DECISION OF OPTIMIZATION PROBLEMS USING SYMMETRIC ALGORITHM OF HEAVY BALL METHOD

Kharlamova Y.N.
National Mining University, Dnipro, Ukraine
E-mail: harlamyshka@gmail.com

Abstract: *The algorithm of the heavy ball method, based on the principle of symmetry, to find a global extremum is described. The computer simulation of the method for the three test functions (Ackley, Griewank and Schwefel) is carried out. The results of the study of the efficiency of this algorithm are given. The results of mathematical modeling in the graphs, describing the process of convergence of representative points to the global optimum point of test functions, are shown. Conclusions about the efficacy of the described algorithm applied to optimization problems are drawn.*

KEYWORDS*: HEAVY BALL METHOD, ENERGY INTERACTION BETWEEN THE TWO BALLS, EXTREMUM SEEKING PROCESS, CONCEPT OF SYMMETR*

Y.

1. Introduction

One of the main factors that must be considered in solving the problems of synthesis of modern adaptive identification systems and information measuring systems is the inertia of the control object and measuring equipment. The creators of these systems use various optimization methods to ensure high efficiency of taken decisions.

In most cases, the seeking global extremum of multiextremal objective functions is carried out dynamically using the heavy ball method. But a moving heavy ball can have both a lack of kinetic energy and an excess of it. In the first case, it can stop at the one of the local extremes, before reaching the global extremum, and in the second case - to jump over it.

In the current situation, it is actual to organize the boosting of additional kinetic energy of the ball moving to the global extremum or to remove its excess.

2. Analysis of the literature sources

Today, both global optimization algorithms for solving a separate class of problems [1] and for more universal ones have been developed. When considering dynamic optimization problems, the relaxation method [2,3] is often applied, the realization of which is carried out by means of nonstationary processes that are described by vector differential equations of the form

(1) $m \cdot x^{(2)}(t) + r \cdot x^{(1)}(t) + gradf\,(x) = 0, \quad m > 0, \quad r > 0;$

(2) $\frac{dx}{dt} + k\ grad\ f\left(x\right) = 0, \quad k > 0.$

These processes to solve the task are eventually established.

The equation (1) describes the *heavy ball method*, which (with an appropriate choice of the parameters *m* and *r*) is referred to methods of seeking for a global extremum. The relaxation method, realized according to equation (2), is called the *steepest gradient descent method*, it is usually referred to methods of seeking for a local extremum of the function.

It is known that surface elongation of the objective function along one of the directions and its complex relief sharply worsens the effectiveness of these methods. The heavy ball method and the steepest gradient descent method in solving seeking global extremum problems, with an increase of the oscillations amplitude do not give a positive result, and the process of motion of the representing point stops at the first local extremum.

The purpose of the article is to construct algorithms for seeking for a global extremum of functions based on the principle of symmetry of the interaction of two heavy balls and to justify the advantage of this algorithm in seeking for a global extremum of the multiextremal function in comparison with other relaxation algorithms.

3. Results and discussion

The problem of seeking for the minimum of a multiextremal function can be solved by using the concept of symmetry, which has proven itself in such one-dimensional optimization methods such as methods of dichotomy, Fibonacci and golden section. In these methods, two *representative* points move symmetrically to the extremum of the function, significantly reducing the interval of uncertainty (localization).

Let us consider multidimensional methods improving process of function extremum seeking by applying the concept of symmetry [4,5].

Let us represent the expression of a downward-convex function *f(x)* (*x* is a vector argument), which extremum is sought, in the form

(3) $f\left(x\right) = 0.5\Big(\left(x - x\right)^T Q\left(x - x\right) + f\left(x\right) + f\left(x\right)\Big),$

where *Q* is a positive definite symmetric matrix.

Then, replacing one of the vectors *x* with the vector *y* and the other with the vector *z* in the expression (3), we obtain an auxiliary function

(4) $F\left(y,z\right) = 0.5\Big[\left(y - z\right)^T Q\left(y - z\right) + f\left(y\right) + f\left(z\right)\Big],$

the extremum of which will take place under the condition that *y=z=x**, where *x** is the value of the vector argument at which the function *f(x)* takes an extreme value.

The motion to a minimum of the auxiliary function F (y, z) is ensured as a result of a simultaneous coherent change of vector arguments y and z by any of the known extremum seeking algorithms.

The algorithm (1) of a heavy ball method [6.7] when working with an auxiliary function *F(y,z)* will be such as

(5)
$$m\frac{d^2y}{dt^2} + r\frac{dy}{dt} + \frac{\partial F\left(y,z\right)}{\partial y} = 0,$$
$$m\frac{d^2z}{dt^2} + r\frac{dz}{dt} + \frac{\partial F\left(y,z\right)}{\partial z} = 0.$$

It should be noted that for the extremum seeking of the auxiliary function (5), both continuous and discrete algorithms of several converging points can be used. This allows them to overcome local extremes.

Let us consider the efficiency of the heavy ball method, based on the principle of symmetry, on the example of three standard test functions: *Ackley*, *Griewank* and Schwefel [8].

The *Ackley* function (Fig. 1, a) has many local extremums near the global optimum. On the interval [-7; 3], the function takes a minimum value at the point 0, where *x*=0 and corresponds to the following description

(6) $f(x) = 20 + e - 20 \cdot \exp(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}) - \exp(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i))$.

Then the auxiliary symmetric function *F(y,z)* according to the equation (4) for the *Ackley* function (6) will have the following form:

$$(7)\quad \begin{aligned} F(y,z) = 0.5(2(20+e) - 20\exp(-0.2y^2) - \exp(\cos(2\pi y)) - \\ -20\exp(-0.2z^2) - \exp(\cos(2\pi z)) + 0.5q(y-z)^2, \end{aligned}$$

and the corresponding algorithm (5) for the function (7)

$$(8)\quad \begin{cases} y_1'(t) = y_2, \\ y_2'(t) = -\frac{r}{m}y_2 - \frac{1}{m}\left[\begin{aligned} q(y_1 - z_1) + 2y_1\sqrt{y_1^2}\exp\left(-0.2\sqrt{y_1^2}\right) + \\ +\pi\sin(2\pi y_1)\exp(\cos(2\pi y_1)) \end{aligned}\right], \\ z_1'(t) = z_2, \\ z_2'(t) = -\frac{r}{m}z_2 - \frac{1}{m}\left[\begin{aligned} q(z_1 - y_1) + 2z_1\sqrt{z_1^2}\exp\left(-0.2\sqrt{z_1^2}\right) + \\ +\pi\sin(2\pi z_1)\exp(\cos(2\pi z_1)) \end{aligned}\right]. \end{cases}$$

where $y_1 = y, \quad z_1 = z$.

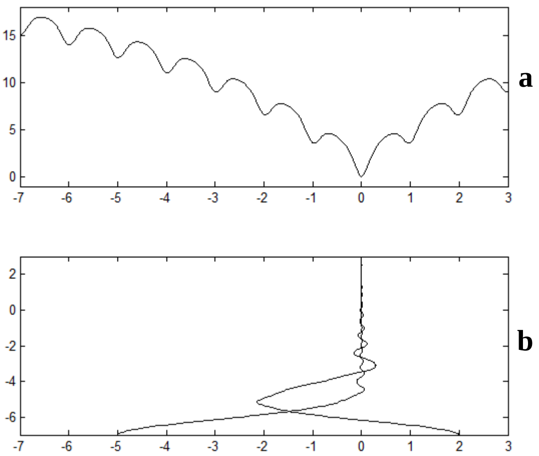

Fig.1. *Graphs illustrating the process of the global extremum seeking of a function with use of the concept of symmetry: a - the type of the Ackley test function; b - solution of the system (8)*

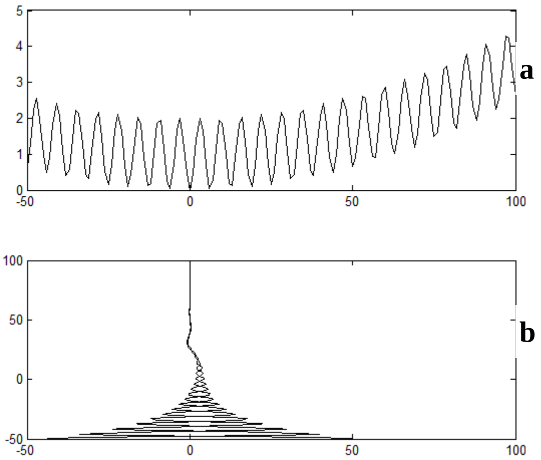

Fig.2. *Graphs illustrating the process of the global extremum seeking of a function with the use of the concept of symmetry: a - the type of the Griewank test function; b - solution of the system (11)*

For the solution of the system (8) initial conditions were set: $y_1(0) = -5, z_1(0) = 2$ and the following parameters were chosen: a mass of a heavy ball *m = 1.5*, damping coefficient *r = 2* and *q = 2*. Fig. 1(b) shows the result of the solution of the system (8), which describes the process of convergence of the *representative* points to the point of the global optimum ($x^* = 0$). The following values are obtained: $y_1 = 0.000002$ and $z_1 = -0.001472$. The solution of the problem (6) with the usual heavy ball method according to the expression (1), with constant values of the parameters *m* and *r*, gave the following results: x = - 4.9862 (from the initial point *x* (0) = -5) and *x* = 1.9745 (from the point *x* (0) = 2).

There are many local minimums in the *Griewank* function (Fig. 2, a). On the interval [-50; 100], the function takes a minimum value at the point 0, where *x* = 0 and has such description

$$(9)\quad f(x) = 1 + \sum_{i=1}^{n}\frac{x_i^2}{4000} - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right).$$

Then the auxiliary symmetric function *F(y,z)* for the *Griewank* function (9) is written so

$$(10)\quad \begin{aligned} F(y,z) = 0.5(2 + \frac{y^2}{4000} - \cos(y) + \frac{z^2}{4000} - \\ -\cos(z)) + 0.5q(y-z)^2 \end{aligned},$$

and the corresponding algorithm (5) for the function (10)

$$(11)\quad \begin{cases} y_1'(t) = y_2; \\ y_2'(t) = -\frac{r}{m}y_2 - \frac{1}{m}\left[q(y_1 - z_1) + \frac{y_1}{4000} + 0.5\sin(y_1)\right], \\ z_1'(t) = z_2; \\ z_2'(t) = -\frac{r}{m}z_2 - \frac{1}{m}\left[q(z_1 - y_1) + \frac{z_1}{4000} + 0.5\sin(z_1)\right]. \end{cases}$$

For the solution of the system (11) initial conditions were set: $y_1(0) = -47$, $z_1(0) = 53$ and the following parameters were chosen: a mass of a heavy ball *m = 7*, damping coefficient *r = 1* and *q = 2*. Fig. 2(b) shows the result of the solution of the system (11), which describes the process of convergence of the *representative* points to the point of the global optimum ($x^* = 0$). The calculated values were: y_1 = - \0.000254 and $z_1 = -0.000862$. The solution of the problem (9) with the usual heavy ball method according to the expression (1) gave the following results: *x* = –43.9603(from the initial point *x* (0) = -47) and *x* = 50.2403 (from the point *x* (0) = 53).

The *Schwefel* function (Fig. 3(a)) as well as the two above-mentioned functions, has a local minimums near global optimum. On the interval [-500; 500], the function takes a minimum value at the point 0, where *x* = 420 and corresponds to the following description

$$(12)\quad f(x) = 418.9829 \cdot n + \sum_{i=1}^{n}(-x_i \cdot \sin(\sqrt{|x_i|})).$$

For the Schwefel function (12), the auxiliary symmetric function *F(y,z)* will be

$$(13)\quad \begin{aligned} F(y,z) = 0.5(2 \cdot 418.9829 - y\sin(\sqrt{|y|}) - \\ -z\sin(\sqrt{|z|})) + 0.5q(y-z)^2 \end{aligned},$$

and the corresponding algorithm (5) for the function (13)

$$(14)\quad \begin{cases} y_1'(t) = y_2; \\ y_2'(t) = -\dfrac{r}{m} y_2 - \dfrac{1}{m}\left[\begin{array}{l} q(y_1 - z_1) - 0.5\sin(\sqrt{|y_1|}) - \\ -\dfrac{y_1 \cos(\sqrt{|y_1|}) \cdot \mathrm{sign}(y_1)}{4\sqrt{|y_1|}} \end{array} \right], \\ z_1'(t) = z_2; \\ z_2'(t) = -\dfrac{r}{m} z_2 - \dfrac{1}{m}\left[\begin{array}{l} q(z_1 - y_1) - 0.5\sin(\sqrt{|z_1|}) - \\ -\dfrac{z_1 \cos(\sqrt{|z_1|}) \cdot \mathrm{sign}(z_1)}{4\sqrt{|z_1|}} \end{array} \right]. \end{cases}$$

For the solution of the system (14) initial conditions were set: $y_1(0) = -200, z_1(0) = 480$ and the following parameters were chosen: a mass of a heavy ball $m = 57.8$, damping coefficient $r = 0.67$ and $q = 2$. Fig. 3(b) shows the result of the solution of the system (14), which describes the process of convergence of the *representative* points to the point of the global optimum ($x^* = 420$).The following values are obtained: $y_1 = 420.0194$ and $z_1 = 420.0862$.The solution of the given problem (12) by the heavy ball method according to the equation (1) from the initial point $x(0)$ 0= -4.5 did not give a positive result, the representing point completed its motion at x = -124.8345, from the point $x(0) = 3.75$ the global optimum point is found with a low accuracy $x = 420.8741$.

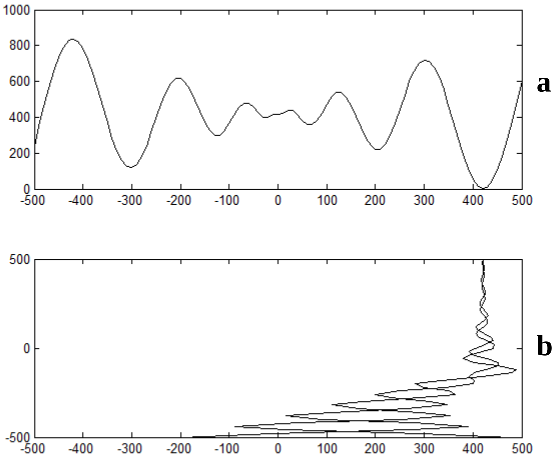

***Fig.3**. Graphs illustrating the process of the global extremum seeking of a function with the use of the concept of symmetry: a - the type of the Schwefel test function; b - solution of the system (14)*

Let us analyze the work of the heavy ball method by comparing it with the method of coordinate descent and the gradient method with a constant step, applying the concept of symmetry to them.

The algorithm of the steepest gradient descent method of when working with $F(y,z)$ is:

$$(15)\quad \begin{array}{l} \dfrac{dy}{dt} = -kQ(y-z) - 0{,}5k\ grad\ f(y), \quad y(0) = y_0 \\ \dfrac{dz}{dt} = -kQ(z-y) - 0{,}5k\ grad\ f(z), \quad z(0) = z_0, \end{array} \quad y_0 \neq z_0.$$

In contrast to the differential equation (1), which may be represented as a system of two first order differential equations and which describes the motion of one representative point, the algorithm (15) describes the energy interaction of two representative points. These points form a single system.

Let us study the work of the principle of the concept of symmetry by applying it to a function:

$$(16)\quad f(x) = k \cdot (x-a)^2 - c \cdot \cos(2\pi x) + b,$$

the graph of which is shown in Fig.4(a). It can be seen from Fig. 4(a) that the function under consideration has local extremums, which are located near the global extremum, which is at the point with the coordinate $x = 4$.

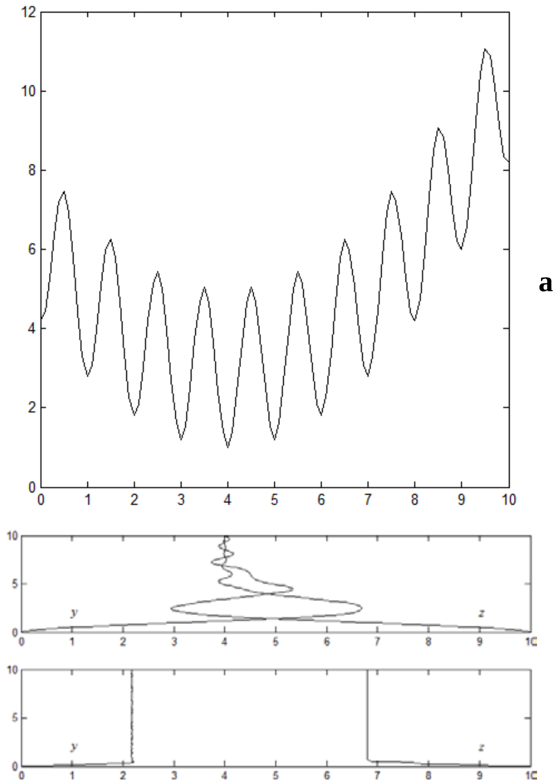

***Fig. 4.** a - the graph of the function (16); b - the process of mot of the representative points to the global extremum according to the algorithm of the heavy ball method, c – according to the gradient method.*

Using equation (4), for the function (16), we obtain the auxiliary function $F(y,z)$:

$$(17)\quad \begin{array}{l} F(y,z) = 0{,}5\Big[k(y-a)^2 - c \cdot \cos(2\pi y) + k(z-a)^2 - \\ -c \cdot \cos(2\pi z) + 2b \Big] + 0.5q(y-z)^2. \end{array}$$

The motion to the minimum of the function *f(x)* of two representative points with the initial values of the coordinates $y = 0$ and $z = 10$, belonging to the initial function (16), in accordance with the algorithm of the heavy ball method is shown in Fig. 4(b), of the gradient method - in Fig. 4(c).

To minimize the function (17) we`ll apply the following methods: coordinate descent, gradient method with constant pitch and heavy ball method. Lines of the function level (17) with the motion path of representative points to its minimum, in accordance with the algorithms of the abovementioned methods (with $k = 0.2$, $a = 4$, $b = 3$, $c = 2$, and $q = 1$) are presented in Fig. 5-7.

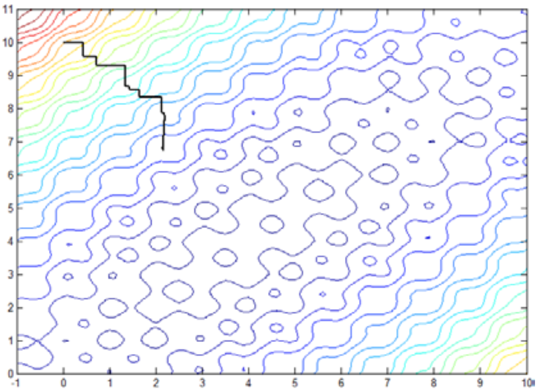

Fig. 5. *Graphical illustration of the movement to a minimum by the coordinate descent method*

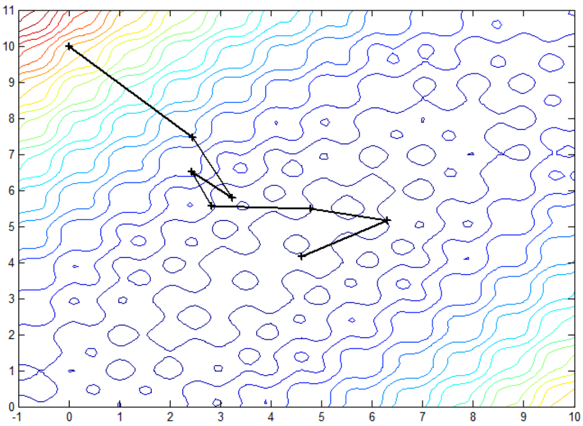

Fig. 6. *Graphical illustration of the movement to a minimum by the gradient method with constant pitch*

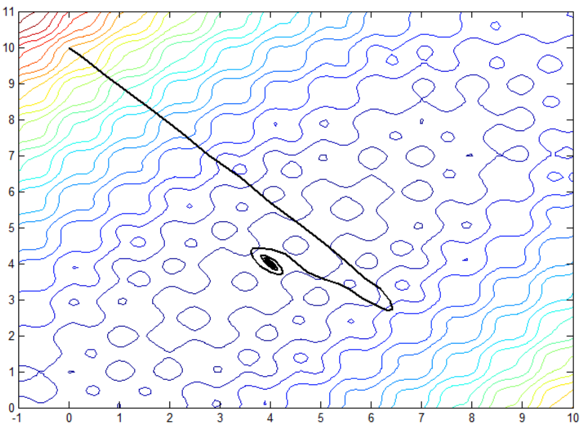

Fig. 7. *Graphical illustration of the movement to a minimum by the heavy ball method*

When the amplitude is increased in methods of coordinate descent and gradient descent with a constant pitch, the representative points end their motion in local extremums ($y = 6.8420$ and $z = 2.1491$ - coordinate descent method, $y = 4.1595$ and $z = 4.6124$ - gradient descent method with constant pitch). In the heavy ball method the process of motion of the representative points ends at the point of the global optimum $x^* = 4$. The calculated values were: $y = 4.0006$ and $z = 3.9973$, with a mass of heavy balls $m = 2.3$ and a damping factor $r = 2$.

4. Conclusion

The work of the algorithm of the modified heavy ball method, based on the principle of symmetry was investigated. The results of the research showed that for the test functions of *Ackley*, *Griewank* and Schwefel the seeking for a global extremum of a function using the usual heavy ball method does not give a positive result. However, the application of the concept of symmetry to the algorithm of the heavy ball method gave the desired result: the algorithms converge over time to the global minimum of multiextremal functions. When the amplitude of the oscillations of the multiextremal function (16) is increased, gradient methods of zero and first order are unsuitable for seeking for the global extremum of the function, since the representative points terminate at local minima. In addition, the considered algorithm of the heavy ball method with the application of the symmetry principle has good operability and high accuracy. Thus, the parallelization of the process of seeking for an extremum of a function based on the use of the concept of symmetry with regard to optimization problems will subsequently yield a number of positive results for estimating unknown parameters of objects in solving the problems of metrology of dynamic measurements and the synthesis of adaptive identification systems.

Bibliography

1. Стоян Б.Г. Решение некоторых многоэкстремальных задач методом сужающихся окрестностей / Б.Г. Стоян, В.З. Соколовский. – К.: Наук. думка, 1980. – 208 с.

2. Бахвалов Н.С. Численные методы (анализ, алгебра, обыкновенные дифференциальные уравнения) / Н.С. Бахвалов. -М.: Наука, 1973. – 632 с.

3. Васильев Ф.П. Численные методы решения экстремальных задач / Ф.П. Васильев. – М.: Наука, 1988. – 552 с.

4. Корсун В.И. Использование симметрии для распараллеливания процесса поиска экстремума целевой функции в задачах оптимального проектирования и адаптивной идентификации / В.И. Корсун// Математические модели и современные информационные технологи: сб.науч.тр. НАН Украины, Ин-т математики. – К.: 1998.– С.66-68.

5.Kharlamova Y.N. Study of the process of seeking global extremum of function by symmetric algorithms with parallel space / Y. N. Kharlamova, V.I. Korsun// Systems technologies — 2015. — № 5(100). — С. 151-161.

6. KorsunV. Paralleling of extremum seeking process in technical problems. / V. Korsun, Y. Harlamova // Science, Education and Culturein Eurasiaand Africa: The 5th International Academic Congress (23-25 March 2015, Paris, France). – “Paris University Press”, 2015. – С. 112-116. – ISBN 978-2-547-75283-5.

7. Корсун В.И. Расширение возможностей методов установления при порске глобального экстремума функции на основе концепции симметрии / В.И.Корсун, Ю.В.Жихарев, В..Л Галюта // Матеріали міжнародної конференції «Математичні проблеми технічної механіки». – Д.: ДНВП «Системні технології», – 2005. – С.160.

8. Харламова Ю.Н. Оценка эффективности симметричного алгоритма метода тяжелого шарика при поиске глобальных экстремумов тестовых функций / Ю.Н. Харламова // Гірнича електромеханіка та автоматика – 2015. — №95 – С 31-35.