

Accelerating Evolutionary Computation with Elite Obtained in Projected One-Dimensional Spaces

Pei, Yan
Graduate School of Design, Kyushu University

Takagi, Hideyuki
Faculty of Design, Kyushu University

<https://hdl.handle.net/2324/1808904>

出版情報 : International Conference on Genetic and Evolutionary Computing (ICGEC). 2011, pp.89-92, 2011-09-01. IEEE

バージョン :

権利関係 :

Accelerating Evolutionary Computation with Elite Obtained in Projected One-Dimensional Spaces

Yan Pei

Graduate School of Design, Kyushu University
Fukuoka, Japan
peiyan@kyudai.jp

Hideyuki Takagi

Faculty of Design, Kyushu University
Fukuoka, Japan
<http://www.design.kyushu-u.ac.jp/takagi>

Abstract—We propose a method for accelerating evolutionary computation (EC) searches using an elite obtained in one-dimensional space and use benchmark functions to evaluate the proposed method. The method projects individuals onto n one-dimensional spaces corresponding to each of the n searching parameter axes, approximates each landscape using Lagrange polynomial interpolation or power function least squares approximation, finds the best coordinate for the approximated shape, obtains an elite by combining the best n found coordinates, and uses the elite for the next generation of the EC. The advantage of this method is that the elite may be easily obtained thanks to their projection onto each one-dimensional space and there is a higher possibility that the elite will be located near the global optimum. Experimental tests with differential evolution and eight benchmark functions show that the proposed method accelerates EC convergence significantly, especially in early generations.

Keywords—evolutionary computation, fitness landscape, convergence acceleration, dimensionality reduction, elite combination search

I. INTRODUCTION

Accelerating evolutionary computation (EC) is necessary for many EC applications to improve the performance of the target systems. For example, as interactive EC (IEC) optimizes the target system based on the IEC user's subjective evaluation, user fatigue is a serious problem for its practical use. Multiple trials for accelerating EC have therefore been proposed [5].

One such acceleration approach is to use the EC landscape information directly. Some methods that use this approach include the combination of the EC search with gradient methods, hill-climbing methods, or other metaheuristic methods. Other approaches include landscape approximation with simpler shapes [4]. As the complexity of the approximated landscape is lower than that of its original searching space, it becomes easier to reach to the neighbor area around the global optimum. Although the global optimum of the approximated simple space is different from that of its original space, it provides useful information for finding the real global optimum.

The objectives of this paper are to propose a method for finding elite obtained in regression spaces and to evaluate the acceleration performance of the elite with benchmark functions. Here, *regression space* refers to a lower dimensional space consisting of k dimensional (k -D) axes of the original n -D axes, where we use $k = 1$ in this paper. In other words,

individuals are projected onto the k -D space to simplify finding the elites. This elitism does not destroy the original EC search space by approximation, but can accelerate the EC convergence with less computational cost by conducting the elite search in the n regression spaces of k -D ($k = 1$). This is the originality contributed by this paper.

Following this introductory section, in Section 2, the elite combination search approach enabled by a technique for reducing dimensionality of the searching space is explained in detail, and we shows how elite can be obtained from the regression search space. In Section 3, experimental evaluations with eight benchmark functions [1][3] are conducted, and their results are discussed. Finally, we discuss our proposed methods and obtained results in Section 4, and conclude on the performance of the proposed methods in Section 5.

II. OBTAINING ELITE FROM A REGRESSION SEARCH SPACE

A. Dimensionality Reduction Approach

It is not easy for conventional interpolation or approximation approaches to find an accurate regression search space corresponding to the original multi-dimensional space. An alternative approach is to reduce the dimensionality of the original search space and find approximation curve expressions in the lower dimensional space.

Our approach for reducing the dimensionality of the searching space uses only one of the n parameter axes at a time instead of all n parameter axes, and projects individuals onto each 1-D regression space. The landscape of the n -D parameter space is given by a fitness function, $y = f(x_1, x_2, \dots, x_n)$, and the fitness value of the m -th individual are given by Equation (1):

$$y_m = f(x_{1m}, x_{2m}, \dots, x_{nm}) \quad (m = 1, 2, \dots, M) \quad (1)$$

There are M individuals with n -D parameter variables. We project the individuals onto the n 1-D spaces as follows:

$$(x_1, y) \quad (x_2, y) \quad \dots \quad (x_n, y)$$

Each of the n 1-D regression spaces has M projected individuals.

B. Method for Simplifying the Regression Space Landscape to Select Elite

We interpolate or approximate the landscape of each 1-D regression space using the projected M individuals and select elite from the n approximated 1-D landscape shapes. In this paper, we test two approaches for approximating the 1-D regression search spaces; in the first approach we use a Lagrange two-degree polynomial interpolation and in the other linear least squares approximation. Elite are generated from the resulting approximated shapes (see Figure 1).

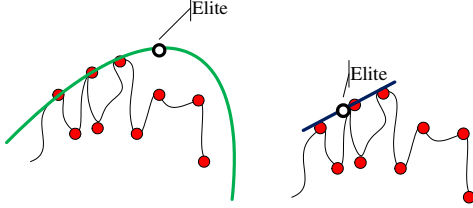


Figure 1. New elite selection approaches from a regression search space. Left: by Lagrange interpolation approach; Right: by least squares approximation approach.

Finding the elite corresponds to a kind of local search in the area where relatively better individuals exist in the original search space. The global optimum is expected to be near this area and we may be able to find it with a probability higher than chance [4]. So the elite selection approach is a critical step in the proposed acceleration processes. As the elite obtained by the two different elite selection approaches is different, it is expected that the acceleration performance will also differ. Further, the regression EC search space obtained by approximation or interpolation has its own characteristics and particularities, and we must use an efficient approach to obtain an elite from this simplified search space after analyzing its characteristics. Lagrange two-degree polynomial interpolation simplifies a regression space with a nonlinear curve, but it is easy to obtain its stationary point from its gradient, using the stationary point as the elite (see Figure 1). Linear least square approximation uses a linear function to approximate the regression space. Its gradient is either descent or ascent. A safer approach, taking into account both descent and ascent, is to select the middle point of the linear approximation line as the elite (see Figure 1). The actual interpolation polynomial and least square regression functions used are given by the Equations (2) and (3).

$$L(x) = \sum_{k=1}^3 \left\{ \prod_{i=1, i \neq k}^3 \frac{(x - x_i)}{(x_k - x_i)} \right\} y_k \quad (2)$$

$$\begin{pmatrix} (\varphi_0 \varphi_0) (\varphi_0 \varphi_1) \\ (\varphi_1 \varphi_0) (\varphi_1 \varphi_1) \end{pmatrix} \begin{pmatrix} D \\ E \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \quad (3)$$

where x_i, x_k, y_i are the best projected individual and their fitness values among (x_{im}, y_m) in the i -th 1-D regression space for $(i = 1, 2, \dots, n)$ and $(m = 1, 2, \dots, M)$, D and E are the parameters obtained through a least squares method, and φ_0 , and φ_1 are the vectors of the least square equation.

The proposed methods replace the worst individual in each generation with the selected elite. Although we cannot deny the small possibility that the global optimum is located

near the worst individual, the possibility that the worst individual will become a parent in the next generation is also low; removing the worst individual therefore presents the least risk and is a reasonable choice.

C. Method for Synthesizing Elite

The methods in section II select n elite points in n 1-D regression spaces, respectively: $x_{1-elite}, x_{2-elite}, \dots$, and $x_{n-elite}$. The n -D elite used for accelerating EC convergence in the next generation is obtained as follows:

$$\text{new elite} = (x_{1-elite}, x_{2-elite}, \dots, x_{n-elite}).$$

It is easier to calculate elite in a lower dimensional space than a higher dimensional space. Although we use 1-D as the lower dimensional space in this paper, the method need not be restricted to 1-D in general.

Once a new elite has been obtained, there are two approaches for how it can be handled. In one cautious approach, the fitness value of this elite is calculated to determine whether the new elite is really useful for acceleration, and in the other straightforward approach, the new elite is inserted into the next EC iteration process without any prior consideration or judgment.

Our proposed method is based on the hypothesis that elite calculated by interpolation or approximation from relatively better individuals will also have good fitness; synthesizing an n -D elite from n elite points in n 1-D regression spaces will also produce good elite; the probability of the global optimum being located near to the synthesized elite is high. In general, this proposed method represents a novel local search approach for accelerating EC convergence, and it is this approach that represents this paper's contribution.

III. EXPERIMENTAL EVALUATIONS

We use the DeJong five functions (F1 - F5) [1], Rastrigin function (F6), Schwefel function (F7) and Griewank function (F8) [3] as benchmark functions to evaluate the proposed approaches. Their dimensions and search ranges of all parameters are 3-D F1 in $[-5.12, 5.12]$, 2-D F2 in $[-2.048, 2.048]$, 5-D F3 in $[-5.12, 5.12]$, 30-D F4 in $[-1.28, 1.28]$, 2-D F5 in $[-65.536, 65.536]$, 5-D F6 in $[-5.12, 5.12]$, 5-D F7 in $[-512, 512]$, and 5-D F8 in $[-512, 512]$. All these function optimization tasks are posed as minimization problems with the optimal solution being the point with the lowest value. Their landscapes have a variety of characteristics; they include both continuous and discontinuous, convex and non-convex, unimodal and multimodal, and low dimensional and high dimensional shapes.

We use differential evolutionary (DE/best/1/bin) as the optimization method to evaluate the proposed approaches. We test each benchmark function with 50 trial runs of 50 generations and apply sign tests to evaluate the performance of our proposed approaches. Here, we abbreviate the DE approach where the search space is regressed by a two-degree Lagrange interpolation as DE-LR, where it is regressed by a line power function least squares approximation as DE-LS, and normal DE as DE-N. These abbreviations are also used in Figures 2 and 3. Figure 2 shows the

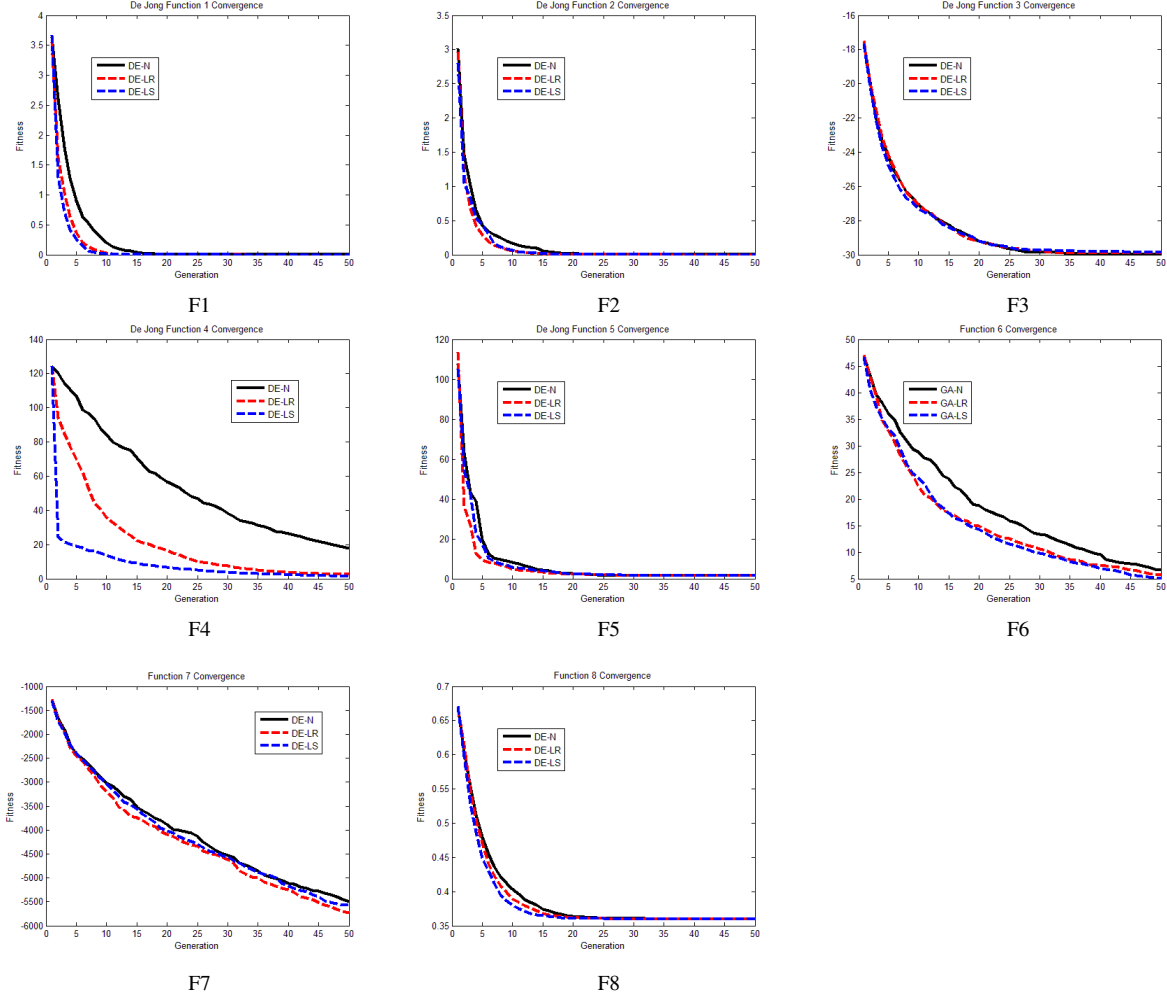


Figure 2. Average convergence curves of 50 trial runs for (F1) DeJong Function1, (F2) DeJong Function2, (F3) DeJong Function3, (F4) DeJong Function4, (F5) DeJong Function5, (F6) Rastrigin Function, (F7) Schwefel Function, and (F8) Griewank Function.

average convergence curves of the best fitness values for 50 trial runs of (DE-N vs. DE-LR) and (DE-N vs. DE-LS) for eight benchmark functions, and Figure 3 shows their sign test results at each generation. From these results, we can say that (1) our proposed methods could accelerate all benchmark functions well except F3 and F7, (2) they did not accelerate DE convergence well for F3 and F7 though there was no significant difference between normal DE and DE with either of our proposed methods (DE-LR or DE-LS), and (3) their performance look similar and, if there is any difference, the superiority depends on the task being performed. In most of cases, DE using the proposed methods are better than normal DE in latter generations, i.e. when search points approach the global optimum.

IV. DISCUSSIONS

Figure 4 shows rough landscapes for eight benchmark functions. We analyze the performance of our proposed methods from the experimental results of Figures 2 and 3 and landscape shapes of Figure 4.

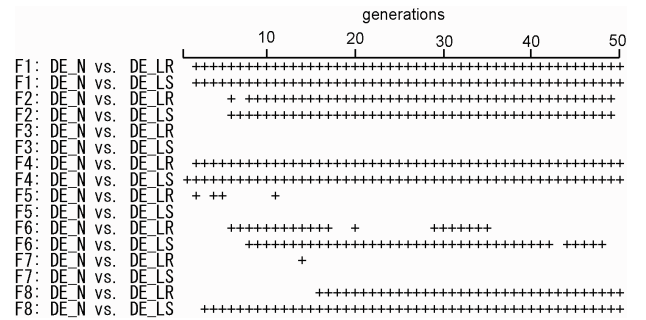


Figure 3. Sign test results for 50 trial runs of (DE-N vs. DE-LR) and (DE-N vs. DE-LS) per generation, where DE-N, DE-LR, and DE-LS mean normal DE, DE with Lagrange interpolation, and DE with least squares approximation. See F1 - F8 in Figure 2. The + mark means that a propose method is converges significantly better than normal DE ($p < 0.05$). There were no cases where the proposed methods were significantly poorer than normal DE.

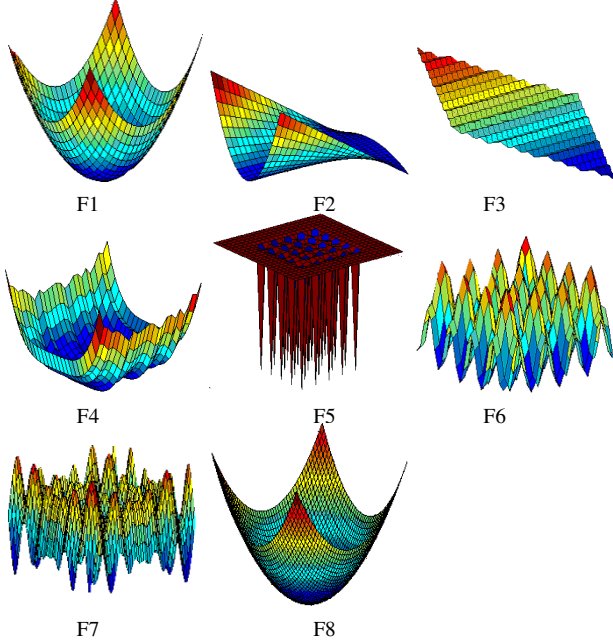


Figure 4. 2-D landscapes of (F1) DeJong Function1, (F2) DeJong Function2, (F3) DeJong Function3, (F4) DeJong Function4, (F5) DeJong Function5, (F6) Rastrigin Function, (F7) Schwefel Function, and (F8) Griewank Function.

As the global optimum of F3 is on the edge of a searching space, the elite obtained around the global optimum by function approximation may be located outside of the searching range. Our experiment did not use elite in this case, and DE using our proposed methods became identical to normal DE then. This would explain why there was no significant difference among DE-N, DE-LR and DE-LS.

F7 is made by putting multiple sine curves on a sloped plane that is otherwise similar to F3. The difficulty with this task is that the minimum point, i.e. the global optimum, is located near the maximum point of the plane and there are local minimums around the global optimum. When local minimums and the global optimum are located close together, the elite obtained from an approximated curve may be located in a gap between these peaks and its fitness may be poor. This may be a reason for why the performance of our proposed methods was the same as with normal DE.

F4 is a quartic function with Gaussian noise and has its global optimum in the center of the searching space. Although it is a multimodal function due to the quartic function and fluctuates due to noise, their influence is relatively small and its whole shape is close to a quadratic function such as in F1. Elite from an approximated function using individuals that fluctuate slightly should be located in the center, and the performance of DE-LR and DE-LE should be better, as with F1, while that of normal DE is negatively influenced by the fluctuations. This may explain the good performance of our proposed method with F4.

Reducing computational cost is the biggest feature of our methods. This approach of approximating landscapes and using the peak of the approximated function is in the same

category as the reference [4], but we extended the work by introducing projection onto a 1-D space, synthesizing an elite from the elites on multiple 1-D functions, and reducing computational cost. Let's examine the efficiency of the approaches by comparing the number of algebraic operations required.

Let us consider a multi-variable regression model, $y = X\beta + \varepsilon$, where y, X, β and ε are respectively a fitness vector, a parameter vector matrix, a coefficient vector, and an error vector. The dimension of X is decided by the dimension of the search space and the number of sampled data that we select to obtain a regression model. These variables become scalars in a 1-D space. When we use least square approach to obtain the unknown coefficient vectors and use Gaussian elimination to solve system equations by M pieces of (individual, fitness) pairs of sampled data, the number of algebraic operations is $n(n-1)(n+m) + mn(n+1) + n^3$ for an n -D space, while it is $n(2m+1)$ for our proposed method because we perform n times the number of operations required for a 1-D space, i.e. $1(1-1)(1+m) + m*1(1+1) + 1^3$. This demonstrates how our proposed methods can reduce computational cost significantly.

V. CONCLUSION

We proposed that the EC landscape be approximated with a two-degree Lagrange polynomial interpolation or a linear power function least squares approximation to accelerate EC convergence. The novel feature in these acceleration methods is to use an elite synthesized from elite points in lower dimensional spaces. Our experimental evaluation of these two approaches with eight benchmark functions showed that the proposed methods can accelerate EC search especially when the landscape of the EC tasks takes on a roughly big valley structure. We also analyzed the relationship between the proposed approaches' performance and the landscape shapes.

Acknowledgement: Yan Pei would like to thank Yoshida Scholarship Foundation for its support of his doctoral research.

REFERENCES

- [1] De Jong, Kenneth A., "An Analysis of the Behavior of a Class of Genetic Adaptive System", Ph.D Dissertation, University of Michigan, USA (1975).
- [2] Jin, Yaochu, "A Comprehensive Survey of Fitness Approximation in Evolutionary Computation", Soft Computing, Springer, Vol.9, No.1, pp.3-12 (2005).
- [3] Whitley, D., Mathias, K., Rana, S., and Dzubera, J., "Building Better Test Functions", 6th Int. Conf. on Genetic Algorithms (ICGA95), Pittsburgh, PA, USA, pp.239-246 (July, 1995).
- [4] Takagi, H., Ingu, T., and Ohnishi, K., "Accelerating a GA Convergence by Fitting a Single-Peak Function", J. of Japan Society for Fuzzy Theory and Intelligent Informatics, vol.15, no.2, pp.219-229 (2003) (in Japanese).
- [5] Takagi, Hideyuki, "Interactive Evolutionary Computation: Fusion of the Capabilities of EC Optimization and Human Evaluation", Proceedings of the IEEE, Vol.89, No.9, pp.1275-1296 (2001).