



PROBABILISTIC ANALYSIS OF THE CONVERGENCE OF THE DIFFERENTIAL EVOLUTION ALGORITHM

R. Knobloch, J. Mlýnek**

Abstract: Differential evolution algorithms represent an efficient framework to tackle complicated optimization problems with many variables and involved constraints. Nevertheless, the classic differential evolution algorithms in general do not ensure the convergence to the global minimum of the cost function. Therefore, the authors of the article designed a modification of these algorithms that guarantees the global convergence in the asymptotic and probabilistic sense. The modification consists in adding a certain ratio of random individuals to each generation formed by the algorithm. The random individuals limit the premature convergence to the local minimum and contribute to more thorough exploration of the search space. This article concentrates specifically on the role of random individuals in the identification of the global minimum of the cost function. Besides, the paper also contains some useful estimates of the probability of finding the global minimum of the corresponding cost function.

Key words: *optimization, differential evolution algorithm, random sampling of the search space, probabilistic analysis, binomial distribution*

Received: November 26, 2019

DOI: 10.14311/NNW.2020.30.017

Revised and accepted: August 30, 2020

1. Introduction

Optimization tasks represent a principally important part of application mathematics. The optimization algorithms are frequently used in science, economy, engineering, and production practice. In general, to optimize a system means to maximize the system's desirable properties minimizing at the same time its unfavourable characteristics. For an example of the optimization task in the automotive industry see articles [12], and [13]. In these articles the optimization of the infrared heaters setting over the complex shell metal mould is realized by a differential evolution algorithm with good results.

The differential evolution algorithms represent an efficient framework to solve complicated optimization tasks with many variables and diverse constraints. These algorithms are sufficiently universal and versatile and provide good results even

*Roman Knobloch – Corresponding author; Jaroslav Mlýnek; Technical University of Liberec, Studentská 1402/2, CZ-461 17 Liberec 1, Czech Republic, E-mail: roman.knobloch@tul.cz, jaroslav.mlynek@tul.cz

with problems characterized by non-differentiable cost functions with many local minima. The conventional gradient optimization methods do not usually provide in these cases satisfactory results either because they use the derivatives of the cost function (that need not be available) or since they end up at a local minimum. Furthermore, their results may strongly depend on the point from which the search of the cost function minimum starts.

The topic of efficient utilization of the differential evolution algorithms in optimization tasks is treated in a row of articles. The articles [1, 4–7], and [14] deal with possible modifications of the differential evolution algorithms focused on limiting the stagnation of the algorithms and convergence acceleration to the global minimum of the cost function. Simultaneously, some articles concentrate on the division of the optimization process into several local optimization tasks which are subsequently solved using the differential evolution algorithms (see articles [15] and [20]). The topic of combining the differential evolution algorithms with another optimization or approximation technique is studied in articles [3] and [19]. The topics of convergence of the differential evolution algorithm itself are handled in article [8].

The authors of the current article demonstrated (see [9]) that even the differential evolution algorithms as introduced in the original works of Price, Storn, and Lampien [16], [18] do not guarantee in general the convergence to the global minimum of the cost function. To be specific we focus on the classic differential evolution algorithm *DE/rand/1/bin* (further referenced to as CDEA) but our remarks apply mostly to all alternatives of the original differential evolution algorithms. The principal problem with this algorithm is its too fast convergence to the local minimum of the cost function in some cases. This phenomenon is usually called premature convergence. The premature convergence consists in the rapid reduction of the size of generations formed by the algorithm around a local minimum excluding in this way the possibility to attain a better local minimum or the global minimum of the cost function.

This was the reason why we proposed a modification of the CDEA that ensures the global convergence in the asymptotic sense based on relatively general assumptions imposed on the cost function. The modification consists in replacing some individuals (with higher values of the cost function) in the current generation by fully random individuals. The modified differential evolution algorithm was denoted as MDEA. For the algorithm MDEA it was possible to prove the global convergence to the minimum of the cost function in the asymptotic sense. For details see the article [9]. In article [10] we further investigated the properties of MDEA, primarily its convergence speed.

In the current paper we concentrate more closely on the role of random individuals in the convergence process. The random individuals limit the premature convergence and provide more thorough exploration of the search space. In reality, they usually slow down partially the convergence speed of the algorithm, but they provide much better chance to identify the global minimum of the cost function.

2. Optimization task

Let us have an optimization task to find a minimum of a cost function $f(x_1, x_2, \dots, x_n)$ of n real variables defined on a bounded domain in the Euclidean space E_n . For brevity we denote by the symbol x the ordered n -tuple of all variables x_1, x_2, \dots, x_n of the optimized function f . That is, we should find the minimum of the function $f(x)$. In principle, this function may have many local and global minima. We would like to find the minimum with the lowest cost function value

$$\min\{f(x) : x \in S\}, \quad (1)$$

where $S \subset E_n$ is a measurable search space of a finite measure representing all possible configurations of the variables x . The solution set can be defined as

$$S^* = \{x^* : f(x^*) = \min\{f(x) : x \in S\}\}, \quad (2)$$

where x^* represents the global minimum of the cost function.

To be able to find an approximate solution of the optimization task (1), we introduce a concept of an expanded solution set

$$S_\varepsilon^* = \{x \in S : |f(x) - f(x^*)| < \varepsilon\}, \quad (3)$$

where $\varepsilon > 0$ is a small positive real number. Denoting by μ the Lebesgue measure, we suppose that for each ε it holds $\mu(S_\varepsilon^*) > 0$.

In the following part we concentrate on finding an $x^* \in S^*$ defined by the relation (2) or at least on finding an approximate solution of the task (1) $x \in S_\varepsilon^*$ defined by the relation (3) for a sufficiently small value $\varepsilon > 0$. We will solve the task (1) utilizing the differential evolution algorithm.

3. Differential evolution algorithms

In this part we briefly describe the functioning of CDEA and MDEA.

3.1 Classic differential evolution algorithm

In general, CDEA seeks for the minimum of the cost function by constructing whole generations of potential solutions. These solutions are usually called individuals. Each individual is an ordered set of specific values from the cost function domain. The quality of the individual is determined by the evaluation of the cost function. Further we refer to the cost function domain as the search space.

The next generation is formed from the existing generation by means of mutation and crossover operators. Specifically, we go successively through all individuals in the generation G . To each individual y_i^G (termed as the *target individual*) we select randomly three other (different) individuals $y_{r1}^G, y_{r2}^G, y_{r3}^G$ from the current generation. We form in a specific way (including randomness) a combination of these three individuals and the target individual. This combination is termed the *trial individual* and denoted y_i^{trial} . Then we evaluate the cost function for the target y_i^G and trial individual y_i^{trial} and compare the results. The individual with lower value of the cost function advances to the position of the target individual of the

next generation y_i^{G+1} . When this procedure is completed for all target individuals in generation G , we have the new generation of individuals numbered $G + 1$.

The next part illustrates CDEA operation more specifically in the form of the pseudocode.

Input: Optimization task parameters:

f denotes the cost function, D is the dimension of the cost function domain, $\langle x_{i \min}, x_{i \max} \rangle$ is a domain of each cost function variable x_i .

CDEA parameters:

NP denotes the generation size (the number of individuals in each generation), NG is the total number of generations, F stands for the mutation factor, $F \in \langle 0, 2 \rangle$, and CR denotes the crossover probability, $CR \in \langle 0, 1 \rangle$. The symbol G stands for the generation number, index i is the number of the individual in a specific generation, index j describes the j -th component of a specific individual y_i .

Computation:

1. create the initial generation ($G = 1$) of NP individuals y_i^G , $1 \leq i \leq NP$, randomly or according to a prescribed scheme
 2. (a) evaluate all individuals y_i^G of the generation G (calculate $f(y_i^G)$ for each individual y_i^G)
 - (b) store the individuals y_i^G and their evaluations $f(y_i^G)$ into the i -th row of matrix \mathbf{A} with NP rows and $D + 1$ columns
 3. **repeat until** $G \leq NG$
 - (a) **for** $i = 1$ **to** NP **do**
 - i. randomly select three different indices $r_1, r_2, r_3 \in \{1, 2, \dots, NP\}$, $r_m \neq i, m \in \{1, 2, 3\}$
 - ii. randomly select an index $k_i \in \{1, \dots, D\}$
 - iii. **for** $j = 1$ **to** D **do**
 - if** ($\text{rand}(0, 1) \leq CR$ **or** $j = k_i$)
 - then** $y_{i,j}^{\text{trial}} = y_{r_3,j}^G + F(y_{r_1,j}^G - y_{r_2,j}^G)$
 - else** $y_{i,j}^{\text{trial}} = y_{i,j}^G$
 - endif**
 - endfor(j)**
 - iv. **if** $f(y_i^{\text{trial}}) \leq f(y_i^G)$
 - then** $y_i^{G+1} = y_i^{\text{trial}}$
 - else** $y_i^{G+1} = y_i^G$
 - endif**
 - endfor(i)**
 - (b) store the individuals y_i^{G+1} and their evaluations $f(y_i^{G+1}), 1 \leq i \leq NP$, of the new generation $G + 1$ into the matrix \mathbf{A} , set $G = G + 1$
- endrepeat.**

Output:

The matrix \mathbf{A} with NP rows and $D + 1$ columns contains the final generation of individuals including their evaluations. The row of matrix \mathbf{A} that contains the lowest cost function value represents the best found individual y_{\min} .

3.2 Modified differential evolution algorithm

In this part we describe the modification of CDEA denoted by MDEA. Since there is in principle necessary to make one simple change in the algorithm, we present only the differences with respect to CDEA.

Input: We add another parameter R that determines the ratio of random individuals in each generation, $R \in \langle 0, 1 \rangle$, e.g., $R = 0.1$ means that 10% of individuals in each generation are generated randomly.

Computation: We add another procedure to the part (3), specifically:

(c) determine in matrix \mathbf{A} the quantity $\lfloor NP \cdot R \rfloor$ of individuals with the highest cost function values and replace these individuals by random individuals $(y_{\text{rand}})_i$ from the search space, $1 \leq i \leq \lfloor NP \cdot R \rfloor$, including their evaluations. Here the symbol $\lfloor x \rfloor$ denotes the integer part of the real number x . This straightforward modification makes possible to prove the asymptotic convergence in probability. For more details see the article [11].

3.3 Convergence in probability

To examine the global convergence of MDEA we need to introduce a concept of the convergence in probability introduced in article [21].

Definition: Let $\{Y(G) : G = 1, 2, \dots\}$ be a generation sequence formed by a differential evolution algorithm to solve the optimization problem (1). We say that the algorithm converges to the optimal solution set in probability if and only if

$$\lim_{G \rightarrow \infty} p\{Y(G) \cap S_\varepsilon^* \neq \emptyset\} = 1, \quad (4)$$

where p denotes the probability of an event. Now we can prove the following theorem.

Theorem 1. *Let us suppose that for each population $Y(G)$ of a differential evolution algorithm there exists at least one individual y such that*

$$p\{y \in S_\varepsilon^*\} \geq \alpha > 0,$$

where α is a small positive value. Then the algorithm converges to the optimal solution set S_ε^ in probability. That is the relation (4) holds.*

Here $p\{y \in S_\varepsilon^*\}$ denotes the probability that y belongs to the optimal solution set S_ε^* . For the proof of the theorem and relevant details see the article [11].

4. Sampling of the search space by random individuals

The operations of algorithms CDEA and MDEA are relatively straightforward to describe. For details see articles [9] or alternatively [10]. On the other hand their exact theoretical analysis is relatively demanding and up to now not available in any publications. Nevertheless, the analysis of the role of random individuals in MDEA is relatively simple. By a random individual we mean here a random point from the search space S . Since these random individuals are the part of the algorithm that ensures the asymptotic convergence to the global minimum of the cost function in probability, it definitely has sense to have a clear idea about how this feature of MDEA works. That is the reason we focus in this part on the mechanism how random individuals contribute to the identification of the global minimum of the cost function.

From the description of CDEA and MDEA (see subsection 3.1 pseudocode part (3)(a)(iv)), it is apparent that both algorithms are greedy in the following sense: When they attain a point in the search space with a small value of the cost function they will not lose it unless they replace it by another point with even smaller value of the cost function.

This feature complements conveniently with the random sampling part in the MDEA (see subsection 3.2 pseudocode part (3)(c)). The random sampling explores the search space and can be characterized by the fact that more random individuals provide more detailed exploration of the search space. Technically speaking, when performing a practical calculation, the random sampling can bring the MDEA close to the global or very low (acceptable) local minimum and the mechanism of the differential evolution ensures the effective convergence to this minimum.

4.1 Probabilistic convergence analysis

Let us suppose we have identified before a local minimum of the cost function f at the point x_L with the cost function value $f(x_L)$ (index L stands here for Local). We would like to try the possibility to find a better local minimum or preferably the global minimum of the cost function f .

We formulate a null hypothesis H_0 that there exists a part of the search space S with cost function values lower than $f(x_L)$. We denote this part of the search space S as the target region σ

$$\sigma = \{x \in S : f(x) < f(x_L)\}. \quad (5)$$

We denote the measure of the target region σ as $\mu(\sigma)$. Additionally, we suppose $\mu(\sigma) > 0$. By the symbol $\mu(S)$ we understand the measure of the whole search space S . The target region σ is thus the part of the search space S that according to our assumption contains a minimum or minima with lower cost function values than $f(x_L)$.

Subsequently, we use the algorithm MDEA that includes the generation of random individuals, that is a random sampling. The probability p_0 that we hit the target region σ with one random individual is given by the ratio of $\mu(\sigma)$ and $\mu(S)$

$$p_0 = \frac{\mu(\sigma)}{\mu(S)} \quad (6)$$

according to the usual concept of geometrical probability introduced for instance in book [2].

The hypothesis H_0 can be expressed symbolically as

$$H_0 : p = p_0, \quad (7)$$

where p denotes the probability to hit the target region σ with one random individual. In principle, we can confirm or reject the hypothesis H_0 by generating a quantity of random individuals and monitoring whether some of them hit the target region σ or not.

The number k of random individuals that should hit the target region σ is described by the binomial distribution (for details see [2], page 140)

$$Bi(n, k, p_0) = \binom{n}{k} p_0^k (1 - p_0)^{n-k}. \quad (8)$$

After we generated n random individuals there exist two possible principally different results:

1. We got some individuals from the target region σ . This means we have found some points x_B with cost function values $f(x_B)$ lower than $f(x_L)$ (index B here stands for Better). This result is considered positive and we can decide whether the lowest value $f(x_B)$ is acceptable or whether to continue with the search for even better cost function values.
2. No random individuals hit the target region σ , that is we did not find any points x_B . What conclusion can we make based on this negative result? There are in principle three possible alternatives:
 - The target region σ does not exist at all.
 - The target region σ does exist but its measure $\mu(\sigma)$ is smaller than we have supposed. This implies that the probability to hit the target region σ with one random individual is smaller than p_0 .
 - The target region σ does exist and the estimate of its measure $\mu(\sigma)$ is in principle right but we did not hit it accidentally since we used exclusively random individuals that missed the target region.

The probability that the target region σ should not be hit even once after generating n random individuals is according to (8)

$$Bi(n, 0, p_0) = (1 - p_0)^n. \quad (9)$$

It is obvious that this probability converges to 0 with the increasing n for $p_0 > 0$. Even if the random individuals indicate that the probability p to hit the target region σ is close to 0 or at least significantly smaller than the value p_0 we do not know this for sure just because the generated individuals are random. We can claim this only with a relative certainty.

At this point, it is suitable to utilize some terminology from the area of statistical hypotheses testing. It is necessary to set a fixed significance level α_S . Quantity α_S represents the probability that we decline the hypothesis H_0 although it is true. Symbolically,

$$\alpha_S = P(H_0 \text{ true, but declined}), \tag{10}$$

where P denotes the probability of the event.

The value of α_S depends on circumstances and in particular on the fact what quantity of risk to reject the right hypothesis is for us acceptable. It is usually chosen 0.05, 0.01, 0.005, ... but in fact it can be set arbitrarily. The significance level α_S is sometimes termed as the error of the first kind in statistical hypotheses testing.

The logic of the hypotheses testing is that we determine some part of the binomial distribution $Bi(n, k, p_0)$ with a small probability α_S and declare this part of the distribution as critical. Since the probability α_S is small it is relatively improbable that we hit the critical part. In our case the critical part corresponds to the "no hit of the target region σ " area that is to the result $Bi(n, 0, p_0)$. Since the value $Bi(n, 0, p_0)$ is according to (9) decreasing with increasing n and $p_0 > 0$ we have to take the number n in such a way that the result $Bi(n, 0, p_0)$ is a part of the binomial distribution $Bi(n, k, p_0)$ with probability equal to or less than α_S .

Now, it is possible to determine the smallest such n that we denote by n_0 . Number n_0 can be calculated from the equation

$$(1 - p_0)^{n_0} = \alpha_S$$

as the minimal number guaranteeing to attain the significance level α_S which gives

$$n_0 = \left\lceil \frac{\log(\alpha_S)}{\log(1 - p_0)} \right\rceil, \tag{11}$$

where $\log(x)$ stands here for the decadic logarithm of the positive real number x and the delimiters $\lceil \cdot \rceil$ stand for the upper integer part of the enclosed real number.

From the considerations above it follows that if we generate $n \geq n_0$ random individuals and none of them hits the target region σ we can claim that the probability p to hit the target region σ is $p < p_0$ with the risk α_S at most. We finally introduce the concept of the relative certainty C_R by the relation

$$C_R = 1 - \alpha_S. \tag{12}$$

Now, we can formulate the following proposition.

Proposition 2. *Let us assume that we generate $n \geq n_0$ random individuals, where n_0 is defined by the relation (11), and not even one hits the target region σ . Then we can claim that with relative certainty $C_R \geq 1 - \alpha_S$ the actual probability p to hit the target region with one random individual is less than p_0 . This implies that the measure of the target region is smaller than $\mu(\sigma)$ with the same relative certainty C_R .*

Proof. The proof follows easily from the meaning of the significance level α_S and the value n_0 given by the relation (11). Suppose we generated $n \geq n_0$ random individuals and not even one hits the target region σ . Since it definitely holds

$$n \geq n_0 = \frac{\log(\alpha_S)}{\log(1 - p_0)},$$

we have

$$n \log(1 - p_0) \leq n_0 \log(1 - p_0) = \log(\alpha_S),$$

which implies

$$(1 - p_0)^n \leq (1 - p_0)^{n_0} = \alpha_S.$$

The last relation expresses the fact that when the hypothesis H_0 is true then the probability that no random individual out of n hits the target region σ is smaller than α_S . Since the value α_S is relatively small we declare this situation as relatively improbable and accept the risk α_S to decline the hypothesis H_0 with a relative certainty $C_R = 1 - \alpha_S$. In other words we say that since this situation is on the assumption $p = p_0$ relatively improbable its occurrence is caused rather by the fact that $p < p_0$. \square

Remark. It is apparent from the relation (11) that the number n_0 is dependent on the significance level α_S and on the assumed probability p_0 .

4.2 More probabilistic estimates

In spite of the fact that the formula (11) is relatively simple, it can be further simplified. Let us introduce an auxiliary quantity

$$\psi = np_0 \tag{13}$$

that expresses the relation between the number n of generated random individuals and the assumed probability to hit the target region p_0 defined by (6). Now, we can express the probability that not even one random individual out of n hits the target region σ in the following way

$$(1 - p_0)^n = (1 - p_0)^{\frac{\psi}{p_0}} = \left[(1 - p_0)^{\frac{1}{p_0}} \right]^\psi.$$

When we again make use of the significance level α_S defined by the relation (10) we can write

$$\left[(1 - p_0)^{\frac{1}{p_0}} \right]^\psi \leq \alpha_S.$$

Introducing the quantity ψ_0 as a minimal value of ψ complying with the previous inequality we get and equation

$$\left[(1 - p_0)^{\frac{1}{p_0}} \right]^{\psi_0} = \alpha_S.$$

We need to estimate in a suitable way the term in the square brackets. The task can be reformulated in the following way: Find the supreme of the function

$$g : y = (1 - x)^{\frac{1}{x}}$$

with the domain $D(g) = (0, 1)$. Elementary analysis gives us the estimation (see for instance [17], page 63)

$$(1 - p_0)^{\frac{1}{p_0}} \leq \frac{1}{e}.$$

Using this estimate gives

$$\left(\frac{1}{e}\right)^{\psi_0} = \alpha_S.$$

From this relation we immediately get

$$\psi_0 = -\ln(\alpha_S), \tag{14}$$

where $\ln(x)$ stands for the natural logarithm of the positive real number x . The quantity ψ_0 gives us a useful representation for the term

$$\psi_0 = N_0 p_0.$$

The last relation expresses the fact that when we assume the hypothesis $H_0 : p = p_0$ and no random individual hits the target region σ we need to take N_0 in such a way that

$$N_0 p_0 \geq -\ln(\alpha_S)$$

to be entitled to claim that $H_0 : p = p_0$ does not hold implying $p < p_0$. In case we need explicitly to express the number of individuals not hitting the target region σ we can use the formula (13) and get

$$N_0 = \left\lceil -\frac{1}{p_0} \ln(\alpha_S) \right\rceil. \tag{15}$$

Analogously to Proposition 2 we can state the following:

Proposition 3. *Let us assume we generate $n \geq N_0$ random individuals where N_0 is defined by the formula (15) and not even one individual hits the target region σ . Then we can claim with the relative certainty $C_R \geq 1 - \alpha_S$ that the actual probability p to hit the target region is less than p_0 . This implies that the measure of the target region is smaller than $\mu(\sigma)$ with the same relative certainty C_R .*

Proof. The proof follows from the preceding considerations. □

It is not complicated to verify that the formulas (11) and (15) for the numbers n_0 and N_0 are equivalent under the condition $p_0 \ll 1$.

Lemma 4. *The formulas (11) and (15) are equivalent on the assumption $p_0 \ll 1$.*

Proof. The following formula holds for logarithms with different bases

$$\log_a(x) \cdot \log_b(a) = \log_b(x).$$

Using this formula in our context gives

$$\log_{10}(\alpha_S) \cdot \ln(10) = \ln(\alpha_S),$$

where $\ln(x)$ denotes the natural logarithm of a positive real number x . Substituting the last formula into the (11) gives

$$n_0 = \frac{\log(\alpha_S)}{\log(1-p_0)} = \frac{\ln(\alpha_S)}{\log(1-p_0) \cdot \ln(10)} = \frac{\ln(\alpha_S)}{\ln(1-p_0)}.$$

When we suppose $p_0 \ll 1$ we can approximate using the Taylor's expansion (for details see [17], page 110)

$$\ln(1-p_0) \approx -p_0.$$

Using the last approximation gives

$$n_0 \approx \frac{\ln(\alpha_S)}{-p_0} = N_0,$$

which was to prove. □

It is interesting to note that the Proposition 2 and its equivalent Proposition 3 can be reformulated in another rather unexpected way.

Corollary. *Let us assume that the probability p to hit the target region σ with one random individual is equal to p_0 . When we consider n random individuals we denote by P_1 the probability to hit the target region σ with at least one individual. The probability P_1 is obviously*

$$P_1 = 1 - (1 - p_0)^n.$$

When none of the n random individuals hits the target region σ , we can claim with relative certainty $C_R = P_1$ that the probability $p < p_0$ implying that the measure of the target region is smaller than $\mu(\sigma)$.

Proof. The probability P_0 that none out of n random individuals hits the target region σ is apparently

$$P_0 = (1 - p_0)^n.$$

The probability P_1 is obviously given by

$$P_1 = 1 - P_0 = 1 - (1 - p_0)^n.$$

But by considerations in Proposition 2 it holds $(1 - p_0)^n = \alpha_S$ and the relative certainty C_R is given by $C_R = 1 - \alpha_S$ implying

$$C_R = P_1.$$

□

5. Utilization of the probabilistic estimates of the MDEA convergence

In general, using MDEA for an optimization task provides two principle advantages compared to utilization of CDEA.

1. From the theoretical point of view, according to Theorem 1 the asymptotic convergence to the global minimum of the cost function is assured provided relatively weak requirements on the cost function are satisfied.
2. From the practical point of view, the Propositions 2 and 3 give useful estimates of the measure of the target region. More specifically, when more and more generations of MDEA are constructed and these generations do not provide any improvement of the attained minimum of the corresponding cost function, we can use this negative result as a base for the probabilistic estimates of the size of the target region according to Propositions 2 and 3.

To illustrate the dependence of the quantities ψ_0 (see 14) and N_0 (see 15) on values α_S for three different levels of the probability p_0 we summarized these quantities in Tab. I.

α_S	C_R	ψ_0	N_0		
			$p_0 = 10^{-6}$	$p_0 = 10^{-9}$	$p_0 = 10^{-12}$
0.1	0.9	2.302 586	2 302 586	2 302 586 000	2 302 586 000 000
0.05	0.95	2.995 733	2 995 733	2 995 733 000	2 995 733 000 000
0.01	0.99	4.605 171	4 605 171	4 605 171 000	4 605 171 000 000
0.005	0.995	5.298 318	5 298 318	5 298 318 000	5 298 318 000 000
0.001	0.999	6.907 756	6 907 756	6 907 756 000	6 907 756 000 000
0.000 5	0.999 5	7.600 903	7 600 903	7 600 903 000	7 600 903 000 000
0.000 1	0.999 9	9.210 341	9 210 341	9 210 341 000	9 210 341 000 000
0.000 05	0.999 95	9.903 488	9 903 488	9 903 488 000	9 903 488 000 000
0.000 01	0.999 99	11.512 926	11 512 926	11 512 926 000	11 512 926 000 000
0.000 005	0.999 995	12.206 073	12 206 073	12 206 073 000	12 206 073 000 000
0.000 001	0.999 999	13.815 511	13 815 511	13 815 511 000	13 815 511 000 000

Tab. I Values ψ_0 and N_0 corresponding to different values α_S and probabilities p_0 .

The Tab. I demonstrates the fact that when we suppose a fixed value of probability $p = p_0$ and perform N_0 experiments with a negative result we can claim with a relative certainty C_R that $p < p_0$. The higher relative certainty we require the more experiments have to be performed, in other words the more random individuals have to be generated.

6. Conclusions

During the study and utilization of differential evolution algorithms we revealed the principle limitations of the standard differential evolution algorithm CDEA. This fact was the base to construct an improved algorithm MDEA with substantially better prospective to identify successfully the global minimum of the cost function. In this paper we concentrate specifically on the role of random individuals in MDEA. We can recapitulate the main results in the following summary:

- CDEA can converge to a local minimum of the cost function as a result of the premature convergence. This fact constitutes the principal weakness of CDEA.
- MDEA converges asymptotically to the global minimum of the cost function in probability on the assumption that the cost function satisfies the condition (3) where $\mu(S_\varepsilon^*) > 0$.
- In the real optimization task we are usually able to find a minimum of the cost function. Let us denote this minimum as the optimized minimum. But the principal question is whether this optimized minimum is only local or global (there can be more equivalent global minima).
- MDEA solves this question by random individuals formed in each generation of the algorithm. The random individuals explore the search space of the optimization task and not only the neighbourhood of the optimized minimum in each generation.
- If the random individuals identify a region with lower (better) minimum than the optimized minimum we made a useful step towards the global minimum. Of course this new minimum can be just another local minimum. Nevertheless, we found a more optimized state and this can be considered a partial success.
- In the negative case when more and more random individuals are generated and we do not get any improved optimized minimum we do not achieve any improved state (better optimized minimum). Nevertheless, we can state with a determined probability that any minimum with lower value than the optimized minimum does not exist. This probability can be made arbitrarily close to 1 if we generate sufficiently large number of random individuals.

We believe that the presented statements contribute to better understanding of the convergence process of MDEA and subsequently to its broader utilization in practical optimization tasks.

Acknowledgement

This work was supported by project Modular platform for autonomous chassis of specialized electric vehicles for freight and equipment transportation, Reg. No. CZ.02.1.01/0.0/0.0/16_025/0007293.

References

- [1] ALSWAIITI M., ALBUGHDADI M., ISA N.A. Variance-based differential evolution algorithm with an optional crossover for data clustering. *Applied Soft Computing*. 2019, 80, pp. 1–17, doi: [10.1016/j.asoc.2019.03.013](https://doi.org/10.1016/j.asoc.2019.03.013).
- [2] BARTOSZYŃSKI R., NIEWIADOMSKA-BUGAJ M. *Probability and Statistical Inference*. Wiley – Interscience, A John Wiley & Sons, Inc., Publication, 2008, ISBN 978-0-471-69693-3.
- [3] CHOI T.J., AHN C.W. An Improved Differential Evolution Algorithm and Its Application to Large-Scale Artificial Neural Networks. *Journal of Physics: Conference Series*, 2017, 806(1).
- [4] ČERVENKA M., BOUDNÁ H. Visual Guide of F and CR Parameters Influence on Differential Evolution Solution Quality. In: *24th International Conference Engineering Mechanics 2018*, Svratka, Czech Republic, 2018, pp. 141–144, doi: [10.21495/91-8-141](https://doi.org/10.21495/91-8-141).
- [5] GOU X., HUANG T., YANG S., SU M., ZENG F. Optimized Differential Evolution Algorithm for Software Testing. *International Journal of Computational Intelligence Systems*. 2018, 12(1), pp. 215–226.
- [6] HU Z., XIONG S., FANG Z., SU Q. A Convergent Differential Evolution Algorithm with Hidden Adaptation Selection for Engineering Optimization. Hindawi Publishing Corporation, *Mathematical Problems in Engineering*. 2014, 11 pages, doi: [10.1155/2014/135652](https://doi.org/10.1155/2014/135652).
- [7] ISLAM S.M., DAS S., GHOSH S., ROY S., SUGANTHAN P.N. An Adaptive Differential Evolution Algorithm with Novel Mutation and Crossover Strategies for Global Numerical Optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*. 2012, 42(2), pp. 482–500.
- [8] JEYAKUMAR G., SHANMUGEVALAYUTHAM C. Convergence Analysis of Differential Evolution Variants on Unconstrained Global Optimization Functions. *International Journal of Artificial Intelligence & Applications (IJAIA)*. 2011, 2(2), doi: [10.5121/ijaia.2011.2209116](https://doi.org/10.5121/ijaia.2011.2209116).
- [9] KNOBLOCH R., MLÝNEK J., SRB R. Improving Convergence Properties of a Differential Evolution Algorithm. In: *Proceedings of 42nd International Conference Applications of Mathematics in Engineering and Economics 2016*, published by the American Institute of Physics, 1789, 030005 (2016), doi: [10.1063/1.4968451](https://doi.org/10.1063/1.4968451).
- [10] KNOBLOCH R., MLÝNEK J., SRB R. Convergence Rate of the Modified Differential Evolution Algorithm. In: *Proceedings of 43rd International Conference Applications of Mathematics in Engineering and Economics 2017*, published by the American Institute for Physics, Conf. Proc. 1910, 030005-1-030005-8, 2017, ISBN 987-0-7354-1602-4, doi: [10.1063/1.5013964](https://doi.org/10.1063/1.5013964).
- [11] KNOBLOCH R., MLÝNEK J., SRB R. The Classic Differential Evolution Algorithm and Its Convergence Properties. *Journal Applications of Mathematics*, published by Institute of Mathematics of the Czech Academy of Sciences, Prague. 2017, 62(2), pp. 197–208, doi: [10.21136/AM.2017.0274-16](https://doi.org/10.21136/AM.2017.0274-16).
- [12] MLÝNEK J., KNOBLOCH R. Model of Shell Metal Mould Heating in the Automotive Industry. *Journal Applications of Mathematics*, published by Institute of Mathematics of the Czech Academy of Sciences, Prague, 2018, 63(2), pp. 111–124.
- [13] MLÝNEK J., KNOBLOCH R., SRB R. Optimization of a Heat Radiation Intensity and Temperature Field on the Mould Surface. *Proceedings of 30th European Conference on Modelling and Simulation*, Regensburg, Germany, 2016, ISBN 987-0-9932440-2-5, doi: [10.7148/2016-0425](https://doi.org/10.7148/2016-0425).
- [14] MOHAMED A.W. RDEL: Restart Differential Evolution Algorithm with Local Search Mutation for Global Numerical Optimization. *Egyptian Informatics Journal*. 2014, 15(3), pp. 175–188, doi: [10.1016/j.eij.2014.07.001](https://doi.org/10.1016/j.eij.2014.07.001).
- [15] PIOTROWSKI A.P., KICZKO A., NAPIÓRKOWSKI J.J. Differential Evolution Algorithm with Separated Groups for Multidimensional Optimization Problems. *European Journal of Operational Research*. 2012, 216(1), pp. 33–46, doi: [10.1016/j.ejor.2011.07.038](https://doi.org/10.1016/j.ejor.2011.07.038).
- [16] PRICE K.V., STORN R.M., LAMPIEN J.A. *Differential Evolution, A Practical Approach to Global Optimization*. Springer-Verlag, Berlin Heidelberg, 2005.

- [17] RUDIN W. *Principles of Mathematical Analysis*. McGraw-Hill Book Company, Third Edition, 1976.
- [18] STORN R.M., PRICE K.V. Differential Evolution — A Simple and Efficient Heuristics for Global Optimization over Continuous Spaces. *Journal of Global Optimization*. Kluwer Academic Publishers. 1997, 11, pp. 341–359.
- [19] WANG J., ZHANG C., JIANG T., DUAN Q. Research on Wireless Network Node Localization Algorithm for Internet of Things. *Proceedings of 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference*, ITAIC 2019, Chongqing, China, 2019, pp. 548–558, doi: [10.1109/ITAIC.2019.8785836](https://doi.org/10.1109/ITAIC.2019.8785836).
- [20] ZHANG B., SUN X., LIU S., DENG X. Adaptive Differential Evolution-Based Receding Horizon Control Design for Multi-UAV Formation Reconfiguration. *International Journal of Control, Automation and Systems*. 2019, doi: [10.1007/s12555-018-0421-2](https://doi.org/10.1007/s12555-018-0421-2).
- [21] HU Z., XIONG S., SU Q., ZHANG X. Sufficient Conditions for Global Convergence of Differential Evolution Algorithm, *Journal of Applied Mathematics*, Article ID 139196, 2013.