Sampling in Latent Space for a Multiobjective Estimation of Distribution Algorithm

Bing Dong

East China Normal University, Shanghai, China

April 21, 2016

Outline

1 Background

2 DES

3 Experiment results

4 Conclusions

Definition

Multiobjective optimization problems (MOPs) can be defined as follows in this paper:

$$\begin{aligned} \min \quad & F(x) = (f_1(x), \cdots, f_m(x)) \\ \text{s.t} \quad & x \in \Omega \end{aligned} \tag{1}$$

- $x = (x_1, \cdots, x_n)^T \in R^n$ is a decision variable vector,
- $\Omega = \Pi_{i=1}^{n}[a_i, b_i] \subset R^n$ is the feasible region of the search space,
- $f_i : R^n \to R, i = 1, \cdots, m$, is a continuous mapping,
- $F(x)$ is an objective vector.

The algorithm framework of EDA

Initialization: Initial the *Pop*(t) randomly, and t is the generation.
while *not terminate* **do**
  Modelling: Build a probabilistic model $p(x)$ according to the statistical information of the *Pop*(t).
  Sampling: Generate a new solution set Q by sampling from the built probabilistic model $p(x)$.
  Selection: Select from $Q \cup Pop(t)$ to construct the next population $Pop(t+1)$. The selection criterion is the objective function value.
  $t = t + 1$
end

RM-MEDA

Under mild smoothness conditions, the PS of a continuous MOP is a piecewise continuous $(m-1)$-dimensional manifold (m is the number of objectives). This regularity property has been applied in the *regularity model based multiobjective estimation of distribution algorithm (RM-MEDA)*

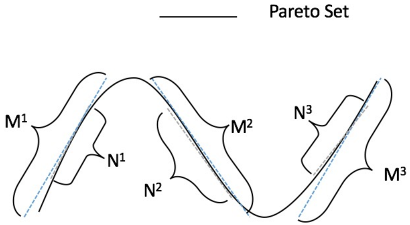


Figure 1: An illustration of model building and sampling in RM-MEDA.

RM-MEDA

EAST CHINA NORMAL UNIVERSITY

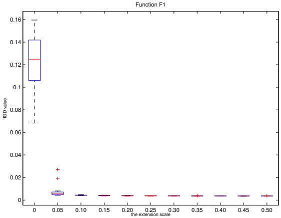
Algorithm 1: RM-MEDA Framework

Initialize a population $Pop(0)$, and set $t = 0$.
while *not terminate* **do**
    Modeling: Build the probabilistic model δ in order to model the distribution of the solutions in $Pop(t)$.
    Reproduction: Generate a new solution set Q from the probabilistic model.
    Selection: Select N solutions from $Q \bigcup Pop(t)$ to construct a new population $Pop(t+1)$.
    $t = t + 1$
end
Return the solutions in $Pop(t)$.

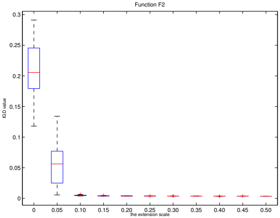
Figure 2: The framework of RM-MEDA

- t is the generation counter.
- $Pop(t) = \{x^1, x^2, \cdots x^N\}$ is the population at generation t.
- $\delta = \zeta + \varepsilon$ is the probabilistic model built in RM-MEDA.
- Q is a new solution set sampled from the probabilistic model.

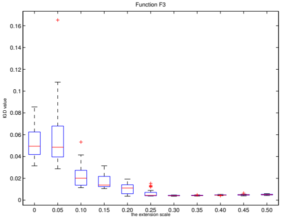
Extension scale in RM-MEDA



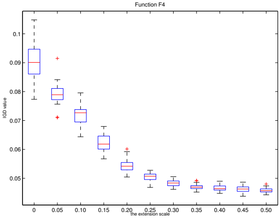
(a) F1



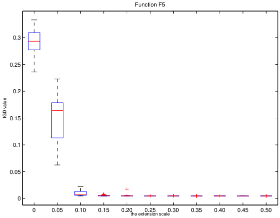
(b) F2



(c) F3

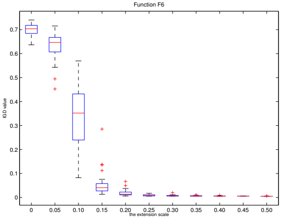


(d) F4

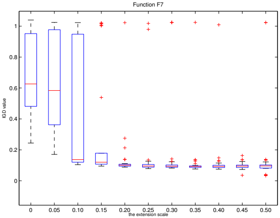


(e) F5

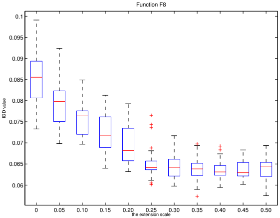
Extension scale in RM-MEDA



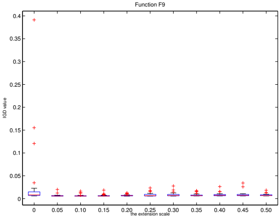
(f) F6



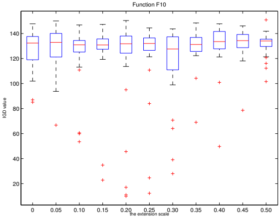
(g) F7



(h) F8



(i) F9



(j) F10

mutation scheme

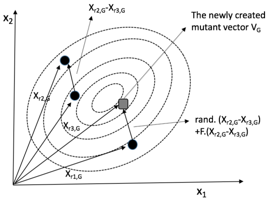


Figure 3: The schema of the mutation

$$X = X_{r_1} + rand \cdot (X_{r_2} - X_{r_3}) + F \cdot (X_{r_2} - X_{r_3}) \tag{2}$$

- X_{r_1}, X_{r_2}, X_{r_3} are sampled randomly from the population
- r_1, r_2 and r_3 are mutually exclusive integers selected randomly from 1 to NP
- F is the scaling factor

华东师范大学 EAST CHINA NORMAL UNIVERSITY

Algorithm 2: DES

1. Compute the covariance matrix C of the given cluster, and decompose it as
$$C = EDE^T$$
where E is the eigenvector matrix of C, and D is a diagonal matrix composed of eigenvalues.
2. For each solution x in the cluster, project it to a latent space as
$$y = x \cdot R.$$
where R is a matrix that contains the first $(m-1)$ components of the eigenvector matrix E.
3. Generate M points in the latent space as
$$y' = y_{r_1} + rand \cdot (y_{r_2} - y_{r_3}) + F \cdot (y_{r_2} - y_{r_3})$$
where r_1, r_2 and r_3 are integers randomly selected from $\{1, 2, 3, \cdots, M\}$, M is the size of the cluster.
4. Map y' back to the decision space.
$$x' = y' \cdot R^T.$$
5. Return the generated solution.
$$x'' = x' + \varepsilon'$$
where ε' is the Gaussian noise subjects to the distribution $\mathcal{N}(0, \sigma_\tau I)$ ($\tau \in \{1, 2, \cdots, K\}$ is a randomly generated integer).

Figure 4: The framework of DES

Algorithm 3: DES-RM-MEDA

Initialize a population $Pop(0)$, and set $t = 0$.
while *not terminate* **do**
    Modeling: Build the probabilistic model δ in order to model the distribution of the solutions in $Pop(t)$.
    Reproduction: Partition the population into different clusters C_i according to the probabilistic model. For each cluster, use DES to generate a set of candidate solutions Q_i. Set $Q = \cup_i Q_i$.
    Selection: Select N solutions from $Q \bigcup Pop(t)$ to construct a new population $Pop(t+1)$.
    $t = t + 1$
end
Return the solutions in $Pop(t)$.

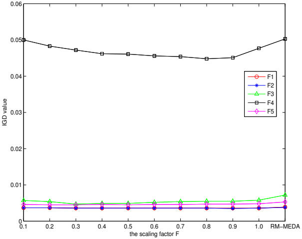
Figure 5: The framework of DES-RM-MEDA

Experimental settings

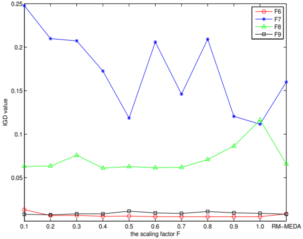
The 10 test instances, $F1 - F10$, introduced in RM-MEDA are used as the benchmark problems.

- *Initialization of the population*: The initial population in algorithms is randomly generated.
- *The number of new trial solutions generated*: 100 for instances ($F3$, $F7$, $F9$), and 200 for other instances.
- *The number of decision variables*: 30
- *The number of clusters*: 5
- *The scaling factor F*: 0.4
- *The number of runs*: 30
- *The number of generation*: 100 for instances ($F1$, $F2$, $F5$, $F6$), 200 for instances $F4$ and $F8$, and 1000 for instances $F3$, $F7$ and $F9$.

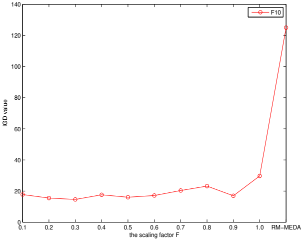
The sensitivity of the F



(a) F1-F5



(b) F6-F9



(c) F10

Wilcoxon's rank sum test

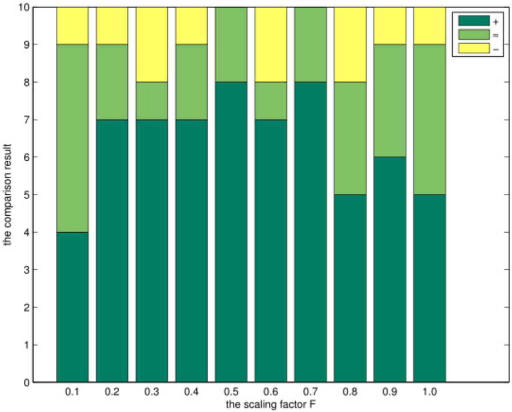
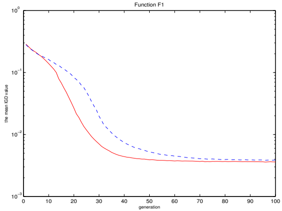
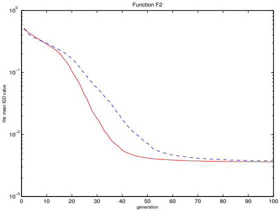


Figure 6: The compariosn of RM-MEAD and RM-MEDA-DES with different settings of F

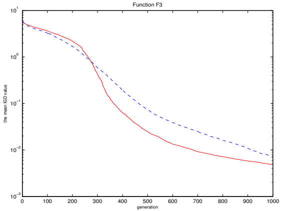
Comparison study



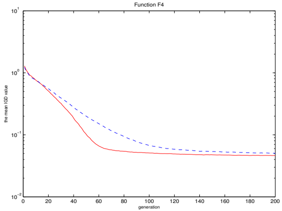
(a) F1



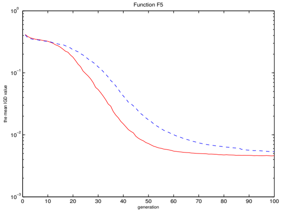
(b) F2



(c) F3



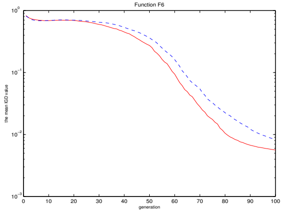
(d) F4



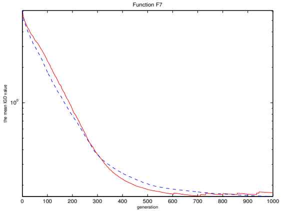
(e) F5

Figure 7: The dashed line is RM-MEDA, and the solid line is RM-MEDA-DES

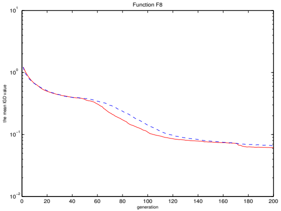
Comparison study



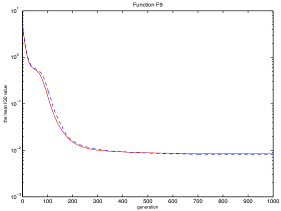
(a) F6



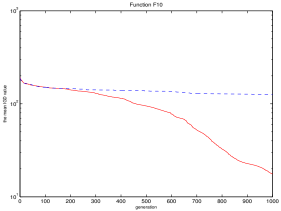
(b) F7



(c) F8



(d) F9



(e) F10

Figure 8: The dashed line is RM-MEDA, and the solid line is RM-MEDA-DES

Comparison study

EAST CHINA NORMAL UNIVERSITY

TABLE I
STATISTICAL RESULTS OF THE IGD VALUES OF THE FINAL POPULATIONS OBTAINED BY RM-MEDA AND DES-RM-MEDA ON THE 10 TEST INSTANCES OVER 30 RUNS.

instance	RM-MEDA				DES-RM-MEDA			
	mean	std.	best	worst	mean	std.	best	worst
$F1$	$3.90e-03$	$1.39e-04$	$3.70e-03$	$4.20e-03$	**3.60e-03**	$9.16e-05$	**3.43e-03**	**3.77e-03**
$F2$	$3.80e-03$	$1.43e-04$	$3.50e-03$	$4.10e-03$	**3.60e-03**	$1.01e-04$	**3.35e-03**	**3.82e-03**
$F3$	$7.20e-03$	$3.90e-03$	**3.60e-03**	$1.55e-02$	**4.90e-03**	$8.09e-4$	$3.90e-03$	**7.31e-03**
$F4$	$5.03e-02$	$1.30e-03$	$4.82e-02$	$5.35e-02$	**4.62e-03**	$9.32-04$	**4.44e-02**	**4.85e-02**
$F5$	$5.30e-03$	$3.00e-03$	$4.40e-03$	$2.12e-02$	**4.60e-03**	$1.49e-04$	**4.33e-03**	**4.96e-03**
$F6$	$8.30e-03$	$2.10e-03$	$5.70e-03$	$1.50e-02$	**5.60e-03**	$9.04e-04$	**4.52e-03**	**8.30e-03**
$F7$	**1.60e-01**	$2.35e-01$	$7.96e-02$	$1.03e+0$	$1.73e-01$	$2.28e-01$	**3.20e-02**	**1.02e+0**
$F8$	$6.59e-02$	$3.50e-03$	$6.05e-02$	$7.69e-02$	**6.10e-02**	$2.03e-03$	**5.67e-02**	**6.39e-02**
$F9$	**8.00e-03**	$2.80e-03$	$5.80e-03$	**1.48e-02**	$8.40e-03$	$3.20e-03$	**5.51e-03**	$2.12e-02$
$F10$	$1.25e+02$	$2.35e+01$	$2.27e+01$	$1.44e+02$	**1.76e+0**	$1.29e+01$	**4.73e+0**	**7.15e+01**

[1] The bolder ones mean better.

Conclusions

Main contributions

- A new sampling strategy for multiobjective estimation of distribution algorithm.
- Addressing to the issue of the extension scale setting in RM-MEDA.

This paper proposed a DES scheme to generate points in the latent space. The basic idea is to project the parent solutions into the latent space, and use a DE mutation operator to generate new points in the latent space based on the projected points, and finally map the points back to to the decision space added with Gaussian noise to generate offspring solutions. The DES is implemented into RM-MEDA to improve the performance. The results are impressive.

The future work

- To exploit the deeper application of the DES. There is the possibility to apply the DES to other MOEAs.
- It is valuable to explore the hybrid menthod of the DE and EDA. Though DE/EDA has been proposed, it is still interesting to explore the potential of this method. And the arrange of the resources of DE and EDA is also an interesting topic.

Thanks!

- B. Dong, A. Zhou, and G. Zhang, Sampling in Latent Space for a Multiobjective Estimation of Distribution Algorithm, 2016 IEEE Congress on Evolutionary Computation (CEC), 2016.