

進化計算における OneMax 問題の Markov 連鎖を用いた収束時間解析

古賀 仁信^{†1} 安永 和馬^{†2} 馬 青蓮^{†3}
坂本 真人^{†2} 古谷 博史^{†2}

進化計算を実際の問題に適用する場合、その計算時間を予測することは非常に重要である。本論文では、OneMax 問題の収束時間を Markov 連鎖モデルを用いて解析した結果について報告する。最初に、OneMax 問題の進化過程を Markov 連鎖モデルの一つである Wright-Fisher モデルを用いて記述し、その遷移行列を計算する。次に、集団が連鎖平衡にある場合、OneMax 問題が非対称突然変異モデルと同等であることを示す。非対称突然変異モデルは集団遺伝学でよく研究されており、遷移行列の固有値はすでに知られている。我々は、その結果から OneMax 問題における遷移行列の固有値の解析的な形を導いた。Markov 連鎖の定常状態への収束時間は、遷移行列の 2 番目に大きい固有値を用いることで理論的に予測することができる。得られた理論的予測値と平均適応度の収束時間を実験的に比較し、よい一致が得られること示す。

Analysis of convergence time using Markov chain for OneMax problem in evolutionary computation

KIMINOBU KOGA,^{†1} YASUNAGA KAZUMA,^{†2} QINGLIAN MA,^{†3}
MAKOTO SAKAMOTO^{†2} and HIROSHI FURUTANI^{†2}

In applying Evolutionary Computation (EC) to realistic problems, it is very important to estimate their computational times. In this paper, we report the results obtained by Markov chain model for the convergence time of OneMax problem. First, we describe the evolution process of OneMax problem within the framework of Wright-Fisher model, which is a version of Markov chain, and calculate its transition matrix. Next, we show that, if a population is in linkage equilibrium, OneMax problem is equivalent to the asymmetric mutation model. The asymmetric mutation model is well studied in population genetics, and the eigenvalues of its transition matrix were already known. From this result, the analytical form of eigenvalues of transition matrix for OneMax problem can be obtained. In Markov chain theory, the convergence time of Markov chain can be estimated by using the second largest eigenvalue of the transition matrix. We compare this theoretical estimation with the convergence time of average fitness, and show that the agreement is very good.

1. はじめに

進化計算は、これまで最適化などの様々な問題に適用され、多大の成果を上げてきた。しかし応用面での成功に比べ、計算機構の理論的解明については大きく遅れているのが現状である¹⁾。進化計算の理論的研究において、最近注目されている確率論的解析手法の一つにマルコフ連鎖がある²⁾。進化計算では多くの場合、マルコフ性（未来の状態は現在の状態のみによって予

測できる）が成り立ち、マルコフ連鎖の理論が適用可能である。マルコフ連鎖理論は豊富な理論的内容をもち、理学、工学の分野に応用され、その有効性は広く認められている。特に集団遺伝学では多くのマルコフ連鎖による研究がなされている^{3),4)}。近年、進化計算の分野でもマルコフ連鎖を利用した研究が次第に増加しつつある。

遺伝的アルゴリズム (GA) におけるマルコフ連鎖理論の適用として、Nix と Vose による GA の進化方程式 (マルコフ連鎖の遷移行列) の導出がある⁵⁾。また彼らとは独立に Davis と Principe は同様な方程式を導いた。彼らの方法は、選択、突然変異、交叉を含む GA の確率論的枠組みを与え、GA の進化過程を正確に記述することができるが、その遷移確率行列の次元は、集団の個体数 N や個体を表現するビット長 l の増加とともに指数関数的に増加する。そのため

^{†1} 宮崎大学工学研究科
Graduate School of Engineering, University of Miyazaki

^{†2} 宮崎大学工学部
Faculty of Engineering, University of Miyazaki

^{†3} 宮崎大学農学工学総合研究科
Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki

程度の個体数をもつ現実の問題に彼らのモデルを適用することはほぼ不可能である。

Evolutionary Algorithm (EA) において, Drift Analysis と呼ばれるマルコフ連鎖を応用した手法が開発され, 研究者の注目を集めている¹⁾. Drift Analysis は, 1982 年の Hajek の論文⁶⁾ から始まると言われている. しかし実際に EA の分野に応用したのは Yao らの Birmingham 大学のグループである⁷⁾. その後, ヨーロッパの研究者を中心に急速に Drift Analysis を行う研究者が増加し, 様々な問題に適用されつつある.

GA の基礎的研究において, マルコフ連鎖などの確率過程理論を適用して例はあまり多くない. それは一般に, 問題の次元数が大きくなりすぎて, 数値計算がほとんど実行不可能なほどになってしまうことによる. しかし, GA では従来からスキーマ理論とよばれる解析法がある. 最近になり, より厳密で定量的取り扱いができる新しいスキーマ理論が開発された^{8), 9)}. スキーマ理論は進化計算における問題の次元数を大幅に縮小することができるという特徴をもつ. 本研究の特色は, 定量的なスキーマ理論に着目し, 対象問題の次元数を大幅に縮小することにより, 理論的, 数値的研究を可能とすることにある. また, 確率的揺らぎの効果は, 従来の研究では副次的なものにすぎないと考えられてきたが, 我々のこれまでの研究によれば, 本質的に重要な働きをもつ. その計算に及ぼす影響を総合的に理解することは, 今後の進化計算の発展において非常に重要である.

本研究では, OneMax 問題を取り上げ, その進化過程の定量的な分析を行った. 計算過程を Markov 連鎖としてモデル化し, 連鎖が定常状態に収束する世代数を予測する手法を提案する. まず集団が連鎖平衡にある場合, OneMax 問題が非対称突然変異モデルと同等であることを示す. 非対称突然変異モデルは集団遺伝学でよく研究されている¹²⁾. モデルの様々な確率論的性質が得られており, 遷移行列の固有値も知られている. 我々は, その結果から OneMax 問題における遷移行列の固有値の解析的な形を導いた. Markov 連鎖の定常状態への収束時間は, 遷移行列の 2 番目に大きい固有値を用いることで理論的に予測することができる. この理論的予測値と実験的に得られた平均適応度の収束時間を比較検討した. 本研究の一部は, 国際会議で発表した¹⁰⁾.

2. 数学モデル

進化のモデルとして Goldberg の Simple Genetic Algorithm (SGA) を用い, OneMax 問題における標準的な突然変異モデルと非対称突然変異モデルとの関係を導出する. 集団の個体数は世代に依らず常に一定とする. 個体の遺伝子型を固定長 ℓ の 2 進ビット列で表す. このとき, 遺伝子型の総数は $N = 2^\ell$ となる. 遺

伝子型の番号 $i (0 \leq i \leq n-1)$ と 2 進ビット列を

$$i = \langle i(\ell), i(\ell-1), \dots, i(2), i(1) \rangle,$$

のように対応させる. $i(k)$ は第 k ビットの値である. 世代 t における遺伝子型 i の個体数を N_i とすると,

$$N = \sum_{i=0}^{n-1} N_i(t),$$

となる. 遺伝子型 i の相対頻度を

$$x_i = \frac{N_i(t)}{N},$$

と表す. OneMax 問題の適応度関数は

$$f_i = \sum_{k=1}^{\ell} i(k),$$

として定義される. 適応度は全てのビットが 1 のとき最大となる.

連鎖は GA の計算効率に大きな影響を及ぼす重要な概念である. 交差や突然変異はビット間の連鎖を弱め, 集団を連鎖平衡の状態へ導く¹¹⁾. もし, 集団が連鎖平衡の状態であれば, 個体の分布は 1 次スキーマの頻度のみ依存する. 1 次スキーマの進化方程式は¹¹⁾

$$h_1(t+1) = ah_1(t) + b, \quad (1)$$

である⁹⁾. $h_1(t)$ はビット 1 をもつ 1 次スキーマの相対頻度を表す. 係数 a と b は

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

によって与えられ, p_m は突然変異率を表す. 非対称突然変異の GA において, ビット $1 \rightarrow 0$, $0 \rightarrow 1$ の突然変異率を, それぞれ p_α , p_β とする. 1 次スキーマの決定論的進化方程式は

$$\begin{aligned} h_1(t+1) &= (1 - p_\alpha)h_1(t) + p_\beta h_0(t), \\ h_0(t+1) &= p_\alpha h_1(t) + (1 - p_\beta)h_0(t), \end{aligned}$$

となる. この両式に $h_1(t) + h_0(t) = 1$ を用いると

$$h_1(t+1) = (1 - p_\alpha - p_\beta)h_1(t) + p_\beta, \quad (3)$$

を得る. 式 (1) と式 (3) を比較し, 次の関係を得る.

$$p_\alpha = p_m, \quad p_\beta = p_m + \frac{1}{\ell}(1 - 2p_m). \quad (4)$$

3. Markov モデル

集団遺伝学における Markov モデルの 1 つに, Wright-Fisher モデルがある. Wright-Fisher モデルでは, 1 つの遺伝子座と 2 つの対立遺伝子を持つ染色体を取り扱う. これは遺伝子型が $i \in \{0, 1\}$ で長さ $\ell = 1$ の GA に相当する³⁾. 遺伝子型 1 の個体数を

$$N_1 = 0, 1, \dots, N,$$

と表すと, 遺伝子型 0 の個体数は $N_0 = N - N_1$ で与えられる.

ここでは, Wright-Fisher モデルを OneMax 問題に適用する. 世代 t において遺伝子型 1 の個体数が $N_1 = i$ のとき, 次世代 $t+1$ で $N_1 = j$ となる条件付き確率は

$$\begin{aligned} P_{i,j} &= P(j|i) = \binom{N}{j} p_i^j (1 - p_i)^{N-j} \quad (5) \\ p_i &= ay + b = a \left(\frac{i}{N}\right) + b, \end{aligned}$$

と表される．

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 $\mu_i(t)$ をベクトル形式で次のように表す．

$$\mu^T = (\mu_0, \mu_1, \dots, \mu_N)$$

また，正規化条件は

$$\sum_{i=0}^N \mu_i(t) = 1, \quad (6)$$

となる．このとき，進化の過程は

$$\mu_j(t+1) = \sum_{i=0}^N \mu_i(t) P_{i,j}, \quad (7)$$

で表される．

進化方程式はベクトル形式

$$\mu(t+1)^T = \mu(t)^T P, \quad (8)$$

で与えられる． P は $(N+1) \times (N+1)$ の行列である．この行列の固有値は分かっており¹²⁾

$$\lambda_0 = 1, \quad \lambda_1 = a, \quad \lambda_2 = a^2(1-1/N), \dots, \quad (9)$$

となる．このうち，2 番目に大きな固有値 $\lambda_1 = a$ が集団サイズ N に依存しないことは重要なことである．行列 P の左固有ベクトルと右固有ベクトルを

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

とする．これらの固有ベクトルは直交性の条件を満たす．

$$\mathbf{u}_i^T \cdot \mathbf{v}_j = 0 \quad (i \neq j) \quad (10)$$

右固有ベクトル \mathbf{v}_0 は自明であり，

$$\mathbf{v}_0 = (1, 1, \dots, 1)^T \quad (11)$$

によって与えられ，左固有ベクトル \mathbf{u}_0 のすべての要素は正である．正規化条件を適用すると

$$\sum_{i=0}^N |u_i| = 1, \quad \mathbf{u}_i^T \cdot \mathbf{v}_i = 1,$$

を得る．

突然変異率 pm が正の場合，遷移行列 P のすべての要素は正となり，スキーマ進化における Markov 連鎖は既約かつ非周期的である．既約かつ非周期的な Markov 連鎖は定常分布 π に収束する．

$$\lim_{t \rightarrow \infty} \mu(t) = \pi. \quad (12)$$

また，すべての要素 π_i は正である¹²⁾．

1 次スキーマの初期分布は左固有ベクトルによって

$$\mu(t=0)^T = \sum_{i=0}^N C_i \mathbf{u}_i^T, \quad (13)$$

と表される． \mathbf{u}_i は固有ベクトルなので，(8) 式より

$$\mu(t)^T = \sum_{i=0}^N \lambda_i^t C_i \mathbf{u}_i^T \quad (14)$$

を得る．ここで右側から右固有ベクトル \mathbf{v}_0 を掛けると

$$\mu(t)^T \cdot \mathbf{v}_0 = \sum_{i=0}^N N \lambda_i^t C_i \mathbf{u}_i^T \cdot \mathbf{v}_0 = \lambda_0^t C_0 = C_0,$$

を得，さらに (6) 式と (11) 式から $C_0 = 1$ を得る． t が十分大きい場合

$$\mu(t) \approx \mathbf{u}_0 + a^t C_1 \mathbf{u}_1, \quad (15)$$

となり， \mathbf{u}_0 は定常分布 π を表す．

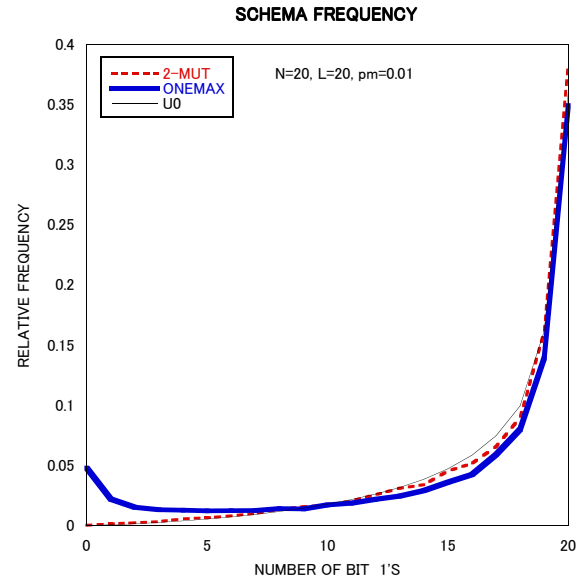


図 1 $N = 20$, $p_m = 0.01$ の場合の 1 次スキーマの分布．
Fig. 1 The distribution of the number of bit 1's with $N = 20$ and $p_m = 0.01$. The dotted line shows the result of asymmetric mutation model. The thick solid line for OneMax GA calculation and the thin solid line for the left eigenvector \mathbf{u}_0 , respectively.

(15) 式から，Markov 連鎖の収束時間に a^t が大きな影響を与えていることが分かる． a は遷移確率行列の 2 番目に大きな固有値である．この値を用いて収束時間の理論予測値を

$$T_c(\text{eigenvalue}) = \min_t \{a^t < 0.05\}, \quad (16)$$

と定義する．また，実験的な推定として平均適応度を利用し，定常状態への収束時間を

$$T_c(\text{fitness}) = \min_t \left\{ \frac{\bar{f}(\infty) - \bar{f}(t)}{\bar{f}(\infty) - \bar{f}(0)} < 0.05 \right\}, \quad (17)$$

と定義する．ただし， $\bar{f}(t)$ は世代 t での平均適応度， $\bar{f}(\infty)$ は定常状態での平均適応度を表す．

4. 数値計算

これまでの考察を基に，実際に数値計算を行った結果と理論を比較する．選択は適応度比例選択，交叉は一様交叉を用い，交叉率は 1 とした．突然変異はビット値を反転させる方法を用いた．初期集団は各ビットごとに確率 0.5 でビット 0 と 1 を発生させ，初期個体を生成した．よって，初期集団は連鎖平衡の状態にある．乱数を変えながら同じ計算を 1000 回繰り返し，結果の平均を比較した．

4.1 OneMax 問題と非対称突然変異の GA との関係

まず，OneMax 問題と非対称突然変異の GA が同等の振る舞いをするところを確かめる．

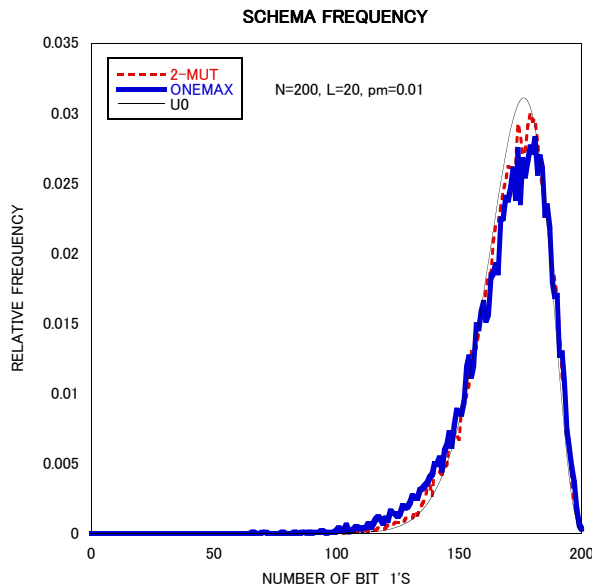


図 2 $N = 200$, $p_m = 0.01$ の場合の 1 次スキーマ分布。
 Fig. 2 The distribution of the number of bit 1's with $N = 200$ and $p_m = 0.01$.

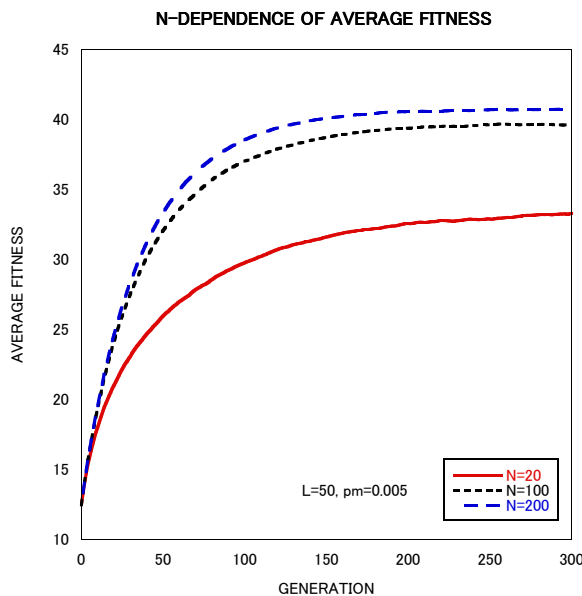


図 3 平均適応度の N 依存性。
 Fig. 3 Dependence of the average fitness on N . $p_m = 0.005$. The horizontal axis represents generation.

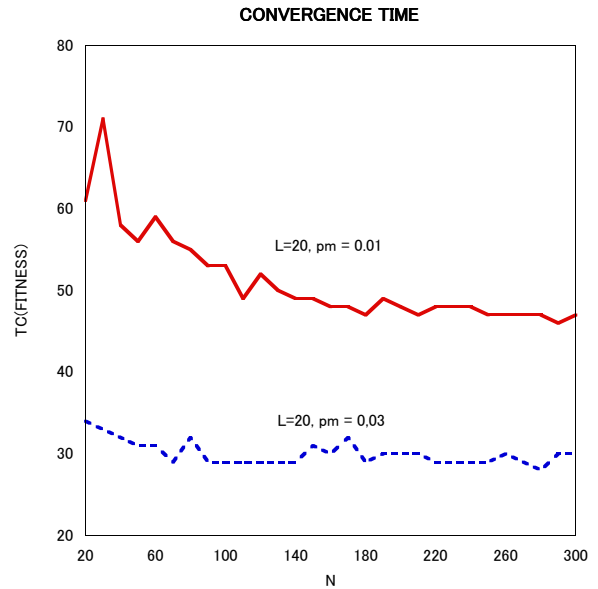


図 4 収束時間の N 依存性。
 Fig. 4 The population size dependence of the convergence time $T_c(\text{fitness})$ with $p_m = 0.01$ and 0.03 . $\ell = 20$.

図 1 と図 2 に OneMax 問題と非対称突然変異の GA の 1 次スキーマの分布を示した。 u_0 は推移確率行列 P の固有ベクトルであり、同時に定常分布 π である。したがって、OneMax 問題も非対称突然変異の GA も、定常分布に収束していることが分かる。 N や p_m の値により分布の形は異なるものの、いずれの場合においても OneMax 問題と非対称突然変異の GA の 1 次スキーマ分布はほとんど一致している。

図 3 は平均適応度の集団サイズ N 依存性を示している。 N の値が大きくなると平均適応度も高くなるが、収束する速さはあまり変化がない。

4.2 収束時間

図 4 は収束時間の N 依存性を示している。 N の値が小さい場合を除けば、収束時間は N の値に関わらずほぼ一定となっている。本研究では、収束時間を N によらない固有値 a を用いて予測している。図 4 に示した結果は、理論予測と矛盾しない結果となっている。 N が小さい場合については、図 1 も $N = 20$ の場合であるが、他の分布と比べると不一致の部分が目立つことから、集団が連鎖平衡の状態にないことが考えられる。本研究では、集団が連鎖平衡の状態にあることを前提としてるため、このことが不一致の原因の可能性はある。

図 5 と図 6 は収束時間のビット長 ℓ に対する依存性を示している。ここでは p_m の値を 0.01 と 0.03 に取った場合を示すが、いずれの場合においても、計算結果と予測値は良い一致をしている。式 (2) より、 ℓ が大きくなるほど、 a の値も大きくなる。式 (16) に

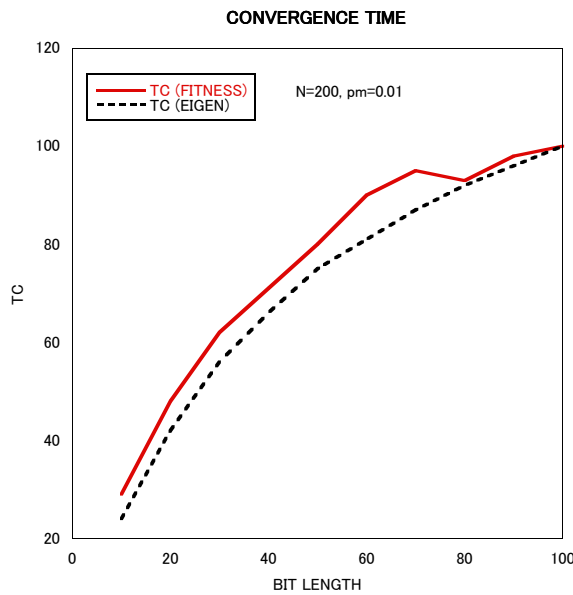


図 5 収束時間のビット長 ℓ 依存性 .
 Fig. 5 The bit length dependence of the convergence time $T_c(\text{fitness})$ with $N = 200$, $p_m = 0.01$.

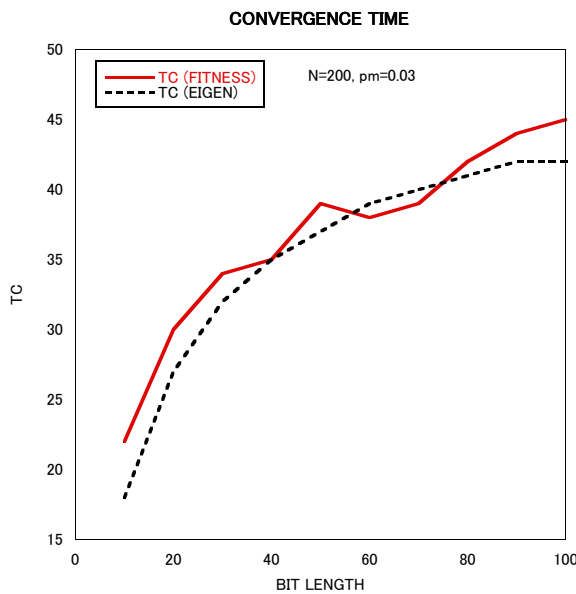


図 6 収束時間のビット長 ℓ 依存性 .
 Fig. 6 The bit length dependence of the convergence time $T_c(\text{fitness})$ with $N = 200$, $p_m = 0.03$.

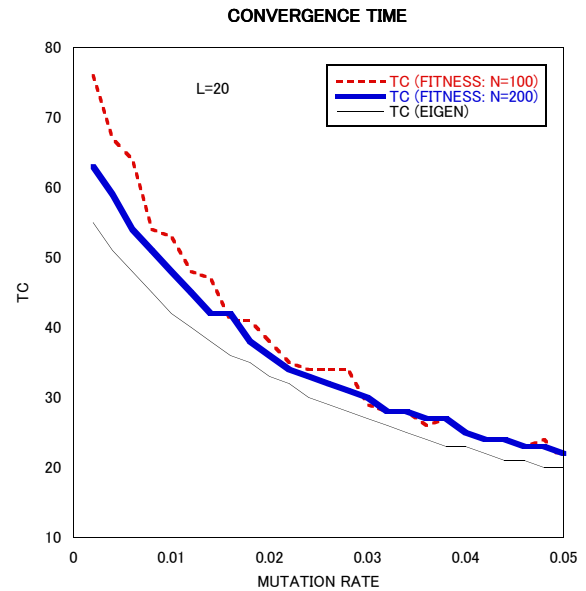


図 7 収束時間の p_m 依存性 .
 Fig. 7 The mutation rate dependence of the convergence time $T_c(\text{fitness})$ with $N = 100$ and $N = 200$. $\ell = 20$.

よると、 a が大きいほど収束時間は遅くなるが、確かに数値計算の結果と一致している。

図 7 は収束時間の p_m 依存性を示している。数値計算の結果が予測値よりもやや大きな値を取ったが、上手く予測できていると言える。 ℓ の場合とは逆に、式 (2) より p_m が大きくなるほど a は小さくなるので、それに伴って収束は速くなる。実際の計算結果も p_m の増加とともに T_c が小さくなっていることが分かる。

図 8 は初めて最適解が出現した時間の N 依存性を示している。集団サイズが大きくなるほど最適解が早く出現することが分かる。

5. まとめと今後の課題

本研究では、OneMax 問題と非対称突然変異の GA が同等であることを示し、OneMax 問題の収束時間を定量的に予測する方法を提案した。実際に計算した結果が示す通り、提案した予測法である程度の収束時間の予測ができることが分かった。特に、収束時間が集団サイズ N によらない事は重要なことである。また、ビット長 ℓ が短いほど、突然変異率 p_m が大きいほど速く収束することが分かった。

今回は平均適応度の収束時間を予測する手法を提案したが、最適解が初めて出現する時間を予測する方法について今後の研究課題として考えていきたい。

この他、確率的計算手法として乱択アルゴリズムが注目を集めている^{13),14)}。乱択アルゴリズムでは、計算効率の予測のため Chernoff 限界, Azuma 不等式など

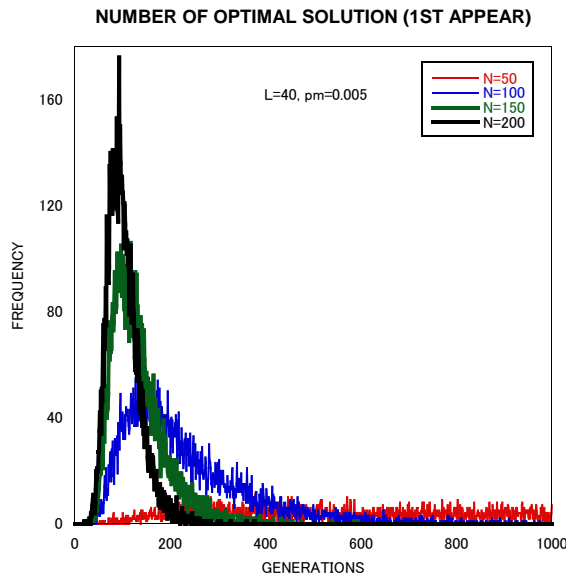


図 8 初めて最適解が出現する時間の N 依存性 ($pm = 0.005$, $L = 40$).

Fig. 8 The N dependence of number of generations that the optimal solution will appear for the first time. $\ell = 40$, $p_m = 0.005$.

様々な数学的手法が取り入れられており、進化計算の分野での応用も考えていきたい。

進化計算においてマルコフ連鎖などの確率論的方法を取り入れたものとして分布推定にもとづく進化計算がある。これは Estimation of Distribution Algorithm(EDA) などと呼ばれるが様々なバリエーションがある。分布推定法は明示的に統計的計算を行って集団を生成するため、従来の GA などと比べて確率論的性質が比較的単純であり、研究の対象として取り上げてみたい。

参 考 文 献

- 1) A. Auger and B. Doerr, eds.: “Theory of Randomized Search Heuristics”, (World Scientific, New Jersey, 2011)
- 2) J. G. Kemeny and J. L. Snell: “Finite Markov chains”, Springer-Verlag, 1976
- 3) Crow, J. F., Kimura M.: An Introduction to Population Genetics Theory. Harper and Row, New York (1970)
- 4) 福島正俊, 石井一成: 自然現象と確率過程 . 増補版 . 日本評論社, 東京 (1996)
- 5) S. Droste, T. Jansen and I.Wegener: “On the analysis of the (1+1) evolutionary algorithm”, Theoretical Computer Science, **276**, pp.51–81, 2002
- 6) B. Hajek: “Hitting-time and Occupation-time

Bounds Implied by Drift Analysis with Applications”, Advances in Applied Probability, **13**, pp.502–525, 1982

- 7) J. He and X. Yao: “Drift Analysis and Average Time Complexity of Evolutionary Algorithms”, Artificial Intelligence, **127**, pp.57–85, 2001
- 8) 古谷博史: Walsh 変換による突然変異と交叉に対するスキーマ定理の導出, 情報処理学会論文誌「数理モデル化と応用」, **43**, pp.1050–1060, 2002
- 9) 古谷博史: 選択における連鎖不平衡の効果-OneMax 問題のスキーマ解析, 情報処理学会論文誌「数理モデル化と応用」, **45**, pp.12–21, 2004
- 10) Q. Ma, Y. Zhang, K. Koga, M. Sakamoto, and H. Furutani: “Analysis of first order schema for OneMax problem”, Proceedings of the 17th International Symposium on Artificial Life and Robotics, 2012
- 11) 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
- 12) T. Wai-Yuan: “Stochastic Models with Applications to Genetics, Cancers, AIDS and Other Biomedical Systems”, (World Scientific, New Jersey, 2002)
- 13) R. Motwani and P.Raghavan: “Randomized Algorithms”, (Cambridge University Press, Cambridge, 1995)
- 14) 玉木久夫: 乱択アルゴリズム, (共立出版, 2008)