

遗传算法数学机理分析

韩玉兵

(河海大学常州校区 ,江苏 常州 213022)

摘要 对遗传算法(或基因算法)进行了较详细的分析与论述 ,包括 SGA 一般模型 ,GA 模式定理 .并
利用齐次有限 Markov 链证明了 SGA 的非全局收敛性和 OMSGa 的全局收敛性 .

关键词 遗传算法 ,模式定理 ,全局收敛 ,Markov 链

中图分类号 :O29 文献标识码 :A 文章编号 :1000-1980(2001)03-0092-03

遗传算法(Genetic Algorithm-GA)是一种新近发展的随机优化算法 ,由美国密西根大学 Holland 教授等创立 .它是一种自适应启发式群体型 ,概率性迭代式全局收敛算法 ,是基于自然选择和基因遗传学原理的随机的搜索算法 .它将“ 适者生存 ”这一基本的达尔文进化理论引入串结构 ,并且在串之间进行有组织但又随机的信息交换 .伴随着算法的进行 ,优良的品质被逐渐保留并加以组合 ,从而不断产生出更佳的个体 .这一过程就如生物进化一样 ,好的特性被继承 ,坏的特性被淘汰 .新代个体中包含着上代的大量信息 ,又不断地在总体特性上胜过上代 ,从而使整个群体向前进化发展 .遗传算法也就是不断地接近最优解 ,它以其解决非线性问题的鲁棒性 ,全局收敛性具有独特的吸引力 ,正引起越来越大的研究及应用热潮 .

1 标准遗传算法及其运行过程

考虑如下的非线性优化问题 $\max\{f(x) | x \in D\}$ 遗传算法是把该问题中的自变量当作生物体 ,将其转化为基因构成的染色体 ,相应的函数定义为适应度(fitness) ,优化的目标是进化成具有最佳适应度的基因型 .下文称染色体为串 ,标准遗传算法的步骤如下 :

(a)基因编码 .选择确定的编码策略将参数转化为串 ,这里采用二进制编码 ,每个二进位即为一基因 ,如

$x \in [a, b]$ 则 $x = a + (b - a) \sum_{i=0}^{l-1} g_i 2^i / (2^l - 1)$,其中 l 为码长 , g_i 为第 i 个基因 .

(b)定义串的适应度函数 $F(x) = f(x)$.

(c)确定算法参数 种群大小 n ,交叉概率 p_c ,变异概率 p_m ,随机生成 n 个串 $s_i (i = 1, \dots, n)$ 构成初始群体 .

(d)计算各串的适应度 $f(s_i)$.

(e)根据各串的复制概率 $P(s_i) = f(s_i) / \sum_j f(s_j)$ 随机选择两串进行复制(reproduction) .

(f)以概率 p_c 进行交叉(crossover)操作 ,交叉位置随机生成 .

(g)对两串中的基因以变异概率 p_m 进行变异(mutation)取反操作 .

(h)转(e)重复进行 ,直至新群体个数为 n .

(i)转(d)重复进行 ,直至满足某一指标或规定的迭代次数 .

(j)找出最佳串 ,解码得问题的最优解 .

2 遗传算法的模式定理^[1]

定义 1 基因串中某些特征位相同的结构称为模式(schema) .

二进制串中模式是 $\{a_1 \dots a_i \dots a_l | a_i \in \{0, 1, * \}\}$,其中 $*$ 为任意符(0 或 1) ,显然模式是串的集合 ,这

些串与模式在所有基因位上与不是 * 的基因匹配 ,因此模式中的 * 越多 ,则可描述的串越多 ,通常一个模式 H 包括 4 个参数 :原码长度 l ,模式长度 $\delta(H)$,阶数 $\alpha(H)$,维数 $D(H)$. l 即为串的长度 , $\delta(H)$ 为最前、最后固定位间的距离 , $\alpha(H)$ 是模式中固定位的个数 , $D(H)$ 表示模式中包含串的个数.

定理 1 对于 SGA 模式 H 的数量可估计为

$$m(H,t) \geq m(H,0) \left(\frac{\bar{f}(H)}{f} \right)^{\alpha(H)} (1 - p_c \frac{\delta(H)}{l-1} - \alpha(H)p_m)$$

其中 : $m(H,t)$ ——时刻 t 时种群中包含特定模式 H 的个数 ; $\bar{f} = (\sum f_i)/n$ 为整个种群的平均适应度 ; $\bar{f}(H)$ ——时刻 t 时对应于模式 H 的串的平均适应度 . 证见文献 [1].

此即为遗传算法的 Holland 模式定理 ,其说明模式长度短 ,阶次低且适应度高于平均适应度的模式的数量在遗传过程中将以指数形式增长.

3 遗传算法的全局收敛性分析^[2,3]

传统的基于模式定理的对全局收敛性的定性分析认为 ,遗传算法是全局收敛的 ,最近基于 Markov 链的定量数学证明认为 ,带有复制、交换、变异操作的标准遗传算法不是全局收敛的 ,不适合于静态函数的优化问题 ,而最优保存算法 (OMSGA) 是概率性全局收敛的 ,下面对此进行严格的数学证明.

定义 2 设离散随机过程 $\{X_k, k \in N\}$ 的状态空间 S 为有限集 ,若 X_k 满足马氏性 :

$$P(X_{k+1} = i_{k+1} | X_0 = i_0, X_1 = i_1, \dots, X_k = i_k) = P(X_{k+1} = i_{k+1} | X_k = i_k)$$

称此随机过程为 Markov 链 ,条件概率 $p_{ij}(k, k+m) = P(X_{k+m} = j | X_k = i)$ 为 Markov 链在时刻 k 处于状态 i 条件下 ,在时刻 $k+m$ 转移到状态 j 的转移概率 ,若对于 $\forall i, j \in S, p_{ij}(k, k+m) = p_{ij}(m)$ 与 k 无关 ,则称此 Markov 链为齐次 . 其中 $p_{ij}(1) = p_{ij}$ 为一步转移概率 ,矩阵 $P = [p_{ij}]$ 称为 Markov 链的转移概率矩阵 .

定义 3 设 $\{X_k\}$ 为齐次 Markov 链 ,若 $\forall i, j \in S$,有 $\lim_{m \rightarrow \infty} p_{ij}(m) = \pi_j (\sum_{j \in S} \pi_j = 1)$,则称 $\{X_k\}$ 具有遍历性 .

定义 4 A 为一方阵 .

(a) 若所有 $i, j, a_{ij} > 0$ 称 A 严格正矩阵 ,记 $A > 0$;

(b) 若 $A \geq 0$ 且 $\exists k \in N$ 使 $A^k > 0$ 称 A 正则矩阵 ;

(c) 若 $A \geq 0$ 且 $\forall i$ 使 $\sum_{j=1}^n a_{ij} = 1$ 称 A 随机矩阵 .

定理 2 若 P 是正则的齐次 Markov 链的转移矩阵 ,则此齐次 Markov 链是遍历的 . 证见文献 [3].

定义 5 称 $H(s_i, s_j) = \sum_{k=1}^l |g_{ik} - g_{jk}|$ (其中 g_{ik} 是串 s_i 的第 k 位基因) 为串 s_i 和 s_j 的 Hamming 距离 .

SGA 用齐次有限 Markov 链描述如下 :状态空间为群体空间 Λ , Λ 的维数为 2^{ln} , Λ 的元素 λ_i 是一群体 ,包含 n 个串长为 l 的串 $s_j (j = 1, \dots, n)$,群体空间的概率变化由复制、交换和变异三种基因操作引起 ,它们分别用矩阵 R, C, M 描述 ,显然 SGA 的 Markov 转移阵为 $P = RCM$.

引理 1 比例复制操作的概率矩阵 R 是随机的 .

证明 比例复制操作的作用是按与适应度相关的概率映射到自身或其他状态 ,对所有的 $i \in [1, 2^{ln}]$,

$$\sum_{j=1}^{2^{ln}} r_{ij} = 1, \text{ 所以 } R \text{ 为随机阵.}$$

引理 2 交换概率为 p_c 的交换操作概率矩阵 C 是随机的 .

证明 交换操作的作用是按一定的概率把 λ_i 映射到 λ_j ,对所有的 $i \in [1, 2^{ln}]$, $\sum_{j=1}^{2^{ln}} c_{ij} = 1$,所以 C 是随机矩阵 .

引理 3 变异概率为 p_m 的变异操作概率矩阵 M 是严格正的随机矩阵 .

证明 群体状态 λ_i 和 λ_j 之间的距离为 $H(\lambda_i, \lambda_j) = \sum_{a=1}^n \sum_{b=1}^l |g_{iab} - g_{jab}|$,所以状态 λ_i 变异为 λ_j 概率为 $m_{ij} = p_m^{H(\lambda_i, \lambda_j)} (1 - p_m)^{nl - H(\lambda_i, \lambda_j)} > 0$,又对所有的 $i \in [1, 2^{ln}]$, $\sum_{j=1}^{2^{ln}} m_{ij} = 1$ 所以 M 是严格正的随机阵 .

定理 3 SGA 的转移矩阵 P 是正则随机的, 构成的 Markov 链是遍历的.

证明 因 R, C, M 是随机矩阵, P 也是随机矩阵, 令 $A = RC$, 因为 $M > 0$, 所以对于任意的 $i, j \in [1, 2^{ln}]$, $p_{ij} = \sum_{k=1}^{2^{ln}} a_{ik} m_{kj} > 0$, 即 $P > 0$, 所以 SGA 构成的 Markov 链是遍历的.

此结论说明不论群体的初始分布如何, 从任意的状态 λ_i 出发可在有限时间内到达任意的状态 λ_j , 即能遍历整个状态空间, 马氏链中的任何状态均有一大于零的唯一极限分布, 但这并不意味着收敛于全局最优解, 为此定义如下全局收敛概念.

定义 6 令 $F_k = \max\{f(s_{ik}) | j \in [1, n]\}$ 是时刻 k 状态 λ_i 时群体中的最大适应度, 令 $F^* = \max\{f(s_j) | j \in [1, 2^l]\}$ 是所求问题的全局最优适应度, 当且仅当 $\lim_{k \rightarrow \infty} P(F_k = F^*) = 1$ 成立时, GA 是全局收敛的.

定理 4 SGA 不是全局收敛的.

证明 设 λ_i 是满足 $F_k < F^*$ 的一任意状态, 并设 $P(k)$ 是 SGA 在时刻 k 处于状态 λ_i 的概率, 显然 $P(F_k < F^*) \geq P(k)$, 则 $P(F_k = F^*) \leq 1 - P(k)$, 由 SGA 的遍历性 (见定理 3) 知, 任意状态 λ_i 均有一大于零的唯一极限分布, 即 $\lim_{k \rightarrow \infty} P(k) > 0$, 因此 $\lim_{k \rightarrow \infty} P(F_k = F^*) \leq 1 - \lim_{k \rightarrow \infty} P(k) < 1$, 所以 SGA 不是全局收敛的.

本定理说明 SGA 虽然发现最优解, 但却不是全局收敛的, 它有全空间搜索能力但不能实现全局收敛的原因是发现的最优解不能保持, 下面的 OMSGA 则避免了这方面的缺点.

定义 7 设 λ_k 是 SGA 马氏链在时刻 k 具有最优适合度 F_k 的状态, 若 $F_{k+1} < F_k$, 则 $\lambda_{k+1} = \lambda_k$, $F_{k+1} = F_k$, 以上加入最优保存操作的 SGA 称为 OMSGA (optimum maintaining simple genetic algorithm).

定理 5 OMSGA 是全局收敛的.

证明 设 Λ^* 为包含全局最优解的状态集合, 由 OMSGA 的定义, Λ^* 是自封闭的吸引子, 即一旦状态转移到此空间则最优个体就能保存下来. 把 GA 的整个状态空间划分为 $\Lambda_1, \Lambda_2, \dots, \Lambda_s = \Lambda^*$, 其中 $\Lambda_i (i = 1, \dots, s)$ 为具有相同最优适应度的状态集合, 且 Λ_{i+1} 中的最优适应度比 Λ_i 中的大. 则 OMSGA 的转移阵为, $P =$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ 0 & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_{ss} \end{bmatrix} \quad \text{其中 } a_{ij} \text{ 为由状态空间 } \Lambda_i \text{ 转移到 } \Lambda_j \text{ 的概率, 显然 } P \text{ 满足 (a) } P \text{ 为随机阵, 即 } P \geq 0, \forall i,$$

$$\sum_{j=i}^s a_{ij} = 1 \quad \text{(b) } 1 \leq i < s \text{ 时, } a_{ii} < 1. \text{ 又因为 } \lim_{k \rightarrow \infty} P^k = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}, \text{ 即对于 OMSGA 不论何种初始状态, 经}$$

过一定时间后必会转移到 Λ^* 空间, 又因 Λ^* 是自吸引的, 所以 OMSGA 为全局收敛.

参考文献:

- [1] Holland J H. Adaptation in nature and artificial system[M]. Ann Arbor: The University of Michigan Press, 1975.
- [2] Rudolph G. Convergence properties of canonical genetic algorithms[J]. IEEE Trans. on Neural Networks, 1994, 5(1): 96 ~ 101.
- [3] Emanuel Parzen. 随机过程[M]. 北京: 高等教育出版社, 1987. 200 ~ 289.

Analysis on Mathematical Mechanism of Genetic Algorithm

HAN Yu-bing

(Hohai Univ. Changzhou Branch, Changzhou 213022, China)

Abstract: In this paper, the genetic algorithm is discussed in detail, including the general model of GA and GA schema theorem. Then the non-global convergence of simple genetic algorithm (SGA) and the global convergence of (OMSGA) are proved by use of homogeneous finite Markov chain.

Key words: genetic algorithm; schema theorem; global convergence; Markov chain