



RS Global
Journals

Scholarly Publisher
RS Global Sp. z O.O.
ISNI: 0000 0004 8495 2390

Dolna 17, Warsaw, Poland 00-773
Tel: +48 226 0 227 03
Email: editorial_office@rsglobal.pl

JOURNAL	World Science
p-ISSN	2413-1032
e-ISSN	2414-6404
PUBLISHER	RS Global Sp. z O.O., Poland
ARTICLE TITLE	SEARCH FOR AN EXTREMUM USING THE STEEPEST DESCENT METHOD UNDER THE CONDITIONS OF EXPERIMENTAL ERRORS
AUTHOR(S)	Nona Otkhozoria, Vano Otkhozoria, Shorena Khorava
ARTICLE INFO	Nona Otkhozoria, Vano Otkhozoria, Shorena Khorava. (2022) Search for an Extremum Using the Steepest Descent Method Under the Conditions of Experimental Errors. World Science. 2(74). doi: 10.31435/rsglobal_ws/28022022/7785
DOI	https://doi.org/10.31435/rsglobal_ws/28022022/7785
RECEIVED	18 January 2022
ACCEPTED	19 February 2022
PUBLISHED	24 February 2022
LICENSE	 This work is licensed under a Creative Commons Attribution 4.0 International License .

© The author(s) 2022. This publication is an open access article.

SEARCH FOR AN EXTREMUM USING THE STEEPEST DESCENT METHOD UNDER THE CONDITIONS OF EXPERIMENTAL ERRORS

Nona Otkhзорia, PhD, Georgian Technical University, Tbilisi, Georgia

Vano Otkhзорia, PhD, Georgian Technical University, Tbilisi, Georgia

Shorena Khorava, PhD student, Georgian Technical University, Tbilisi, Georgia

DOI: https://doi.org/10.31435/rsglobal_ws/28022022/7785

ARTICLE INFO

Received: 18 January 2022

Accepted: 19 February 2022

Published: 24 February 2022

KEYWORDS

extremum, steepest descent method

ABSTRACT

One of the spread first level methods of optimum search is learned by the steepest descent method in conditions when there are mistakes in the experiment. The steepest descent method is investigated and is successfully applied in situations, when, there are no mistakes of experiment. However, in real situations the used means of measurement always have determined errors owing to what the appropriate meanings of the response receives with mistakes. The model of the steepest descent algorithm is created, when the length of the step does not depend on the meaning of the purpose functioning. Stepping process realization algorithm and program provision in MathCAD, computer mathematic, system is designed. The realization outcome mistakes for different meaning are presented, the step movement of the optimum dot direction is shown according to function meaning and argument meaning as well. The amount needed for the tactics necessary to approach the minimum is established, the quake amplitude in the surrounding of different level experiment mistakes at the optimum search efficiency in different step conditions.

Citation: Nona Otkhзорia, Vano Otkhзорia, Shorena Khorava. (2022) Search for an Extremum Using the Steepest Descent Method Under the Conditions of Experimental Errors. *World Science*. 2(74). doi: 10.31435/rsglobal_ws/28022022/7785

Copyright: © 2022 Nona Otkhзорia, Vano Otkhзорia, Shorena Khorava. This is an open-access article distributed under the terms of the **Creative Commons Attribution License (CC BY)**. The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Introduction. The steepest descent algorithm (in the vicinity of the value of the function argument) is an alternative procedure for the purpose of moving in the direction of the minimum from a given point to the lowest value. Such a direction is the opposite direction given by the gradient vector of the optimization function under the conditions of experimental error.

$$\Delta f(x) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right]^y.$$

The general formula of the steepest descent method for finding the argument $x(k+1)$ with value $x(k)$ on the k -step is as follows:

$$x^{k+1} = x^k + \lambda^k + x^k,$$

Where $x(k)$ is a unit of length vector at a point $x(k)$ defined in the opposite direction of the gradient.

$$S^k = - \frac{\nabla f(x^k)}{\|\nabla f(x^k)\|}.$$

$\|\nabla f(x^k)\|$ is $\nabla f(x^k)$ The length of the gradient vector

$$\|\nabla f(x^k)\| = + \sqrt{\left(\frac{\partial f(x_1^k)}{\partial x_1}\right)^2 + \left(\frac{\partial f(x_2^k)}{\partial x_2}\right)^2 + \dots + \left(\frac{\partial f(x_n^k)}{\partial x_n}\right)^2}.$$

λ^k Gradient procedure step.

If the step is constant and does not depend on the minimization function, then there will be a constant oscillation in the vicinity of the maximum, the amplitude of which will depend on the value of and the shape of the minimization function. One way to perfect a method with a constant step without the existing complication of the algorithm is to use a step whose magnitude decreases according to the iterative process, i.e. depends on the k step. Such an attitude can be expressed in several different formulas:

$$\lambda^k = \frac{a}{b + k^\beta}$$

$$\lambda^k = d \cdot b^\alpha$$

Where a, b, β , α are “positive constants”.

The algorithm for this modification reduces and suppresses the oscillation around the extremum when $k \rightarrow \infty$ but, at the same time, the algorithm has a downside if the extremum of the function has a small gradient. Then the step λ^k precedes the reduction already away from the extreme. Therefore the extreme approach can be reduced quite significantly or a much larger number of iterations may be required, making it necessary to determine the step according to the specific function [1].

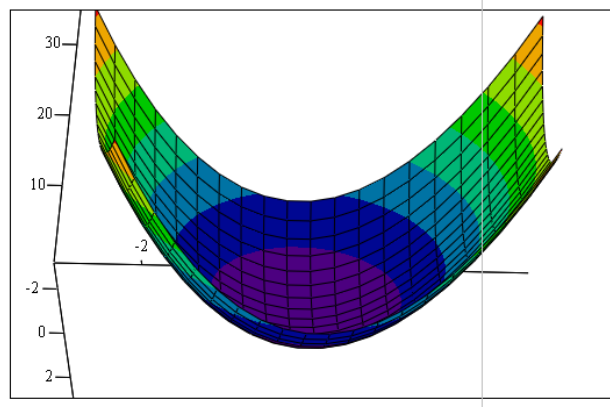


Fig.1. Concave quadratic function

Consider a concave quadratic function $f(x, y) = x^2 + \mu \cdot y^2$ (Fig. 1). Its inclination is determined by the parameter: when $\mu = 1$, the function $f(x,y)$ is a circular parabola, when $\mu > 1$ - The paraboloid becomes elliptical, stretching along the x-axis; And when $\mu < 1$ - Stretched towards the axis.

Materials and Methods. We carried out the software implementation of the fastest climbing algorithm in the mathematical software Mathcad system. To implement the algorithm, we defined the formulas and parameters to be calculated by the gradient method: $v_{max} = 20$ maximum number of iterations; $v = 0 \dots v_{max}$ - range of iteration change, $x_0 = 2$ - initial value of the argument x; $y_0 = -1$ - the initial value of the argument y; $f_0 = f(x_0, y_0)$ - the significance of the optimization function at the starting points; $\lambda_0 = 0.3$ - the initial value of the step, $q_x(x, y) = 2 \cdot x$ - with respect to the private derivative of the objective function x; $q_y(x, y) = 2\mu y$ - with respect to the private derivative of the objective function y.

Vector length $L(x, y) = \sqrt{g_x(x, y)^2 + g_y(x, y)^2}$
 $s_x(x, y)$ and $s_y(x, y)$ planes on the x,y axis in the opposite direction of the gradient vector

$$S_x(x, y) = \frac{-g_x(x, y)}{L(x, y)}; \quad S_y(x, y) = \frac{-g_y(x, y)}{L(x, y)}.$$

Step determination parameters

$$\alpha:=1; \quad \beta:=1; \quad \gamma:=0$$

$$\text{Step}(v) := \frac{\alpha \cdot \lambda_0}{\beta + \gamma \cdot v}$$

$$\lambda(x, y) := \text{Step}(v)$$

Vector initial values of the gradient procedure

$$\begin{bmatrix} x_0 \\ y_0 \\ f_0 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ f(x_0, y_0) \end{bmatrix}$$

Formula for changing vector components during a gradient procedure

$$\begin{bmatrix} x_{v+1} \\ Y_{v+1} \\ ff_{v+1} \end{bmatrix} = \begin{bmatrix} x_v + \lambda(x_v, Y_v) \cdot S_X(x_v, Y_v) \\ Y_v + \lambda(X_v, Y_v) \cdot S_X(X_v, Y_v) \\ f(x_v + \lambda(X_v, Y_v) \cdot S_X(X_v, Y_v)Y_v + \lambda(x_v, Y_v) \cdot S_Y(X_v, Y_v)) \end{bmatrix}$$

With the initial value of the parameter, self-oscillation starts from the seventh measure (Fig.2)

Based on the obtained value, we can determine how many step became necessary to approximate the minimum, With the initial value of the parameter, self-oscillation starts from the seventh step (Fig.2)

$$\sqrt{(0,122 + 0,147)^2(0,061 + 0,073)^2} = 0,301.$$

The average value of the objective function at these coordinates is equal to 0.023 (as the result shows, the step length matches the value of the amplitude $\lambda = 0.3$)

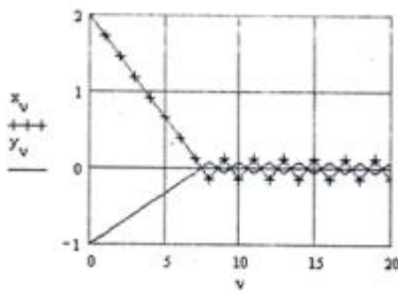


Fig. 2. Approach the minimum point

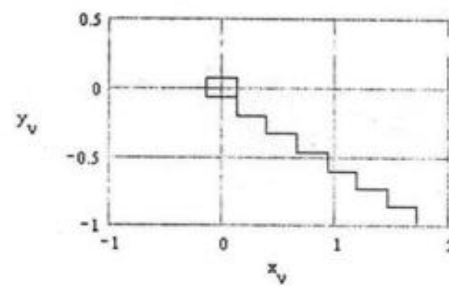


Fig. 3. Stepping towards the minimum

As can be seen from Fig. 3, at least in the vicinity there is a self-oscillation. If the paraboloid is a circular ($\mu=1$), then the amplitude with respect to both variables will be equal. Under error conditions [2] we implemented the algorithm μ different ($\mu = 0,25; 0,5; 0,75; 1; 1,25; 1,50; 1,75$) and the initial step λ ($\lambda = 0,1; 0,2$) for different values, the results obtained are given in the table

Table 1. Research Results.

$\mu=1; \lambda=0,1$					$\mu=1; \lambda=0,2$			
	Interaction N	X	Y	Mean value of the function	Interaction N	X	Y	Mean value of the function
0	11	0.032	-0.016	10.014	22	0.032	-0.016	10.0025
		-0.147	0.07			-0.057	0.029	
10%	14	-0.154	0.077	10.067	26	-0.12	0.06	9.862
		0.062	-0.031			-0.012	0.0061	
20%	14	-0.161	0.08	10.1225	26	-0.038	0.019	10.029
		0.091	-0.046			0.061	-0.03	
$\mu=1,25; \lambda=0,2$					$\mu=1,25; \lambda=0,2$			
	Interaction N	X	Y	Mean value of the function	Interaction N	X	Y	Mean value of the function
0	24	0.044	-0.011	10.0025	11	0.047	0.00011	10.0125
		-0.052	0.018			-0.153	-0.00051	
10%	27	0.05	-0.057	9.93	14	-0.152	0.055	10.057
		0.0076	0.025			0.068	-0.045	
20%	28	0.096	0.0064	9.701	14	-0.143	0.091	10.124
		0.047	0.0023			0.079	-0.085	

Continuation of table 1.

$\mu=1,5; \lambda=0,1$					$\mu=1,5; \lambda=0,2$			
	Interaction N	X	Y	Mean value of the function	Interaction N	X	Y	Mean value of the function
0	23	-0.038	0.0019	10.0025	17	0.0021	0.072	10.0167
		0.062	0.0057			-0.0016	-0.128	
10%	29	-	0.025	9.842	15	0.098	0.0092	10.146
		0.0038	-0.05			0.077	-0.0115	
20%	29	-0.039	-0.004	9.67	16	0.00334	0.034	10.23
		0.033	0.0073			-0.0098	0.166	
$\mu=1,75; \lambda=0,1$					$\mu=1,75; \lambda=0,2$			
	Interaction N	X	Y	Mean value of the function	Interaction N	X	Y	Mean value of the function
0	23	-0.019	0.00039	10.0025	16	-0.00712	-0.108	10.0175
		0.081	-0.0032			0.00043	0.092	
10%	29	-0.023	-0.031	9.84	15	0.095	0.079	10.153
		0.011	0.049			-0.00542	-0.067	
20%	27	0.032	0.00254	9.862	15	0.152	-0.019	10.292
		-0.067	-0.011			-0.0018	0.014	

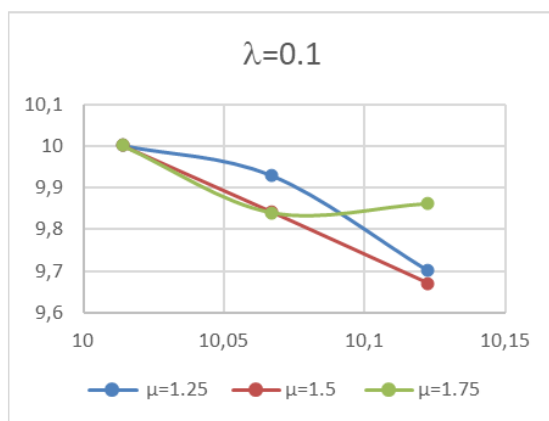


Fig. 4.

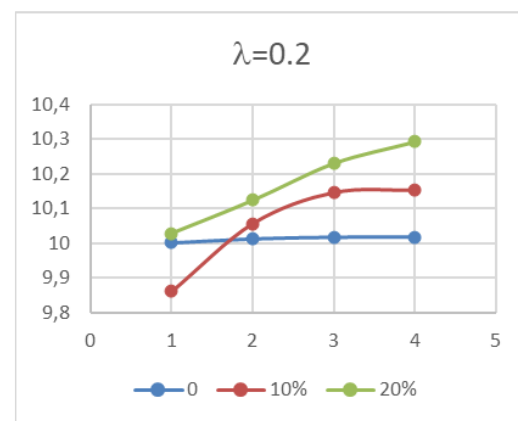


Fig. 5.

Conclusions. Graphical images are constructed according to the results of the table for visuals (as shown in Fig. 5, when the step length $\lambda = 0.2$, the effect of the errors on the value of the objective function is significant at $\xi = 20\%$. Ranges from 10.096 to 10.284 For the values $\mu = 1.75$ and $\mu = 0.75$, the distance to the extreme point averages 0.28, a relatively efficient search $\mu = 0.25$, $\mu = 1.25$, and $\mu = 1$ In this case, the mean deviation is 0.12, the reduction in efficiency is noteworthy for $\mu = 0.5$ for the value as shown in the figure. -Times are decreasing. When the step length $\lambda = 0,1$ (Fig. 4), then the largest deviation from the minimum point is equal to 0.33 and this is fixed $\mu = 0,25$; $\mu = 1,25$; And for $\mu = 1$ values. Search efficiency $\mu = 0.5$ is relatively effective; for the values of $\mu = 0.75$ and $\mu = 1.75$

Interestingly, autocorrelation under error conditions, for the value $\mu = 1$, i.e. when the function is a circular paraboloid, starts relatively quickly. However, a more effective search value than μ value has already been achieved for this iteration than μ for other values.

REFERENCES

1. Éric Walter, Numerical Methods and Optimization, Springer, Cham, Number of PagesXV, 476, Springer International Publishing Switzerland 2014.
2. Otkhozoria n, Zedhinidze I. Extreme search by simplex method under experimental error conditions Scientific-Periodical Journal "Intellect", N2, (16), 2003.