第 29 卷 第 12 期
Vol. 29 No. 12

控 制 与 决 策
Control and Decision

2014 年 12 月
Dec. 2014

文章编号: 1001-0920 (2014) 12-2201-06 **DOI:** 10.13195/j.kzyjc.2013.1228

思维进化算法的转移概率分析及几乎处处收敛性证明

郭 红 戈

(太原科技大学 电子信息工程学院，太原 030024)

摘 要: 思维进化算法已有的收敛性分析均是在依概率收敛意义下考虑的, 而几乎处处收敛强于依概率收敛. 在详细分析思维进化算法趋同算子和异化算子转移概率的基础上, 利用种群最大适应度值函数描述思维进化算法的演化过程, 将最大适应度值函数的进化过程转化为下鞅数列, 并根据数学期望的性质和最大适应度值函数的特点, 利用下鞅收敛定理严格证明了思维进化算法的几乎处处收敛性.

关键词: 思维进化算法；几乎处处收敛；转移概率；下鞅收敛定理

中图分类号: TP301 **文献标志码:** A

Migration probabilities analysis and almost sure convergence proof of mind evolutionary algorithm

GUO Hong-ge

(College of Electronic and Information Engineering，Taiyuan University of Science and Technology，Taiyuan 030024，China．E-mail：ghg1977@163.com)

Abstract: The convergence in probability of mind evolutionary algorithm (MEA) is proved by using the Markovian chain. The almost sure convergence is stronger than the convergence in probability. The migration probabilities of the similar-taxis and the dissimilation are analyzed in detail, and the evolution process of the maximal fitness function is described as a sub-martingale series. Based on the characteristic of the conditional expectation and the maximal fitness function, the almost sure convergence of MEA is proved rigorously by using the sub-martingale convergence theorem.

Key words: mind evolutionary algorithm；almost sure convergence；migration probabilities；sub-martingale convergence theorem

0 引 言

思维进化算法 (MEA)[1]是进化计算的一种, 用群体寻优代替个体寻优, 用趋同和异化两个算子来模仿人类思维进化的过程. 趋同发生在子群体范围内, 是个体竞争成为胜者的过程. 异化发生在整个解空间内, 是各子群体为成为胜者而竞争, 不断探索解空间中新的点的过程. 趋同和异化在算法运行过程中反复进行直到满足算法终止运行条件. MEA 的优点是具有正负反馈机制, 由于有利条件的正反馈使进化总是向着有利于群体生存的方向发展, 进化成果能得以巩固和发展. 负反馈机制有助于避免陷入局部最优解, 防止算法早熟[2]. MEA 最显著的特点是, 各子群体间采用并行寻优. 工程实践[3-6]和理论应用[2, 7]已经证明 MEA 具有很高的搜索效率和收敛特性.

很多学者[8-10]已证明 MEA 的全局收敛性, 但他们均采用传统的 Markov Chain 理论分析 MEA 的收敛性, 证明过程复杂, 且收敛性结果属于依概率收敛, 弱大数律范畴, 而几乎处处收敛性明显强于依概率收敛性. 对此, 下面将在分析 MEA 转移概率的基础上证明 MEA 的几乎处处收敛性.

1 MEA 优化模型

考虑一个优化问题

$$\max A(x),\ A(x) > 0. \tag{1}$$

利用 MEA 求解问题, MEA 模型的具体描述如下.

Step 1: MEA 初始化.

Step 2: 在解空间中随机散布 N 个个体作为初始群体数列.

Step 3: 在初始群体数列中选出 N_S 个适应度值最高的个体作为优胜子群体的中心, 其余 $N_T(N =$

收稿日期: 2013-09-05；**修回日期:** 2014-02-07.

基金项目: 山西省自然科学 (青年科技研究) 基金项目 (2012021012-4).

作者简介: 郭红戈 (1977−), 女, 讲师, 博士, 从事预测控制、智能控制等研究.

N_S+N_T) 个个体作为临时子群体的中心.

Step 4: 趋同. 分别以各子群体中心作为均值, 服从正态分布产生新的 S_G-1 (S_G 是子群体中包含的个体数, 在此令所有子群体规模相同) 个个体, 形成 N 个子群体; 计算每个子群体中所有个体的适应度值, 选出优胜者作为子群体新的中心.

Step 5: 判断子群体是否成熟. 若是, 则转 Step 6, 否则转 Step 4.

Step 6: 异化. 在所有子群体中选出 N_S 个适应度值高的子群体, 形成优胜子群体; 其余 N_T 个临时子群体释放掉, 同时在解空间中重新均匀散布 N_T 个新的个体, 作为新的临时子群体中心.

Step 7: 在优胜子群体和临时子群体中选出适应度值最大的子群体.

Step 8: 判断是否满足终止条件. 若是, 则停止进化, 否则重复 Step 4 ~ Step 7.

依照 MEA 的术语, 环境是所求优化问题的解空间, 用 F 表示; 子群体用 $G_1, G_2, \cdots, G_N$ 表示; 描述优化问题的任意解称为个体, 用 g 表示; 子群体的元素用 g_{ij} 表示, g_{ij} 的含义是 F 中子群体 G_i 的第 j 个个体, $i=1,2,\cdots,N$, $j=1,2,\cdots,S_G$; 子群体 G_i 的中心用 $\hat{g}_i$ 表示.

考虑式 (1) 的优化问题, 在 F 上定义的任意非负实值函数 f 称为一个适应度函数, 即

$$\max f(g),\ f(g)>0. \tag{2}$$

令全局最大适应度值

$$f^*=\max\{f(g):g\in F\};$$

全局最优个体集

$$B^*=\{g\in F:f(g)=f^*\};$$

子群体 G_i 的最大适应度值

$$\hat{f}_i=\max\{(f(g):g\in G_i),\ 1\leqslant i\leqslant N\};$$

子群体 G_i 的最优个体集

$$\hat{B}_i=\{g\in G_i:f(g)=\hat{f}_i,\ 1\leqslant i\leqslant N\};$$

群体最大适应度值

$$\hat{f}=\max\Big\{\hat{f}(g):g\in\bigcup_{i=1}^{N}G_i\Big\};$$

群体最优个体集

$$\hat{B}=\Big\{g\in\bigcup_{i=1}^{N}G_i\in F:f(g)=\hat{f},\ 1\leqslant i\leqslant N\Big\}.$$

用 t 表示 MEA 的当前进化代数, 子群体趋同成熟后, G_i 的中心用 $\hat{g}_i^t$ 表示, $\hat{g}_i^t$ 的适应度值称作子群体 G_i^t 的适应度值, 胜者 $\hat{g}_i^t$ 代表子群体 G_i 参与异化操作. 异化时, 计算每个子群体的适应度值, 得到当前群体的最大适应度值 $\hat{f}^t$ 和对应群体最优个体集 $\hat{B}^t$. 从 $\hat{B}^t$ 中随机选出一个个体, 用 $\hat{g}^t$ 表示, 那么数列 $\{\hat{g}^t:f(\hat{g}^t)=\hat{f}^t, t=1,2,\cdots\}$ 是由 MEA 进化的每一代的最优个体所组成的数列, 由此易知 $\hat{g}^t$, $\hat{g}^{t+1}\in B^*$ 的充分必要条件为 $f(\hat{g}^t)=f(\hat{g}^{t+1})=f^*$.

群体局部最优个体集为

$$H=\Big\{g\in\bigcup_{i=1}^{N}G_i:f(g^t)=\hat{f}(g^{t+1})<f^*\Big\}.$$

当 H 存在时, 说明群体陷入局部最优. 由此易知 $\hat{g}^t$, $\hat{g}^{t+1}\in H$ 的充分必要条件为 $f(\hat{g}^t)=f(\hat{g}^{t+1})<f^*$.

2 MEA 操作过程的转移概率分析

MEA 模拟人类的思维进化过程与机制, 用趋同和异化两个算子来定义. 从 t 代的 $\hat{g}^t$ 到 $t+1$ 代的 $\hat{g}^{t+1}$ 的转移通过以下两个过程实现:

$$\hat{g}^t\xrightarrow{\text{趋同}S_t}Y'\xrightarrow{\text{异化}D_t}\hat{g}^{t+1}.$$

其中: $Y^t=\{y_i^t:1\leqslant i\leqslant N\}$, 是趋同结束后所有子群体中心所组成的数列.

下面将在概率空间上对这两个过程作详细的说明. 设所有随机映射和随机过程定义在概率空间 (Ω, F, P) 上, 为简化记号, 略去 Ω 的元素 ω.

2.1 趋同算子

趋同算子 S_t 是子群体经若干代的进化而成熟的过程, 是寻找子群体中适应度值最大的个体的过程, 是由 F 到 F 的映射. 以子群体 G_i 为例, 令 k 为 MEA 子群体趋同过程中的当前进化代数, 则 $\hat{g}_i^k$ 为趋同的第 k 代时子群体 G_i^k 的中心.

由 $\hat{g}_i^k$ 到 $\hat{g}_i^{k+1}$ 的转移, 为了符号表示方便, 令 $\hat{g}_i^k$ 作为子群体 G_i^k 的第一个个体 ($\hat{g}_i^k$ 在子群体中的位置对算法并没有影响), 子群体 G_i^k 中的个体为 $\{\hat{g}_i^k, g_{i2}^k,\cdots,g_{iS_G}^k\}$; 然后, 从 S_G 个个体中选出一个适应度值最大的个体作为子群体 G_i^k 新的中心 $\hat{g}_i^{k+1}$; 重复上述过程, 直到子群体成熟. 趋同将每个子群体的最优个体作为子群体中心, 保留到下一代, 不参与竞争.

子群体的中心个体选择概率为

$$a_k=P_s^{k^*}(\hat{g}_i)=|\hat{g}_i|/|\hat{B}_i|,\ \hat{g}_i\in\hat{B}_i\in G_i. \tag{3}$$

其中: $|\hat{B}_i|$ 为 $\hat{B}_i$ 所包含的最优个体数, $|\hat{g}_i|$ 为 $\hat{B}_i$ 中包含 $\hat{g}_i$ 的个数. 子群体中心个体的最大选择概率为

$$a_m=\max\{P_s^{k^*}\}=1, \tag{4}$$

子群体中心个体的最小选择概率为

$$a_n=\min\{P_s^{k^*}\}=1/S_G. \tag{5}$$

趋同时, 子群体中除中心个体之外的其他个体常用的分布方式是多维正态分布 $g_{ij}\sim N(u=\hat{g}_i,\Sigma)$, 选择概率为

$$P_s^k(g_{ij})=P\{S_t(g_{ij}):g_{ij}\in G_i,\ 2<j\leqslant S_G\}. \tag{6}$$

由于正态分布是连续分布, 在某一点的概率为零, 在此近似用 $g_{ij}\ (2 < j \leqslant S_G)$ 为中心的任意小邻域上的正态分布值来表征个体散布到 $g_{ij}\ (2 < j \leqslant S_G)$ 的概率. 因此, 在趋同中中心个体之外的任意个体产生的概率为

$$b_k = P_s^k(g_{ij}) = \int\int_C \cdots \int f_1(g_{ij}^1, g_{ij}^2, \cdots, g_{ij}^l)\mathrm{d}C. \quad (7)$$

其中: C 为 l 维解空间 F 上的以 g_{ij} 为中心, ε 为半径的开球, 记为 $C(g_{ij}, \varepsilon)$, 这里 ε 为任意小的正实数; $f_1(\cdot)$ 为 l 维正态分布的密度函数.

设子群体内, 与子群体中心 $\hat{g}_i$ 的欧氏距离最大的个体是 g'_{ij}. 因为解空间是有限的, 所以 g'_{ij} 一定存在, 当 $\hat{g}_i$ 确定时, g'_{ij} 即被确定. 子群体中按正态分布产生个体的最小概率为

$$b_n = P_s^k(g'_{ij}),\ g'_{ij} \in G_i. \quad (8)$$

综合式 (3) 和 (7), 可得中心个体之外的个体参与下一代进化的概率为

$$c_k = a_k \cdot b_k. \quad (9)$$

$\hat{g}_i$ 以概率 1 保存到下一代, 因此个体参与下一代进化的最大概率为

$$c_m = a_m \cdot 1 = 1, \quad (10)$$

个体参与下一代进化的最小概率为

$$c_n = \min\{c_k\} = a_n \cdot b_n = \frac{1}{S_G} b_n. \quad (11)$$

MEA 进化的第 t 代, 子群体 G_i 趋同进化成熟后, 中心从 $\hat{g}_i$ 转移到 y_i 的概率为

$$d_t = P_S^t(\hat{g}_i, y_i) = \prod_{k=1}^{K_i^t} c_k, \quad (12)$$

其中 K_i^t 表示在 MEA 进化的第 t 代子群体 G_i 趋同成熟所需的进化代数.

根据式 (12), 当中心 $\hat{g}_i$ 在趋同的每一代都以概率 1 参与下一代趋同进化时, $\hat{g}_i$ 转移到 y_i 的概率最大, 即

$$d_m = 1, \quad (13)$$

子群体中心从 $\hat{g}_i$ 转移到 y_i 的最小概率为

$$d_n = \min\{d_t\} = (b_n/S_G)^{\max\{K_i^t\}}. \quad (14)$$

2.2　异化算子

异化算子 D_t 包括以下两个过程.

1) 以概率 $e_t = \{P_{D'}^t(y_i, \hat{g}_i^{t+1}) : \hat{g}_i^{t+1} \in G_S\}$ 选取优胜子群体, 各优胜子群体的中心可以重复, G_S 是优胜子群体组成的集合, 包含的子群体数目是 N_S.

定理 1　子群体 G_i 被选取为优胜子群体的概率 e_t 的最大值为 $e_m = 1$, 最小值为 $e_n = \dfrac{1}{N_T + 1}$.

证明　为了叙述方便, 先将子群体按适应度值从高到低排序.

设异化时, 优胜子群体适应度值的数列为 $\{f(\hat{g}_i) : \hat{g}_i \in F, 1 \leqslant i \leqslant N_s\}$, 临时子群体适应度值数列为 $\{f(\hat{g}_j) : \hat{g}_j \in F, N_S + 1 \leqslant j \leqslant N\}$, 且总有 $\{f(\hat{g}_i) \geqslant f(\hat{g}_j)\}$, $\hat{g}_i$ 的数目为 $|\hat{g}_i|$.

当存在 $f(\hat{g}_i)$ 对所有 $f(\hat{g}_j)$ 满足 $f(\hat{g}_i) > f(\hat{g}_j)$ 时, $f(\hat{g}_i)$ 对应的子群体 G_i 以概率 1 强制性被选取为优胜子群体中心, 因此 e_t 的最大值 $e_m = 1$ 成立.

当存在 $f(\hat{g}_i)$ 和 $f(\hat{g}_j)$ 满足 $f(\hat{g}_i) = f(\hat{g}_j)$ 时, 即优胜子群体的最小适应度值等于临时子群体的最大适应度值时, 从 $f(\hat{g}_i)$ 和 $f(\hat{g}_j)$ 对应的个体集 $\{\hat{g}_i\} \bigcup \{\hat{g}_j\}$ 中选出 $|\hat{g}_i|$ 个子群体作为优胜子群体, 所以有

$$e_t = \frac{|\hat{g}_i|}{|\hat{g}_i| + |\hat{g}_j|}. \quad (15)$$

其中: $1 \leqslant |\hat{g}_i| \leqslant N_S$, $1 \leqslant |\hat{g}_j| \leqslant N_T$.

当只有一个优胜子群体的适应度值与临时子群体的适应度值相同, 且所有临时子群体的适应度值均相同时, 此优胜子群体被选择的概率最小, 为

$$e_n = \min\left\{\frac{|\hat{g}_i|}{|\hat{g}_i| + |\hat{g}_j|}\right\} = \frac{1}{N_T + 1}. \quad (16)$$

由此定理得证. □

当前代优胜子群体被选择后, 以概率 1 参与下一代进化, 所以下一代优胜子群体生成的概率与当前代被选取的概率相同.

2) 临时子群体被释放并按照均匀分布, 在 F 中散布 N_T 个个体作为新的临时子群体中心 $\hat{g}_j^t$, 参与下一代进化. 因为均匀分布是连续分布, 在某一点的概率为零, 所以在此近似用 $\hat{g}_j^t$ 为中心的任意小邻域 C 上的分布函数来表征临时子群体中心散布到 $\hat{g}_j^t$ 的概率, 则临时子群体生成的概率为

$$h_t = \{P_{D''}^t(\hat{g}_j) : \hat{g}_j \in G_T\} = \int\int_C \cdots \int f_2(g_i^1, g_i^2, \cdots, g_i^l)\mathrm{d}C. \quad (17)$$

其中: $f_2(\cdot)$ 为 l 维均匀分布的密度函数, 当 MEA 搜索空间确定后, $f_2(\cdot)$ 即被确定; G_T 为临时子群体组成的集合, 包含的子群体数目是 N_T.

因为均匀分布在解空间上产生的各子群体中心的概率相同, 故临时子群体生成的概率为常数, 用 h 表示, 即

$$h = h_t. \quad (18)$$

在 MEA 进化过程中, 由于适应度值最大的优胜子群体的中心总被保留下来, 参与下一代的进化, 所以下一代群体的最大适应度值总不小于当前代的最大适应度值, 即

$$f(\hat{g}^{t+1}) - f(\hat{g}^t) = \alpha_t f^* \geqslant 0 \quad (19)$$

恒成立. 其中: $\alpha_t \in [0,1]$, $\lim\limits_{T\to\infty}\sum\limits_{t=1}^{T}\alpha_t \leqslant \dfrac{f^*-f(\hat{g}^1)}{f^*}$.

由于解空间有无穷多个可能解, 由式 (12) 易知

$$\sum_{t=1}^{\infty} d_t=\infty. \tag{20}$$

当 $\hat{g}_i \notin B^*$ 且 $\hat{g}_i \notin H$ 时, 因为 $\alpha_t \in (0,1]$, 所以有

$$\sum_{t=1}^{\infty} d_t\alpha_t=\infty. \tag{21}$$

3 MEA 的几乎处处收敛性证明

定理 2 描述 MEA 的最大适应度值函数过程 $\{f(\hat{g}^t):\hat{g}^t\in\hat{B},t\geqslant 1\}$ 是非负有界下鞅数列[11], 即 $E\{f(\hat{g}^{t+1})/\hat{g}^t\}\geqslant f(\hat{g}^t), t\geqslant 1$.

证明 根据 MEA 算法模型, 趋同算子总是以上一代子群体的最大适应度值个体作为中心进行分布, 保留了上一代子群体的最大适应度值个体到下一代, 因此下一代子群体的最大适应度值不会小于上一代子群体的最大适应度值. 异化算子总是强制性以概率 1 让上一代优胜子群体的中心直接作为下一代子群体的中心参与进化, 即上一代最大适应度值的子群体中心必然参与下一代进化. 因此下一代群体的最大适应度值不会小于上一代群体的最大适应度值, 即

$$E\{f(\hat{g}^{t+1})/\hat{g}^t\}\geqslant f(\hat{g}^t)>0. \tag{22}$$

综上所述, 描述 MEA 的最大适应度值函数过程 $\{f(\hat{g}^t):\hat{g}^t\in\hat{B},\ t\geqslant 1\}$ 是非负有界下鞅数列. □

定理 3 如果描述 MEA 的最大适应度值函数的非负有界下鞅数列 $\{f(\hat{g}^t):\hat{g}^t\in\hat{B},\ t\geqslant 1\}$ 满足 $\sup E(f(\hat{g}^t))<\infty$, 则 $\lim f(\hat{g}^t)$ (a.s.) 有意义, 记为

$$\lim_{t\to\infty} f(\hat{g}^t)=f(\hat{g}^\infty)\,(\text{a.s.}).$$

证明 由于优化问题的全局最大适应度值 f^* 存在 (否则优化问题无意义), 有

$$f(\hat{g}^t)\leqslant f^*<\infty, \tag{23}$$

根据数学期望的性质有

$$E(f(\hat{g}^t))\leqslant f^*<\infty, \tag{24}$$

所以 $\sup E(f(\hat{g}^t))<\infty$, 那么对于描述 MEA 的最大适应度值函数的非负有界下鞅数列必有 $\lim f(\hat{g}^t)=f(\hat{g}^\infty)$(a.s.) 成立. 换言之, 在 MEA 中, 群体中心几乎处处收敛于全局最优解. □

定理 4 在 MEA 中, $f(\hat{g}^\infty)=f^*$ (a.s.), 即 MEA 几乎处处收敛于全局最优点.

证明 证明过程分以下两个步骤进行.

Step 1 证明 $\sum\limits_{t=1}^{\infty} d_t\alpha_t P(\hat{g}^t\notin B^*)<\infty$.

MEA 从所有子群体中选择适应度值最大子群体的概率为

$$P_D^{t\,*}(y_i^t,\hat{g}^{t+1})=|\hat{g}^{t+1}|/|\hat{B}|,\ \hat{g}^{t+1}\in\hat{B}\in F. \tag{25}$$

其中: $|\hat{B}|$ 为 $\hat{B}$ 所包含的最优个体数, $|\hat{g}^{t+1}|$ 为 $\hat{g}^{t+1}$ 在 $\hat{B}$ 中出现的次数, $\hat{g}^{t+1}$ 满足

$$f(\hat{g}^{t+1})=\max\{f(y_i^t),\ 1\leqslant i\leqslant N\}.$$

令 $P_D^t(\hat{g}_i,\hat{g}_j)$ 表示异化操作中子群体中心由 $\hat{g}_i$ 转移到 $\hat{g}_j$ 的概率. 由 MEA 模型和条件期望的定义有

$$\begin{aligned}
&E\{f(G^{t+1})/G^t\}=\\
&\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)\sum_{y_j^t}P_S^t(\hat{g}_j^t,y_j^t)\times\\
&\sum_{\hat{g}^{t+1}}P_D^{t\,*}(y_j^t,\hat{g}^{t+1})f(\hat{g}^{t+1})=\\
&\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)\sum_{y_j^t}P_S^t(\hat{g}_j^t,y_j^t)\times\\
&\sum_{\hat{g}^{t+1}}|\hat{g}^{t+1}|/|\hat{B}|f(\hat{g}^{t+1})=\\
&\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)\sum_{y_j^t}P_S^t(\hat{g}_j^t,y_j^t)f(y_j^t)\geqslant\\
&d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t).
\end{aligned} \tag{26}$$

由条件期望的性质和式 (26) 可得

$$\begin{aligned}
&E\{f(G^{t+1})\}-E\{f(G^t)\}=\\
&E\{E\{f(G^{t+1})/G^t\}\}-E\{f(G^t)\}=\\
&\sum_{\hat{g}^t}P(G^t=\hat{g}^t)\{E\{f(G^{t+1})/G^t\}-f(G^t)\}\geqslant\\
&\sum_{\hat{g}^t}P(G^t=\hat{g}^t)\Big\{d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\Big\}\geqslant\\
&\Big[\sum_{\hat{g}^t\in B^*}+\sum_{\hat{g}^t\in H/B^*}+\sum_{\hat{g}^t\notin H}\Big]P(G^t=\hat{g}^t),\\
&\Big\{d_t\sum_{\hat{g}_j}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\Big\}=\\
&P(\hat{g}^t\in B^*)\Big[d_t\sum_{\hat{g}_j}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\Big]+\\
&P(\hat{g}^t\in H/B^*)\Big[d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\Big]+\\
&P(\hat{g}^t\notin H)\Big[d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\Big].
\end{aligned} \tag{27}$$

分 3 种情况讨论 $d_t\sum\limits_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)$ 的取值.

1) 当 $\hat{g}^t\in B^*$ 时, $f(\hat{g}_j^t)=f(\hat{g}^t)=f^*$, 可得

$$\begin{aligned}
&d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)=\\
&d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f^*-f^*.
\end{aligned} \tag{28}$$

假如 $\hat{g}^t\in B_t^*\bigcup\hat{g}^t\in G_S$, 且优胜子群体集中包含子群

体的适应度值不全相同, 由定理 1 可得

$$d_t\sum_{\hat{g}_j^t}P_{D'}^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\geqslant d_tf^*-f^*;\tag{29}$$

假如 $\hat{g}^t\in B_t^*\bigcup\hat{g}^t\in G_S$, 且所有优胜子群体的适应度值均相同, 由定理 1 可得

$$d_t\sum_{\hat{g}_j^t}P_{D'}^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\geqslant-\Big(1-\frac{d_t}{N_T+1}\Big)f^*;\tag{30}$$

假如 $\hat{g}^t\in B^*\bigcup\hat{g}^t\in G_T$, 则临时子群体集中包含适应度值最大的子群体, 由式 (17) 和 (18) 可得

$$d_t\sum_{\hat{g}_j^t}P_{D''}^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\geqslant-(1-d_th)f^*.\tag{31}$$

在初始化 MEA 时, 由于考虑到进化的实时性, N_T 不能选得太大, 所以有

$$\frac{1}{N_T+1}>h.\tag{32}$$

因此当 $\hat{g}^t\in B^*$ 时, 综合式 (28) ∼ (32) 可得

$$d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\geqslant-(1-d_th)f^*.\tag{33}$$

2) 当 $\hat{g}^t\in H\backslash B^*$ 时, $f(\hat{g}_j^t)=f(\hat{g}^t)$, 适应度值最大的子群体陷入局部最优, 所以有

$$\begin{aligned}&d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)=\\&d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}^t)-f(\hat{g}^t),\end{aligned}\tag{34}$$

分析过程同 $\hat{g}^t\in B^*$, 且有 $f(\hat{g}^t)<f^*$. 因此当 $\hat{g}^t\in H\backslash B^*$ 时, 有

$$d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)>-(1-d_th)f^*.\tag{35}$$

3) 当 $\hat{g}^t\notin H$ 时, $f^*>f(\hat{g}_j^t)>f(\hat{g}^t)$ 成立.

假如 $\hat{g}^t\in G_S/H$ 且优胜子群体集中包含子群体的适应度值不全相同, 那么由定理 1 可得

$$d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\geqslant d_tf(\hat{g}_j^t)-f(\hat{g}^t).\tag{36}$$

由式 (19) 和 (36) 可得

$$d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)>d_t\alpha_tf^*-(1-d_t)f^*.\tag{37}$$

假如 $\hat{g}^t\in G_S/H$ 且优胜子群体集中包含子群体的适应度值均相同, 那么由定理 1 可得

$$d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\geqslant\frac{d_tf(\hat{g}_j^t)}{N_T+1}-f(\hat{g}^t).\tag{38}$$

由式 (19) 和 (38) 可得

$$\begin{aligned}&d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)>\\&\frac{d_t\alpha_tf^*}{N_T+1}-\frac{(1+N_T-d_t)f^*}{N_T+1}.\end{aligned}\tag{39}$$

假如 $\hat{g}^t\in G_T/H$, 临时子群体集中包含适应度值最大的子群体, 由式 (17) 和 (18) 可得

$$d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)\geqslant d_thf(\hat{g}_j^t)-f(\hat{g}^t).\tag{40}$$

由式 (19) 和 (40) 可得

$$\begin{aligned}&d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)>\\&d_th\alpha_tf^*-(1-d_th)f^*.\end{aligned}\tag{41}$$

因此当 $\hat{g}^t\notin H$ 时, 综合式 (32), (37), (39) 和 (41) 可得

$$\begin{aligned}&d_t\sum_{\hat{g}_j^t}P_D^t(\hat{g}_i^t,\hat{g}_j^t)f(\hat{g}_j^t)-f(\hat{g}^t)>\\&(d_th\alpha_t-1+d_th)f^*;\end{aligned}\tag{42}$$

综合式 (33), (35) 和 (42) 可得

$$\begin{aligned}&E\{f(G^{t+1})\}-E\{f(G^t)\}>\\&-3f^*+hf^*[d_tP(\hat{g}^t\in B^*)+d_tP(\hat{g}^t\in H/B_t^*)+\\&P(\hat{g}^t\notin H)(d_t\alpha_t+d_t)].\end{aligned}\tag{43}$$

又因为恒有

$$f^*>E\{f(G^{T+1})\}-E\{f(G^1)\},\tag{44}$$

所以由式 (44), 在式 (43) 中对 t 从 1 到 T 求和可得

$$\begin{aligned}4/h>&\sum_{t=1}^Td_tP(\hat{g}^t\in B^*)+\sum_{t=1}^Td_tP(\hat{g}^t\in H/B^*)+\\&\sum_{t=1}^T(d_t\alpha_t+d_t)P(\hat{g}^t\notin H).\end{aligned}\tag{45}$$

令 $T\to\infty$, 由式 (45) 可推出

$$\sum_{t=1}^{\infty}d_t\alpha_tP(\hat{g}^t\notin H)<\infty,$$

$$\sum_{t=1}^{T}d_tP(\hat{g}\in H/B^*)<\infty,$$

从而有

$$\sum_{t=1}^{\infty}d_t\alpha_tP(\hat{g}^t\notin B^*)<\infty.\tag{46}$$

Step 2　证明 $\lim\limits_{t\to\infty}f(\hat{g}^t)=f(\hat{g}^\infty)=f^*$, a.s..

为了书写方便, 令 $\varsigma=\lim\limits_{t\to\infty}f(g^t)=f(g^\infty)$. 因为 $\varsigma\leqslant f^*$, a.s., 所以只需证明当 $t\geqslant T$ 时, $P(\varsigma<f^*)=0$ 即可, 用反证法.

假设 $P(\varsigma<f^*)>0$, 则对于任何 $\beta>0$, 由文献 [12] 可知, 对于一切 $t\geqslant T$ 有

$$P\{\hat{g}^t\notin B^*\}\geqslant\beta.\tag{47}$$

由式 (21) 和 (47) 可得

$$\sum_{t=1}^{\infty} P(\hat{g}^t \notin B^*) d_t \alpha_t \geqslant \beta \sum_{t=1}^{\infty} d_t \alpha_t = \infty. \tag{48}$$

式 (48) 与 (46) 相矛盾, 因此假设 $P(\varsigma < f^*) > 0$ 不成立. 由此定理 4 得证, 即 MEA 几乎处处收敛于全局最优解 f^*. □

从定理 4 的证明过程可以看出, MEA 初始群体的选择对算法的收敛性没有影响, 但会影响收敛速度, 因为初始群体的选择方法决定了初始群体的个体的取值情况, 从而决定了 $f(\hat{g}^1)$ 的值. 如果 $f(\hat{g}^1)$ 小, 则收敛速度慢, 如果 $f(\hat{g}^1)$ 大, 则收敛速度快.

4 结 论

本文根据 MEA 的优化模型, 详细地分析了趋同和异化算子的转移概率, 为 MEA 的数学研究奠定了基础.

本文利用下鞅理论取代传统的 Markov chain 理论, 分析了 MEA 的收敛性, 避开了求解转移矩阵及特征值的复杂性, 反映了 MEA 的收敛速度不仅与趋同和异化算子的操作方法有关, 而且与初始群体的选择有关, 由此说明算法参数和问题参数控制着 MEA 的优化进程.

本文利用下鞅理论分析了思维进化过程, 为思维进化算法的理论研究开拓了一条新的研究途径. 今后进一步研究的重点是定量地给出 MEA 进化速度的表达式.

参考文献(References)

[1] Sun Chengyi, Sun Yan. Mind-evolution-based machine learning: Framework and the implementation of optimization[C]. Proc of IEEE Int Conf on Intelligent Engineering Systems. Piscataway: IEEE Press, 1998: 355-359.

[2] 王芳, 谢克明, 刘建霞. 基于群体智能的思维进化算法设计[J]. 控制与决策, 2010, 25(1): 145-148.
(Wang F, Xie K M, Liu J X. Swarm intelligence based MEA design[J]. Control and Decision, 2010, 25(1): 145-148.)

[3] 刘建霞, 王芳, 谢克明. 基于改进的思维进化算法的宽带阻抗变换器设计[J]. 中北大学学报: 自然科学版, 2008, 29(5): 435-438.
(Liu J X, Wang F, Xie K M. Design of broadband impedance transformer based on improved mind evolutionary algorithm[J]. J of North University of China: Natural Science Edition, 2008, 29(5): 435-438.)

[4] Yan Gaowei, Xie Gang, Qiu Yuxia. MEA based nonlinearity correction algorithm for the VCO of LFMCW radar level gauge[C]. The 10th Int Conf on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing. Regina: University of Regina, 2005: 461-470.

[5] 谢刚, 盛彬, 王芳. 一种基于相容粒度空间的图像分割方法[J]. 控制与决策, 2013, 28(2): 317-320.
(Xie G, Sheng B, Wang F. An image segmentation method based on tolerance granular space[J]. Control and Decision, 2013, 28(2): 317-320.)

[6] Liu Jianxia, Li Nan, Xie Keming. Application of chaos mind evolutionary algorithm in antenna arrays synthesis[J]. J of Computers, 2010, 5(5): 717-724.

[7] 刘洋. 基于思维进化计算和蚂蚁算法的网格资源分配[J]. 计算机工程, 2007, 33(7): 172-174.
(Liu Y. Grid resource allocation based on mind evolutionary[J]. Computer Engineering, 2007, 33(7): 172-174.)

[8] Wang Chuanlong, Xie Keming. Convergence of a new evolutionary computation algorithm in continuous state space[J]. Int J of Computer Mathematics, 2002, 79(1): 27-37.

[9] Zhang Zhijun, Qiu Yuxia, Xie Keming. Convergence analysis on an improved mind evolutionary algorithm[C]. Proc of the 6th Int Conf on Natural Computation. Yantai: Yantai University, 2010: 2316-2320.

[10] 谢克明, 邱玉霞. 基于数列模型的思维进化算法收敛性分析[J]. 系统工程与电子技术, 2007, 29(2): 308-311.
(Xie K M, Qiu Y X. Convergence analysis of mind evolutionary algorithm based on sequence model[J]. Systems Engineering and Electronics, 2007, 29(2): 308-311.)

[11] 程士宏. 高等概率论[M]. 北京: 北京大学出版社, 1996, 12: 77-126.
(Cheng S H. Advanced probability theory[M]. Beijing: Peking University Press, 1996, 12: 77-126.)

[12] 徐宗本, 聂赞坎, 张文修. 遗传算法的几乎必然强收敛性—— 鞅方法[J]. 计算机学报, 2002, 25(8): 785-793.
(Xu Z B, Nie Z K, Zhang W X. Almost sure convergence of genetic algorithms: A martingale approach[J]. Chinese J of Computers, 2002, 25(8): 785-793.)

(责任编辑: 滕 蓉)