

基于爬山算子和适应值共享的改进遗传算法

涂井先, 刘伟

(广东工业大学 应用数学学院, 广东 广州 510006)

摘要: 提出了一种基于爬山算子和适应值共享的改进遗传算法, 将局部搜索算法与遗传算法有效结合, 增强了遗传算法的搜索能力. 爬山算子以黄金分割法为基础, 依次对个体每一维进行优化. 数值实验证明, 改进后的新算法优于当前一些较好的遗传算法. 新算法既有较快的收敛速度, 又能以较大概率搜索到优化问题的全局最优解.

关键词: 遗传算法; 适应值共享; 爬山算子; 黄金分割法

中图分类号: TP18

文献标志码: A

文章编号: 1007-7162(2011)01-0078-04

遗传算法是一种模拟生物进化的智能优化算法. 在解决实际优化问题上, 遗传算法是最有效的优化算法之一^[1-3].

传统的优化算法具有局部搜索能力强但全局搜索能力弱的特点, 而遗传算法具有全局搜索能力强但局部搜索能力弱的特点, 所以如何将传统的优化算法和遗传算法有效结合以改善遗传算法的搜索性能一直是研究的热点^[4-6]. 本文定义了爬山算子, 该算子是以黄金分割法为基础的, 依次对个体每一维进行优化. 黄金分割法是一种解决一维函数极值问题的方法. 在此基础上, 本文将该算子与基于适应值共享的遗传算法有效结合, 提出了基于爬山算子和适应值共享的改进遗传算法. 数值实验证明该算法优于当前一些较好的遗传算法, 增强了遗传算法的局部搜索能力. 新算法既有较快的收敛速度, 又能以较大概率搜索到约束优化问题的全局最优解.

1 基于爬山算子和适应值共享的改进遗传算法

优化问题模型描述如下:

$$\min f(x),$$

其中 $x \in R^n$, R 代表实数域, $x = (x_1, x_2, \dots, x_n)$, $x_i \in [a_i, b_i]$, $i = 1, 2, \dots, n$.

1.1 编码方式

算法采用的是实数编码.

1.2 杂交

杂交算子采用的是算术杂交, 两个父代个体分别为 X_i 和 X_j , r 为杂交系数, 是 $(0, 1)$ 区间上的随机数, 杂交后的两个子代分别是 Y_1, Y_2 . 其中 $Y_1 = r \times X_i + (1-r) X_j$, $Y_2 = (1-r) X_i + r X_j$.

1.3 变异

变异算子采用的是非均匀变异, 它是由 Kusum Deep 和 Manoj Thakur^[7] 提出的.

$X = (x_1, x_2, \dots, x_n)$ 为一个个体, 假设需要变异的基因是 x_i , $i = 1, 2, \dots, n$, 变异后的基因为

$$x'_i = \begin{cases} x_i + \Delta(t, x_i^u - x_i) & \text{如果 } r < 0.5, \\ x_i - \Delta(t, x_i - x_i^l) & \text{其它.} \end{cases}$$

其中 t 为当前种群代数, r 是区间 $[0, 1]$ 上的随机数, x_i^u 和 x_i^l 分别为 x_i 的上下界, 函数 $\Delta(t, y)$ 的值是在 $[0, y]$ 区间上, 具体定义如下:

$$\Delta(t, y) = y \left(1 - u^{(1 - \frac{t}{T})}\right)^b,$$

其中 u 是区间 $(0, 1)$ 上的随机数, T 为最大进化代数, b 为大于零的常数.

1.4 适应值共享

适应值共享是用来维持种群多样性的, 它可以降低聚集在一起的个体的适应值, 它首先被 Holland^[8] 提出, 然后被 Goldberg 和 Richardson^[9] 改进. 给定一个小生境半径 r , 设当前种群代数为 t , 种群规模为 M , 父代种群个体表示为 X_i , $i = 1, 2, \dots, M$, a 为共享参数, 它是大于零的. d_{ij} 为个体 X_i 到 X_j 的欧氏距离, f_i 为个体在未进行适应值共享时的适应值, 它总是大于零的, 且适应度越高说明个体越好, 个体 X_i 和 X_j 的共享值为

$$s_{ij} = \begin{cases} 1 - \left(\frac{td_{ij}}{r}\right)^a & \text{如果 } d_{ij} < \frac{r}{t}, \\ 0 & \text{其他.} \end{cases}$$

个体 X_i 的总共享值为 $s_i = \sum_{j=1}^M s_{ij}$, 适应值共享

后的适应度为 $f_i^* = \frac{f_i}{s_i}$.

收稿日期: 2010-09-30

作者简介: 涂井先 (1987-), 男, 硕士研究生, 主要研究方向为优化算法理论及其应用.

1.5 爬山算子

爬山算子是在每一代的进化操作之前以一定概率对个体进行一次爬山操作得到较好的新个体. 黄金分割法也叫 0.618 法,它是一种基于区间收缩的极小点搜索算法.

在爬山操作中,先把个体的第一维当作变量,其它 $n-1$ 维的值固定,然后确定第一维的搜索区间,搜索区间的选取方法为

$$e = \frac{(b_1 - a_1)n}{2M}, a = \max(a_1, x_1 - e), b = \min(b_1, x_1 + e).$$

$[a, b]$ 就是搜索区间. 确定好搜索区间后,则在这个搜索区间内,对个体的第一维进行黄金分割法以得到一个较好的新个体,新个体替代原来的个体,然后把个体第二维当作变量,其它维的值是固定的,接着确定搜索区间,只要在确定第一维搜索区间的方法中把下标 1 变为 2 即可,然后在这个区间内,对个体的第二维进行黄金分割法以改善个体,依此类推,直到对个体的第 n 维进行黄金分割法.

爬山算子的定义如下:

假设个体为 $X = (x_1, x_2, \dots, x_n)$.

步骤 1: $i = 1$;

步骤 2: 确定搜索区间 $[a, b]$;

其中 $e = \frac{(b_i - a_i)n}{2M}$, $a = \max(a_i, x_i - e)$, $b = \min(b_i, x_i + e)$;

步骤 3: 对一维函数 $f(x_1, x_2, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$ 在区间 $[a, b]$ 上进行一次黄金分割法;

步骤 4: 如果进行黄金分割法后的新个体优于未进行该次局部搜索之前的个体,则将新个体代替旧个体, $i = i + 1$;

步骤 5: 如果 $i \leq n$, 则返回步骤 2, 否则算法停止.

爬山算子是以黄金分割法为基础的,依次对个体的每一维进行黄金分割法. 在进化前期,它能很好地提高个体的适应值. 为了提高算法的效率,搜索区间既不能取变量的整个取值区间,也不能是固定,它必须是包含个体值的某个小区间,个体值不同,搜索区间也不同.

1.6 算法的步骤

步骤 1: 参数的初始化. 设定种群规模 M , 进化代数 t , 最大进化的代数 T , 杂交概率 p_c , 变异概率 p_m .

步骤 2: 设定爬山概率 P_{cl} , 对种群进行爬山算子操作.

步骤 3: 对种群进行算术杂交和非均匀变异操作.

步骤 4: 计算种群每个个体的适应值,然后用适应值共享来降低比较密集的个体的适应值.

步骤 5: 保留最好的个体,用转轮法从父代种群

及其后代中选择 $M-1$ 个个体到下一代父代种群.

步骤 6: 如果 $t \leq T$, $t = t + 1$, 返回第 2 步, 否则运算结束.

2 数值实验

本文选择了 7 个经典复杂的多峰函数优化问题,测试了 3 种算法的收敛速度和 200 代内达到的最好解.

f_1 是 Robachevsky 测试函数,最小值约为 -0.24 .

$$\min f_1(x, y) = x^2 + y^2 - 0.3 \cos(3\pi x) + 0.3 \cos(4\pi y) + 0.3, x, y \in [-1, 1].$$

f_2 是 Schubert 测试函数,有 760 个点达到局部最小值,其中只有 18 个点达到全局最小值,全局最小值约为 -186.73 .

$$\min f_2(x, y) = \left\{ \sum_{i=1}^5 i \cos[(i+1)x + i] \right\} \times \left\{ \sum_{i=1}^5 i \cos[(i+1)y + i] \right\}, x, y \in [-10, 10].$$

f_3 是多峰函数,有 4 个点达到全局最大值,最大值约为 2.118 8, 有大量的点达到局部最大值.

$$\max f_3(x, y) = 1 + x \sin(4\pi x) - y \sin(4\pi y) + \frac{\sin(6\sqrt{x^2 + y^2})}{6\sqrt{x^2 + y^2} + 10^{-15}}, x, y \in [-1, 1].$$

f_4 是 Needle-in-a-haystack 问题,当 $a = 0.05$, $b = 3$, 其全局最大值为 3 600, 有 5 个点达到局部最大值,其中只有一个点达到全局最大值.

$$\max f_4(x, y) = \left(\frac{b}{a + x^2 + y^2} \right)^2 + (x^2 + y^2)^2, x, y \in [-5.12, 5.12].$$

f_5 是另一种多峰函数,其全局最小值为 0, 有无穷多个点达到局部最小值.

$$\min f_5(x, y) = (x^2 + y^2)^{0.25} [\sin^2 50(x^2 + y^2)^{0.1} + 1.0], x, y \in [-5.12, 5.12].$$

f_6 是 Schaffer1 函数,全局最小值为 0.

$$\min f_6(x, y) = 20 + x^2 - 10 \cos(2\pi x) + y^2 - 10 \cos(2\pi y), x, y \in [-5.12, 5.12].$$

f_7 是 Schaffer 函数,全局最小值为 0.

$$\min f_7(x, y) = 0.5 + \frac{\sin^2 \sqrt{x^2 + y^2} - 0.5}{[1 + 0.001(x^2 + y^2)^2]^2}, x, y \in [-10, 10].$$

数值实验是在 Matlab 7.0 中完成,本文将基于爬山算子和适应值共享的改进遗传算法(记为 MCFSGA)和另外两个遗传算法进行比较,分别是文献 [10] 中的基于适应值共享的遗传算法(记为 FS-GA)和文献 [7] 中的基于拉普拉斯交叉算子的遗传算法(记为 HX-NUM). 3 种算法的参数均取: 杂交概

率为 0.8,变异概率为 0.1,在 MCFSGA 和 FSGA 中,小生境半径均为 $\frac{1}{10t}$,其中 t 是当前种群代数,在 MCFSGA 中,爬山概率设为 $0.01 + (0.1 - 0.01) \frac{t}{T}$,其中 T 是最大进化代数,一般地,爬山概率设为

表 1 7 个测试函数收敛精度的比较

测试函数	算法	最好解	平均解	最坏解
f_1	FSGA	-0.240 033 369	-0.240 005 186	-0.239 920 811
	HX-NUM	-0.240 034 485	-0.240 027 915	-0.240 013 683
	MCFSGA	-0.240 034 985	-0.240 034 985	-0.240 034 985
f_2	FSGA	-1.867 300 257E+02	-1.862 874 381E+02	-1.844 698 149E+02
	HX-NUM	-1.867 308 992E+02	-1.867 300 029E+02	-1.867 269 083E+02
	MCFSGA	-1.867 309 088E+02	-1.867 309 088E+02	-1.867 309 088E+02
f_3	FSGA	2.118 469 35	2.101 616 13	2.077 137 64
	HX-NUM	2.118 746 74	2.117 574 36	2.109 793 53
	MCFSGA	2.118 763 42	2.118 763 42	2.118 763 42
f_4	FSGA	2.730 454 35E+03	2.676 238 608E+03	2.529 917 537E+03
	HX-NUM	3.599 907 47E+03	2.791 336 004E+03	2.748 765 480E+03
	MCFSGA	3.599 999 99E+03	2.791 343 222E+03	2.748 782 337E+03
f_5	FSGA	1.548 592E-02	5.119 205E-02	8.752 325E-02
	HX-NUM	1.528 757E-02	3.938 835E-02	6.966 568E-02
	MCFSGA	1.353 025E-04	4.115 232E-04	6.937 712E-04
f_6	FSGA	1.885 306E-05	3.186 750E-03	1.112 975E-02
	HX-NUM	3.967 879E-05	1.149 805E-03	5.501 735E-03
	MCFSGA	5.329 070E-15	7.194 245E-14	4.281 019E-13
f_7	FSGA	4.985 876E-06	4.878 572E-03	1.580 867E-02
	HX-NUM	1.437 878E-07	1.034 498E-04	7.431 841E-04
	MCFSGA	3.885 780E-16	1.798 838E-14	5.673 239E-14

表 2 7 个测试函数达到最优解的平均迭代次数的比较

测试函数	最优值	精度	算法	平均迭代次数
f_1	-2.400 349E-01	1E-05	FSGA	92
			HX-NUM	59
			MCFSGA	22
f_2	-1.867 309E+02	1E-02	FSGA	118
			HX-NUM	68
			MCFSGA	17
f_3	2.118 762	1E-03	FSGA	171
			HX-NUM	30
			MCFSGA	71
f_4	3 600	1E-01	FSGA	-
			HX-NUM	-
			MCFSGA	166
f_5	0	1E-03	FSGA	-
			HX-NUM	-
			MCFSGA	61
f_6	0	1E-04	FSGA	195
			HX-NUM	145
			MCFSGA	37
f_7	0	1E-05	FSGA	192
			HX-NUM	161
			MCFSGA	47

$0.01 + (c - 0.01) \frac{t}{T}$,其中 c 为常数,经测试发现,当 c 的取值范围为 $[0.05, 0.1]$ 时,算法效果比较好.

3 种算法对每一个测试函数都独立重复运行 20 次.为了比较求解精度,本节比较了 20 次实验 200 代后达到的最好解、平均值和最差解(见表 1).为了比较收敛速度,本文比较了 20 次实验收敛到最优解的平均迭代次数(见表 2),并给出了一次实验的进化曲线图(图 1 和图 2).

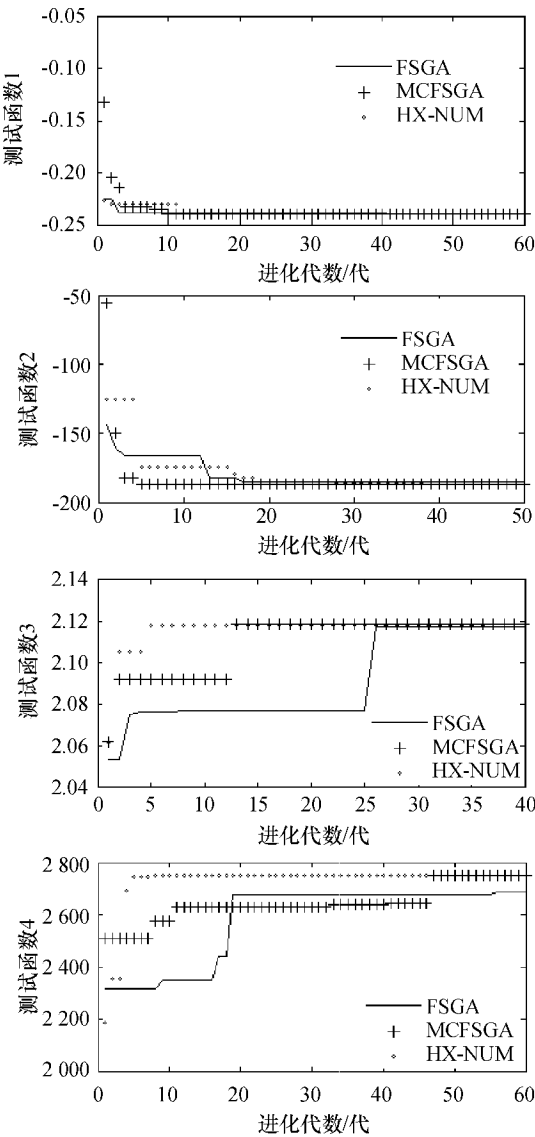
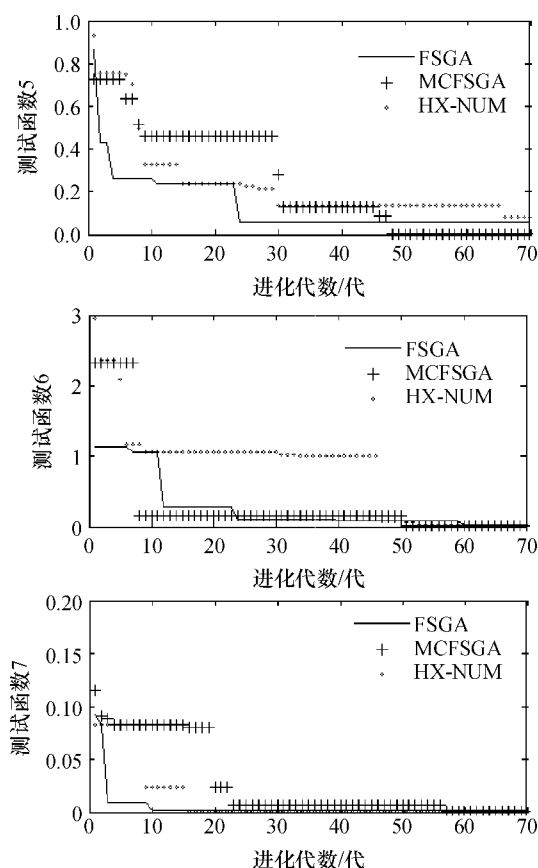


图 1 f_1, f_2, f_3 和 f_4 的进化曲线图

表 2 中的“-”表示测试函数在 20 次实验中均未收敛到最优解.由表 1 可以看出,在优化 7 个测试函数中,新算法的求解精度都比其它两个算法高,说明新算法有很强的稳健性;表 2 表明,除了 f_3 外,新算法的收敛速度都比其它两个算法快,从图 1 和图 2 中的进化过程可以进一步证实这个结论.所以,新算法优于 FSGA 和 HX-NUM,新算法不仅仅能求解到较优的解,而且求解的速度也较快.

图2 f_5 , f_6 和 f_7 的进化曲线图

3 小结

本文提出了一种新的爬山算子,将该算子与基于适应值共享的遗传算法结合,提出了基于爬山算子和适应值共享的改进遗传算法。爬山算子实际上是一个局部搜索算子,它以传统优化算法中的黄金分割法为基础,依次对个体的每一维进行黄金分割

法。它弥补了遗传算法局部搜索能力弱的缺点。数值实验显示基于爬山算子和适应值共享的改进遗传算法是一种便于实现、通用性强、高效稳健的方法,是一种有潜力的全局优化解法。

参考文献:

- [1] Tu Cheng-yuan, Zeng Yan-jun. A New Genetic Algorithm Based Upon Globally-Optimal Choosing and Its Practices [J]. Engineering Science, 2003, 5(2): 28-29.
- [2] Zhou Ming, Sun Shu-dong. Principle of genetic Algorithm and application [M]. Beijing: National Defence Industry Press, 1999.
- [3] Zhao Chuan-xin, Ji Yi-mu. Particle Swarm Optimization for 0/1 Knapsack Problem [J]. Microcomputer Development, 2005 (10): 23-25.
- [4] 唐焕文, 秦学志. 实用最优化方法 [M]. 第3版. 大连: 大连理工大学出版, 1999.
- [5] 周明, 孙树栋. 遗传算法原理及其应用 [M]. 北京: 国防工业出版社, 1999.
- [6] 刘伟, 刘海林. 基于外点法的混合遗传算法求解约束优化问题 [J]. 计算机应用, 2007, 27(1): 238-240.
- [7] Kusum Deep, Manoj Thakur. A new crossover operator for real coded genetic Algorithms [M]. India: Elsevier Inc, 2007.
- [8] Holland J H. Adaptation in natural and artificial systems [M]. Ann Arbor: The MIT Press, 1992.
- [9] Goldberg D E, Richardson J. Genetic algorithms with sharing for multimodal function optimization [M]. Hillsdale: Lawrence Erlbaum, 1987.
- [10] 于歆杰, 王赞基. 对适应值共享遗传算法的分类及评价 [J]. 模式识别与人工智能, 2001, 14(1): 42-47.

An Improved Genetic Algorithm Based on Mountain-climbing Operators and Fitness Sharing

Tu Jing-xian, Liu Wei

(Faculty of Applied Mathematics, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: An improved genetic algorithm, based on mountain-climbing operators and fitness sharing, is proposed. It combines the genetic algorithm with the local searching algorithm effectively, which enhances the searching ability of the genetic algorithm. The mountain-climbing operator, based on the method of golden section, optimizes each dimension of the individual in turn. The results show that the improved algorithm is better than some current algorithms. The new algorithm not only has a rather high convergence speed, but also locates the global optimum with a rather large probability.

Key words: genetic algorithm; fitness sharing; mountain-climbing operator; the method of golden section