An Improved Differential Evolution Scheme for Noisy Optimization Problems

Swagatam Das and Amit Konar

Dept. of Electronics and Telecommunication Engineering, Jadavpur University,
Kolkata-700032, India
swagatamdas19@yahoo.co.in

Abstract. Differential Evolution (DE) is a simple and surprisingly efficient algorithm for global optimization over continuous spaces. It has reportedly outperformed many versions of EA and other search heuristics when tested over both benchmark and real world problems. However the performance of DE deteriorates severely if the fitness function is noisy and continuously changing. In this paper we propose an improved DE scheme which can efficiently track the global optima of a noisy function. The scheme performs better than the classical DE, PSO, and an EA over a set of benchmark noisy problems.

1 Introduction

The problem of optimizing noisy or imprecise (not exactly known) functions occurs in diverse domains of engineering application, especially in the task of experimental optimizations. A number of methods for dealing with these noisy optimization problems have been proposed in the last few years by various experts in the fields of evolutionary programming (EP) [1], evolutionary strategies (ES) [2], genetic algorithms (GA) [3] and particle swarm optimization (PSO) [4]. Although DE is a simple and very fast technique for numerical optimization, it is reported in [5] that the performance of DE becomes poorer in comparison to EA when the function to be optimized is corrupted by noise and accuracy of the results is a vital factor.

In this study we propose an improved DE (DE/Rand1/Exp) algorithm where the scalar factor used to weigh the difference vector has been made completely random. We also introduce a novel threshold based selection strategy for the DE algorithm. Under this scheme an offspring vector replaces its parent in the next generation only if its fitness is greater than the parent's fitness by a certain threshold value. This provides prevention from accepting poor candidate solutions, which may deceptively appear fitter due to noise.

2 The Classical Differential Evolution – A Brief Introduction

DE [7], [8] searches for a global optimum point in an N-dimensional hyperspace. It begins with a randomly initialized population of N-dimensional real-valued parameter vectors. Each vector forms a candidate solution to the multi-dimensional optimization problem. Unlike the conventional GA, the reproduction scheme in DE is maintained as follows. For each individual vector G_k^D belonging to generation D, randomly

sample three other individuals G_i^D, G_j^D and $G_m{}^D$ *from* the same generation (for distinct *i, j, k and m*), calculate the difference of the components (chromosomes) of G_i^D and $G_o{}^d$, scale it by a scalar R ($\in [0,1]$) and create a trial vector by adding the result to the chromosomes of G_k^D.

$$
\left.
\begin{aligned}
G_{k,n}{}^{D+1} &= G_{m,n}{}^D + R.(G_{i,\,n}{}^D - G_{j,\,n}{}^D) \qquad \text{if rand}_n\,(0,\,1) < CR \\
&= G_{k,\,n}{}^D, \qquad\qquad\qquad\qquad\qquad \text{otherwise.}
\end{aligned}
\right\}
\tag{1}
$$

for the n-th component of each parameter vector.

CR ($\in [0,1]$) is the crossover constant. The trial solution is evaluated and replaces its parent G_k^D. deterministically if its fitness is better.

3   The Newly Proposed Scheme

It is an observed fact that DE needs hardly any parameter tuning and converges surprisingly faster than PSO, GA and EA in most of the cases of static optimization problems. But we strongly feel that DE uses a less stochastic and greedier approach which makes the situation difficult for it whenever there is noise in the scenario.

In the original DE (1) the difference vector (G_i-G_j) is scaled by a constant factor 'R'. Usually the most popular choice for this control parameter in DE is in the range of (0.4, 1). However, we set this scale factor to change in a random fashion in the range (0.5, 1) by using the relation [11]

$$
R = 0.5*(1+ \text{rand}\,(0,\,1)) \tag{2}
$$

where rand (0,1) is a uniformly distributed random number within the range [0, 1]. The mean value of the scale factor remains at 0.75. This allows for stochastic variation in the amplification of the difference vector and hence helps retain the population diversity as the search progresses.

Following the work done on EA in [6] we take up a threshold based selection procedure for DE. Here the offspring vector substitutes its parent vector in the new generation if its fitness is lesser than the fitness of the parent (in case of minimization problems) by a threshold margin τ. We keep the threshold margin proportional to the noise strength or variance ($\sigma_n{}^2$) i.e.

$$
\tau = k.\sigma_n{}^2 . \tag{3}
$$

4   Experimental Setup and Simulation Strategy

We have used the noisy versions of the following benchmark functions enlisted in table 1. All of these are minimization problems. The noisy versions of the benchmark functions are defined as:

$$
f_{noisy}(\bar{x}) = f(\bar{x}) + N(0, \sigma^2) \tag{4}
$$

with $N(0, \sigma^2)$ = Normal (or Gaussian) distribution with mean 0 and variance σ^2. To obtain N we take up the Box and Muller method [9] with various values of σ^2. In this

work, we compare the performance of Particle Swarm Optimization (PSO), classical DE, the new scheme and Evolutionary Algorithm [10] over the noisy benchmarks listed above. Due to space limitations, it is not possible to give the brief description of the competitor algorithms.

Table 1. Benchmark functions for simulation

Name of the Function and Dimension	Mathematical Representation
Sphere function (50D)	$f_1(x) = \sum_{i=1}^{n} x_i^2$
Rosenbrock function(50D)	$f_2(x) = \sum_{i=1}^{n} [100 \ (x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$
Griewank function(50D)	$f_4(x) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos(\frac{x_i}{\sqrt{i}}) + 1$
Rastrigin function (50D)	$f_3(x) = \sum_{i=1}^{n} [x_i^2 - 10 \ \cos(2\pi x_i) + 10]$
Levy No. 5 (2D)	$f_5 = \sum_{i=1}^{5} i \cos[(i+1)x_1 + i] \times \sum_{j=1}^{5} j \cos[(j+1)x_2 + j] + (x_1 + 1.42513)^2$ $+ (x_1 + 0.80032)^2$

5 Results of Simulation

Although we have experimented with noise variance values 0, 0.1, 0.2 ...1, due to space limitations we tabulate the final result in table 2 only for $\sigma_n^2 = 1.0$ i.e. when the noise strength is highest.
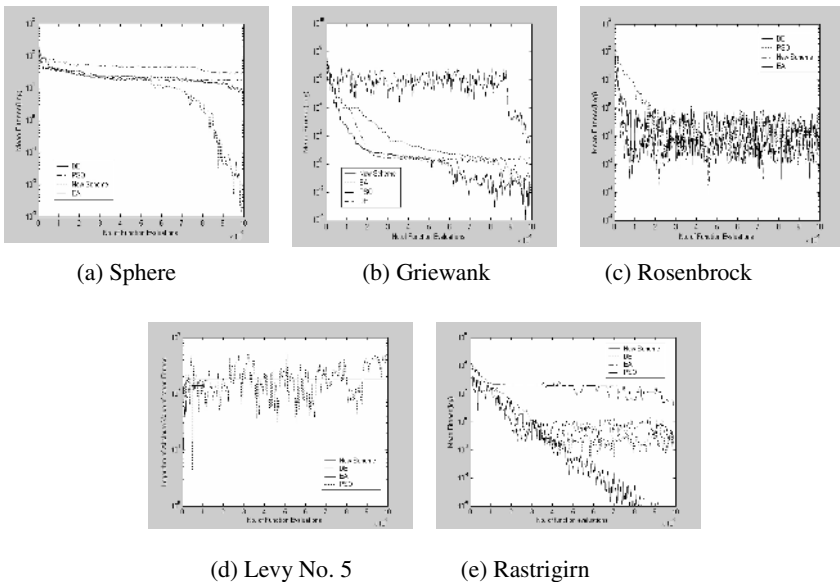

(a) Sphere   (b) Griewank   (c) Rosenbrock

(d) Levy No. 5   (e) Rastrigirn

Fig. 2. Progress to the optimum solution

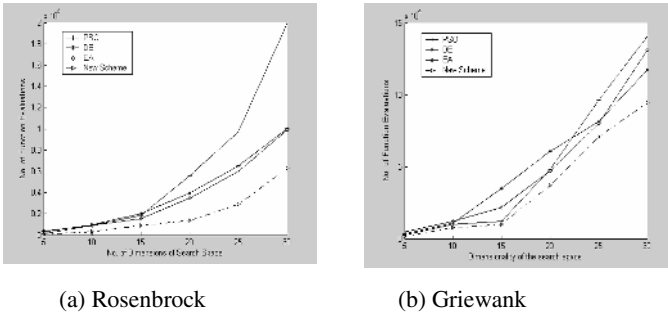(a) Rosenbrock   (b) Griewank

Fig. 3. Variation of mean convergence time with increase in dimensionality of the search space. (Shaded graph accounts for the new algorithm).

Table 2. Mean and Standard Error (±SE) of the Final Results after 10^5 Function Evaluations

Function	DE	PSO	EA	New Scheme
f_1	0.2145±0.02408	0.3562±0.06021	0.05002±0.0056	**6.123E-6±0.0002**
f_2	42.2748±0.1741	4559.75±867.89	114.25±12.68	**1.27±7.363**
f_3	2.4169±0.04591	51.4508±3.0067	32.7144±1.656	**2.6148±0.0474**
f_4	3.7142±0.0766	13.6445±1.4435	1.0975±0.0026	**0.2113±0.00054**
f_5	0.1386±0.0309	0.7765±0.1137	0.0982±0.02573	**0.02215±0.00663**

6   Conclusion

In this paper we firstly point out that, DE due to employing a deterministic selection scheme and having no random mutation process becomes 'greedy' and fails to optimize noisy functions satisfactorily. As a remedy to this problem, we equip the basic algorithm with a randomly changing scale factor and a threshold based selection scheme. Substantial empirical evidence has been provided to justify the usefulness of the proposed approach. The new method has been compared against (a) the basic DE, (b) the PSO, and (c) EA using a five-function test suite.

References

1. Fogel, L.J, Owens, A.J, Walsh, M.J.: Artificial Intelligence through Simulated Evolution. John Wiley & Sons, New York, (1966).
2. Rechenberg, I.: Evolution Strategy: Optimization of technical systems by means of biological evolution. Fromman-Holzboog, Stuttgart, (1973).
3. Holland, J.H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, MI, (1975)
4. Kennedy, J, Eberhart R.: Particle swarm optimization. Proc. IEEE Int. conf. Neural Networks. (1995) 1942-1948
5. Krink, T, Bodgan, F, Fogel, G.B., Thomson, R.: Noisy Optimization Problems – A Particular Challenge for Differential Evolution? Proc. Sixth Congress on Evolutionary Computation (CEC-2004). IEEE Press.

6. Markon, S, Arnold, V, D, Bäck, T, Beislstein, T, Beyer, G.-H.: Threshholding – a Selection Operator for Noisy ES. Proc. Congress on Evolutionary Computation (CEC-2001). 465-472
7. Storn, R., Price, K.: Differential evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, Journal of Global Optimization, 11(4) (1997) 341–359.
8. Price K. V.: An introduction to differential evolution. In New Ideas in Optimization. Corne, D, Dorigo, M, Glover, F, Eds. McGrow-Hill, London. (1999) 79-108.
9. Box, G. E. P, Muller, M. E.: A note on the generation of random deviates, Ann. Math. Statistics, vol. 29, (1958) 610-611.
10. Michalewicz, Z, Fogel, D. B.: How to Solve It: Modern Heuristics. Springer, Berlin, 2000.
11. Das, S., Konar, A., Chakraborty, U.K., Two improved differential evolution schemes for faster global search, to appear in the ACM-SIGEVO Proceedings of GECCO, Washington D.C., June 2005.