分布估计算法中差分采样策略的研究

董兵
周爱民　副教授

华东师范大学
计算机科学软件工程学院

January 2017

华东师范大学 EAST CHINA NORMAL UNIVERSITY

演化算法是一种基于种群的启发式优化算法，是隶属于演化计算的人工智能算法。

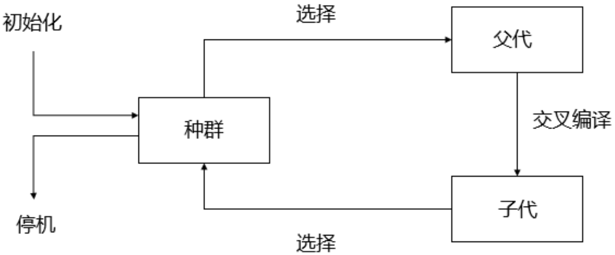


Figure 1: 演化算法流程图

分布估计算法是一种新型的演化算法，传统的分布估计算法主要由三个步骤组成：建模、采样、选择。
本文的研究目的是利用**差分采样策略**来取代分布估计算法传统的建模采样，从而提高算法的性能。

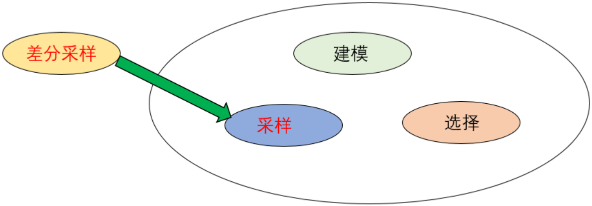


Figure 2: 分布估计算法

- 受差分进化启发，提出一种基于差分进化的采样策略，即差分采样策略。
- 对于单目标优化问题，利用基于特征向量的差分进化去改进采样，通过expensive Local Search进一步改进解集质量，提出基于差分采样的单目标分布估计算法。
- 对于多目标优化问题，利用差分采样策略去改进RM-MEDA的采样，提出基于差分采样的多目标分布估计算法。

华东师范大学 EAST CHINA NORMAL UNIVERSITY

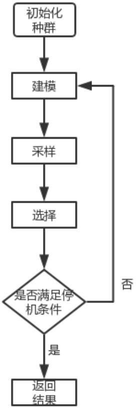


Figure 3: 分布估计算法流程

- 分布估计算法主要有三个主要步骤：建模、采样、选择
- 传统的分布估计算法通过从建立的概率模型中采样产生新的个体

研究现状:

根据问题的不同类型，可以将分布估计算法分成以下几类:

- 基于离散变量的分布估计算法
- 基于遗传编程的分布估计算法
- 多目标分布估计算法

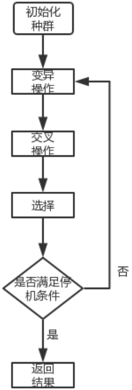


Figure 4: 差分进化流程

- 变异操作：产生变异向量
  常用的变异策略：
  - DE/rand/1
  - DE/best/1
  - DE/rand/2
  - DE/best/2
  - DE/current-to-best/1
  - DE/current-to0rand/1
- 交叉操作：利用交叉算子结合变异向量和目标向量来产生实验向量。
- 选择：利用一对一的竞争机制从目标向量和试验向量中挑选个体。

差分进化自从提出后，就受到了工业界以及学业界的广泛专注。根据差分进化解决的问题类型，可以将差分进化分为以下几种类型：

- 针对连续单目标优化问题的差分进化算法
- 针对约束优化的差分进化算法
- 针对多目标优化问题的差分进化算法

单目标优化问题几乎出现在科学和工程应用的各个领域中。单目标优化问题的目标函数一般是非凸函数，并且在可行区域内具有很多的局部极小值或者极大值。本文研究的单目标优化问题针对的是连续空间的全局优化问题，即是求目标函数的最小值或者最大值。
对于全局优化问题在本文做出以下定义：

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

其中$x = (x_1, x_2, \cdots, x_n)^T \in R^n$ 是决策变量向量，$[a_i, b_i]^n$是搜索空间区域，$f : R^n \to R$ 则是目标函数。

```
初始化种群Pop(t) = {x_1, x_2, x_3, ··· , x_N} (N是种群大小)
while not terminate do
    v_{i,G} = x_{r1,G} + F · (x_{r2,G} − x_{r3,G})
    if rand() < p then
        if rand() ≤ CR then
            u_{i,G} = v_{i,G}
        else
            u_{i,G} = x_{i,G}
        end
    else
        求得x_{i,G}的特征向量矩阵E，令E'为特征向量矩阵的逆矩阵。
        x'_{i,G} = E' · x_{i,G}
        v'_{i,G} = E' · v_{i,G}
        if rand() ≤ CR then
            u'_{i,G} = v'_{i,G}
        else
            u'_{i,G} = x'_{i,G}
        end
        u_{i,G} = E · u'_{i,G}
    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
    t = t + 1
end
```

Figure 5: DE-EIG算法

- DE-EIG算法是基于特征向量的差分进化算法，其主要贡献是在一个旋转的坐标空间对个体进行交叉操作，这样可以利用旋转空间中种群的协方差矩阵的特征向量信息。
- 这样在保证种群的多样性的同时，能够有效地引导种群向全局最优演化。

在可行的搜索空间内构建一个随机种群$Pop(t)$
while *not terminate* **do**
　构建概率模型：
　$p_k(x) = \prod_{i=1}^{n} \mathcal{N}(x_i; \mu_i, \sigma_i)$
　对于所有$j = 1, 2, \cdots, n$，产生一个试验向量$u = (u_1, u_2, \cdots, u_n)$
　if $rand() < CRP$ **then**
　　$u = \frac{(x_i)_j + (x_d)_j}{2} + F \cdot [(x_d)_j - (x_i)_j + (x_b)_j - (x_c)_j]$
　else
　　u根据$\mathcal{N}(x_i; \mu_i, \sigma_i)$采样产生
　end
　if $f(u) < f(x_i)$ **then**
　　$x_i^{t+1} = u$
　else
　　$x_i^{t+1} = x_i^t$
　end
　$t = t + 1$
end

Figure 6: DE/EDA算法

- DE/EDA是一种结合差分进化和分布估计算法来解决全局连续优化问题。
- DE/EDA通过利用分布估计算法可以提取种群全局信息和差分进化可以提取种群查分信息的优点，是一种非常具有研究前景的算法。

expensive Local Search

利用 $Converage(\theta, t, t_e)$ 函数来判断解是否收敛

$$\Delta f = \frac{|f^1_{t-50} - f^1_t|}{max\{|f^1_{t-50}|, |f^1_t|\} + \varepsilon} \quad (2)$$

$$\Delta x = \frac{|c_{t-50} - c_t|}{max\{c_t, c_{t-50}\} + \varepsilon} \quad (3)$$

- Δf 表示在近50代中最好的评价值的比率的降低
- $f^1_t = min_{x \in pop_t}$ 是在代数t时最好的目标函数值。
- Δx 表示在近50代中种群覆盖区域的变化比率
- $c_t = \frac{1}{n}\sum{}^n{}_{i=1}(max_{x \in pop} - min_{x \in pop})$，$\varepsilon = 1.0 \times 10^{-50}$。
- 通过将 $min\{\Delta f, \Delta x\}$ 和给定的阈值 θ 进行比较，皆可以判断解是否收敛。为了提高算法的运行效率，每两次expensive LS至少间隔50代。

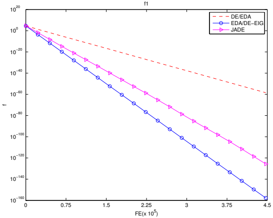
初始化种群$Pop(t) = \{x_1, x_2, x_3, \cdots, x_N\}$ (N是种群的大小)
while *not terminate* **do**
  构建概率模型：
  $p(x) = \prod_{i=1}^{n} \mathcal{N}(x_i; \mu_i, \sigma_i)$
  根据下面的流程产生新的实验向量$u_{i,G}$：
  if $rand() < CRP$ **then**
    根据DE-EIG采样得到$u_{i,G}$
  else
    根据概率模型$p(x)$采样得到$u_{i,G}$
  end
  if $f(u_{i,G}) < f(x_{i,G})$ **then**
    $x_{i,G+1} = u_{i,G}$
  else
    $x_{i,G+1} = x_{i,G}$
  end
  if $Coverage(\theta, G, G_e)$ **then**
    执行expensive LS
  end
  $t = t + 1$
end

Figure 7: 基于差分采样的单目标分布估计算法

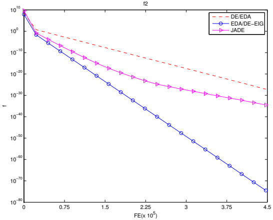
- 为了增加种群的多样性，通过利用随机参数来设置*CRP*，可以提高算法对于大多数问题的鲁棒性。
- 在基于DE/EDA算法框架的基础上，利用DE-EIG进行采样，同时结合expensive LS进一步提高解集质量，提出基于差分采样的单目标分布估计算法。

实验设置：

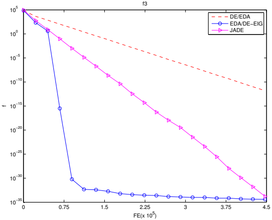
- 测试题中所有种群维度都设置为30.所有算法都会在每一个测试题上独立运行50 次，体积条件是450000函数评估。
- JADE：参数设置为：$N = 150, p = 0.05, c = 0.1, F = 0.5$ and $CR = 0.9$
- DE/EDA: $N = 150, F = 0.5$ and $CRP = 0.9$。
- EDA/DE-EIG: $CRP = 0.5, F = 0.5, CR = 0.6, \theta = 0.1$；控制坐标旋转的参数$p$ 设置为0.5；种群的大小N设置为150。对于expensive LS的相关参数的设置，与EDA/LS中的参数设置相同。



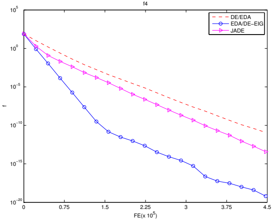
(a) $f1$



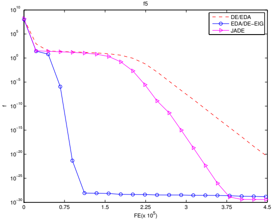
(b) $f2$



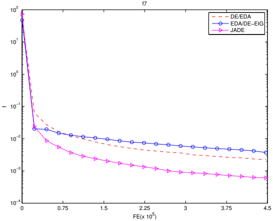
(c) $f3$



(d) $f4$

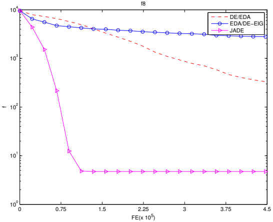


(e) $f5$

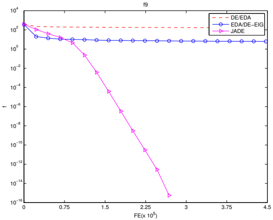


(f) $f7$

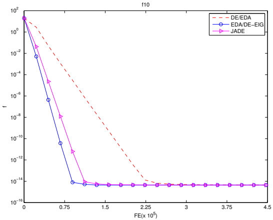
Figure 8: 在$f1 - f13$中除了$f6$的12个测试题上目标函数平均值的折线图



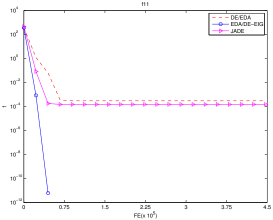
(a) $f8$



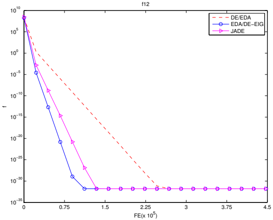
(b) $f9$



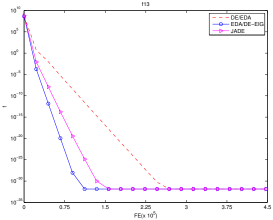
(c) $f10$



(d) $f11$



(e) $f12$



(f) $f13$

Figure 9: 在$f1-f13$中除了$f6$的12个测试题上目标函数平均值的折线图

	EDA/DE-EIG	JADE	DE/EDA
$f1$	**1.54e-159 ± 5.11e-159**	3.90e − 127 ± 2.74e − 126(+)	1.39e − 59 ± 2.58e − 59(+)
$f2$	**1.02e-75 ± 7.46e-76**	2.60e − 35 ± 1.64e − 34(+)	5.15e − 28 ± 4.68e − 28(+)
$f3$	**4.01e-35 ± 8.47e-35**	7.79e − 35 ± 2.51e − 34(∼)	1.23e − 12 ± 1.20e − 12(+)
$f4$	**5.01e-20 ± 3.06e-19**	3.15e − 14 ± 6.42e − 14(+)	9.90e − 12 ± 2.69e − 11(+)
$f5$	1.46e − 29 ± 2.62e − 29	**3.85e-30 ± 9.58e-30(−)**	3.37e − 21 ± 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 ± 1.00e − 03	**6.01e-04 ± 2.23e-04(−)**	2.20e − 03 ± 5.59e − 04(−)
$f8$	2.79e + 03 ± 5.02e + 02	**4.74e+00 ± 2.34e+01(−)**	1.82e + 03 ± 6.72e + 02(−)
$f9$	6.23e + 00 ± 2.21e + 00	**0.00e+00 ± 0.00e+00(−)**	1.54e + 02 ± 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 ± 1.05e − 03(∼)	2.96e − 04 ± 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(−)

Figure 10: 基于差分采样的单目标分布估计算法

华东师范大学 EAST CHINA NORMAL UNIVERSITY

现实生活中的问题往往具有多个目标，而每个目标之间的最优值往往是冲突的。在本文中，假设多目标问题中的每个问题都是最小化问题，则做出以下定义：

$$\begin{aligned} \min \quad & F(x) = (f_1(x), \cdots, f_m(x)) \\ \text{s.t} \quad & x \in \Omega \end{aligned} \tag{4}$$

其中，$x = (x_1, \cdots, x_n)^T \in R^n$ 是决策变量向量，$\Omega = \Pi_{i=1}^{n}[a_i, b_i] \subset R^n$ 表示可能的搜索空间区域，$f_i : R^n \to R, i = 1, \cdots, m$ 是一个连续的目标函数，$F(x)$则是相应的目标函数向量。

在多目标优化问题中，多个目标相互之间往往是冲突的，从而导致无法在满足所有约束条件下使得所有目标函数都能够达到全局最优解，但是存在一组Pareto最优解 [?]。对于此，做出以下定义：令$a, b \in R^n$，当$a_i \leq b_i \wedge a \neq b$，且$i = 1, \cdots, n$，则称$a$ 支配b。向量$x^* \in \Omega$ 即是公式 4的Pareto最优解，如果不存在$x \in \Omega$使得$F(X)$ 支配$F(x^*)$。$F(x^*)$被称为Pareto最优目标向量。所有的Pareto 最优解的集合就是Pareto最优解集（PS），对应的最优向量的集合则成为Pareto前端（PF）。

在一般条件下，根据Karush-Kuhn-Tucker 可以推导出：连续多目标问题在决策空间中的Pareto set 是一个连续分段的（m-1）维的流形体（m是目标数）。对于一个成功的多目标演化算法（multiobjective evolutionary algorithm, MOEA）来说，独立的个体应该是在决策空间中分散在Pareto set 附近，如图 11 所示。

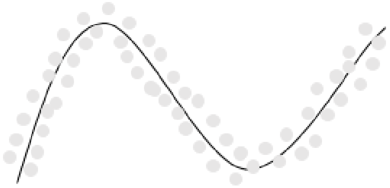


Figure 11: 连续多目标问题决策空间中个体的分布情况

对于连续多目标问题，在决策空间中，种群体中的个体如果越接近Pareto set，则越容易进行问题的求解。因此，假设种群中的个体为随机向量$\xi \in R^D$ 的观测值，ξ的中央部分就是Pareto set。并且因为在连续多目标问题中，Pareto set是一个m-1维的流体，那么ξ 则可以由公式 5表示：

$$\xi = \zeta + \epsilon \tag{5}$$

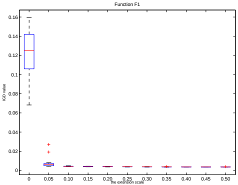
ζ相当于是均匀分布在m-1维流体附近的个体，ϵ是均值为0的n维的噪音向量。

华东师范大学 EAST CHINA NORMAL UNIVERSITY

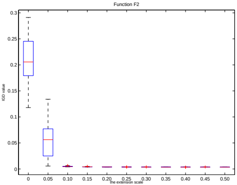
初始化一个随机种群$Pop(0)$，并且设置$t=0$。
while 没有达到停机条件 **do**
　建模： 建立一个概率模型ξ来表示在随机种群$Pop(t)$中的个体。
　采样： 通过上述的概率模型进行采样得到新的解集Q。
　选择： 从$Q\bigcup Pop(t)$中挑选出N个个体来组成一个新的种群$Pop(t+1)$。
　$t=t+1$
end

Figure 12: RM-MEDA

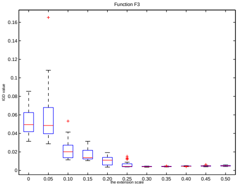
- 根据连续多目标问题以上的特性，基于规律模型的多目标分布算法（RM-MEDA）算法被提出用于解决连续多目标问题。
- 在每一次迭代中，通过局部主成分分析（LPCA）在决策空间中的区域建立概率分布模型，然后通过使用拉丁实验设计采样得到新的子代种群。
- RM-MEDA 采用基于非劣排序的方法来挑选个体来产生新的种群。



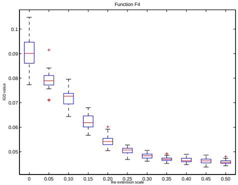
(a) F1



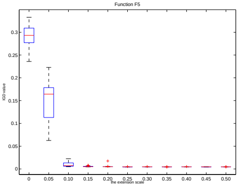
(b) F2



(c) F3

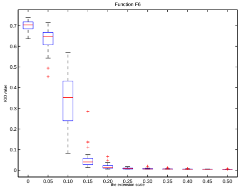


(d) F4

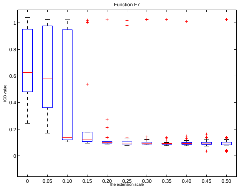


(e) F5

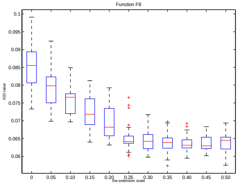
Figure 13: 不同缩放因子下，RM-MEDA的IGD指标的箱线图



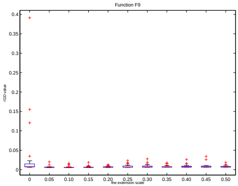
(a) F6



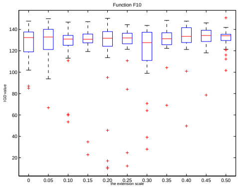
(b) F7



(c) F8



(d) F9



(e) F10

Figure 14: 不同缩放因子下，RM-MEDA的IGD指标的箱线图

华东师范大学 EAST CHINA NORMAL UNIVERSITY

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.
- 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.