

## 变异区间自适应调整的遗传算法

赖鑫生,冷明伟,谭国律,周玉林

(上饶师范学院 数学与计算机系,江西 上饶 334001)

(xsl2001\_jx@163.com)

**摘要:**在对遗传算法进化概率进行深入分析的基础上,提出了变异区间随演化进程自适应调整的遗传算法,旨在通过变异区间的自适应调整提高遗传算法的性能。实验结果表明变异区间自适应调整的遗传算法不仅比基本遗传算法有更快的收敛速度,而且有更高的求解精度。

**关键词:**遗传算法;变异区间;进化概率;早熟

**中图分类号:**TP301 **文献标志码:**A

## Genetic algorithm with variable mutation intervals

LAI Xin-sheng, LENG Ming-wei, TAN Guo-lü, ZHOU Yu-lin

(Department of Mathematics and Computer, Shangrao Normal University, Shangrao Jiangxi 334001, China)

**Abstract:** Based on thorough analysis of the evolution probability of Genetic Algorithm (GA), an improved genetic algorithm with variable mutation interval was presented, in which mutation intervals were adjusted adaptively with evolution process. Experimental results show that the proposed algorithm not only converges faster than the original algorithm, but also has a higher precision.

**Key words:** Genetic Algorithm (GA); mutation interval; probability of evolution; premature

### 0 引言

遗传算法(Genetic Algorithms, GA)自Holland等于1975年提出以来<sup>[1]</sup>,因其普适性和较强的函数优化能力而得到广泛应用。它是基于群体的优化算法。选择、交叉、变异等遗传算子作用于当前群体产生新的群体。一般过程是通过选择操作获得双亲个体,以一定概率 $p_c$ 做交叉操作获得两个子代个体,再以一定概率 $p_m$ 对子代个体在一定区间实行变异得到最终的子代个体进入下一代。该区间被称为变异区间。

由于变异算子是一个重要的搜索算子,所以一直是一个研究热点。针对变异算子本身的改进有正态性变异<sup>[2]</sup>,非一致性变异<sup>[3]</sup>,自适应性变异<sup>[4]</sup>,多级变异<sup>[5]</sup>等。然而,变异区间是影响变异算子性能的重要因素。有关如何确定变异区间 $[-A_i, A_i]$ 的研究不多。文献[6]提出自适应伸缩搜索空间的方法。文献[7]提出把变异区间设为整个搜索空间 $[a_i, b_i]$ 的某个百分比,如 $A_i = 0.1(b_i - a_i)$ 。

本文提出变异区间随演化进程自适应调整的遗传算法。对几种常用函数的测试试验表明,变异区间自适应调整的遗传算法不仅能提高收敛速度,而且能有效提高搜索精度。

### 1 群体进化概率分析

遗传算子作用于当前群体产生新的群体。如果新群体中的最优个体比当前群体中的最优个体好,那么当前群体就获得进化。如果新群体中的最优个体比当前群体中的最优个体差,那么当前群体就被退化。

在遗传算法中通常采用精英个体保留法,保留目前为止

找到的最优个体,避免群体退化。只有采用了精英个体保留法的遗传算法,马尔可夫链才具备收敛性<sup>[8]</sup>。

以下讨论的遗传算法采用:实数编码,精英保留法,单点算术交叉操作<sup>[9]</sup>,整个搜索空间中的均匀变异<sup>[7]</sup>。

**定义1** 进化概率 $P$ 。进化概率 $P$ 是指当前群体在遗传操作算子的作用下找到比当前群体中最优个体更好的解个体的概率。

**定理1**

$$P \leq \sum_{g=1}^N \sum_{h=1}^N p_c p_s^g p_s^h \sum_{i=1}^m \prod_{j=1}^n \left[ (1 - p_m) \frac{\Delta_j^{i,g,h}}{L_j^{g,h}} + p_m \frac{\Delta_j^i}{(b_j - a_j)} \right] \quad (1)$$

**证明** 遗传算法群体中的个体可以看作求解空间中的点。以下讨论中的点、个体和解表示同一概念。现假设由 $m$ 个极小的 $n$ 维超立方涵盖了所有适应值比当前群体最优个体更好的点。我们称之为进化超立方。第 $i$ 个进化超立方由 $\Delta_j^i$ 围成( $i = 1, 2, \dots, m$ )。 $\Delta_j^i$ 表示平行于第 $j$ 维的一条线段( $j = 1, 2, \dots, n$ )。单点算术交叉产生的子个体一定位于由 $L_j^{g,h}$ 围成的超立方内,称之为交叉超立方。 $L_j^{g,h}$ 是双亲个体连线在第 $j$ 维上的投影, $g, h$ 表示双亲编号, $g$ 和 $h$ 的取值均为 $[1, N]$ ( $N$ 为群体规模)。共有 $C_N^2$ 个交叉超立方。一个交叉超立方与 $m$ 个进化超立方交叠得到的 $m$ 个超立方称为交叉可进化超立方。如果交叉超立方不与某个进化超立方相交叠,那么相应的交叉可进化超立方的边长就为零。设由父个体 $g$ 和 $h$ 组成的交叉超立方与 $m$ 个进化超立方交叠得到的第 $i$ 个交叉可进化超立方由边 $\Delta_j^{i,g,h}$ 围成, $j = 1, 2, \dots, n$ ,属于第 $i$ 个进化超立方的一部分,且 $0 \leq \Delta_j^{i,g,h} \leq \Delta_j^i$ 。图1表示了在二维空间交叉的情况。

收稿日期:2008-12-29;修回日期:2009-02-25。

基金项目:江西省教育厅科技项目(GJJ09377; GJJ08466);江西省教育厅青年科学基金资助项目(GJJ09616)。

**作者简介:**赖鑫生(1972-),男,江西广昌人,讲师,硕士,主要研究方向:人工智能;冷明伟(1979-),男,山东东明人,讲师,硕士,主要研究方向:模式识别;谭国律(1957-),男,江西余干人,教授,主要研究方向:计算机算法;周玉林(1962-),男,江西上饶人,副教授,主要研究方向:算法分析与设计。

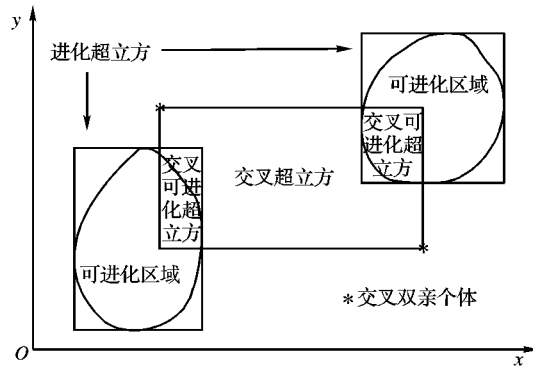


图1 交叉超立方与进化超立方叠的二维示例

只有当前群体子代个体中至少有一个个体出现在  $m$  个进化超立方的任何一个中, 群体才可能进化。因为遗传算法的变异操作对通过交叉产生的子代个体的每一维都以相同的概率  $p_m$  进行变异。因而子个体要进入进化超立方有两种可能。一是通过交叉操作进入进化超立方, 且之后不发生变异; 二是不管之前的交叉操作产生的子个体有没有进入进化超立方, 通过变异操作进入了进化超立方。考虑由父体  $g, h$  产生的一个子代个体进入第  $i$  个进化超立方中的概率为:

$$\prod_{j=1}^n \left[ \frac{(1-p_m)\Delta_j^{i,g,h}}{L_j^{g,h}} + \frac{p_m\Delta_j^i}{(b_j-a_j)} \right] \quad (2)$$

其中  $b_j$  和  $a_j$  分别是搜索空间第  $j$  维的上界和下界,  $p_m$  为变异概率。进入  $m$  个进化超立方中的任何一个的概率为:

$$\sum_{i=1}^m \prod_{j=1}^n \left[ \frac{(1-p_m)\Delta_j^{i,g,h}}{L_j^{g,h}} + \frac{p_m\Delta_j^i}{(b_j-a_j)} \right] \quad (3)$$

设  $g, h$  的被选择概率分别是  $p_s^g$  和  $p_s^h$ ,  $p_c$  为交叉概率。因为群体规模为  $N$ , 且两个双亲个体可以产生两个子代个体, 则下一代群体进入进化超立方的概率为:

$$2 \times \frac{1}{2} \sum_{g=1}^N \sum_{h=1}^N p_c p_s^g p_s^h \sum_{i=1}^m \prod_{j=1}^n \left[ (1-p_m) \frac{\Delta_j^{i,g,h}}{L_j^{g,h}} + p_m \frac{\Delta_j^i}{(b_j-a_j)} \right] = \sum_{g=1}^N \sum_{h=1}^N p_c p_s^g p_s^h \sum_{i=1}^m \prod_{j=1}^n \left[ (1-p_m) \frac{\Delta_j^{i,g,h}}{L_j^{g,h}} + p_m \frac{\Delta_j^i}{(b_j-a_j)} \right] \quad (4)$$

由于进化超立方中只是包含所有的比当前群体中最优个体更好的解, 而不是所有位于超立方中的解都优于当前群体的最优个体。所以, 当前群体获得进化的概率  $P$  满足:

$$P \leq \sum_{g=1}^N \sum_{h=1}^N p_c p_s^g p_s^h \sum_{i=1}^m \prod_{j=1}^n \left[ (1-p_m) \frac{\Delta_j^{i,g,h}}{L_j^{g,h}} + p_m \frac{\Delta_j^i}{(b_j-a_j)} \right] \quad (5)$$

## 2 变异区间自适应收缩的遗传算法

由定理1可以看出, 随着算法的运行, 进化超立方逐渐收缩,  $\Delta_j^i$  逐渐变小,  $\Delta_j^{i,g,h}$  随之变小, 群体的进化概率  $P$  逐渐变小。尽管以整个搜索空间为变异区间的变异操作仍能保证有一定的概率继续进化, 但是由于  $\Delta_j^i$  变小而变异区间不变, 使得  $\frac{\Delta_j^i}{(b_j-a_j)}$  迅速减小, 从而使得变异操作对群体进化概率的贡献急速下降, 同时也产生大量无效搜索, 降低收敛速度。我们试图通过收缩变异区间来提高搜索的收敛速度。

定义2 群体第  $j$  维的最大半径  $d_j$ :

$$d_j = \max_{1 \leq i, k \leq N} \{ind_{ij} - ind_{kj}\} \quad (6)$$

其中:  $ind_{ij}$  和  $ind_{kj}$  表示个体  $i$  和  $k$  的第  $j$  维分量。

变异区间自适应调整的遗传算法 (VMIGA) 取  $d_j$  为当前第  $j$  维的变异区间 ( $j = 1, 2, \dots, n$ )。这样变异区间就会随着群

体的收敛而自适应地收缩调整。当然, 可能导致变异发生在进化区间之外。但是它能保证收敛速度, 这对于快速搜索次优解是非常有效的。

## 3 试验结果

本文采用最为常用的三个测试函数, 均为高维函数, 包含一个单峰和两个多峰函数。三个函数在可行域中的最小值均为或接近零点, 具体如下:

$$f_1 = \sum_{i=0}^{n-1} |x_i| + \prod_{i=0}^{n-1} x_i; x_i \in [-10, 10] \quad (7)$$

$$f_2 = \sum_{i=0}^{n-1} [x_i^2 - 10 \cos(2\pi x_i) + 10]; x_i \in [-5.12, 5.12] \quad (8)$$

$$f_3 = -20 \exp \left[ -0.2 \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} x_i^2} \right] - \exp \left[ \frac{1}{n} \sum_{i=0}^{n-1} \cos(2\pi x_i) \right] + 20 + e; x_i \in [-100, 100] \quad (9)$$

三种函数的解空间均设置为 30 维和 100 维, 种群规模为 10,  $p_c = 0.8$ ,  $p_m = 0.005$ , 每个测试函数独立测试 30 次, 最大函数评价次数设置为 50000 次, 凡是精度达到  $10^{-4}$  则为成功。具体测试结果如表 1 ~ 4 所示。

表1 VMIGA 的测试结果 ( $D=30$ )

函数	最好值	最差值	平均值	成功率/%
$f_1$	8.8868E-26	6.0848E-23	3.5842E-24	100.00
$f_2$	0	0	0	100.00
$f_3$	4.9057E-13	5.1544E-13	4.9531E-13	100.00

表2 GA 的测试结果 ( $D=30$ )

函数	最好值	最差值	平均值	成功率/%
$f_1$	0.0029	0.0876	0.0262	0
$f_2$	0.0023	2.0034	0.1649	0
$f_3$	0.0301	3.2355	0.7896	0

表3 VMIGA 的测试结果 ( $D=100$ )

函数	最好值	最差值	平均值	成功率/%
$f_1$	7.6578E-13	1.0249E-11	4.2051E-12	100.00
$f_2$	5.3433E-12	3.6733	0.6101	53.33
$f_3$	3.9616E-9	20.0104	6.1043	33.33

表4 GA 的测试结果 ( $D=100$ )

函数	最好值	最差值	平均值	成功率/%
$f_1$	0.8248	1.7055	1.2342	0
$f_2$	12.8304	27.3114	20.0270	0
$f_3$	3.2669	20.0273	16.7875	0

图2~4分别描绘了VMIGA和GA在30维和100维函数  $f_1, f_2, f_3$  上独立运行30次的最好解的平均收敛趋势。

从表1~4的数据和图2~4可以看出在对高维空间的一个单峰函数和两个复杂多峰函数的优化上, 变异区间自适应收缩遗传算法在效果上明显优于基本的遗传算法。变异区间的自适应收缩不仅提高了收敛速度, 而且能提高优化精度。

## 4 结语

本文在研究遗传算法进化概率的基础上提出了变异区间自适应调整的遗传算法。变异区间随着群体的演化而自适应

地调整。试验数据表明,变异区间自适应调整的遗传算法在几种测试函数上均优于基本遗传算法,能有效提高遗传算法的搜索速度和精度。变异区间自适应调整的遗传算法为解决高维函数的优化问题及优化算法的改进和发展提供了一条新的思路。

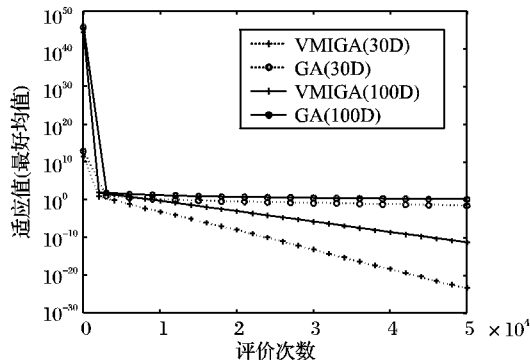


图2 VMIGA 和 GA 在 30 维和 100 维的  $f_1$  上的比较

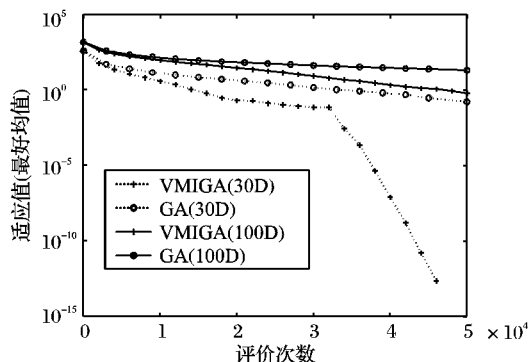


图3 VMIGA 和 GA 在 30 维和 100 维的  $f_2$  上的比较

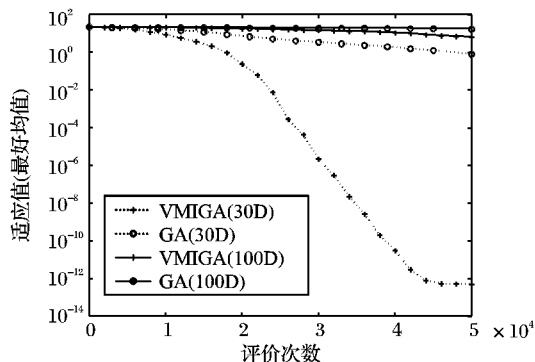


图4 VMIGA 和 GA 在 30 维和 100 维的  $f_3$  上的比较

#### 参考文献:

- [1] JOHN H. Adaptation in natural and artificial systems [M]. Ann Arbor, MI: The University of Michigan Press, 1975.
- [2] PAUL S H. Evolution and optimization seeking [M]. New York: John Wiley & Sons, 1995.
- [3] ZBIGNIEW M. Genetic algorithms + data structures = Evolutionary programs [M]. Berlin: Springer-Verlag, 1992.
- [4] PANZHENG - JUN, KANG LI - SHAN, NIE SI - XIANG. Evolving both the topology and weights of neural networks [J]. Parallel Algorithms and Applications, 1996, 9(3/4): 299 - 307.
- [5] HEINZ M, DIRK S V. Predictive models for the breeder genetic algorithm [J]. Evolutionary computation, 1993, 1(1): 25 - 49.
- [6] 钟伟才, 刘静, 焦李成. 遗传算法中自适应伸缩搜索空间的方法 [J]. 系统工程与电子技术, 2004, 26(2): 245 - 247.
- [7] 潘正君, 康立山, 陈毓屏. 演化计算 [M]. 北京: 清华大学出版社, 1998.
- [8] 刘静, 钟伟才, 刘芳, 等. 组织进化数值优化算 [J]. 计算机学报, 2004, 27(2): 157 - 167.
- [9] 王欣星, 周晖. 基于混合遗传-模拟退火算法的日前购电策略研究 [J]. 水电能源科学, 2007, 25(5): 122 - 126.

(上接第 1528 页)

保持为 11 周,无法利用加班压缩项目总工期。当加班系数为 0.25 时,所得最优解如图 3 所示,其项目总工期为 10 周。可见该项目如在第 1~3 周及第 6 周安排加班 10 个工时,则能保证项目提前 1 周完成。

因此,对于 FWCPSP 问题而言,在允许适当加班的情况下,如何加速项目进度以实现最小化工期就成为一个新的问题。本文所建立的模型与算法有助于求解此类问题。

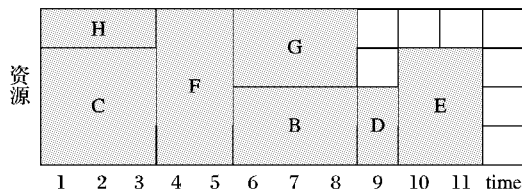


图2 最优项目进度计划

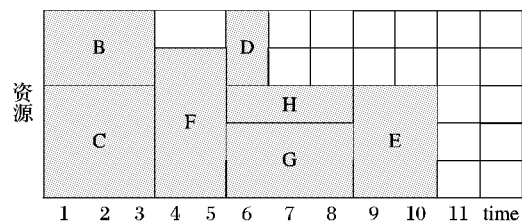


图3 允许加班情况下的项目进度计划

## 4 结语

本文结合软件工程项目实践情况,提出了一类新的项目调度问题,即在考虑加班的情况下如何合理调配人力资源使得项目在最短工期内完工。同时基于并行项目进度机制设计

了将蚁群算法用于求解这一类柔性工时约束下的项目调度问题。本文研究尚可以进一步深入,例如尚未考虑加班费,如果考虑到加班费则还需要在支付加班费和提前完工之间进行权衡,这也将是后续研究的主要方向。

#### 参考文献:

- [1] KOLISCH R, SPRECHER A, DREXL A. Characterization and generation of a general class of resource constrained project scheduling problems [J]. Management Science, 1995, 41(10): 1693 - 1703.
- [2] VALIS V, BALLESTIN F, QUINTANILLA S. Justification and RCPS: A technique that pays [J]. European Journal of Operational Research, 2005, 165(2): 375 - 386.
- [3] ALVAREZ-VALDES R, TAMARIT J M. Heuristic algorithm for a resource-constrained project scheduling: A review and an empirical analysis [C]// Advances in Project Scheduling. Amsterdam: Elsevier, 1989: 113 - 134.
- [4] 胡明娟, 丁建定. 全球化、SA8000 与我国企业社会责任标准的建立 [J]. 华中科技大学学报: 社会科学版, 2005, 19(5): 115 - 120.
- [5] 刘士新, 宋健海, 唐加福. 蚁群最优化: 模型、算法及应用综述 [J]. 系统工程学报, 2004, 19(5): 496 - 502.
- [6] HARTMANN S, KOLISCH R. Experimental evaluation of state-of-the-art heuristics for the resource-constrained project scheduling problem [J]. European Journal of Operational Research, 2000, 127(2): 394 - 407.
- [7] MERKLE D, MIDDENDORF M, SCHMECK H. Ant colony optimization for resource-constrained project scheduling [J]. IEEE Transactions on Evolutionary Computation, 2002, 6(4): 333 - 346.
- [8] PATTERSON J H. A comparison of exact approaches for solving the multiple constrained resource, project scheduling problem [J]. Management Science, 1984, 30(7): 854 - 867.