文章编号: 1001-0920(2018)12-2213-05　　DOI: 10.13195/j.kzyjc.2017.0938

区间参数高维多目标集合进化优化方法

季新芳[1†], 张　凤[1], 王彩君[2], 严海领[1], 李　娜[1]

(1. 中国矿业大学银川学院 机电动力与信息工程系，银川 750001；
2. 中国矿业大学银川学院 土木工程系，银川 750001)

摘　要: 区间参数高维多目标优化问题是现实生活中常见的一类优化问题,但其有效的求解方法并不是很多. 对此,利用集合的概念,提出一种求解此类问题的新方法. 首先,利用衡量解集收敛性、分布性、多样性的3种性能指标将原优化问题降为3目标优化问题;其次,采用集合Pareto占优关系和不确定测度来区分转化后优化问题解的优劣;再次,设计自适应变化的交叉、变异概率以提高种群的全局和局部搜索能力;最后,利用4种基准函数优化问题,对所提出方法和对比方法进行测试.测试结果显示,除了收敛性,所提出方法得到的Pareto解集的不确定性、多样性、分布性均优于对比方法.

关键词: 高维多目标优化；自适应；区间参数；集合进化

中图分类号: TP301.6　　　**文献标志码:** A

Optimizing many-objective problems with interval parameters using set-based evolutionary algorithms

JI Xin-fang[1†], *ZHANG Feng*[1], *WANG Cai-jun*[2], *YAN Hai-ling*[1], *LI Na*[1]

(1. Department of Mechanical Power and Information Engineering，China University of Mining and Technology Yinchuan College，Yinchuan 750001，China；2. Department of Civil Engineering，China University of Mining and Technology Yinchuan College，Yinchuan 750001，China)

Abstract: There is few effective solving methods for the high-dimensional multi-objective optimization problem with interval parameters in real life. Therefore a new method is proposed using to concept of set. Firstly , three performance indicators including convergence, distribution and diversity of a solution set are used to transform the original optimization problem into a tri-objective one. Then, the set-based Pareto dominant relationships and a measure of uncertainty are used to distinguish the solutions of the converted optimization problem. Additionally, the adaptive crossover and mutation rates are suggested to improve the global and local search ability of the population. Finally, four benchmark function optimization problems are used to test the performance of the proposed method and comparison method. The test results show that, the uncertainty, diversity, and distributivity of a Pareto solution set obtained by the proposed method is superior to that by the contrast method, in addition to convergence.

Keywords: high-dimensional multi-objective optimization；adaptive；interval parameter；set-based evolution

0　引　言

在处理实际优化问题时,需要考虑该问题中包含的多个相互冲突的目标,如汽车驾驶室设计问题[1]和井下巷道无线射频阅读器布局问题[2]. 当目标函数中含有不确定参数,如模糊数、随机变量,或者区间时,这类问题称为不确定多目标优化问题.

由于随机变量通过置信水平[3]、模糊数通过截集水平[4]均能够转化为区间,从而随机和模糊参数优化问题可以转化为区间参数优化问题. 当相互冲突的目标函数多于3个时,称为区间参数高维多目标优化问题.

目前,区间参数高维多目标优化问题的求解面临如下问题[1]: 1) 目标函数的高维,使得种群中Parteo最优解增多,无法快速地选择优势解; 2) 进化过程中,某些性能指标的运算时间会随着目标维数的增多而呈指数增长; 3) 目标维数超过3维后,无法图示出优化问题的Pareto前沿; 4) 目标函数的不确定性,增加了优势解选择的难度.

收稿日期: 2017-07-13；**修回日期:** 2017-10-13.
基金项目: 宁夏高等学校科研项目(NGY2016226).
责任编委: 魏秀琨.

作者简介: 季新芳 (1987−), 女, 讲师, 硕士, 从事多目标优化、智能算法的研究；张凤 (1987−), 女, 讲师, 硕士, 从事智能控制的研究.
†通讯作者. E-mail: mimosa_615615@126.com

鉴于以上区间参数高维多目标优化问题求解时遇到的问题,巩敦卫等[5]不再基于传统的占优准则求解该类问题,而是将平时作为评价解集优劣的性能指标,如超体积、不确定度等作为问题的目标,将高维多目标优化问题降成低维进行求解. 由于此时的解不是一个单个的个体,而是一个含有多个个体的解集,进而还提出了相关的解集个体的变异和重组策略.

本文在文献[5]的基础上进行深入研究,给出一种求解区间参数高维多目标优化问题的自适应集合进化优化方法. 首先,利用衡量解集收敛性的超体积测度、分布性测度、多样性测度,使原区间参数高维多目标优化问题变成3目标确定型优化问题;然后,套用NSGA-II范式,利用基于集合的Pareto占优关系以及解集的不确定度,对转化后优化问题的解进行排序;最后,根据种群信息设计自适应交叉、变异概率,以产生高性能的子代个体.

1 算 法

考虑如下优化问题:

$$\begin{aligned} &\min\ f(\boldsymbol{x},\boldsymbol{c})=(f_1(\boldsymbol{x},\boldsymbol{c}),f_2(\boldsymbol{x},\boldsymbol{c}),\cdots,f_m(\boldsymbol{x},\boldsymbol{c})).\\ &\text{s.t. } \boldsymbol{x}\in S\subset R^n;\\ &\qquad \boldsymbol{c}=(c_1,c_2,\cdots,c_l)^{\mathrm{T}};\\ &\qquad c_k=[\underline{c_k},\overline{c_k}],\ k=1,2,\cdots,l. \end{aligned}\tag{1}$$

其中: $\boldsymbol{x}$为n维决策变量; S为$\boldsymbol{x}$的决策空间; $f_i(\boldsymbol{x},\boldsymbol{c})=[\underline{f_i}(\boldsymbol{x},\boldsymbol{c}),\overline{f_i}(\boldsymbol{x},\boldsymbol{c})](i=1,2,\cdots,m)$为第$i$个目标函数; m为目标函数的个数,且$m>3$; $\boldsymbol{c}$为区间向量参数, $\overline{c}_k$和$\underline{c}_k$分别为$\boldsymbol{c}$的第k个分量c_k的上下限; $\overline{f_i}(\boldsymbol{x},\boldsymbol{c})$和$\underline{f_i}(\boldsymbol{x},\boldsymbol{c})$分别为第$i$个目标函数值的上下限.

根据式(1),本文提出一种求解区间参数高维多目标优化问题的集合进化优化方法. 首先,对原高维多目标优化问题进行降维,利用设计的区间分布性测度,结合集合的超体积、多样性测度来实现(详见1.1节);然后,在求解降维后优化问题时,套用NSGA-II范式,利用基于集合的Pareto占优关系,以及解集的不确定度,比较个体的优劣(详见1.2节);通过设计的自适应交叉、变异概率,提高种群的搜索能力,产生高性能的子代个体(详见1.3节). 最后,通过求解基准函数优化问题来验证算法的性能(详见2.3节).

1.1 优化问题降维转化

本节延续文献[5-7]的思想,利用评价解集优劣的性能指标对问题(1)进行降维,使其转化为以集合为决策变量的确定型多目标优化问题.

对于区间多目标优化问题, Pareto优化解集一般要求具有如下4个特性: 1) 得到的Pareto前沿具有较好的收敛性,即离真实Pareto前沿越近越好; 2) 构成Pareto前沿的点在目标空间上能够均匀分布; 3) Pareto前沿上的点在解空间中的差异性越大越好,即希望具有较好的多样性; 4) Pareto前沿在目标空间上具有较小的不确定度. 其中:特性1),收敛性可以利用文献[8]提出的超体积来衡量;特性3),多样性可以利用文献[9]提出的多样测度来衡量; 特性2),分布性可以参考文献[10],但是它只适用于求精确解的分布性. 基于此,这里给出求解区间解的分布性测度$D(X)$为

$$D(X)=\sqrt{\sum_{i=1}^{N}(d^*-d_i)^2/(N-1)}.\tag{2}$$

其中

$$\begin{aligned} d(x_i)=&\min_{j=1,2,\cdots,N,i\neq j}\Big(\sum_{l=1}^{m}\Big|\frac{\overline{f_l}(x_i)+\underline{f_l}(x_i)}{2}-\\ &\frac{\overline{f_l}(x_j)+\underline{f_l}(x_j)}{2}\Big|\frac{1}{1+\mathrm{e}^{-\frac{2k}{r+\xi}}}\Big),\\ r=&\frac{\overline{f_l}(x_i)-\underline{f_l}(x_i)}{2}+\frac{\overline{f_l}(x_j)-\underline{f_l}(x_j)}{2},\\ d^*=&\frac{1}{N}\sum_{i=1}^{N}d(x_i). \end{aligned}$$

这里: $d(x_i)$为个体x_i在原目标空间中最小的拥挤距离; $\overline{f_l}(x_i)$为个体x_i第l个目标的上限值; $\underline{f_l}(x_i)$为个体x_i第l个目标的下限值; r为个体x_i与个体x_j的不确定度之和,用于调节$d(x_i)$,不确定度r越大,个体间的距离$d(x_i)$越小; $\xi=0.000\,1$,保证分母不为0; k为调节系数,这里取1; d^*为所有个体最小拥挤距离的平均值; N为集合中原优化问题解的数目. $D(X)$越小,得到的解的分布性越好.

有了区间分布性测度后,利用Pareto解集的前3个特性,即收敛性、多样性、分布性,将式(1)描述的优化问题转化为确定型3目标优化问题

$$\begin{aligned} &\max\ F_1(X)=\varLambda\Big(\bigcup_{x_j\in X}\{h|\underline{f}(\boldsymbol{x}_j)<h<\overline{f}(\boldsymbol{x}_{\mathrm{ref}})\}\Big),\\ &\qquad F_2(X)=1-H/H_{\max},\\ &\qquad F_3(X)=-D(X)=-\sqrt{\sum_{i=1}^{N}(d^*-d_i)^2/(N-1)};\\ &\text{s.t. } \boldsymbol{X}\in P(S). \end{aligned}\tag{3}$$

其中: $H=\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}H(x_i,x_j)\Big/\sum_{i=1}^{N}(N-i)$; $H_{\max}=\max\Big(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}H(x_i,x_j)\Big)$; $P(S)$为解集个体的集合; $X=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N\}$为由原优化问题(1)的$N$个解构成的集合; $\varLambda(\cdot)$为勒贝格测度; $\boldsymbol{x}_{\mathrm{ref}}$为参考点; $H(X)$为X中个体在解空间中的平均距离; $H_{\max}$为X中距离最大的两个个体间的距离. $F_1(X)$、$F_2(X)$、$F_3(X)$分别用于计算解集X的收敛性、多样性和分布性. 注意, $F_3(X)$是对式(2)取负得到的,使得式(3)均为最大化问题.

1.2 转化后问题解的排序策略

采用NSGA-II中个体比较的模式,先通过集合Pareto占优关系[5]得到个体的序值. 对于相同序值的个体,再采用其他测度进一步比较. 鉴于式(3)描述的优化问题只能体现原优化问题的前3个特性,即收敛性、分布性、多样性,缺少了不确定性,因此,这里将采用不确定度进一步比较具有相同序值的个体.所设计的不确定度表达式为

$$BQ=\sum_{i=1}^{N}\sum_{l=1}^{m}\frac{\overline{f_l}(x_i)-\underline{f_l}(x_i)}{2}\Big/N. \tag{4}$$

其中: $\overline{f_l}(x_i)$为个体x_i第l个目标的上限值, $\underline{f_l}(x_i)$为个体x_i第l个目标的下限值, N为集合中包含的原优化问题解的个数.

1.3 集合进化策略设计

集合进化策略包括集合交配选择操作、集合交叉操作的设计、交叉概率的设计、集合变异操作的设计、变异概率的设计. 鉴于文献[5]已对选择操作和交叉、变异操作进行了相关研究,本文将借鉴文献[5]中的一些基本操作,着重设计自适应交叉、变异概率.

1.3.1 自适应交叉概率

鉴于转化后优化问题的一个解是原优化问题多个解的集合,因此求解转化后优化问题时,其解的交叉方式分为两种,即:不同解之间的交叉和同一解内部的交叉.具体的交叉操作可以参见文献[5].

由于交叉概率的大小会对算法的全局搜索能力和收敛速度产生影响,对于不同解之间的交叉,本文采用固定的交叉概率;而对于同一解内部的交叉则采用自适应交叉概率. 这里将利用集合解的多样性和收敛进度等信息,动态调整交叉概率,进一步使得集合解具有较好的多样性和收敛性. 需要注意的是,种群中每个集合解内部的交叉概率都是不同的.

本文设计的自适应交叉概率为

$$P_c=\frac{1}{1+\exp\Big(-\dfrac{k_1}{F_2(X)F_1(X)+k_2t+\xi}\Big)}. \tag{5}$$

其中: $F_1(X)$、$F_2(X)$分别表示集合个体X的收敛性和多样性, t为进化代数, $k_1=0.1, k_2=0.03, \xi=0.000\,1$,保证分母不为0.

由式(5)可以看出:集合解的收敛进度$F_1(X)$和多样性$F_2(X)$均与交叉概率成反比关系,其值减小时, P_c变大,使算法搜索新解的能力增强;反之,当集合解的收敛性或多样性较好时, P_c变小,使好的解的基因能够保留下来. 同时,为了使算法能够收敛,交叉概率应随着进化过程的推进整体呈下降趋势.

1.3.2 自适应变异概率

对种群中的每个集合解都进行变异操作,对进行变异操作的个体X_i,采用单点变异的方式,该操作进行N次.

变异概率的大小会影响算法的局部搜索能力,因此,依据式(5)的设计思路,这里设计了随集合解收敛性和多样性自适应变化的变异概率,即

$$P_m=1-\frac{1}{1+\exp(-F_2(X)F_1(X)-k_2t)}. \tag{6}$$

由式(6)可以看出: 与交叉概率类似,多样性$F_2(X)$或收敛进度$F_1(X)$越小, P_m越大;反之,多样性$F_2(X)$或收敛进度$F_1(X)$越大, P_m越小. 同时,为了使算法能够收敛,变异概率也应随着进化过程的推进整体呈下降趋势.

2 应 用

将本文提出的方法记为SET-IMOEA(Optimizing many-objective problems with interval parameters using set-based evolutionary algorithms),并与SET-MOEA[6]进行性能比较. 采用Matlab7.0编程实现,每种方法均独立运行20次,记录运行结果,并计算这些结果的平均值. 选基准优化问题DTLZ1′、DTLZ2′、DTLZ3′及DTLZ5′[5,11]进行性能验证. 实验中4个基准优化问题的目标函数个数分别取5和15,且均为最大化问题.

2.1 参数设置

将SET-IMOEA和SET-MOEA的种群规模和每一个个体包含原优化问题解的个数均取为10. SET-MOEA中的集合变异操作采用规模为2的联赛选择、模拟二进制交叉以及多项式变异. 其中:交叉和变异概率分别为0.9和0.1,交叉和变异操作的分布系数均为20. SET-IMOEA中采用规模为2的联赛选择,集合个体间采用单点交叉,交叉概率取为0.9,集合个体内的交叉采用模拟二进制交叉,交叉概率由式(5)得到;集合个体采用单点变异,变异概率由式(6)得到. 两种方法的最大进化代数取为$T_{\max}=100$. 两种方法均采用文献[12]提出的Monte Carlo方法近似计算超体积,采样次数取10 000. 由于优化问题为最大化问题,且目标函数经过归一化处理,超体积计算参考点为$r=[0,0,\cdots,0]$.

2.2 性能指标

实验将从以下5个方面比较不同方法的性能:

1) 最大超体积[12],简称H测度. H测度值越大,算法的收敛性越好.

2) 多样性,简称D测度,采用式(3)的第2个式子计算. 某方法所得Pareto前沿的多样性值越大,该前沿的多样性越好.

3) 分布度,简称SP测度,采用式(3)的第3个式子计算. 某方法所得Pareto前沿的分布性值越小,该前沿的分布性越好.

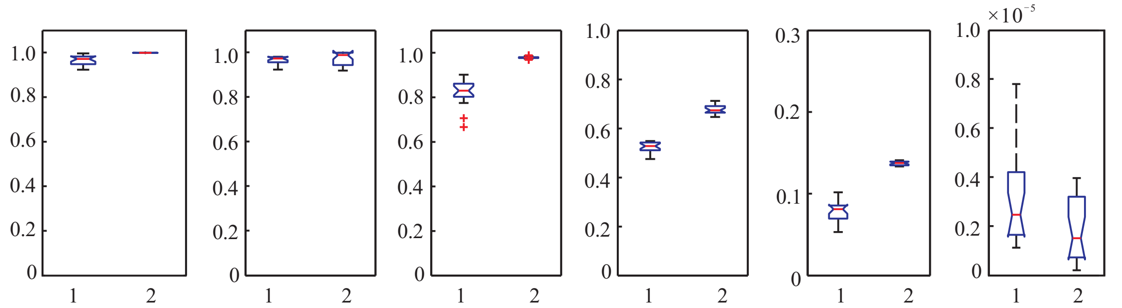
4) 不确定度,简称BQ测度,采用式(4)计算. 某方法所得Pareto前沿的不确定度越小越好.

5) CPU时间,简称T测度. T测度越小,运行效率越高.

2.3 实验结果与分析

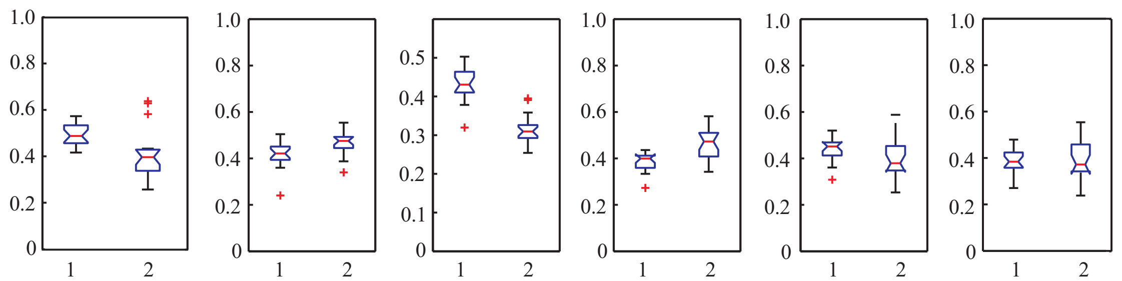
下面将SET-IMOEA与SET-MOEA得到的第1个优势个体进行比较,验证所提出方法能够更快地得到收敛性、不确定性、多样性、分布性均衡的原优化问题的Pareto解集.

图1～图4分别给出了SET-IMOEA和SET-MOEA求解各优化问题时,得到的Pareto前沿的H、D、SP以及BQ测度的箱图. 其中:横坐标中1代表SET-IMOEA,2代表SET-MOEA. 由于DTLZ3$'$的各测度值与DTLZ1$'$类似,这里其箱图不再给出. 采用Mann-Whitney U检验这两种方法在相同测度上差异的显著性,显著性水平取0.05,检验结果见表1. 其中:“+”表示SET-IMOEA显著优于SET-MOEA,“−”表示SET-IMOEA显著劣于SET-MOEA,“0”表示SET-IMOEA与SET-MOEA之间无明显差异. 表2给出了不同方法求解优化问题的CPU耗时,即T测度值.



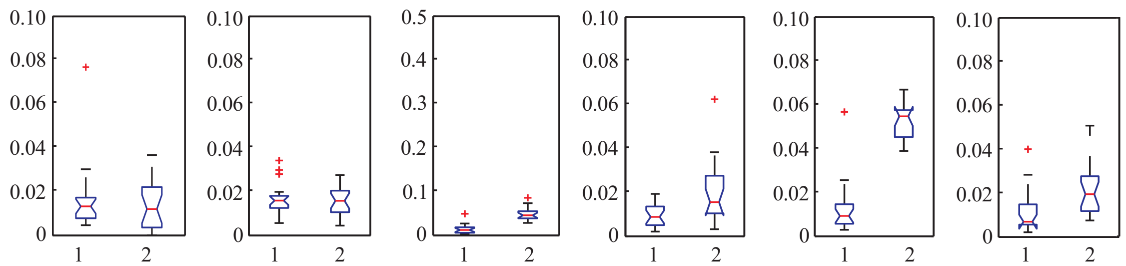
(a) H DTLZ1 5D (b) H DTLZ1 15D (c) H DTLZ2 5D (d) H DTLZ2 15D (e) H DTLZ5 5D (f) H DTLZ5 15D

图 1 两种方法得到的Pareto前沿的H测度值



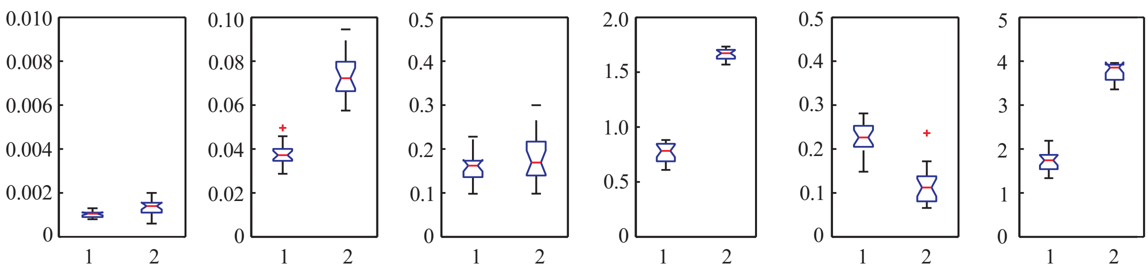
(a) D DTLZ1 5D (b) D DTLZ1 15D (c) D DTLZ2 5D (d) D DTLZ2 15D (e) D DTLZ5 5D (f) D DTLZ5 15D

图 2 两种方法得到的Pareto前沿的D测度值



(a) SP DTLZ1 5D (b) SP DTLZ1 15D (c) SP DTLZ2 5D (d) SP DTLZ2 15D (e) SP DTLZ5 5D (f) SP DTLZ5 15D

图 3 两种方法得到的Pareto前沿的SP测度值



(a) BQ DTLZ1 5D (b) BQ DTLZ1 15D (c) BQ DTLZ2 5D (d) BQ DTLZ2 15D (e) BQ DTLZ5 5D (f) BQ DTLZ5 15D

图 4 两种方法得到的Pareto前沿的BQ测度值

表1　两种方法所得各测度的Mann-Whitney U检验结果

目标函数个数	DTLZ1′				DTLZ2′				DTLZ3′				DTLZ5′			
	H	D	SP	BQ	H	D	SP	BQ	H	D	SP	BQ	H	D	SP	BQ
5	−	+	0	+	−	+	+	0	0	+	0	+	−	+	+	−
15	0	0	0	+	−	0	+	+	0	0	0	+	0	0	+	+

表2　不同方法的T测度　　s

方法	目标函数个数	DTLZ1′	DTLZ2′	DTLZ3′	DTLZ5′
SET-IMOEA	5	134	122	123	125
	15	285	273	272	282
SET-MOEA	5	150	250	265	401
	15	1 726	1 752	1 182	1 219

由图1～图4和表1可以看出: 1) 对比方法SET-MOEA得到的H测度值显著优于本文方法SET-IMOEA得到的, 说明SET-IMOEA的收敛性不如SET-MOEA; 2) SET-IMOEA得到的Pareto前沿的D测度不劣于SET-MOEA, 这是因为, 所提出方法将D测度作为优化目标之一, 而SET-IMOEA没有单独考虑个体的多样性, 同时, 可能也因为SET-IMOEA采用的是自适应的交叉、变异概率, 而SET-IMOEA采用的是固定的交叉、变异概率; 3) 对于优化问题DTLZ1′和DTLZ3′, SET-IMOEA得到的Pareto前沿的SP测度与SET-MOEA无显著性差异, 而对于优化问题DTLZ2′和DTLZ5′, SET-IMOEA得到的Pareto前沿的SP测度显著优于SET-MOEA; 4) SET-IMOEA得到的Pareto前沿的BQ测度不劣于SET-MOEA, 原因是SET-IMOEA将BQ测度用于相同Pareto秩的个体比较.

由表2可以看出, 对于相同的优化问题: 1) 随着目标的增多, 运行时间会不断增加; 2) SET-IMOEA的运行时间远少于SET-MOEA.

通过上述实验得到如下结论: SET-IMOEA得到的Pareto解集的收敛性劣于SET-MOEA, 但是, 它的运行时间却远少于SET-MOEA; SET-IMOEA得到的Pareto优化解集的多样性、分布性、不确定性均优于SET-MOEA.

3　结　论

本文提出了一种用于解决区间高维多目标优化问题的集合进化优化方法. 该方法的创新之处主要体现在以下4个方面: 1) 设计了衡量含区间参数集合个体的分布性测度; 2) 利用3种衡量Pareto优化解集优劣的性能指标, 将区间参数高维多目标优化问题进行了降维转化; 3) 采用不确定性测度进一步比较相同Pareto序值的集合进化个体; 4) 设计了随种群多样性和收敛性自适应变化的交叉、变异概率.

采用基准优化函数对SET-IMOEA和SET-MOEA进行了测试, 实验结果表明, SET-IMOEA综合性能优于SET-MOEA. 如何使种群的收敛性得到进一步的提高, 是接下来需要研究的工作.

参考文献(References)

[1] Deb K, Jain H. An improved NSGA-II procedure for many-objective optimization Part I: Solving problems with box constraints[R]. Indian Institute of Technology, 2012.

[2] Yan G C, Shen B, Wang J H, et al. Way to the multi-objective constrained optimization site of RFID reader in roadway of mine[J]. J of China Coal Society, 2010, 35(9): 1581-1586.

[3] Eskandari H, Geiger C D, Bird R. Handling uncertainty in evolutionary multiobjevtive optimazation:SPGA[C]. Proc of IEEE Congress on Evolutionary Computation. Singapore, 2007: 4130-4137.

[4] Wang H F, Wang M L. A fuzzy multiobjective linear programming[J]. Fuzzy Sets and Systems, 1997, 86(1): 61-72.

[5] Gong D W, Ji X F. Optimizing interval higher-dimensional multi-objective problems using set-based evolutionary algorithms incorporated with preference[J]. Control Theory & Applications, 2013, 30(11): 1369-1383.

[6] Bader J, Brockhoff D, Welten S, et al. On using population of sets in multiobjective optimization[C]. Lecture Notes in Computer Science. Heidelberg: Springer, 2009, 5467: 140-154.

[7] 巩敦卫, 季新芳, 孙晓燕. 基于集合的高维多目标优化问题的进化算法[J]. 电子学报, 2014, 42(1): 77-83.
(Gong D W, Ji X F, Sun X Y. Solving many-objective optimization problems using set-based evolutionary algorithms[J]. Acta Electronica Sinica, 2014, 42(1): 77-83.)

[8] Limbourg P, Aponte D. An optimizaiton algorithm for imprecise multi-objective problem function[C]. Proc of the IEEE Int Conf on Evolutionary Computation. New York: IEEE, 2005: 459-466.

[9] 巩敦卫, 潘凤萍. 自适应遗传算法理论及应用[M]. 徐州: 中国矿业大学出版社, 2003: 20-28.
(Gong D W, Pan F P. Theory and applications of adaptive genetic algorithms[M]. Xuzhou: China University of Mining and Technology Press, 2003: 20-28.)

[10] Schoot J R. Fault tolerant design using single and multicriteria genetic algorithms optimization[D]. Cambridge: Massachusetts Institute of Technology, 1995.

[11] Deb K, Thiele L, Laumanns M, et al. Scalable test problems for evolutionary multiobjective optimization[C]. Evolutionary Mul-tiobjective Optimization: Theoretical Advances and Applications. London: Springer, 2005: 105-145.

[12] Bader J, Zitzler E. HypE: An algorithm for fast hypervolumebased many-objective optimization[R]. Computer Engineering and Networks Laboratory Zurich, Zurich: 2008.

(责任编辑: 李君玲)