Journal of Computational and Applied Mathematics 268 (2014) 121–134

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Finite Markov chain analysis of classical differential evolution algorithm

ZhongBo Hu [a,b,*], ShengWu Xiong [a], QingHua Su [a,b], ZhiXiang Fang [c]

[a] *School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, PR China*

[b] *School of Mathematics and Statistics, Hubei Engineering University, Xiaogan 432000, Hubei, PR China*

[c] *State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430072, PR China*

ARTICLE INFO

Article history:
Received 7 April 2012
Received in revised form 7 December 2012

Keywords:
Convergence in probability
Continuous optimization
Differential evolution algorithm
Markov chain

ABSTRACT

Theoretical analyses of algorithms are important to understand their search behaviors and develop more efficient algorithms. Compared with the plethora of works concerning the empirical study of the differential evolution (DE), little theoretical research has been done to investigate the convergence properties of DE so far. This paper focuses on theoretical researches on the convergence of DE and presents a convergent DE algorithm. First of all, it is proved that the classical DE cannot converge to the global optimal set with probability 1 by using the property that it cannot escape from a local optimal set. Inspired by the characteristics of the elitist genetic algorithm, this paper proposed a modified DE to overcome the disadvantage. The proposed algorithm employs two operators that assist it in escaping from a local optimal set and enhance the diversity of the population. And it is then verified that the proposed algorithm is capable of converging to global optima with probability 1. The theoretical research of this paper is undertaken in a finite discrete set, and the analysis tool used is the Markov chain. The numerical experiments are conducted on a deceptive function and a set of benchmark functions. The experimental results support the theoretical analyses on the convergence performances of the classical and modified DE algorithm.

1. Introduction

Theoretical analyses of the convergence properties of differential evolution (DE) benefit understanding their search behaviors and developing more efficient algorithms. However, few studies focused on theoretical analyses of guiding to modify the DE. Inspired by Markov chain performance in the theoretical analyses of genetic algorithm (GA), this paper gives the convergence analyses of the classical DE, as well as present a modified DE, which converges to the global optimum with probability 1.

1.1. Background

DE, proposed by Storn and Price in 1995 [1], is a population-based stochastic parallel evolutionary algorithm. The first book on DE was published in 2005 [2]. Das and Suganthan [3] surveyed DE in detail in 2011, which includes basic concepts,

* Corresponding author at: School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, PR China. Tel.: +86 13971978456.

E-mail addresses: huzbdd@126.com, huzbdd@yahoo.com.cn (Z. Hu).

http://dx.doi.org/10.1016/j.cam.2014.02.034
0377-0427/

major variants, application and theoretical studies. Das et al. [4] identified the four main directions of DE research, which include the basic DE research, the application-general research, the application-specific research and the computational environment specific research.

DE [1–7] emerged as a powerful stochastic real-parameter optimization algorithm. And DE takes few control parameters, which makes it easy to implement. Perhaps these advantages triggered the popularity of DE among researchers within a short time. In recent years, DE has also got many real-world applications [8–10] and a lot of improved DE algorithms have been proposed. Ref. [11] subdivided these modified versions of DE into two classes:

- DE integrating an extra component. This class includes those algorithms [12–16] which use DE as an evolutionary framework which is assisted by additional algorithmic components.
- DE with modified structures. This class includes those algorithms [17–23] which make a substantial modification within the DE structure, in the search logic or the selection etc.

No matter DE integrating an extra component or DE with modified structures, the motivations of these improved DE algorithms are either based on a certain complementarity of the biological mechanism or based on the exploiting and exploring capability of the classical DE. Few motivations improving DE algorithms are based on theoretical analyses on the convergence properties of the classical DE. Perhaps the main reason is that little theoretical research has been presented on investigating the convergence properties of DE so far.

In 2005, Xu et al. [24] performed the mathematical modeling and convergence analysis of continuous multi-objective differential evolution (MODE) under certain simplifying assumptions. The authors assumed that the DE-population is initialized by sampling from a Gaussian distribution with given mean and standard deviation. They then proved that the initial population is Gaussian distributed and contains the Pareto optimal set, and the subsequent populations generated by the MODE without the selection operator are also Gaussian distributed and the population mean converges to the center of the Pareto optimal set. This work was extended in [25].

In 2006, Ter Braak C.J.F. [26] proposed the Differential Evolution Markov Chain algorithm (DE-MC). Its reproduction operator was defined by adding a uniformly random number to the mutation operator of the classical DE. DE-MC abandons crossover operator and uses Metropolis selection operator instead of the greedy one used in the classical DE. This paper also proved DE-MC yields a unique joint stationary distribution. However, it has not been theoretically proven whether DE-MC holds certain asymptotic convergence. In fact, it does not seem difficult to prove that DE-MC holds with convergence in distribution. However, only if its selection operator meets certain conditions, DE-MC can converge to the global optimum with probability 1. For example, the selection pressure [27,28] of Metropolis selection operator satisfies certain conditions and the probability of selecting bad individuals approaches 0 when iteration times tends to infinity. DE-MC's selection operator does not seem to meet the condition.

Whether the classical DE holds certain asymptotic convergence or not, these papers do not analyzed. The area on theoretical analyses, especially for guiding to modify the DE, remains largely open. Some improved DE algorithms, which are based on theoretical analyses on the convergence properties, will need to be further research.

1.2. Motivation and contribution

Like GA, DE uses the similar computational steps (mutation, crossover, and selection operators) at each generation to move its population toward the global optimum. DE and GA are both typical evolutionary algorithm. Through analyzing the reproduction operators of two algorithms, we can get that there are several differences between GA and DE as follows.

(1) DE is usually used to dealing with the continuous optimization problem. The search space of the continuous optimization problem is a continuous closed region (or a union of countable continuous closed regions). Therefore, the solution space of the classical DE with real code is generally continuous, which is opposite to the canonical GA with binary code. In contrast with a discrete solution space, the numbers of individuals in a continuous space is infinite.
(2) Using the mutation operator of the canonical GA, the GA's mutation probability from an individual to another which is an arbitrary point in a search space is greater than 0. However, the DE's mutation probability from an individual to another may equal to 0. In fact, the DE's mutation probability from an individual to another equals to 0 if the distance of the two individuals is enough great.
(3) Compared with the canonical GA which eliminates the best solution with a little-probability, the selection operator of the classical DE always maintains the best solution(s) in the population.

Inspired by Markov chain performance in the theoretical analyses of GA [29–31], this paper will employ the Markov chain as main tool. According to 1st difference, we must discretize the search space. Considering the limitations of the calculation accuracy in computer, this paper maps the continuous search space to a finite discrete set. A Markov chain model is then developed to investigate the classical DE. The theoretical analyses show that the classical DE cannot escape from a local optimal set in case of trapping in. This makes the classical DE not converge in probability to the global optimum of a continue optimization problem.

According to the 2nd difference, the classical DE's populations are not always accessible to each other in population space. And DE, like the elitist GA, always remains the best solution(s) in the current population according to the 3rd difference. Taking these two points above into account, this paper presents two operators, called *uniform mutation* and *diversity selection*