A Hybrid Estimation of Distribution Algorithm with Differential Evolution for Global Optimization

Bing Dong, Aimin Zhou, Guixu Zhang

East China Normal University, Shanghai, China

December 2016

Outline

Outline for Section 1

Definition

The box-constrained continuous global optimization can be stated in the following:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t} \quad & x \in [a_i, b_i]^n \end{aligned} \tag{1}$$

- $x = (x_1, x_2, \cdots, x_n)^T \in R^n$ is a decision vector
- $[a_i, b_i]^n$ is the search space
- $f : R^n \to R$ is the objective function

Differential Evolution(DE)

DE is a simple but powerful optimization algorithm. Classical DE algorithm consists of three steps:

- mutation: Utilize mutation operator to generate mutant vector.
- crossover: Utilize crossover operator to generate trial vector.
- selection: Target vector and trial vector competes to enter the next generation.

EDA is a recent stochastic optimization algorithm which mainly includes three steps:

- modeling: Build a probabilistic model.
- sampling: Generate individuals according to the built probabilistic model.
- selection: Select individuals from the generated individuals and parent population to the next generation.

Outline for Section 2

DE-EIG

DE-EIG is a novel DE which utilize eigenvector to rotate the original coordinate system. It is significant to extract the statistical information form the population.

Crucial work:

- crossover in a rotated coordinate system
- utilize a appropriate parameter to control the crossover in the original coordinate system or the rotated coordinate system

DE/EDA

DE/EDA is a algorithm combining DE and EDA.
Its main work:

- combine the differential information from DE and global information from EDA
- make a parameter to control the sampling of EDA

EDA/DE-EIG

Based on the framework of DE/EDA, we propose EDA/DE-EIG.
Our thoughts:

1. Import DE-EIG to improve the sampling of EDA.
2. Utilize a random parameter to control the resource allocations of DE-EIG and EDA.
3. Expensive local search is applied to refine the solutions further more.

Algorithm Framework

Initial the population $Pop(t) = \{x_1, x_2, x_3, \cdots, x_N\}$ (N is the size of the population)

while *not terminate* **do**

  Construct the probabilistic model:

  $p(x) = \prod_{i=1}^{n} \mathcal{N}(x_i; \mu_i, \sigma_i)$

  Generate a trial solution $u_{i,G}$ as follows:

  if $rand() < CRP$ **then**

    $u_{i,G}$ is produced by DE-EIG.

  else

    $u_{i,G}$ is sampled from the probabilistic model $p(x)$.

  end

  if $f(u_{i,G}) < f(x_{i,G})$ **then**

    $x_{i,G+1} = u_{i,G}$

  else

    $x_{i,G+1} = x_{i,G}$

  end

  if *Converage*(θ, G, G_e) **then**

    Operate the expensive local search.

  end

  $t = t + 1$

end

Figure 1: The algorithm framework of EDA/DE-EIG

Outline for Section 3

Compared algorithms and experimental settings

In this paper, EDA/DE-EIG is compared with JADE and DE/EDA on the first 13 test instances form YYL test instances.

- The dimension of the population is 30. All algorithms are run independently 50 times and stopped after 450,000 function evaluations.
- JADE: $N = 150, p = 0.05, c = 0.1, F = 0.5$ and $CR = 0.9$.
- DE/EDA: $N = 150, F = 0.5$ and $CRP = 0.9$.
- EDA/DE-EIG: $N = 150, CRP = 0.5, f = 0.5, CR = 0.6, P = 0.5, \theta = 0.1$.

All the algorithms are implemented by Matlab and executed at the same computer.

TABLE I

STATISTICAL RESULTS ($mean \pm std$) FOR THE THREE ALGORITHMS ON INSTANCES $f1 - f13$.

instances	EDA/DE-EIG	JADE	DE/EDA
$f1$	**1.54e-159 ± 5.11e-159**	$3.90e-127 \pm 2.74e-126(+)$	$1.39e-59 \pm 2.58e-59(+)$
$f2$	**1.02e-75 ± 7.46e-76**	$2.60e-35 \pm 1.64e-34(+)$	$5.15e-28 \pm 4.68e-28(+)$
$f3$	**4.01e-35 ± 8.47e-35**	$7.79e-35 \pm 2.51e-34(\sim)$	$1.23e-12 \pm 1.20e-12(+)$
$f4$	**5.01e-20 ± 3.06e-19**	$3.15e-14 \pm 6.42e-14(+)$	$9.90e-12 \pm 2.69e-11(+)$
$f5$	$1.46e-29 \pm 2.62e-29$	**3.85e-30 ± 9.58e-30**$(-)$	$3.37e-21 \pm 8.66e-21(+)$
$f6$	**0.00e+00 ± 0.00e+00**	**0.00e+00 ± 0.00e+00**$(\sim)$	**0.00e+00 ± 0.00e+00**$(\sim)$
$f7$	$3.60e-03 \pm 1.00e-03$	**6.01e-04 ± 2.23e-04**$(-)$	$2.20e-03 \pm 5.59e-04(-)$
$f8$	$2.79e+03 \pm 5.02e+02$	**4.74e+00 ± 2.34e+01**$(-)$	$1.82e+03 \pm 6.72e+02(-)$
$f9$	$6.23e+00 \pm 2.21e+00$	**0.00e+00 ± 0.00e+00**$(-)$	$1.54e+02 \pm 1.96e+01(+)$
$f10$	**4.44e-15 ± 0.00e+00**	**4.44e-15 ± 0.00e+00**$(\sim)$	**4.44e-15 ± 0.00e+00**$(\sim)$
$f11$	**0.00e+00 ± 0.00e+00**	$1.48e-04 \pm 1.05e-03(\sim)$	$2.96e-04 \pm 1.46e-03(\sim)$
$f12$	**1.57e-32 ± 5.53e-48**	**1.57e-32 ± 5.53e-48**$(\sim)$	**1.57e-32 ± 5.53e-48**$(\sim)$
$f13$	**1.35e-32 ± 1.11e-47**	**1.35e-32 ± 1.11e-47**$(\sim)$	**1.35e-32 ± 1.11e-47**$(\sim)$
		$3(+)6(\sim)4(-)$	$6(+)5(\sim)2(-)$

[1] The bold ones mean the best.

TABLE I

STATISTICAL RESULTS ($mean \pm std$) FOR THE THREE ALGORITHMS ON INSTANCES $f1 - f13$.

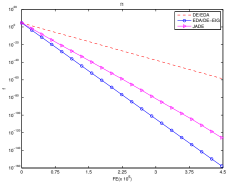
instances	EDA/DE-EIG	JADE	DE/EDA
$f1$	**1.54e-159 ± 5.11e-159**	$3.90e-127 \pm 2.74e-126$(+)	$1.39e-59 \pm 2.58e-59$(+)
$f2$	**1.02e-75 ± 7.46e-76**	$2.60e-35 \pm 1.64e-34$(+)	$5.15e-28 \pm 4.68e-28$(+)
$f3$	**4.01e-35 ± 8.47e-35**	$7.79e-35 \pm 2.51e-34$(∼)	$1.23e-12 \pm 1.20e-12$(+)
$f4$	**5.01e-20 ± 3.06e-19**	$3.15e-14 \pm 6.42e-14$(+)	$9.90e-12 \pm 2.69e-11$(+)
$f5$	$1.46e-29 \pm 2.62e-29$	**3.85e-30 ± 9.58e-30**(−)	$3.37e-21 \pm 8.66e-21$(+)
$f6$	**0.00e+00 ± 0.00e+00**	**0.00e+00 ± 0.00e+00**(∼)	**0.00e+00 ± 0.00e+00**(∼)
$f7$	$3.60e-03 \pm 1.00e-03$	**6.01e-04 ± 2.23e-04**(−)	$2.20e-03 \pm 5.59e-04$(−)
$f8$	$2.79e+03 \pm 5.02e+02$	**4.74e+00 ± 2.34e+01**(−)	$1.82e+03 \pm 6.72e+02$(−)
$f9$	$6.23e+00 \pm 2.21e+00$	**0.00e+00 ± 0.00e+00**(−)	$1.54e+02 \pm 1.96e+01$(+)
$f10$	**4.44e-15 ± 0.00e+00**	**4.44e-15 ± 0.00e+00**(∼)	**4.44e-15 ± 0.00e+00**(∼)
$f11$	**0.00e+00 ± 0.00e+00**	$1.48e-04 \pm 1.05e-03$(∼)	$2.96e-04 \pm 1.46e-03$(∼)
$f12$	**1.57e-32 ± 5.53e-48**	**1.57e-32 ± 5.53e-48**(∼)	**1.57e-32 ± 5.53e-48**(∼)
$f13$	**1.35e-32 ± 1.11e-47**	**1.35e-32 ± 1.11e-47**(∼)	**1.35e-32 ± 1.11e-47**(∼)
		3(+)6(∼)4(−)	6(+)5(∼)2(−)

[1] The bold ones mean the best.

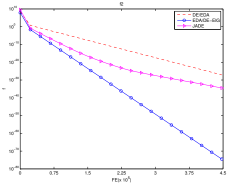
TABLE I
STATISTICAL RESULTS ($mean \pm std$) FOR THE THREE ALGORITHMS ON INSTANCES $f1 - f13$.

instances	EDA/DE-EIG	JADE	DE/EDA
$f1$	**1.54e-159 ± 5.11e-159**	$3.90e-127 \pm 2.74e-126(+)$	$1.39e-59 \pm 2.58e-59(+)$
$f2$	**1.02e-75 ± 7.46e-76**	$2.60e-35 \pm 1.64e-34(+)$	$5.15e-28 \pm 4.68e-28(+)$
$f3$	**4.01e-35 ± 8.47e-35**	$7.79e-35 \pm 2.51e-34(\sim)$	$1.23e-12 \pm 1.20e-12(+)$
$f4$	**5.01e-20 ± 3.06e-19**	$3.15e-14 \pm 6.42e-14(+)$	$9.90e-12 \pm 2.69e-11(+)$
$f5$	$1.46e-29 \pm 2.62e-29$	**3.85e-30 ± 9.58e-30(−)**	$3.37e-21 \pm 8.66e-21(+)$
$f6$	**0.00e+00 ± 0.00e+00**	**0.00e+00 ± 0.00e+00(∼)**	**0.00e+00 ± 0.00e+00(∼)**
$f7$	$3.60e-03 \pm 1.00e-03$	**6.01e-04 ± 2.23e-04(−)**	$2.20e-03 \pm 5.59e-04(-)$
$f8$	$2.79e+03 \pm 5.02e+02$	**4.74e+00 ± 2.34e+01(−)**	$1.82e+03 \pm 6.72e+02(-)$
$f9$	$6.23e+00 \pm 2.21e+00$	**0.00e+00 ± 0.00e+00(−)**	$1.54e+02 \pm 1.96e+01(+)$
$f10$	**4.44e-15 ± 0.00e+00**	**4.44e-15 ± 0.00e+00(∼)**	**4.44e-15 ± 0.00e+00(∼)**
$f11$	**0.00e+00 ± 0.00e+00**	$1.48e-04 \pm 1.05e-03(\sim)$	$2.96e-04 \pm 1.46e-03(\sim)$
$f12$	**1.57e-32 ± 5.53e-48**	**1.57e-32 ± 5.53e-48(∼)**	**1.57e-32 ± 5.53e-48(∼)**
$f13$	**1.35e-32 ± 1.11e-47**	**1.35e-32 ± 1.11e-47(∼)**	**1.35e-32 ± 1.11e-47(∼)**
		3(+)6(∼)4(−)	6(+)5(∼)2(−)

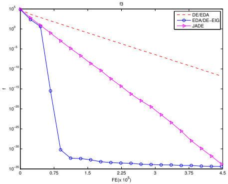
[1] The bold ones mean the best.



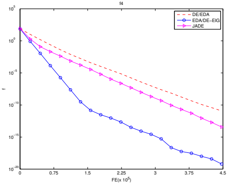
(a) $f1$



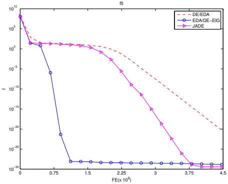
(b) $f2$



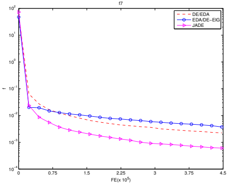
(c) $f3$



(d) $f4$

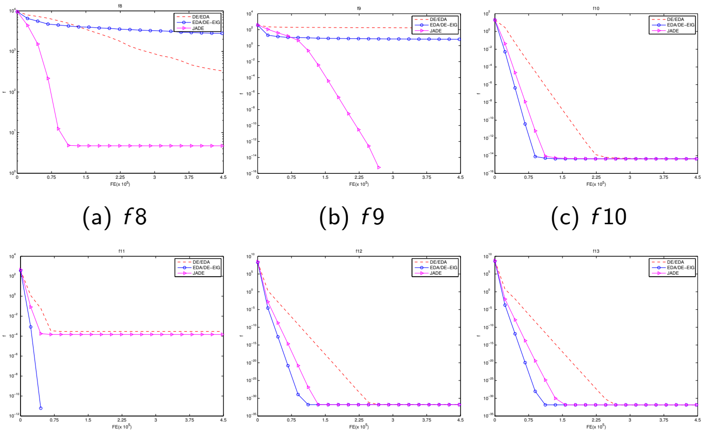


(e) $f5$



(f) $f7$

Figure 2: The mean function value versus on $f1 - f7$ except $f6$.



(a) $f8$ (b) $f9$ (c) $f10$

(d) $f11$ (e) $f12$ (f) $f13$

Figure 3: The mean function value versus on $f8 - f13$

According to figure 2 and figure 3, the following conclusions are obtained:

1. obtain best results on 8 out of 12 test instances
2. better than DE/EDA except $f7$ and $f8$
3. has a similar performance in comparison with JADE

Outline for Section 4

Conclusions

1. DE/EDA is a promising algorithm framework utilizing global and local information.
2. DE-EIG is significant to improve the sampling.
3. EDA/DE-EIG has a impressive performance comparing with JADE and DE/EDA.

Future work

The results reported in this paper is preliminary and there are several ways to improve the algorithm performance. The future work includes:

- simplify the algorithm framework of EDA/DE-EIG
- investigate the resources allocation of DE-EIG and EDA

Thanks!

- B. Dong, A. Zhou, and G. Zhang, A Hybrid Estimation of Distribution Algorithm with Differential Evolution for Global Optimization, 2016 IEEE Symposium Series on Computational Intelligence (SSCI), 2016.