

基于参考点选择策略的改进型 NSGA-III 算法

耿焕同^{1,2} 戴中斌¹ 王天雷¹ 许 可¹

摘 要 针对多目标进化算法忽视种群在决策空间的分布信息,未考虑待优化问题 Pareto 前沿形状的问题,文中提出基于参考点选择策略的改进型 NSGA-III 算法.首先,根据种群在决策空间的分布特征,借助信息论中的熵思想,计算相邻两代种群的熵差,判定种群的进化阶段.然后,根据种群在目标空间的分布特征,借助参考点关联个体数目的统计信息,评估参考点的重要性.最后,在种群进化的中后期,依据参考点的重要性特征剔除冗余的无效参考点,使保留的参考点适应种群规模与 Pareto 前沿面,利用筛选后的参考点引导种群进化方向,加快算法收敛及优化效率.在测试函数集上的对比实验表明,文中算法在收敛性和分布性上均较优.

关键词 多目标优化,参考点,决策空间分布,目标空间分布

引用格式 耿焕同,戴中斌,王天雷,许 可.基于参考点选择策略的改进型 NSGA-III 算法.模式识别与人工智能, 2020, 33(3): 191–201.

DOI 10.16451/j.cnki.issn1003-6059.202003001

中图法分类号 TP 183

Improved NSGA-III Algorithm Based on Reference Point Selection Strategy

GENG Huantong^{1,2}, DAI Zhongbin¹, WANG Tianlei¹, XU Ke¹

ABSTRACT The traditional multi-objective evolutionary algorithm ignores distribution information of the population in the decision space and Pareto front shape of the optimization problem is not taken into account. To solve the problems, an improved NSGA-III algorithm based on reference point selection strategy is proposed. Firstly, according to the entropy thought in information theory, the entropy difference between two adjacent generations is calculated in line with the distribution characteristics of the population in the decision-making space, and the evolutionary status of the population is determined. Then, in the light of the distribution characteristics of the population in the target space, the importance of reference points is evaluated via statistical information of the number of the individuals associated with reference points. Finally, redundant and invalid reference points are removed according to the importance characteristics of reference points in the middle and late period of population evolution. Reserved reference points can adapt to the population size and Pareto frontier, and the selected reference points are exploited to guide the evolution direction of the population and accelerate the convergence and optimization efficiency. Experiments on test function sets indicate the significant advantages of the proposed algorithm in convergence and distribution.

Key Words Multi-objective Optimization, Reference Point, Decision Space Distribution, Target Space Distribution

收稿日期: 2020-01-03; 录用日期: 2020-03-13

Manuscript received January 3, 2020;

accepted March 13, 2020

南京大气科学联合研究中心(No.NJCAR2018MS05)、国家自然科学基金项目(No.51977100) 资助

Supported by Nanjing Joint Center of Atmospheric Research(No. NJCAR2018MS05), National Natural Science Foundation of

China(No.51977100)

本文责任编辑 何清

Recommended by Associate Editor HE Qing

1.南京信息工程大学 计算机与软件学院 南京 210044

2.江苏省气象局 南京大气科学联合研究中心 南京 210009

1.School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044

2.Nanjing Joint Center of Atmospheric Research, Meteorological Bureau of Jiangsu Province, Nanjing 210009

Citation GENG H T, DAI Z B, WANG T L, XU K. Improved NSGA-III Algorithm Based on Reference Point Selection Strategy. Pattern Recognition and Artificial Intelligence, 2020, 33(3): 191-201.

多目标优化问题 (Multi-objective Optimization Problem, MOP) [1] 由两个及以上相互冲突的优化目标组成, 往往一个目标性能得到提升时会导致其它数个目标的性能下降。当问题的目标数大于 3 时, 称为高维多目标优化问题 (Many-Objective Optimization Problem, MaOP)。在高维空间中, 非支配个体的比例会随目标数的增加而迅速上升, 难以评估个体间的优劣关系, 这成为当前多目标进化算法领域公认的研究难点与热点。近些年, 学者们提出许多高维多目标进化算法 (Many-Objective Optimization Evolutionary Algorithm, MaOEA), 按其进化机制的不同, 大致可分为如下 3 类。

第一类为基于支配关系的 MaOEA [2-6]。传统 Pareto 支配关系在高维问题中丧失对个体的选择压力, 为此, 研究者们提出许多新型的强支配关系。Yang 等 [3] 提出基于网格的方法, 用于选择优先级更高的优势解, 通过调整网格大小控制 Pareto 最优解的比例。Zhang 等 [4] 提出基于凹点驱动的进化算法 (Knee Point-Drive Evolution Algorithm, KnEA), 以二次选择的方式提高选择压力。Zhu 等 [5] 扩大支配区域。Tian 等 [6] 提出基于目标向量角度的小生境技术, 判断不同解之间的支配关系。上述方法都是在整个进化过程中保持对 Pareto 前沿的选择压力, 为 MaOP 寻找帕累托最优集。

第二类为基于分解的 MaOEA [7-13]。将待求解的多目标问题分解为若干个目标数更低的多目标问题或单目标问题, 用于协同解决。Zhang 等 [7] 基于分解思想, 结合数学规划方法和进化算法, 将多目标优化问题转化为一组单目标优化问题, 提出基于分解的多目标进化算法 (Multi-objective Evolutionary Algorithm Based on Decomposition, MOEA/D)。继而, 研究者们提出多种 MOEA/D 的变体 [8-11], 进一步提升算法性能。

第三类为基于指标的 MaOEA [14-17]。利用算法性能评价指标指导个体选择和种群搜索。反转世代距离 (Inverted Generational Distance, IGD) 和超体积 (Hyper Volume, HV) 是此类多目标进化算法 (Multi-objective Evolutionary Algorithm, MOEA) 常用的两个指标 [14]。Tian 等 [15] 提出基于指标的参考点自适应多目标进化算法 (Indicator Based MOEA with Reference Point Adaption, AR-MOEA), 采用基于增强的

IGD 指标 (IGD-NS) 作为主要的环境选择策略, 利用一组均匀分布的参考点作为计算增强的 IGD 指标的参考点集。Sun 等 [16] 提出基于 IGD 的 MaOEA (MaOEA-IGD), 在解决 MaOP 方面性能良好。黎明等 [17] 提出基于改进的 IGD (Improved IGD, IGD⁺S) 指标的高维多目标进化算法 (MOEA Based on IGD⁺S Indicator, MOEA/IGD⁺S), 采用 IGD⁺S 指标, 在 IGD 指标的基础上提出无贡献个体的概念。这类算法通常比其它两类 MaOEA 更有效, 但缺点是计算开销较大。

尽管上述算法能有效解决多目标优化问题, 但是, 算法的性能评估指标都是由最终种群的目标值计算得到, 且目前大多数算法仅从目标空间分析种群, 并设计相应策略提高优化效果, 忽视从决策空间分析种群。Li 等 [18] 指出 MOEA 的性能依赖于待解决问题的 Pareto 前沿的形状, 即大多数 MOEA 的性能对 Pareto 前沿形状敏感。部分 MOEA 只能处理常规和有针对性的 Pareto 前沿问题, 难以处理复杂的不规则 Pareto 前沿的问题。

为了解决上述两个问题, 本文提出基于参考点选择策略的改进型 NSGA-III 算法 (Improved NSGA-III Algorithm Based on Reference Point Selection Strategy, AR-NSGA-III)。利用种群在决策空间分布特征信息, 根据四分位差统计量可以客观反映数据在空间中离散程度的原理, 通过计算相邻两代熵值之差判别种群的进化阶段, 将种群在进化过程中划分为“探索”与“探究”两个阶段。利用种群在目标空间分布特征信息, 伴随种群进化, 个体会逐渐偏向关联穿过真实 Pareto 前沿面的参考线的特性。在种群的“探索”阶段, 统计参考点关联每代种群个体的总数目, 评估参考点的重要性, 保留重要性较高的参考点, 剔除冗余无效参考点, 使参考点适应种群规模与 Pareto 前沿面。

1 NSGA-III 及问题分析

1.1 多目标优化问题

不失一般性, 以最小化问题为例, 一个由 n 个决策变量 m 个目标函数 p 个不等式约束 q 个等式约束组成的多目标优化问题定义如下:

$$\begin{aligned} \min y &= F(\mathbf{x}) = \{f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})\}, \\ \text{s.t. } g_i(\mathbf{x}) &\leq 0 \quad i = 1, 2, \dots, p, \\ h_j(\mathbf{x}) &= 0 \quad j = 1, 2, \dots, q. \end{aligned}$$

其中: $\mathbf{x} = (x_1, x_2, \dots, x_n) \in X$ 称为 n 维决策向量, X 为 n 维决策空间; $\mathbf{y} = (y_1, y_2, \dots, y_m) \in Y$ 称为 m 维目标向量, Y 为 m 维目标空间. $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})$ 为相互冲突的 m 个优化目标; $g_i(\mathbf{x}) \leq 0$ 定义 p 个不等式约束, $h_j(\mathbf{x}) = 0$ 定义 q 个等式约束. 若 $m > 3$, 该问题称为高维多目标优化问题.

1.2 基于参考点的非支配排序进化算法

非支配排序遗传算法 II (Nondominated Sorting Genetic Algorithm-II, NSGA-II) 仅能较好处理低维多目标问题 (目标数 $M \leq 3$). 究其原因, 随着待优化问题目标维数的增加, Pareto 支配关系的选择压力急剧下降, 种群中的非支配个体呈指数增长, 个体间的优劣难以区分. 为此, Deb 等^[19] 提出基于参考点的非支配排序进化算法 (Reference-Point Based Many-Objective, NSGA-III), 求解高维多目标问题.

算法 1 NSGA-III

输入 t 代种群 P_t , 参考点 Z , 种群规模 N

输出 $t+1$ 代种群 P_{t+1}

1. $Q_t = \text{make_new_pop}(P_t)$; // 生成子代种群
2. $R_t = P_t \cup Q_t$; // 父代子代合并
3. $F = \text{nondominated_sort}(R_t)$; // 快速非支配排序
4. $P_{t+1} = \emptyset$;
5. $i = 1$;
6. while($|P_{t+1}| + |F_i| \leq N$) {
7. $P_{t+1} = P_t \cup F_i$;
8. $i = i + 1$;
9. }
10. 标准化目标空间;
11. 根据参考点 Z 计算参考线;
12. 个体与参考线进行关联操作;
13. 统计每条参考线关联的个体数;
14. $k = N - |P_{t+1}|$;
15. 根据参考线关联的个体数, 从临界非支配层 F_i 中选择 k 个个体进入新一代种群 P_{t+1} 中;

NSGA-III 与 NSGA-II 总框架十分相似, 唯一不同点在于精英策略中临界非支配层的选择. 为了保证种群分布性, NSGA-III 采用参考线关联个体的方法, NSGA-II 引入个体拥挤度的概念.

1.3 NSGA-III 中参考点存在的问题

NSGA-III 采用边界交叉构造权重的方法构造参考点. 在 M 维的标准化超平面上, 每维目标的划分为 p 均匀产生 $H(M, p)$ 个参考点:

$$H(M, p) = \frac{(M + p - 1)!}{p! \cdot (M - 1)!}. \quad (1)$$

如图 1 所示, $M = 3, p = 13$, 参考点数 $H(3, 13) = 105$. NSGA-III 在环境选择过程中, 除了利用 Pareto 支配关系, 还强调各参考线在种群中关联的个体数. 参考线均匀贯穿整个标准化超平面, 被选择的种群个体均匀分布在真实的 Pareto 前沿面上.

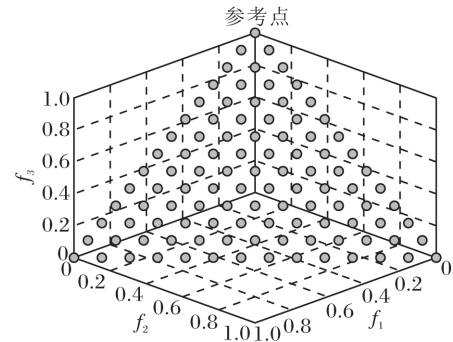
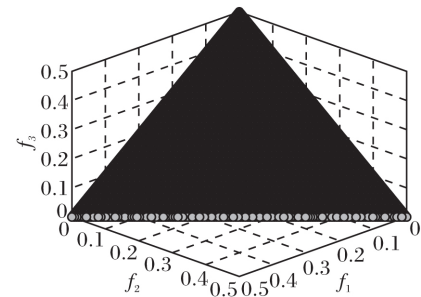


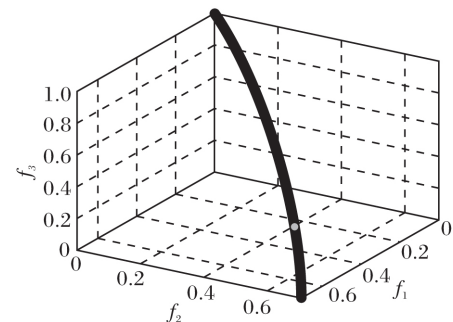
图 1 参考点分布情况

Fig.1 Reference points distribution

在实际问题中, Pareto 前沿面并非都是均匀和连续分布的. 图 2 为测试函数 DTLZ1、DTLZ5 的 Pareto 前沿面 ($M = 3$ 时).



(a) DTLZ1



(b) DTLZ5

图 2 不同测试函数的 Pareto 前沿面

Fig.2 Pareto fronts for different test functions

NSGA-III 中构造的均匀分布在超平面上的参考点可用于较好地求解 DTLZ1、UF8^[20]、WFG1^[21] 等问题,这是因为此类 Pareto 前沿面是均匀连续分布的.而处理 DTLZ5、UF9、WFG3 等问题时,此类 Pareto 前沿面不连续,导致效果欠佳.与此同时,对于不同的问题,每个参考点对个体的选择并非具有同等重要性.

NSGA-III 通过 Pareto 支配关系及个体与参考线的关联进行环境选择,均衡种群个体的收敛性与分布性.然而,参考点数 H 与目标数 M 、划分 p 呈指数级关系,尤其在多维多目标问题中,划分 p 的少量增加也会导致参考点数目急剧上升.种群规模在进化中是恒定的,不当的划分会产生大量冗余的参考点,影响算法性能与效率.

NSGA-III 中参考点的数目决定算法的运行时间,呈正相关.算法最终进化的种群并不会随参考点的增加而变优,过多、过少的参考点都不利于临界非支配层的选择.不妨选取两个特殊情况进行分析,当参考点数趋近 $+\infty$ (所有个体与最近参考线的距离必定为 0) 时和当参考点数为 1 (所有个体必定关联到该参考点上) 时,参考点筛选群个体的功能丧失.由此可见,参考点的选择对算法性能具有重要影响.

1.4 种群分布特征的定量分析

种群进化过程中仅考虑目标空间对个体进行选择,具有局限性.为了方便问题分析,以单目标最大化问题为例.如图 3 所示, X 为决策空间, Y 为目标空间. A 个个体 A, B, C, D , 其中 $y_A = y_B = y_C = y_D$. 从 $\{A, B, C, D\}$ 种群中选择 2 个个体作为子代种群.若仅从目标空间分析种群,子代种群 $\{A, B\}$ 、 $\{C, D\}$ 、 $\{A, C\}$ 、 $\{B, D\}$ 等价.而实际上子代种群 $\{A, C\}$ 、 $\{B, D\}$ 分布性明显优于子代种群 $\{A, B\}$ 、 $\{C, D\}$. 种群 $\{A, B\}$ 在进化过程中极有可能陷入局部最优解.

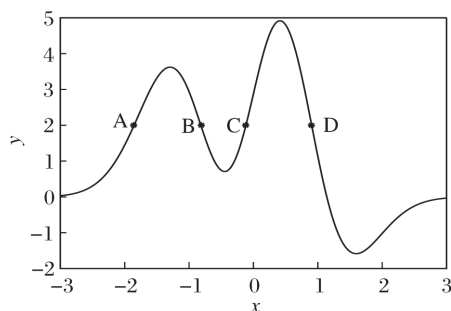


图 3 单目标问题上的个体

Fig.3 Individuals on single objective problem

同理,在多目标进化算法中,决策空间映射于目标空间,个体可以由决策向量映射到目标向量,但仅

有目标向量无法映射到决策向量.因此,在种群进化过程中同时从目标空间和决策空间分析种群,可更好地保持种群的收敛性与分布性.

统计学上的四分位差 (Inter-Quartile Range, IQR) 反映数列中间 50% 数据的离散程度,由有序数列的第 3/4 位 (上四分位 Q_3) 与第 1/4 位 (下四分位 Q_1) 相减得到,有效避免极值的干扰,数值越大表示数据离散度越高.通过计算四分位差与空间上下界差 ($ub - lb$) 的比值,实现数值的归一化.具体计算如下^[22]:

$$inf = \frac{Q_3 - Q_1}{ub - lb}.$$

例 1 二维数组

$\{(0.59, 0.2), (0.63, 0.3), (0.57, 0.5), (0.63, 0.6)\}$

由 4 个点组成, x 轴上的上四分位数 $X_{Q_3} = 0.63$, 下四分位数 $X_{Q_1} = 0.59$, 该数组在 X 轴上的四分位差 $iqr_x = 0.04$. 同理,在 Y 轴上的四分位差 $iqr_y = 0.2$. 若定义 X, Y 轴上界 $ub = 1$, 下界 $lb = 0$, 则

$$inf_x = 0.04, inf_y = 0.2,$$

两者的巨大差距直接反映该数列在 X, Y 轴上分布性的巨大差异.然而,四分位差难以反映数列的整体平移,为此本文引入标准化中位数差的概念:

$$\Delta mid^t = \frac{|mid^t - mid^{t-1}|}{ub - lb},$$

Δmid^t 表示数组 x 在前后两代 $t-1$ 和 t 的中位数差值与空间上下界差的比值.

2 基于参考点选择策略的改进型 NSGA-III

针对 NSGA-III 未能利用种群决策空间信息,参考点划分难以确定,对不同形状 Pareto 前沿自适应性不够等问题,本文提出基于参考点选择策略的改进型 NSGA-III 算法 (AR-NSGA-III). 通过种群在决策空间的四分位差分布特征信息判别种群的进化阶段,通过种群在目标空间的分布特征进行超平面上参考点的选择.

2.1 基于决策空间的进化阶段判定策略

在种群进化过程中,种群由无序变成有序,逐渐收敛,这其实是一个熵减的过程.本文采用相邻两代熵值之差描述种群的进化阶段^[22].熵值 e^t 由种群的标准化的四分位差 inf_i 和标准化中位数差 Δmid_i^t 计算求得,定义如下:

$$e^t = -\sum_{i=1}^n inf_i \lg inf_i - \sum_{i=1}^n \Delta mid_i^t \lg \Delta mid_i^t, \text{cnki.net}$$

相邻两代熵值之差

$$\Delta e^t = |e^t - e^{t-1}|. \quad (2)$$

当种群更新时 e^t 放大决策空间中种群的变化。 Δe^t 值越大,表明种群中个体的变化越大,说明种群正在积极探索未知的潜在解空间。反之 Δe^t 值越小,表明种群越稳定,种群正在收敛。因此,可使用 Δe^t 判定种群的进化状态 S^t ,其中 t 为种群当前进化次数, μ 为阈值。

若 $|\Delta e^t| > \Delta\mu$,表明算法处于“探索”阶段,种群正在探索解空间。若 $|\Delta e^t| < \mu$,表明算法进入“探究”阶段,种群可能开始收敛。

进化阶段 S^t 的判定依赖阈值 μ 的设定十分关键。本文阈值 μ 的设定依据为:对于任一维有限区间 E ,最理想的情况是种群中 N 个个体均匀分布在该区间内,此时种群个体与相邻个体的最短距离为 $|E|/N$,标准化后为 $1/N$ 。

计算个体均匀分布时的标准化四分位差:

$$\overline{inf} = \frac{0.75E - 0.25E}{E} = \frac{1}{2}.$$

选取 \overline{inf} 的变化小于 $1/N$ 且种群未发生偏移 $\Delta mid \approx 0$ 时,根据式(2)得到 $|\Delta e|$ 作为阈值 μ ,并作为参照。进一步考虑阈值 μ 与决策空间维数存在线性关系,最终设计

$$\mu = -D \overline{inf} \lg \overline{inf} - D \left(\overline{inf} + \frac{1}{N} \right) \lg \left(\overline{inf} + \frac{1}{N} \right),$$

其中 D 为决策空间维数, N 为种群规模。

以标准测试函数 ZDT2 为例,设种群规模 $N = 100$, ZDT2 的决策空间维数 $D = 30$,则

$$\mu = 30 \left| -0.51 \lg 0.5 + 0.51 \lg 0.51 \right| \approx 0.0413.$$

图4为 NSGA-III 在 ZDT2 函数上进化 300 代的熵值 e 和熵差 Δe 变化趋势。

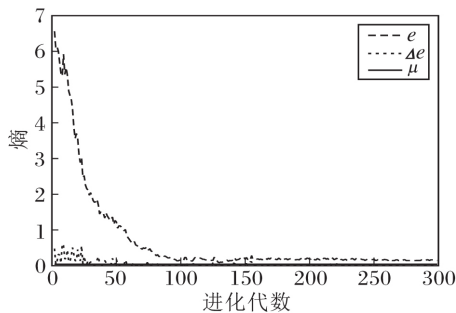


图4 NSGA-III 在 ZDT2 函数上的 e 及 Δe 变化趋势

Fig.4 Trend of e and Δe of NSGA-III on ZDT2 function

在种群进化前期,由于算法探索导致种群大幅度更新 e 和 Δe 变化较剧烈。在进化后期,算法趋于收敛,导致 e 和 Δe 变化趋于平稳。每当 Δe 出现剧烈波动时,说明种群发现新的潜在解区域并进行探索。

在种群进化前期,种群个体变化剧烈 Δe^t 普遍较高,但也会少量出现 $|\Delta e^t| < \mu$ 的情况。为了避免此情况对判定种群进化阶段的干扰,设定当 $|\Delta e^t| < \mu$ 的出现次数大于最大进化代数 G_{\max} 的 10% 时,认为种群进入“探究”阶段。

2.2 基于目标空间的参考点选择策略

NSGA-III 采用精英策略,合并父代种群和子代种群后进行环境选择,其中非临界层的选择取决于 Pareto 支配关系,临界非支配层的选择依赖于标准化超平面上参考点关联个体的数目及距离。由参考点生成的参考线均匀贯穿整个目标空间,许多问题的真实 Pareto 前沿面并非均匀、连续分布,因此每个参考点对个体的选择并非具有同等重要性。

贯穿真实 Pareto 前沿面的参考线,个体关联的可能性较高。反之远离前沿面的参考线,关联可能性较低。随着种群进化,个体会逐渐偏向关联那些穿过真实 Pareto 前沿面的参考线。因此,虽然求解问题的 Pareto 前沿面未知,但可通过统计参考点关联每代种群个体的总数目评价参考点的重要性,保留部分关联数较多的参考点,测算 Pareto 前沿面所在范围。剔除冗余参考点,既减少无效参考点对种群个体选择的干扰,加快收敛,又减少计算量,提升算法效率。

以标准测试函数 DTLZ5 为例,真实前沿面如图 2(b) 所示,设定目标数 $M = 3$,决策数 $D = 12$,种群规模 $N = 100$,种群进化代数 $It = 200$,参考点每维划分 $p = 13$,参考点的数目 $H = 105$ 。统计 105 个参考点(如图 1)在种群 200 代进化中关联个体的总数,105 项数据总和为 $100 \times 200 = 20000$ 。105 个参考点关联个体数目具有明显区别:距离真实前沿面越近的参考点关联的个体数越多;否则,距离越远,关联越少。保留参考点中关联个体数最多的 21 个,得到图 5,显然,保留下的参考点分布与 DTLZ5 真实前沿面一致。

本文提出参考点策略,参考点数目自适应种群规模 N ,根据种群在目标空间进化阶段动态选取重要性较高的参考点。该策略无需用户设定参数 p 。为了保证筛选效果,淘汰的参考点数目应大于等于 $20\%N$ 。

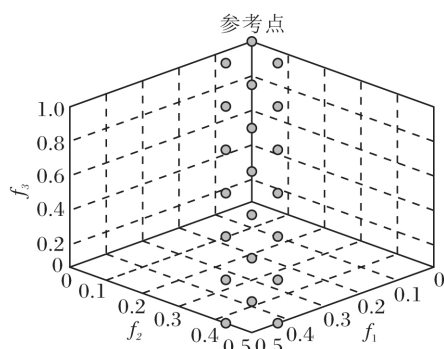


图 5 关联个体数较高的参考点

Fig.5 Reference points with higher number of association individuals

例 2 若目标数 $M = 3$ 种群规模 $N = 100$ 根据式 (1) 得

$H(M, 12) = 91$, $H(M, 13) = 105$, $H(M, 14) = 120$, 所以初始参考点数选择 120, 即 $p = 14$. 经参考点选择操作后保留 100 个参考点.

参考点筛选的详细步骤如下.

1) 根据种群规模 N 选取每维划分为 p 的参考点集 Z , 参考点数目为 H_p , p 满足

$$H_p \geq 1.2N, H_{p-1} < 1.2N.$$

2) 根据 2.1 节方案确定种群进化阶段.

3) 当种群处于“探索”阶段时, 统计参考点集 Z 每代关联个体数的总和 Z_{sum} .

4) 当种群刚进入“探究”阶段时, 根据 Z_{sum} 保留 Z 中关联数最多的 N 个参考点, 构成新的参考点集 Z_n .

2.3 算法流程

结合上述两个策略, 基于参考点选择策略的改进型 NSGA-III 算法的具体步骤如下.

算法 2 AR-NSGA-III

step 1 确定最大进化代数 G_{max} , 种群规模 N , 约束边界 ub , lb , 当前进化代数 $t = 0$.

step 2 初始化种群 P , 计算阈值 μ , 由 2.2 节计算确定初始参考点集 Z , 进化阶段 S' 默认为“探索”, 参考点集历史关联总数目 $Z_{\text{sum}} = 0$.

step 3 如果 S' 为“探索”阶段, 种群关联参考点, 更新 Z_{sum} .

step 4 如果 S' 为“探究”阶段, 种群数 N 与参考点集 Z 不等, 根据 Z_{sum} 删除 Z 中关联数最少的 $|Z| - N$ 个参考点.

step 5 根据参考点集 Z 进化种群 P , $t = t + 1$.

step 6 由 2.1 节计算进化阶段 S' .

step 7 如果 $t < G_{\text{max}}$ 转至 step 3; 否则 结束进化 输出种群.

具体源代码参见 <https://github.com/1209805090/AR-NSGA-III>.

2.4 时间复杂度分析

设目标数为 M , 决策数为 D , 种群规模为 N , NSGA-III 最坏时间复杂度^[19] 为 $O(N^2 \log^{M-2} N)$ 或 $O(N^2 M)$. 由 NSGA-III 改进得到 AR-NSGA-III, 增加如下操作“探索”阶段计算相邻两代熵值之差 $\Delta e'$ 的时间复杂度为 $O(DN \lg N)$. 统计参考点关联操作与原 NSGA-III 个体关联参考点同时进行, 无额外的时间开销.“探究”阶段参考点筛选的时间复杂度为 $O(N \lg N)$, 此项操作有且只进行一次. 综上所述, AR-NSGA-III 的每代循环最坏时间复杂度为

$$\max\{O(N^2 \log^{M-2} N), O(N^2 M), O(DN \lg N)\}.$$

3 实验及结果分析

为了验证 AR-NSGA-III 的有效性, 实验选取 DTLZ、MaF^[23] 测试函数集上 12 个标准测试函数, 分别为 IDTLZ1、IDTLZ2、SDTLZ2、CDTLZ2、DTLZ2、DTLZ4、DTLZ5、MaF1、MaF2、MaF5、MaF7、MaF8.

选取 4 种高维多目标进化算法进行对比, 分别为 NSGA-III、MaOEA-IGD、MaOEA-IT^[24] 和 MaOEA-R&D^[25]. MaOEA-R&D 为基于分解的 MaOEA, 首先沿 m 个方向寻找解并构造目标空间边界, 再采用多样性改善策略提高目标空间边界内种群的多样性. MaOEA-IGD 为基于指标的 MaOEA, 采用贪心策略, 通过 IGD 指标衡量当前种群的质量, 迭代删除对种群指标价值贡献最小的解. MaOEA-IT 为多阶段的 MaOEA, 提出非支配的动态权重集合, 进化过程中的两个阶段分别独立处理种群的收敛性和多样性.

为了保证公平性, 对于所有测试问题, 种群规模 $N = 100$, 最大进化代数 $G_{\text{max}} = 300$, 进化算子采用模拟二进制交叉和多项式突变, 交叉率均为 0.9, 变异率为 $1/D$, 分布指数设为 20, 其中 D 为决策变量维数. 每种算法在各测试函数上独立运行 30 次, 实验环境为目标进化算法开源平台 PlatEMO^[26].

为了评估算法有效性, 本文采用可同时反映算法收敛性与分布性的 IGD 和 HV 作为主要的性能评估指标.

IGD 为真实 Pareto 前沿面上点到多目标优化算法获得的解集个体的最短距离之和的均值, 值越小, 表示对应算法的性能越好. 计算公式如下:

$$IGD(P, Q) = \frac{1}{|P|} \sum_{v \in P} d(v, Q),$$

其中, Q 为多目标进化算法获得的解集, P 为真实 Pareto 前沿面上均匀采样的点集, $|P|$ 为点集 P 的个体数目, $d(v, Q)$ 为 P 中个体 v 到 Q 中个体的最小欧氏距离。

HV 为多目标优化算法获得的非支配解集与参照点围成的目标空间中的区域体积, 值越大, 表示对应算法的性能越好。计算公式如下:

$$HV = \delta \left(\bigcup_{i=1}^{|S|} v_i \right),$$

其中 δ 表示 Lebesgue 测度, 用于测量体积, $|S|$ 表示非支配解集的数目, v_i 表示参考点 z^* 与解集中第 i 个解构成的超体积。HV 指标的计算无需测试问题的 PF_{true} 。

计算 IGD 时, 每个测试实例的 Pareto 前沿使用 Das 和 Dennis 的两层方法采样 10 000 个参考点。计算 HV 时, 将目标值归一化, 参考点设置为 $(1, 1, \dots, 1)$ 。

AR-NSGA-III、NSGA-III、MaOEA-IGD、MaOEA-IT、MaOEA-R&D 在 12 个测试函数上的 IGD 和 HV 统计结果如表 1 和表 2 所示。

表 1 各算法在 12 种测试函数上的 IGD 值

Table 1 IGD metric values of different algorithms on 12 test functions

测试函数	M	D	AR-NSGA-III	NSGA-III	MaOEA-IGD	MaOEA-IT	MaOEA-R&D
IDTLZ1	4	8	8.2519e-2 (7.88e-3)	8.6017e-2 (5.66e-3) -	1.7534e-1 (1.15e-1) -	1.2168e-1 (4.78e-2) -	2.1309e-1 (1.60e-2) -
IDTLZ2	4	13	1.5996e-1 (5.12e-3)	1.7107e-1 (2.75e-3) -	3.7550e-1 (1.73e-2) -	3.9144e-1 (8.32e-2) -	6.3356e-1 (1.20e-2) -
SDTLZ2	4	13	4.8603e-1 (1.22e-2)	5.2956e-1 (5.42e-4) -	7.5963e-1 (5.67e-3) -	1.2697e+0 (3.03e-1) -	4.8189e-1 (1.29e-2) ≈
CDTLZ2	4	13	6.8347e-2 (1.37e-3)	7.6333e-2 (7.04e-4) -	2.0614e-1 (2.89e-2) -	4.3339e-1 (1.54e-1) -	7.3029e-2 (3.19e-3) -
DTLZ2	4	13	1.3443e-1 (1.44e-3)	1.4033e-1 (1.36e-5) -	1.6067e-1 (9.11e-4) -	4.2673e-1 (7.94e-2) -	1.4481e-1 (3.38e-3) -
DTLZ4	4	13	2.2232e-1 (1.42e-1)	2.4435e-1 (2.03e-1) -	3.5519e-1 (1.94e-1) -	8.5600e-1 (2.11e-1) -	1.8537e-1 (1.07e-1) +
DTLZ5	4	13	4.9866e-2 (1.27e-2)	5.6200e-2 (1.01e-2) -	5.9376e-1 (1.40e-1) -	3.4818e-1 (1.64e-1) -	7.4209e-1 (1.43e-7) -
MaF1	4	13	1.5529e-1 (7.57e-3)	1.6472e-1 (1.23e-2) -	2.4470e-1 (9.30e-4) -	3.2772e-1 (1.12e-1) -	4.1548e-1 (3.34e-2) -
MaF2	4	13	8.9716e-2 (3.02e-3)	9.8637e-2 (3.73e-3) -	2.6623e-1 (1.38e-1) -	1.3188e-1 (5.14e-2) -	4.1478e-1 (5.74e-2) -
MaF5	4	13	1.3650e+0 (7.40e-1)	1.4106e+0 (6.70e-1) -	2.8694e+0 (1.01e+0) -	6.4794e+0 (1.85e+0) -	1.3169e+0 (7.38e-1) +
MaF7	5	24	3.5759e-1 (8.83e-3)	3.9104e-1 (5.25e-2) -	7.6101e-1 (6.56e-2) -	3.3425e+0 (2.68e+0) -	1.7979e+0 (5.28e-1) -
MaF8	5	2	2.2068e-1 (2.23e-2)	2.6990e-1 (3.27e-2) -	9.6198e-1 (2.75e-1) -	1.6633e+2 (2.45e+2) -	9.9209e-1 (1.68e-1) -
+/-/≈				0/12/0	0/12/0	0/12/0	2/9/1

表 2 各算法在 12 种测试函数上的 HV 平均值

Table 2 HV metric values of different algorithms on 12 test functions

测试函数	M	D	AR-NSGA-III	NSGA-III	MaOEA-IGD	MaOEA-IT	MaOEA-R&D
IDTLZ1	4	8	3.2764e-2 (4.38e-3)	3.0296e-2 (3.43e-3) -	2.2088e-2 (1.21e-2) -	3.1265e-2 (1.13e-2) ≈	8.4316e-3 (1.17e-3) -
IDTLZ2	4	13	2.2234e-1 (5.25e-3)	2.1598e-1 (4.33e-3) -	1.7831e-1 (1.29e-2) -	5.6556e-2 (4.02e-2) -	7.9877e-2 (6.55e-3) -
SDTLZ2	4	13	6.9642e-1 (2.55e-3)	6.9125e-1 (4.19e-4) -	5.8551e-1 (1.97e-3) -	1.5814e-1 (9.60e-2) -	6.1993e-1 (1.01e-2) -
CDTLZ2	4	13	9.9409e-1 (2.45e-4)	9.9276e-1 (1.88e-4) -	9.1799e-1 (3.66e-2) -	5.8245e-1 (1.88e-1) -	9.9106e-1 (7.96e-4) -
DTLZ2	4	13	6.9641e-1 (2.63e-3)	6.9117e-1 (5.97e-4) -	6.9053e-1 (7.39e-4) -	1.7228e-1 (6.98e-2) -	6.1621e-1 (1.13e-2) -
DTLZ4	4	13	6.5576e-1 (7.14e-2)	6.3449e-1 (1.23e-1) -	5.8615e-1 (1.13e-1) -	1.7909e-1 (9.79e-2) -	6.2466e-1 (3.81e-2) -
DTLZ5	4	13	1.3746e-1 (1.77e-3)	1.3498e-1 (2.74e-3) -	7.6043e-2 (4.22e-2) -	5.4104e-3 (7.04e-3) -	9.0909e-2 (3.51e-7) -
MaF1	4	13	3.5197e-2 (1.75e-3)	3.2724e-2 (2.33e-3) -	2.4672e-2 (3.32e-4) -	1.0871e-2 (9.95e-3) -	9.0905e-3 (1.04e-3) -
MaF2	4	13	2.2369e-1 (2.46e-3)	2.1806e-1 (2.58e-3) -	1.3711e-1 (4.84e-2) -	1.5033e-1 (3.41e-2) -	5.7935e-2 (1.39e-2) -
MaF5	4	13	6.4806e-1 (7.73e-2)	6.4011e-1 (8.00e-2) -	4.5908e-1 (1.00e-1) -	1.9618e-1 (6.35e-2) -	6.0931e-1 (4.67e-2) -
MaF7	5	24	2.3326e-1 (4.04e-3)	2.2881e-1 (4.58e-3) -	1.2077e-1 (4.05e-2) -	2.7447e-2 (3.81e-2) -	1.6161e-1 (1.86e-2) -
MaF8	5	2	9.3321e-2 (4.17e-3)	8.6442e-2 (7.80e-3) -	3.3823e-2 (1.81e-2) -	5.2107e-3 (9.58e-3) -	3.0047e-2 (1.08e-2) -
+/-/≈				0/12/0	0/12/0	0/11/1	0/12/0

对于同个测试问题, 每种算法独立运行 30 次, 计算结果的平均值与标准差, 每组中的最优值使用黑体数字表示。采用非参数统计 Wilcoxon 秩和检验分别对比 5 种算法, 显著性水平设为 5%。表 1 和表 2 中符号“+”表示其它算法显著优于 AR-NSGA-III,

“-”表示显著劣于 AR-NSGA-III, “≈”表示 2 种算法无差别。

对于 IGD 性能指标, 由表 1 可知, 本文算法在 12 个测试函数上的 9 个函数取得最优的 IGD 值。在 SDTLZ2、DTLZ4 和 MaF5 上的最优 IGD 值由 MOEA-

R&D 获得,其它对比算法均未获取最优的 IGD 值.本文算法与对比算法在 IGD 指标上对比 48 次,显著优于(+)为 45 次,显著劣于(-)为 2 次,无差别(\approx)为 1 次.

由表 2 可知,AR-NSGA-III 在 12 个测试函数上的 HV 指标全部取得最优平均值,在 48 项对比中,显著优于(+)为 47 次,显著劣于(-)为 0 次,无差别(\approx)为 1 次.由此可知,AR-NSGA-III 的收敛性和多样性在 12 个测试函数上表现最优.

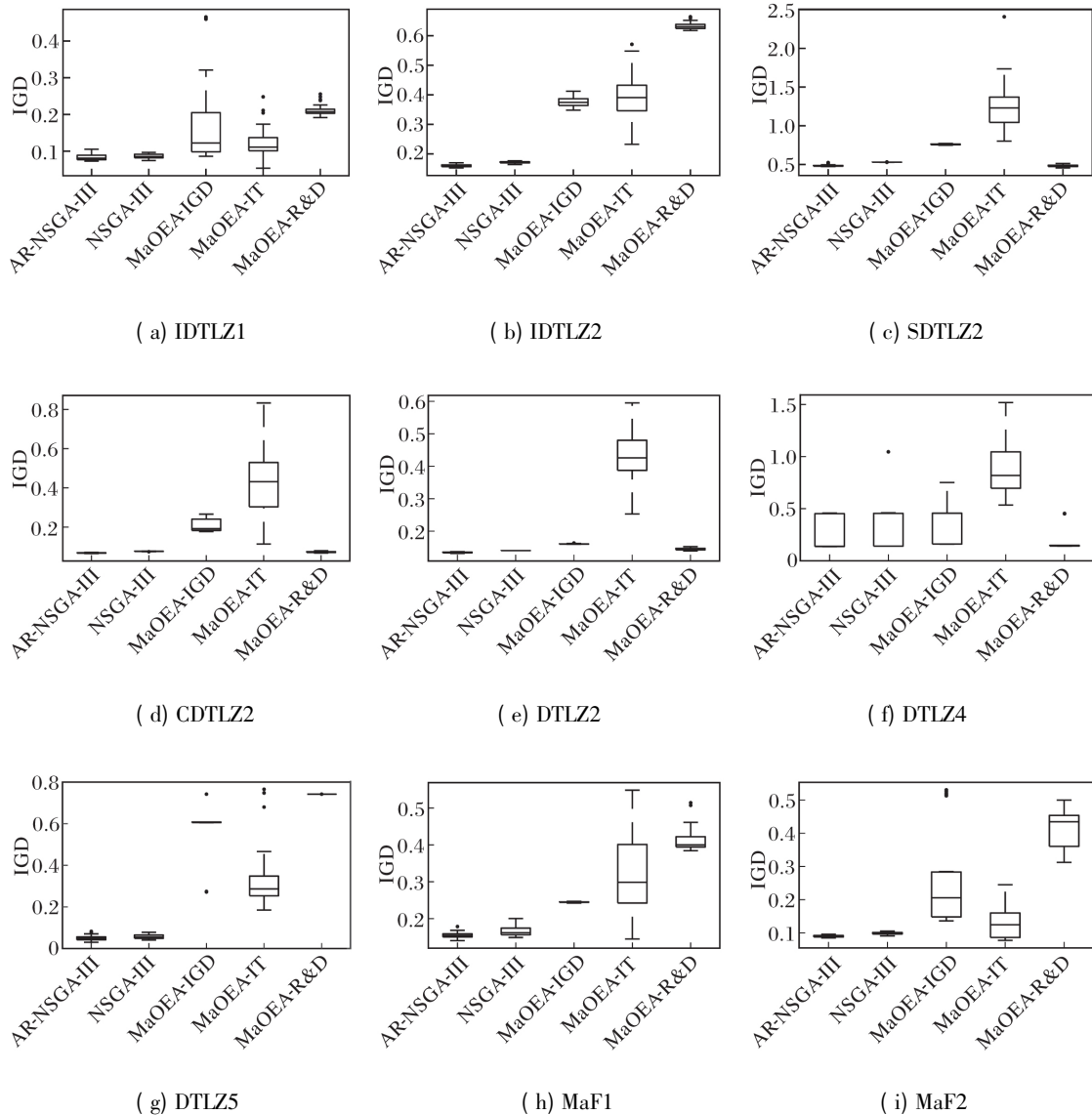
在 IGD 和 HV 指标的对比中,本文算法无论是在四目标问题还是五目标问题上,相比 NSGA-III 具有一定优势,这说明基于决策空间的进化阶段判定策略和基于目标空间的参考点选择策略的有效性.

本文利用箱形图表示算法在每个测试函数上的

统计结果.箱形图的外层上下线表示样本的最大值和最小值,内层上下线表示上下四分位数,中间线表示中位数,“ \cdot ”表示异常值.各算法在 12 个函数上的箱型图如图 6 和图 7 所示.

在 IGD 指标箱形图方面,如图 6 中 (b)、(c)、(e)、(g)、(i) 所示.在 HV 指标箱形图方面,如图 7 中 (b)、(c)、(e)、(g)、(i)、(k) 所示.本文算法不仅在均值上占优,在中位数、上下四分位、异常值、方差上也具有明显优势,表明算法相对稳定,具有良好的鲁棒性.

综上所述,AR-NSGA-III 利用种群在决策空间的分布信息及根据问题前沿面选择参考点的策略,有效平衡算法的分布性和收敛性.此外,在处理不同类型的问题时,本文算法的鲁棒性也较优.



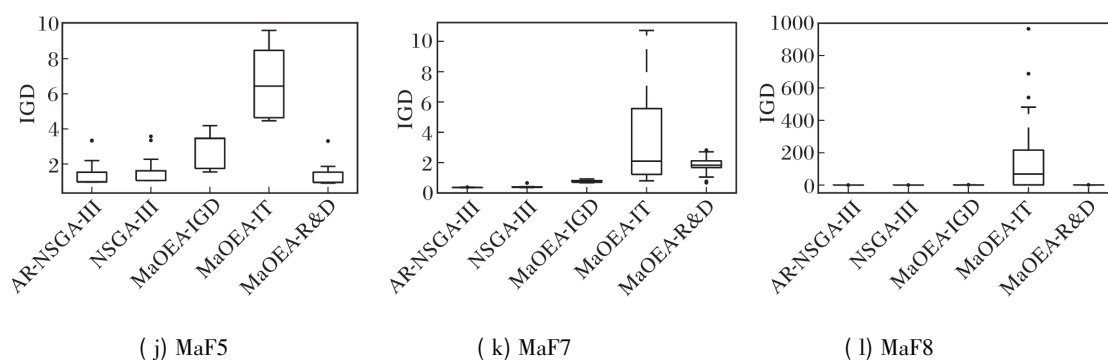
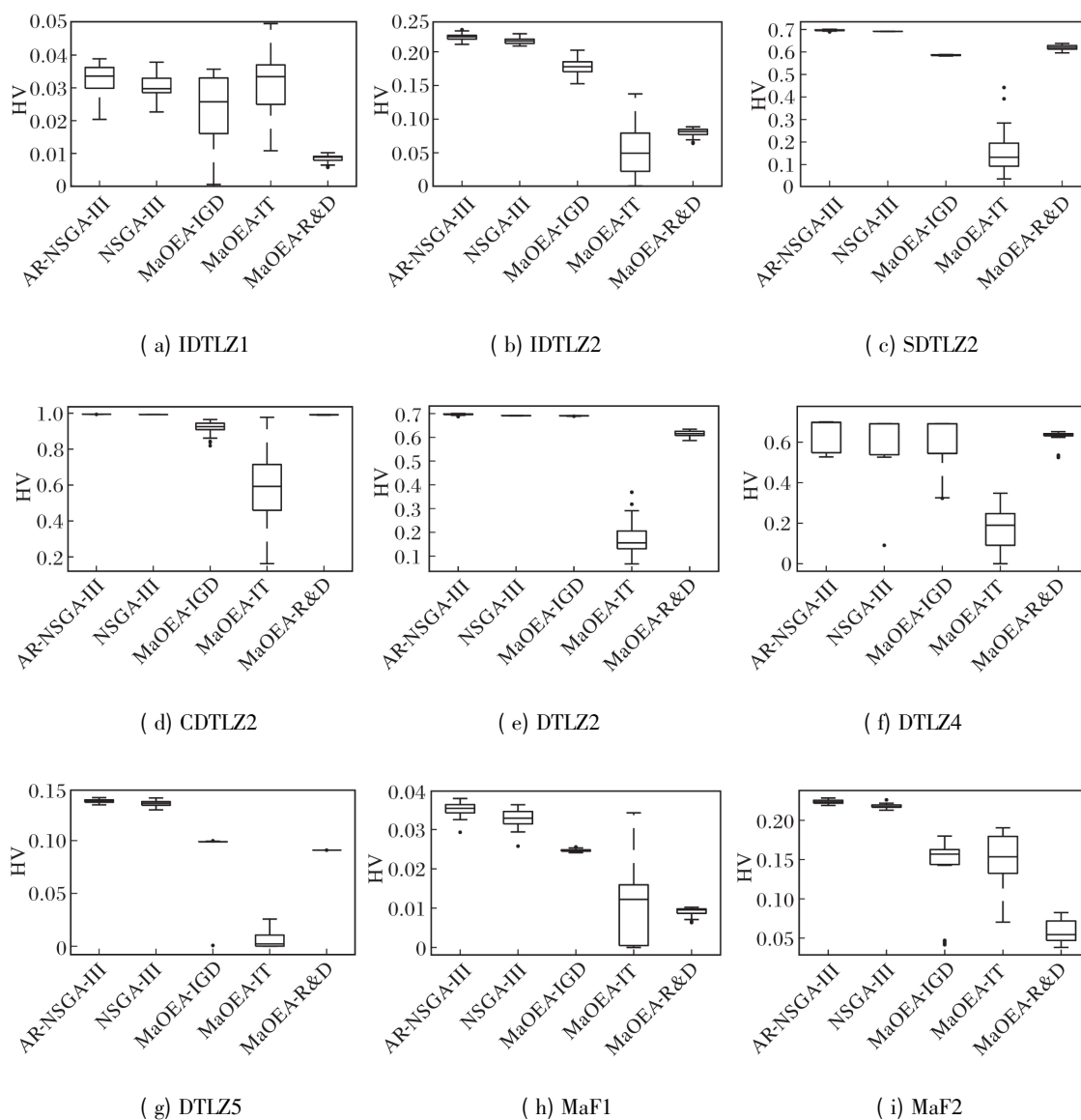


图6 各算法在12个测试函数上的IGD指标箱形图

Fig.6 Box plots of IGD metric values of different algorithms on 12 functions



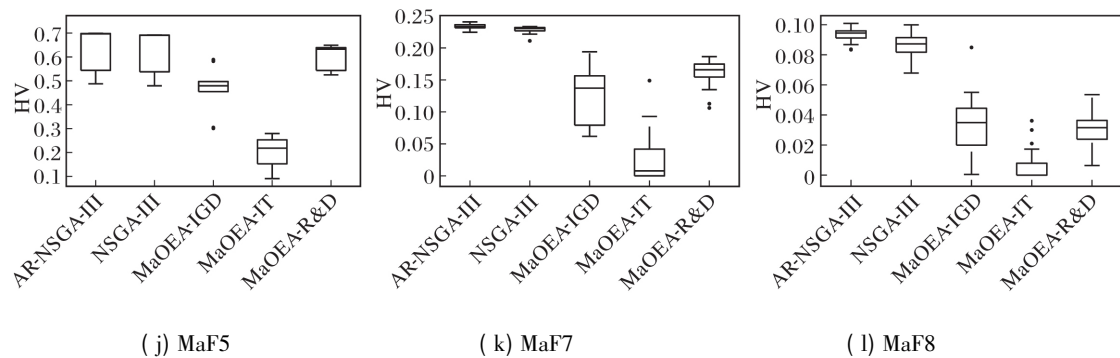


图 7 各算法在 12 个测试函数上的 HV 指标箱形图

Fig.7 Box plots of HV metric values of different algorithms on 12 test functions

4 结 束 语

本文针对 NSGA-III 未利用种群决策空间信息与参考点冗余没有自适应应解决问题 Pareto 前沿形状的不足,提出基于参考点选择策略的改进型 NSGA-III 算法(AR-NSGA-III).在进化过程中利用种群在决策空间四分位的信息,以熵的差值进行量化,判定种群进化阶段.根据种群在目标空间的分布信息,关联参考点,评估各参考点的重要性,进行参考点筛选.在不同的进化阶段利用不同的参考点对种群进行选择,减少无效参考点对种群个体选择的干扰,提升算法效率与性能.实验表明,本文算法可以较好地收敛到可行域,并保证可行域内的分布性.在一些特定现状的 Pareto 前沿面问题上的表现较优.今后将对初始参考点规模的设定进行分析,拓展参考点筛选策略,使保留参考点具有更好的分布性.下一步还应对进化阶段的判定进行讨论,采用更合理的阈值,使阶段划分更准确.

参 考 文 献

- [1] ISHIBUCHI H, AKEDO N, NOJIMA Y, *et al.* Behavior of Multiobjective Evolutionary Algorithms on Many-Objective Knapsack Problems. *IEEE Transactions on Evolutionary Computation*, 2015, 19(2): 264-283.
- [2] PIERRO F, KHU S, SAVIC D A, *et al.* An Investigation on Preference Order Ranking Scheme for Multiobjective Evolutionary Optimization. *IEEE Transactions on Evolutionary Computation*, 2007, 11(1): 17-45.
- [3] YANG S X, LI M Q, LIU X H, *et al.* A Grid-Based Evolutionary Algorithm for Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 2013, 17(5): 721-736.
- [4] ZHANG X Y, TIAN Y, JIN Y C, *et al.* A Knee Point-Driven Evolutionary Algorithm for Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 2015, 19(6): 761-776.
- [5] ZHU C W, XU L H, GOODMAN E D, *et al.* Generalization of Pareto-Optimality for Many-Objective Evolutionary Optimization. *IEEE Transactions on Evolutionary Computation*, 2016, 20(2): 299-315.
- [6] TIAN Y, CHENG R, ZHANG X Y, *et al.* A Strengthened Dominance Relation Considering Convergence and Diversity for Evolutionary Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 2019, 23(2): 331-345.
- [7] ZHANG Q F, LI H. MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition. *IEEE Transactions on Evolutionary Computation*, 2007, 11(6): 712-731.
- [8] HO-HUU V, HARTJES S, VISSER H G, *et al.* An Improved MOEA/D Algorithm for Bi-objective Optimization Problems with Complex Pareto Fronts and Its Application to Structural Optimization. *Expert Systems with Applications*, 2018, 92: 430-446.
- [9] LI H, ZHANG Q F, DENG J D, *et al.* Biased Multiobjective Optimization and Decomposition Algorithm. *IEEE Transactions on Cybernetics*, 2017, 47(1): 52-66.
- [10] ZHENG W, TAN Y Y, MENG L L, *et al.* An Improved MOEA/D Design for Many-Objective Optimization Problems. *Applied Intelligence*, 2018, 48(10): 3839-3861.
- [11] ZHOU Z B, LIU X H, XIAO H L, *et al.* A DEA-Based MOEA/D Algorithm for Portfolio Optimization. *Cluster Computing*, 2019, 22(6): 14477-14486.
- [12] YI J H, DEB S, DONG J Y, *et al.* An Improved NSGA-III Algorithm with Adaptive Mutation Operator for Big Data Optimization Problems. *Future Generation Computer Systems*, 2018, 88: 571-585.
- [13] YI J H, XING L N, WANG G G, *et al.* Behavior of Crossover Operators in NSGA-III for Large-Scale Optimization Problems. *Information Sciences*, 2020, 509: 470-487.

- [14] LOPEZ E M , COELLO C C A. IGD⁺-EMOA: A Multi-objective Evolutionary Algorithm Based on IGD⁺ // Proc of the IEEE Congress on Evolutionary Computation. Washington , USA: IEEE , 2016: 999–1006.
- [15] TIAN Y , CHENG R , ZHANG X Y , *et al.* An Indicator-Based Multiobjective Evolutionary Algorithm with Reference Point Adaptation for Better Versatility. IEEE Transactions on Evolutionary Computation , 2018 , 22(4) : 609–622.
- [16] SUN Y N , YEN G G , YI Z , *et al.* IGD Indicator-Based Evolutionary Algorithm for Many-Objective Optimization Problems. IEEE Transactions on Evolutionary Computation , 2019 , 23(2) : 173–187.
- [17] 黎明, 段茹茹, 陈昊, 等. 基于 IGD⁺S 指标的高维多目标进化算法. 模式识别与人工智能, 2019 , 32(9) : 800–810.
(LI M , DUAN R R , CHEN H , *et al.* Multi-objective Evolutionary Algorithm Based on IGD⁺S. Pattern Recognition and Artificial Intelligence , 2019 , 32(9) : 800–810.)
- [18] LI M Q , YANG S X , LIU X H. Pareto or Non-pareto: Bi-criterion Evolution in Multi-objective Optimization. IEEE Transactions on Evolutionary Computation , 2016 , 20(5) : 645–665.
- [19] DEB K , JAIN H. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach , Part I: Solving Problems with Box Constraints. IEEE Transactions on Evolutionary Computation , 2014 , 18(4) : 577–601.
- [20] ZHANG Q F , ZHOU A M , ZHAO S Z , *et al.* Multiobjective Optimization Test Instances for the CEC 2009 Special Session and Competition. Technical Report , CES-487. Colchester , UK: University of Essex , 2008.
- [21] HUBAND S , HINGSTON P , BARONE L , *et al.* A Review of Multiobjective Test Problems and a Scalable Test Problem Toolkit. IEEE Transactions on Evolutionary Computation , 2006 , 10(5) : 477–506.
- [22] 耿焕同, 陈哲, 陈正鹏, 等. 一种基于群体分布特征的自适应多目标粒子群优化算法. 控制与决策, 2017 , 32(8) : 1386–1394.
(GENG H T , CHEN Z , CHEN Z P , *et al.* A Self-adaptive Multi-objective Particle Swarm Optimization Algorithm Based on Swarm Distribution Characteristic. Control and Decision , 2017 , 32(8) : 1386–1394.)
- [23] CHENG R , LI M Q , TIAN Y , *et al.* A Benchmark Test Suite for Evolutionary Many-Objective Optimization. Complex and Intelligent Systems , 2017 , 3(1) : 67–81.
- [24] SUN Y N , XUE B , ZHANG M J , *et al.* A New Two-Stage Evolutionary Algorithm for Many-Objective Optimization. IEEE Transactions on Evolutionary Computation , 2019 , 23(5) : 748–761.
- [25] HE Z N , YEN G G. Many-Objective Evolutionary Algorithm: Objective Space Reduction and Diversity Improvement. IEEE Transactions on Evolutionary Computation , 2016 , 20(1) : 145–160.
- [26] TIAN Y , CHENG R , ZHANG X Y , *et al.* PlatEMO: A MATLAB Platform for Evolutionary Multi-objective Optimization. IEEE Computational Intelligence Magazine , 2017 , 12(4) : 73–87.

作者简介



耿焕同(通讯作者), 博士, 教授, 主要研究方向为计算智能、多目标优化、气象数据挖掘. E-mail: htgeng@nuist.edu.cn.

(**GENG Huantong**(Corresponding author) , Ph.D. , professor. His research interests include computational intelligence , multi-objective optimization and meteorological data mining.)



戴中斌, 硕士研究生, 主要研究方向为多目标优化. E-mail: 1209805090@qq.com.

(**DAI Zhongbin** , master student. His research interests include multi-objective optimization.)



王天雷, 硕士研究生, 主要研究方向为深度学习. E-mail: 609294510@qq.com.

(**WANG Tianlei** , master student. His research interests include deep learning.)



许可, 硕士研究生, 主要研究方向为多目标优化. E-mail: 1327381035@qq.com.

(**XU Ke** , master student. His research interests include multi-objective optimization.)