International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 1, January 2017)

The Adaptation of Random Search Algorithms

George Filatov

Professor, Doctor of Techn. Sciences, Ukrainian State University of Chemical Technology, Ukraine

Abstract **− The paper presents the algorithms of adaptation of random search: adaptation of the value of working search step and adaptation of the volume of the accumulation of information about the behavior of the optimization object. Adduced the evaluation of the direction of the descent of statistical gradient. The effectiveness of adaptation of search algorithms is illustrated on a number of numerical examples.**

Keywords **− Random Search, Adaptation of step value, Volume of Accumulation, Probabilistic Properties of Search Algorithms.**

I. Introduction

Efficiency of random search (speed, convergence) depends on the degree of adaptation. Issues of adaptation of random search devoted a significant amount of research. Almost every algorithm is a random search, offered by different researchers, is an adapted system, because it uses some form of self-learning search engine during its operation. To the theoretical basics, associated with the problem of adapting of random search, is devoted relatively little research, in particular, the monograph [1], in which has consistently set out the objectives and goals of adaptation, analytically and experimentally investigated the methods to adaptation of the value of working step, of the volume of accumulation and probabilistic properties of search algorithms. The problem of adaptation in the work 1] is formulated as a management task, which consists to maintain the functioning of the adapted object at the required level.

II. The Statement Of The Problem Of Adaptation Algorithms Of Random Search

Following [1], we'll consider the adaptation of random search, like a randomized optimization procedure for the implementation of which is not required to know the analytic properties of the optimized object. In general, this type of optimization problem is solved in form

$$Q(\mathbf{X}) \to \min_{X \in \Omega} , \qquad (1)$$

Where: $\mathbf{X}$ − the state of object;

$$\Omega = \begin{cases} G(\mathbf{X}) = 0 \\ H(\mathbf{X}) \geq 0 \end{cases}.$$

Here:

$$G(\mathbf{X}) = \{g_1(\mathbf{X}), \ldots\ldots, g_q(\mathbf{X})\} \qquad (2)$$

$G(X)$ − the stabilizing criteria:

$$g_j(\mathbf{X}) = k_{k+1}(\mathbf{X}) \quad (j = 1, \ldots., q) \qquad (3)$$

In the process of adaptation are deduced, for definiteness on the zero level, i.e. $G(\mathbf{X}) = 0$.

$H(\mathbf{X})$ − the constraint criteria:

$$H(\mathbf{X}) = \{h_1(\mathbf{X}), \ldots., h_p(\mathbf{X})\} , \qquad (4)$$

where:

$h_i(\mathbf{X}) = k_{k+q+1}(\mathbf{X}) \quad (l = 1, \ldots., p = m - k - q)$ − in the process of adaptation must exceed a given, for definiteness of the zero threshold, i.e. $H(\mathbf{X}) \geq 0$.

The problem (1) is performed recurrently:

$$\mathbf{X}_{N+1} = \mathbf{X}_N + \Delta\mathbf{X}_{N+1} , \qquad (5)$$

where the step $\Delta\mathbf{X}_{N+1}$ is determined by the *p*-stepper random search algorithm:

$$\Delta\mathbf{X}_{N+1} = B_\xi(\mathbf{X}_N, \ldots\ldots, \mathbf{X}_{N-p-1}, \mathbf{A}) , \qquad (6)$$

where B_ξ − a random search algorithm; $\mathbf{A}$ − the adapted parameters of the algorithm.

By the adaptive random search parameters are as follows:

1. Work step, or the value $|\Delta\mathbf{X}_N|$, the amount of which should change where the system of search approaching close to the target $\mathbf{X}^*$ so as to maintain the effectiveness of search on the extreme level.
2. The volume of accumulation, i.e. the number of random trials in the area $\mathbf{X}_N$ needed in order, that the step $\Delta\mathbf{X}_N$ was the best in a certain sense.
3. Distribution of random steps Ξ :

$$p(\Xi, \mathbf{U}) \qquad (7)$$

where $\mathbf{U}$ − the adapted memory vector.

If the first two factors are used for deterministic search algorithms, the latter (the parameters of the random distribution of the step) is typical only for random search that gives it an undoubted advantage over regular methods [2].

IJETAE EXPLORING RESEARCH AND INNOVATIONS

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 1, January 2017)

III. THE ADAPTATION OF THE MAGNITUDE OF WORKING STEP

The efficiency of existing algorithms of the regular and wandering random search depends on the choice of the initial step length and its adaptation in the search process. In the most of performed researches with random search algorithms or in general had not been considered the issue about the adaptation of step or had been proposed to introduce changes in stride length in response to a priori stipulated number of failed attempts. In the linear field the length of step remained constant, installed priori. In the central field, near the extremum, we decrease the length of step till a priory selected limit value. Such approach was connected with unnecessary losses on search. Therefore, we have the problem of the manage of size of step depending on the situation in which is turned out the search system. Obviously, in the search process is necessary to link the value of the step length with the kind of optimized function and with the nature of its changes on each step of the search. Maximum efficiency can be achieved if the control step size is optimal. In practical cases construct an algorithm with optimal step length at each iteration is impossible, but it is possible to construct an algorithm which is close to the optimum. For this purpose, the adaptation is introduced in the search process, which allows to adjust the step size at each iteration, depending on the information received.

Consider one of the heuristic algorithms of adaptation and test of working steps adduced in [3]. For this purpose, we consider a random search algorithm, in which the length of trial step and the working step are generated according to a certain law of distribution with known parameters. In the base of such algorithm we'll lay the idea of statistical gradient algorithm [4]:

$$\mathbf{X}_{N+1} = \mathbf{X}_N + \frac{a_N \mathbf{Y}_N}{\|\mathbf{Y}_N\|}, \qquad (8)$$

Where

$$\mathbf{Y}_N = \sum_{j=1}^{mN} \left[F\left(\mathbf{X}_N + g_N \Xi_j^N\right) - F\left(\mathbf{X}_N\right)\right] \cdot \Xi_j^N$$

Ξ_j^N – independent random vectors which is distributed, for example, uniformly over the unit sphere; g_N – the value of the trial step on the N-th iteration. We will analyze the information about the nature of the change of objective function at each step. The information obtained is then used to training the parameters of the law of distribution of the length of next steps.

Let the range of the control variables described by the vector $\mathbf{X}(x_1, x_2, ..., x_n)$, a priori can not be determined. As a possible step size distributions in such field we take logarithmical normal distribution [5] with a density:

$$p(h) = \frac{1}{\sqrt{2\pi}\sigma h} \exp\left[-\frac{(\ln h - \mu)^2}{2\sigma^2}\right], \quad h > 0 \qquad (9)$$

The length of the raffled step is defined as follows:

$$h = \exp\left(\sigma \cdot R_N + \mu\right), \qquad (10)$$

Where R_N – normalized random variable with zero expectation and unit variance.

In the presence of constraints on the range of the changing of control variables the distribution of SB Johnson can be used as a step-length distribution law [5]:

$$p(h) = \frac{\eta}{\sqrt{2\pi} h(1-h)} \exp\left[-\frac{\left\{\gamma + \eta \ln\left(\frac{h}{1-h}\right)\right\}^2}{2}\right] \qquad (11)$$

Considering that $a_N = g_N = h_N$ and assuming $\sigma = \eta^{-1}$, and $\mu = -\gamma$, the expression for determining the length of raffled step we obtained in the form:

$$h = H \frac{\exp\left[\sigma(\xi + \mu_N)\right]}{1 + \exp\left[\sigma(\xi + \mu_N)\right]} \qquad (12)$$

where H – is the maximum allowable step, ξ – normally distributed random variable with parameters.

Let in the *n*-dimensional Euclidean space E^n the unimodal continuous function is defined. This function $F(\mathbf{X})$ should be minimized. Let us set the initial parameters of distribution μ_0 and σ_0 arbitrarily or on the basis of a priori information about the behavior of the optimized function in the vicinity of the starting point of the search.

Change in average step size in the search process takes place by means of learning the distribution parameter μ. The learning is offered to make according to general rule with forgetting [6]:

$$\mu_{N=1} = \mu_N - d_N^{(1)} + d_N^{(2)} \frac{h_N^+}{h_N(0)}, \qquad (13)$$

IJETAE
EXPLORING RESEARCH AND INNOVATIONS

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 1, January 2017)

 where $d_N^{(1)} > 0$ – the coefficient of the forgetting of prehistory; $d_N^{(2)} > 0$ – the coefficient of the rate of learning ; h_N^+ – the greatest value of a series of successful test of trial steps and working steps in the transition from point $\mathbf{X}_N$ to point $\mathbf{X}_{N+1}$.

The kind of the rule (13) m is determined by the fact that the parameter μ for the law (11) is not an mathematical expectation. For the law, in which the parameter μ is the mathematical expectation, the rule in terms of learning (13) has the form:

$$\mu_{N+1} = \mu_N d_N^{(1)} + \mu_N d_N^{(2)} \frac{h_N^+}{h_n(0)}, \qquad d_N^{(1)} < 1, \quad (14)$$

In each of the proposed rules is planned to increase the value of the mathematical expectation of the length of step on a subsequent search stage, if the previous stage information reliably reflects the character of decrease of the function. On the other hand, if a series of trial steps is unrepresentative or the direction of calculated statistical gradient at the corresponding point is not close to the gradient [7], the attempt is considered as unsuccessful and the value of the mathematical expectation of step length on the subsequent step does not increase.

In the process of search is stipulated an increase of the coefficient of the rate of learning in dependence on the speed of the change of optimized function:

$$d_{N+1}^{(2)} = d_N^{(2)} + K\delta_{N,N-1}\varphi(\tilde{F}'_N), \qquad (15)$$

Where $K \geq 0$ the maximum permissible amount of change in the speed of learning on one step;

$$\delta_{N,N-1} = \begin{cases} 1, & \text{if on N-th and N-1-th stages the working steps were "successful"} \\ 0 & \text{otherwise} \end{cases}$$

$\tilde{F}'_N$ – the greatest value $\left|\frac{\Delta F_N}{h_N}\right|$ of a series of "successful" trial step and of working step on the *N*-th stage; φ – continuous function such that $0 \leq \varphi(t) < 1$ at $0 \leq t < \infty$ and $\lim_{t\to\infty} \varphi(t) = 1$.

For a smooth convex functions the function $\varphi(\mathbf{X})$ can be determined, for example, in the form $\varphi(\mathbf{X}) = 1 - \exp(-\alpha\mathbf{X})$ where α – the smallest radius of curvature of the curves obtained by all kinds different sections the function F by the two-dimensional planes, which are the orthogonal to subspace of parameters $\mathbf{X}(x_1, x_2, ..., x_n)$ and passing through the point of extreme. For the rule (12) increase in the rate of learning is as follows:

$$d_{N+1}^{(2)} = d_N^{(2)} K^{\delta_{N,N-1}\varphi(F_N)}, \quad K \geq 1. \quad (16)$$

The constructed rules of learning parameters in the formula (11) involve more than the exponential increase in the mathematical expectation step length in a linear field and when the distance to the target is much longer than the average step size. This leads to the fact that when the search system hits with a large step in the region of extremum or turns out in ravine, the search system during of certain period can not do the successful steps.

The duration of this period depends on the distance till the target out of the last point, which was found, the value of μ at this point and the parameter of forgetting $d_N^{(1)}$. In order to reduce the losses on the search for a specified period, we introduce a transitional stage of search.

The learning of parameters of the law of distribution step length in this point is described by the following rule:

$$\mu_{N+1} = \mu_N - d_N^{(1)} m, \qquad (17)$$

where: $m = N - N_0 + 1$, and N_0 – a step number, corresponding to the beginning of the transitional period.

Equation (17) is used by search system until is found such value of parameter μ, at which the system may move to a point with a smaller value of the quality criterion. After this the rule (13) is used .

For the rule (13), a transitional stage of learning of parameter μ is as follows:

$$\mu_{N+1} = \mu_N d_N^{(2)^m}, \quad d_N^{(2)} < 1. \quad (18)$$

Consider the question of the appointment of the initial parameters of the step size distribution law, as well as the question of changing the parameters of learning and forgetting on the example of distribution S_B Johnson (11). We show that the proposed algorithm of adaptation step in the simplest case can be reduced to a known system of dividing step after N of unsuccessful attempts.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 1, January 2017)

Let the maximum and minimum values of step, respectively, equal H and ε. For the algorithm of dividing step size the value

$$\varepsilon = HQ^{-s}, \quad q > 1 \qquad (19)$$

where $s > 1$ – a priori designated number of dividing of step.

Using the expression (12) with known values of H and ε, we'll choose the value of parameter σ and we'll change the interval for parameter μ out of the condition:

$$h = H\frac{\exp[\mu_k \sigma]}{1+\exp[\mu_k \sigma]}. \qquad (20)$$

Considering that the value of ε is rather small compared to unity, we obtain

$$\mu_k \sigma = -\ln \varepsilon . \qquad (21)$$

In the expressions (20) and (21) μ_k – the finite modulo the value of parameter μ, corresponding to the minimum step size. Interval for the changing of parameter μ, you can choose symmetrical with respect to zero $[-\mu_k, \mu_k]$. In practical calculations depending on the value of ε the value of μ_k can be chosen in the range from 50 to 100, but so that the value σ (21) was less than 1.

Suppose that N -th step was successful. As a result, the parameter μ receives the value $\mu = \mu_{N+1}$. Then the dependence for determining forgetting factor takes the following form:

$$qh_{[\mu_{i+1} - Nd_N^{(1)}]} = h_{[\mu_{i+1}]}$$

or

$$q\frac{\exp\left[\sigma\left(\mu_{N+1} - Nd_N^{(1)}\right)\right]}{1+\exp\left[\sigma\left(\mu_{N+1} - Nd_N^{(1)}\right)\right]} = \frac{\exp(\sigma\mu_{N+1})}{1+\exp(\sigma\mu_{N+1})} \qquad (22)$$

From (22) we obtain the value of the forgetting parameter

$$d_N^{(2)} = \frac{1}{N\sigma}\ln\left[q + \frac{q-1}{\exp(-\sigma\mu_{N+1})}\right]. \quad (23)$$

Described step adaptation algorithm is illustrated by several examples.

Search of extreme values of functions for each of the model was carried out twice: with the adaptation of step using the Johnson distribution in accordance with (21) and by the way of simple step of dividing of step (19) at $p = 10$ and $q = 2$. The calculations were performed on a computer using a standard program of receiving the initial pseudo-code, uniformly distributed on the interval [0,1]. The calculation results are shown in Table 1.

Example №1. Is searched a minimum of function

$$F(\mathbf{X}) = 1 - \exp \beta\left(-\sum_{i=1}^{n} x_i^2\right) \quad (i = 1,2,\ldots,n).$$

Are considered the cases where the number of control variables have such values: n = 2, 5 and 8. Extremum search is carried out of different points of the permissible area with the distance from the target $\rho = 50$. Parameters of search are adopted such: $H = 10;\ \varepsilon = 10^{-2};$ $d_N^{(1)}(0) = 5;\ d_N^{(2)}(0) = 5;\ k = 5;\ N = 1.$

Analysis of the results allows us to conclude that the at the calculation of the central smooth models the adaptation of step in quantitative terms does not win, although the method does not yield to a simple procedure of dividing the step.

Example №2. It is necessary to find the minimum of the function [7] $F(\mathbf{X}) = 20(x_2 - x_1)^2 + (1 - x_1)^2$. Considered function is the curved shallow ravine, on the bottom of which in the point lies the minimum. As a starting point is chosen the point with coordinates; $x_1 = -2$; $x_1 = 1$. Search parameters are adopted such: $H = 0{,}5;\ \varepsilon = 5\cdot 10^{-4};$ $d_N^{(1)}(0) = 5;\ d_N^{(2)}(0) = 5;\ k = 5;\ N = 1.$

The Table 1 shows the results of a calculation corresponding to 1000 and 2000-th trials. A further search using the algorithm of dividing step of proved to be ineffective, and the search was stopped.

This example illustrates the tangible effectiveness of the algorithm with a continuous adaptation step when using it in a gullied situations.

We show another situation where the training step length substantially reduces loss search.

Example 3. Is required minimize the function $F(\mathbf{X}) = \frac{(x_1+5)^2}{10} + x_2^2$ at the performance of restrictions:

$$x_2^2 - \left[\sin\left(\frac{x_1}{2} + \pi \right) + 1{,}05 \right]^2 \leq 0 \cdot$$

The parameters of search are adopted such: $H = 1$; $d_N^{(1)}(0) = 5$; $d_N^{(2)}(0) = 5$; $k = 5$; $N = 1$. In Fig. 1 shows the trajectory of the search system in the permissible region for the case of continuous adaptation step (20).

It is clearly seen that the system is suitable to the "canyon" with a big step, passes "canyon" with a small step, then a step is restored when you exit the "gorges" and again is reduced upon its getting into the area of the extremum. A significant difference in losses on search of extremum (Table 1) is explained by the fact that when reducing the step in the "valley", the search system without learning of step moved to the target with a small step, without the possibility of its restoration.

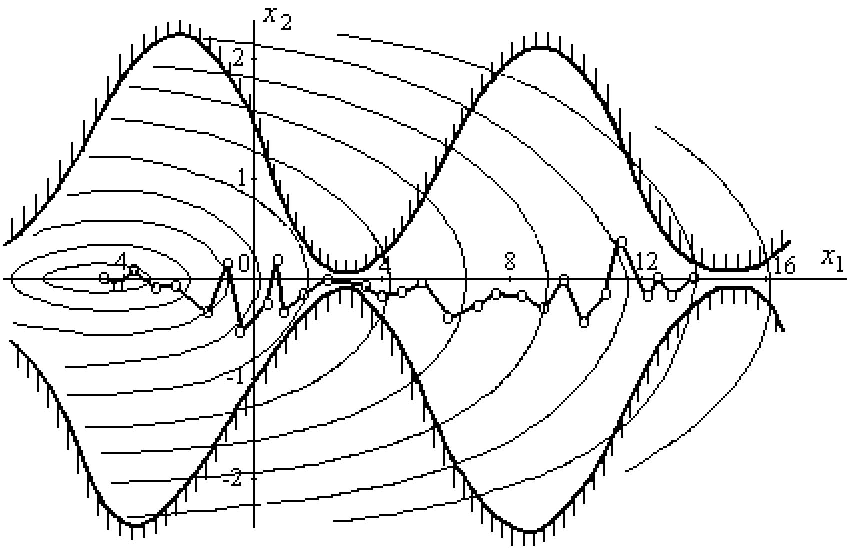

Fig. 1. The trajectory of the motion search system in the permissible region for the case of continuous adaptation of step

Table 1
Comparison of the results of applying the adaptation algorithms of step and a simple dividing of step

№ of example	The algorithm of adaptation with continuous law step length distribution				The algorithm of dividing of step			
	Number of attempts	The value of optimized function	The coordinates of point		Number of attempts	The value of optimized function	The coordinates of point	
			x_1 x_1	x_2			x_1	x_2
1	71	$5 \cdot 10^{-7}$	$3 \cdot 10^{-6}$	$7 \cdot 10^{-6}$	70	$2 \cdot 10^{-7}$	$8 \cdot 10^{-6}$	$2 \cdot 10^{-6}$
2	1000	$7 \cdot 10^{-6}$	0,74	0,54	1000	3,02	-0,72	0,56
3	2000	$3{,}5 \cdot 10^{-3}$	0,94	0,88	2000	1,48	-0,20	0,07
4	566	$2{,}2 \cdot 10^{-6}$	-5,0	$10 \cdot 10^{-4}$	2575	$2{,}8 \cdot 10^{-6}$	-4,99	$3 \cdot 10^{-4}$

IJETAE EXPLORING RESEARCH AND INNOVATIONS

International Journal of Emerging Technology and Advanced Engineering

Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 1, January 2017)

IV. Adaptation In The Process Of Accumulation Of Information About The Behavior Of The Optimization Object. Estimation Of Direction Of Descent By The Method Of Statistical Gradient

In assessing the effectiveness of random search algorithms essential importance belongs of such indicator as the loss on search. As mentioned above, step adaptation of the search process can significantly reduce the loss search. However, equally important is the selection of the optimum number of samples at each stage of research at the determining the direction of the working steps. For the solution of this problem is followed use adaptation in the process of accumulation of information about the behavior of the optimization object. Let us dwell briefly on this problem.

The theoretical justification for adaptation in the process of accumulation is contained in [1], where as the object of the study was accepted method of statistical gradient [4], which is characterized by the fact that out of the initial point of search are performed the trial random steps, and then on the basis of thus obtained information, is fulfilled the working step. The using this procedure of constructing a working search step is possible to create a search algorithm, for which the likelihood of false step less of beforehand set of arbitrarily small value. This was may be achieved in the process of the accumulation volume in the process of building a working step.

We investigate in more details the problem of evaluation of the direction of descent by the method of statistical gradient [8]. As noted in [9], in the algorithms of random search with the "accumulation" is advantageous to use a procedure of random m-gradient. Indeed, according to Theorem 1 [9] in this case, each trial step approaches the resulting direction to gradient direction. However, the using of the procedures of random gradient is associated with the need for additional calculations in the process of constructing of the random orthonormal basis at each step.

To determine the characteristics of the algorithm, which use the statistical gradient method, it is necessary to find the density distribution of the angle ψ_m between the statistical gradient and the gradient of the optimized function. With a value of ψ_m, it is easy to determine the average displacement along the gradient and the dispersion, as well as the losses on the search.

As the direction of descent for the method of statistical gradient is adopted by the weighted average direction out of m random directions. Each of these directions are selected, taking into account the quality index increments along this direction:

$$\mathbf{Y} = -\mathrm{dir}\sum_{i=1}^{m} \Xi\left[F(\mathbf{X}+g\Xi_i)-F(\mathbf{X})\right]$$

where dir denotes the unit vector defining direction: $\mathrm{dir}\mathbf{Z} = \frac{\mathbf{Z}}{|\mathbf{Z}|}$; $\mathbf{Z}$ – unit vector uniformly distributed in the space of parameters to be optimized $\mathbf{X}$; F – quality index; g – the value of trial step.

The estimation of gradient direction $\overrightarrow{\mathrm{grad}}\, F(\mathbf{X})$, which was determined thus, in general case does not coincide with the exact direction of the vector gradient of optimized function: $\mathrm{dir}\,\overrightarrow{\mathrm{grad}}\, F(\mathbf{X}) \neq \mathrm{dir}\,\mathrm{grad}\, F(\mathbf{X})$. The angle between the direction of the gradient $\mathrm{grad}\, F(\mathbf{X})$ and the statistical assessment $\overrightarrow{\mathrm{grad}}\, F(\mathbf{X})$ is a random variable, depending on the number of samples m and the search space dimension N.

In this paper, for the linear objective function is analyzed the dynamics of the gradient vector estimation depending on m and N, are adduced the approximate expressions for the distribution of densities of the angle ψ_m, allowing to determine the losses on the search under a given deviation from the direction of descent from the gradient direction and under given losses on search to evaluate the deviation of chosen direction from the gradient direction.

As noted in the monograph [6], the probability density function of the angle ψ_m analytically very difficult to obtain. In this regard, here these distribution function were obtained in numerical form on the computer and for them was selected the approximating expression. For the density distribution of angle ψ_m the choice was performed over the asymmetry and with account of the presence of acute vertex of distribution using the plane Pearson (β_1, β_2) [10]:

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}};\quad \beta_2 = \frac{\mu_4}{\mu_2^2};$$

$$\mu_k = \int_0^{\pi/2} (x-\mu_1) f(\psi_m) d\psi_m ,$$

where $f(\psi_m)$ – the function of the density distribution of angle.

In the result of calculations is turned out, that the function of the angle density distribution for a variety of species distribution can be described [11]:

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 1, January 2017)

$$\varphi(x) = \begin{cases} \dfrac{1}{\sigma\sqrt{2\pi}} \left\{ \exp\left[\dfrac{(x-\bar{x})^2}{2\sigma^2} \right] + \exp\left[\dfrac{(x+\bar{x})^2}{2\sigma^2} \right] \right\}, & \text{if} \quad x < 0 \\ & \text{if} \quad x \geq 0 \end{cases}$$

At this the mathematical expectation and variance of the angle ψ_m are described, respectively, by functions of the form (Fig. 2, 3.):

$$\mu(m,N) = A(N)\exp[-B(N)m]; \quad D(N,m) = A_1 \exp[-B_1(N)m]m^{P(N)} \quad , \qquad (24)$$

where:

$$3 \leq N \leq 20; \quad 2 \leq m \leq 1{,}5N;$$
$$A(N) = -1{,}029N^{-4{,}341} + 0{,}021N + 1{,}022;$$
$$B(N) = 1{,}099\exp(-0{,}958N - 0{,}004N + 0{,}068);$$
$$A_1(N) = 1{,}07\exp(-0{,}708N) - 0{,}001N + 0{,}041;$$
$$B_1(N) = 0{,}99\exp(-1{,}001N) + 0{,}251N^{0{,}109} - 0{,}228;$$
$$P(N) = -1{,}07N^{-4{,}425} + 0{,}059N + 0{,}030.$$

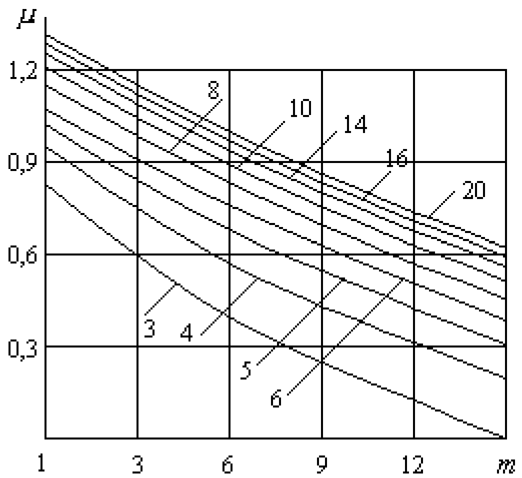

Fig. 2. The graphs of the dependence of $\mu(N)$ for different values m.

The numbers on the curves – values N

The distribution parameters $\bar{x}, \sigma$ (24) are determined out of solving the system of equations:

$$\left. \begin{aligned} \mu(m,N) &= 2[\bar{x}h_0(\bar{x}/\sigma) + |\,\sigma\varphi(\bar{x}/\sigma)]; \\ D(m,N) &= \bar{x}^2 + \sigma^2 - \mu(m,N), \end{aligned} \right\}$$

where: $h_0(t) = \dfrac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{x^2}{2}} dx$; $\varphi(t) = \dfrac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}}$.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 1, January 2017)

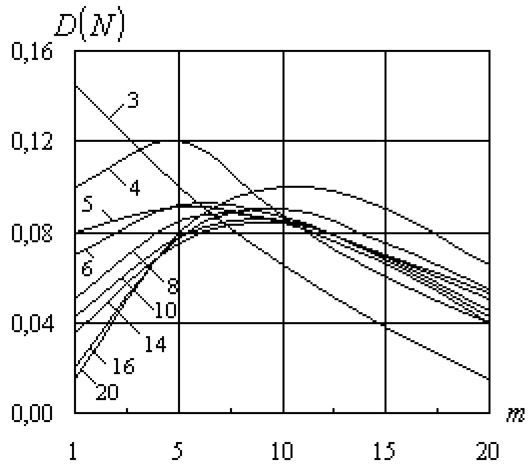

Fig. 3. The graphs of the dependence of $D(N)$ for different values m.

The numbers on the curves – values N

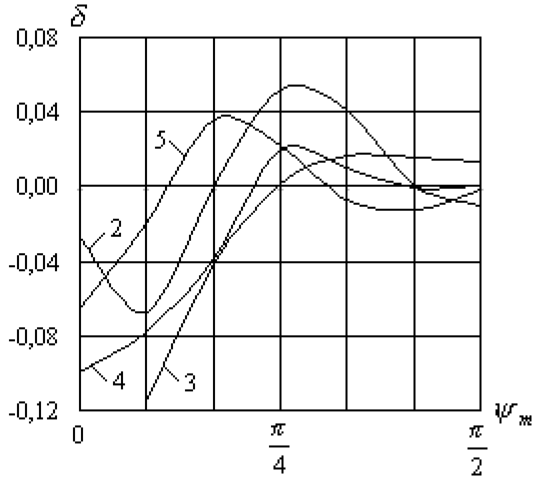

Fig. 4. The graphs of the dependence of $\delta(\psi)$ for different values m.

The numbers on the curves – values m

The accuracy of approximation was estimated by the value of the relative error (Fig. 4.):

$$\delta(\psi)=\frac{\int_0^{\psi} d[F(\varphi)-h(\varphi)]}{\int_0^{\psi} dF(\varphi)}=\frac{[F(\psi)-h(\psi)]}{F(\psi)},$$

 where $F(\psi)$ – the empirical and $h(\psi)$ – the theoretical distribution function angle ψ. At the same time it is easy to select a rectangular area on the plane Pearson:

$$\left(\sqrt{\beta_1}-\sqrt{\beta_2};\;\; \beta_2-\Delta\beta_2; \sqrt{\beta_1}+\sqrt{\beta_2};\;\; \beta_2+\Delta\beta_2\right)\qquad .$$

IJETAE
EXPLORING RESEARCH AND INNOVATIONS

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 1, January 2017)

in which the error

$$\left|\delta(\psi)\right| \le \bar{\delta}\,,\ \text{where: } \Delta\sqrt{\beta_1} \le \frac{\bar{\delta}(\mu_3)}{\mu_3^{3/2}};\ \ \Delta\beta_2 \le \bar{\delta}\,\frac{\mu_4}{\mu_2^2}\,.$$

In the next article, we present several algorithms of pseudo-gradient class, which realizes the idea of accumulation of information.

REFERENCES

[1] Rastrigin L.A., Ripa K.K., Tarasenko G.S. Adaptation of random search. − Riga, Zinatne, 1978. □ 243 p..

[2] Rastrigin L.A. Random search -□ specifics, steps, history and prejudices // In .: Questions of kibernetik. Problems random search.- M .: Science. - 1978. □ P.3-□17.

[3] Volynskij E.I., Filatov G.V. On the question of adaptation step in the algorithms of random search // Automation and Computer tehnika.□ 1974.□ № 1. □ p.66-□70.

[4] Rastrigin L.A. Statistical methods of search.− M .: Science, 1968.□ 376 p.

[5] Khas'minskii R.3. Application of Random Noise to Optimization and identification // Coll. "Problems of Information Transmission". □ № 3. □ 1965. □ p.49-□55.

[6] Rastrigin L.A. Random search with linear tactics. □ Riga Zinatne, 1971. − 310 p.

[7] Zipkin 3.Y. ,Polyak B.T. Pseudo-gradient adaptation and learning algorithms // Automation and telemehanika. − 1973. □ №3. □ p.48-□56.

[8] Volynskij E.I., Malyshev S.A., Filatov G.V. About the estimate of the direction of descent by the method of statistical gradient// Automation and Computer Science. − Riga: Zinatne. □1976. □ №6. □ p.53□-55.

[9] Nikolaev E.G. Random and gradient selection of the descent direction in the problem of finding a minimum of a function of many variables. Kand. Dis. on soisk. scientists. PhD degree. Sci. Sciences. M., 1971. − 106 p.

[10] Han G., Shapiro S. Statistical models in engineering problems. − M .: Mir, 1969. − 395p.

[11] Leone F. C., Nelson L. S. and Nottingham R. B. The folded normal distribution. □ Technometrics, 1961, − No. 3. −p. 543.