

## OneMax 問題における収束時間予測法の検討

古賀 仁 信<sup>†1</sup> 坂本 真人<sup>†2</sup> 古谷 博 史<sup>†2</sup>

本研究では遺伝的アルゴリズム (GA) における OneMax 問題について、確率論的解析を行った。まず遷移確率行列を求め、その固有値の明示的な形を与える。さらに、2 番目に大きな固有値  $\nu_1$  を用いて収束時間を予測する。その結果、収束時間は突然変異率に依存することが分かる。

### Study of prediction method for the convergence time for OneMax problem

KIMINOBU KOGA,<sup>†1</sup> MAKOTO SAKAMOTO<sup>†2</sup>  
and HIROSHI FURUTANI<sup>†2</sup>

In this study, we carried out probabilistic analysis for Onemax problem in Genetic Algorithms(GA). First we derive the transition probability matrix and obtain the explicit form of eigenvalues. Then we use the second largest eigenvalue  $\nu_1$  for predicting convergence time and indicate that it depends on the mutation rate.

#### 1. はじめに

遺伝的アルゴリズム (GA) の計算時間を見積もることは、実際の問題に GA を適用する場合に非常に重要である。GA では扱う個体数は有限である。しかし、有限の個体数を仮定した理論では確率論的取扱いが必要であるため、理論が複雑になる。本研究では GA における OneMax 問題について、1 次スキームを用いて解析した。集団が連鎖平衡にある場合

OneMax 問題と非対称突然変異モデルが同等な振る舞いをする事が分かっている<sup>1)</sup>。本研究では非対称突然変異モデルについての広汎な研究成果<sup>2)</sup>を OneMax 問題に適用した。その結果得られた収束時間の予測法について報告する。

#### 2. 理 論

ここでは、Goldberg の Simple Genetic Algorithm (SGA) を進化のモデルとし、OneMax 問題における決定論的進化方程式を導く。さらに、確率論的モデルの OneMax 問題への適用を検討する。個体は長さ  $\ell$  の固定長 2 進ビット列  $\langle i(\ell), \dots, i(1) \rangle$  で表し、第  $k$  ビットの値を  $i(k)$  とする。OneMax 関数の適応度  $f_i$  を  $f_i = \sum_{k=1}^{\ell} i(k)$  と定義する。突然変異率を  $p_m$  と表す。

遺伝学では、連鎖の概念が重要であり、GA においても連鎖の有無が計算効率に大きな影響を及ぼす。交叉と突然変異は、ビット間の連鎖を弱め、集団を連鎖平衡状態に導く働きがある<sup>3)</sup>。集団が連鎖平衡にあることを仮定した場合、1 次スキームの進化方程式は

$$h_i(t+1) = ah_1(t) + b, \quad (1)$$

となる<sup>3)</sup>。ここで  $h_1$  はビット 1 をもつ 1 次スキームの相対確率、係数  $a, b$  は

$$a = (1 - \frac{1}{\ell})(1 - 2p_m), \quad b = \frac{1}{\ell}(1 - 2p_m) + p_m,$$

で与えられる。

集団遺伝学における代表的な確率論的モデルの一つに、Markov モデルを応用した Wright-Fisher モデルがある。ここでは Wright-Fisher モデルを OneMax 問題へ適用する。まず、OneMax 問題における 1 次スキームの遷移確率行列を求める。1 次スキームの進化方程式 (1) において、 $h_1(t) \rightarrow y = i/N$  と置き換え、遷移確率行列

$$P_{i,j} = P(j|i) = \binom{N}{j} u(y)^j \{1 - u(y)\}^{N-j}, \quad u(y) = a \frac{i}{N} + b,$$

を得る。遷移確率行列  $P$  の固有値  $\nu_0, \dots, \nu_N$  は

$$\nu_0 = 1, \quad \nu_1 = a, \quad \nu_k = \prod_{i=0}^{k-1} a(1 - \frac{i}{N}), \quad (2)$$

となり、2 番目に大きい固有値  $\nu_1 = a$  は  $N$  に依存しない<sup>1)</sup>。

ここで、世代  $t$  において遺伝子型 1 の個体数が  $i$  個となる確率を  $\mu_i(t)$  とする。このとき進化方程式は

$$\mu_j(t+1) = \sum_{i=0}^N \mu_i(t) P_{i,j}, \quad (3)$$

<sup>†1</sup> 宮崎大学工学研究科

Graduate School of Engineering, University of Miyazaki

<sup>†2</sup> 宮崎大学工学部

Faculty of Engineering, University of Miyazaki

と表され、ベクトル形式では  $\mu^T = (\mu_0, \mu_1, \dots, \mu_N)$  とすると次のように与えられる。

$$\mu(t+1)^T = \mu(t)^T P. \quad (4)$$

遷移確率行列  $P$  の固有値は (2) 式で与えられるが、 $P$  の左固有ベクトルを  $\mathbf{u}_i^T$ 、右固有ベクトルを  $\mathbf{v}_i^T$  と表す。

$$\mathbf{u}_i^T P = \nu_i \mathbf{u}_i^T, \quad P \mathbf{v}_i = \nu_i \mathbf{v}_i, \quad (0 \leq i \leq N).$$

このとき、1 次スキーマの初期分布は左固有ベクトルを用いて

$$\mu(t=0)^T = \sum_{i=0}^N C_i \mathbf{u}_i^T, \quad (5)$$

と表される。  $\sum_{i=0}^N \mu_i(t) = 1$ 、かつ、 $\mathbf{v}_0 = (1, 1, \dots, 1)^T$  から  $C_0 = 1$  を得る。また、十分大きな  $t$  に対して  $\mu(t)$  は

$$\mu(t) \approx \mathbf{u}_0 + a^t C_1 \mathbf{u}_1, \quad (6)$$

となり、 $\mathbf{u}_0$  は定常分布である<sup>1)</sup>。

(6) 式から、 $a^t$  が収束の速さを決めていていると考えられる。そこで、収束時間の理論値を

$$T_c(\text{eigenvalue}) = \min_t \{a^t < 0.05\}, \quad (7)$$

と定めた。比較のために、平均適応度による収束時間を

$$T_c(\text{fitness}) = \min_t \left\{ \frac{\bar{f}(\infty) - \bar{f}(t)}{\bar{f}(\infty) - \bar{f}(0)} < 0.05 \right\}, \quad (8)$$

とした。ただし、 $\bar{f}(t)$  は世代  $t$  における平均適応度、 $\bar{f}(\infty)$  は定常状態における平均適応度である。

定常分布への収束の速さの指標として mixing time  $T_m$  がよく用いられる<sup>4)</sup>。

$$T_m = \sum_{j=0}^N m_{ij} u_{0j} \quad (9)$$

(9) 式は Hunter の計算法で、 $m_{ij}$  は初期状態  $i$  から状態  $j$  に初めて到達する時間の平均である。また、 $u_{0j}$  は定常分布  $u_0$  の成分である。

### 3. 数値計算

OneMax 問題の収束時間について  $T_c(\text{eigenvalue})$  と  $T_c(\text{fitness})$  を比較した。選択は適応度比例選択、交叉は一様交叉を用いた。交叉率  $p_c = 1$  とし、 $\ell = 40$ 、 $N = 100$ 、 $T = 300$  とした。

図 1 は収束時間の突然変異率依存性を示した。横軸は突然変異率であり、縦軸は収束時間である。TC(EIGEN)、TC(FITNESS) はそれぞれ、(7) 式と (8) 式から求めた値である。

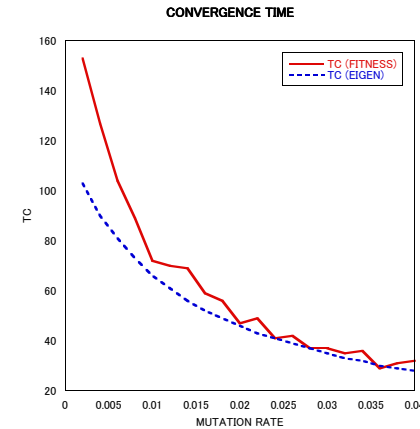


図 1 The mutation rate dependence of the convergence time  $T_c(\text{fitness})$  and  $T_c(\text{eigenvalue})$  with  $N = 100$ .

### 4. おわりに

本研究では GA における OneMax 問題の収束時間が、1 次スキーマの遷移確率行列の固有値によって予測可能であることを導いた。また、この固有値は集団サイズ  $N$  に無関係であるため、 $N$  に対して収束時間がほとんど依存しないことも分かった。本報告で示した予測法以外にも収束時間を予測する方法がいくつかある。今後は他の予測法との比較を行いたい。

### 参考文献

- 1) Q.Ma, Y.Zhang, K.Koga, K.Yamamori, M.Sakamoto and H.Furutani: "Analysis of first order schema for OneMax problem", Proceedings of seventeenth International Symposium on Artificial Life and Robotics (AROB2012).
- 2) T. Wai-Yuan: "Stochastic Models with Applications to Genetics, Cancers, AIDS and Other Biomedical Systems", (World Scientific, New Jersey, 2002)
- 3) H. Furutani: "Schema Analysis of OneMax Problem - Evolution Equation for First Order Schemata". in Foundations of Genetic Algorithms 7, (Morgan Kaufmann, San Francisco, 2003) 9-26
- 4) Hunter, J.J.: "Coupling and mixing times in Markov chain". Linear Algebra and its Applications, Vol.430. pp.2607-2621(2009).