

基于实数编码的多种群演化遗传算法*

刘淳安, 杨建宏

(宝鸡文理学院 数学系, 陕西 宝鸡 721007)

摘 要:针对标准遗传算法(SGA)在实际应用中存在早熟收敛、精度较差及运算速度慢的缺点,提出了一种基于实数编码的多种群演化遗传算法(RMGA)。实验结果表明该算法能够有效的提高全局搜索能力和局部快速搜索能力,对改进 SGA 的缺点是十分有效的。

关键词:遗传算法;实数编码;多种群

中图分类号:TP18

文献标识码:A

文章编号:1007-1261(2005)02-0085-03

Multi-population evolution genetic algorithm based on real coding

LIU Chun-an, YANG Jian-hong

(Dept. Math., Baoji Univ. Arts & Sci., Baoji 721007, Shaanxi, China)

Abstract: In order to improve the problems of simple genetic algorithm (SGA) such as premature convergence, rough result and low speed of calculation, a multi-population evolution genetic algorithm based on real coding (RMGA) is presented. Application examples show that RMGA is more efficient and robust than SGA.

Key words: genetic algorithm; real coding; multi-population

MSC 2000: 90C30; 90C29

1 引 言

遗传算法(Genetic Algorithm, 简称 GA)是模拟生物在自然环境中的遗传和进化而形成的一种全局优化概率搜索算法,以其简单、通用、鲁棒、不依赖函数模型、适用于处理较为复杂的优化问题,在函数优化、自动控制、人工智能等众多领域得到了广泛而成功的应用^[1, 2]。

标准遗传算法(Simple Genetic Algorithm, 简称 SGA)是采用二进制编码方式,利用固定的选择、交叉、变异遗传操作模拟自然进化,完成对问题的寻优求解。但是它对于连续优化问题局部搜索能力较差,对多变量、高精度优化问题由于个体串码较长往往会导致计算量大,精度低,计算时间长及过早收敛^[3]。基于此,本文提出了一种基于实数编码的多种群遗传算法(Multi-Population Evolution Genetic Algorithm Based on Real Cod-

ing,以下简称 RMGA),有效地提高了全局搜索能力和局部快速搜索能力,降低了计算量,最后的实验结果表明了 RMGA 算法的有效性。考虑优化问题的一般形式:

$$\min_{\mathbf{x} \in [L, U] \subset \mathbb{R}^n} f(\mathbf{x})$$

其中:超立方体 $[L, U] = \{\mathbf{x} = (x_1, x_2, \dots, x_n) \mid l_i \leq x_i \leq u_i, i = 1, 2, \dots, n\}$ 为问题的搜索空间, $[l_i, u_i]$ 为可行解 $\mathbf{x}$ 的第 i 个分量 x_i 的变化区间, n 为优化变量数目。

2 RMGA 算法

2.1 适应度评价

设 $f_t(\mathbf{x})$ 表示第 t 代种群中个体 $\mathbf{x}$ 对应的目标值,定义适应度函数为 $F_t(\mathbf{x}) = C_{\max} - f_t(\mathbf{x})$,

* 收稿日期:2004-11-2. E-mail:chunanliu@yahoo.com.cn

基金项目:国家自然科学基金项目(60374063);

宝鸡文理学院院级科研计划项目(JK2439);宝鸡文理学院中青年科研计划项目(QK2407)

作者简介:刘淳安(1972-),男,陕西淳化人,讲师,硕士,研究方向:最优化、进化算法。

其中 $C_{\max}$ 是一适当的相对比较的常数。

2.2 杂交算子

文献[4]定义的算术交叉算子中有一个交叉系数 α , 其取值范围为 $[0, 1]$, 文献[5]对其进行了改进, α 取值不再限制于 $[0, 1]$, 取得了较好效果, 但在文献[5]中, α 仍然采用定值, 本文给出一个随进化代数增加而变化的实变数 δ , 促使交叉系数 α 随 δ 变化而自适应地调节, 设以概率 p_c 对第 t 代的第 j 个种群 $P_j(t)$ ($j = 1, 2, \dots, m$) 中个体 $x_i^j = (x_1^j, x_2^j, \dots, x_n^j)$, $y_i^j = (y_1^j, y_2^j, \dots, y_n^j)$, 进行杂交, 则杂交后所得后代分别为

$$\begin{cases} x_{i+1}^j = \alpha x_i^j + (1 - \alpha) y_i^j \\ y_{i+1}^j = \alpha y_i^j + (1 - \alpha) x_i^j \end{cases}$$

其中 $\alpha = 0.5r^{(1-\delta)^\lambda}$, $\delta = 1 - \frac{|F_t(x_i^j) - F_t(y_i^j)|}{\max_{x \in P_j(t)} F_t(x)}$,

r, λ 均为决定影响 α 程度的参数(本文后面的计算取 $r = 6, \lambda = 0.5$), $\max_{x \in P_j(t)} F_t(x)$ 表示从第一代到第 t 代为止第 j 个种群 $P_j(t)$ 中所得个体适应度的最大值, 从上式可以看出, 在进化初期, 个体分布比较分散, α 逼近于 3 (当 $r = 6$ 时), 这样可以增大算法的搜索空间, 在进化后期, 种群分布集中于最优解附近, α 逼近于 0.5, 这样可以加快收敛速度。

2.3 变异算子

设第 t 代以概率 p_m 对第 j 个种群 $P_j(t)$ 中个体 x 进行变异, 则对 x 执行非均匀变异^[4]操作, 如果 $x = (x_1, \dots, x_k, \dots, x_n)$ 的元素 x_k 被选择变异, $x_k \in [l_k, u_k]$, 则变异结果为 $x' = (x_1, \dots, x'_k, \dots, x_n)$, 其中 $k = 1, 2, \dots, n$ 。

$$x'_k = \begin{cases} x_k + \Delta(t, u_k - x_k) & \text{rand}(0, 1) \leq 0.5 \\ x_k - \Delta(t, x_k - l_k) & \text{rand}(0, 1) > 0.5 \end{cases}$$

$\Delta(t, y) = y \cdot (1 - r^{(1-\frac{t}{T})^\lambda})$, 式中: T 为最大代数, λ 为决定非均匀度的系数参数, $\Delta(t, y) \in [0, y]$, 且随着代数的增加, $\Delta(t, y) \rightarrow 0, r \in \text{rand}(0, 1)$, 这样算子在初始阶段可以均匀地搜索空间, 而在后面阶段非常局部化。

2.4 RMGA 算法的提出与实现

2.4.1 RMGA 算法思想

该算法从问题的搜索空间入手构造了多个种群和一个外存储器, 分两个阶段进行, 第一阶段对产生的多个种群并行独立进化, 当进化到一定代数时, 把各种群中最好的个体输入到一个外存储器, 第二阶段对外存储器中的个体以同样的方式单独进化, 进化到一定代数后结束, 因为问题的最优解

有可能出现在第一阶段的某一种群, 因此对第二阶段进化结束后所得最优解与外存储器中最优解作比较, 其中适应度值最大者即为问题的最优解。

2.4.2 RMGA 算法流程

(1) (编码) 采用实数编码。

(2) (初始多种群) 给定杂交概率 p_c , 变异概率 p_m , 多个种群的规模 N , 对超立方体 $[L, U]$ 进行 m 等份划分, 不妨设 $[L, U] = \bigcup_{j=1}^m [L_j, U_j]$ 且 $[L_j, U_j] \cap [L_i, U_i] = \emptyset, \forall i \neq j, (i, j = 1, 2, \dots, m)$ 。这样在划分得到的 m 个小超立方体 $[L_j, U_j]$ 上分别各自独立均匀产生 N 个个体 $x_{j1}, x_{j2}, \dots, x_{jN}$ 共同组成 m 个并行的初始多种群 $P_j(0)$ ($j = 1, 2, \dots, m$), 令 $t = 0$ 。

(3) (杂交) 对每个种群 $P_j(t)$ ($j = 1, 2, \dots, m$) 中的父代个体以概率 p_c 按 2.2 的方式并行进行杂交, 用杂交后所得的后代代替其父代, 并与 $P_j(t)$ 中没有参与杂交的父代个体共同组成种群 $P_j(t)$ 的临时父代。

(4) (变异) 对上步产生的各种群 $P_j(t)$ ($j = 1, 2, \dots, m$) 的临时父代以概率 p_m 用 2.3 的变异策略并行进行变异, 用变异后的个体代替其父代, 并与没有参与变异的个体组成种群 $P_j(t)$ 的临时后代 $P'_j(t)$ 。

(5) (多种群选择并保优) 从各个种群 $P_j(t)$ ($j = 1, 2, \dots, m$) 的临时后代 $P'_j(t)$ 中采用适应度比例选择法选出 N 个个体, 同时把选出的 N 个个体中最差的个体 $x_j^*(t) = \arg \max_{x \in P_j(t)} F_t(x)$ 用种群 $P_j(t)$ 中最好个体 $x'_j(t) = \arg \min_{x \in P_j(t)} F_t(x)$ 替换掉, 这样生成种群 $P_j(t)$ 的下一代 $P_j(t+1)$, 令 $t = t + 1$, 当满足预设的第一阶段迭代次数时, 转(6), 否则, 转(3)。

(6) (临时储存器) 从步(5)产生的每个种群 $P_j(t)$ ($j = 1, 2, \dots, m$) 的后代 $P_j(t+1)$ 中求出适应度最大的个体 $x_j^{t+1} = \arg \min_{x \in P_j(t+1)} F_{t+1}(x)$ (共 m 个) 存入临时储存器 $c(t)$ 。

(7) 把 $c(t)$ 看作过渡种群, 对其依次执行 2.2, 2.3 的进化操作, 所得子代个体组成临时过渡种群 $c'(t)$, 再次采用适应度比例选择法从 $c'(t)$ 中选出 m 个个体组成 $c(t)$ 的下一代 $c(t+1)$ 。

(8) 当满足预设的第二阶段迭代次数时, 停, 输出 $c(t) \cup c(t+1)$ 中适应度值最大的个体 x^* , 否则, 令 $t = t + 1$, 继续执行(7)的进化操作。

2.4.3 RMGA 算法分析

对于标准遗传算法(SGA),其优化变量的搜索空间的大小影响着算法的收敛速度,且初始解不容易分散到整个解空间,容易陷入局部最优。对多种群演化遗传算法(RMGA),由于每个子种群同时扫描搜索区域的不同部分,这样加强了其局部快速搜索能力,同时 RMGA 是通过 2 个阶段来实现,在第一阶段各子种群把不同区域上搜索到的局部最优解传递给第二阶段共同组成一个新种群,显然该新种群中的个体具有很大差异性,且携带好的基因,这样又加强了全局搜索能力,避免了早熟收敛。从这两方面说明 RMGA 具有较好全局搜索能力和局部快速搜索能力是可行的。

3 数值模拟

为了测试 RMGA 的有效性,本文采用 3 个不同性质的测试函数对其进行数值测试,并与标准遗传算法(SGA)的测试结果进行了比较。

3.1 试验函数

$$\text{例 1} \quad \min f(\mathbf{x}) = 100(x_1 - x_2)^2 + (1 - x_1)^2$$

$$s. t.: -2.048 \leq x_1, x_2 \leq 2.048。$$

$$\text{例 2} \quad \min f(\mathbf{x}) = \sin^2(3\pi x_1) + (x_1 - 1)^2(1 + \sin^2(3\pi x_2)) + (x_2 - 1)(1 + \sin^2(2\pi x_2))$$

$$s. t.: -10 \leq x_1, x_2 \leq 10。$$

$$\text{例 3} \quad \max f(\mathbf{x}) = 0.5 -$$

$$\frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{[1 + 0.001(x_1^2 + x_2^2)]^2}$$

$$s. t.: -100 \leq x_1, x_2 \leq 100。$$

其中,例 1 是 Rosenbrok 函数,单峰非凸病态(如图 1 所示),难以极小化,在(1,1)点达到最小值 0。例 2 是一多峰呈皱褶状函数(如图 2 所示),理论分析有 900 个局部极小点,在点(1,1)达到最小值 0。例 3 称作 Schaffer 函数,它是一个多峰函数,函数以最小值为中心有多个起伏(如图 3 所示),其中只在(0,0)取得最大值 1。

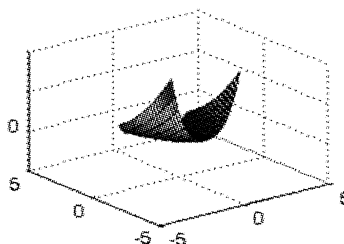


图 1 例 1 函数图像

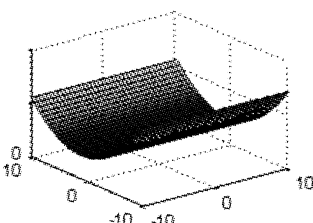


图 2 例 2 函数图像

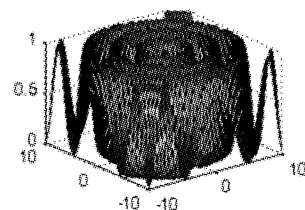


图 3 例 3 函数在[-20, 20] 区间图像

3.2 模拟结果

在计算中 RMGA 操作参数设计如下,各个种群的规模 $N = 40$,杂交概率 $p_c = 0.8$,变异概率 $p_m = 0.05$,多种群的个数 $m = 64$,记录所求函数最好值(记作 Best Result),函数值的平均值

(Mean),函数的最差值(记作 Worst Result),及平均迭代次数(记作 num)(RMGA 的迭代次数包括 2 个阶段之和),并和 SGA 的计算结果进行了比较,结果见表 1 和表 2。

表 1 RMGA 的计算结果

Problem	num	Optimal Value	Best Result	Mean	Worst Result
1	15	0	0.000 000	0.000 015	0.000 028
2	20	0	0.000 000	0.000 064	0.000 086
3	29	1	1.000 000	0.845 228	0.485 479

表 2 SGA 的计算结果

Problem	num	Optimal Value	Best Result	Mean	Worst Result
1	35	0	0.000 012	0.000 087	0.00 0213
2	48	0	0.000 104	0.000 246	0.000 487
3	67	1	0.853 426	0.524 876	0.165 236

比较表 1、表 2 可以看出, RMGA 算法比 SGA 算法能用更少的迭代次数找到更精确的最优值。

4 结 语

本文提出了一种基于实数编码的多种群演化遗传算法(RMGA),对标准遗传算法(SGA)在实

际应用中存在的早熟收敛、精度较差及运算速度慢等缺点进行了极大的改进,有效的提高了全局搜索能力和局部快速搜索能力,最后的实验结果表明 RMGA 算法是非常有效的。

(下转第 103 页)

$$w^{a+c} - w^{a+d} - w^{b+c} - w^{b+d} - w^{c+d+k} - w^{a+b+k},$$

$$R = -2w^{2m+2} - 4w^{2m} + 4w^{m+1} + 4w^m + w^3.$$

根据对称性可设 $a = \min\{a, b, c, d\}$ 。

情况 1: $m = 2$, 则 $R = -2w^6 - 4w^4 + 5w^3 + 4w^2$, 最低次项为 $4w^2$, S 的最低次项为 w^a, w^b, w^c, w^d 和 w^{k+1} 中的一项或几项之和, 由 $S = R$, 可分下面几种情况:

1^0 : $a = b = c = d = m$, 由 $a + b + c + d + k = 4m + 2$ 有 $k = 2$, 这时 $H \cong G$ 。

2^0 : $k + 1 = 2$, 即 $k = 1$, 不妨设 $a = b = c = 2$, 由 $a + b + c + d + k = 4m + 2$ 有 $d = 3$, 去掉 R 和 S 相同的项后, 右边的最高次项为 $-2w^6$, 左边的最高次项为 $-w^6$, 与 $R = S$ 矛盾。

3^0 : $k \geq 3$, 这时必有 $a = b = c = d = 2$, 但 R 和 S 中 w^3 的系数不相等与 $R = S$ 矛盾, 因此当 $m = 2$ 时, 若 H 和 G 色等价, 则 $H \cong G$ 。

情况 2: $m \geq 3$,

1^0 : $m = 3$, 此时 R 的最低次项为 $5w^3$, 所以只有 $a = b = c = d = m = 3, k = 2$ 时, $R = S$, 这时 $H \cong G$;

2^0 : $m > 3$, R 的最低次项为 w^3 , S 的最低次项为 w^{k+1} 或 w^a , 即 $k + 1 = 3$ 或 $a = 3$ 。

$2^0.1$: 若 $k + 1 = 3$, 即 $k = 2$ 时, $a + b + c + d = 4m$, 去掉 w^3 后, R 的最低次项为 $4w^m$, 在 R 内部不能被抵消, 所以只有 $a = b = c = d = m$ 时, $R = S$, 因而 $H \cong G$ 。

$2^0.2$: $a = 3, k + 1 \geq 4, b, c, d \geq 4$, 则 $3 + b + c + d + k = 4m + 2$, 即: $b + c + d + k = 4m - 1$,

去掉 w^3 后, S 的最低次项为 w^4 所在项, 在 S 内部不能被抵消, R 的最低次项为 $4w^m$, 所以 $m = 4$, 由 $R = S$, 有 $b = c = d = 4$ 或 $k + 1 = 4, b, c, d$ 中有 2 个值为 4。

若 $b = c = d = 4$, 由 $b + c + d + k = 4m - 1$, 有 $k = 3$, 而这时 S 的最低次项为 $5w^4$ 与 R 的最低次项不相等, 与 $R = S$ 矛盾。

若 $k + 1 = 4, k = 3$, 这时 b, c, d 中必有 2 个值为 4, 但由 $b + c + d + k = 4m - 1$ 可得另一个值也为 4, 这时 $R \neq S$ 。

因此 $a \neq 3$ 。

综合上面 2 种情况: $H \cong G$ 。

参考文献:

- [1] CHAO C Y, ZHAO L C. Chromatic polynomials of a family of graphs[J]. *Ars Combinatoria*, 1983, 15:111-129.
- [2] KOH K M, TEO K L. The search for chromatically unique graphs[J]. *Graphs and Combinatorics*, 1990, (6):259-285.
- [3] XU S J, LIU J J, PENG Y H. The chromaticity of s-bridge graphs and related graphs[J]. *Discrete Mathematics*, 1994, 135:349-358.
- [4] PENG Y H. Another family of chromatically unique graphs[J]. *Graphs and Combinatorics*, 1995, (11):285-291.
- [5] KOH K M, TEO K L. The search of chromatically unique graphs-II [J]. *Discrete Mathematics*, 1997, 172:59-78.

(编校:李哲峰)

(上接第 87 页)

参考文献:

- [1] GOLDBERY D E. Genetic Algorithms in Search, Optimization and Machine Learning [M]. New York: Addison-Wasley, 1989. 1-18.
- [2] ELIAS J G. Genetic Generation of Connection Patterns for A Dynamic Artificial Neural Network[C]. Combinations of Genetic Algorithms and Neural networks International Workshop, 1992. 38-54.
- [3] 周明, 孙树林. 遗传算法原理及应用[M]. 北京: 国防工业出版社, 1999.
- [4] MICHAŁ WICZ Z, JANIKOW C Z, KRAWCZYK J B. A modified genetic algorithm for optimal control problems[J]. *Computers Math Appl*, 1992, 23(12):83-94.
- [5] 黄晓峰, 潘立登, 陈标华. 用改进的实数编码遗传算法估计反映动力学参数[J]. 高校化学工程学报, 1999, 13(1):55-55.

(编校:李哲峰)