

遗传算法的平均收敛速度及其估计

喻寿益¹, 郭观七^{1,2}

(1. 中南大学 信息科学与工程学院, 湖南 长沙 410083; 2. 湖南理工学院 机电工程系, 湖南 岳阳 414000)

摘要: 给出了独立于表示的变异算子和交叉算子的数学描述, 建立了遗传算法种群的精确马尔可夫链模型, 导出了种群中最佳个体的马尔可夫链及其随机矩阵. 将遗传算法的平均收敛速度定义为最佳个体转移至吸收态的平均吸收时间的数学期望, 提出了应用最佳个体的随机矩阵估计遗传算法平均收敛速度的理论方法和计算步骤.

关键词: 遗传算法; 平均收敛速度; 马尔可夫链

中图分类号: TP301

文献标识码: A

Estimating average convergence velocity of genetic algorithms

YU Shou-yi¹, GUO Guan-qi^{1,2}

(1. College of Information Science and Engineering, Central South University, Hunan Changsha 410083, China;

2. Department of Electromechanical Engineering, Hunan Institute of Science and Technology, Hunan Yueyang 414000, China)

Abstract: The description of mutation and crossover operators independent of representation of solutions is formulized. The precisely quantitative Markov chain model of populations of a genetic algorithm is presented. Based on this model, the stochastic matrix of Markov chain of the best-so-far individual with the highest fitness in populations is derived. The average convergence velocity of a genetic algorithm is defined as the mathematical expectation of the mean absorbing time that the best-so-far individual transfers to the absorbing state. The theoretic method and computing process of estimating the average convergence velocity of a genetic algorithm are proposed.

Key words: genetic algorithms; average convergence velocity; Markov chain

1 引言 (Introduction)

迄今为止, 遗传算法 (GAs: genetic algorithms) 的收敛速度问题无严格的理论结论, 缺乏有效的分析方法. 虽然模式定理^[1]一直被认为是最经典的理论成果, 但该定理仅证明了种群中定义长度短的低阶模式按指数规律增长的趋势, 不能用来分析 GAs 的一般动力学特性. Tomas Bäck^[2]基于简化的 GAs (采用最小种群规模, 忽略交叉算子) 和最简单的优化问题 (二进制 1 的计数问题), 应用顺序统计学对收敛速度进行了分析, 其分析方法很难甚至不可能推广到标准 GAs 和一般的优化问题.

马尔可夫链理论对 GAs 的动力学特性分析提供了有效的理论工具, 但现有的 GAs 马尔可夫链模型^[3-5]都是简化的定性分析模型, 只能用来分析 GAs 的极限特性. 本文首先描述基于精英 (elitist) 保留机制的 GAs, 给出了独立于表示方法的变异算子和交叉算子的数学定义, 建立了 GAs 的精确马尔可

夫链模型, 提出了构造种群中最佳个体随机矩阵的方法和应用该随机矩阵对 GAs 的平均收敛速度进行估计的理论方法.

2 算法描述 (Descriptions of the algorithm)

设最优化问题有 k 个候选解, k 为大于 0 的整数. 为了便于引用, 将 k 个候选解按照它们的适应值的降序分别映射为整数 $1, \dots, k$, 并直接用相应的整数来引用它们, 那么, 全体解空间的 k 个适应值满足 $f_1 \geq f_2 \geq \dots \geq f_{k-1} \geq f_k$. 设种群规模为 $N+1$, 则种群 $i = (i_0, i_1, \dots, i_N)$, 其中 i_0 称为超级个体, 该个体不参与标准遗传算法的选择、交叉和变异操作, 遗传算子仅作用于 $i_1, \dots, i_N$. 每一代遗传操作结束后, 仅当 $i_1, \dots, i_N$ 中的最佳个体的适应值大于超级个体的适应值时, 才用最佳个体更新超级个体. 这种超级个体更新策略实现精英 (elitist) 保留机制, 以保证 GAs 的全局收敛性. 进一步设交叉的配对策略为均匀随机配对策略, 一对父代个体仅生成一个子代个体. 为

了引用的方便,将引入超级个体更新机制的遗传算法简称为 EGAs. EGAs 首先随机地初始化 $i_1, \dots, i_N$, 并将其中的最佳个体设置为初始的超级个体, 然后, 循环地进行选择、变异、交叉和超级个体更新操作, 直到超级个体收敛到全局最优解或满足停止规则.

3 变异算子和交叉算子的定义 (Formulations of mutation and crossover)

为了建立 EGAs 的马尔可夫链模型, 首先给出独立于表示方案的变异和交叉算子的形式化描述.

定义 1 设 $\delta = (\delta_{xy})$ 表示一个行随机的变异概率矩阵, 元素 δ_{xy} 表示由解 x 转移到 y 的变异概率,

并满足对全体 $x, y = 1, \dots, k, \delta_{xy} > 0, \sum_{y=1}^k \delta_{xy} = 1$.

例如, 当解用长度为 l 位的二进制编码表示并采用均匀变异策略时

$$\delta_{xy} = p_m^{H(x,y)}(1-p_m)^{l-H(x,y)}, \delta_{xy} = \delta_{yx} = 0. \quad (1)$$

式(1)中, p_m 表示变异概率参数, $H(x, y)$ 表示 x, y 的编码之间的 Hamming 距离.

定义 2 设 $\eta = (\eta_{xy,z})$ 表示一个行随机的交叉概率矩阵, 此处, $x \leq y$, 其中的元素 $\eta_{xy,z}$ 表示由解 x 和 y 经交叉操作生成解 z 的概率, 对全体 $1 \leq x \leq y \leq k$, 满足 $\sum_{z=1}^k \eta_{xy,z} = 1$, 并且规定, 当 $x > y$ 时, $\eta_{xy,z} = \eta_{yx,z}$, 当 $x = y$ 时, $\eta_{xy,x} = 1$.

强制 $x \leq y$ 是为了保证对父代个体引用的唯一性. 从理论的观点看, 构造不同的变异概率和交叉概率矩阵, 可使 GAs 获得不同的动力学特性.

4 种群的马尔可夫链模型 (Markov chain model of the populations)

鉴于超级个体不参与选择、变异和交叉操作, 在建立种群的马尔可夫链时可不考虑超级个体的存在. 因此, 将种群重新定义为 $i = (i_1, \dots, i_N)$, 全体种群的状态空间 $I = \{i = (i_1, \dots, i_N) | 1 \leq i_m \leq k, 1 \leq m \leq N\}$, 显然, 状态空间的基数 $|I| = k^N$. 设 $X_{i,e}$ 表示在种群 i 中解 e 的个数, 当采用按适应值比例复制的选择策略时, 解 e 被选中进入下一代的概率为

$$r_e = \frac{X_{i,e} f_e}{\sum_{d=1}^k f_d}, 1 \leq e \leq k. \quad (2)$$

于是, 在选择算子的作用下, 由状态 $i \in I$ 转移到任意状态 $j = (j_1, \dots, j_N) \in I$ 的转移概率为

$$s_{ij} = P(j = (j_1, \dots, j_N) | i = (i_1, \dots, i_N)) = \prod_{e=1}^k r_e^{Y_{i,e}}. \quad (3)$$

式(3)中, $Y_{i,e}$ 表示在种群 j 中解 e 的个数. 对于不同的选择策略, 选择概率 r_e 的计算方法可能不同于式(2), 但这并不影响式(3)的一般性.

根据定义 1, 可立即得到与变异算子相对应的状态转移概率为

$$m_{ij} = P(j = (j_1, \dots, j_N) | i = (i_1, \dots, i_N)) = \prod_{d=1}^N \delta_{i_d, j_d}. \quad (4)$$

当交叉算子采用均匀分布的随机配对策略时, 种群中任意的个体被选中的概率为 $1/N$, 因此, 根据定义 2 可得, 由种群中任意的一对父代个体 i_a, i_b 经交叉操作生成子代个体 j_c 的概率等于

$$\frac{\eta_{i_a i_b, j_c}}{N^2}, 1 \leq a, b, c \leq N.$$

于是, 与交叉算子相对应的状态转移概率为

$$c_{ij} = P(j = (j_1, \dots, j_N) | i = (i_1, \dots, i_N)) = \prod_{c=1}^N \left(\sum_{a=1}^N \sum_{b=1}^N \frac{\eta_{i_a i_b, j_c}}{N^2} \right) = \frac{1}{2^{2N}} \prod_{c=1}^N \sum_{a=1}^N \sum_{b=1}^N \eta_{i_a i_b, j_c}. \quad (5)$$

对全体 $i, j \in I$, 令矩阵 $S = (s_{ij}), M = (m_{ij}), C = (c_{ij})$, 那么, 种群的进化过程可用如下马尔可夫链动力学方程

$$i(t+1) = C \cdot M \cdot S \cdot i(t) = P \cdot i(t) \quad (6)$$

描述, 式(6)中, $P = C \cdot M \cdot S$. 那么, 按照 Rudolph^[3]的分析方法, P 必为正的随机矩阵, 该马尔可夫链存在唯一的、非零的极限分布, 因此, 链中不存在任何吸收态, 不能收敛到全局最优解.

5 平均收敛速度的估计 (Estimation of the average convergence velocity)

在种群中引入超级个体后, 更新机制保证了它进化的定向性, 其进化过程也可用马尔可夫链描述. 由于每一代种群中的超级个体 i_0 实为 $i_1, \dots, i_N$ 中的最佳个体, 因此, 可用种群中最佳个体的定向转移概率来导出超级个体的转移概率矩阵 U , 并应用 U 阵对收敛速度进行估计.

为了计算 U 阵, 从种群状态空间 I 中提取 k 个分组, 第 n 个分组中的各个种群均包含最佳个体 n , $1 \leq n \leq k$. 显然, 超级个体的状态转移概率可通过分组间(内)的状态转移概率来计算. 设第 n 个分组

的状态集用 G_n 表示,即

$$G_n = \{i | (i \in I) \wedge (e^* = \arg \max \{f_e | e \in i\} = n)\}. \quad (7)$$

对全体 $e^* \in G_n$ 和 $1 < n < k-1$, 满足 $e^* \in G_{n-1} \leq e^* \in G_n \leq e^* \in G_{n+1}$. 容易证明

$$|G_n| = \begin{cases} \sum_{d=1}^N \binom{N}{d} (k-n)^{N-d}, & 1 \leq n \leq k-1, \\ 1, & n = k. \end{cases} \quad (8)$$

因此,对全体 $x \in \{1, \dots, k\}$, 可得

$$u_{xy} = \begin{cases} 0, & x < y, \\ \sum_{i \in G_x} \sum_{j \in G_y} p_{ij}, & x \geq y. \end{cases} \quad (9)$$

其中 p_{ij} 按式(6)计算. 显然, 矩阵 U 具有如下下三角矩阵和分块矩阵形式

$$U = (u_{xy}) = \begin{pmatrix} 1 & & & 0 \\ u_{21} & u_{22} & & \\ \vdots & \cdots & \ddots & \\ u_{k1} & u_{k2} & \cdots & u_{kk} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ R & Q \end{pmatrix},$$

$$Q = \begin{pmatrix} u_{22} & & 0 \\ \vdots & \ddots & \\ u_{k2} & \cdots & u_{kk} \end{pmatrix}. \quad (10)$$

U 阵的结构形式说明, 超级个体 $i_0 = 1$ 为其马尔可夫链的唯一吸收态, 其它状态则构成了非常返状态集 $T = \{2, \dots, k\}$. 令 μ_x 表示超级个体从任意状态 $x \in T$ 转移至吸收态的平均吸收时间步数, 则根据马尔可夫链的平均吸收时间定理^[6] 可得, 对于全体 $x \in T$, μ_x 满足方程组

$$\mu_x = 1 + \sum_{y=2}^k u_{xy} \mu_y. \quad (11)$$

由于超级个体的马尔可夫链只有有限个非常返状态, 此方程组有唯一解.

定义 3 将超级个体从任意的初始状态转移至吸收态的平均吸收时间步的数学期望称为 EGAs 的平均收敛速度 $E(t)$.

定理 1 当初种群中的超级个体服从均匀分布时, 运行控制参数不随时间变化的 EGAs 的平均收敛速度为

$$E(t) = \frac{1}{k} \sum_{x=2}^k \mu_x. \quad (12)$$

证 因为初始种群中的超级个体服从均匀分布, 超级个体的初始值为任意解的概率 $P^0(i_0 = x | x = 1, \dots, k) = 1/k$, 又因为 μ_1 表示从吸收态转移

至吸收态的平均吸收时间, 故 $\mu_1 = 0$, 所以

$$E(t) = \sum_{x=1}^k P^0(i_0 = x) \mu_x = \frac{1}{k} (\mu_1 + \sum_{x=2}^k \mu_x) = \frac{1}{k} \sum_{x=2}^k \mu_x.$$

证毕.

对于特定的优化问题和 EGAs 实现, 对平均收敛速度的估计可按下面的步骤进行:

- 按式(2)~(6)分别求出随机矩阵 S, M, C 和 P ;
- 按式(7)将种群状态空间进行分组;
- 根据式(9)构造 U 阵;
- 解方程组(11);
- 按式(12)计算 EGAs 的平均收敛速度.

对于二进制编码、均匀变异和一点交叉的 EGAs, 应用本文所介绍的方法和步骤对具体优化问题的平均收敛速度进行估计和分析的实例将在另文中给出, 在此不在赘述.

6 结论(Conclusions)

本文给出了变异算子和交叉算子的数学描述, 建立了 EGAs 定量的、精确的马尔可夫链模型, 在此基础上导出了种群中最佳个体状态转移的随机矩阵, 提出了根据最佳个体的随机矩阵估计 EGAs 平均收敛速度的理论方法和计算步骤. 尽管这种方法在实现时显得冗长烦琐, 但是应用该方法估计 EGAs 的平均收敛速度具有理论上的严密性, 借助于计算机仿真软件的支持, 该方法具有可推广性.

参考文献(References):

- [1] HOLLAND J H. *Adaptation in Nature and Artificial Systems* [M]. Cambridge, Massachusetts: The MIT Press, 1992.
- [2] BÄCK T. Order statistics for convergence velocity analysis in simplified evolutionary algorithms [A]. WHITLEY D, VOSE M. *Foundations of Genetic Algorithms-3* [M]. San Mateo, CA: Morgan Kaufmann, 1995: 91-102.
- [3] RUDOLPH G. Convergence analysis of canonical genetic algorithms [J]. *IEEE Trans on Neural Network*, 1994, 5(1): 96-101.
- [4] SUZUKI J. Markov chain analysis on simple genetic algorithms [J]. *IEEE Trans on Systems, Man, and Cybernetics*, 1995, 25(4): 655-659.
- [5] REYNOLDS D, GOMATAM J. Stochastic modelling of genetic algorithms [J]. *Artificial Intelligence*, 1996, 82(2): 303-330.
- [6] 申鼎煊. 随机过程[M]. 武汉: 华中理工大学出版社, 1990: 54-80. (SHEN Dingxuan. *Stochastic Processes* [M]. Wuhan: Press of Huazhong University of Science and Technology, 1990: 54-80.)

作者简介:

喻寿益 (1940—), 男, 教授, 博士生导师. 主要研究方向: 自适应控制, 进化算法, 人工智能. E-mail: s_yushouyi@sina.com;

郭观七 (1963—), 男, 副教授, 博士研究生. 主要研究方向: 遗传算法, 人工智能, 最优化方法. E-mail: guanqi-guo@hotmail.com.