

基于两种新型遗传算子的优化组合遗传算法

帅训波¹ 马书南² 邵艳伟¹ 王建忠¹ (1.中国石油勘探开发研究院廊坊分院 地球物理与信息研究所
河北 廊坊 065007; 2.北京工业大学 计算机学院 北京 100022)

摘要: 针对遗传算法中全局搜索与局部搜索之间的矛盾,应用二进制编码对搜索空间描述精细、容易位值计算的特点,从矩阵遗传算子和布尔遗传算子的角度,分别对全局搜索和局部搜索的性能进行改进,并将二者组合应用,构造基于这两种新型遗传算子的优化组合遗传算法,避免了传统遗传算法中杂交率和变异率参数的选取,保证了算法的全局收敛性。实验结果表明,该算法具有更好的整体搜索性能,对应用二进制编码遗传算法求解复杂非线性优化问题具有重要借鉴意义。

关键词: 全局搜索;局部搜索;矩阵遗传算子;布尔遗传算子

Optimization Combination Genetic Algorithm Based on Two New Operators

SHUAI Xun-Bo¹, MA Shu-Nan², SHAO Yan-Wei¹, WANG Jian-Zhong¹

(1.Institute of Geophysics and Information, Langfang Branch of Research Institute of Petroleum Exploration and Development, PetroChina, Langfang 065007, China; 2.College of Computer Science and Technology, Beijing University of Technology, Beijing 100022, China)

Abstract: In order to resolve the contradiction between global searching and local searching of genetic algorithm, matrix genetic operator and Boolean genetic operator are proposed based on binary code describing search space fine and calculating fitness easier. Global searching of genetic algorithm is improved by the matrix genetic operator making good use of colony information to find new space, while local searching of genetic algorithm is improved by the Boolean genetic operator making good use of chromosome information. An optimization combination genetic algorithm is constructed through combining the two new genetic operators. It does not require parameters of crossover probability and mutation probability, which are important to general genetic algorithm. The constructed genetic algorithm is proved to be convergent, and its better searching efficiency is shown by experiments results. It is a significant reference for other optimum problem resolved by genetic algorithm based on binary coding.

Keywords: global searching; local searching; matrix genetic operator; boolean genetic operator

1 引言

应用遗传算法解决大规模、复杂和精度要求高的优化问题时,常陷入局部最优^[1,2],这一问题的原因一方面是遗传搜索过程存在全局搜索和局部搜索之间的矛盾,即遗传算子所具有的全局搜索和局部搜索性能各异^[3,4];另一方面是遗传算法自身存在局部搜索能力不足^[4]。针对全局搜索和局部搜

索间的矛盾,喻寿益等从小生境技术角度,改善了全局搜索性能^[5]。马书南等从遗传算子组合角度,改善了遗传算法的搜索性能^[6-8]。同时,近几年的研究表明,从构造遗传算子的角度改善遗传算法的搜索性能成为研究热点之一^[6-11]。

本文从应用二进制编码方案对搜索空间描述精细、容易位值计算等特点,模仿生物群体智能

收稿时间:2009-10-23;收到修改稿时间:2009-11-27

性，提出矩阵遗传算子，增强全局搜索性能；模拟基因的显隐性，提出布尔遗传算子，改善局部搜索性能；将二者有机结合，构造了一种优化组合遗传算法，从理论上证明了它的良好全局收敛性。大量实验结果表明，该算法具有更好的整体寻优性能。

2 矩阵遗传算子

2.1 矩阵算子提出^[8]

定义 1. (群体矩阵). 在二进制编码遗传算法中，编码长度为 n ，根据给定方法选取 n 个染色体，并随机存储在矩阵 $A_{n \times n}$ 矩阵的各行中，该 $n \times n$ 矩阵称为一个群体矩阵。

定义 2. (矩阵遗传算子). 矩阵遗传算子是实现将群体矩阵 $A_{n \times n}$ 转置，得到转置矩阵 A^T ，以 A^T 每行基因组成新染色体，将 A^T 中 n 个新染色体与 A 中 n 个染色体组成新种群，依据适应度值选取最优的个染色体参与遗传运算。

上述定义是以二进制编码为例进行说明，也适合于基因在给定范围内取值概率均等的编码方案。矩阵遗传算子从群体矩阵的各个染色体中均选一个基因构成新染色体，产生新的 n 个染色体，最后在 $2n$ 个染色体中选优 n 个染色体，与常规算子的遗传运算方式相比，具有更强的全局搜索能力。

2.2 矩阵遗传算子性能分析

应用模式定理分析^[12]，经过矩阵算子运算后，新种群的形成，符合遗传算法的搜索方向，向距离搜索目标更近的样本空间逼近。由算子定义可知，以原种群中 n 个染色体为基础，“勘探”出 n 个样本空间，最后从 $2n$ 个染色体中优选出 n 个染色体，等效于隐形增加了一倍小群体规模。研究结果表明种群规模越大，遗传算法的全局搜索性能越好^[13,14]。因此，应用矩阵遗传算子可以提高遗传算法的全局搜索能力。

下面给出矩阵遗传算子具有良好全局搜索性的实验。

$$f_1 = 4 - (x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1) - 0.4\cos(4\pi x_2))$$

其中 $(x_1, x_2) \in [-1.024, 1.024]$ ，求函数极大值；

$$f_2 = 2x_1^2 + 3x_2^2 - 0.8\sin(2\pi x_1) - 1.2\cos(3\pi x_2)$$

其中 $(x_1, x_2) \in [-1.024, 1.024]$ ，求函数极小值；

$$f_3 = 4 - (x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1) \times \cos(4\pi x_2))$$

其中 $(x_1, x_2) \in [-1.024, 1.024]$ ，求函数极大值；

上述函数极值分别为 4.7、-1.889084 和 4.3，被视为测试算法全局搜索性能的典型问题^[15]。

实验过程是对标准遗传算法改进，应用矩阵遗传算子替代常规杂交算子进行全局搜索，选用单点变异遗传算子进行局部搜索，所构造的矩阵算子遗传算法(MGA)与标准遗传算法(GA)分别对函数 $f_1 - f_3$ 进行求解，构成两组对比实验。均采用二进制编码，锦标赛选择策略，标准遗传算法中杂交概率 $p_c = 0.45$ ，两种算法的变异率 $p_m = 0.01$ ，种群规模 $m = 80$ ，最大进化 500 代。随机抽取 20 次实验，求平均值、二阶样本矩、搜索最好值，应用二阶样本矩计算结果搜索离散度，求解结果对比如表 1 所示：

表 1 MGA 算法和 GA 算法的求解结果对比

算法	函数			
	f_1	f_2	f_3	
GA	平均值	4.678614	-1.828779	4.252583
	二阶矩	0.000129	0.002149	0.000539
	最好值	4.693702	-1.880394	4.291556
MGA	平均值	4.501295	-1.761623	4.132761
	二阶矩	0.042459	0.005827	0.012791
	最好值	4.699737	-1.888849	4.294286

由表 1 的结果对比，GA 和 MGA 均存在不同程度“早熟”。但在二阶样本矩方面，GA 的实验结果小于 MGA 的结果，说明 MGA 算法搜索的结果在平均值周围分布相对离散，MGA 算法可以搜索到更广的样本空间；在搜索到最好值的方面，MGA 算法优于 GA 的搜索结果，这说明 MGA 算法比 GA 算法有更好的全局搜索性能；在搜索平均值方面，MGA 的结果稍差于 GA 的搜索结果，说明 MGA 算法在局部搜索方面差于 GA 算法，这是因为矩阵遗传算子“发现”较广的样本空间，而局部搜索性能对其“开发”不足。

3 布尔遗传算子

3.1 布尔遗传算子提出

定义 3. (异或算子). 设二进制编码长度为 n 的染色体 $s = \alpha_1 \alpha_2 \dots \alpha_{n-1} \alpha_n$ ，按照编码顺序依次取基因 a_i ， $c_i = a_i \oplus a_j$ ，其中 $j = (i+1) \bmod n$ ，将 c_i 与 a_i 一一对应组成染色体 $c = c_1 c_2 \dots c_{n-1} c_n$ ，染色体记为的异或染色体。

定理 1. 给定二进制编码染色体集合 $A \neq \emptyset$ ， $\forall t \in A$ ， t 的异或染色体为 c ，如果 $c \notin A$ ， $A = A \cup \{c\}$ ， A 那么在异或算子运算下是可闭合。

证明：假设在异或算子运算下非闭合，则存在染色体 $x \in A$ ， c 为 x 异或染色体，且 $c \notin A$ ；这与

$A = A \cup \{c\}$ 相矛盾, 故染色体集合在异或算子运算下是可闭合的。

定义 4. (逆序算子)^[6]. 若染色体的反序列基因组有意义, 逆序算子实现由染色体顺序基因组 $s = \alpha_1 \alpha_2 \dots \alpha_{n-1} \alpha_n$ 到其逆序基因组 $s' = \alpha_n \alpha_{n-1} \dots \alpha_2 \alpha_1$, 如果 s' 的适应度优于 s , 则 s' 替代 s , 否则被淘汰。

定义 5. (对偶算子)^[7]. 对偶算子实现依次对基因组 $s = \alpha_1 \alpha_2 \dots \alpha_{n-1} \alpha_n$ 的每个基因位 j , 将 j 取值 α 被其任意等位基因 β 替代, 以形成新等位基因组 $s' = \beta_1 \beta_2 \dots \beta_{n-1} \beta_n$, 如果 s' 的适应度优于 s , 则 s' 替代 s , 否则被淘汰。

定义 6. (布尔遗传算子). 设二进制编码染色体 $s = \alpha_1 \alpha_2 \dots \alpha_{n-1} \alpha_n$, 布尔遗传算子对实现 4 步操作:

- (1) 染色体集合初始化: $B = \{s\}$;
- (2) $\forall t \in B$, t 异或染色体为 c , 如果 $c \notin B$, $B = B \cup \{c\}$ 直到在异或算子运算下为闭集;
- (3) 对集合 B 的每个染色体, 分别进行逆序算子和对偶算子运算, 提高 B 的整体适应度;
- (4) 检索集合 B 中最优染色体 x , 如果 x 的适应度优于 s , 则 s 被 x 所替代, 否则 x 被淘汰。

布尔遗传算子以初始的一个染色体为样本空间的“代表”, 进行该样本空间内的“全局搜索开发”, 寻找此空间的最好“代表”, 实现整个遗传操作的一次局部搜索。为了避免已经是样本空间最优的染色体进行不必要的“全局搜索式开发”, 在染色体编码时, 在首位增加标识基因, 初始为“假”, 经过布尔遗传算子运算后, 其标识基因值为“真”。

3.2 布尔遗传算子性能分析

对于整个种群而言, 应用布尔遗传算子, 对于染色体所代表的每个样本空间进行“开发”, 最后形成种群规模不变的适应度更优的群体, 向各个样本空间的最优个体逼近, 符合遗传算法搜索方向。

选取对测试局部搜索性能有代表性的函数 $f_4 - f_6$, 对布尔遗传算子搜索性能进行实验。

$$f_4 = 0.5 - \frac{\sin^2(x_1^2 + x_2^2) - 0.5}{[1 + 0.01(x_1^2 + x_2^2)]^2}$$

其中 $(x_1, x_2) \in [-2.048, 2.048]$, 求函数极大值;

$f_5(x) = |(1-x) \cdot x^2 \cdot \sin(200\pi x)|$, 其中 $x \in [0, 1]$, 求函数最大值;

$$f_6 = 100(x_1^2 - x_2)^2 + (1 - x_1)^2$$

其中 $(x_1, x_2) \in [-2.048, 2.048]$, 求函数极大值;

上述三个函数均有多个局部极值点, 最优值分别是 1.0、0.148147 和 3905.926227, 被视为测试算法局部搜索能力的典型问题^[16,17]。实验过程中, 选用单点杂交方式, 应用布尔遗传算子替代常规变异算子进行局部搜索, 构成算法(BGA)与标准遗传算法(GA)的求解结果进行比较。杂交概率, 种群规模, 最大进化 500 代, GA 的变异率。随机抽取各自 20 次的求解结果, 计算平均值、二阶样本矩和搜索到的最好值进行比较如表 2 所示:

表 2 BGA 算法和 GA 算法的求解结果对比

算法	函数			
	f_4	f_5	f_6	
GA	平均值	0.999904	0.143328	3905.904724
	二阶矩	0.000001	0.000021	0.000421
	最好值	1.0	0.148147	3905.925008
BGA	平均值	0.999957	0.148136	3905.918155
	二阶矩	0.000000	0.000001	0.000003
	最好值	0.999998	0.148147	3905.921227

由表 2 可知, 在二阶样本矩方面, GA 的实验结果大于 BGA 的结果, 说明 BGA 算法搜索的结果在平均值周围分布相对集中; 在搜索平均值方面, BGA 的结果均优于 GA 的搜索结果, 说明 BGA 算法在局部搜索方面优于 GA 的搜索结果。对于 f_5 函数, 因为布尔遗传算子具有对样本空间“深度开发”功能, 使得 BGA 算法对一些函数可以搜索到全局最优解。而对于 f_4 和 f_6 二个函数, BGA 算法搜索的最好值均稍差于 GA 算法的搜索结果, 这是 BGA 算法应用强的局部搜索策略, 使样本空间模式过于集中, 在一定程度上影响了全局搜索性能。将其与具有良好全局搜索性能的矩阵遗传算子组合应用, 进行优势互补, 具有理想的搜索性能, 而且避免了杂交率与变异率等敏感性参数的选择分析。

4 两种遗传算子组合

常规遗传算法搜索性能与杂交概率和变异概率等敏感性参数有较大关系, 因此杂交率和变异率的选取分析常常是遗传算法设计的重要内容^[18,19]。本文基于该两种新型遗传算子构造的优化组合遗传算法, 避免了传统遗传算法中杂交率和变异率的参数选取, 使遗传算法高效和简洁。

4.1 算法构造

本文构造的优化组合遗传算法(MB_GA)描述:

步骤 1. 根据约束条件, 生成二进制串编码的初始

化种群，初始全局最优解为 p ；

步骤 2. 随机选择与染色体编码长度相等的数个染色体，应用矩阵遗传算子进行全局搜索；

步骤 3. 对矩阵遗传算子搜索得到的个新染色体，分别应用布尔遗传算子进行局部搜索；

步骤 4. 计算种群中各个染色体的适应度，如果本代的最优解优于 p ，则 p 被本代最优解替代；

步骤 5. 判断是否满足算法结束条件，如果满足，则输出最终解 p ，否则转向步骤 2；

步骤 6. 输出最终求解结果。

4.2 收敛性分析

引理 1^[7]. 如果变异概率 $p_m \in (0,1)$ ，交叉概率 $p_c \in (0,1)$ ，同时采用排序选择和局部搜索策略，遗传算法最终收敛到全局最优解。

定理 2. 设染色体的编码长度为 n ，当种群规模大于时 $2n$ ，算法 MB_GA 以概率 1 收敛到全局最优解。

证明：根据 2.2 节的矩阵遗传算子全局搜索性能分析，等效于 n 个染色体“矩阵”方式杂交；由 3.2 节的布尔遗传算子分析，等效于矩阵遗传算子搜索过程中的 $2n$ 个染色体的“布尔遗传”方式变异。设染色体的编码长度为 n ，当种群规模大于时 $2n$ ，本文的优化组合遗传算法交叉概率 $p_c \in (0,1)$ 且变异概率 $p_m \in (0,1)$ ，又因为算法采用选择排序策略，故由引理 1 可得，本文的优化组合遗传算法 MB_GA 以概率 1 收敛到全局最优解。

4.3 实验及结果分析

为了验证 MB_GA 算法具有良好的寻优能力，对 $f_1 - f_6$ 函数进行求解，种群规模 $m = 80$ ，搜索到的最优解与理论极值差别小于 10^{-6} 时，认为找到全局最优解，最大进化代数为 500 代。同样随机抽取计算 20 次的结果，求取各自平均收敛代数、平均收敛次数、最好值和平均值，如表 3 所示：

表 3 MB_GA 算法求解结果

函数	最好值	平均值	收敛次数	收敛代数
f_1	4.7	4.7	20	2.4
f_2	-1.889084	-1.889083	20	55.65
f_3	4.3	4.3	20	2.35
f_4	1.0	0.999999	20	3.8
f_5	0.148147	0.148147	20	15
f_6	3905.926227	3905.926227	20	7.1

由表 3 的实验结果可以看出，MB_GA 算法能在

较少的进化代数内快速搜索到最优解；与表 1 和表 2 的实验结果分别对比，MB_GA 算法具有理想的整体搜索性能。

5 实例应用

本文选用文献[16]中的经典 Camel 函数 f_7 求极值，以测试算法 MB_GA 的搜索性能，分别与文献[20]的算法 1、文献[21]中算法 2、文献[15]中算法 3 和算法 4 求解结果比较。

$$f_7 = (4 - 2.1x_1^2 + \frac{1}{3}x_1^4)x_1^2 + x_1x_2 + (4x_2^2 - 4)x_2^2, \quad (1)$$

其中 $(x_1, x_2) \in [-2.048, 2.048]$ ，求函数极小值；

算法 MB_GA 的种群参数设置，结合文献资料的推荐值进行多次试算，找到比较合适的值。其它算法的种群规模大小、交叉概率和变异概率等参数选择，参考文献[15]中的实验。各算法的参数如表 4 所示：

表 4 5 种算法的参数设置

名称	N	p_m	p_c	a	b
算法 1	100	0.4	0.06	0.8	0.6
算法 2	100	0.4	0.06	1.2	0.6
算法 3	100				
算法 4	100				
MB_GA	100				

在遗传运算过程中，搜索到的最优解与理论极值差别小于 10^{-6} 时，认为找到全局最优解。记录此时的进化代数，以此作为衡量算法收敛速度的指标，最大进化代数为 500 代，每种算法计算 1000 次，统计各自得到最优解的次数及平均收敛代数，对比求解结果如表 5 所示：

表 5 5 种算法的实验结果对比

实验结果	算法 1	算法 2	算法 3	算法 4	MB_GA
收敛次数	830	844	978	1000	1000
收敛代数	175.7	167.5	124.3	24	20.25

由表 4 参数设置和表 5 实验结果对比，MB_GA 算法，给定较小的种群规模参数，具有良好的全局收敛性和较快的收敛速度。

6 结论

本文利用二进制编码容易值计算优点，借鉴群体智能类算法机制，构造了具有良好全局搜索性

能的矩阵遗传算子;模拟基因的显隐性,构造了具有良好局部搜索性能的布尔遗传算子。对这两种新型遗传算子优化组合应用,提高了算法整体寻优能力,而且避免了常规遗传算法中交杂率和变异率等参数选取。采用最优保留策略,从理论上证明了算法全局收敛性。通过对大量典型函数的求解结果分析,并且与其它改进的遗传算法相比较,本文基于两种新型遗传算子的优化组合遗传算法具有良好全局收敛性和较快收敛速度。

参考文献

- 1 周克民,胡云昌.遗传算法计算效率的改进.控制理论与应用,2002,19(5):812 - 814.
- 2 Lin F, Yang QW. Improved genetic operator for genetic algorithm. Journal of Zhejiang University Science, 2002,3(4):431 - 434.
- 3 吴少岩,许卓群.遗传算法中遗传算子的启发式构造策略.计算机学报,1998,21(11):1003 - 1008.
- 4 张文,李祥.基于优化组合的遗传算子的研究与应用.数值计算与计算机应用,2005,26(3):208 - 214.
- 5 喻寿益,郭观七.一种改善遗传算法全局搜索性能的小生境技术.信息与控制,2001,30(6):526 - 530.
- 6 马书南,帅训波,曹凤雪.一种基于逆序算子的优化组合遗传算法.电子技术应用,2006,32(6):19 - 21.
- 7 帅训波,周相广,马书南.一种基于阈值对偶算子的优化组合遗传算法.西华大学学报,2008,27(4):56 - 59.
- 8 帅训波,马书南,周相广,等.一种基于矩阵遗传算子的优化组合遗传算法.小型微型计算机系统,2009,30(5):951 - 954.
- 9 严心池,安伟光,赵维涛,等.遗传算法中“免疫算子”的构造与性能.哈尔滨工程大学学报,2005,26(6):732 - 735.
- 10 汪定伟,王俊伟,王洪峰,等.智能优化算法.北京:高等教育出版社,2007.
- 11 熊伟清,魏平,赵立煜.遗传算法中一个调节算子的研究.小型微型计算机系统,2003,24(3):531 - 533.
- 12 Holland J. Adaptation in Nature and Artificial Systems. 2nd edition, Cambridge, MA: MIT Press, 1992.
- 13 徐晓华,陈陵,陈宏建.可变种群的遗传算法.系统仿真学报,2006,18(4):870 - 872.
- 14 郑金华,史忠植,谢勇.基于聚类的快速多目标遗传算法.计算机研究与发展,2004,41(7):1081 - 1087.
- 15 王成栋,张优云.基于实数编码的自适应伪并行遗传算法.西安交通大学学报,2003,37(3):707 - 710.
- 16 王小平,曹立明.遗传算法---理论、应用与软件实现.西安:西安交通大学出版社,2002.
- 17 李敏强,寇纪淞,林丹等.遗传算法的基本理论与应用.北京:科学出版社,2002.
- 18 Scrinivas M, Patnaik LM. Adaptive probabilities of crossover and mutation in genetic algorithm. IEEE Transactions on Systems, Man and Cybernetics. 1994,24(4):656 - 667.
- 19 曹阳,方强,王国仁,等.基于遗传算法的多连接表达式并行查询优化.软件学报,2002,13(2):250 - 257.
- 20 Michalewicz Z, Janikow CZ, Krawczyk JB. A Modified Genetic Algorithm for Optimal Control Problems. Comp Math Aplic, 1992,23(12):83 - 94.
- 21 黄晓峰,潘文登,陈标华,等.用改进的实数编码遗传算法估计反应动力学参数.高校化学工程学报,1999,13(1):50 - 55.