in some sense to its initial point? Is the algorithm stable with respect to its initial point? The answers to the both questions are negative. The algorithm is, of course, entirely deterministic and knowledge of its initial point determines the point of its convergence uniquely. However, even a small change of the initial point can lead the algorithm very far from the point of its former convergence. Fig 3 illustrates this situation. Two different colours here mark the points of two different local searches started from two close points.

The phenomenon of deterministic chaos is not determined by the size of an initial simplex used in the Nelder-Mead algorithm. Any decrease of the simplex size does not eliminate the deterministic chaos.

Some pictures illustrating the deterministic chaos situations are presented here. They are formed in the following way. Fixed parameters for the family F are set (as in section 2.1), and the corresponding random parameters are chosen, fixing a test function f from F . Additionally, the conditions (3) must be checked, thus generating test functions may require much time in case of many local minima and high dimensions.

Then the local optimisation algorithm is applied many times to the function f . The initial points for these optimisations are taken from the uniform cubic grid U consisting of $N \times N$ points, $U = \{\mathbf{u}_{ij} = (u_i, u_j), i, j = 1, \dots, N\} \in K$, where

$$u_k = \frac{1}{2N} + \frac{k-1}{N}, k = 1, \dots, N.$$

A unique colour h_i is assigned to each local minimum $\mathbf{p}_i, i = 1, \dots, m$. A point is painted with the colour h_j , if the local algorithm, starting from this point, has converged to

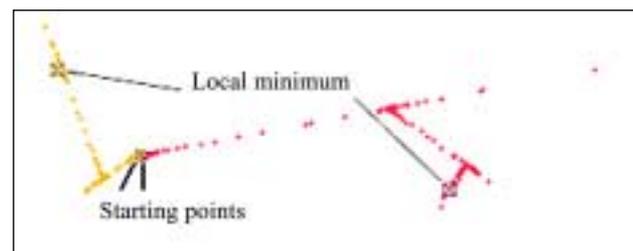


Fig 3. Instability of Nelder-Mead algorithm path due to a small change of its initial point

$\mathbf{p}_j, j \in 1, \dots, m$ (ie the point belongs to RA of \mathbf{p}_j). Points from the complementary RA (starting from which, the local search does not converge to any of the points $(\mathbf{p}_1, \dots, \mathbf{p}_m)$) are painted by the colour h_{m+1} . So the region K is splitted into $m + 1$ RAs. Crosses mark the local minima points. Situation of deterministic chaos in Fig 4, 5 is obvious.

3.2. Approximate gradient method

A variant of gradient algorithm is also considered for performing multiple local optimisation searches,

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_k \mathbf{g}(\mathbf{x}^k), \quad \alpha_k = \frac{1}{k + k_0}, \quad k = 1, 2, \dots, k_0 = 9.$$



Fig 4. Regions of attraction for the Nelder-Mead algorithm ($\rho = \rho_1, N = 700$)

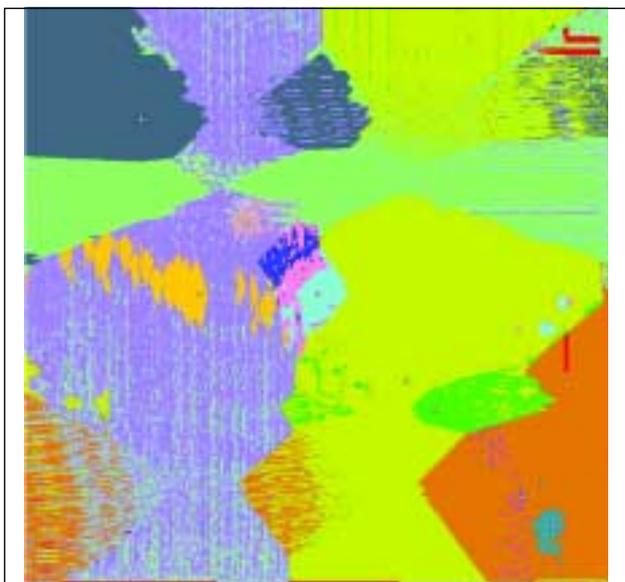


Fig 5. Regions of attraction for the Nelder-Mead algorithm ($\rho = \rho_2, N = 700$)

Here the function gradient $\text{grad } f(\mathbf{x})$ is replaced by its approximation $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x})), \mathbf{x} = (x_1, x_2) \in K$,

$$g_1(\mathbf{x}) = \frac{f(x_1 + \delta, x_2) - f(x_1 - \delta, x_2)}{2\delta},$$

$$g_2(\mathbf{x}) = \frac{f(x_1, x_2 + \delta) - f(x_1, x_2 - \delta)}{2\delta}, \quad \delta > 0.$$

Figs 6 and 7 are made for the same function from the considered family F . We can see that the phenomenon of deterministic chaos is observed for various local optimisation algorithms, although shapes of the corresponding RAs are different.

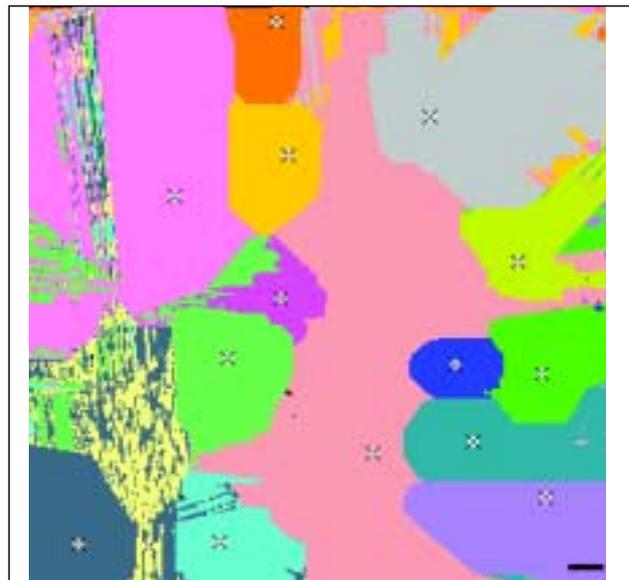


Fig 6. RAs for the Nelder-Mead algorithm

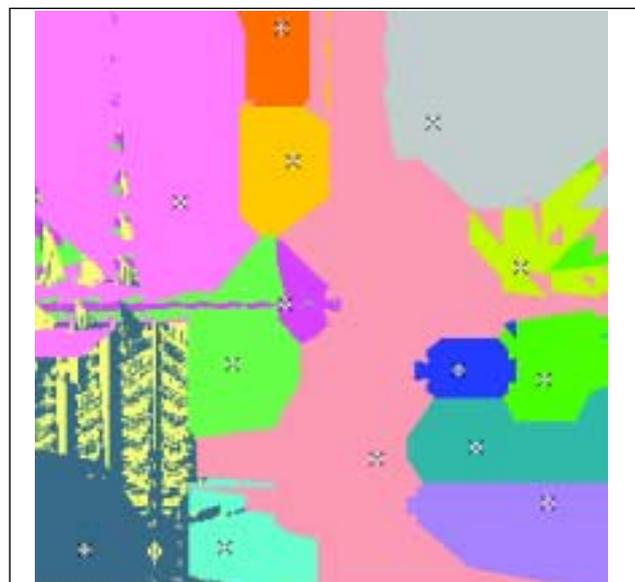


Fig 7. RAs for the approximate gradient algorithm for the same function as in Fig 6

4. RAs properties

Some hypotheses about RAs are expressed in the literature dealing with RAs in optimisation algorithms [1–4, 9–11]:

- 1) RA is a contiguous, convex set;
- 2) RA containing the global optimum usually has the largest size;
- 3) Size of RA (diameter, Lebegue measure) reflects its significance;
- 4) With a reasonable number of restarts, the global optimum's RA will be found.

Our experiments with creating chaotic pictures of RAs show that these hypotheses can be wrong. Instead, we discover next properties of RAs:

- 1) RA generally is neither convex, nor even connected set;
- 2) Points of different RAs can alternate, forming a deterministic chaos picture;
- 3) RA's size does not reflect the significance of the corresponding local optimum;
- 4) RAs depend on a local optimisation algorithm;
- 5) RAs depend rather from a distance between local optima than from the values in these points;
- 6) Clustering based on a distance between the points is not an adequate technique for producing RAs;
- 7) Notion of RA can be non-productive and ineffective even for simple local optimisation algorithms, such as the approximate gradient or Nelder-Mead methods.

5. Software

Special software is developed, that allows:

- *Setting the test family parameters* (like in section 2.1);
- *Generating test functions*. A special code is assigned to each combination of parameters, which uniquely defined a test function. This code is later used for reconstructing the function, in order to draw its

graph, apply another local optimisation algorithms to it, etc;

- *Storing all information about actual generated functions in the database;*
- *Performing multiple local optimisations from different initial points;*
- *Storing and visualising the optimisation results;*
- *Compute the characteristics of the RAs, such as radius, diameter, percentage of boundary points etc.*

Process of preparing the data for the pictures like Fig 4, 5 (performing of 490 000 local Nelder-Mead searches and storing the results in the MS SQL Server database) requires nearly 25 min on the IBM PC/ Intel Celeron 1,8 GHz computer.

Fig 8 demonstrates the screen for setting the parameters of the test family, optimisation algorithm and visualisation method.

6. Conclusions

Local optimisation algorithms must be checked for chaotic character of RAs generated by them. Regularity of RAs must be a special characteristic of such algorithms.

References

1. SOLOMATINE, D. P. Two strategies of adaptive cluster covering with descent and their comparison to other algorithms. *J. of Global Optimization*, 1999, Vol 14, No 1, p. 55–78.
2. TÖRN A.; ZILINSKAS A. *Global optimization. Lecture Notes in Comp. Science*, Vol 350. Berlin: Springer Verlag, 1989. 250 p.
3. GAVIANO, M.; LERA D. Test functions with variable attraction regions for global optimization problems. *J. of Global Optimization*, 1998, Vol 13, Iss 2, p. 207–223.
4. GAVIANO, M.; KVASOV, D. E.; LERA, D.; SERGEYEV, Y. D. Algorithm 829: Software for generation of classes of test functions with known local and global minima for global optimization. *ACM Transactions on Mathematical Software*, 2003, Vol 29, Iss. 4, p. 469–480.
5. PODOBEDOV, V. E.; SUKHAREV A. G. Global maximum seeking algorithm for a function of several variables. *Computational Mathematics and Modeling*, 1992, Vol 3, No 1, p. 49–55.
6. EPUREANU, B. I.; GREENSIDE H. S. Fractal basins of attraction associated with a damped Newton's method. *SIAM Review*, 1998, Vol 40, No 1, p. 102–109.
7. PODOBEDOV, V. E. Efficiency Analysis of Optimization Algorithms in the Deterministic Chaos Conditions. In *Proc of Intern. Conf on Artificial Intelligence Systems AIS'04*, Vol 2. Moscow: Fizmatlit, 2004, p. 329–334 (in Russian).
8. NELDER J. A.; MEAD R. A simplex method for function minimization. *Computer J.*, 1965, No 7, p. 308–313.
9. ADDIS, B.; LOCATELLI, M.; SCHOEN, F. *Local optima smoothing for global optimization*. Technical Report DSI 5-2003, Dipartimento di Sistemi e Informatica, Univ. degli Studi di Firenze. Firenze, 2003.

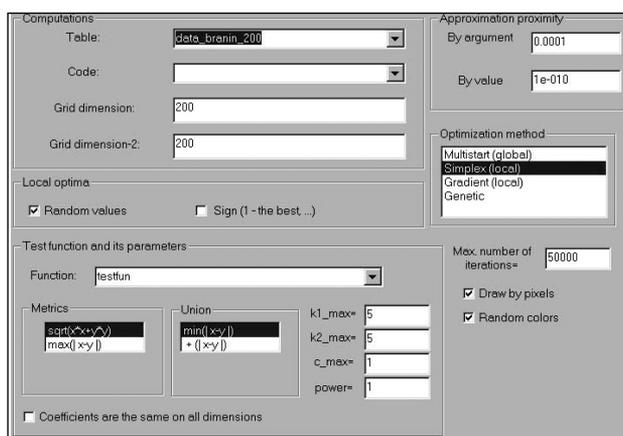


Fig 8. Screen for setting the test family parameters

10. TÖRN A.; ALI, M.; VIITANEN, S. Stochastic global optimization: problem classes and solution techniques. *J. of Global Optimization*, 1999, Vol 14, p. 437–447.
11. SEUNG HWAN OH, BYUNG CHAI LEE. Development of an efficient algorithm for global optimization by simplex elimination. *Engineering Optimization*, 2003, Vol 35, No 6, p 607–625.

APIBRĖŽTO CHAOSO SUFORMAVIMAS LOKALIAISIAIS OPTIMIZAVIMO ALGORITMAIS

V. E. Podobedov

Santrauka

Straipsnyje nagrinėjamos lokaliųjų optimizavimo algoritmų suformuotos apibrėžto chaoso situacijos. Pasiūlyta daugiaekstreminių funkcijų šeima, kuri leidžia nustatyti iš anksto pasirinktų lokaliųjų opimumų reikšmes. Sukurta speciali programinė įranga, kuri iš funkcijų šeimos kuria testavimo funkcijas, atlieka kartotinį lokalinį optimizavimą pagal įvairius pradinius taškus ir vaizduoja chaoso zonas. Pateikos tų zonų savybės. Zonų taisyklumas lemia lokalojo optimizavimo algoritmų ypatybes.

Reikšminiai žodžiai: globalusis ir lokalusis optimizavimas, chaoso zona, apibrėžtas chaosas.

Vitaly E. PODOBEDOV. PhD. Dept of Computational Mathematics and Cybernetics. Moscow State University, Russia.

Degree in applied mathematics, Moscow State University, Russia (1992). Associated researcher in Bordeaux Business School, France, since 2005, Member of the EUROPT Working Group within EURO since 2004. Author of about 30 research articles.

Research interests: global optimisation, approximation, numerical analysis, data analysis.