解无约束非线性全局优化的一种新进化算法及其收敛性

王宇平[1], 焦永昌[2], 张福顺[2]
(1. 西安电子科技大学理学院, 陕西西安 710071; 2. 西安电子科技大学天线研究所, 陕西西安 710071)

摘 要: 进化算法是解复杂非线性规划问题的一种新型有效方法, 但现有方法的计算量通常较大. 为减小计算量, 提高算法的效率, 本文利用均匀设计来构造新的高效进化算法, 新的进化算法本身具有类似于传统优化技术中的局部搜索功能, 因此它能非常有效地搜索解空间, 保持种群的多样性, 减小计算量. 文中还证明了新算法的全局收敛性. 最后的模拟结果表明, 新算法计算量小且收敛速度快.

关键词: 进化计算; 均匀设计; 非线性规划

中图分类号: TP18 **文献标识码**: A **文章编号**: 0372-2112 (2002) 12-1867-03

A New Evolutionary Algorithm for Unconstrained Nonlinear Global Optimization Problems and Its Convergence

WANG Yu-ping[1], JIAO Yong-chang[2], ZHANG Fu-shun[2]
(1. *Faculty of Science, Xidian University, Xi'an, Shaanxi 710071, China*;
2. *Institute of Antennas and EM Scattering, Xidian University, Xi'an, Shaanxi 710071, China*)

Abstract: Evolutionary algorithm is a new kind of effective algorithms for very complex nonlinear programming problems. However, the amount of their computation is often large. To reduce the amount of the computation and make the algorithms more effective, the uniform design is used to design new effective evolution operator. The new evolution operator has the local-search property similar to that in traditional optimization techniques, thus it can generate a diversity of population and explore the search space effectively with little computation. Moreover, the new algorithm is globally convergent. The numerical results also show the effectiveness of the new algorithm with its less computation, higher convergent speed for all test functions.

Key words: evolutionary computation; uniform design; nonlinear programming

1 引言

在电子工程等领域有很多复杂的全局优化问题, 它们通常很难用传统的优化方法求解. 进化算法是求解这些问题最有效的一类方法[1]. 它通常能跳出局部最优解而求出全局最优解[1~3]. 但其计算量一般较大. 为减少计算量, 提高算法效率, 已有一些新的方法出现[4~6], 然而, 它们的效果仍不尽理想. 均匀设计是统计和数论结合而产生的一种新的试验设计方法[7], 其主要思想是在所考虑的空间用最少的计算量设计一个均匀散布的点集. 本文利用它来构造新的高效进化算法, 新的进化算法产生的后代均匀散布在其父母附近, 因而具有类似于传统优化技术中的局部搜索功能, 因此, 新的进化算法能非常有效地搜索解空间, 保持种群的多样性. 另外, 由于新算法本身具有局部搜索的功能, 运行时不需再进行局部搜索, 因而和其它方法相比, 不仅有效, 而且计算量小. 文中还对维数均为 30 的 12 个困难的标准试验函数进行了计算, 与文献中已有结果[4~6]的比较表明, 新算法不仅能找出全局最优解或更好的近似全局最优解, 而且计算量比已有方法小得多.

2 新的杂交算子

下面介绍如何用均匀设计[7]构造高效的杂交算子. 均匀设计可产生 q 个 n 维整数点:

$$(a_{k1}, a_{k2}, \cdots, a_{kn}) = k(1, b, \cdots, b^{n-1}) \pmod q \qquad (1)$$

$k = 1 \sim q$, 其中 q 为素数, b 为适当正整数(可由[7]中给出的表查出)满足: $n \le q-1$, $1 < b < q$, 且 $b^1, b^2, \cdots, b^{n-1} \pmod q$ 是不同的数. 假设被选定进行杂交的两点为 $\boldsymbol{X} = (x_1, x_2, \cdots, x_n)^T$ 和 $\boldsymbol{Y} = (y_1, y_2, \cdots, y_n)^T$, 新的杂交算子步骤为:

步 1 令 $w_i = \min\{x_i, y_i\}$, $z_i = \max\{x_i, y_i\}$, $i = 1, 2, \cdots, n$, 定义

$$\boldsymbol{W} = (w_1, w_2, \cdots, w_n)^T, \boldsymbol{Z} = (z_1, z_2, \cdots, z_n)^T \qquad (2)$$

步 2 选适当大小的素数 q, 查文[7]中表可得对应的 b, 用式(1)求 q 个点.

步 3 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的 q 个后代 $\boldsymbol{P} = \{\boldsymbol{X}^1, \boldsymbol{X}^2, \cdots, \boldsymbol{X}^q\}$ 可由下面两种情形产生:

情形 1 若 $n \le q-1$, 则 $\boldsymbol{X}^i = (x_1^i, \cdots, x_n^i)^T$ 的分量可定义为:

$$x_j^i = w_j + (2a_{ij} + 1/2q)(z_j - w_j), j = 1, 2, \cdots, n$$

情形 2 若 $n > q-1$, 将任一 n 维向量 $\boldsymbol{H} = (h_1, h_2, \cdots, h_n)^T$ 分成 $(q-1)$ 子向量, $\boldsymbol{H} = (\boldsymbol{H}^1, \boldsymbol{H}^2, \cdots, \boldsymbol{H}^{q-1})^T$, 其中, $\boldsymbol{H}^j = (h_{r_{j-1}+1}, h_{r_{j-1}+2}, \cdots, h_{r_j})$ 是 $(r_j - r_{j-1})$ 维子向量, $j = 1, 2, \cdots, q-1$, $r_1, r_2, \cdots, r_{q-2}$ 是在 $[1, n]$ 中随机产生的互不相同的整数, $0 = r_0 < r_1 < \cdots < r_{q-1} = n$. 类似地, 将 $\boldsymbol{X}$、$\boldsymbol{Y}$、$\boldsymbol{W}$ 和 $\boldsymbol{Z}$ 也按如上方式分成 $(q-1)$ 个子向量, 它们的第 j 个子向量分别记为 $\boldsymbol{A}^j$, $\boldsymbol{B}^j$, $\boldsymbol{W}^j$ 和 $\boldsymbol{Z}^j$, $j = 1 \sim q-1$, 则 $\boldsymbol{X}^i = (\boldsymbol{V}_1^i, \boldsymbol{V}_2^i, \cdots, \boldsymbol{V}_{q-1}^i)^T$ 可定

收稿日期: 2001-06-20; 修回日期: 2002-04-01
基金项目: 国家自然科学基金(No.60171045); 陕西省自然科学研究计划项目(No.2001SL06)

义为：$V_j^i = W^i + (2a_{ij} + 1/2q)(Z^i - W^i), j = 1,2,\cdots,q-1$.

3 新的变异算子

定义 1 假设解的精度要求是 ε，若点 $\boldsymbol{X}$ 满足：$\| \boldsymbol{X} - \boldsymbol{X}^* \|_1 \le \varepsilon$，则称 $\boldsymbol{X}$ 为一个 ε-精度的最优解．其中 ε 是一个非常小的正数，$\boldsymbol{X}^*$ 是问题的一个全局最优解．

设搜索空间为 $[\boldsymbol{L}, \boldsymbol{U}] = \{(x_1, x_2, \cdots, x_n)^T\}, l_i \le x_i \le u_i, 1 \le i \le n$，则对 $\boldsymbol{X} = (x_1, \cdots, x_n)^T$ 进行变异的方法如下：

步 1 将 $[l_i, u_i]$ 分成若干子区间 $[w_1^i, w_2^i], [w_2^i, w_3^i], \ldots, [w_{n_i-1}^i, w_{n_i}^i]$，使得 $|w_j^i - w_{j-1}^i| < \varepsilon, j = 2 \sim n_i$，其中 $w_1^i = l_i, w_{n_i}^i = u_i$.

步 2 对 $\boldsymbol{X}$ 的每个分量 x_i，随机产生一个数 $r \in [0,1]$，若 $p_m > r$，随机在 $\{w_1^i, \cdots, w_{n_i}^i\}$ 中选一个 $w_{j_i}^i$ 代替 x_i；否则，x_i 保持不变，其中 p_m 为变异概率．经过这样的变异后，$\boldsymbol{X}$ 变为 $\boldsymbol{O} = (o_1, \cdots, o_n)^T$.

4 一个新的进化算法

算法 1 (1)给定杂交概率 p_c，变异概率 p_m 及种群大小 pop．随机产生初始种群 $\boldsymbol{P}(1) = \{\boldsymbol{X}^1, \cdots, \boldsymbol{X}^{pop}\}$，令 $k = 1$.

(2)(杂交)对每一代 k，产生 pop 个随机数 $r_1, \cdots, r_{pop} \in [0,1]$．若 $r_i < p_c$，则 $\boldsymbol{P}(k)$ 中 $\boldsymbol{X}^i$ 参加杂交，$i = 1 \sim pop$．对参加杂交的点随机配对，每一对用新的杂交算子进行杂交产生临时后代．

(3)(变异)对上步中产生的每个临时后代用新的变异算子进行变异产生 $\boldsymbol{P}(k)$ 的后代．

(4)(选择)从 $\boldsymbol{P}(k)$ 和步 3 产生的其所有后代中选择最好的 pop 个点作为下一代的种群 $\boldsymbol{P}(k+1)$，令 $k = k+1$，转 2.

(5)若终止条件满足，停止．

5 全局收敛性

定义 2 称 $\boldsymbol{X}'$ 是从 $\boldsymbol{X}$ 通过杂交和变异可达的，若 $\boldsymbol{X}'$ 可由 $\boldsymbol{X}$ 通过杂交和变异产生，即：

$Prob\{\boldsymbol{MC}(\boldsymbol{X}) = \boldsymbol{X}'\} > 0$，$\boldsymbol{MC}(\boldsymbol{X})$ 表示由 $\boldsymbol{X}$ 通过杂交和变异算子产生的点，$Prob\{\}$ 表示随机事件 $\{\}$ 的概率．

定义 3 若 $Prob\{\| \boldsymbol{MC}(\boldsymbol{X}) - \boldsymbol{X}' \|_1 \le \varepsilon\} > 0$，则称 $\boldsymbol{X}'$ 是从 $\boldsymbol{X}$ 通过杂交和变异为 ε-精度可达的．

考虑极小化问题 $\min_{X \in \Omega} f(\boldsymbol{X})$，$\Omega$ 为可行域．Bäck 已经证明[2] (pp.129)，若一个进化算法满足下面两个条件，则其以概率 1 收敛到全局最优解；

(1)对可行域中任两个点 $\boldsymbol{X}'$ 和 $\boldsymbol{X}$，$\boldsymbol{X}'$ 是从 $\boldsymbol{X}$ 通过杂交和变异可达的．

(2)种群序列 $\boldsymbol{P}(1), \boldsymbol{P}(2), \cdots, \boldsymbol{P}(k), \cdots$ 是单调的，即：对 $\forall k$，有 $\min\{f(\boldsymbol{X}) | \boldsymbol{X} \in \boldsymbol{P}(k)\} \le \min\{f(\boldsymbol{X}) | \boldsymbol{X} \in \boldsymbol{P}(k+1)\}$.

下证：将第 1 个条件中的"可达"改为："ε-精度可达"，本文进化算法以概率 1 收敛到 ε-精度的最优解．

定理 1 设 $\varepsilon > 0$ 充分小，若 $f(\boldsymbol{X})$ 在搜索空间 $[\boldsymbol{L}, \boldsymbol{U}] \supseteq \Omega$ 上连续，则算法 1 以概率 1 收敛到 ε-精度的最优解，即：

$Prob\{\lim_{k \to \infty} \| \boldsymbol{P}(k) - \boldsymbol{P}^* \|_1 \le \varepsilon\} = 1$，其中 $\boldsymbol{P}^*$ 是全局最优解集，$\{\| \boldsymbol{P}(k) - \boldsymbol{P}^* \|_1 \le \varepsilon\}$ 表示

$$\{\| \boldsymbol{X} - \boldsymbol{X}^* \|_1 \le \varepsilon \mid \forall \boldsymbol{X} \in \boldsymbol{P}(k), \exists \boldsymbol{X}^* \in \boldsymbol{P}^*\}.$$

证明 先证对任两个点 $\boldsymbol{X}' = (x'_1, \cdots x'_n)^T$ 和 $\boldsymbol{X} = (x_1, \cdots x_n)^T \in [\boldsymbol{L}, \boldsymbol{U}]$，$\boldsymbol{X}'$ 是从 $\boldsymbol{X}$ 通过杂交和变异为 ε-精度可达的，即：$Prob\{\| \boldsymbol{MC}(\boldsymbol{X}) - \boldsymbol{X}' \|_1 \le \varepsilon\} > 0$，其中 $\boldsymbol{MC}(\boldsymbol{X})$ 是 $\boldsymbol{X}$ 通过杂交和变异产生的点．事实上，设 $\bar{\boldsymbol{X}}$ 是 $\boldsymbol{X}$ 通过杂交产生的任一点，即：$C(\boldsymbol{X}) = \bar{\boldsymbol{X}}$，则只要证 $\bar{\boldsymbol{X}}$ 是从 $\boldsymbol{X}$ 通过变异为 ε-精度可达的即可，即：$Prob\{\| \boldsymbol{M}(\bar{\boldsymbol{X}} - \boldsymbol{X}') \|_1 \le \varepsilon\} > 0$．注意到，对 $\boldsymbol{X}'$ 的每个分量 x'_k，记 $|w_{i_k}^k - x'_k| = \min\{|w_j^k - x'_k| : j = 1 \sim n_k\}$，则 $|w_{i_k}^k - x'_k| \le \varepsilon$，其中 w_j^k 由变异算子确定．记 $\tilde{\boldsymbol{X}} = (w_{i_1}^1, w_{i_2}^2, \cdots, w_{i_n}^n)^T$，则 $\| \tilde{\boldsymbol{X}} - \boldsymbol{X}' \|_1 \le \varepsilon$．因此，若能证 $Prob\{\boldsymbol{M}(\bar{\boldsymbol{X}}) = \tilde{\boldsymbol{X}}\} > 0$，则 $Prob\{\| \boldsymbol{M}(\bar{\boldsymbol{X}}) - \boldsymbol{X}' \|_1 \le \varepsilon\} > 0$ 成立．设 $\boldsymbol{M}(\bar{\boldsymbol{X}})$ 和 $\tilde{\boldsymbol{X}}$ 的 Hamming 距离为 h，不失一般性，设 $\boldsymbol{M}(\bar{\boldsymbol{X}})$ 和 $\tilde{\boldsymbol{X}}$ 的后 $(n-h)$ 个分量相同，于是有

$$Prob\{\boldsymbol{M}(\bar{\boldsymbol{X}}) = \tilde{\boldsymbol{X}}\}$$
$$= (p_m^h \prod_{i=1}^{h} \frac{1}{n_i})[1 - p_m]^{n-h} \prod_{j=h+1}^{n} (1 - \frac{1}{n_j})] > 0$$

即 $\boldsymbol{X}'$ 是从 $\boldsymbol{X}$ 通过杂交和变异为 ε-精度可达的．

由算法步 4 知，对每一代 k，算法总是保持 pop 个最好的点作为种群 $\boldsymbol{P}(k)$，故种群序列 $\boldsymbol{P}(1), \boldsymbol{P}(2), \cdots, \boldsymbol{P}(k), \cdots$ 是单调的．因为 $f(\boldsymbol{X})$ 在 $[\boldsymbol{L}, \boldsymbol{U}]$ 上连续，故对充分小的 $\varepsilon > 0$，当 $\boldsymbol{X}_1 \in \boldsymbol{P}^*(\varepsilon) = \{\boldsymbol{X} : \| \boldsymbol{X} - \boldsymbol{X}^* \|_1 > \varepsilon\}$，$\boldsymbol{X}_2 \notin \boldsymbol{P}^*(\varepsilon)$ 时，有 $f(\boldsymbol{X}_1) > f(\boldsymbol{X}_2)$，其中 $\boldsymbol{X}^*$ 是任一全局最优解．

算法 1 产生的种群可看成是具有两个状态的 Markov 链：状态 1：种群满足 $\| \boldsymbol{P}(k) - \boldsymbol{P}^* \|_1 \le \varepsilon$；状态 2：种群不满足 $\| \boldsymbol{P}(k) - \boldsymbol{P}^* \|_1 \le \varepsilon$．由种群序列的单调性知，从状态 1 转移到状态 2 的概率为 0，因此状态 1 为吸收态(absorbing state)．由任两点的 ε-精度可达性知，从状态 2 转移到状态 1 的概率恒大于 0，因此状态 2 为瞬时态(transient state)．由 Markov 链理论[8]即知结论成立．

表 1 NEA 和 FES 对计算结果的比较

f	MNF		Mean solution(Standard deviation)	
	NEA	FES	NEA	FES
1	454412	900030	-12568.3423 (3.110×10^{-3})	-12556.4 (32.53)
2	359341	500030	1.317×10^{-2} (5.220×10^{-3})	0.16 (0.33)
3	139931	150030	4.456×10^{-10} (1.581×10^{-10})	1.2×10^{-2} (1.8×10^{-3})
4	164165	200030	0 (0)	3.7×10^{-2} (5.0×10^{-2})
5	142182	150030	5.003×10^{-6} (1.102×10^{-6})	2.8×10^{-6} (8.1×10^{-7})
6	134814	150030	1.002×10^{-4} (8.213×10^{-5})	4.7×10^{-5} (1.5×10^{-5})

6 数值模拟

本文选了 12 个困难的标准试验函数 $f_1 \sim f_{12}$，它们依次对应于文[4]中的 $f_8 \sim f_{13}, f_5, f_1, f_7, f_2 \sim f_4$．我们用本文的算法(简记为 NEA)对它们进行了计算，并和 FES[4]，MMO，AMMO[5]

及 PSO,EO[6]对这些函数计算的结果进行了比较.在运行 NEA 时,各参数取值如下:$pop = 150$,$q = 5$,$p_c = 0.1$,$p_m = 0.02$,$\epsilon = 10^{-6}$.对每个试验函数独立运行 30 次,每次运行,若连续 50 代最好解没有改进,则算法停止.记录最优值的平均值、标准差和函数值的平均计算次数(在表 1 ~ 3 中记为 MNF),并和文[4 ~ 6]中记录的结果比较,结果见表 1 - 3,表中第 1 列数字为函数编号,NA 表示此量文献中未提供.

表 2　NEA 和 PSO、EO 对 4 个函数计算结果的比较

f	MNF		Mean solution(Standard deviation)		
	NEA	PSO&EO	NEA	PSO	EO
2	359341	250000	1.317×10^{-2} (5.220×10^{-3})	47.1354 (3.5277)	46.4689 (6.0244)
4	164165	250000	0 (0)	0.4498 (0.00320)	0.4033 (0.00190)
7	144914	250000	3.447×10^{-3} (1.223×10^{-3})	1911.598 (140095.6)	1610.359 (86188.23)
8	125386	250000	1.139×10^{-7} (1.013×10^{-7})	11.175 (1.7444)	9.8808 (0.8919)

表 3　NEA 和 MMO、AMMO 对 9 个函数计算结果的比较

f	MNF		Mean solution(Standard deviation)		
	NEA	MMO&AMMO	NEA	MMO $b = 10^{-4}$	AMMO $b = 10^{-4}$
2	359341	250000	1.317×10^{-2} (5.220×10^{-3})	9.52 (NA)	11.90 (NA)
3	139931	150000	4.456×10^{-10} (1.581×10^{-10})	7.49×10^{-4} (NA)	9.43×10^{-4} (NA)
4	164165	250000	0 (0)	6.99×10^{-7} (NA)	1.02×10^{-6} (NA)
7	144914	250000	3.447×10^{-3} (1.223×10^{-3})	63.8 (NA)	144.00 (NA)
8	125386	150000	1.139×10^{-7} (1.013×10^{-7})	9.81×10^{-7} (NA)	1.61×10^{-6} (NA)
9	137325	250000	3.523×10^{-6} (1.182×10^{-7})	9.53 (NA)	9.64 (NA)
10	139985	250000	7.219×10^{-7} (4.882×10^{-8})	3.23×10^{-3} (NA)	3.99×10^{-3} (NA)
11	139109	250000	3.215×10^{-8} (2.529×10^{-9})	11.80 (NA)	6.12×10^{-1} (NA)
12	134531	250000	8.453×10^{-7} (2.246×10^{-7})	1.88 (NA)	3.23×10^{-1} (NA)

从表 1 可看出,NEA 对 $f_1 \sim f_6$ 中每个函数都能求出非常接近全局最优解的解;而 FES 对 f_1, f_2 不能求出满意的解.NEA 对 $f_1 \sim f_4$ 求出的解远远好于 FES 求出的解.虽然 NEA 对 f_5, f_6 求出的解比 FES 求出的稍差,但 NEA 所需的函数值的平均计算次数比 FES 要少得多.从表 2、3 可看出,NEA 对比较的每个函数都能求出非常接近全局最优解的解,且 NEA 对所比较的函数求出的解远远好于 PSO,EO[6]和 MMO,AMMO[5]求出的解,更进一步,NEA 所需的函数值的平均计算次数也比这些算法少得多.

7　结论

本文结合均匀设计构造了一个新的进化算法,从理论上证明了其收敛性,模拟结果也表明算法非常有效.由于篇幅所限,文中没有包含该算法的应用实例.实际上,它适合于电子工程领域,特别是电磁学领域中的复杂优化设计问题.

参考文献:

[1] Z Michalewicz. Genetic Algorithms + Data Structures = Evolution Programs [M]. Berlin: 3rd Edition, Springer-Verlag, 1996.

[2] T Bäck. Evolutionary Algorithms in Theory and Practice [M]. New York: Oxford University Press, 1996.

[3] Y W Leung, Yuping Wang. An orthogonal genetic algorithm with quantization for global numerical optimization [J]. IEEE Trans. on Evolutionary Computation, 2001, 5(1): 41 - 53.

[4] Yao X, Liu Y. Fast evolution strategies [A]. Peter J. Angeline, et al. Proceedings of 6th International Conference on Evolutionary Programming VI [C]. Berlin: Springer-Verlag, 1997. 151 - 161.

[5] K Chellapilla. Combining mutation operators in evolutionary programming [J]. IEEE Trans. on Evolutionary Computation, 1998, 2(3): 91 - 96.

[6] P J Angeline. Evolutionary optimization versus particle swarm optimization: philosophy and performance differences [A]. V W Porto, et al. Proceedings of 7th International Conference on Evolutionary Programming VII [C]. Berlin: Springer-Verlag, 1998. 601 - 610.

[7] K T Fang, Y Wang. Number-Theoretic Methods in Statistics [M]. London: Chapman & Hall, 1994.

[8] A O Allen. Probability, Statistics and Queuing Theory with Computer Science Applications [M]. 2nd Edition, Boston, MA: Academic Press, 1990.

作者简介:

王宇平　教授,博士生导师,1961 年生于甘肃兰州,1993 年获博士学位,主要从事优化理论、方法及应用,进化算法及应用的研究,曾应邀赴香港中文大学和香港浸会大学做访问学者,已在国内外刊物发表论文 30 多篇,主持和参加多项科研项目,曾获省部级科技进步奖 3 项.

焦永昌　教授,博士生导师,1964 年生于山西,1990 年获博士学位,主要从事优化算法及应用、天线优化设计、天线工程等方面的研究,先后应邀赴日本、香港做访问学者,近年来,发表学术论文 50 余篇,其中 20 余篇被 SCI、EI 检索,主持和参加十余项科研项目,曾获国家级和省部级科技进步奖 4 项,现为陕西省第九届人大代表,中国电子学会高级会员、青年工作委员会委员.