社団法人　情報処理学会　研究報告
IPSJ SIG Technical Report

2005－MPS－56（18）
2005／9／22

遺伝的アルゴリズムで、突然変異確率を小さくして選択圧力を大きくすると、定常確率が最良個体ばかりの一様集団に集中する

鈴木譲　　チャンディデシルバ [*2]
Joe Suzuki[*1]　　U. Chandimal de Silva[*2]

大阪大学 大学院理学研究科
Graduate School of Science, Osaka University

Abstract: 遺伝的アルゴリズム (GA) に限らず、数値例で結論を出そうとすると、その数値例の選び方にとっては、正反対の結論が得られる。本論文では、GA で、突然変異確率を小さくして選択圧力を大きくすると、定常確率が最良個体ばかりの一様集団 (複数可) に集中する、という現象を数学的に証明する。本論文の結果は、一部 2005 年 6 月に Washington DC で開催された GECCO 2005 で発表済みである。

紙面の都合上、要点だけを述べる。

1 準備

1.1 遺伝的アルゴリズムの設定

各個体は、一定の長さ L の 2 進列で表現されているものとする。$J := \{\underbrace{0\cdots0}_{L}, \underbrace{0\cdots1}_{L}, \cdots, \underbrace{1\cdots1}_{L}\}$ を個体全体の集合とする ($N := 2^L$ 種類の個体がある)。本論文で扱う遺伝的操作を以下のように定義する (Vose 1999)。

突然変異: 各個体 $k \in J$ を確率 μ_i で $k \oplus i \in J$ に置き換える操作である $i \in J$。ここで、$x \oplus y$ は等しい長さの 2 進列 2 個 x, y をビットごとに排他的論理和をとって得られる 2 進列である。ただし、$(\mu_i)_{i\in J}$ は $\sum_{i\in J}\mu_i = 1, \mu_i \geq 0$ をみたすものとする。以下では、$\mu = 1 - \underbrace{0\cdots0}_{L}$ とおく。たとえば、$\mu_i = \mu^{|i|}(1-\mu)^{L-|\mu|}$ という形の突然変異がよく用いられている。ただし、$|i|$ は $i \in J$ における 1 の個数である。

交叉: 2 個の個体 $k, j \in J$ を、確率 χ_i で 2 個の個体 $k \otimes i \oplus \bar{i} \otimes j, j \otimes i \oplus \bar{i} \otimes k \in J$ に置き換えて、そのどちらかの個体を確率 1/2 で選択する操作である ($i \in J$)。ただし、$\chi \geq 0$ および $c_i \geq 0$ $\sum_{i\in J} c_i = 1$ を満足する $(c_i)_{i\in J}$ を用いて、$(\chi_i)_{i\in J}$ は、

$$\chi_i = \begin{cases} \chi c_i, & i \neq \underbrace{0\cdots0}_{L} \\ 1 - \chi + \chi c_0, & i = \underbrace{0\cdots0}_{L} \end{cases} \tag{1}$$

とかけるものとする。一様交叉 $c_i = 2^{-l}$、一点交叉

$$c_i = \begin{cases} 1/(L-1), & i = 1\underbrace{0\cdots0}_{L-1}, 11\underbrace{0\cdots0}_{L-2}, \cdots, \underbrace{1\cdots1}_{L-1}0 \\ 0, & \text{otherwise} \end{cases}$$

などもこれに含まれる。

M を集団中に含まれる個体の数とする。世代交代を以下のように定義する。

世代交代: 現在の現世代の集団から次世代の集団に変更する。

1. 現世代の集団から確率
$$\frac{c[i]f(i)}{\Sigma_{j\in J} c[j]f(j)}, \tag{2}$$
で個体 $i \in J$ を選択する。ただし、$f : J \to \mathbf{R}^+$ を適合度関数、$c[i]$ を現世代中の $i \in J$ の頻度とする。
2. 1 をもう 1 度繰り返して、個体 $j \in J$ を得る。
3. 個体 i, j に交叉を施して、個体 $k \in J$ を得る。
4. 個体 k に突然変異を施して、$k' \in J$ を得る。
5. 1.-4. を M 回繰り返して、M 個の個体を得る (次世代の集団を得る)。

鈴木譲, 大阪大学大学院理学研究科数学専攻
573-1194 豊中市待兼山町 1-1
E-Mail: suzuki@math.sci.osaka-u.ac.jp

ランダムに初期世代の集団を発生した後、世代交代を有限回繰り返す処理を遺伝的アルゴリズム (GA) とよぶ。

1.2 有限マルコフ連鎖

現世代の集団から次世代の集団への確率は、

$$(c[i])_{i\in J},\ \sum_{i\in J} c[i] = M,\ c[i] \geq 0$$

を用いて記述されることがわかる。個体長 L および集団サイズ M が与えられたもとで、$(c[i])_{i\in J}$ が同じ集団を等価な集団とよぶ。たとえば、$L=2, M=3$ であれば、2 個の集団 $\underbrace{10,11,10}_{M}$ および $\underbrace{10,10,11}_{M}$ は、$(c[00],c[01],c[10],c[11])$ が同じ $(0,0,2,1)$ であるので同値である。そのような同値な類を状態とよび、その $(c[i])_{i\in J}$ で代表させる。状態の集合 S は、有限集合で、L, M が決まれば一意である。例えば、$L=2, M=3$ の場合、$J=\{00,01,10,11\}$ であり、考えられる $(c[00],c[01],c[10],c[11])$ は、

$$(0,0,0,3),(0,0,1,2),(0,0,2,1),(0,0,3,0),(0,1,0,2),$$
$$(0,1,1,1),(0,1,2,0),(0,2,0,1),(0,2,1,0),(0,3,0,0),$$
$$(1,0,0,2),(1,0,1,1),(1,0,2,0),(1,1,0,1),(1,1,1,0),$$
$$(1,2,0,0),(2,0,0,1),(2,0,1,0),(2,1,0,0),(3,0,0,0).$$

となる。

以下では、状態と集団は同じ意味で用いるものとする。

状態集合が S で、状態 $s\in S$ から状態 $t\in S$ への推移確率が $Q=(Q(t|s))_{t,s\in S}$ で表された場合、状態 $s\in S$ の定常確率は、行列 Q の固有値 1 に対応する固有ベクトルとして定義される。$S=\{1,2,\cdots,\sigma\}$ であれば、

$$\begin{aligned}([q[1],\cdots,q[\sigma]])&\begin{pmatrix} Q(1|1) & \cdots & Q(\sigma|1)\\ \cdots & \cdots & \cdots \\ Q(1|\sigma) & \cdots & Q(\sigma|\sigma)\end{pmatrix}\\ = \ &([q[1],\cdots,q[\sigma]])\ .\end{aligned}$$

の解となる。Cramer の公式を用いれば、$s\in S$ の定常確率は

$$q[s] = \frac{\det(Q[s]-I)}{\sum_{t\in S}\det(Q[t]-I)} \tag{3}$$

で表される。ただし、推移行列 Q の $s\in S$ に対応する行をゼロベクトル $\underbrace{(0,...0)}_{\sigma}$ に置き換えた行列を $Q[s]$ とおいた。

2 従来の研究

2.1 $\mu\to 0, \chi=0$

U を同じ個体ばかりからなる集団 (一様集団) の集合とする。T. Davis は、$\lim_{\mu\to 0} q[s]$ の値の計算を試みた。しかし、一般に $\lim_{\mu\to 0}\det(Q[s]-I)=0$ が成立するので、(3) は直接用いることはできない (分子分母とも 0 になる)。そのかわりに、他の記法 $Q^*[u], u\in U$ を用いて、$\lim_{\mu\to 0}\det(Q[s]-I), s\in S$ ではなく、$\lim_{\mu\to 0}\det(Q^*[u]-I), u\in U$ の形で $\lim_{\mu\to 0} q[s]$ を表現した。

まず、$Q^*[u](t|s), u\in U, s,t\in S-U\cup\{u\}$ を

$$Q^*[u](t|s) = \begin{cases} 0, & s=u \\ Q(t|s) + \displaystyle\sum_{v\in U-\{u\}, t\in S[v]} \frac{Q(v|s)}{L}, & s\neq u\end{cases}$$

で定義する。ここで、$S[v]$ は $v\in U$ の 1 個体を 1 ビットだけ 0,1 を反転させた集団の集合である。定義から、各 $s\in S-U$ について

$$Q^*[u](u|s) = Q(u|s) \tag{4}$$

が成立する。また、各 $u\in U$ について、$S[u]$ には L 個の要素があるので、

$$\sum_{t\in S-U\cup\{u\}} Q^*[u](t|s) \tag{5}$$
$$\begin{aligned} &= \sum_{t\in S-U\cup\{u\}} Q(t|s) + \sum_{v\in U-\{u\}}\sum_{t\in S[v]}\frac{Q(v|s)}{L}\\ &= \sum_{t\in S-U\cup\{u\}} Q(t|s) + \sum_{v\in U-\{u\}} Q(v|s)\\ &= \sum_{t\in S} Q(t|s)\\ &= 1\ . \end{aligned} \tag{6}$$

が成立する。

記法 $\{Q^*[u](t|s)\}_{t,s\in S}$ を用いて、Davis は $\chi=0$ のもとで、

$$\lim_{\mu\to 0} q[s] = \begin{cases} \dfrac{\lim_{\mu\to 0}\det(Q^*[s]-I)}{\sum_{t\in U}\lim_{\mu\to 0}\det(Q^*[t]-I)}, & s\in U \\ 0, & s\in S-U\end{cases} \tag{7}$$

が成立することを示した。

2.2 $\mu\to 0$, $p\to\infty$, $\chi=0$

U^* $(\subseteq U)$ を、$f(i)$ が最大の値 f_{max} をもつ $i\in J$ からなる一様集団の集合とする。

$g_p : \mathbf{R}^+ \to \mathbf{R}^+$ を $p\in\mathbf{R}^+$ および $0<a<b$ に関して、

1. $g_p(a) < g_p(b),\, p > 0$
2. $\dfrac{g_p(b)}{g_p(a)} \to \infty,\, p \to \infty$

となる任意の関数とする。たとえば、$g_p(x) = x^p$ は、上記 2 条件を満足している。そして、以降、適合度関数 f を $g_p \circ f$ で置き換える。したがって、(2) は

$$\frac{c[i]g_p(f(i))}{\Sigma_{j\in J}c[j]g_p(f(j))}\ . \tag{8}$$

となる。また、以降、p を選択圧力とよぶ。

鈴木 (1998) は、$\chi = 0$ のもとで、

$$\lim_{p\to\infty}\lim_{\mu\to 0} q[s] \qquad \begin{cases} > 0, & s \in U^* \\ = 0, & s \in S - U^* \end{cases} \tag{9}$$

を示した。

Albuquerque,Mazza (2000) も、$\chi = 0$ のもとで同様の結果を証明した。

Davis, Suzuki, Albuquerque-Mazza の結果はいずれも $\mu \to 0$、もしくは $\mu \to 0, p \to \infty$ の場合の定常確率に関するものであった。しかしながら、それらは $\chi = 0$ のもとでの導出であって、GA では特に交叉が本質的な役割を果たすと信じられているので、強い制約のもとでの結果といわざるを得ない。

3 主結果

本論文では、$\chi = 0$ の制約を除去し、第 1 節で定義した一般的な GA について、(9) を示す。GA では、望ましい集団に

1. 高い定常確率が割り当てられ、
2. 速く収束する

ことの 2 条件が要求される。本論文では、このうち前者について検討している。数学的には以下のように表現できる: 各 $s \in S - U^*$ と任意の $\epsilon > 0$ について、

$$\mu < \gamma(\epsilon), \delta(\epsilon, \mu) < p \Longrightarrow q[s] < \epsilon$$

なる、$\gamma : \mathbf{R}^+ \to \mathbf{R}^+, \delta : \mathbf{R}^+ \times \mathbf{R}^+ \to \mathbf{R}^+$ が存在する。すなわち、$\epsilon > 0$ に応じて、μ, p の限界を設定できる。

また、本論文では、第 1 節で定義したような一般的な GA を扱っている点に注意したい。SA 的な GA ではなく、極限は定常確率の極限をとっている。μ, χ の各値は、世代交代の間、固定である。

GA のポジティブな性質を数学的に証明した結果は、非常に少ない。数値例だけで結論を出す場合は、その数値例についての結論になる。本論文では、一般的な結論を出している。

定理 1 突然変異確率を *0* に近づけ、選択圧力を十分大きくとると、適合度最大の個体ばかりの一様集団のいずれかが生じる定常確率が *1* となる。

4 証明

$J^* \subseteq J$ を $f(i) = f_{max}$ なる個体 $i \in J$ の集合とする。

$i \in J^*$ を含むが、それ以外の $j \in J^*$ は含まない非一様集団 s の集合を V と書く。そのような i だけからなる一様集団を $v(s)$ と書くと、 $s \in V$ について $Q(v(s)|s) \to 1$ が成立する。

$H(i,j) = 1$ となる $i,j \in J^*$ を含むが、それ以外の $k \in J^*$ を含まない非一様集団 s の集合を W と書く。ただし、$H(i,j)$ は $i,j \in J$ のハミング距離である。$s \in W$ について、$T(s)$ をそのような i,j のみからなる集団の集合 (一様でも非一様でもよい) をとし、$v(s), w(s)$ をそれぞれそのような i,j からなる一様集団とする。このとき、$s \in W$ について、$\displaystyle\sum_{t\in T(s)} Q(t|s) \to 1$ が成立する。

以下、$Q^*[u](t|s) \to 0, t \in S - U - V - W, s \in V \cup W$ を示す。

1. $s \in V, u = v(s)$.
$$\sum_{t\in V\cup W\cup\{u\}} Q^*[u](t|s) \geq Q(v(s)|s) \to 1$$
2. $s \in V, u \neq v(s)$. $S[v(s)] \subseteq V \cup W$ であるので、
$$\begin{aligned}\sum_{t\in V\cup W\cup\{u\}} Q^*[u](t|s) &\geq \sum_{t\in V\cup W}\frac{1}{L}\sum_{t\in S[v],v\neq u} Q(v|s) \\ &\geq \sum_{t\in V\cup W}\frac{1}{L}\sum_{t\in S[v(s)]} Q(v(s)|s) = Q(v(s)|s) \to 1\ .\end{aligned}$$
3. $s \in W,\, u = v(s)$ or $u = w(s)$. 一般性を失うことなく、$u = v(s)$ を仮定すると、
$$\begin{aligned}&\sum_{t\in V\cup W\cup\{u\}} Q^*[u](t|s) \\ \geq\ &\sum_{t\in V\cup W\cup\{v(s)\}} Q(t|s) + \sum_{t\in V\cup W}\frac{1}{L}\sum_{t\in S[w(s)]} Q(w(s)|s) \\ =\ &\sum_{t\in V\cup W\cup\{v(s)\}} Q(t|s) + Q(w(s)|s) \\ \geq\ &\sum_{t\in T(s)} Q(t|s) \to 1\end{aligned}$$

4. $s \in W$, $u \neq v(s), u \neq w(s)$.

$$\begin{aligned}
&\sum_{t \in V \cup W \cup \{u\}} Q^*[u](t|s) \\
\geq &\sum_{t \in V \cup W} Q(t|s) + \sum_{t \in V \cup W} \frac{1}{L} \sum_{t \in S[v(s)]} Q(v(s)|s) \\
&+ \sum_{t \in V \cup W} \frac{1}{L} \sum_{t \in S[w(s)]} Q(w(s)|s) \\
= &\sum_{t \in V \cup W} Q(t|s) + Q(v(s)|s) + Q(w(s)|s) \\
\geq &\sum_{t \in T(s)} Q(t|s) \to 1
\end{aligned}$$

したがって、

$$W_{\mu,p}[u] = (Q^*[u](t|s))_{s,t \in V \cup W - U},\ W[u] = \lim_{\mu,p} W_{\mu,p}[u]$$

$$D_{\mu,p}[u] = (Q^*[u](t|s))_{s,t \in S-U-V-W},\ D[u] = \lim_{\mu,p} D_{\mu,p}[u]$$

$\det(Q^*[u] - I) = \det(W[u] - I)\det(D[u] - I)$ が成立する。そこで、

1. $u \in U^*$ について、
   (a) $\det(W[u] - I) \neq 0$
   (b) $\det(D[u] - I) \neq 0$
2. $u \in U - U^*$ について、$\det(W[u] - I) = 0$

をいえば十分である。

1.(a) および 2. の導出については、Perron-Frobenius の定理 ”$A = (a_{i,j})$ について

$$\forall i \sum_j a_{i,j} = 1 \Longleftrightarrow \det(A - I) = 0$$

が成立”を適用する。

1.(a) については、$u \in U^*$ および $v(s) = u$ となる $s \in V$ について、$Q[u](u|s) = Q(u|s) \to 1$ が成立する。、そのような u, s について、$\sum_{t \in V \cup W} Q[u](t|s) \to 0$ が成立するので、1.(a) が成立する。

2. については、$\sum_{t \in V \cup W \cup \{u\}} Q^*[u](t|s) \to 1$ が成立するので、$u \in U - U^*, s \in V \cup W$ について $Q[u](u|s) = Q(u|s) \to 0$ が成立する。したがって、$\sum_{t \in V \cup W} Q^*[u](t|s) \to 1$ となり、2. が成立する。

1.(b) については、$|J - J^*| = 0, 1$ のときまたは $L = 1$ のとき、$S - U - V - W = \{\}$ より、$D[u], u \in U$ が 0×0 の行列となる。したがって、$|J - J^*| \geq 2$ かつ $L \geq 2$ を仮定してよい。そして、以下の命題を適用する。

命題 1 $|J - J^*| \geq 2$ かつ $L \geq 2$ のとき、$H(i,j) = 1$ なる $i \in J^*$ が存在するような $j \in J - J^*$ が 2 個以上存在する。

1.(b) の証明に戻ると、仮定 $|J - J^*| \geq 2$ と命題 1 から、ある $i, h \in J^*$ について $H(j,i) = H(k,h) = 1$ となるような $j, k \in J - J^*$ が存在する。一般性を失うことなく、$f(j) \geq f(k)$ を仮定し、

v: そのような j からなる一様集団

t: $c[i] = 1, c[j] = M - 1$ なる集団

s: $1 \leq c[j] \leq M - 1, 1 \leq c[k] \leq M - 1$, and $c[j] + c[k] = M$ なる集団

とおく。このとき、$u \in U^*, v \in U - U^*$ であるので、$Q(v|s)$ は正の値に収束し、$Q^*[u](t|s) \geq \frac{1}{L} Q(v|s)$ も正の値に収束する。$t \in V \cup W$ であるので、$\sum_{r \in S-U-V-W} Q^*[u](r|s)$ は 1 より小さな値に収束する。

参考文献

1. Albuquerque, P. and Mazza, C.,(2000) Foundations of Genetic Algorithms-6, San Francisco, CA. Morgan Kaufmann.
2. Davis, T. and Principe, J.C. (1991) A simulated annealing like convergence theory for the simple genetic algorithm. In Belew, R. and Bookers, L., editors, Proc. of the Fourth International Conference on genetic Algorithms, pages 174-181, San Mateo, CA. Morgan Kaufmann.
3. De Jong, K A. (1975) An analysis of the behaviour of a class of genetic adoptive systems. Ph.D. dissertation, University of Michigan.
4. Schmitt, Lothar M. (2001) Theory of genetic algorithms. Theoretical Computer Science, 259(1-2):1-61
5. Suzuki, J. (1995) A Markov Chain Analysis on Simple Genetic Algorithms. IEEE Transactions on Systems, Man and Cybernatics, 25(2) pages 655-659, April 1995.
6. Suzuki, J. (1998) A further result on the Markov chain model of genetic algorithms and its application to a simulated annealing-like strategy. IEEE Transactions on Systems, Man and Cybernatics, 28(1)
7. Vose, M. D. (1999) The Simple Genetic Algorithm, Foundations and Theory. The MIT Press.